The Aspergilli
Genomics, Medical Aspects, Biotechnology, and Research Methods

Edited by
Gustavo H. Goldman
Stephen A. Osmani

CRC Press
Taylor & Francis Group
The Aspergilli
Genomics, Medical Aspects, Biotechnology, and Research Methods
MYCOLOGY SERIES

Editor
J. W. Bennett
Professor
Department of Plant Biology and Pathology
Rutgers University
New Brunswick, New Jersey

Founding Editor
Paul A. Lemke

1. Viruses and Plasmids in Fungi, edited by Paul A. Lemke
2. The Fungal Community: Its Organization and Role in the Ecosystem, edited by
   Donald T. Wicklow and George C. Carroll
3. Fungi Pathogenic for Humans and Animals (in three parts), edited by
   Dexter H. Howard
5. Secondary Metabolism and Differentiation in Fungi, edited by Joan W. Bennett
   and Alex Ciegler
6. Fungal Protoplasts, edited by John F. Peberdy and Lajos Ferenczy
7. Viruses of Fungi and Simple Eukaryotes, edited by Yigal Koltin
   and Michael J. Leibowitz
8. Molecular Industrial Mycology: Systems and Applications for Filamentous Fungi,
   edited by Sally A. Leong and Randy M. Berka
9. The Fungal Community: Its Organization and Role in the Ecosystem,
    Second Edition, edited by George C. Carroll and Donald T. Wicklow
10. Stress Tolerance of Fungi, edited by D. H. Jennings
11. Metal Ions in Fungi, edited by Gü‘çnter Winkelmann and Dennis R. Winge
    and Colin G. Orpin
    Richard A. Calderone and Ronald L. Cihlar
    Dexter H. Howard
17. Fungi in Ecosystem Processes, John Dighton
18. Genomics of Plants and Fungi, edited by Rolf A. Prade and Hans J. Bohnert
19. Clavicipitalean Fungi: Evolutionary Biology, Chemistry, Biocontrol, and Cultural
    Impacts, edited by James F. White Jr., Charles W. Bacon, Nigel L. Hywel-Jones,
    and Joseph W. Spatafora
21. Fungal Biotechnology in Agricultural, Food, and Environmental Applications,
    edited by Dilip K. Arora
22. Handbook of Industrial Mycology, edited by Zhiqiang An
The Aspergilli
Genomics, Medical Aspects, Biotechnology, and Research Methods

Edited by
Gustavo H. Goldman
Universidade de São Paulo
São Paulo, Brazil

Stephen A. Osmani
The Ohio State University
Columbus, Ohio, U.S.A.
Contents

Preface .......................................................... xi
Editors ........................................................... xiii
Contributors ..................................................... xv

I. Genomics of the Aspergilli

1. An Overview of the Genus Aspergillus ......................... 3
   Scott E. Baker and Joan W. Bennett

2. A First Glance into the Genome Sequence of Aspergillus flavus 15
   Gary A. Payne, Jiujiang Yu, William C. Nierman, Masayuki Machida,
   Deepak Bhatnagar, Thomas E. Cleveland, and Ralph A. Dean

3. A Comparative View of the Genome of Aspergillus fumigatus 25
   Natalie D. Fedorova, William C. Nierman, Geoffrey Turner, Vinita Joardar,
   Rama Maiti, Michael J. Anderson, David W. Denning, and Jennifer R. Wortman

4. Aspergillus nidulans Genome and a Comparative Analysis of Genome Evolution in Aspergillus 43
   Antonis Rokas and James E. Galagan

5. Aspergillus nidulans Linkage Map and Genome Sequence:
   Closing Gaps and Adding Telomeres .......................... 57
   A. John Clutterbuck and Mark Farman

6. Genome Sequence of Aspergillus oryzae ......................... 75
   Masayuki Machida

II. Basic Biology of the Aspergilli

7. Signal Transduction in Aspergilli ............................... 87
   Jae-Hyuk Yu and Christophe d’Enfert

8. Gene Regulation ................................................. 103
   Mark X. Caddick and Chris Dobson

9. Mitogen-Activated Protein Kinase Pathways in Aspergilli ........ 121
   Gregory S. May
10. Gluconeogenic Carbon Metabolism .......................................................... 129
    Michael J. Hynes

11. Amino Acid Supply of Aspergillus .......................................................... 143
    Oliver W. Draht, Silke Busch, Kay Hofmann, Susanna Braus-Stromeyer,
    Kerstin Helmstaedt, Gustavo H. Goldman, and Gerhard H. Braus

12. Endocytosis ......................................................................................... 177
    Juan C. Sánchez-Ferrero and Miguel A. Peñalva

13. RNA Silencing in the Aspergilli ............................................................... 197
    Thomas M. Hammond and Nancy P. Keller

14. Hyphal Morphogenesis in Aspergillus nidulans ...................................... 211
    Steven D. Harris

15. Cytoskeleton, Polarized Growth, and the Cell Cycle
    in Aspergillus nidulans ........................................................................... 223
    Reinhard Fischer, Norio Takeshita, and John Doonan

    Eduardo A. Espeso and Stephen A. Osmani

17. Sexual Development in Aspergillus nidulans ............................................ 279
    Dong-Min Han, Keon-Sang Chae, and Kap-Hoon Han

18. Aspergillus Transporters ....................................................................... 301
    George Diallinas

19. Chromatin in the Genus Aspergillus ......................................................... 321
    Claudio Scazzocchio and Ana Ramón

20. Transposable Elements and Repeat-Induced Point Mutation in
    Aspergillus nidulans, Aspergillus fumigatus, and Aspergillus oryzae ........ 343
    A. John Clutterbuck, Vladimir V. Kapitonov, and Jerzy Jurka

III. Medically Important Aspects of the Genus

    William J. Steinbach

22. Pathogenicity Determinants and Allergens ............................................. 377
    Sven Krappmann

23. Mammalian Models of Aspergillosis ....................................................... 401
    Aimee K. Zaas and William J. Steinbach

24. Emerging Role of Mini-Host Models in the Study of Aspergillosis .......... 413
    Georgios Chamilos and Dimitrios P. Kontoyiannis
IV. Biotechnological Aspects of the Genus

25. Food Products Fermented by *Aspergillus oryzae* ........................................... 429
   Keietsu Abe and Katsuya Gomi

26. Aspergillus as a Cell Factory for Protein Production: Controlling Protease Activity in Fungal Production .......................................................... 441
   Machtelt Braaksma and Peter J. Punt

27. Mycotoxin Production and Prevention of Aflatoxin Contamination in Food and Feed .............................................................. 457
   Jiujiang Yu, Gary A. Payne, Bruce C. Campbell, Baozhu Guo, Thomas E. Cleveland,
   Jane F. Robens, Nancy P. Keller, Joan W. Bennett, and William C. Nierman

V. Methods: Techniques and Resources

28. Microarrays in *Aspergillus* Species ................................................................. 475
   Andrew Breakspear and Michelle Momany

29. Chemostats and Microarrays ..................................................................... 483
   Manda Gent and Karin Lanthaler

30. Advances in Gene Manipulations Using *Aspergillus nidulans* .......... 493
   Stephen A. Osmani, Hui-Lin Liu, Michael J. Hynes, and Berl R. Oakley

31. Fluorescent Labels for Intracellular Structures and Organelles ........ 513
   Berl R. Oakley and Xin Xiang

32. *Aspergillus* at the Fungal Genetics Stock Center .......................... 527
   Kevin McCluskey

Index ........................................................................................................... 537
Preface

As a group of organisms, the filamentous fungi impact mankind in both very positive and very negative ways. Yet as a whole they are poorly understood and surprisingly understudied organisms. However, within the genus *Aspergillus* exist some of the most intensely studied filamentous fungal species. One only has to consider the fact that over eight *Aspergillus* species have recently been sequenced to gain some understanding of the level of interest regarding this genus. These recent genomic advances provided the main impetus to produce this book as there has been an avalanche of information and insights generated recently and the time was ripe for a comprehensive volume covering the Aspergilli. We have tried to include chapters that deal with all aspects of the biology of the Aspergilli, ranging from new insights gleaned from genome sequence data, to basic biology, medical issues, biotechnology and evolving experimental approaches. The chapters, therefore, cover many aspects of the biology of filamentous fungi and include insights from numerous *Aspergillus* species. We firmly believe that groups studying different aspects of the biology of fungi, be it industrial, medical, agricultural, or academic in its nature, will benefit greatly from the collective integrated efforts of us all. This sentiment is eloquently expressed in the first chapter of the book by Baker and Bennet, “Using the resources currently in hand and strongly advocating coordination of research and development of new resources, we will move forward into a new “golden age” of *Aspergillus* research.” We hope that this book will help foster further collaborative research amongst groups working with different Aspergilli and filamentous fungi in general.

In the opening section, Genomics of the Aspergilli, the stage is set in Chapter 1 by Baker and Bennett who provide a global overview of the Aspergilli from a historical perspective. They clearly convey the magnitude of the impact that different species of *Aspergillus* have, and will continue to have, on the world in general and mankind in particular. This chapter provides an outstanding introduction to the Aspergilli and we feel should be required reading for all working on this most fascinating genus. In the next five chapters an extensive analysis of the genomes of *A. flavus*, *A. fumigatus*, *A. nidulans* and *A. oryzae* is presented by world experts on the genomics of the Aspergilli. These chapters provide new and interesting insights derived from the individual genomes of these species and also provide clear evidence of the power of comparative genomics. Comparative genomics is an exiting arena in which research utilizing the Aspergilli should blossom and excel. This message comes through loud and clear when reading these chapters.

In section II, Basic biology of the Aspergilli, numerous aspects of *Aspergillus* biology, and biology in general, are addressed. This is the largest section of the book and as would be expected is somewhat dominated by studies that utilize *A. nidulans*, the model genetic organism representative of the Aspergilli. These chapters provide a wide breadth of insights that are reflective of the historical strengths of research areas of *A. nidulans*. These include studies of gene expression and chromatin, metabolic regulation, developmental regulation, growth control, cell cycle regulation and cell biology. In addition, new areas of interest are covered, such as endocytosis, RNA silencing, transporters and transposable elements. In each area authors have incorporated new ideas and highlight the huge potential afforded by the genome sequences of the Aspergilli noting how this information can be harnessed for future research.

Section III touches on the medically important aspects of the Aspergilli. Many *Aspergillus* researchers concentrate their research to try and understand the pathological aspects of the Aspergilli with an eye to development of better diagnosis and treatments for which there are dire needs. With the advent of the age of *Aspergillus* genomics, and the methodologies this affords researchers in their arsenal to understand and treat disease caused by the Aspergilli, there is a need to have effective and standard models for these
diseases. Therefore in addition to two inclusive chapters on disease caused by Aspergilli and their pathogenic determinants, two chapters are included dealing with experimental models of Aspergillosis.

Several species of Aspergilli, rather than being enemies of mankind, are allies utilized in the biotechnology and food industries. Two chapters in section IV, Biotechnological aspects of the genus, describe the beneficial use of *Aspergillus* species for food production and as cell factories for commercial heterologous protein production. A final chapter delves back into the darker side of the Aspergilli and deals with the ability of some species to cause food spoilage and generate toxins that can cause death and cancer.

In the final section V, Methods: Techniques and Resources, several chapters are presented that outline technologies that can be utilized to experimentally leverage the genomes of the Aspergilli to maximum effect. It is generally true that any new technique or protocol generated using a specific species of *Aspergillus* works equally well in all other *Aspergillus* species and often other filamentous fungi. From the work of many *Aspergillus* labs, and those of others working with different filamentous fungi, it is safe to say that the main technical bottle necks that existed not long ago when working with the Aspergilli have all largely been overcome. At this point in time researchers in all fields of *Aspergillus* biology face few if any technical barriers and the “Golden Age” of *Aspergillus* research is ours for the taking. We hope that this book will help in some small way in realizing this vast potential.
Editors

Gustavo H. Goldman is Professor of Molecular Biology at the Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo. He received the BS degree in Biology from the Universidade Federal do Rio de Janeiro, Brazil, in 1983, the Master Science in Microbiology at the Universidade de São Paulo, Brazil, in 1988, and the PhD degree in Molecular Biology from the University of Gent, Belgium, in 1993. He joined the Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Brazil, in 1994 and was appointed as Professor in 2002. He is also a Researcher of the National Scientific Council, Brazil (CNPq).

His research interests include the study of human genetic instability syndromes using Aspergillus nidulans as a model system, the calcineurin pathway in the opportunistic fungus A. fumigatus, and the identification of genes that are expressed upon the dimorphic transition in the pathogenic fungus Paracoccidioides brasiliensis.

He has served on numerous program committees and chaired many international conferences and workshops. He has also served on various advisory committees of granting agencies, such as Wellcome Trust, Australian Research Council, and CNPq, CAPES, and FAPESP (Brazil). He is currently Associate Editor of “Fungal Genetics and Biology” and fellow of the John Guggenheim Memorial Foundation.

Stephen A. Osmani received an Higher National Diploma in Applied Biology, with Distinction in Microbiology, from the Polytechnic of the South Bank in 1977. He gained Membership of the Institute of Biology from Trent Polytechnic in 1979. He earned his PhD degree from the Department of Biochemistry, Kings College, London in 1984. Dr Osmani then moved to the United States and completed four years of Postdoctoral training at the Department of Pharmacology, Robert Wood Johnson Medical School. He joined the faculty at the Department of Cell Biology, Baylor College of Medicine as an Assistant Professor in 1988. He was a Senior Staff Scientist at the Weis Center for Research, Geisinger Clinic and Professor at the Penn State College of Medicine. Since 2001 he has been a Professor and Ohio Eminent Scholar in the Department of Molecular Genetics at the Ohio State University.

His main research interests are in the general area of cell cycle regulation with specific focus on the role of protein phosphorylation during regulation of mitosis and the dynamic changes that occur to nuclear structure during division. His research has been funded by the National Institutes for Health since 1989. He reviews widely for international granting agencies and journals and is currently on the editorial board of Eukaryotic Cell and a Member of the Nuclear Dynamics and Transport Study Section of the NIH.

In 2006 Dr Osmani was elected Fellow of the American Association for the Advancement of Science.
Contributors

Keietsu Abe  Graduate School of Agricultural Science, Tohoku University, Sendai, Japan

Michael J. Anderson  School of Medicine and Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Scott E. Baker  Pacific Northwest National Laboratory, Richland, Washington

Joan W. Bennett  Department of Plant Biology and Pathology, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, New Jersey

Deepak Bhatnagar  USDA/ARS, Southern Regional Research Center, New Orleans, Louisiana

Machtelt Braaksma  Department of Microbiology, TNO Quality of Life, Zeist, The Netherlands

Gerhard H. Braus  Department of Microbiology and Genetics, Institute of Microbiology and Genetics, Georg-August-University Goettingen, Goettingen, Germany

Susanna Braus-Stromeyer  Institute of Microbiology and Genetics, University of Goettingen, Goettingen, Germany

Andrew Breakspear  Department of Plant Biology, University of Georgia, Athens, Georgia

Silke Busch  Institute of Microbiology and Genetics, University of Goettingen, Goettingen, Germany

Mark X. Caddick  School of Biological Sciences, University of Liverpool, Liverpool, United Kingdom

Bruce C. Campbell  USDA/ARS, Western Regional Research Center, Albany, California

Keon-Sang Chae  Division of Biological Sciences, Chonbuk National University, Jeonju, South Korea

Georgios Chamilos  Department of Infectious Diseases, Infection Control and Employee Health, University of Texas, M. D. Anderson Cancer Center, Houston, Texas
Thomas E. Cleveland    USDA/ARS, Southern Regional Research Center, New Orleans, Louisiana

A. John Clutterbuck    Institute of Biomedical and Life Sciences, Division of Molecular Genetics, University of Glasgow, Glasgow, Scotland, United Kingdom

Ralph A. Dean    Department of Plant Pathology, North Carolina State University, Raleigh, North Carolina

David W. Denning    School of Medicine and Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

George Diallinas    Facult of Biology, Department of Botany, University of Athens, Athens, Greece

Chris Dobson    School of Biological Sciences, University of Liverpool, Liverpool, United Kingdom

John Doonan    John Innes Centre, Norwich, Norfolk, United Kingdom

Oliver W. Draht    Institute of Microbiology and Genetics, University of Goettingen, Goettingen, Germany

Christophe d’Enfert    Unité Biologie et Pathogénicité Fongiques, Département Génomes et Génétique, Institut Pasteur, Paris, France

Eduardo A. Espeso    Department of Microbiología Molecular, Centro de Investigaciones Biológicas, Ramiro de Maeztu, Madrid, Spain

Mark Farman    Department of Plant Pathology, University of Kentucky, Lexington, Kentucky

Natalie D. Fedorova    The Institute for Genomic Research, Rockville, Maryland

Reinhard Fischer    Department of Applied Microbiology, Institute for Applied Biosciences, University of Karlsruhe, Karlsruhe, Germany

James E. Galagan    Microbial Genome Analysis and Annotation Group, The Broad Institute of MIT and Harvard, Cambridge, Massachusetts

Manda Gent    Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Gustavo H. Goldman    Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

Katsuya Gomi    Graduate School of Agricultural Science, Tohoku University, Sendai, Japan

Baozhu Guo    USDA/ARS, Crop Protection and Management Research Unit, Tifton, Georgia
Thomas M. Hammond  Department of Plant Pathology, University of Wisconsin, Madison, Wisconsin

Dong-Min Han  Division of Life Sciences, Wonkwang University, Iksan, South Korea

Kap-Hoon Han  Department of Pharmaceutical Engineering, Woosuk University, Samnye, South Korea

Steven D. Harris  Plant Science Initiative and Department of Plant Pathology, University of Nebraska, Lincoln, Nebraska

Kerstin Helmstaedt  Institute of Microbiology and Genetics, University of Goettingen, Goettingen, Germany

Kay Hofmann  Bioinformatics Group, Miltenyi Biotec GmbH, Cologne, Germany

Michael J. Hynes  Department of Genetics, University of Melbourne, Parkville, Victoria, Australia

Vinita Joardar  The Institute for Genomic Research, Rockville, Maryland

Jerzy Jurka  Genetic Information Research Institute, Mountain View, California

Vladimir V. Kapitonov  Genetic Information Research Institute, Mountain View, California

Nancy P. Keller  Department of Plant Pathology, University of Wisconsin, Madison, Wisconsin

Dimitrios P. Kontoyiannis  Department of Infectious Diseases, Infection Control and Employee Health, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Sven Krappmann  Department of Molecular Microbiology and Genetics, Institute of Microbiology and Genetics, Georg August-University Göttingen, Goettingen, Germany

Karin Lanthaler  Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Hui-Lin Liu  Department of Molecular Genetics, Ohio State University, Columbus, Ohio

Masayuki Machida  National Institute of Advanced Industrial Science and Technology, Higashi, Tsukuba, Ibaraki, Japan

Rama Maiti  The Institute for Genomic Research, Rockville, Maryland

Gregory S. May  Division of Pathology and Laboratory Medicine, University of Texas M. D. Anderson Cancer Center, Houston, Texas
Kevin McCluskey  School of Biological Sciences, University of Missouri, Kansas City, Kansas

Michelle Momany  Department of Plant Biology, University of Georgia, Athens, Georgia

William C. Nierman  The Institute for Genomic Research, Rockville, Maryland, USA, and Departments of Biochemistry and Molecular Biology, George Washington University School of Medicine, Washington, D.C.

Berl R. Oakley  Department of Molecular Genetics, The Ohio State University, Columbus, Ohio

Stephen A. Osmani  Department of Molecular Genetics, The Ohio State University, Columbus, Ohio

Gary A. Payne  Department of Plant Pathology, North Carolina State University, Raleigh, North Carolina

Miguel A. Peñalva  Department of Microbiologia Molecular, Centro de Investigaciones Biológicas, Ramiro de Maeztu, Madrid, Spain

Peter J. Punt  Department of Microbiology, TNO Quality of Life, Zeist, Utrecht, The Netherlands

Ana Ramón  Departamento de Biología Celular y Molecular, Universidad de la República, Montevideo, Uruguay

Jane F. Robens  USDA/ARS, Beltsville, Maryland

Antonis Rokas  Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee

Juan C. Sánchez-Ferrero  Department of Microbiología Molecular, Centro de Investigaciones Biológicas, Ramiro de Maeztu, Madrid, Spain

Claudio Scazzocchio  Institute de Génétique et Microbiologie, UMR CNRS 8621, and UMR CNRS 8080, Développement, Morphogénèse et Evolution, Université Paris XI-Sud, Orsay, France

William J. Steinbach  Department of Pediatrics and Department of Molecular Genetics and Microbiology, Duke University, Durham, North Carolina

Norio Takeshita  Department of Applied Microbiology, Institute for Applied Biosciences, University of Karlsruhe, Karlsruhe, Germany

Geoffrey Turner  Departments of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, United Kingdom

Jennifer R. Wortman  The Institute for Genomic Research, Rockville, Maryland
Xin Xiang  Department of Biochemistry and Molecular Biology, Uniformed Services University of the Health Sciences, Bethesda, Maryland

Jae-Hyuk Yu  Department of Food Microbiology and Toxicology, University of Wisconsin–Madison, Madison, Wisconsin

Jiujiang Yu  USDA/ARS, Southern Regional Research Center, New Orleans, Louisiana

Aimee K. Zaas  Department of Medicine, Duke University, Durham and National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina
Genomics of the Aspergilli
1

An Overview of the Genus Aspergillus

Scott E. Baker and Joan W. Bennett

CONTENTS
1.1 Early History and Taxonomy ................................................................. 3
1.2 Industry ................................................................................................. 6
   1.2.1 Koji .................................................................................................. 6
   1.2.2 Citric Acid and Other Aspergillus niger Products ......................... 7
   1.2.3 Aspergillus Secondary Metabolites ................................................... 8
      1.2.3.1 Lovastatin ................................................................................... 8
      1.2.3.2 Aflatoxin ................................................................................... 8
1.3 Aspergillus as an Animal Pathogen ....................................................... 9
1.4 Genetics and Aspergillus ................................................................. 10
1.5 Genomics and the Future ................................................................. 11
References ............................................................................................... 11

Few fungi are as important as members of the genus Aspergillus. This taxonomic group encompasses organisms whose characteristics are of high pathological, agricultural, industrial, pharmaceutical, scientific, and cultural importance. Superb agents of biodeterioration, aspergilli have been isolated from sources as varied as alligator nesting material, aviation fuel, Egyptian mummies, electrical fuses, plastic products, and old sauna boards. Indeed, this large and cosmopolitan group of molds is a major player in the ecosystem, involved in the degradation of a wide range of natural organic substrates, particularly plant materials. Aspergillus species are generalists in that they will grow and reproduce on many different carbon sources; they have an amazing nutritional flexibility. The diversity of enzymes and organic acids used in nutrition is complemented by the metabolic capacity to secrete numerous low molecular weight secondary metabolites believed to be important in ecological signaling. Because these molds can be found almost everywhere on the planet, degrading both natural and human-made substrates, Aspergillus and human history have been intertwined intimately for centuries.

1.1 Early History and Taxonomy

People have known about mushrooms and molds since the beginning of recorded history, but it is easier to find historical references to mushrooms—which are easily visible to the naked eye—than it is to find mention of mold. Microscopic fungi such as Aspergillus are usually referenced indirectly by their metabolic action as agents of rot and decay. The Greek physician Nicander of Colophon (ca. 185 BCE) wrote of “the evil ferment of the earth which men generally call by the name of fungi” (quoted in Ainsworth 1976), while one of the earliest books on the genus Aspergillus opened with the following words: “Historically, the Aspergilli, as part of the moldiness of things, have always been a factor in man’s
environment...” (Thom and Raper 1945). One possible mention of molds in antiquity concerns a description of ritual defilement in the Hebrew Bible. The noun tsara’at appears over 20 times, mostly in the book of Leviticus Chapters 13 and 14, where it is used to describe scaliness, rash, and discoloration of skin, garments, and the walls of houses. In the early Greek and Latin translations of the Bible, tsara-at was translated as a “plague of leprosy” Modern scholars have suggested that the word should be translated as “mold” or “mildew” (Heller, Heller et al. 2003). For example, in Leviticus Chapter 14, the priest is instructed on how to behave if the plague “be in the walls of the house” with, “greenish or reddish color.” Almost certainly, this is an early description of the kind of mold infestation that modern people associate with “sick building syndrome.”

The first known human exploitation of *Aspergillus* for beneficial purposes was for the transformation of rice, soybeans, and other plant foods to improve their palatability and to make them available for further fermentation by yeasts and bacteria. The domestication of *Aspergillus* for food production is thought to have originated in China close to 2000 years ago. Subsequently, similar food fermentations were adopted in Indonesia, Japan, Korea, and other parts of Asia. Koji is the Japanese name for mold-fermented grains and/or soybeans. Koji translates roughly as “bloom of the mold” and the modern Japanese ideograph shows conjoined symbols for “rice” and “flower.” The English language does not have a comparable word. These fermentations are now collectively known as koji processes and are the basis of robust commercial processes in the modern world (see section on industrial fermentations later). In koji, a filamentous fungus secretes a variety of enzymes as it invades and degrades its substrate from proteins to peptides and amino acids, and from starch to simple sugars.

The koji process was developed centuries before the microbiology was understood. Scientific study of *Aspergillus* and other molds began only when the microscope became available. Pier A. Micheli (1679–1737), an Italian botanist, first observed the distinctive spore-bearing structure of this fungus growing on a herbarium specimen. Micheli derived the name for the genus from the similarity in appearance between the microscopic anatomy of the spore-bearing structure and an aspergillum, the instrument used for sprinkling holy water in the Roman Catholic Church. Using an *Aspergillus* and a *Mucor* species, Micheli was the first to demonstrate asexual reproduction in molds by spores (Schaechter 1999). Other early mycologists such as Haller, Persoon, and Link described fungi that were probably *Aspergillus*, however, it is difficult, if not impossible, to interpret their identifications and nomenclature. It was DeBary, working in the early 1850s, who is usually credited with the beginning of modern studies. Most importantly, in examining herbarium specimens he realized that a fungus producing sexual cleistothecia and ascospores, named as *Eurotium herbariorum* by Link, was part of the same mycelium as an organism producing conidiophores and conidia and previously identified as *Aspergillus glaucus* (Ainsworth 1976). Thus, he had connected the perfect (now also known as the sexual, teleomorphic, or meiosporic) phase of the mold life cycle to the imperfect (also known as the asexual, anamorphic, or mitosporic) phase.

DeBary’s discovery has led to a nomenclatural conundrum. Currently, the naming of fungi is governed by the rules of Botanical Nomenclature. In contrast to all other codes of taxonomic nomenclature, wherein each species is known by one name and one name only, it is “legal” (and required) under the rules of Botanical Code for species of *Aspergillus* that have a sexual phase to have two names. Thus, the genetic model, *Aspergillus nidulans*, is also called *Emericella nidulans*. For scientists not trained in fungal systematics (which comprises most scientists) this scheme of dual nomenclature is confusing and impedes information retrieval (Bennett 1985). In the past, the major *Aspergillus* taxonomists defended the use of the asexual name (i.e., *Aspergillus*) to refer to both asexual and species asking, Why should the worker have to deal with multiple genera when a single one, and the oldest, will suffice? (Thom and Raper 1945). They answered their own rhetorical question with a suggestion for nomenclatural stability: “International recognition of *Aspergillus* for both ascosporic and conidial forms would constitute the logical and, we feel, correct solution” (Thom and Raper 1945). Many contemporary taxonomists do not agree; they recommend use of the teleomorph genus names for ascosporic species (e.g., *Emericella*, *Eurotium*, *Sterigmatocystis* (Samson, Hong et al. 2006), and recent revisions of the Botanical Code increasingly have privileged the sexual names over the asexual names. By definition, *Aspergillus* is a name referring to the asexual phase and, therefore, according to current rules of nomenclature, any *Aspergillus* with a sexual stage (teleomorph) no longer should be called *Aspergillus*. Indeed, in GenBank, *A. nidulans* is found listed as *Emericella nidulans*. Since the vast majority of molecular biologists who work with this...
organism call it “Aspergillus nidulans;” there is a disconcerting gap between common usage and the legalism of the Botanical Code.

In the early part of the twentieth century, Charles Thom and James N. Currie began with a study of oxalic acid production by *Aspergillus niger* and its relatives and ended with a comprehensive treatment of the black-spored aspergilli. Subsequently, with Mable Church, Thom studied a large and diverse taxonomic literature and attempted to bring the existing published work together with direct observations of molds grown under controlled conditions in the laboratory. Their studies resulted in the publication in the 1926 of a monograph entitled *The Aspergilli* (Thom and Church 1926). This work formed the basis for subsequent taxonomic treatments of the genus by Thom and Raper in 1945 (*A Manual of the Aspergilli*) and by Raper and Fennell in 1968 (*The Genus Aspergillus*). The centerpiece of *Aspergillus* taxonomy focuses on the morphology of the aspergillum (i.e., the spore bearing structure or conidiophore). It has a long stalk (stipe) that ends in a swollen apex. On the surface of the expanded apical region are a series of spore-bearing cells called phialides. Repeated mitotic division in the phialide nucleus yields a chain of asexual spores usually called conidiospores or conidia. The conidiospore varies in shape from spherical to elongate and may be smooth or echinulate. Conidia are extremely hydrophobic and are easily dispersed by air. Modern taxonomic works also rely on characteristics of the whole colony (color, size, presence or absence of sclerotia and pigments) when strains are grown under standardized culture conditions (Klich 2002; Samson, Hong et al. 2006). Figure 1.1 is an example of typical *Aspergillus* conidia and a conidiophore.

**FIGURE 1.1** *Aspergillus caespitosus* Raper and Thom. (a) Colonies of *A. caespitosus* grown on Czapek Yeast agar with 20% sucrose. (b) Conidiophores of *A. caespitosus*. (c) Light micrograph of *A. caespitosus* conidia, which are 3.5–4.5 μm in diameter. (d) 8000× scanning electron micrograph of *A. caespitosus* conidia. (Picture and micrograph credit: Maren A. Klich, USDA-ARS, New Orleans, Louisiana.)
While the unwieldy double nomenclatural scheme for the asexual and sexual states of fungi is still debated, the field of *Aspergillus* systematics has become more complex. In addition to morphological characters, profiling of molecular markers and “extrolites,” and extensive sequencing of conserved DNA sequences has increased the resolving power of *Aspergillus* systematics researchers. As increasing numbers of genome sequences for the aspergilli are generated, the ability further to resolve the characters used to define the lines of speciation will continue to increase. Nevertheless, unless we reach agreement on the type and number of morphological, metabolic, and genomic characters that define a species, the delineation of the species concept in this genus will remain contentious.

### 1.2 Industry

The economic footprint of *Aspergillus* is enormous; many different industrial processes have harnessed members of the genus. Aspergilli are important in the beverage, pharmaceutical, and enzyme industries. As mentioned earlier, the oldest processes are associated with traditional food fermentations used in a number of Asian cultures. These fermented foods and beverages include miso (soybean paste), shoyu (soy sauce), sake (rice wine), shochu (spirits), and yonezu (rice vinegar).

#### 1.2.1 Koji

For over 1500 years, *Asperillus oryzae*, *Aspergillus sojae*, and other closely related species have been used throughout Asia for koji food and beverage processes. Koji is a mixture of wheat, rice, or other grain, with or without soy (depending on what is being made), and the appropriate *Aspergillus* species. The mold mycelium grows through the substrate, putting out enzymes and organic acids across its cell walls. The secreted mold enzymes break down carbohydrates, proteins, and other organic polymers. In a sense, the mold enzymes “digest” the substrate, making it more flavorful for human consumption and available for further fermentations with yeasts and lactic acid bacteria. Koji processes for soy sauce and miso are primarily soy fermentations while sake is primarily a rice fermentation. Japanese foods, drinks, and condiments have become popular in western culture so people have adopted the Japanese names into English. It is important to point out that similar koji processes are used widely in other parts of Asia where they are known by different names. For example, miso is called “chiang” in China, “jang” in Korea, and “tao-tjo” in Indonesia and Thailand, while soy sauce has a similar multiplicity of names: “chiang-yi” in China, “kanjang” in Korea, and “kecap” in Indonesia (Reddy, Pierson et al. 1986). The *A. oryzae* strain chosen for genome sequencing by a Japanese consortium of government, academic and industry research groups is called RIB40 (National Research Institute Culture Stock; ATCC 42149). *A. oryzae* RIB 40 is a wild type strain similar to those used for sake brewing, but it also has a strong capacity for proteinase production, an attribute characteristic of strains used in soybean fermentations.

Not surprisingly, it was a Japanese scientist who adapted the koji process for broader commercial use. Called the “forgotten father of American biotechnology” Japanese-born Jokichi Takamine brought a modification of the koji process to the United States in the late nineteenth century (Bennett 1985). Using alcohol precipitation, he isolated a crude cocktail of extracellular mold enzymes of which the starch degrading amylases were the most important to his process. He first applied his commercialized enzyme to the whiskey trade while working in Peoria, Illinois. Labor unrest, intimations of arson, and ill health combined to make Dr. Takamine unsuccessful in this venture. However, he found commercial success for the diastatic (i.e., amylolytic) koji enzymes with treatment for indigestion in a formulation called Taka-diastase. He was issued a patent in 1894, thought to be the first U.S. patent for a microbial enzyme (Bennett 1988; Bennett 2001). An excerpt from the patent is given below:

> Be it known that I, Jokichi Takamine, a subject of the Emperor of Japan, residing at Peoria, in the county of Peoria and State of Illinois, have invented certain new and useful Improvements in Processes of Making Diastatic Enzyme…
An Overview of the Genus Aspergillus

The object of this invention is to prepare and manufacture diastatic enzyme, or soluble ferment in a concentrated form which possesses the power of transforming starch into sugar for use in various industries, by a process not hitherto practiced, and in a very economical and practical manner. My invention is based upon the utilization of the property possessed by certain fungi during their growth on proper media of producing diastatic enzyme.

In a current twist on the idea behind the use of Taka-diatase, amylase is included as an ingredient of some brands of dried infant grain cereals to aid in digestion. Today the modern enzyme industry, which Takamine was central in creating, crosses into many diverse fields including food and beverage preparation, detergents, as well as degradation of biomass for bioethanol and bioproducts (Archer and Peberdy 1997). Moreover, many of the enzymes used in these preparations are derived from and/or produced by Aspergillus species. In 1958, the Federal Food, Drug, and Cosmetic Act (FC&C) was passed by the U.S. Congress. It recognized that foods, organisms, and additives that had a long, safe history of common use in foods (e.g., yeast, salt, cinnamon) did not have to be tested individually for safety. Such substances are categorized as Generally Recognized as Safe (GRAS). GRAS status is a highly desired category in the food industry. Because of the long history of use of A. oryzae and A. niger for the preparation of human foods and beverages, new products from these fungi find easier approval than products from organisms that do not have a history of human food use. Genome sequence analyses of aspergilli and other filamentous fungi indicate that they have many hitherto undiscovered and unexploited enzymatic capacities. It is likely that the number of industrial and food processes that make use of Aspergillus-derived enzymes will grow.

1.2.2 Citric Acid and Other Aspergillus niger Products

Citric acid is the primary acidulant in the food and beverage industries, used in products such as candy, fruit juices, jams, jellies, and many other commodities. Further, in the pharmaceutical industry, citric acid is used as a preservative for stored blood and in the cosmetics industries it is used as a buffer and antioxidant. Industrial applications include detergent manufacture, electroplating, and leather tanning. In 1917, James N. Currie, a chemist working for the United States Department of Agriculture (USDA), published his research on the production of citric acid by A. niger (Currie 1917). In the decades since then, this species has been the biological factory on which the food and beverage industry bases its massive supply of citric acid. Currie took his process to Pfizer Corp., with the idea of eliminating the dependence of citric acid producers on the import of citrus fruits from Europe for the production of citric acid. Currie and his colleagues worked tirelessly for a number of years, improving the process and laying the foundation for modern fermentation technologies (Rodengen 1999). The original Pfizer process used surface culture methods but after the discovery of penicillin and the concomitant development of submerged culture systems for filamentous fungi, the citric acid process also was adapted for large tanks and submerged culture.

The citric acid process conducted today is much more efficient than the original Pfizer fermentation. Still, there is room for improvement. Low levels of manganese must be maintained in order to maintain the tight-pelleted morphology and high level of citric acid production. Additionally, strains in use today are the product of several rounds of random mutagenesis and have accumulated “collateral” damage in addition to mutations beneficial to the citric process. Current research in the field of A. niger citric acid production is focused on improving the process through use of the genome sequences available for the organism and reverse genetics approaches that have the potential to dramatically decrease the numbers of undesirable mutations (Karaffa and Kubicek 2003; Baker 2006). Today, almost the entire world supply of citric acid is still made by A. niger. Because this mold is nonpathogenic, nontoxicogenic, and has a long history of safe use in food and beverages, it has been classified as GRAS. In addition to citric acid production, A. niger has been harnessed to make gluconic and fumaric acids, organic acids of relatively minor economic importance.

It is in the production of enzymes used in the food industry that A. niger has seen some of its most versatile applications. A. niger-derived amylases are used for the hydrolysis of starch in bread and beer
production, and in the removal of sizing from fabrics. Invertase is used in confections, while pectinases are applied in the pretreatment of fruit juices to remove turbidity as well as for the reduction of cloudiness in wines. With the advent of industrial-style feed lots in animal husbandry, phytases have become a big market. They are used in animal feeds as an additive for liberation of phosphate from plant material. Finally, *A. niger* proteases are used in both meat tenderizing and to reduce elasticity of gluten proteins in bread (Godfrey and West 1996).

### 1.2.3 *Aspergillus* Secondary Metabolites

*Aspergillus* and other filamentous fungi are known for their ability to secrete a variety of biologically active chemical compounds. Formally, the term “secondary metabolite” is used to describe these low molecular weight, “nonessential” natural products usually produced after primary growth has stopped. These compounds include antibiotics, mycotoxins, immunosuppressants, and cholesterol-lowering agents. They are classified chemically by their biosynthetic origin as polyketides, nonribosomal peptides, sesquiterpenes, and so forth (Keller, Turner et al. 2005). In particular, the genes controlling the biosynthetic steps for lovastatin and aflatoxin, two model polyketides produced by *Aspergillus* species, have received considerable research attention. Genome sequence databases for aspergilli and other filamentous fungi have revealed an amazing diversity in the number of genes, predominately found clustered, that are putatively involved in secondary metabolism (Keller, Turner et al. 2005).

#### 1.2.3.1 Lovastatin

While *Aspergillus* infections and toxins detract from the quality of human life, some of the products of *Aspergillus* metabolism have made positive contributions to the human health as drugs. In an age when cardiovascular disease is a leading cause of death worldwide, the *Aspergillus*-derived statins are a case in point. Statins are drugs used to lower cholesterol levels. They bind to the active site of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR). HMGR catalyzes the conversion of HMG-CoA to mevalonate, an early and rate-limiting step in the cholesterol biosynthetic pathway. Statins are widely used for lowering serum cholesterol levels in the treatment of hypercholesterolemia, reducing the risk of heart attacks.

Statins were first discovered in Japan from *Penicillium citrinum* by Akira Endo. His compound was called compactin (also known as ML236 and mevastatin). However, the first statin to be developed for human drug use, lovastatin (also called mevinolin) came from *A. terreus*. The Merck Research Laboratories patented lovastatin in 1980 (Alberts-Schonberg, Alberts et al. 1980; Tobert 2003). In 1987, it was approved by the U.S. Food and Drug administration and became the first prescribed statin used in humans for lowering cholesterol. Lovastatin was a blockbuster drug for Merck, with over $1 billion of sales at its peak (Tobert 2003). The lovastatin biosynthetic pathway in *A. terreus* has been well described. This pathway was the first example of a polyketide synthetic pathway in which two fungal type I polyketide synthases work in combination to produce a product (Hendrickson, Devis et al. 1999).

#### 1.2.3.2 Aflatoxin

In 1960, Turkey X disease killed over 100,000 young turkeys and smaller numbers of duck and pheasant on farms in and around London in the United Kingdom (Blount 1961). Epidemiologists quickly found that all the dead poultry had been fed from the same lot of *A. flavus*-contaminated peanut meal (Sargeant, A. Sheridan et al. 1961). Not long later, chemists isolated a toxic principle from *A. flavus*-inoculated peanut meal, later named aflatoxin, and inferred that aflatoxin had poisoned these animals (Nesbitt, O’Kelly et al. 1962; vanderZijden, Koelensmid et al. 1962). Aflatoxin is actually a mixture of four or more closely related bisfururan polyketides. The major aflatoxins are called B₁, B₂, G₁, and G₂ based on their blue or greenish-blue florescence under ultraviolet light and their relative chromatographic mobility using thin layer chromatography. Aflatoxin B₁ is usually the major metabolite produced by toxigenic strains (Goldblatt 1969). It is also one of the most potent carcinogens known, inducing liver tumors in a wide variety of experimental animals. Further, epidemiological and molecular biological data implicate
An Overview of the Genus Aspergillus

Aflatoxin as a cause of human liver cancers. Acute human aflatoxin poisoning, however, is rare and usually only occurs when starvation forces people to subsist on moldy foods. In contrast, veterinary aflatoxicosis is a major problem, especially for animals fed in large feedlots (Eaton and Groopman 1994). Many agricultural commodities support growth and aflatoxin production by A. flavus, A. parasiticus, A. nomius, and other aflatoxigenic species (Bennett and Klich 2003). In the years since Turkey X disease ravaged the British poultry population, aflatoxin has become one of the best-studied mycotoxins. The genes for aflatoxin biosynthesis are linked together in a cluster; this pathway has become a model for studying fungal secondary metabolites, especially polyketides (Yu, Chang et al. 2004). Moreover, even after over 45 years of intense research, many questions remain unanswered: What evolutionary pressures keep the aflatoxin biosynthetic genes clustered? What are the genetic factors involved in aflatoxin biosynthesis that are not associated with the cluster? Where in the cell is aflatoxin produced? These same questions also could be asked about most of the secondary metabolites that are produced by Aspergillus and other filamentous fungi. It is hoped that microarray and other postgenomic studies will make it possible to gain insight into some of these unresolved mycological puzzles.

1.3 Aspergillus as an Animal Pathogen

The first five sentences from the publication of the A. fumigatus genome analysis serve as a summary of the impact this organism has on human health.

Aspergillus fumigatus is exceptional among microorganisms in being both a primary and opportunistic pathogen as well as a major allergen. Its conidia production is prolific, and so human respiratory tract exposure is almost constant. A. fumigatus is isolated from human habitats and vegetable compost heaps. In immunocompromised individuals, the incidence of invasive infection can be as high as 50% and the mortality rate is often about 50%. The interaction of A. fumigatus and other airborne fungi with the immune system is increasingly linked to severe asthma and sinusitis (Nierman, Pain et al. 2005).

British physician John Hughes Bennett is credited with the first published description of an Aspergillus infection, an aspergilloma (“fungus ball”) in 1842. Subsequently, it has been learned that most aspergillomas are caused by A. fumigatus, which is a thermotolerant species often resident in compost heaps. (Brakhage and Langfelder 2002). It is one of the most common airborne fungi, and humans and other animals regularly inhale numerous conidia. In healthy organisms, the respiratory tract eliminates these spores. For decades A. fumigatus was considered a weak pathogen, associated mostly with allergic conditions such as “farmer’s lung” and bronchopulmonary aspergillosis (Latge 1999). The first case of invasive human aspergillosis, in an immunocompromised patient, was made in 1953. Over the subsequent decades there has been a dramatic rise in the percentage of fungal infection by Aspergillus species. The vast majority of these infections are due to A. fumigatus. Systemic aspergillosis infections are also caused by A. flavus, A. nidulans, and, more rarely, other members of the genus capable of growing at 37°C. All of these species are usually nonpathogenic. They become opportunistic pathogens when confronted with an individual with a compromised immune system. In particular, bone marrow and organ transplant patients; lymphoma and leukemia patients receiving chemotherapy; and people receiving steroid treatments are at risk of systemic aspergillosis.

In the medical mycology literature, the term “aspergillosis” is used broadly to describe a spectrum of disease states that include both localized and systemic conditions. Allergic bronchopulmonary aspergillosis usually is found associated with asthma and cystic fibrosis. Some patients experience little permanent loss of respiratory function while others develop irreversible, obstructive lung diseases. Another form of localized aspergillosis occurs in patients with preexisting lung cavitations (e.g., tuberculosis patients). Aspergillus hyphae grow in the lung cavity and form brownish masses usually called aspergillomas (“fungus balls”). There are no good treatments for systemic aspergillosis. Unfortunately, amphotericin B, a compound known for its high level of negative side effects, is not well tolerated by immunocompromised patients yet it remains one of the drugs of choice (Denning 1996; Denning 1998). Itraconazole is
another antifungal agent active against *Aspergillus* species, but as with amphotericin B there are major
drawbacks (Latge 1999). Mortality from invasive aspergillosis remains extremely high. Thus, with the
rise in the number of cases of systemic aspergillosis, there has also been an increase in *A. fumigatus*
research. The *A. fumigatus* genome project has served as the focal point of the development of the growing
international *A. fumigatus* research community (Nierman, Pain et al. 2005).

In a recent intersection of popular culture and medical mycology, research supported by the National
Geographic Society indicated that *A. flavus* and other mold species could be isolated from ancient mum-
mies. The identification of these potentially disease and toxin-producing molds has reactivated stories of
“the mummy’s curse” also known as “King Tut’s curse,” which have circulated since 1922 when Lord
Carnavon, a British sponsor of archeological digs, died shortly after his involvement in the opening of
King Tutankhamun’s tomb. Headlines such as, “Egypt’s King Tut Curse caused by tomb toxins?” have
captured the public imagination and are circulating widely on the internet (National Geographic, 2005,

1.4 Genetics and *Aspergillus*

*Aspergillus nidulans* has been, and continues to be, one of the most important model organisms in eukary-
otic genetics. The father of *Aspergillus* genetics, Guido Pontecorvo, began research on *A. nidulans*
genetics over 50 years ago. Pontecorvo is credited with the discovery of the parasexual cycle. In this
process, two haploid nuclei fuse and form a mitotic diploid (Roper 1952; Pontecorvo and Roper 1953).
In a small percentage of nuclei in the resultant diploid organism, recombination occurs during mitosis.
The parasexual diploid can be analyzed for the appearance of diploid mitotic recombinants and/or haploid
segregates created through the application of haploidization agents that break down the diploid nuclei.
When differentially marked haploid mutants are put through this process, genetic analysis is possible in
the absence of meiosis (Pontecorvo and Kafer 1958). Parasexual cycle genetics has been most useful for
studying fungi whose sexual cycle is unknown, or may not exist, in species such as *A. flavus, A. niger,*
and *A. parasiticus* (Pontecorvo 1956; Papa 1973; Papa 1978; Bennett 1979; Bennett, Vinnett et al. 1980;
Bos, Debets et al. 1988; Debets, Swart et al. 1990). Additionally, in the decades before recombinant DNA
research, Pontecorvo recognized that parasexual cycle genetics had the potential to be a major tool for
human genetics. Indeed, during the 1960s and early 1970s, somatic cell fusions in tissue culture, followed
by haploidization, became a major strategy for localizing human genes to specific chromosomes
(Martinelli and Kinghorn 1994).

*A. nidulans* continues to be an important model in eukaryotic genetics and a veritable laundry list of
biological processes has been researched using a combination of sexual and parasexual approaches
(Smith, Pateman et al. 1977). For example, some of the first genetic studies on the cell cycle were
conducted during the 1970s under the leadership of Ronald Morris. He described a group of temperature
sensitive mutants blocked in different stages of mitosis (Morris 1975). Mutants affected in septation,
nuclear movement, interphase, and mitosis served as a treasure trove for the discovery of genes that affect
stages of the mitotic process and nuclear migration and have laid a foundation for a fertile area of research
that continues today. The first genetics based research on tubulins, the building blocks of microtubules,
was also done in the Morris lab using *A. nidulans* as the experimental organism (Sheir-Neiss, Lai et al.
1978; Morris, Lai et al. 1979). This research led to the discovery of γ tubulin (Oakley and Oakley, 1989).
*A. nidulans* has played a central role in the genetic study of, among other topics, fungal development,
both asexual and sexual, gene regulation by extracellular pH, nitrogen metabolism, and gene regulation
by nitrogen and carbon source.

Thirty years after the 1953 publication of the seminal Pontecorvo et al. description of genetics
in *A. nidulans* (Pontecorvo, Roper et al. 1953), the age of molecular genetic analysis for this organism
An Overview of the Genus Aspergillus

(and Aspergillus) was ushered in with the first report of transformation (Tilburn, Scazocchio et al. 1983). Two years later, in 1985, came the publication of the first A. nidulans homologous transformation, wherein TrpC+ and ArgB+ containing constructs were used to replace TrpC− and ArgB− loci (Miller, Miller et al. 1985). Twenty years later, the genome sequence of A. nidulans was published alongside manuscripts describing the A. fumigatus and A. oryzae genomes (Galagan, Calvo et al. 2005; Machida, Asai et al. 2005; Nierman, Pain et al. 2005).

1.5 Genomics and the Future

The paths of human and Aspergillus history have crossed many times over the last 2000 years and a number of species from this genus have become important experimental organisms. Genome sequence data is available for species that span the phylogenetic spectrum of the genus. At its most basic level the genome sequence of an organism can paint a high-level picture of the biology of an organism. More importantly, genome sequence enables large-scale, reverse genetic, global transcriptomic and proteomic experiments. The catalog of genes whose products are involved in the synthesis of primary and secondary metabolites is now known in each organism whose genome has been sequenced (An 2005). The simultaneous publication of three aspergilli genome manuscripts established Aspergillus as the leading fungal genus for comparative genomic studies. Scattered across the phylogenetic diversity of the Aspergillus genus are a number of organisms with active genome-sequencing projects. In addition, comparative genomics projects “clustered” around A. fumigatus (Neosartorya fischeri) and A. oryzae (A. flavus) are ongoing while a project to sequence close A. niger relatives (A. aculeatus and A. carbonarius) was recently initiated. The analyses of the genomic data “clustered” around specific species will be crucial to the identification of gene regulatory elements and will facilitate improved gene annotation.

Aspergillus researchers must be able to put available genome sequence to work to understand the biology of this important genus. Indeed, our current ability to generate nucleic acid sequence and proteomic data surpasses our ability to understand these data and to apply them to basic and applied problems. We have more information than we can use. To answer these challenges, the Aspergillus research community requires organization, coordination, and sharing of experimental design and data. We must continue to work together and communicate effectively between research groups. Furthermore, we must not be content with the genomic sequences currently available, but we also must continue to advocate genome sequencing deeper into each major clade of the genus. Finally, we need to attract a large coterie of young scientists to Aspergillus research to ensure that these versatile molds can continue to contribute to our understanding and exploitation of the unique features of fungal physiology. Using the resources currently in hand and strongly advocating coordination of research and development of new resources, we will move forward into a new “golden age” of Aspergillus research.

References


An Overview of the Genus Aspergillus


A First Glance into the Genome Sequence of Aspergillus flavus

Gary A. Payne, Jiujiang Yu, William C. Nierman, Masayuki Machida, Deepak Bhatnagar, Thomas E. Cleveland, and Ralph A. Dean

CONTENTS
2.1 Introduction ............................................................................................................. 15
  2.1.1 Ecology of Aspergillus flavus ................................................................. 15
  2.1.2 Aspergillus as a Plant Pathogen .............................................................. 16
  2.1.3 Aspergillus as an Animal Pathogen ......................................................... 17
  2.1.4 Aspergillus oryzae as an Ecotype of Aspergillus flavus ......................... 17
2.2 Structural Genomics ............................................................................................... 18
  2.2.1 Sequencing and Annotation ....................................................................... 18
  2.2.2 Physical Structure ..................................................................................... 18
  2.2.3 Database .................................................................................................... 19
2.3 Comparative Genomics ......................................................................................... 19
  2.3.1 Physical Structure ..................................................................................... 19
  2.3.2 Unique Genes and Features ...................................................................... 19
  2.3.3 Population Analysis .................................................................................. 19
2.4 Functional Genomics ........................................................................................... 20
  2.4.1 Design of Affymetrix Arrays ..................................................................... 20
  2.4.2 Integrated Database for Functional Analysis ............................................ 20
2.5 Summary and Prospects ...................................................................................... 21
Acknowledgments ........................................................................................................ 21
References ..................................................................................................................... 21

2.1 Introduction

2.1.1 Ecology of Aspergillus flavus

Aspergillus flavus is a competitive saprobe, a plant pathogen, and an animal pathogen; thus, its ecology encompasses all of the known trophic phases of the genus Aspergillus. In addition, A. flavus has a rich profile of secondary metabolites, including several mycotoxins. Known mycotoxins produced by A. flavus include cyclopiazonic acid, aspertoxin, aflatem, aspergillic acid, and aflatoxins. Aflatoxins, which are both toxic and carcinogenic, accumulate in food and feeds. Because of health concerns related to ingestion of aflatoxin, a large research community has developed with the focus of understanding the population biology, ecology, and secondary metabolism in this fungus. As a result, the population structure and ecology of A. flavus is better described than for any other Aspergillus species and aflatoxin biosynthesis is one of the best-characterized pathways of secondary metabolism in filamentous fungi. Couple this with the recently available genome sequence and whole genome DNA microarrays, and
A. flavus emerges as an appealing model organism to study the biology, ecology, and pathogenicity of the genus. 

A. flavus belongs to the subgenus Circumdati, section Flavi. Species within this section are most common in the subtropical and warm temperate zones. A. flavus is most frequently found between the latitudes of 26 and 35 degrees. Although most commonly isolated from cultivated soils, it can be found in several biomes including forest soils.2 A. flavus, and its close relative A. parasiticus, are the most commonly occurring aflatoxin-producing species. While both species can be found on developing seeds of cotton, corn, peanuts, and nut trees, A. flavus is the more prevalent species. Aspergillus parasiticus has a slightly lower optimum temperature for growth and aflatoxin production, which may explain why it is found on peanuts seeds and rarely on aerially produced seeds.3 It is important to note that while these species can and often do colonize plant seeds, they are most often associated with the soil where they must compete with the microflora and microfauna of soil.

The aflatoxin biosynthetic cluster is highly conserved in these two species, and the regulation of the aflatoxin pathway appears to be very similar. Gene order in the biosynthetic cluster is conserved, and the cluster genes share 96% DNA identity. The most striking difference is the absence of the aflatoxin cluster genes norB and cypA in A. flavus, two genes required for the biosynthesis of the G family of aflatoxins.6,7 With the exception of one clade, which is discussed later, A. flavus produces only the B family of aflatoxins while A. parasiticus produces both B and G aflatoxins. The biosynthesis and regulation of aflatoxin have been studied extensively in these two strains and a review of these studies is presented by Yu et al. in this book.

A. flavus populations are genetically diverse,8 and there are two clades that are described within the species. Strains within these clades are distinguished based on the size of sclerotia and referred to as either L (for large sclerotia) or S (for small sclerotia). Phylogenetic analysis has shown that clades L and S represent deep divergence within A. flavus and that each group is monophyletic.9 Most L strains produce B aflatoxins and none produce G aflatoxins, whereas some of the S strains produce B and G aflatoxin.9 In general S strains produce higher amounts of aflatoxin. L strains produce abundant conidia and sclerotia that are usually larger than 400 μm in diameter.10,11 S strains produce fewer conidia and numerous sclerotia usually smaller than 400 μm. The aflatoxin biosynthetic gene clusters of the L and S groups of A. flavus are 99% identical at the nucleotide level.4 Interestingly, these two strains appear to differ in their geographic distribution,11–13 although the reasons for these differences are not clearly understood.

Aflatoxin biosynthesis has been most extensively studied in L strains, and in the United States the L strain is the predominant strain. For these reasons a representative L strain (NRRL 3357) that has been used for over 30 years in field and laboratory experiments14–18 was chosen for the whole genome-sequencing project.

2.1.2 Aspergillus as a Plant Pathogen

A. flavus is an opportunistic pathogen of developing seeds, particularly corn, peanuts, cottonseed, and tree nuts. It is not an aggressive pathogen, and successful colonization often requires a host plant whose defenses are compromised, often by drought and temperature stress.19 Most often associated with wounded kernels, A. flavus can invade adjacent intact maize kernels and contaminate these kernels with aflatoxin. Direct infection of maize kernels after inoculation of silk tissue with A. flavus also occurs.19 Even though A. flavus is an opportunistic pathogen, it is well adapted to colonizing seeds. Few other fungi infect and colonize developing seeds susceptible to A. flavus. Such adaptation implies that A. flavus has a unique suite of genes necessary for seed colonization or, alternately, superior competitive abilities. Recent research suggests that fungi in section Flavi are successful in the colonization of peanut seeds because they are superior competitors. Horn found that A. flavus and some other members of section Flavi are dominant fungi on wounded peanut seeds even though they represent <1% of the soil flora.20 From these observations, he has proposed that species in section Flavi have high competitive saprophytic ability and thus can out-compete the other saprophytes.

In general, only a few seeds of a commodity are infected with A. flavus. Unfortunately, even limited colonization by the fungus can result in concentrations of aflatoxin that exceed guidelines established
A First Glance into the Genome Sequence of Aspergillus flavus

by the Food and Drug Administration. In developed countries, aflatoxin remains predominately an economic problem resulting from the inability to sell the crop nationally or internationally. A guideline of 20 parts aflatoxin per billion parts of food or feed substrate (ppb) is the maximum allowable limit imposed by the U.S. Food and Drug Administration for interstate shipment. In developing countries, where detection and decontamination policies are impractical, aflatoxin contamination is predominately a food safety issue. Aflatoxins have been associated with liver cancer and many veterinary toxic syndromes.21–23 Aflatoxin contamination is a recurring problem in Kenya, and in 2005 at least 125 people were killed by consuming aflatoxin-contaminated maize (http://www.cdc.gov/mmwr/preview/mmwr.html/mm5334a4.htm).

2.1.3 Aspergillus as an Animal Pathogen

Human infections caused by members of the genus Aspergillus are another emerging health problem in developed countries. Just as A. flavus is an opportunistic pathogen of plants, it can also be an opportunistic pathogen of animals. New medical treatments for diseases that require suppression of the immune system predispose patients to an increased risk of succumbing to aspergillosis, a condition that encompasses a variety of diseases caused by members of the genus Aspergillus, including invasive aspergillosis, pulmonary aspergilloma, allergic bronchopulmonary aspergillosis, and others.24 Nosocomial infections represent a significant threat for these individuals.25 A. flavus is the second leading cause of aspergillosis in humans and the leading causative agent of chronic indolent invasive sinonasal infection in immunocompetent patients.24 Mortality from Aspergillus infections is high because only a limited number of antifungal drugs are available and resistant strains have been identified.25 In a recent surveillance project that monitored concentrations of Aspergillus spores in a large hospital, A. flavus colony forming units ranked third behind A. niger and A. candidus.27

2.1.4 Aspergillus oryzae as an Ecotype of Aspergillus flavus

An interesting debate in the Aspergillus community is whether A. flavus and A. oryzae represent different species or ecotypes of the same species. Unlike A. flavus, which is a plant and animal pathogen, A. oryzae is commonly used in food fermentation and has GRAS (Generally Regarded as Safe) status. A. oryzae is not considered to be either a plant or animal pathogen, nor do strains produce aflatoxin. It has been used for centuries in the food fermentation industry and is known as a koji mold because of its use in the production of oriental koji products such as miso.28

While these two species have been traditionally considered taxonomically distinct, there always have been unease about the taxonomic uniqueness of A. oryzae. The morphological differences between the two strains can be subtle;30–31 however, consistent morphological features allow these strains to be distinguished from one another.32 A. oryzae strains are described as more floccose than A. flavus and they tend to have a pale yellow color. Also, A. oryzae has longer conidiophores with thinner, smoother walls than found in A. flavus. These conidiophores bear conidia of A. oryzae that are larger than those for A. flavus.31 Wicklow31 states that these traits are rarely observed in wild populations of A. flavus. Thom and Raper,30 on the other hand, reported that within the yellow Aspergillus group one can find any variant from dwarf and dark green characteristic of A. parasiticus to the longest stalked and palest yellow characteristics of A. oryzae. The review by Wicklow provides a nice historical perspective on these two species. He presents the argument that selection during the process of preparing the koji starter culture selects for artificial forms of A. flavus and that over time this selection leads to the phenotype described as A. oryzae. Such a common selection process could lead to the grouping of strains that show varying degrees of cultural and morphological relatedness. If this is true one may expect these strains to have an altered metabolism and possibly to have lost some of the traits that make it competitive in natural ecosystems. This selection scheme may explain why A. oryzae is not commonly found in nature.

Molecular evidence has long supported the contention that A. flavus and A. oryzae are very similar if not the same species. Early research with DNA reassociation experiments predicted 100% DNA complementarity between the two species.33 More recent studies using isozyme profiling34 and tests for DNA polymorphism have failed to clearly distinguish between the two species.35–37
Careful phylogenetic analysis of these two strains further supports the hypothesis that *A. oryzae* is a domesticated *A. flavus*. *A. flavus* is polyphyletic. If *A. oryzae* is a domesticated form of *A. flavus*, and the domestication occurred several times, one may expect *A. oryzae* to also be polyphyletic. In fact, *A. oryzae* is polyphyletic. Further, Chang et al. found *A. oryzae* strains RIB 40 (the sequenced strain), SRRC 2044, SRRC 2098, and SRRC 2103 to be in a clade that also contains *A. flavus* isolates.

No strain of *A. oryzae* has been reported to produce aflatoxin, but many contain the entire aflatoxin biosynthetic cluster. Tominaga et al. compared several strains of *A. oryzae* with *A. flavus* and found that compared to the *A. flavus* sequence the *A. oryzae* genes contained deletions, frameshift mutations, and base-pair substitutions. Just as has been observed in *A. flavus* many strains of *A. oryzae* have deletions within the 75 kb cluster region.

Now that the genome sequence of *A. oryzae* is available to the public and a whole genome sequence of *A. flavus* is near completion, we have the opportunity to examine these two fungi in detail. These studies provide a rare opportunity to look at the effect of domestication on genome organization and structure of a filamentous fungus. A comparison of the genomes of these two fungi will likely reveal information on changes that have occurred during the domestication of *A. oryzae*, and help identify pathogenicity factors in *A. flavus*.

### 2.2 Structural Genomics

#### 2.2.1 Sequencing and Annotation

The USDA/NRI Microbial Genome Sequencing Project provided funding for whole genome sequencing of *A. flavus*. Additional funds for fine finishing to close small gaps were provided by the USDA/ARS/SRRC. The project has greatly benefited from the advice of a steering committee composed of Dr. Charles Woloshuk, Dr. Greg May, Dr. Nancy Keller, Dr. Heather Wilkinson, Dr. Joan Bennett, Dr. Deepak Bhatnagar, and Dr. T. E. Cleveland.

Sequencing to 5x coverage was done at The Institute for Genomic Research (TIGR). A multiple library strategy with different insert sizes was used to attain maximal genome coverage and maximal linkage of the assembled contigs. A combination of 3–4 kb and 10 kb insert size libraries and a 50 kb linking library were used. Automated annotation of the genome was done at TIGR using tools trained on the available genomic sequence of *A. oryzae* as well as *A. flavus* and *A. oryzae* ESTs. A total of 35,959 *A. flavus* ESTs and 33,930 *A. oryzae* ESTs were used for annotation. Sequence reads are available at NCBI and at the website www.aspergillusflavus.org.

#### 2.2.2 Physical Structure

The DNA scaffold sizes for the sequenced genome range from 4.5 Mb to 1.0 kb, and over 75% of the genome is represented in the 10 largest scaffolds. More importantly, over 99.6% of the genome is represented in the 17 largest scaffolds. Early linkage group assignment by K. E. Papa using parasexual analysis and subsequent karyotypes of *A. flavus* strains suggested that *A. flavus* had eight chromosomes. An optical map was developed for *A. oryzae*, which showed that its genome is organized into eight chromosomes. Because of the high degree of DNA correspondence between *A. flavus* and *A. oryzae* we were able to place the *A. flavus* genome on the physical map of *A. oryzae* and in effect assign a physical location to 99.6% of the *A. flavus* predicted genes. The 16 largest scaffolds of *A. flavus* map to the 16 predicted arms of the 8 chromosomes.

Since the final annotation of the *A. flavus* genome is not complete, the data presented here represent a preliminary prediction of gene number. The estimated genome size of 36.8 Mb is similar to that for *A. oryzae* (36.7 Mb), but larger than that for *A. nidulans* or *A. fumigatus*. Further, *A. flavus* and *A. oryzae* are enriched in genes for secondary metabolism. *A. flavus*, for example, is predicted to have 34 polyketide synthases, 24 nonribosomal peptide synthases, 77 ABC transporters and greater than 122 cytochrome p450 enzymes. We have not yet manually annotated the cytochrome p450 enzymes.
2.2.3 Database

A web browser based on GBrowse, a generic genome browser produced by the GMOD project has been developed at North Carolina State University to display annotations on the \textit{A. flavus} genome. Currently, it allows users to see regions of the \textit{A. flavus} genome that have BLAST matches to genes, proteins, and genomic sequence of other \textit{Aspergillus} species. It also displays alignments of ESTs, regions of repetitive DNA sequences, and the locations of predicted \textit{A. flavus} genes together with annotation for these genes. This annotation includes a brief description of the predicted gene function, GO terms, and Pfam and InterPro matches. The web browser allows a user to visualize the genome sequence, and see, for example, whether a gene they are interested in has any supporting EST evidence, which genes are located close by, and whether it is in a region with any repetitive DNA. The Gbrowse system supports third-party annotation, which will allow other members of the \textit{A. flavus} community to contribute to the website by adding information on genes or gene families for which they have specialist knowledge. Additional functionality will be added to this browser as more sequence information becomes available. Links to the web browser and to other information on the sequencing project can be found at www.aspergillusflavus.org.

2.3 Comparative Genomics

2.3.1 Physical Structure

Due to the high degree of similarity between \textit{A. flavus} and \textit{A. oryzae}, we have begun a more extensive comparison of these two genomes. A body of information is emerging to suggest that these strains are indeed very similar if not members of the same species. There is a high degree of DNA correspondence and synteny between these two species. Each species has an entire aflatoxin biosynthetic cluster and in both species it is located near the telomere on chromosome III. The cluster in \textit{A. flavus} resides within 70 kb of the end of the DNA scaffold containing the cluster. This contig, which is syntenic with chromosome III of \textit{A. oryzae}, is slightly longer in \textit{A. flavus}, likely due to differences in assembly. The major physical difference observed between the two species is a translocation event in \textit{A. flavus}. There has been an exchange of DNA between chromosomes II and VI. The break site is associated with a family of uncharacterized repeat elements.

2.3.2 Unique Genes and Features

In addition to the translocation event, there are other differences between the two genomes. Each species has approximately 300 unique genes, and most of these are in regions where the synteny between the two species breaks down. Most of the insertions and deletions involve small clusters of genes, but one deletion in \textit{A. flavus} spans 250 kb. These regions of apparent instability of the genome are most often associated with repetitive DNA sequences. We have not extensively analyzed the repetitive DNA in \textit{A. flavus}, but our initial examination has revealed that \textit{A. oryzae} has more repetitive DNA than \textit{A. flavus}. This observation is interesting because it was reported that \textit{A. oryzae} has less repetitive DNA than \textit{A. nidulans} or \textit{A. fumigatus}. While it is unclear what role repetitive DNA may have had in the domestication of \textit{A. oryzae}, the differences in the content between the two fungi likely reflect differing selections pressures applied to these species. Retroelements and retrotransposons were identified in both species. These included a Ty-3/group of retrotransposons and the Fot5-like transposons. Overall there appears to be more copies of each of these transposable elements in \textit{A. oryzae} than in \textit{A. flavus}.

2.3.3 Population Analysis

An intriguing question in this comparison is whether the unique features that we observed between the two sequenced strains represent true species differences or if they are within the natural variation of the population. Because these two fungi are so similar, we can use a directed approach to address this question. We are currently examining four additional strains each of \textit{A. flavus} and \textit{A. oryzae}. Our approach
is to use comparative genome hybridization (using both Affymetrix GeneChip arrays and oligo-based arrays) to identify gene differences among these additional eight strains. The high degree of similarity between the two fungi will allow us to use either the available whole genome oligo arrays prepared for *A. oryzae* or the two whole genome arrays developed for *A. flavus* that are described in the next section. For these comparisons we have intentionally included the predicted unique *A. oryzae* genes in RIB40 on the whole genome Affymetrix GeneChip (described later) and the oligo whole genome array that is being developed. Results from these studies will allow us to predict the prevalence of these predicted genes in *A. flavus* and *A. oryzae*.

### 2.4 Functional Genomics

#### 2.4.1 Design of Affymetrix Arrays

Gene expression analysis using DNA microarrays is a powerful tool to dissect complex regulatory circuits as one can simultaneously measure the expression of thousands of genes. It also can be used to identify new genes involved in a pathway or regulatory network. To further our understanding of the biology and pathogenicity of *A. flavus* and the regulation of secondary metabolism, a whole genome Affymetrix GeneChip was developed for *A. flavus*. The array contains 12,834 predicted genes and 397 predicted antisense transcripts of *A. flavus*. No biological function of any of the antisense transcripts is yet known, although one of the antisense transcripts overlaps the 5' end of the pathway regulatory gene, *afl*.

All of the putative antisense transcripts have ESTs in either the *A. flavus* or *A. oryzae* libraries. Because this array will be employed in plant and animal pathogenicity studies, we included 8895 maize seed genes and 25 human and mouse innate response genes. To determine whether there are transcriptionally active regions within the aflatoxin biosynthetic cluster that we have not predicted, we have also tiled across the 26 intergenic regions of the cluster (both strands) with one probe approximately every 25 bp. This array will be a powerful tool to help understand the pathogenicity of *A. flavus* and the regulatory elements involved in secondary metabolism. We also included over 300 genes from *A. oryzae* that appear to be absent in the sequenced strain of *A. flavus*.

Our initial experiments with these multiple species arrays indicate that they are reliable for measuring gene expression in pure cultures and in host parasite interactions. We have used these arrays to monitor gene expression of *A. flavus* during aflatoxin biosynthesis in defined media and during infection of developing maize seeds. A parallel study comparing this array with a 5002 element cDNA array showed the same expression pattern profile for the aflatoxin biosynthetic genes when *A. flavus* was grown on defined media at conducive and nonconducive temperatures for aflatoxin production. Further, the profile of aflatoxin gene expression six days after inoculation of dough stage kernels was similar to that observed on the Affymetrix slides in the conducive temperature experiment. Thus, the presence of maize seed DNA did not appear to affect the detection of the aflatoxin pathway genes. In none of the experiments with the Affymetrix arrays did we observe species cross hybridization. These initial observations show that these multispecies arrays will be very powerful tools for studying the complex ecology and metabolism of *A. flavus*.

#### 2.4.2 Integrated Database for Functional Analysis

To facilitate the research efforts in the *A. flavus* community we are developing a database to integrate multiple categories of data. This database resource will serve two important functions: (1) it will provide a platform for the deposition of data from individual experiments; and (2) it will permit the ready analysis of composite data from all experiments enabling researchers to mine a larger data set. It will include phenotypic measurements, gene expression data from microarrays, and metabolic profile information, and will be flexible enough to allow the addition of new types of measurement in the future. Users will interact with the database through a web-based interface and will be able to: describe experiments; upload data gathered during those experiments; run analyses on the data; select and download raw data; select and download the results of analyses. This database can be accessed from www.Aspergillusflavus.org.


A First Glance into the Genome Sequence of Aspergillus flavus

2.5 Summary and Prospects

*Aspergillus flavus* is a cosmopolitan fungus with the ability to colonize diverse ecological niches. It is an extremely competitive saprotroph and an opportunistic pathogen of both plants and animals. What mechanisms have lead to this adaptability? Has *Aspergillus flavus* accumulated a battery of tools that allows it to colonize diverse ecosystems? If so, what are these gene products? Does *Aspergillus flavus* have a “plastic genome” that allows individuals or populations of the fungus to adapt rapidly to new niches? If this is true, are there signatures of this plasticity in the genome? These are the types of questions that we hope can be addressed now that the complete genome sequences of *Aspergillus flavus* and *A. oryzae* are available. The ability to compare two fungi with different ecological niches yet with a high degree of synteny should allow us to focus on unique difference between the fungi. An added advantage is the availability of whole genome arrays for *Aspergillus flavus*, which will allow us to compare gene expression between the two fungi. Findings from these studies will have broad implications for understanding the genus as a whole.

Acknowledgments

The authors express their sincere thanks to Beth Prichard at North Carolina State University and Dr. Jennifer Wortman at The Institute for Genomic Research for their assistance in the bioinformatic analysis. Sequencing of the genome was funded by the USDA/NRI Competitive Grants Program (USDA-CSREES Award Number: USDA-2003-05428) through a grant awarded to Payne and Dean. Supplemental funding was provided by the USDA/ARS/SRRC in New Orleans. The integrated database for functional analysis is being funded by USDA/NRI/CGP 2006-35604-16666.

References


A First Glance into the Genome Sequence of Aspergillus flavus


3

A Comparative View of the Genome of Aspergillus fumigatus

Natalie D. Fedorova, William C. Nierman, Geoffrey Turner, Vinita Joardar, Rama Maiti, Michael J. Anderson, David W. Denning, and Jennifer R. Wortman

CONTENTS

3.1 Introduction ............................................................................................................. 25
  3.1.1 Invasive Aspergillosis Is a Significant Human Health Problem ...................... 26
  3.1.2 Innate Immune System and Invasive Aspergillosis Prevention and Resolution .................................................. 26
  3.1.3 Application of Comparative Genomics to the Study of Aspergillus fumigatus .................................................. 27
3.2 Genome Sequences ......................................................................................................... 27
  3.2.1 Aspergillus fumigatus Af293 ........................................................................ 27
  3.2.2 Neosartorya fischeri NRRL181 (Aspergillus fischerianus) .................................................. 28
  3.2.3 Aspergillus clavatus NRRL1 ........................................................................ 28
  3.2.4 Sequence Divergence, Synteny, and Chromosomal Location ......................... 28
3.3 Virulence Genes .......................................................................................................... 29
  3.3.1 Genes Implicated in Virulence ........................................................................... 30
  3.3.2 Function of Virulence Genes ........................................................................... 30
  3.3.3 What Is Virulence? ......................................................................................... 32
3.4 Secondary Metabolite Biosynthetic Gene Clusters ...................................................... 32
  3.4.1 Clusters in Aspergillus fumigatus Af293 ......................................................... 32
  3.4.2 Associating Clusters with Products .................................................................... 32
  3.4.3 Evolution of Clusters ...................................................................................... 34
3.5 Sex and Sex Genes ....................................................................................................... 35
  3.5.1 Mating loci in Aspergillus fumigatus, Aspergillus clavatus, and Neosartorya fischeri .................................................................................. 36
3.6 Status of Aspergillus fumigatus Genome Annotation ............................................... 37
3.7 Summary and Prospects .............................................................................................. 38
References ......................................................................................................................... 38

3.1 Introduction

The genus Aspergillus was named by P. A. Micheli in 1729 after a holy water sprinkler, or aspergillum, which resembled the genus-characteristic conidia-forming structure of these fungi.\textsuperscript{1} Aspergillus is an extremely diverse and widely distributed genus of filamentous ascomycete fungi. It includes over 200 species of mostly asexual fungi found ubiquitously in soil as well as in forage products, food, dust, organic debris, and decomposing vegetation.\textsuperscript{2} Being supreme opportunists, the aspergilli have adapted to various chemical, physical, and biological stresses and have repeatedly changed their lifestyle and reproductive mode in the course of evolution. While most of them are thought to be saprophytes, a surprising
number of species are able to infect wounded plants and animals. The advent of immunosuppressive agents and other medical advances created a new biological niche for aspergilli, the immunocompromised human host.

3.1.1 Invasive Aspergillosis Is a Significant Human Health Problem

The human respiratory tract is constantly exposed to the opportunistic fungal pathogen *A. fumigatus*—a prolific producer of asexual spores (or conidia) and a ubiquitous inhabitant of vegetable matter composts and human habitats.\(^3\)\(^-\)\(^5\) *Aspergillus fumigatus* is the cause of highly lethal invasive disease in people with compromised immune function, several million of whom are at risk each year. While the interaction of *A. fumigatus* spores with the human respiratory mucosa is understood to an extent, the basic biology of the organism has only recently been actively investigated by more than a few groups. Its thermotolerance, its apparent loss of sexuality, its numerous secondary metabolic pathways producing toxic (e.g., gliotoxin) or pharmacologically useful secondary metabolites (e.g., fumagillin), its mechanisms of antifungal drug resistance, and its remarkable intrinsic growth rate are poorly understood. *A. fumigatus* is the cause of highly lethal invasive disease in people with compromised immune function, several million of whom are at risk each year. IA is the most common manifestation of *A. fumigatus* infection in immunocompromised patients, having incidence rates of 10–15% in allogenic bone marrow recipients, 7% in acute leukemia, and 40% in inherited chronic granulomatous disease.\(^6\) IA undermines the success of advanced and expensive conventional therapies such as bone marrow and organ transplantation, and cancer chemotherapy. The therapeutic management options for invasive aspergillosis are limited and even with antifungal therapy the mortality rate is approximately 50%. Current antifungal treatments are not generally rapidly fungicidal to aspergilli and have other limitations such as poor bio-availability, some toxicity, and interaction with other treatment regimens. There is also a problem of emerging resistance among the aspergilli to existing antifungals.\(^7\)\(^,\)\(^8\) The situation is further complicated by the lack of reliable diagnostic criteria that can delay unambiguous diagnosis, further compromising clinical outcomes.\(^9\) In addition to invasive disease, *A. fumigatus* causes allergic disease in the form of allergic bronchopulmonary aspergillosis and fungal sinusitis and may be of major significance in many adults with severe asthma.\(^9\)\(^-\)\(^12\) Although this fungus is clinically important, it has not been the subject of intensive biological investigation. It is necessary to study the biological and pathogenic processes of this fungus as such studies will lead to an enhanced understanding of virulence and ultimately to the development of novel and effective therapies for IA.

*A. fumigatus* clinical isolates vary in pathogenicity, although most of them appear to belong to one phylogenetic subspecies with a global distribution.\(^13\) The molecular mechanisms responsible for generating genetic variability and differences in pathogenicity in *A. fumigatus* are yet unknown. It has been explained by frequent subtelomeric exchanges between heterologous chromosomes and by intrapopulation genetic recombination. Being an asexual fungal species, *A. fumigatus* reproduces clonally generating haploid conidia from aerial conidiospores. Nonetheless evidence was found to support intrapopulation genetic recombination in this species attributed to the presence of an undetected sexual stage, parasexuality (mitotic cross-over), transposable elements, or past meiotic exchanges.\(^14\)

3.1.2 Innate Immune System and Invasive Aspergillosis Prevention and Resolution

Macrophages and neutrophils of the innate phagocytic immune system play fundamental roles in preventing and resolving IA. Macrophages in lung alveoli phagocytose and kill *A. fumigatus* conidia or small germings and thus are the first line of defense against infection. Neutrophils are very important for killing fungal hyphae and clearing established infections. The importance of both macrophages and neutrophils in combating IA is evident from several observations. First, in animal models of IA, treatments that inhibit or kill both cell types are required to successfully establish disease.\(^13\) Second, IA is a very rare disease in healthy people and is increasingly seen in patient populations with severe neutropenia\(^4\)\(^-\)\(^5\) or neutrophil dysfunction such as chronic granulomatous disorder, where incidence of IA reaches 40%.\(^6\)\(^,\)\(^16\) Finally, there is a positive correlation between recovery of neutrophil counts in infected patients and clearing of the invading fungi, and this observation is also true for animal models of IA.\(^17\)
These data demonstrate an absolute requirement for phagocytes in preventing and resolving IA and exemplify how studies of the interactions of these cells with Aspergillus fumigatus conidia and hyphae, using molecular genetic tools, will provide information on the biology of this host-pathogen interaction and lead to greater understanding of the disease process.

The other element of the innate immune system that is critical to first line defense against Aspergillus spp. is the system of innate immune molecules such as mannose binding protein (lectin) and pentraxin 3. Mannose binding protein is a hexameric protein synthesized in the liver, which avidly binds numerous microorganisms including A. fumigatus. Genetic defects, both heterozygous and homozygous, appear to be important in the pathogenesis of chronic pulmonary aspergillosis, a slowly destructive disease of the lungs caused by A. fumigatus. In 2002, Garlanda et al. showed that a KO mouse deficient in pentraxin 3 died rapidly of invasive aspergillosis. Additional work suggests that surfactant proteins are also important in pulmonary defense against Aspergillus. Thus another key element protecting mammals against overwhelming infection by Aspergillus is the innate immune system.

3.1.3 Application of Comparative Genomics to the Study of Aspergillus fumigatus

Recently we presented the genomic sequence of A. fumigatus isolate Af293. Its comparison with the two distantly related genomes, Aspergillus nidulans and Aspergillus oryzae, resulted in many unexpected discoveries, including the remarkable genetic variability of this genus. Proteome comparison of these three Aspergillus species revealed an average amino acid identity of less than 70% between each species pair, suggesting that they are as evolutionarily distant from each other as humans and fish. The significant phylogenetic distances hindered some aspects of comparative genomic analysis, such as studies of the genome organization and niche adaptation strategies present in the aspergilli. To examine the differences in gene content and regulatory elements responsible for the differences in virulence, sexual, and physiological properties of A. fumigatus, a very closely related sexual species, Neosartorya fischeri NRRL181 (A. fischerianus), and a more distantly related asexual species, Aspergillus clavatus NRRL1 were chosen for complete sequencing.

Despite the phylogenetic proximity of these two species, they are different enough at the phenotypic level from A. fumigatus Af293 to maximize the resolving power of the whole-genome comparative analysis. Both N. fischeri and A. clavatus are rarely identified as human pathogens with only a couple of medical cases reported in literature. The differences in pathogenicity have been attributed to their relative scarcity in environment, misidentification in the laboratory, or relative lack of virulence in N. fischeri and A. clavatus. Like most aspergilli N. fischeri is saprophytic and plays a role in food spoilage. Its thermoresistant ascospores allow it to survive heat processing and cause spoilage of processed foods and juices. Although not an invasive pathogen, A. clavatus may be an important allergenic fungus and has been shown to be the cause of an extrinsic allergic alveolitis known as malt worker’s lung. A. clavatus grows more slowly at 37°C than A. fumigatus and has a bigger spore size, which may prevent lung penetration. It produces a number of mycotoxins including patulin, kojic acid, cytochalasins, and tremorgenic mycotoxins.

3.2 Genome Sequences

3.2.1 Aspergillus fumigatus Af293

The genome of A. fumigatus Af293 was sequenced by the whole genome random-sequencing method augmented by optical mapping. Af293 contains eight chromosomes ranging in size from 1.8 to 4.9 Mb, for a total of 29.4 Mb of genomic sequence (Table 3.1). Additionally, there are at least 12 mitochondrial copies per nuclear genome. At the time of this publication, there were 9632 predicted protein-coding genes with a mean gene length of 1478 bp. About one-third of these predicted genes are of unknown function. Comparisons to the genomes of N. fischeri and A. clavatus revealed ~1000 genes specific to A. fumigatus, which have no detectable orthologs in other genomes, including several mycotoxin islands and paralogous gene families. Other notable findings include a complete gene complement for heterothallic sex; a cell wall assembly process that is quite different in structural detail from yeast; at least...
23 predicted gene clusters encoding proteins involved in secondary metabolism and mycotoxin production; evidence of cell death pathway components; at least 168 efflux pumps for drugs, toxins, and macromolecules; and genes encoding arsenate resistance similar to those found in bacteria. The genome sequence provides an unparalleled resource for the future understanding of this extremely prevalent fungus.

### 3.2.2 Neosartorya fischeri NRRL181 (Aspergillus fischerianus)

Although *N. fischeri* is rarely an invasive pathogen, it is a very close homothallic sexual relative of *A. fumigatus*. The genome of *N. fischeri* NRRL181 was sequenced by the whole genome random-sequencing method. The *N. fischeri* genome (32.6 Mb) is 10–13% larger than the *A. clavatus* and *A. fumigatus* AF293 genomes (Table 3.1). There are currently 10,407 predicted protein-coding genes with a mean gene length of 1466 bp. Comparisons to the genomes of *A. fumigatus* AF293 and *A. clavatus* revealed ~1600 genes unique to *N. fischeri*, including several mycotoxin islands and paralogous gene families. Other notable findings include a large number of transposable elements that may have contributed to the genome size expansion observed in this species.

### 3.2.3 Aspergillus clavatus NRRL1

*A. clavatus* is a very rare human pathogen with only one invasive medical case reported, that of postoperative endocarditis, and occasional external otitis. It does appear to be potentially allergenic in humans and can cause neurotoxicosis in sheep and cattle fed with infected grain. The genome of *A. clavatus* was sequenced by the whole genome random sequencing method. The *A. clavatus* genome (27.9 Mb) is the smallest seen to date among the aspergilli. There are currently 9125 predicted protein-coding genes with a mean gene length of 1483 bp (Table 3.1). Comparisons to the genomes of *A. fumigatus* AF293 and *N. fischeri* revealed ~1200 genes unique to *A. clavatus*, including patulin biosynthesis and other mycotoxin clusters as well as paralogous gene families. Other notable findings include a complete gene complement for heterothallic sex.

### 3.2.4 Sequence Divergence, Synteny, and Chromosomal Location

Orthologous genes were used to determine syntenic regions between genomes, by correlating computed orthologs with their genomic locations and relative gene order. Orthologs are defined as genes in different species that evolved from a common ancestral gene through speciation. In general, it is assumed that orthologs retain the same function in the course of evolution. Putative orthologs were identified as mutual best hits between genomes. We identified 8959 ortholog clusters, of which 7494 represent all three analyzed genomes (Fig. 3.1). The average protein identity between *A. fumigatus* and *N. fischeri* orthologs is 94%, while that between *A. fumigatus* and *A. clavatus* is 80%, supporting the published phylogeny of
A Comparative View of the Genome of Aspergillus fumigatus

these genomes. Comparison with other fungal taxa shows that orthologs encoded by central chromosomal regions have the highest average sequence identity. These genes often have important functional roles. The chromosomal core genes are most likely to be involved in information processing including nuclear and chromatin structure, RNA metabolism, translation, transcription, and cell cycle control. Genes from these categories are six times more likely to be found in the central chromosomal regions than within 300 Kb from chromosome ends. Genes involved in other important cellular processes as signal transduction, cytoskeleton organization, intracellular trafficking, energy conversion, and protein or nucleotide metabolism are also more likely to be found in core chromosomal regions. On the other hand, orthologs from subtelomeric regions have lower sequence identity and may, therefore, evolve faster than those encoded by the core genome.

Species-specific genes have been the focal point of several comparative genomic studies, because of the assumption that they may be responsible for differences in pathogenicity among different species. Several pathogenic species such as Leishmania major, Plasmodium falciparum, and Candida glabrata vary presentation of an array of species-specific pathogenicity factors via recombination in subtelomeric regions. This suggests that, at least in some eukaryotic genomes, the subtelomeric locations are associated with virulence and higher evolutionary rates of substitution and recombination.

Comparative analysis showed that species-specific genes comprise 10–20% of the A. fumigatus, A. clavatus, and N. fischeri genomes. A. fumigatus species-specific genes, while located predominantly in subtelomeric regions, do not appear to be involved in virulence. Rather, those genes with a discernible function have been linked to more general niche adaptation processes such as detoxification, transport, carbohydrate metabolism, transcriptional regulation, and secondary metabolism. These genes often belong to large paralogous gene families, such as MSF transporters, P450 oxidoreductases, and Zn(2)C(6)-type transcription factors. While some secondary metabolites such as gliotoxin, spore pigment, hydroxamate siderophores, and unknown products of NRPS Pes1 have been associated with A. fumigatus virulence, their corresponding clusters are not unique to this species.

Species-specific genes are predominantly found in subtelomeric and intrasyntenic “plasticity zones,” a characteristic common among sequenced genomes across the evolutionary spectrum. The plasticity zones appear to be rather large relative to S. cerevisiae and closely related species, suggesting that this may be the genetic characteristic responsible for their extreme environmental adaptability.

### 3.3 Virulence Genes

While A. fumigatus clearly is well adapted as an environmental saprophyte, it also is capable of establishing invasive infection in immunocompromised human hosts. The ability to do this is not particularly prevalent among filamentous fungi, and so it is hoped that the availability of the genome sequence of this fungus will provide clues as to why it can act as a pathogen.
3.3.1 Genes Implicated in Virulence

A limited number of candidate pathogenicity genes and components of pathogenicity have been identified in *A. fumigatus* by assaying mutants in cultured macrophages or in animal models of invasive aspergillosis. These include the enzymes involved in pyrimidine biosynthesis (PyrG) and pigment biosynthesis (PksP); a histidine kinase (fos-1); mycelial catalases (Cat1, Cat2); a Ras-related protein (RhbA); cAMP signaling pathway components (AcyA, GpaB); a folate biosynthesis pathway component (PabaA); secreted proteases; a chitin synthase (ChsG); and a nutrient sensing system component (CpcA). Immunosuppressive substances in culture filtrates such as gliotoxin have also been suggested to be pathogenicity factors. Differential display was used to compare gene expression in fungal cells grown on endothelial cells with that of cells grown without the endothelial cell contact. Two up-regulated genes subsequently characterized encoded a regulatory subunit of a cAMP-dependent protein kinase and a *ras* gene family protein. Both proteins are involved in cAMP-mediated signaling, a result that validates the potential of transcription profiling to reveal pathogenicity-related genes.

One of the goals of our comparative study is the identification of differential genetic traits associated with differences in virulent properties of *A. fumigatus* and the two closely related aspergilli. As noted earlier, approximately 1000 *A. fumigatus* genes are putatively species-specific as they do not have detectable orthologs in *N. fischeri* or *A. clavatus*. Unexpectedly, none of these genes (except the gliotoxin biosynthesis cluster) have been implicated in *A. fumigatus* virulence. Most previously identified virulence-associated genes have orthologs in nonpathogenic aspergilli and in the filamentous ascomycete *Neurospora crassa*. Furthermore, more than half of these genes have apparent orthologs in more distantly related fungi such as the hemiascomycete *S. cerevisiae* and basidiomycete *Cryptococcus neoformans* (Table 3.2).

The average percent identity among orthologs of the virulence genes is higher than the average for the aspergilli genomes consistent with strong Darwinian selection (Table 3.2). One of the most conserved genes is *cnaA*, encoding the calcineurin catalytic subunit, which has been shown to control growth, morphology, and pathogenicity in *A. fumigatus*. Components of the cAMP signaling pathway also appear to be well conserved in ascomycetes and basidiomycetes. Notably, the average percent identity between the *A. fumigatus* and yeast orthologs is 50%, which is higher than the average 39% identity between these species. They are even more conserved than orthologs of *S. cerevisiae* essential genes (median identity 41%; average 44%). Only one *S. cerevisiae* ortholog, adenylate cyclase AcyA/Cyr1p, has been shown to be essential for growth in yeast.

These results are consistent with the view that *A. fumigatus* virulence is a multifactorial process, which depends mostly on host immune system status. Indeed, the innate immune system is crucial in managing the host’s exposure to *A. fumigatus* conidia. Many *A. fumigatus* virulence genes are required for growth, development, or stress response and are, therefore, critical for survival in the hostile host environment. It appears that these proteins are essential for growth in a stressful environment, such as the human host, and, over the course of fungal evolution, can be repeatedly recruited to play a more direct role in pathogenicity. If this is true, similar molecular mechanisms may underpin pathogenesis in diverse fungal species.

3.3.2 Function of Virulence Genes

Although none of the *A. fumigatus* virulence proteins is shared uniquely with other fungal pathogens, several factors have a globally conserved role in the virulence of several animal and plant pathogens. They include components or targets of major signaling pathways such as cAMP/PKA and calcineurin pathways. For example, in many pathogenic fungi, polyketide synthases (PKSs) involved in biosynthesis of melanin-like pigments have been also shown to contribute to virulence by quenching free radicals and/or through the cAMP-dependent signaling pathway. Other globally conserved mediators of fungal pathogenesis include calcineurin, a cAMP-dependent protein kinase (PkaR/Bcy1), and a cell wall biosynthesis protein (Ecm33). If indeed some conserved fungal proteins can be repeatedly recruited for pathogenicity, new virulence factors can be predicted by comparative genomic analysis. This does not mean that all these proteins function as virulence factors in every pathogenic fungus. Nonetheless, at least two *A. fumigatus* proteins
TABLE 3.2
Pairwise Comparison between *A. fumigatus* Af293 Virulence Genes and Their Orthologs from *N. fischeri*, *A. clavatus*, *A. oryzae*, *A. nidulans*, *N. crassa*, *S. cerevisiae*, and *C. neoformans*

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>A. fumigatus</th>
<th>N. fischeri</th>
<th>A. clavatus</th>
<th>A. oryzae</th>
<th>A. nidulans</th>
<th>N. crassa</th>
<th>S. cerevisiae</th>
<th>C. neoformans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcineurin catalytic subunit CnaA</td>
<td>Afu5g09360</td>
<td>99</td>
<td>94</td>
<td>94</td>
<td>92</td>
<td>84</td>
<td>53</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>G protein complex alpha subunit GpaB</td>
<td>Afu1g12930</td>
<td>99</td>
<td>96</td>
<td>91</td>
<td>94</td>
<td>78</td>
<td>55</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Rheb small monomeric GTPase RhbA</td>
<td>Afu5g05480</td>
<td>99</td>
<td>97</td>
<td>91</td>
<td>93</td>
<td>70</td>
<td>46</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Protein kinase regulatory subunit PkaR</td>
<td>Afu3g10000</td>
<td>97</td>
<td>82</td>
<td>78</td>
<td>79</td>
<td>57</td>
<td>50</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Adenylyl cyclase AcyA/Cyr1</td>
<td>Afu6g08520</td>
<td>97</td>
<td>88</td>
<td>74</td>
<td>73</td>
<td>45</td>
<td>34</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Sensor histidine kinase/response regulator TcsA</td>
<td>Afu6g10240</td>
<td>94</td>
<td>80</td>
<td>62</td>
<td>59</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-ornithine N5-oxygenase SidA</td>
<td>Afu2g07680</td>
<td>96</td>
<td>86</td>
<td>77</td>
<td>78</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyketide synthetase PksP</td>
<td>Afu2g17600</td>
<td>96</td>
<td>86</td>
<td>73</td>
<td>68</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulator of secondary metabolism LaeA</td>
<td>Afu1g14660</td>
<td>99</td>
<td>91</td>
<td>83</td>
<td>78</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonribosomal peptide synthase Pes1</td>
<td>Afu1g10380</td>
<td>90</td>
<td>70</td>
<td>54</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rRNA processing protein CgrA</td>
<td>Afu8g02750</td>
<td>99</td>
<td>88</td>
<td>73</td>
<td>75</td>
<td>58</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orotidine 5'-phosphate decarboxylase PyrG</td>
<td>Afu2g08360</td>
<td>98</td>
<td>92</td>
<td>88</td>
<td>78</td>
<td>51</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifunctional catalase-peroxidase Cat2</td>
<td>Afu8g01670</td>
<td>97</td>
<td>86</td>
<td>83</td>
<td>72</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mycelial catalase Cat1</td>
<td>Afu3g02270</td>
<td>98</td>
<td>88</td>
<td>85</td>
<td>78</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bZIP transcription factor CpcA</td>
<td>Afu4g12470</td>
<td>94</td>
<td>77</td>
<td>71</td>
<td>62</td>
<td>71</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Para-aminobenzoate synthase PabaA</td>
<td>Afu6g04820</td>
<td>94</td>
<td>77</td>
<td>65</td>
<td>64</td>
<td>37</td>
<td>30</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Autophagic serine protease Alp2</td>
<td>Afu5g09210</td>
<td>99</td>
<td>93</td>
<td>87</td>
<td>81</td>
<td>66</td>
<td>54</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Alkaline serine protease Alp1</td>
<td>Afu4g11800</td>
<td>95</td>
<td>80</td>
<td>82</td>
<td>80</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chitin synthase ChsG</td>
<td>Afu3g14420</td>
<td>99</td>
<td>95</td>
<td>90</td>
<td>89</td>
<td>69</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chitin synthase ChsE</td>
<td>Afu2g13440</td>
<td>99</td>
<td>93</td>
<td>86</td>
<td>82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell wall organization protein Ecm33</td>
<td>Afu4g06820</td>
<td>96</td>
<td>80</td>
<td>66</td>
<td>60</td>
<td>32</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha-1,3-glucan synthase Ags3</td>
<td>Afu1g15440</td>
<td>97</td>
<td>84</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homoaconitase LysF</td>
<td>Afu5g08890</td>
<td>98</td>
<td>90</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>55</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>97</td>
<td>88</td>
<td>81</td>
<td>78</td>
<td>58</td>
<td>48</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole genome</td>
<td>95</td>
<td>80</td>
<td>73</td>
<td>69</td>
<td>50</td>
<td>42</td>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
are likely to be involved in virulence based on sequence similarity: Afu1g06350, which is orthologous to *Colletotrichum gloeosporioides* pathogenicity protein CAP20 and nonribosomal peptide synthase SidD, which is orthologous to NPS6 involved in virulence and oxidative stress protection in *Cochliobolus heterostrophus*.

### 3.3.3 What Is Virulence?

A revised view of what constitutes virulence genes and virulence in *A. fumigatus* is emerging through the application of comparative analysis. A review of the genes that have been identified by mutant analysis to reduce virulence, as discussed earlier, include pyrG, pksP, sidA, laeA, and several others. The range of functions performed by the products of these genes and their close association to fundamental components of the core metabolic infrastructure of the fungus suggest that the ability to survive in a human host is not the consequence of the presence of true virulence genes but of the metabolic capabilities it has evolved to succeed as a saprophyte, including its temperature versatility, defense mechanisms against oxidative stress, and ability to effectively export potentially harmful chemicals present in its environment. In support of this hypothesis is the observation that no genomic components are shared and exclusively by *A. fumigatus* and other human pathogens such as the *Candida* or *Cryptococcus* species.

### 3.4 Secondary Metabolite Biosynthetic Gene Clusters

Filamentous fungi display many unique characteristics that render them of great interest to the research community. Among these characteristics is the production of natural products, or secondary metabolites. These compounds often have obscure or unknown functions in the producing organism, but have tremendous importance to humankind. Secondary metabolites display a broad range of useful antibiotic and immunosuppressant activities, as well as less desirable phyto- and mycotoxic activities. The distribution of natural products is characteristically restricted to certain fungal taxa, particularly the eurotiomycete branch of filamentous Ascomycetes.

#### 3.4.1 Clusters in *Aspergillus fumigatus* Af293

Secondary metabolism genes appear to be the most fast evolving segment of the fungal genomes. They tend to be found in clusters in many fungal genomes, including the aspergilli. A few low molecular weight fungal metabolites, such as melanin-like conidial pigment and siderophores, are encoded by similar clusters and are, therefore, analyzed here together with secondary metabolites. The accurate identification of clusters producing specific metabolites is not a trivial task.

The sequencing of the *A. fumigatus* Af293 genome revealed potential secondary metabolism clusters, which exceeds the number of secondary metabolites previously identified in this species. Recently, expression studies validated 15 of these clusters, which were shown to be regulated globally by regulator of secondary metabolism, LaeA. We have cataloged 23 secondary metabolism clusters in the *A. fumigatus* genome by leveraging phylogenetic profiles and domain analysis (Table 3.3).

#### 3.4.2 Associating Clusters with Products

Until recently very few *A. fumigatus* genes were assigned a specific metabolite, such as the conidial pigment. Within months after completion of the *A. fumigatus* genome, several new clusters were associated with known metabolites such as gliotoxin, ergot alkaloids, fumigaclavines, fumitremorgin, and ferrichrome-type siderophores and an unknown product of NRPS Pes1 associated with virulence and oxidative stress. Still, many products have yet to be associated with a particular cluster, including fumagillin, fumagiringilin, fumigatin, fumitoxin, helvolic acid, monotrypacidin, phthioic acid, pseurotin, sphingofungins, trypacidin, tryptoquivaline, and verruculogen.
Much less is known about secondary metabolites produced by *N. fischeri* and *A. clavatus*. Fumitremorgins, tryptoquivaline and verruculogen were identified in *N. fischeri*.79 *A. clavatus* has been shown to produce anti-fumicin and its derivatives, brefeldin A, cytochalasins, kojic acid, kotanin, orlandin, patulin, tryptoquivaline, and tryptoquivalone.31 The analysis of the *N. fischeri* and *A. clavatus* genomes identified 28 and 27 putative secondary metabolites clusters, respectively. Three clusters, clusters #1 (Pes1-associated), #3 (conidial pigment) and #7 (siderophore), appear to be shared by all aspergilli, consistent with a role in primary, rather than secondary, metabolism (Table 3.3). *A. fumigatus* and *N. fischeri* share 11 additional orthologous clusters, including gliotoxin, putative ETP toxin, and putative pseurotin biosynthesis clusters, while *A. fumigatus* and *A. clavatus* share only 2 additional clusters.

The discovery of the gliotoxin cluster in *N. fischeri* was unexpected, since this mycotoxin had not been found in this organism before. It is produced by very few species including *A. fumigatus*, *Aspergillus terreus*, *Hypocrea* (*Trichoderma*) *virens*, *Penicillium* spp., and some *Candida albicans* isolates. The gliotoxin cluster is one of the best characterized clusters in *A. fumigatus*, due to the theory that gliotoxin contributes directly to its invasive growth, as it is detected in human and mouse tissue infected with *A. fumigatus*.80 *A. fumigatus* and *N. fischeri* share another ETP toxin cluster with similarity to the gliotoxin cluster, suggesting that one of them might have arisen via a segmental duplication before the divergence of the three species.

In addition to the orthologous clusters shared by the three aspergilli, many unique clusters were identified in *N. fischeri* and *A. clavatus*, including the *A. clavatus* patulin cluster. Patulin is one of the most well-quantified toxic compounds produced by *A. clavatus*.

### TABLE 3.3

**Pairwise Comparison between *A. fumigatus* Af293 Secondary Metabolism Clusters and Their Orthologs and Paralogs from *N. fischeri*, *A. clavatus*, *A. oryzae*, and *A. nidulans***

<table>
<thead>
<tr>
<th>Accessions</th>
<th>Cluster Median % Identity Per Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>N. fischeri</em></td>
</tr>
<tr>
<td>Afu1g10380–Afu1g10390 1</td>
<td>94</td>
</tr>
<tr>
<td>Afu1g17710–Afu1g17740 2</td>
<td>Unknown (NRPS Pes1)</td>
</tr>
<tr>
<td>Afu2g17530–Afu2g17600 3</td>
<td>Pigment (PKS)</td>
</tr>
<tr>
<td>Afu2g17960–Afu2g18060 4</td>
<td>Ergot alkaloid (DMATs FgaPT1, 2)</td>
</tr>
<tr>
<td>Afu3g01400–Afu3g01420 5</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu3g02560–Afu3g02630 6</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu3g03350–Afu3g03470 7</td>
<td>Siderophore (NRPSs SidD and SidE)</td>
</tr>
<tr>
<td>Afu3g12890–Afu3g12960 8</td>
<td>ETP toxin (NRPS, DMAT)</td>
</tr>
<tr>
<td>Afu3g13710–Afu3g13730 9</td>
<td>Unknown (NRPS)</td>
</tr>
<tr>
<td>Afu3g14690–Afu3g14720 10</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu3g15250–Afu3g15310 11</td>
<td>Unknown (NRPS)</td>
</tr>
<tr>
<td>Afu4g00150–Afu4g00235 12</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu4g14520–Afu4g14590 13</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu4g14735–Afu4g14850 14</td>
<td>Unknown (none)</td>
</tr>
<tr>
<td>Afu5g12720–Afu5g12740 15</td>
<td>Unknown (NRPS)</td>
</tr>
<tr>
<td>Afu6g03400–Afu6g03490 16</td>
<td>Unknown (NRPS-like)</td>
</tr>
<tr>
<td>Afu6g09590–Afu6g09740 17</td>
<td>Gliotoxin (NRPS)</td>
</tr>
<tr>
<td>Afu6g12040–Afu6g12080 18</td>
<td>Unknown (NRPSs)</td>
</tr>
<tr>
<td>Afu6g13920–Afu6g14000 19</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu7g00120–Afu7g00180 20</td>
<td>Unknown (PKS, DMAT)</td>
</tr>
<tr>
<td>Afu8g00100–Afu8g00290 21</td>
<td>Fumitremorgin (NRPS, DMAT)</td>
</tr>
<tr>
<td>Afu8g00300–Afu8g00520 22</td>
<td>Unknown (PKS)</td>
</tr>
<tr>
<td>Afu8g00530–Afu8g00580 23</td>
<td>Pseurotin (hybrid)</td>
</tr>
<tr>
<td>All clusters</td>
<td>94</td>
</tr>
<tr>
<td>Whole genome</td>
<td>95</td>
</tr>
</tbody>
</table>

*Conserved, nonsyntenic cluster. May be paralogous.*
researched and widely disseminated mycotoxins found in fruit and animal feed products. Exposure to patulin can result in severe acute and chronic toxicity in animals and humans.\textsuperscript{81} It was recently identified in \textit{A. clavatus} isolates associated with a lethal neurotoxicosis in cattle.\textsuperscript{82} The mechanisms of patulin toxicity are not well understood, but it has been shown to activate protein degradation, sulfur amino acid metabolism, and the defense system for oxidative stress in the yeast transcriptome system.\textsuperscript{83}

In addition to the secondary metabolism clusters, the aspergilli genomes contain large numbers of stand-alone PKS and NRPS-like enzymes, which may or may not be involved in secondary metabolism. One striking example involves a putative NRPS-like enzyme (Afu5g10120) from \textit{A. fumigatus}, which has highly conserved orthologs in most fungal genomes, except yeast, \textit{B. cinerea} and \textit{N. crassa}. Its deletion mutant in \textit{Cochliobolus heterostrophus} (NPS10) did not exhibit a specific phenotype.\textsuperscript{68}

\subsection*{3.4.3 Evolution of Clusters}

Despite years of research, mechanisms involved in cluster assembly and maintenance are not well understood. Several hypothesis have been proposed, including the selfish cluster hypothesis, which suggests that clustering of genes is maintained through horizontal gene transfer (HGT),\textsuperscript{84} and the coregulation hypothesis.\textsuperscript{85} The latter recently received additional support, as functional gene clusters have been shown to be formed by selection for physical proximity of the genes or genetic linkage.\textsuperscript{86} It was hypothesized that regulation at the heterochromatin level a putative nuclear methyltransferase, LaeA, may facilitate cluster assembly through coregulation of all genes inserted within cluster boundaries.\textsuperscript{69} The comparative analysis of the aspergilli performed by our group is consistent with the coregulation theory. It also suggests that once assembled, gene clusters may evolve through duplication and translocation, followed by accelerated evolution and differential gene loss, as described later (Fig. 3.2).

The comparative analysis revealed extreme diversity of secondary metabolism biosynthesis genes in the aspergilli. The \textit{A. fumigatus}, \textit{N. fischeri}, and \textit{A. clavatus} genomes have between 23 and 29 putative secondary metabolism clusters (data not shown). Only five of them are shared among all three aspergilli including the spore pigment biosynthesis, siderophore biosynthesis, and Pes1 clusters, which are conserved in most sequenced aspergilli species (Table 3.3). This sporadic distribution may reflect niche differences among these species and suggests that the clusters evolution involved rampant gene loss.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.2.png}
\caption{Evolution of secondary metabolism clusters. \textit{Note:} This diagram shows evolution of a hypothetical secondary metabolism cluster by segmental duplication, diversification, and differential loss. The star designates cluster loss.}
\end{figure}
Despite this apparent variability, most orthologous clusters are syntenic and well conserved among the three aspergilli. The *A. fumigatus* secondary metabolism genes share 94% and 84% amino acid sequence identity with their orthologs in *N. fischeri* and *A. clavatus*, respectively, which is similar to the average identity between pairs of orthologs for these genomes. Within each cluster most genes have similar percent identity, except genes encoding NRPS and PKS enzymes, catalyzing the first steps in secondary metabolite biosynthesis. They often appear to be the most rapidly evolving genes within a cluster (data not shown).

In addition to these relatively stable orthologous clusters, there are two clusters that appear to be evolving very fast or perhaps are under positive selection at least in some species. Their orthologous relationships are difficult to establish at this time. They share very low average sequence identity per cluster and may in fact be hidden paralogs resulting from differential gene loss (Fig. 3.2). Thus, *A. fumigatus* cluster #18 and related clusters in *N. fischeri* and *A. clavatus* share only 43% and 55% similarity, respectively. The *N. fischeri* cluster is located in the middle of syntenic regions, while the *A. clavatus* cluster seems to be translocated. This cluster has apparently undergone accelerated levels of variation. A putative ETP toxin cluster from *A. clavatus*, which is distantly related to the *A. fumigatus* gliotoxin cluster, might have also undergone accelerated divergence (Fig. 3.2). It is most similar to an *A. oryzae* cluster of unknown function (Table 3.3).

One more observation is consistent with our hypothesis that segmental duplication, accelerated divergence, and differential loss contributed to cluster diversification in aspergilli. The genome of *A. fumigatus* Af293 contains a duplicated and translocated copy of the arsenic resistance gene cluster, which appears to be more divergent than the original syntenic copy. A similar pattern is observed for another duplicated cluster of unknown function in *A. fumigatus* Af293. Similarly the availability of duplicated secondary metabolism clusters may lead to functional redundancy, which allows evolutionary change. This is consistent with the view that segmental duplications have higher evolutionary rates and represent hot spots for chromosome and gene evolution in eukaryotic organisms.87

Besides segmental duplication and accelerated divergence, other events must have contributed to clusters’ evolution. Many species-specific clusters appear to be assembled *de novo*, since they are often surrounded by transposable elements (TEs) and repeats as well as apparent indels and nonsyntenic blocks. Several factors may facilitate cluster *de novo* assembly: availability of “spare parts,” the enriched crossover activity in “plasticity zones,” and coregulation at the heterochromatin level by LaeA as discussed earlier. Most clusters contain genes from a limited number of paralogous families such as Zn(2)C(6) transcription factors, MFS transporters, various oxidoreductases, and FAD binding enzymes. These gene families are often found among species-specific genes and may, therefore, serve as assembly components (spare parts) for future clusters. Finally, with the exception of the pigment biosynthesis and Pes1-associated clusters, *A. fumigatus* secondary metabolism clusters are located in plasticity zones, highly dynamic chromosomal regions known for their high recombination rates in many eukaryotic genomes. This may facilitate both segmental duplication and *de novo* assembly of clusters by placing interacting or coregulated genes in physical proximity through frequent exchanges between nonhomologous chromosomes. Once inserted within cluster boundaries, genes will become coexpressed and genetically linked with the interacting genes from the cluster. Secondary metabolism clusters may encode products that interact with each other in a dose-sensitive manner or are required to prevent cytotoxicity (e.g., transporter and NRPS genes).

### 3.5 Sex and Sex Genes

Despite apparent asexuality of *A. fumigatus* and *A. clavatus*, their genomes possess all the genes required for sexual reproduction including mating-type genes and genes involved in pheromone-dependent signal transduction, meiosis and fruit body development.23,32 The corresponding genes appear to be under negative selection with the signal transduction genes being the most conserved. This lack of substantial sequence divergence is consistent with a very recent loss of sexuality or a hidden sexual stage.88–90 The only exceptions are rosA and mating (mat) locus genes. They encode putative transcription factors and seem to evolve rather rapidly.
3.5.1 Mating loci in *Aspergillus fumigatus*, *Aspergillus clavatus*, and *Neosartorya fischeri*

In filamentous ascomycetes, the mat loci contain one or two different mating type genes that establish sexual compatibility: a high-mobility group (HMG) type gene and an alpha-domain type gene. In different haploid strains of heterothallic species, the loci occupy the same chromosomal location and contain idiomorphic stretches of DNA, which lack apparent sequence similarity. Different *A. fumigatus* isolates contain either an HMG (MAT1-1 or MAT-1) or an alpha box (MAT1-2 or MAT-2) mating-type genes. Analysis of the mat locus in the sequenced strain of *A. clavatus* showed an HMG type gene (MAT1-1), but no alpha-domain type gene (Fig. 3.3). These observations imply that both *A. fumigatus* and *A. clavatus* may be heterothallic species, which require a partner with a different mating type gene (obligate outcrossing). In contrast, *N. fischeri*, which is homothallic (self-fertile), has two unlinked mat loci with the opposing mating type genes.

A closer look at the *A. fumigatus*, *N. fischeri*, and *A. clavatus* mat loci reveals an interesting infrastructure. Although often described as idiomorphs, the MAT-2 genes appear to be highly divergent alleles. Both *A. clavatus* and *N. fischeri* mat loci contain a truncated MAT-2 allele at the same position where there is a functional MAT-2 allele in *A. fumigatus* (Fig. 3.3). On the other hand, MAT-1 genes may be true idiomorphs. The *A. fumigatus* mat locus contains a small uncharacterized gene, which occupies the same positions but shares no apparent sequence similarity with the *N. fischeri* and *A. clavatus* MAT-1 genes. Similar hypothetical genes are found adjacent to MAT-2 genes in *N. fischeri* and the eurotiomycete fungus, *Coccidioides immitis* (CIMG_00407). The organization of the mat locus in the sequenced *N. fischeri* isolate is even more unusual. In addition to the full-length MAT-2 gene and the small uncharacterized gene, it contains a truncated gene coding for DNA lyase. This mat locus is found at the end of a small assembly containing mostly uncharacterized genes, pseudogenes, and transposable elements (TEs). This arrangement is consistent with a recent transposon-assisted translocation or duplication. In contrast, in another homothallic species, *A. nidulans*, the mat loci are separated by an apparent translocating break.

![FIGURE 3.3 Aspergillus mating loci. Note: The diagram shows mating loci from *A. fumigatus* Af923, *N. fischeri*, and *A. clavatus*. APN2, gene-encoding DNA lyase Apn2; unk, protein of unknown function; SLA1, gene-encoding cytoskeleton assembly control protein Sla1; MAT-1, gene-encoding alpha box mating-type protein; MAT-2, gene-encoding HMG mating-type protein.](image-url)
3.6 Status of *Aspergillus fumigatus* Genome Annotation

The identification and annotation of protein-coding genes is one of the primary goals of whole genome sequencing projects, and the accuracy of the predicted proteome is vitally important for in-depth comparative analyses and downstream functional genomic applications. Yet structural annotation of eukaryotic genomes remains a considerable challenge, despite the exponential growth in the number of sequenced genomes and improvements in eukaryotic gene prediction algorithms. Many genomes submitted to public databases, including those of major model organisms, contain significant percentages of misannotated gene structures and unvalidated gene predictions.

Due to the high cost and time required for manual genome annotation, most genomes are annotated via automated gene prediction pipelines. Generically, automated annotation pipelines rely on three types of data: genomic alignment of native cDNA and EST sequences, genomic alignment of available protein sequences from organisms at different evolutionary distances, and gene prediction algorithms that identify putative genes based on statistical patterns characteristic of protein-coding regions. Once all the components of an automated annotation pipeline have been run, preliminary gene structures may be based on the single gene prediction method that performs best when evaluated against alignment data, or generated by an algorithm that computationally combines the alignment evidence and gene predictions into a set of consensus predictions.

Recently, with the increasing availability of evolutionarily related genome sequences, comparative gene prediction has become increasingly important. To date, the most established and successful algorithms have exploited genomic alignments between two genomes of an optimal evolutionary distance to indicate which nucleotides are under negative selection and, therefore, more likely to be coding or otherwise functional. Comparative prediction techniques have been successfully applied to fungi. The dual-genome prediction program TWINSCAN was applied to the primary annotation of the *C. neoformans* Serotype D genome using Serotype A as the comparator. The comparative reannotation of *S. cerevisiae* with three related *Saccharomyces* species resulted in revision of approximately 15% of the predicted gene complement. Comparative techniques hold great promise, but more research and development is needed to effectively leverage the complete spectrum of available sequence.

The genomes described in this chapter, *A. fumigatus*, *N. fischeri*, and *A. clavatus*, were all annotated through automated annotation pipelines configured at TIGR. Despite our best efforts, utilizing state-of-the-art gene prediction programs and leveraging the wealth of fungal genome data available, there were problems and inconsistencies in the data that became evident through the preliminary comparative proteome analysis. There is very little cDNA or EST data available for these genomes, so gene-finding programs were trained on small data sets of manually curated gene models based on protein homology to public databases. This method is suboptimal, since the resulting training set includes only highly conserved shared proteins, ignoring fast-evolving and species-specific genes, which may have different and/or less conserved splicing sites. Common annotation problems include missed or incorrect 5′ exons, missed internal exons, and the inappropriate merging of neighboring genes into single, erroneous gene structures.

The initial *A. clavatus* and *N. fischeri* data sets consisted of the raw output of the automated pipeline without manual review or modification. The *A. fumigatus* gene structures were manually improved prior to publication, but before the public availability of the other Aspergillus genomes. Once these genomes became available, we examined annotation consistency among computed orthologs by assessing sequence similarity and alignment coverage. Although the orthologs shared an average 85–94% amino acid similarity, discrepancies in gene structure and content were identified in approximately 50% of the ortholog groups. In addition, less than 80% had average sequence coverage of 90% or higher, suggesting that merged or truncated protein sequences were prevalent. Additionally, the comparative data suggested that hundreds of new genes were potentially missed in the original annotation of *A. fumigatus*.

These data were consistent with previous annotation comparisons between the published proteomes of *A. fumigatus* and the more distant aspergilli, *A. nidulans* and *A. oryzae*. While there was significant sequence similarity among identified orthologs, only ~20% contained members with consistent
gene structures. Most (80%) orthologous genes differ in length and/or the number of exons. While some differences in exon number appear to be real at these evolutionary distances, examination of protein alignments suggest that the vast majority of these differences are due to annotation problems. It is not surprising that the annotation inconsistency between these genomes is even more marked than that between the more closely related aspergilli. The genomes were annotated at different sequencing centers (TIGR, the Broad Institute and National Institute of Advanced Industrial Science and Technology, Japan) using different gene prediction algorithms, alignment programs, and data sets. In addition, very few EST/cDNA sequences were available for training of gene prediction algorithms used in the original annotation of *A. fumigatus* and *A. oryzae*.

As of fall 2006, the genome annotations of *A. fumigatus*, *A. oryzae*, and *A. nidulans* have all been updated and improved postpublication through computational updates and/or manual gene structure curation. For *A. fumigatus*, comparative genome analyses with *N. fischeri* and *A. clavatus* were used to target gene structures for manual update. Over 1000 gene structures were modified, and approximately 130 genes were added to the annotation. Several hundred unsupported and unconserved open reading frame predictions were removed from the annotation data set. These and any future updates will be available through GenBank under the WGS accession AAHF00000000. In the course of *A. fumigatus* annotation improvement, a number of *N. fischeri* and *A. clavatus* gene structures were also manually reviewed and curated. These improvements will be reflected in the imminent GenBank release of these gene sets.

Even after these iterative annotation improvements, inconsistencies between the data sets remain. All of the *Aspergillus* genomes would still benefit from the generation of additional ESTs and/or the validation of predicted gene models through RT-PCR. There is also a need for more algorithm development to fully leverage comparative genomics data into new and improved gene models, to decrease the need for manual intervention. Ideally, these data sets would continue to evolve as more experimental data becomes available and computational methods improve.

### 3.7 Summary and Prospects

The comparative analysis revealed that the core *Aspergillus* proteomes remained remarkably stable over the last 200 million years. Moreover, many of the proteins have direct orthologs in other ascomycete fungi such as *N. crassa* and *S. cerevisiae*. Other findings include presence of a large number of putative essential genes in the core proteome, which may represent potential new antifungal targets in *A. fumigatus*. Identification of differential genetic traits associated with differences in virulent properties in *A. fumigatus* and closely related species proved to be a difficult task, since proteins associated with *A. fumigatus* virulence are remarkably conserved in other aspergilli. More subtle differences in coding and noncoding sequences may be responsible for these differences. In addition to the highly conserved and syntenic core, each *Aspergillus* genome contains hundreds of species-specific genes, many of which are likely involved in niche adaptation or secondary metabolism. Evolution of some “metabolic adaptation” clusters involved recent segmental duplication and accelerated divergence, suggesting an important role for these factors in cluster diversification in aspergilli.

### References

A Comparative View of the Genome of Aspergillus fumigatus


64. Martinez-Lopez, R., Park, H., Myers, C. L., Gil, C., and Filler, S. G., Candida albicans Ecm33p is important for normal cell wall architecture and interactions with host cells, Eukaryot Cell 5 (1), 140–147, 2006.


Aspergillus nidulans Genome and a Comparative Analysis of Genome Evolution in Aspergillus

Antonis Rokas and James E. Galagan

CONTENTS
4.1 Introduction ............................................................................................................. 43
4.2 Genome Characteristics .......................................................................................... 44
  4.2.1 Genome Assembly ............................................................................................. 44
  4.2.2 Genome Annotation and Analysis ...................................................................... 44
  4.2.3 Repetitive Elements .......................................................................................... 45
  4.2.4 Functional Annotation and Common Protein Domains .................................... 46
  4.2.5 Secondary Metabolites ..................................................................................... 46
  4.2.6 Physiology and Development .......................................................................... 48
4.3 Gene Regulation ...................................................................................................... 48
  4.3.1 Conserved Noncoding Sequences and Prediction of Regulatory Motifs .......... 48
4.4 Genome Evolution .................................................................................................. 48
  4.4.1 Phylogenetic Relationships .............................................................................. 50
  4.4.2 Conserved Synteny ......................................................................................... 50
  4.4.3 Genome Size ..................................................................................................... 51
  4.4.4 Reproductive Strategy .................................................................................... 52
4.5 Future of Aspergillus Genomics ............................................................................ 53
Acknowledgments .......................................................................................................... 53
References ....................................................................................................................... 54

4.1 Introduction

Perhaps no other fungal genus contains species that are so harmful and species that are so beneficial to humans as the genus *Aspergillus*. The genome sequence of the model organism *Aspergillus nidulans* was sequenced by the Broad Institute, and the genomes of a number of other *Aspergillus* species have also been sequenced by several others. The profoundly different lifestyles exhibited by each of the *Aspergillus* species for which genome sequences are available coupled with the varying degrees of evolutionary affinity shared by their genomes makes *Aspergillus* a model clade to address fundamental questions in functional and comparative genomics. Here we review the essential characteristics of the *A. nidulans* genome and place them in the larger evolutionary context of the genus *Aspergillus* and filamentous fungi. *Aspergillus nidulans* is a key fungal model system for genetics and cell biology. Work on *A. nidulans* has led to important advances in our understanding of metabolic regulation, cytoskeletal function, mitosis and cell cycle, and development. *A. nidulans* is evolutionarily related to a large number of other *Aspergillus* species of industrial and medical significance. For example, *A. niger* is widely exploited by the fermentation industry for the production of citric acid, an essential ingredient to the manufacturing of soft drinks, whereas *A. oryzae* plays a key role in the fermentation process of several traditional Japanese...
beverages and sauces. In contrast, *A. flavus* is a plant and animal pathogen that also produces the potent carcinogen aflatoxin, whereas *Aspergillus fumigatus* is an important pathogen of individuals with compromised immune systems as well as a major allergen. Importantly, whereas all these other *Aspergillus* species are only known from the asexual parts of their lifecycle, *A. nidulans* has a well-characterized, conventional genetic system, enabling classical genetic analyses.

The *A. nidulans* genome was sequenced by the Broad Institute and a comparative analysis with the genomes of *A. fumigatus* (sequenced by The Institute for Genomic Research) and *A. oryzae* (sequenced by the National Institute of Technology and Evaluation) has already been published. Since then, the genome of the *A. niger* strain CBS 513.88 has been published, and the genome sequences from *A. clavatus, A. flavus, A. niger* strain ATCC 1015, *A. terreus*, and *Neosartorya fischeri* (teleomorph of *A. fischerianus*) have become publicly available (Table 4.1). Here we review the main characteristics of the *A. nidulans* genome in a comparative context and highlight the progress in our understanding of the genomics of this important filamentous fungus.

### 4.2 Genome Characteristics

#### 4.2.1 Genome Assembly

The *A. nidulans* strain that is sequenced is the strain A4 from the Fungal Genetics Stock Center (http://www.fgsc.net/). The genome was sequenced through the whole-genome shotgun approach and the assembly of all the generated sequence was performed using the Arachne software package. The size of the genome is approximately 31 Mb and is distributed among eight chromosomes. It is estimated that the current assembly release represents 96.3% of the *A. nidulans* genome (30,068,514 bp). Furthermore, the genome assembly is covered to a depth of 13X, and contains 89 supercontigs (scaffolds) from 248 sequence contigs larger than 2 kb. The average base in the genome lies in a contig with length at least 282 Kb (contig N50 value), whereas the average base lies within a supercontig with length at least 2.44 Mb (supercontig N50 value). Very highly conserved repetitive sequences as well as the approximately 45 copies of the ribosomal RNA array are not captured in the current assembly.

#### 4.2.2 Genome Annotation and Analysis

Annotation of the *A. nidulans* genome has been performed in two steps. In the first step, three different gene prediction algorithms [Fgenesh, Fgenesh+ (both available from http://sun1.sofberry.com/berry.phtml), and GeneWise] were used to automatically predict genes. In the second step, the manual
Aspergillus nidulans Genome

45

### Table 4.2
Major Characteristics of the Aspergillus nidulans Genome

<table>
<thead>
<tr>
<th>Contig</th>
<th>Exon</th>
<th>Gene</th>
<th>Coding</th>
<th>Nongenic</th>
<th>Intron</th>
<th>Noncoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>248</td>
<td>35,525</td>
<td>10,701</td>
<td>10,701</td>
<td>10,908</td>
<td>24,824</td>
</tr>
<tr>
<td>Length (bp)</td>
<td>300,68,514</td>
<td>15,47,7748</td>
<td>17,67,1,460</td>
<td>15,47,7748</td>
<td>12,39,7054</td>
<td>21,93,712</td>
</tr>
<tr>
<td>Coverage (%)</td>
<td>100.0</td>
<td>51.5</td>
<td>58.8</td>
<td>51.5</td>
<td>41.2</td>
<td>7.3</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>50.3</td>
<td>53.3</td>
<td>52.4</td>
<td>53.3</td>
<td>47.4</td>
<td>45.8</td>
</tr>
<tr>
<td>Median length (bp)</td>
<td>55,350</td>
<td>219</td>
<td>1413</td>
<td>1233</td>
<td>758</td>
<td>60</td>
</tr>
<tr>
<td>Mean length (bp)</td>
<td>121,244</td>
<td>436</td>
<td>1651</td>
<td>1446</td>
<td>1137</td>
<td>88</td>
</tr>
</tbody>
</table>

The annotation team at TIGR used the PASA pipeline to align all available EST data to the genome and compare it to the annotation data set. Over 1000 gene structure updates were performed by the PASA software, and another ~2000 genes and intergenic regions were flagged for manual review. In addition to the EST data, the manual annotation team also relied on protein alignments and the output of an internal software package, called EvidenceModeler, to help in determining candidates for gene edits, most notably splitting of inappropriately merged gene models. As a result of the manual annotation, 494 original loci were split into two or more loci, 16 original loci were merged into single loci, and 214 new loci were created. Further information about the annotation of the genome can be found in Ref. 2 and in the *A. nidulans* genome website at the Broad Institute (http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/GeneFinding.html).

The current annotation release (Release 4, March 7, 2006) contains 10,701 genes, which correspond to 10,701 transcripts. Of all transcripts, 9227 are spliced, and 1474 are unspliced. There are a total of 35,525 exons and 24,824 introns corresponding to an average 3.3 exons and 2.3 introns per gene.

The major characteristics of the *A. nidulans* genome and its annotation are summarized in Table 4.2. About 58.8% of the genome corresponds to gene sequence, of which 51.5% corresponds to exons and 7.3% to introns. The remaining 41.2% of the genome corresponds to intergenic regions and repeats. The overall genome-wide GC content of the genome is 50.3%, with a slightly higher percentage found in exonic regions (53.3% GC), and a slightly lower one found in introns and intergenic regions (47.4% and 45.8% GC content, respectively).

### 4.2.3 Repetitive Elements

Approximately 4.37% of the genome consists of repeat sequences and duplicated regions larger than 200 bp, with transposable elements (TEs) accounting for 1.4% of the genome. The TEs found in the *A. nidulans* genome cover both major classes (Table 4.3); class I elements (retrotransposons) account for 60% of TEs in the genome, whereas the remaining 40% corresponds to class II elements (DNA transposons). *A. nidulans* TEs are unusual in a number of ways.\(^2\) First, unusual 7–9 kb in length *Mariner*

### Table 4.3
Major Superfamilies of Transposable Elements Identified in the Aspergillus nidulans Genome

<table>
<thead>
<tr>
<th>Size and Percentage of Transposable Elements in the Assembly</th>
<th>Non-LTR Retrotransposons</th>
<th>LTR Retrotransposons</th>
<th>DNA Transposons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size and Percentage of Transposable Elements in the Assembly</td>
<td>Class I</td>
<td>Class I</td>
<td>Class II</td>
</tr>
<tr>
<td>I (1.4%)</td>
<td>SINEs</td>
<td>Others</td>
<td>Gypsy</td>
</tr>
<tr>
<td>417,671</td>
<td>91,149</td>
<td>5650</td>
<td>83,185</td>
</tr>
<tr>
<td>(21.8%)</td>
<td>(1.4%)</td>
<td>(0%</td>
<td>(19.9%)</td>
</tr>
</tbody>
</table>
elements are present, and which encode—in addition to a transposase—a conserved protein of unknown function. Second, the Helitron elements have 5′-TT and CTTG-3′ termini, which differ from the canonical termini. Finally, the Aspergillus genomes (A. nidulans, A. oryzae, and A. fumigatus) are the first fungal genomes found to contain SINE elements (Table 4.3).2

### 4.2.4 Functional Annotation and Common Protein Domains

Approximately 63% of the A. nidulans proteome can be functionally annotated through a combination of a BLAST search17 against Genbank’s nr database and an Interproscan search18 against the Interpro database19 as implemented in the BLAST2GO software20 (Fig. 4.1). A search of the A. nidulans proteome against the PFAM set of hidden Markov models21 using the HMMER program (http://hmmer.janelia.org/), revealed that approximately 51% (5491 proteins) of the A. nidulans proteins contain domains belonging to one of the known protein families in the PFAM database.21 The 10 most commonly found protein domains in the A. nidulans genome and a short description of their function are shown in Table 4.4.

Most of the most common domains and repeats found in the genome represent abundant motifs typically found in any eukaryotic genome, although a few domains represent innovations in the fungal lineage (e.g., the Fungal Zn(2)-Cys(6) binuclear cluster domain and the fungal specific transcription factor domain).

### 4.2.5 Secondary Metabolites

Secondary metabolites are low molecular weight compounds thought to play a key role in fungal niche adaptation and virulence.22 Screens of the A. nidulans genome have revealed an abundance of genes implicated in secondary metabolite production,23 including 27 polyketide synthases, 14 nonribosomal peptide synthases, 6 fatty acid synthases, 1 sesquiterpene cyclase, and 2 dimethylallyl tryptophan synthases. In agreement with previous reports highlighting the cellular dispensability of secondary metabolic gene clusters, the degree of conservation of secondary metabolite clusters found in A. nidulans compared with other Aspergillus species is low4 (Rokas and Galagan, unpublished data). Additionally,
Aspergillus nidulans Genome

Table 4.4

Ten Most Commonly Found Protein Domains in the A. nidulans Genome

<table>
<thead>
<tr>
<th>PFAM Domain</th>
<th>Number</th>
<th>Accession Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD domain, G-beta repeat</td>
<td>569</td>
<td>PF00400.20</td>
<td>Short ~40aa motifs involved in coordinating multiprotein complex assemblies are found across eukaryotes. They are implicated in a variety of biological functions (signal transduction, transcription regulation, apoptosis, etc.).</td>
</tr>
<tr>
<td>Ankyrin repeat</td>
<td>411</td>
<td>PF00023.18</td>
<td>Tandemly repeated modules of ~33aa, occurring in a large number of functionally diverse proteins across eukaryotes. They are implicated in a variety of biological functions (transcription initiation, cell-cycle regulation, signal transduction, etc.).</td>
</tr>
<tr>
<td>Major facilitator superfamily</td>
<td>319</td>
<td>PF07690.5</td>
<td>One of the two transporter families found in all organisms. They are single-polypeptide secondary carriers capable only of transporting small solutes in response to chemiosmotic ion gradients.</td>
</tr>
<tr>
<td>Tetratricopeptide repeat clan</td>
<td>366</td>
<td>CL0020</td>
<td>Tetratricopeptide-like repeats are found in numerous and diverse proteins involved in such functions as cell-cycle regulation, transcriptional control, mitochondrial and peroxisomal protein transport, neurogenesis, and protein folding.</td>
</tr>
<tr>
<td>HEAT repeat</td>
<td>208</td>
<td>PF02985.10</td>
<td>Tandemly repeated, 37–47aa long modules frequently involved in intracellular transport processes.</td>
</tr>
<tr>
<td>Fungal Zn(2)-Cys(6) binuclear cluster domain</td>
<td>166</td>
<td>PF00172.7</td>
<td>A Cys-rich motif involved in zinc-dependent binding of DNA are found in a number of fungal transcriptional regulatory proteins. A wide range of proteins involved in metabolism are known to contain this domain.</td>
</tr>
<tr>
<td>Zinc finger, C_2H_2 type</td>
<td>122</td>
<td>PF00096.15</td>
<td>Composed of 25–30aa residues including 2 conserved Cys and 2 conserved His residues separated by 12 mainly polar basic residues. The domain is found in numerous nucleic acid-binding proteins (both RNA and DNA).</td>
</tr>
<tr>
<td>Short chain dehydrogenase</td>
<td>119</td>
<td>PF00106.14</td>
<td>Part of a large family of enzymes (~250–300 amino acids in length), most of which are known to be NAD- or NADP-dependent oxidoreductases.</td>
</tr>
<tr>
<td>Fungal specific transcription factor domain</td>
<td>119</td>
<td>PF04082.7</td>
<td>Domain found in a number of fungal transcription factors regulating a variety of cellular and metabolic processes.</td>
</tr>
<tr>
<td>Mitochondrial carrier protein</td>
<td>105</td>
<td>PF00153.15</td>
<td>Proteins consisting of up to 3 tandem repeats of a domain of approximately 100 residues, each domain containing 2 transmembrane regions. They are involved in energy transfer and are found in the inner mitochondrial membrane or are integral to the membrane of other eukaryotic organelles such as the peroxisome.</td>
</tr>
</tbody>
</table>

experimental work in A. nidulans identified the LaeA protein as a transcriptional regulator implicated in the regulation of metabolic gene clusters. Importantly, genetic manipulation of this regulator has the potential to reveal novel secondary metabolite compounds. For example, comparison of gene expression in a laeA deletion mutant strain relative to a wild-type strain led to the identification of a gene cluster responsible for the production of the antitumor compound terrequinone A, a secondary metabolite not previously known to be produced by A. nidulans. This finding suggests LaeA-based genome mining will be an invaluable asset to the elucidation of the secondary metabolome of aspergilli and its transcriptional regulation.
4.2.6 Physiology and Development

A number of interesting physiological and developmental characteristics can be inferred by studying the *A. nidulans* genome. The peroxisome is an organelle found in most eukaryotes, which is involved in beta-oxidation, the catabolism of fatty acids. In fungi, peroxisomes have been implicated in diverse processes such as secondary metabolism (penicillin biosynthesis), pathogenesis (aiding in the capturing of nematodes in carnivorous fungi), and growth (sealing septal pores in response to cellular damage as well as in hyphal growth). Examination of the *A. nidulans* proteome through similarity searches and peroxisome localization amino acid signals have led to the identification of several proteins implicated in fatty acid beta-oxidation, such as acyl-CoA oxidase, acyl-CoA dehydrogenase, ketoacyl-CoA thiolase, and Lon protease. Interestingly, beta-oxidation is occurring in both mitochondria and peroxisomes by two sets of paralogous proteins in *A. nidulans*, an organization also found in mammals but unlike *Saccharomyces cerevisiae*, where beta-oxidation occurs only in the peroxisomes.

Polarized hyphal growth is one of the hallmarks of filamentous fungi. The process consists of three molecular steps: the establishment of the positional cues that specify the site of bud emergence; the relay of the positional cue information to the morphogenetic machinery; and finally the recruitment of the morphogenetic machinery required for the remodeling of the cell surface at the bud site. A comparison of the *A. nidulans* proteome with the *S. cerevisiae* proteins involved in polarized hyphal growth revealed that the proteins participating in steps 2 (e.g., Rho-related GTPase signaling molecules) and 3 (e.g., actins, tubulins, etc.) are likely to be functionally conserved, whereas those participating in step 1 (e.g., yeast bud site markers Bud3p, Bud8p, and Bud9p) are either absent from the *A. nidulans* proteome or (more likely) poorly conserved between yeast and *A. nidulans*.

4.3 Gene Regulation

4.3.1 Conserved Noncoding Sequences and Prediction of Regulatory Motifs

A major research theme in genomics research is the identification of functional regulatory elements that govern gene expression. It is widely recognized that regulatory elements are likely to be more conserved in sequence relative to nonfunctional intergenic regions. To identify putative regulatory motifs we identified and aligned orthologous intergenic regions shared between the three distantly related genomes of *A. nidulans*, *A. fumigatus*, and *A. oryzae*. Through this approach, a total of 5801 orthologous intergenic regions that exhibited significantly better alignments than neutrally-evolving sequences were identified, corresponding to approximately 2% of all orthologous intergenic sequences. These regions were used as input for the prediction of putative regulatory motifs, using a modification of the algorithms pioneered by Kellis and colleagues in yeast. A total of 69 putative regulatory motifs were identified, and a subset of these motifs is shown on Figure 4.2. Interestingly, the list of predicted motifs includes several motifs that match known ones from *Aspergillus* or other Ascomycetes. For example, four of the regulatory motifs identified (rows 6–9 in Fig. 4.2) exhibit high similarity to the binding motifs associated with the Puf proteins in *S. cerevisiae*, a protein family implicated in the regulation of the posttranscriptional lives of mRNAs. Furthermore, one of the predicted regulatory motifs (row 2 in Fig. 4.2) was functionally tested in *A. nidulans* and shown to be the actual binding site involved in fatty acid induction of genes regulated by the farA and farB proteins (which encode Zn$_2$-Cys$_6$ binuclear DNA binding domains), highlighting the potential usefulness of good predictive algorithms for regulatory elements in filamentous fungi.

4.4 Genome Evolution

One of the most exciting aspects of the *Aspergillus* genomes that have been sequenced is that they span a range of evolutionary distances (Table 4.5). Whereas some species pairs are very close relatives, others are more distantly related, creating a gradient of species’ comparisons. For example, the similarity at the
amino acid level between A. nidulans and other Aspergillus species is approximately 75% (Table 4.5), an evolutionary distance roughly equivalent to that between the human and the fish genome.2 On the other end of the evolutionary distance spectrum there are species pairs that show a high degree of similarity; A. flavus and A. oryzae are ~99% similar, N. fischeri and A. fumigatus are ~96%, whereas the two A. niger-sequenced strains, strain ATC 1015 (sequenced by the Joint Genome Institute; http://genome.jgi-psf.org/Aspni1/Aspni1.home.html; Table 4.1) and strain CBS 513.88 (sequenced by DSM Food Specialties;7 Table 4.1) are 97% similar7 (the strain sequenced by JGI is the one used for the comparisons reported in this essay). Finally, a number of species exhibit intermediate levels of amino acid similarity (e.g., A. clavatus and A. fumigatus are 84% similar). The combination of the profoundly different lifestyles

### Table 4.5

Percentage of Uncorrected Pairwise Amino-Acid Divergence (Bottom Left) and Similarity (Upper Right) from an Alignment of 2753 Orthologous Proteins Across Aspergillus Genomes

<table>
<thead>
<tr>
<th>C. immitis (%)</th>
<th>A. clavatus (%)</th>
<th>N. fischeri (%)</th>
<th>A. flavus (%)</th>
<th>A. fumigatus (%)</th>
<th>A. niger (%)</th>
<th>A. oryzae (%)</th>
<th>A. terreus (%)</th>
<th>A. nidulans (%)</th>
</tr>
</thead>
</table>
The Aspergilli

exhibited by each of these species coupled with the varying degrees of evolutionary affinity exhibited by their genomes offers a unique opportunity to study key questions in evolutionary genomics.

4.4.1 Phylogenetic Relationships

The phylogenetic analysis of the large subunit of ribosomal DNA from a large number of *Aspergillus* species by Peterson is the most widely accepted hypothesis as to the species relationships within the genus (Fig. 4.3a). We reconstructed the phylogenetic relationships between *A. nidulans* and the *Aspergillus* species for which genome sequences are available, using *Coccidioides immitis* as the outgroup. Maximum likelihood and maximum parsimony analysis of 30 highly conserved genes from a previously used data matrix, and maximum parsimony analyses of the 2753 shared orthologs across all species provide support for a phylogeny (Fig. 4.3b) that differs from Peterson’s phylogeny in two key aspects. First, in the revised phylogeny *A. nidulans* is placed as an early-branching species outside the clade formed by *A. fumigatus*, *A. oryzae*, and their relatives. This finding is in agreement with other published genome-scale analyses, and seems well supported. Second, the revised phylogeny indicates that *A. terreus* is the likely sister taxon to the *A. oryzae* and *A. flavus* clade—not *A. niger* as suggested by Peterson’s phylogeny. Although this finding is in agreement with one other published genome-scale study, in both studies the support (as provided by bootstrap values and the percentage of single-gene parsimonious trees) for this inference is moderate.

4.4.2 Conserved Synteny

To address the extent to which synteny is conserved across *Aspergillus* genomes we developed an algorithm based on hierarchical clustering that identifies regions of conserved synteny while also retaining information about the internal microarrangements. Application of this algorithm to the comparison of the *A. nidulans* genome with those of *A. fumigatus* and *A. oryzae* revealed extensive structural rearrangement between these three species. We applied this algorithm to compare the *A. nidulans* genome to those of the other *Aspergillus* species and *Coccidioides immitis* (Fig. 4.4). A large fraction of the *A. nidulans* genome appears to be conserved in synteny to all the other *Aspergillus* species and, to a lesser extent, to *C. immitis*. Viewed from a genome structure perspective, these areas of conservation can be thought of as representing a core structural *Aspergillus* genome. Interestingly, these conserved syntentic regions are

FIGURE 4.3 Phylogenetic relationships within the genus *Aspergillus*. Note: Only the species for which genome sequences are available are shown. (a) Species relationships according to Peterson’s study. Although *A. oryzae* and *C. immitis* were not included in Peterson’s study, we have included them here for easier comparison with the phylogeny shown in panel B. (b) Phylogenetic relationships according to genome-scale analyses (see main text for details). *C. immitis* was used as the outgroup. Numbers above branches indicate the bootstrap values generated by an analysis of 30 evolutionarily conserved genes by maximum likelihood and parsimony, respectively. From the numbers below branches, the value on the left indicates the bootstrap support from a parsimony analysis of 2753 concatenated orthologs, whereas the value on the right indicates the percentage of individual gene trees that supported that branch. *Note* that the placement of *A. nidulans* in the phylogeny shown on panel B is in agreement with published genome-scale analyses and that the placement of *A. niger* and *A. terreus* relative to each other and other *Aspergillus* taxa in panel B is weakly supported by the individual gene tree analyses.
Aspergillus nidulans Genome

51

frequently interrupted by large regions lacking detectable long-range synteny. As is evident from Figure 4.4, the presence of these regions that lack synteny is typically (but not always) associated with repeats and regions proximal to telomeres.

4.4.3 Genome Size

One notable difference between the currently available Aspergillus genomes is their apparent difference in size (Table 4.1). Specifically, whereas the genomes of A. oryzae, A. flavus, and A. niger range between 34 and 37 Mb in size, the genomes of all the other species range between 28 and 32 Mb. Whether this difference is due to increases in the genome sizes of A. oryzae, A. flavus, and A. niger or due to decreases in some or all other species remains an open question. Interestingly, the phylogeny shown in Figure 4.3b raises the hypothesis that the A. oryzae/A. flavus clade may have increased its genome size independently from A. niger. However, it is also possible that the ancestor of all three species had a large genome and that A. terreus subsequently underwent a decrease in size.

A comparison of the genome of A. oryzae with those of A. fumigatus and A. nidulans found large blocks of sequence specific to A. oryzae, a finding in support of the hypothesis of a genome expansion in the A. oryzae lineage. Furthermore, functional annotation of the three proteomes revealed a significant gene enrichment in the A. oryzae lineage relative to the two other genomes, particularly in the set of genes

FIGURE 4.4 A comparative map of the Aspergillus genomes, using the A. nidulans genome as a reference. Note: The eight large columns correspond to each one of the eight chromosomes of the A. nidulans genome. Within each large column, there are eight thinner columns corresponding to the projections of the genomes of A. fumigatus, A. oryzae, A. clavatus, A. flavus, N. fischeri, A. terreus, A. niger, and C. immitis, respectively, to the genome of A. nidulans. The thin columns are colored by chromosome or genomic scaffold as indicated by the key on the left (only the 10 largest genomic scaffolds are shown for those species whose assemblies are not available at the chromosome level). The line next to the large columns corresponds to the density of repeats in the A. nidulans genome (increasing values on the right).
participating in metabolism and secondary metabolism. Examples of gene families that appear enriched in \textit{A. oryzae} include the cytochrome P450 genes, the WA-like PKS genes, as well as transporters, hydrolytic enzymes, and secretory proteases. However, phylogenetic analyses of some of these protein families are inconsistent with the hypothesis of genome expansion occurring in the \textit{A. oryzae} lineage. Specifically, if the hypothesis of gene family enlargement in the \textit{A. oryzae} lineage is correct, one would expect to observe several pairs of duplicates in \textit{A. oryzae} to exhibit a greater degree of similarity and closer phylogenetic affinity relative to their homologs in \textit{A. nidulans} and \textit{A. fumigatus}. In contrast, the phylogenies for several protein families indicate that the gene duplications observed in \textit{A. oryzae} are likely to have occurred in the ancestor of all three species, followed by retention of the duplicated genes in \textit{A. oryzae} and loss of the duplicates in \textit{A. fumigatus} and \textit{A. nidulans}. It should be noted that increases in gene number are only one of several forces with the potential to affect the genome size of an evolutionary lineage, and that an explanation of the size differences among \textit{Aspergillus} genomes may also involve other factors, such as repetitive elements. Interestingly, the fraction of the genome’s size in each of the species accounted for by repetitive elements is very small (<2%), with TEs in \textit{A. oryzae} accounting for a smaller sequence fraction than those in \textit{A. nidulans} and \textit{A. fumigatus} (Rokas and Galagan, unpublished data).

4.4.4 Reproductive Strategy

Reproductive strategy in \textit{Aspergillus} is governed by two mating-type genes that are implicated in the establishment of mating compatibility; a high-mobility group (HMG) transcription factor and an alpha box transcription factor. Sexual species can be classified into two categories with respect to their sexual compatibility, homothallics and heterothallics. Individuals in homothallic species are self-fertile and possess both genes, often although not always on the same chromosome. In contrast, individuals in heterothallic species are obligate outcrossers (i.e., they self-sterile) and possess only one of the two mating-type genes; as a result, they can only mate with individuals possessing the opposite mating gene. Importantly, the MAT genes in both mates occupy the same genomic locus and thus the MAT genes in heterothallics look like alleles, although they are evolutionarily unrelated, and are generally known as idiomorphs.

\textit{A. nidulans} is a sexual, homothallic species and both the HMG and alpha mating-type genes have been identified in distinct genomic regions in chromosomes 3 and 6, respectively (Fig. 4.5). In contrast, \textit{A. oryzae}, \textit{A. fumigatus}, \textit{A. clavatus}, and \textit{A. niger}—four species only known from their asexual stage—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species—possess a genome structure typical of heterothallic species. Several data points support this possibility. First, population screenings of haploid individuals from three putative asexuals—\textit{A. fumigatus},

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.5.png}
\caption{The genomic location of the mating loci in the \textit{A. nidulans} genome\textsuperscript{2} and their similarity to the mating loci found in the genomes of the putative asexuas: \textit{A. fumigatus}\textsuperscript{4} and \textit{A. oryzae}\textsuperscript{3}. Note: Gray lines indicate regions of alignment between genomic fragments. Genes in genomic fragments are denoted by boxes. The alpha mating locus is shown in dark gray and the HMG mating locus in light gray. The boxes depicting the mating loci have been significantly enlarged relative to the rest of the genes for easier visualization.}
\end{figure}
Aspergillus nidulans Genome

A. oryzae, and A. niger—show that populations of these species possess both mating-type genes in approximately equal frequencies.7,42 Second, the gene machinery involved in the processes of mating, pheromone response, meiosis, and fruiting body development found in the sexually reproducing A. nidulans is also found intact in A. oryzae, A. fumigatus as well as in several other putative asexual Aspergillus species (Rokas and Galagan, unpublished data). Finally, in the only putative asexual species so far examined—A. fumigatus—sequence analysis has provided strong evidence supporting the occurrence of recombination.42 In summary, these data argue that the apparently 114 Aspergillus species that are only known from their asexual stage43 may be in reality sexual species for which the sexual stage of their development remains as yet undiscovered. Importantly, the presence of a sexual stage in certain pathogenic Aspergillus species, such as A. fumigatus, indicates that classical genetic analyses may soon become feasible.

Alignment of the genomic regions containing the MAT genes between A. nidulans, A. fumigatus, and A. oryzae indicates that the relative order of genes in these genomic regions (synteny) has remained largely conserved in all three species (Fig. 4.5). Specifically, the genomic regions of A. oryzae and A. fumigatus exhibit extensive conservation and extend for several hundred Mb upstream and downstream of the MAT genes, suggesting that these regions are homologous. Within this homologous region the MAT genes occupy nearly identical positions, although they are offset with different orientations. Furthermore, the genomic region downstream of the A. oryzae and A. fumigatus genomic region shows extensive synteny with the region downstream of the A. nidulans HMG mating-type gene, whereas the genomic region upstream of the A. oryzae and A. fumigatus genomic region shows extensive synteny with the region upstream of the A. nidulans alpha mating-type gene (Fig. 4.5), suggesting that at least one rearrangement must have occurred since these species diverged from their last common ancestor.

From these data a model of evolution that accounts for the evolution of reproductive strategies can be proposed. The ancestral Aspergillus was a homothallic species with the mating-type genes in physically close linkage. From this ancestor, the existing state of A. nidulans was obtained through a chromosomal translocation, whereas the heterothallic genome structure of A. fumigatus, A. oryzae, and their relatives was obtained through complementary degeneration of each of the two mating-type genes, with some individuals losing the HMG-type gene and others losing the alpha-box type gene. It is also possible that the ancestor of all three species was heterothallic and that the transition to homothallism occurred in the A. nidulans lineage. However, the offset orientation of the MAT loci in A. oryzae and A. fumigatus and the existence of a fragment of the HMG gene neighboring the A. fumigatus alpha locus are consistent with a scenario of gene loss from a homothallic ancestor2,42 and argue against a transition in the opposite direction.

4.5 Future of Aspergillus Genomics

These first analyses of the Aspergillus genomes are but a small step toward a comprehensive understanding of the genomics of this medically, industrially, and agriculturally important fungal genus. Undoubtedly, the available genome sequences have tremendously enhanced our understanding of Aspergillus biology and will continue to do so for a long time. Important questions such as a mechanistic understanding of the dramatic changes in genome size across species or an evolutionary explanation of the sudden origins and losses of the diverse sets of secondary metabolite gene clusters in several lineages, still remain largely unanswered. As new methodology is developed to increase the information retrieved from comparative analyses and assays that facilitate functional genomics studies on Aspergillus become available, it is likely that several further fundamental insights will be gained on the intriguing scientific mysteries underlying the metabolic and physiological diversity of this important fungal genus.

Acknowledgments

Research on Aspergillus at the Microbial Sequencing Center at the Broad Institute is funded by grants from the NIH.
The Aspergilli

References

5

Aspergillus nidulans Linkage Map and Genome Sequence: Closing Gaps and Adding Telomeres

A. John Clutterbuck and Mark Farman

CONTENTS
5.1 Linkage Map .............................................................................................................. 57
5.2 Gap Closure in the Genome Sequence ....................................................................................... 58
5.3 Telomeres ................................................................................................................ .................... 58
5.4 Subtelomere Domains ...................................................................................................... ........... 59
5.5 Telomere-Linked Helicases (TLH) ............................................................................................. 60
5.6 Further Alignment of Genome and Linkage Maps ..................................................................... 61
5.7 Recombination Frequencies ................................................................................................ ........ 62
Acknowledgments ................................................................................................................................ 64
References ............................................................................................................................................ 73

5.1 Linkage Map

G. Pontecorvo, in initiating the use of A. nidulans as a genetic model, was primarily interested in the nature and definition of the gene, and in particular the phenomenon of intragenic recombination [1]. It was therefore logical that his laboratory should build up an intergenic recombination map as a framework for these studies [2]. Since recombination rates are high in A. nidulans (see later), this endeavor was greatly helped by the discovery of the parasexual cycle [3]. This consists of three relatively rare and independent steps: (1) formation of heterozygous vegetative dipoles by fusion of unlike nuclei in heterokaryons, (2) mitotic recombination between homologous chromosomes, which can order markers along a chromosome, irrespective of distance, and (3) random assortment of chromosomes at haploidization, allowing location of a new marker to its chromosome and detection of translocations [4]. Haploidization has been regularly used as a first step in mapping new mutants, and since entire chromosomes assort independently, mutants can be immediately assigned to chromosomes, which are numbered in order of their discovery.

The markers used in constructing the linkage map include auxotrophs, resistance mutants and conidial pigmentation mutants, the last of which are especially valuable for instant recognition of hybrid cleistothecia, and for detection of recombinant sectors in parasexual analysis. Most of these mutants were induced by mutagen treatment, an unwanted side effect of which, discovered later, was the not infrequent induction of chromosome aberrations [5]. Interchromosomal translocations are readily detected during haploidization but other aberrations have only been surmised as explanations for anomalous meiotic linkage. Unfortunately, since Aspergillus chromosomes do not lend themselves to detailed cytogenetic analysis [6], these aberrations remain as inferences.

Despite problems introduced by chromosome aberrations, and the fact that linkage maps are compilations of data from many different laboratories, remarkably self-consistent maps have been published [7,8]. Physical mapping of cloned genes located to cosmids has also been tried [9]: the initial step of
hybridization of cosmids to CHEF gel-separated chromosomes [10] proved very successful and has given results which correlate well with genomic scaffolds. Identification of genes on the same cosmid has also been fruitful, but attempts to order contigs by cross-hybridization [9] have yielded few results conforming to the genome sequence, probably due in part to a wider distribution of repeated elements than had been predicted [11] (see Chapter 20 in this volume).

5.2 Gap Closure in the Genome Sequence

The published Broad Institute genome (http://www.broad.mit.edu/annotation/fungi/aspergillus/) comprises 173 contigs linked by end-sequenced BAC and fosmid bridges into 16 scaffolds. These scaffolds were initially anchored to the linkage map using 91 meiotically linked markers plus 61 markers located to chromosomes only by haploidization or hybridization. A further 75 contigs remained unaligned with the linkage map. However, Blast searches, using these unattached contigs or the ends of aligned contigs as queries, reveals many overlaps, detailed in Table 5.1 through 5.17. As a result, 58 previously solitary contigs can be incorporated into map-aligned scaffolds, leaving only 17 unplaced contigs, and reducing the bridged gaps within scaffolds from 158 to 71. Contig overlaps include 39 duplicated or partially overlapping autocalled gene pairs (Broad Institute gene prediction set 2).

5.3 Telomeres

The genome sequence does not include centromeres, and the telomeres, which in A. nidulans are characterized by the terminally repeated motif (TTAGGG), are poorly represented. Only four aligned scaffolds bear typical telomeric repeats at their ends. A fifth contig (226) containing such repeats is present in the genome assembly but is not yet anchored to the genetic map. Searches of the NCBI sequence trace archive identified a large number of candidate telomere reads that were not used in the genome assembly. Assembly of these sequences using the TERMINUS program led to the identification of 11 contigs that terminate in telomere repeats [12]. These “TelContigs” ranged in length from 973 to 1055 bp, and were very robust because each one was represented by a large number of individual sequence reads (Table 5.18). In fact, the average depth of sequence coverage per telomere (~ 40×) was approximately four times that achieved for the genome as a whole. Five of the TelContigs corresponded to the telomeres already present in the assembly, leaving six that were expected to represent new chromosome ends. These six sequences consisted entirely of terminal reads—none of the constituent reads overlapped with any of the “internal” sequences used in the assembly. As a result, the TelContigs were very short, which explains why they had not been included in the genome sequence in the first place.

Due to the absence of sequence overlaps, it was not possible to identify direct links between the newly-identified telomeres and the genome assembly. Therefore, it was necessary to try to establish links using assembled mate-pair sequences derived from the “sub-telomeric” ends of the relevant clones (SubTelContigs). This was performed using the linking and validation modules built into TERMINUS. Briefly, this involves using the assembled subtelomeric sequences as queries to search the genome sequence using BLAST, employing TruMatch [13] to identify bona fide links and, finally, performing “positional consistency” checks to ensure that each match is within an acceptable distance from the end of the relevant contig and scaffold. Five of the new TelContigs were assigned unique genomic positions in this manner, while the sixth one (TC11) linked up with contigs in six different genomic scaffolds (Table 5.19). These multiple associations were not due to the presence of repetitive sequences, however, because most of the individual SubTelContigs that were associated with TC11 exhibited a unique match to the genome. Instead, this scenario suggested that TC11 comprises reads from as many as six different chromosome ends, and that these ends have the same sequence adjacent to the telomere repeat but diverge in sequence as one moves in toward the respective centromere. Consistent with this interpretation was the finding that TC11 contained almost four times as many reads as the average number of reads in the other TelContigs (151 versus 44). In addition, careful inspection of the individual reads making up TC11 revealed subsets of sequences that differed from the consensus by small numbers of point mutations,
thereby confirming that this TelContig represents multiple telomeres. In total, TERMINUS allowed 10 new telomeres to be unequivocally positioned relative to the genome assembly and, in doing so, identified the genomic contigs that occupy terminal chromosome locations (Table 5.19). In addition, it was possible to estimate the telomere-to-assembly gap sizes using knowledge about the insert lengths of the clones that were sequenced. The approximate gap lengths are also listed in Table 5.19.

Contig 92 exhibited numerous blast matches to SubTelContigs associated with TC11, and the positions of those matches were within an acceptable distance of the contig’s end. However, contig 92 was assembled in the middle of scaffold 6, almost 500 kb from the nearest end. Therefore, it would appear that contig 92 is misassembled. Alternatively, it could contain a sequence duplication that is not represented in the assembly. Tentatively, we have assigned the contig 92, 226-TC11 scaffold as the 16th telomere, which maps by default to the end of the chromosome VR arm (Table 5.19 and Fig. 5.6).

TelContigs and SubTelContigs can be viewed in a web browser at the URL: http://genome.kbrin.uky.edu/fungi_tel/index.html.

5.4 Subtelomere Domains

The general organization of eukaryotic chromosome termini is highly conserved. In particular, detailed characterizations of telomeres in microbes and humans have revealed the ubiquitous presence of distinct subtelomere domains, which consist of sequences that are duplicated at multiple chromosome ends [14–16]. The fact that TC11 can be mapped to the ends of six different \textit{A. nidulans} scaffolds indicates that this TelContig contains the distal end of a distinct subtelomere sequence. Considering that some of the gaps between the TC11 sequence and the assembly were quite small, we suspected that the centromere-proximal ends of the subtelomere domains could be found on the ends of the newly-identified telomeric scaffolds. Therefore, in order to characterize further the \textit{A. nidulans} subtelomere domain, the sequences from the distal portions of the relevant scaffolds were retrieved, transposon sequences were masked and BLAST was then used to search for sequence similarities. All six of the telomeric scaffolds exhibited matches to at least one other and, in every case, at least one of the alignments continued all the way to the telomeric end of the scaffold (Fig. 5.1). In contig 133 the centromere-proximal portion of the subtelomere was duplicated in a tandem fashion. There was no evidence of a discrete border to the subtelomere region, and the point of transition into chromosome-unique sequences varied from one end to another (Fig. 5.1). Taking into account the gaps between the scaffolds and TC11, we estimate that the canonical subtelomere domain of \textit{A. nidulans} is approximately 16.5 kb in length, which is similar to the \textit{Magnaporthe oryzae} subtelomere [16]. In \textit{A. nidulans}, the alignments between different subtelomeres were discontinuous. At first, we suspected that this was due to insertion of transposable elements. Surprisingly, however, we found that, although transposons were often found immediately adjacent to the subtelomeres, there were no insertions within the domains. Instead, the discontinuous alignments were due to the presence in some domains of unique sequence motifs (not shown in Fig. 5.1). In addition to breaks in their alignments, the subtelomeres exhibited significant sequence divergence (up to ~15%), due to the existence of large numbers of G to A and C to T transition mutations. This suggests that the \textit{A. nidulans} subtelomeres have been extensively mutated by a repeat-induced point mutation (RIP)-like process (see Chapter 20 in this volume).

Inspection of the telomere-adjacent sequences in TC11 revealed that the distal portion of the \textit{A. nidulans} subtelomere domain has a high percentage of adenine and thymine residues (80%). Even more striking was the observation that there are only 15 cytosine residues in the terminal 900 base pairs. AT richness is also a feature of the distal region of the subtelomere in \textit{M. oryzae} [16] although there was no obvious bias against C residues in that fungus. In most eukaryotes, the subtelomeres contain blocks of short tandem repeats (STRs) [14–15]. Analysis of the \textit{A. nidulans} subtelomere domains using Tandem Repeats Finder [17] revealed 15 different tandemly-repeated sequences. The consensus sequences for the individual repeat units, and their copy numbers are listed in Table 5.20. Motif F is particularly interesting because it contains the short telomere motif CCTAACCC and is found upstream of telomere-linked...
helicase (TLH) genes (see later). A similar motif, CCTAACC, occurs upstream of the TLH genes in *M. oryzae* [16], suggesting that STRs containing short telomere-like sequences could be involved in TLH regulation. Surprisingly, despite the extensive sequence overlap between subtelomeric sequences at the different chromosome ends, many of the tandem repeats were restricted to a single subtelomere and none of the motifs were ubiquitously present. Careful scrutiny of the BLAST alignments revealed that this uneven distribution is due to the presence in these regions of abundant RIP mutations. Given their patchy distribution among the subtelomeres, it would appear that the tandem repeats are not essential for telomere maintenance and/or function. Alternatively, it may be that inexact repeats are sufficient.

### 5.5 Telomere-Linked Helicases (TLH)

Interrogation of the predicted gene list revealed several candidate genes within the subtelomere regions. The majority of these genes code for conceptual proteins with strong similarity to the telomere-linked helicases—a special class of fungal helicases that are encoded exclusively by telomere-linked genes (C. Rehmeyer and M. Farman, unpublished data). Only one of the predicted TLH genes in *A. nidulans* is full-length, as mutations, many of which are characteristic of RIP (Chapter 20 in this volume), in the other copies are expected to result in truncated, or alternatively spliced versions. The positions of the TLH genes within the subtelomere are shown in Figure 5.1. The other subtelomere-localized genes encode proteins of unknown function.

BLAST searches detected a number of TLH genes that were not identified by the gene-calling programs and which, therefore, are probably pseudogenes. Two of the pseudogenes were found at the ends of the known telomeric scaffolds 146 and 44, and there is also a fragment in contig 1. Copies were also found on contigs 91 and 92, thereby supporting our earlier conclusion that both are linked to telomeres and, as such, may have been misassembled. Another copy was found on contig 177, adjacent to an 11.5 kb segment shared with contigs 92 and 171. Contig 177 has not been definitively
linked to a telomere as yet, but might map in the gap between contig 86 and TC11. Alternatively, given the overlaps with both contigs 92 and 171, it is possible that it should be merged with one of their scaffolds.

5.6 Further Alignment of Genome and Linkage Maps

Over 160 meiotically mapped markers can now be identified with autocalled genes, reinforcing a good correspondence (Fig. 5.2 through 5.9) between the genome sequence and the linkage map [8], updated in the website: http://www.gla.ac.uk/Acad/IBLS/molgen/aspergillus/. The majority of alignment discrepancies can be assigned to inaccuracy of linkages data for new mutants, obtained at a time when no close markers were available; such alignments are indicated by dotted lines in Figure 5.2 through 5.9; e.g., some \textit{acuH} and \textit{uvsJ} (Fig. 5.6), were initially crossed only to other mutants of the same type, yielding only loose linkages. \textit{aromC} (I-R) and the \textit{stc} sterigmatocystin biosynthesis cluster (IV-R) are examples of genes misplaced on the basis of cosmid cross-hybridization.

Chromosome \textit{V} presents the most serious mapping challenge. This chromosome is expected to be interrupted by the nucleolar organizer [10], which is not included in the Broad genome. Furthermore, the linkage map of chromosome \textit{V} has suffered from the regular use of markers \textit{lysB} and \textit{nicA}, which

![FIGURE 5.2](image_url) Genome-linkage map alignments for \textit{A. nidulans} chromosome I. Note: Bold horizontal line: genomic supercontigs, numbered above arrow showing orientation. Bold vertical lines indicate telomeres, numbered T1–T16. Lower thin horizontal line: linkage map, measured in centiMorgans (cM), calculated assuming no interference [7, 18]. Longer intervals were obtained by adding constituent cM values. A broken line indicates >50% meiotic recombination. Diagonal lines connect locations of specified genes. Dashed lines represent linkage map positions based on weak linkages.

![FIGURE 5.3](image_url) Genome-linkage map alignments for \textit{A. nidulans} chromosome II. Note: See Figure 5.2 for explanation and scales.

![FIGURE 5.4](image_url) Genome-linkage map alignments for \textit{A. nidulans} chromosome III. Note: See Figure 5.2 for explanation and scales.
The Aspergilli

were closely linked in all crosses, but could not be reliably ordered with respect to other markers, suggesting the presence of an inversion or other aberration in some mapping strains. Nevertheless, the linkage map agrees with the revised arrangement of contigs resulting from the location of a telomere in association with contig 92. This splits scaffold 6 into two parts, here designated as 6a and 6b (Tables 5.6 and 5.7; Fig. 5.6). The displaced contigs 90-87 (scaffold 6a) include markers lysB, sepB, and pmtA that map in the central region of V, the last two markers now proving to be linked to acuN, that is known to encode enolase and is uniquely associated with gene AN5746 in contig 98. Moreover, the end of contig 90 includes a fragment of the ribosomal rRNA repeat cluster [19] (M.J. Anderson, personal communication), thereby identifying the break between scaffolds 12 and 6a as the position of the nucleolar organizer. This all agrees with mitotic recombination [20] that gives the most likely position of the centromere as between lysB and facA, that is, between contigs 87 and 98 (Fig. 5.6).

Examination of six fosmid and three BAC links found by the Broad Institute between what are now designated as scaffolds 6a and 6b, proves that none are unambiguous: one link is to a Gypsy-1 element, of which there are many copies in the genome (see Chapter 20 in this volume), one could alternatively be an internal link within contig 92, and seven others are between 6b contigs and sequences common to a number of TelContig TC11-associated sequences, supporting the presence of a telomere at the end of this region, although not identifying a specific TelContig.

5.7 Recombination Frequencies

It is evident from Figure 5.2 through 5.9 that recombination per chromosome length varies considerably. In Figure 5.10, recombination frequencies, ignoring outlying values, range from 0.5% to 2.7% per 10 kb. While some of this variation may be due to chromosome aberrations and heterogeneity of data sources, Figure 5.2 through 5.9 suggest that recombination is reduced near centromeres and the nucleolar organizer, and high in the center of each arm. This is especially evident for the well-mapped chromosome II
Aspergillus nidulans *Linkage Map and Genome Sequence*

**FIGURE 5.7** Genome-linkage map alignments for *A. nidulans* chromosome VI. *Note:* See Figure 5.2 for explanation and scales.

**FIGURE 5.8** Genome-linkage map alignments for *A. nidulans* chromosome VII. *Note:* See Figure 5.2 for explanation and scales.

**FIGURE 5.9** Genome-linkage map alignments for *A. nidulans* chromosome VIII. *Note:* See Figure 5.2 for explanation and scales.

**FIGURE 5.10** Recombination per genome interval. *Note:* Each point represents an individual, uncorrected percentage of recombinants obtained in a cross between markers identified in the genome sequence.
The Aspergilli (Fig. 5.3), and has been reported in detail for markers close to centromeres of chromosomes III and IV [21]. Lack of terminal markers, other than the penicillin biosynthesis cluster, \( npeA \) (VI-R) means that little can be said about recombination near telomeres.

Acknowledgments

The telomere analysis was supported by a grant to M.L.F. from the National Science Foundation (MCB0135462) and a subcontract to Chuck Staben from the Kentucky Biomedical Research Infrastructure Network (National Center for Research Resources grant 5P20RR016481-03, awarded to Nigel Cooper of the University of Louisville).

### TABLE 5.1
Supercontig 1, Linkage Group VIII-R

<table>
<thead>
<tr>
<th>Contig*</th>
<th>Length</th>
<th>Overlapb</th>
<th>Mismatchesc</th>
<th>Notesd</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10,851</td>
<td>—</td>
<td></td>
<td>T9</td>
</tr>
<tr>
<td>2</td>
<td>168,814</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>127,977</td>
<td>536</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>230(−)</td>
<td>5336</td>
<td>575</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>155,056</td>
<td>558</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>246(−)</td>
<td>3023</td>
<td>560</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>450,355</td>
<td>513</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>186(+)</td>
<td>10198</td>
<td>483</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>266,891</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>773,302</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>49390</td>
<td>2818, 2817</td>
<td>3e</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>14,258</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>175,166</td>
<td>747, 766</td>
<td>35s</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>61,200</td>
<td>494</td>
<td>2s</td>
<td></td>
</tr>
<tr>
<td>180(−)</td>
<td>11,235</td>
<td>751</td>
<td>1e, 1s</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>174,100</td>
<td>2695</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>249,512</td>
<td>1605</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>364,334</td>
<td>1751</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>178,329</td>
<td>790, 791</td>
<td>18e</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>451,587</td>
<td>1270</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>347,844</td>
<td>616</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>225(−)</td>
<td>5462</td>
<td>601</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>175,793</td>
<td>471</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>192(−)</td>
<td>9030</td>
<td>574</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>71,437</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21,570</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2196</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Orientation of previously unlocated contigs is shown in brackets.

bOverlap with next contig. Where the contigs differ by insertion/deletion, two figures are given, the first for the contig listed, the second for the following contig.

cMismatches in overlapping region; terminal mismatches are indicated by “e” = near end of contig listed or “s” = near start of next contig.

dTelomeres are indicated e.g., “T5” = telomere T5 at end of contig, “→T3” = link to telomere T3. Other notes are numbered and listed below each table.
### TABLE 5.2
Supercontig 2, Linkage Group VII-R

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>175(+)</td>
<td>19,880</td>
<td>728</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>276,296</td>
<td>641</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>224(−)</td>
<td>5623</td>
<td>1214, 1213</td>
<td>11e</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>67,024</td>
<td>943</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>213(−)</td>
<td>6639</td>
<td>226, 227</td>
<td>13e</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>43,525</td>
<td>1205</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>229,102</td>
<td>353</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>182(+)</td>
<td>10,620</td>
<td>402</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>389,371</td>
<td>(−)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>113,764</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>98,898</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>525,887</td>
<td>448</td>
<td></td>
<td>40s</td>
</tr>
<tr>
<td>30</td>
<td>89,802</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2490</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>435,990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>2149</td>
<td>1680</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>244(+)</td>
<td>3100</td>
<td>1316, 1317</td>
<td></td>
<td>23e</td>
</tr>
<tr>
<td>34</td>
<td>177,275</td>
<td>(−)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>84,073</td>
<td>73, 74</td>
<td>29e</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>96,901</td>
<td>709</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>241(−)</td>
<td>3865</td>
<td>566</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>75,346</td>
<td>568</td>
<td>23e</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>233,386</td>
<td>997</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>242(+)</td>
<td>3786</td>
<td>808</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>240,080</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220(+)</td>
<td>6090</td>
<td>400, 401</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>150,763</td>
<td>(−)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>64,302</td>
<td>327</td>
<td>50e</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>99,573</td>
<td>126, 128</td>
<td>21e</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>338,545</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: (1) (C) at end of contig 26, (C) at the start of contig 27; (2) Gypsy element with matching target-site duplications forms probable bridge of 4 kb gap to contig 35; (3) Cloned gene *mnpA* (NCBI accession AF497720) overlaps contig 40 by 1169 bp (0 mismatches) and contig 41 by 584 bp (4 mismatches).

*a–d* See Table 5.1.
### TABLE 5.3

Supercontig 3, Linkage Group VI-R

<table>
<thead>
<tr>
<th>Contig&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Length</th>
<th>Overlap&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Mismatches&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Notes&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>17,771</td>
<td>—</td>
<td>—</td>
<td>→T10</td>
</tr>
<tr>
<td>45</td>
<td>83,158</td>
<td>656</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>235(−)</td>
<td>4760</td>
<td>1009</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>103,482</td>
<td>3366</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>137,873</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>95,913</td>
<td>894, 892</td>
<td>6e</td>
<td></td>
</tr>
<tr>
<td>227(+)</td>
<td>5437</td>
<td>340, 341</td>
<td>13e, 3s</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>206,454</td>
<td>1542</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>47,652</td>
<td>904</td>
<td>1s</td>
<td></td>
</tr>
<tr>
<td>189(−)</td>
<td>10,019</td>
<td>914</td>
<td>3s</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>1114,266</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>82,934</td>
<td>1315</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>51,172</td>
<td>763</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>231(+)</td>
<td>5185</td>
<td>512</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>183,474</td>
<td>1612</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>492,661</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>56</td>
<td>45,063</td>
<td>3028</td>
<td>128s</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>28,094</td>
<td>1366</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>83,422</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a–d</sup>See Table 5.1.

*Note:* Simple sequences at end of contig 55 and start of contig 56.

### TABLE 5.4

Supercontig 4, Linkage Group II-R

<table>
<thead>
<tr>
<th>Contig&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Length</th>
<th>Overlap&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Mismatches&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Notes&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
<td>177,453</td>
<td>740, 741</td>
<td>3s</td>
<td>→T4</td>
</tr>
<tr>
<td>222(+)</td>
<td>5781</td>
<td>474</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>68,571</td>
<td>(−)</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>61</td>
<td>865,047</td>
<td>577</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>191(−)</td>
<td>9200</td>
<td>525</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>201,480</td>
<td>330</td>
<td>4s</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>41,959</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>203,363</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>278,071</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>88,283</td>
<td>486</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>236(−)</td>
<td>4673</td>
<td>616, 617</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>254,221</td>
<td>75</td>
<td>25e</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>250,663</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>80,995</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>41,226</td>
<td>4410, 4407</td>
<td>10e</td>
<td></td>
</tr>
<tr>
<td>174(+)</td>
<td>26,679</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>8494</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a–d</sup>See Table 5.1.

*Note:* 476 nt overlap of I-1 retrotransposons, but 52 mismatches: probably separate elements.
TABLE 5.5
Supercontig 5, Linkage Group III-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>17,816</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>3923</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>12,387</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>213,493</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>153,210</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>132,523</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>179(−)</td>
<td>11,756</td>
<td>657,658</td>
<td>5s, 1e</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>353,456</td>
<td>584</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>238(+)</td>
<td>4644</td>
<td>287,288</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>209,053</td>
<td>1847</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>80</td>
<td>335,440</td>
<td>444</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>201(−)</td>
<td>7389</td>
<td>1628</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>156,263</td>
<td>698</td>
<td>2e</td>
<td></td>
</tr>
<tr>
<td>183(−)</td>
<td>10,063</td>
<td>640</td>
<td>2s</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>44,985</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>55,015</td>
<td>995</td>
<td>2e</td>
<td></td>
</tr>
<tr>
<td>233(+)</td>
<td>4981</td>
<td>953</td>
<td>1s</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>621,812</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>34,531</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>112,247</td>
<td>—</td>
<td>—</td>
<td>→T11</td>
</tr>
</tbody>
</table>

*See Table 5.1.
Note: (1) 1-1 retrotransposon possibly bridges 3467 nt gap; (2) mismatches clustered in middle of overlap.

TABLE 5.6
Supercontig 6a, Linkage Group V (centre)

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>87</td>
<td>12,990</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>183,820</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>282,268</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>3363</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

*See Table 5.1.
Note: Ends in rRNA repeat fragment: indicative of nucleolus organizer.

TABLE 5.7
Supercontig 6b, Linkage Group V-R

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>226(−)</td>
<td>5454</td>
<td>—</td>
<td>1</td>
<td>T15</td>
</tr>
<tr>
<td>92</td>
<td>62,660</td>
<td>4191,4190</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>210(−)</td>
<td>6735</td>
<td>619</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>340,023</td>
<td>1500</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>523,724</td>
<td>4546,4545</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>103,820</td>
<td>1223</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>98,894</td>
<td>1303</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>199(+)</td>
<td>7686</td>
<td>516</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>24,746</td>
<td>620</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>195(+)</td>
<td>8077</td>
<td>825</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>565,485</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

*See Table 5.1.
Note: Contig 92 shows linkage to TelContig complex TC11. It is here shown associated with contig 226, the remaining copy of TC11 for which no specific supercontig linkage has been identified.
### TABLE 5.8
Supercontig 7, Linkage Group I-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>2005</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>337.251</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>209,568</td>
<td>739</td>
<td>1e, 1s</td>
<td></td>
</tr>
<tr>
<td>194(+)</td>
<td>8226</td>
<td>644</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>132,643</td>
<td>1881</td>
<td>2e</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>69,272</td>
<td>1605</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>297,769</td>
<td>1228</td>
<td>3e</td>
<td></td>
</tr>
<tr>
<td>207(+)</td>
<td>6880</td>
<td>754</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>277,183</td>
<td>329</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>204(+)</td>
<td>6994</td>
<td>680</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>67,855</td>
<td>298, 299</td>
<td>1, 4e</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>444,822</td>
<td>15</td>
<td>2e</td>
<td></td>
</tr>
<tr>
<td>208(−)</td>
<td>6838</td>
<td>734, 735</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>352,786</td>
<td>−</td>
<td>−</td>
<td>→T7</td>
</tr>
</tbody>
</table>

*a–d See Table 5.1.

### TABLE 5.9
Supercontig 8, Linkage Group I-R

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>233,950</td>
<td>623</td>
<td>1s</td>
<td></td>
</tr>
<tr>
<td>234(−)</td>
<td>4876</td>
<td>1210</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>298,245</td>
<td>1371</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>62,418</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>353,724</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>316,066</td>
<td>652, 653</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>237(−)</td>
<td>4650</td>
<td>542</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>25,936</td>
<td>1657</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>187(−)</td>
<td>10,087</td>
<td>2576</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>127,320</td>
<td>829, 830</td>
<td>2e</td>
<td></td>
</tr>
<tr>
<td>181(+)</td>
<td>10,660</td>
<td>2263, 2261</td>
<td>38s</td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>56,802</td>
<td>−</td>
<td>−</td>
<td>→T14</td>
</tr>
</tbody>
</table>

*a–d See Table 5.1.

### TABLE 5.10
Supercontig 9, Linkage Group IV-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>117</td>
<td>274,575</td>
<td>−</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>211(+)</td>
<td>6721</td>
<td>711</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>105,880</td>
<td>−</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>137,987</td>
<td>−</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>3924</td>
<td>−</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>13,722</td>
<td>−</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>214,949</td>
<td>649</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>223(+)</td>
<td>5626</td>
<td>815</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>123</td>
<td>128,436</td>
<td>1631</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>205(−)</td>
<td>6930</td>
<td>530</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>55,685</td>
<td>1311</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>32,723</td>
<td>480</td>
<td>6, 30s</td>
<td></td>
</tr>
<tr>
<td>126</td>
<td>40,772</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>127</td>
<td>138,927</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>280,307</td>
<td>−</td>
<td>−</td>
<td>→T2</td>
</tr>
</tbody>
</table>

*a–d See Table 5.1.

Note: (1) I-1 retrotransposon possibly bridges 1050 bp gap; (2) similar I-1 elements at contig 119 end and 120 start; (3) AMA-1 sequence + MATE-1b (DNA-3_AN) transposon [22] bridge 3398 nt gap.
### TABLE 5.11
Supercontig 10, Linkage Group IV-R

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>129</td>
<td>627,835</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>308,583</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>116,788</td>
<td>1713</td>
<td>1s</td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>254,738</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>133</td>
<td>81,213</td>
<td></td>
<td>1</td>
<td>→T13</td>
</tr>
</tbody>
</table>

*See Table 5.1.*

*Note:* Gypsy-1 element with matching target-site duplications forms probable bridge of 2943 bp gap between contigs 132–133.

### TABLE 5.12
Supercontig 11, Linkage Group II-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>134</td>
<td>78,670</td>
<td></td>
<td></td>
<td>T6</td>
</tr>
<tr>
<td>135</td>
<td>301,041</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>136</td>
<td>31,127</td>
<td>1218</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>232(+)</td>
<td>5001</td>
<td>610</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>32,282</td>
<td>654</td>
<td>6s</td>
<td></td>
</tr>
<tr>
<td>193(+)</td>
<td>8744</td>
<td>895, 896</td>
<td>7e</td>
<td></td>
</tr>
<tr>
<td>138</td>
<td>39,791</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>139</td>
<td>347,029</td>
<td>760, 757</td>
<td>3e</td>
<td></td>
</tr>
<tr>
<td>215(+)</td>
<td>6468</td>
<td>613</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>26,834</td>
<td>618</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>203(+)</td>
<td>7055</td>
<td>1060</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>141</td>
<td>203,328</td>
<td>3978, 3977</td>
<td>27e, 1s</td>
<td></td>
</tr>
<tr>
<td>190(+)</td>
<td>9210</td>
<td>170</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>142</td>
<td>43,730</td>
<td>(−)</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>143</td>
<td>30,509</td>
<td>507, 511</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>144</td>
<td>72,448</td>
<td>670</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>229(−)</td>
<td>5381</td>
<td>855</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>176,570</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*See Table 5.1.*

*Note:* (1) Contig 138 end and contig 139 start are both simple sequences; (2) I-1 element putatively bridges 3262 bp gap.

### TABLE 5.13
Supercontig 12, Linkage Group V-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>146</td>
<td>11,447</td>
<td></td>
<td></td>
<td>→T8</td>
</tr>
<tr>
<td>147</td>
<td>5126</td>
<td>(38)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>148</td>
<td>4102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>149</td>
<td>38,673</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>109,819</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>151</td>
<td>72,894</td>
<td>939, 940</td>
<td>3e</td>
<td></td>
</tr>
<tr>
<td>212(+)</td>
<td>6710</td>
<td>1368, 1369</td>
<td>1e</td>
<td></td>
</tr>
<tr>
<td>152</td>
<td>53,749</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>153</td>
<td>403,022</td>
<td>789</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>196(−)</td>
<td>8076</td>
<td>2654, 2657</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>154</td>
<td>34,335</td>
<td>5052, 5053</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>155</td>
<td>40,523</td>
<td>(−)</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>156</td>
<td>11,675</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>157</td>
<td>136,262</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*See Table 5.1.*

*Note:* (1) Doubtful overlap of Gypsy-1 fragments with unmatched target-site duplications; (2) azgA clone (accession no. AJ575188) overlaps contig 155 by 385 bp and contig 156 by 2628 bp, with no mismatches.
### TABLE 5.14
Supercontig 13, Linkage Group III-R

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>158</td>
<td>283,337</td>
<td>605</td>
<td>0</td>
<td>→T16</td>
</tr>
<tr>
<td>209(−)</td>
<td>6773</td>
<td>661</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>159</td>
<td>58,286</td>
<td>1821</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>146,980</td>
<td>864,863</td>
<td>7s</td>
<td></td>
</tr>
<tr>
<td>214(−)</td>
<td>6579</td>
<td>621,620</td>
<td>1s</td>
<td></td>
</tr>
<tr>
<td>161</td>
<td>266,267</td>
<td>918</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>198(+)</td>
<td>7688</td>
<td>772</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>162</td>
<td>65,002</td>
<td>—</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>163</td>
<td>91,924</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*See Table 5.1.

Note: Contigs joined by replacement of 36 incorrect bases with a 34-base joining sequence [21].

### TABLE 5.15
Supercontig 14, Linkage Group VII-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>164</td>
<td>150,254</td>
<td>633</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>176(+)</td>
<td>15,731</td>
<td>537</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>114,598</td>
<td>2911,2907</td>
<td>1s, 5</td>
<td></td>
</tr>
<tr>
<td>178(+)</td>
<td>12,319</td>
<td>663</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>166</td>
<td>41,526</td>
<td>724,727</td>
<td>5s</td>
<td></td>
</tr>
<tr>
<td>167</td>
<td>41,972</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>229,305</td>
<td>1586,1585</td>
<td>4s</td>
<td></td>
</tr>
<tr>
<td>219(+)</td>
<td>6367</td>
<td>504</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>200(+)</td>
<td>7501</td>
<td></td>
<td></td>
<td>T5</td>
</tr>
</tbody>
</table>

*See Table 5.1.

### TABLE 5.16
Supercontig 15, Linkage Group VI-L

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>169</td>
<td>351,151</td>
<td>473</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>206(+)</td>
<td>6924</td>
<td>860</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>170</td>
<td>216,503</td>
<td></td>
<td></td>
<td>T3</td>
</tr>
</tbody>
</table>

*See Table 5.1.
Aspergillus nidulans *Linkage Map and Genome Sequence* 71

**TABLE 5.17**

Supercontig 16, (Linkage Group VIII-L ?)

<table>
<thead>
<tr>
<th>Contig</th>
<th>Length (bp)</th>
<th>Overlap</th>
<th>Mismatches</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>216 (−)</td>
<td>6436</td>
<td>3848, 3852</td>
<td>5, 15e</td>
<td>T1</td>
</tr>
<tr>
<td>91</td>
<td>6778</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>171</td>
<td>24,092</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>172</td>
<td>523,550</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>173</td>
<td>11,633</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

See Table 5.1.

**TABLE 5.18**

TelContigs Identified by TERMINUS

<table>
<thead>
<tr>
<th>TelContig</th>
<th>Length (bp)</th>
<th>No. of Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>973</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>1021</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>1029</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>1012</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>1012</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>1007</td>
<td>48</td>
</tr>
<tr>
<td>7</td>
<td>1055</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>999</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>979</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>1022</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>1023</td>
<td>151</td>
</tr>
</tbody>
</table>

**TABLE 5.19**

Links Between Telomeric Contigs and the *A. nidulans* Genome Assembly

<table>
<thead>
<tr>
<th>Chromosome Arm</th>
<th>Telomere</th>
<th>TelContig</th>
<th>Gap Length (bp)</th>
<th>Linked Contigs</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL</td>
<td>7</td>
<td>TC7</td>
<td>200</td>
<td>108, 208, 107...</td>
</tr>
<tr>
<td>IR</td>
<td>14</td>
<td>TC11</td>
<td>8800</td>
<td>116, 181, 115...</td>
</tr>
<tr>
<td>IIL</td>
<td>6</td>
<td>TC6</td>
<td>—</td>
<td>134, 222, 135...</td>
</tr>
<tr>
<td>IIR</td>
<td>4</td>
<td>TC4</td>
<td>800</td>
<td>59, 60...</td>
</tr>
<tr>
<td>IIIL</td>
<td>11</td>
<td>TC11</td>
<td>18,400</td>
<td>86, 85...</td>
</tr>
<tr>
<td>IIIR</td>
<td>16</td>
<td>TC10</td>
<td>700</td>
<td>158, 209, 159...</td>
</tr>
<tr>
<td>IVL</td>
<td>2</td>
<td>TC2</td>
<td>600</td>
<td>128,127...</td>
</tr>
<tr>
<td>IVR</td>
<td>13</td>
<td>TC11</td>
<td>6300</td>
<td>133, 132...</td>
</tr>
<tr>
<td>VL</td>
<td>8</td>
<td>TC8</td>
<td>1800</td>
<td>146, 147...</td>
</tr>
<tr>
<td>VR</td>
<td>15</td>
<td>TC11</td>
<td>1700</td>
<td>226, 92, 210, 93...</td>
</tr>
<tr>
<td>VII</td>
<td>3</td>
<td>TC3</td>
<td>—</td>
<td>170, 206, 169...</td>
</tr>
<tr>
<td>VIR</td>
<td>10</td>
<td>TC11</td>
<td>3800</td>
<td>44, 45...</td>
</tr>
<tr>
<td>VIIL</td>
<td>5</td>
<td>TC5</td>
<td>—</td>
<td>200, 219, 168, 167...</td>
</tr>
<tr>
<td>VIIR</td>
<td>12</td>
<td>TC11</td>
<td>8700</td>
<td>43, 42...</td>
</tr>
<tr>
<td>VIIIL</td>
<td>1</td>
<td>TC1</td>
<td>700</td>
<td>216, 91, 171, 172...</td>
</tr>
<tr>
<td>VIIIR</td>
<td>9</td>
<td>TC9</td>
<td>—</td>
<td>1, 2...</td>
</tr>
</tbody>
</table>

* Telomere numbers assigned by TERMINUS (http://fungus.kbrin.uky.edu/cgi-bin/gbrowse/A_nidulans_broad_tel_100k).
* Gaps were estimated based on average clone insert size.
* The contig listed first is nearest to the telomere. Adjacent contigs are listed to indicate supercontig orientation.
9080_C005.indd 72

226

44, 226
44, 226
44, 226
44, 226
226
226
TC11
TC11
116, 86
116
133
216
86

B

C
D
E
F
G
H
I
J
K
L
M
N
O

12,141, 2321
12,563, 2741
12,916, 3094
13,061, 3239
3541
4061
678
879
54,491, 111,497
54,874
76,316
5597
111,944

289

10,123

Position(s)

17
39
20
53
15
10
20
18
14
12
11
15
21

98

223

Consensus Sequencea

TGCTGGCGCTCCTCTGGCATTAAACCATAATCCTGGTGCATGACTGGTGCTTCCCCGGCGCGGCCCCGGGA
ATGCTCCTGGCCTGCTTGTCCCGGGCGCATAGCCAGCGCACTGCTGGTATTGTGCCTTGGCTGGCTCCCCCAC
AGGGCTGCCCGGGTGCGACGCGGGCTGTGCACCTATGCAGGCTCTGGGCTGGCGCCACCCAGGCATTGGC
CAGGAGTTC
GGAATGCTCCTGGCCTGCTTGTCCGGGCTGCATAGCCAGCGCACTGCTGCTGTTGTGCCTTTGTCGGCTCCCCC
1.9
CCAGTGCTGCCCTGGCGCGGCCCA
3.5
GCTGGCTTGGGGCTTG
2.9
TCTATATCTGGCCTATCTGCCTATATCTGGCTATATCTA
CTGCCTCTGTACTGGCTGCC
2.3
CCTAACCCTGACTATGACCTTGCCTGCCTATTGCCTGCC AACTACTGCATGAC
1.9
8
ATGATATTACCTGGC
6.7
GGCTCTGCTT
2.9
TTTTCTTATTATTATTATCC
4
CCCTTATTCTTACCCTTA
2.4, 2.1 TCGGGGTCTCAGAC
2.3
GTCGCTGTCCGC
2.5
AGCATAATACC
2
TATTATTATCATTGT
2.1
TTGTTGTTGTTGTGGGAGTTG

2.3

Unit Length Repeats

Sequence shared between motif A and motif B are italicized. A telomere motif is underlined.

44

A

a

Contig(s)

Motif

Sequences of Short Tandem Repeat Motifs Found in the A. nidulans Subtelomeres

TABLE 5.20

72
The Aspergilli

10/22/2007 5:16:01 PM


Aspergillus nidulans Linkage Map and Genome Sequence

References

6

Genome Sequence of Aspergillus oryzae

Masayuki Machida

CONTENTS
6.1 Introduction ................................................................. 75
6.2 Overview of the Aspergillus oryzae Genome ......................... 76
6.3 Gene Prediction ............................................................... 78
6.4 Expansion of Genes on the Aspergillus oryzae Genome .......... 78
6.5 Notable Characteristics of the Aspergillus oryzae Genome .......... 81
6.6 Aiming for Application of the Genome Sequence to Industries .... 82
References ........................................................................... 82

6.1 Introduction
Aspergillus oryzae has been widely used in Japanese fermentation industries, sake (Japanese alcohol beverage), miso (soy bean paste), shoyu (soy sauce), and su (vinegar) for longer than a thousand of years. A. oryzae produces large amounts of various hydrolytic enzymes including amylases, proteinases, lipases, and nucleases. High potential of secretory production of proteins and enzymes has led A. oryzae to be applied to modern biotechnology. The extensive use of A. oryzae in the food fermentation industries over the years prompted the description of industrial application of A. oryzae on the list of Generally Recognized as Safe (GRAS) of Food and Drug Administration (FDA) in the United States of America. The safety of this organism is also supported by World Health Organization.2

The major focus of the basic research of A. oryzae has been the analysis of hydrolytic enzymes, which are important players in fermentation. This includes purification, characterization, and amino acid sequence analysis of various enzymes of interest. Cloning of the corresponding genes facilitated the analysis. However, existence of multiple conidia in a single cell makes it difficult to obtain mutants and lack of sexual generation in life cycle of A. oryzae prevents A. oryzae from being analyzed by traditional genetics.

Introduction of heterologous genes is a key technique to study A. oryzae both for basic research and for industrial applications. Development of transformation system of A. oryzae has made it possible to produce number of enzymes from filamentous fungi in the industrially applicable amounts. (Mucor renin under the control of the α-amylase promoter reached approximately 3.3 g from 1 L of medium, for example.4) Some proteins derived from higher eukaryotes, human lactoferrin,5 human lysozyme,6 calf chymosin,7,8 or plant thaumatin9 were produced from only 50 μg to 25 mg from 1 L culture. Fusion of fungal carrier proteins such as the carboxyl terminus of the A. niger glucoamylase or A. awamori α-amylase and the introduction of KEX2-like processing signal between the carrier and the protein to be secreted improved the secretory production of some eukaryotic proteins, calf chymosin,10,11 calf phospholipase A2,12 and human interleukin 6.13

Remarkable feature of the Japanese traditional fermentation industries is the use of solid-state cultivation (koji). A. oryzae is grown on the surface of steamed rice grain for the sake production. Starch
in the steamed rice grain is degraded by amylases produced by *A. oryzae* and successively converted into alcohol by yeast. For soy sauce brewing, *A. oryzae* is grown on the surface of grounded soy beans to degrade proteins to peptides and amino acids. It is known that the secretory production of enzymes by *A. oryzae* is significantly enhanced by the solid-state cultivation. Ishida et al. found that *glaB*, one of the genes encoding the solid-state culture-specific enzymes was strongly induced when grown on a nitrocellulose membrane on a Czapek-Dox medium plate with low water activity in the presence of 50% maltose. These findings suggest that the solid-state specific induction is mainly due to low water activity and physical barrier. The high productivity in the solid-state cultivation seems to induce the potential for secretion of extracellular proteins.

The secretion mechanism of *A. oryzae* has been also extensively studied to enhance the productivity of fermentation and the amounts of enzymes. The secretory pathway includes various steps of protein sorting between endoplasmic reticulum (ER), golgi, plasma membrane, endosome, and vacuole. Studies of hyphal tip, where the protein secretion is most prominent, are also important in relation to the enhancement of enzyme secretion in the solid-state cultivation.

Large-scale EST sequencing was completed in 2001 by the collaboration of National Research Institute of Brewing (NRIB) (Higashi-Hiroshima, Japan), National Institute of Advanced Industrial Science and Technology (AIST) (Tsukuba, Japan), National Food Research Institute (NFRI) (Tsukuba, Japan), Nagoya University (Nagoya, Japan), Food Research Institute of Aichi Prefectural Government (Nagoya, Japan), The University of Tokyo (Tokyo, Japan), Tokyo University of Agricultural Technology (Tokyo, Japan), and Tohoku University (Sendai, Japan). *A. oryzae* strain RIB40 (ATCC-42149) was selected for the sequencing. The strains used in the soy sauce companies have their own strains that have been obtained by extensive breeding. The *sake* brewing companies purchase the strains suitable for the brewing from the companies that breed the strains. *A. oryzae* RIB40 is a wild-type strain, most similar to those used for *sake* brewing but still has ability of strong production of proteinases, which is important characteristic for soy sauce fermentation. The mRNAs were prepared from *A. oryzae* mycelia grown in several different culture conditions, rich medium, heat shock, the medium without any carbon source and so on. The library included the mRNA from the mycelia grown in the solid-state culture. Number of genes including those encoding the solid-state culture-specific enzymes already found were detected from the library. The 5'-terminus of each cDNA was specifically analyzed by constructing the libraries by directional cloning technique so that the protein coding region might be effectively analyzed. Total number of ESTs and the total length analyzed reached 21,550 and 12.24 Mb, respectively, by the analysis. After clustering, the total number of the nonredundant sequences (unigenes) was approximately 7700. The BLAST search showed significant similarities to the deposited sequences of known function in the public database for 38.6% of the EST contigs. The contigs of the *A. oryzae* ESTs are made searchable at the websites (http://www.nrib.go.jp/ken/EST/db/index.html, http://www.aist.go.jp/RIODB/fldb/index.html). The EST results showed that the number of highly expressed ESTs (highly expressed genes in the other word) are less than only 500, which is less than 5% of the total *A. oryzae* genes.

The genome size of *A. oryzae* had been estimated to be 35 Mbp consisting of eight chromosomes ranging from 2.8 Mbp to 7 Mbp in length from the pulse field gel electrophoresis before sequencing the entire genome of *A. oryzae*. No genetic and physical maps were available at that time. The shortest band at 2.8 Mbp has approximately two times stronger intensity than expected, indicating that the band derives from two chromosomes, VII and VIII. The second shortest band (VI) is obviously weaker than the others and is smeared. The band was found to hybridize with the DNA fragment having ribosomal DNA (rDNA) sequence from *A. oryzae*. These results strongly suggest that chromosome VI possesses rDNA that repeats in variable numbers. The positional information available at that time was the mapping of approximately 15 genes on chromosomes. The GC content of the genome was estimated to be approximately 46% from the random sequencing of small number of shotgun library clones.

### 6.2 Overview of the *Aspergillus oryzae* Genome

*Aspergillus oryzae* RIB40 (National Research Institute Culture Stock; ATCC42149), which was used for EST sequencing, was selected for genome sequencing. Although neither genetic nor physical map was
available when the genome sequencing was launched, the genes and DNA sequences of RIB40 had been most well characterized among those from the Aspergillus oryzae strains.18,20

The genome sequencing of A. oryzae was accomplished by the whole genome shotgun (WGS) approach by the collaboration of National Institute of Technology and Evaluation (NITE) and The Consortium for A. oryzae Genomics*. High-quality sequence reads from both ends of over 500,000 WGS clones generated by arrayed capillary DNA sequencers reached approximately 7× depth of coverage of the A. oryzae genome. Cosmid and BAC clones of approximately 5000 for each were used to connect the contigs obtained by WGS. Consequently, 36.9 Mb of the A. oryzae genome was sequenced with approximately 9× depth of coverage and greater than 99% of the sequence assembly was supported by two or more than two independent BAC or cosmid clones. Chromosome assignment and relative position of the contigs were analyzed by Southern hybridization against the chromosomes separated by PFGE (pulse field gel electrophoresis) and by the fingerprinting method, respectively.22 The final assembly consisted of 6 scaffolds and 10 contigs (or 24 contigs in total). There were 11 physical and 5 sequence gaps remaining in the assembly. The total contig length of the assembly was 37,047,050 bases.

The A. oryzae genome consists of eight chromosomes numbered from 1 to 8 in decreasing size. The length of each chromosome is 6.3 Mb, 6.2 Mb, 5.0 Mb, 4.8 Mb, 4.4 Mb, 4.1 Mb, 3.4 Mb (including 0.7 Mb rDNA repeats) and 3.3 Mb, resulting in total genome size of 37.6 Mb (Fig. 6.1).22 The mitochondrial DNA is 28.9 kb in size. In accordance with the existence of 8 chromosomes, 16 DNA fragments containing the telomeric repeats (TTAGGGTCAACA), which are 6 nt longer than that of A. nidulans, were identified. The chromosomes were renumbered according to the sequencing results (Fig. 16.1). The longest DNA fragment on PFGE, which was previously assigned to chromosome I, was found to consist

![FIGURE 6.1 Structure of A. oryzae genome. Note: Ovals and black boxes represent presumed centromeres and telomeres, respectively. rDNA designates ribosomal DNA repeats. Dark gray, light gray, and white boxes represent a sequence gap, a physical gap linked by Southern hybridization and an unlinked physical gap, respectively.](image)

* The Consortium for A. oryzae Genomics consisted of National Institute of Advanced Institute of Industrial Science and Technology (AIST), National Research Institute of Brewing (NRIB), National Food Research Institute (NFRl), Tohoku University, The University of Tokyo, Tokyo University of Agricultural Technology, Nagoya University, Amano Enzyme, Gekkeikan Sake, Higeta, Intec Web and Genome Informatics (Tokyo, Japan), Kikkoman, Kyowa-Hakko Kogyo (Tokyo, Japan), Ozeki and Brewing Society of Japan (Tokyo, Japan) as a representative.
of chromosomes 1 and 2. The shortest fragment previously shown to contain two chromosomes (chromosomes VII and VIII) consisted of a single chromosome (chromosome 8). Centromeric sequences have not yet been analyzed because cloning or sequencing of the centromeric fragments was unsuccessful in spite of extensive efforts by trying several different vectors and PCR amplification protocols, probably due to extremely high AT-content and/or DNA curvature. The A. oryzae genome size is very close to those of A. niger and A. flavus, and 20–30% larger than those of A. nidulans and A. fumigatus.27

The sequence assembly was further validated by the optical mapping. The restriction map by the optical mapping with AflII and that predicted from the sequence assembly perfectly matched. The total length of the physical gaps was estimated to be approximately 0.8 Mb by the optical mapping. Therefore, the genome size of A. oryzae is estimated to be 38.4 Mb consisting of 37.0 Mb, 0.6 Mb, and 0.8 Mb, for sequenced, ribosomal, and physical gap lengths, respectively.

The A. oryzae genome contained numerous stretches of AT rich sequence as compared to the other two aspergilli: 1759, 197, and 308 AT-stretches with >90% ATs longer than 50 b were found in A. oryzae, A. fumigatus, and A. nidulans genomes, respectively. The sequence center, NITE, made huge efforts to sequence the AT-rich segments applying various techniques of PCR for the preparation of corresponding DNA fragments and for the sequencing.

6.3 Gene Prediction

Combination of several gene-calling programs was employed for the prediction of genes in the A. oryzae genome. The homologs of the proteins of aspergilli, Neurospora crassa, Magnaparth grisea, Gibberella zeae, Penicillium, and Paecilomyces were searched by BLASTX. The resulting candidates of homologs were evaluated by ALN by aligning the candidates and the protein sequences. ALN takes into account frameshift errors, coding potentials and signals for translational initiation, termination, and splicing. Consequently, the 489 highly reliable genes were adopted as a learning set for GeneDecoder and Glimmer software, the ab initio gene finders. GeneDecoder also integrates the information for splice sites provided by the A. oryzae and A. flavus ESTs, which are aligned to the genome sequence by SIM. The gene models were manually corrected by referring to the alignment of the predicted genes and the genes of strong similarity in the public database and by comparing orthologs from A. fumigatus and A. nidulans.26 The total number of the A. oryzae genes, encoding proteins longer than 100 amino acids, has reached 12,074 from the sequenced genome of 36.7 Mb. Thus, an average gene density is 3.0 kb/gene, which is approximately 1.5 times longer than that of Saccharomyces cerevisiae (2 kb/gene). All the predicted protein sequences were annotated by searching against COG database using BLASTP, followed by a manual correction. Identification of transfer RNAs was based on tRNAscan-SE, and repeated sequences were detected using RepeatMasker.

6.4 Expansion of Genes on the Aspergillus oryzae Genome

Although A. oryzae and S. cerevisiae are closely related to each other in a phylogenetic tree, A. oryzae has about 2.5 times bigger genome size and approximately twice more genes than S. cerevisiae. The A. oryzae genome has even 25–30% bigger genome size and the number of the predicted genes as compared to the A. fumigatus and the A. nidulans genomes.22 Figure 6.2 shows comparison of the number of genes in each COG functional category. All the microorganisms analyzed in Figure 6.2 show very close numbers of genes belonging to the COG functional categories for “Information Storage and Processing.” The genes for most of the COG functional categories, “Cellular Processes and Signaling,” also show similar numbers except for Y (nuclear structure), V (defense mechanisms), T (signal transduction mechanisms), and M (cell wall/membrane/envelope biogenesis) functional categories. On the other hand, the genes for all the COG functional categories belonging to “Metabolism” (C, G, E, F, H, I, P, and Q; energy production and conversion, carbohydrate transport and metabolism, amino acid transport and metabolism, nucleotide transport and metabolism, coenzyme transport and metabolism, lipid transport and
metabolism, inorganic ion transport and metabolism, secondary metabolites biosynthesis, transport and catabolism, respectively) are significantly expanded, among which expansion of the Q genes are most prominent. Expansion of the functionally unknown genes (R, S, and X; general function prediction only, function unknown, and no homology to any COGs, respectively) is also significant to the same extent. Considering the distribution of the absolute number of genes in each COG category, the expansion of metabolic genes most significantly contributes to the expansion of the total gene number in A. oryzae.

These observations are consistent with the previous suggestion based on the comparison between eubacterial and archaeal genomes that metabolic flexibility may depend on the genome size. By the analysis of number of the orthologs conserved between A. oryzae and A. fumigatus or A. nidulans, the metabolic gene expansion in A. oryzae appeared to occur in several specific metabolic pathways rather than randomly distributed. The most highly expanded metabolic genes are those for phenylalanine/tryptophan degradation, toluene/m-cresol/p-cymene degradation, and biosynthesis/degradation of hydrophobic amino acids. Pyruvate decarboxylase and alcohol dehydrogenase genes involved in phenylalanine/tryptophan degradation are known as the most highly expressed genes in the presence of glucose. Although genes in all the COG categories related to metabolism are expanded in A. oryzae relative to A. nidulans and A. fumigatus, the most highly expanded genes are those involved in phenylalanine/tryptophan degradation (ARO10, PDC6, PDC5, PDC1, SFA1, ADH5, ADH4, ADH3, ADH2, ADH1) and those participating in toluene/m-cresol/p-cymene degradation (AAD15, AAD14, AAD10, AAD6, AAD4, and AAD3). The other remarkable feature concerning the primary metabolism is the significant expansion of the ATP-binding cassette (ABC), the amino acid-polyamine-organocation (APC), and the major facilitator superfamily (MFS) transporter genes. In fermentation, A. oryzae is grown on the surface of steamed rice or steamed ground soybean, where large amounts of amino acids and sugars are supplied but are deficient at the beginning. These specific gene family expansions of metabolism and transporters are thus consistent with the need for A. oryzae to more effectively access external nitrogen resources and to degrade protein and starch.

Of the three aspergilli, A. oryzae had the largest number of aspartic proteinase, metalloproteinase, and serine-type carboxypeptidase genes. Roughly half of the metalloproteinases and most of the aspartic
proteinases and the serine-type carboxypeptidases had signal sequences. The phylogenetic analysis of the putative secretory proteinases from the three aspergilli showed extra copies of genes existing specifically in the *A. oryzae* genome. Interestingly, the extra copy of the *A. oryzae* homolog is phylogenetically more distant to the *A. oryzae* gene in each orthologous cluster of the three genes from the three species than the orthologs from the other two species. It is unlikely that the extra copies of the genes derive from gene duplication as described later; so, they are described as “extra homologs.” In contrast to the secreted proteinases, the number of protease genes encoding intracellular enzymes is consistent among the three aspergilli. Three α-amylase genes (*amyA*, *amyB*, and *amyC*), which have almost identical nucleotide sequences each other, have been found. Together with the association of the three amylase genes with transposon-like or incomplete transposon sequence, strongly suggests gene duplication mechanism for the generation of the three genes. In contrast, the maltase genes showed the expansion manner similar to that of proteinases.

As described earlier, the largest gene family expansion in *A. oryzae* relative to *A. nidulans* and *A. fumigatus* occurs in genes predicted to play a role in secondary metabolism. Comparison of the three *Aspergillus* genomes has revealed a large increase in number of cytochrome P450s (149, 102, and 65 for *A. oryzae*, *A. nidulans*, and *A. fumigatus*, respectively) and a slight increase in number of polyketide synthase genes (30, 27, and 14 for *A. oryzae*, *A. nidulans*, and *A. fumigatus*, respectively). Resistance of aspergilli to various chemicals might be due to the remarkable redundancy and pleiotropy of the cytochrome P450s and the transporters described earlier. Several genes homologous to those characteristic to plant pathogenic fungi (e.g., *Nectria haematococca*, *Fusarium* spp.), trichothecene hydroxylases, isorotichodermia hydroxylases, trichodiene oxygenases, and pisatin demethylases, were found in the *A. oryzae* genome. A moderate expansion of NRPS was also observed in the *A. oryzae* genome as compared to the other two *Aspergillus* genomes.

Although the predicted genes homologous to aflatoxin production are present in *A. oryzae*, no ESTs to these genes were detected except for *aflJ* and *nora*. In contrast, ESTs for all the 25 genes were found in *A. flavus*. These results suggest that the long history of industrial use of *A. oryzae* has selected strains favorable for human consumption or that *A. oryzae* may have been selected as a safe mutant from the beginning. The silencing of the genes may reflect the mechanism similar to that observed in the regulation of aflatoxin biosynthesis in *Aspergillus sojae* and/or the mutation of a global regulator of secondary metabolism genes such as *laeA*.

*A. oryzae* has almost never been recognized as a human pathogen in spite of its extensive, high inoculum use (in traditional-style sake breweries, the craftsmen sprinkle *A. oryzae* conidia onto steamed rice through a sieve and cultivate them in a small wooden room). The clear contrast in pathogenicity between *A. oryzae* and the other two aspergilli, highly redundant secondary metabolism genes, and the completion of sequencing the three genomes should have made aspergilli more attractive for medical and therapeutic research.

Although genes concerning basic cellular functions including “Information Storage and Processing” (J-B) and “Cellular Processes and Signaling” (D-O) except nuclear structure (Y), defense mechanism (V), cell wall/membrane/envelope biogenesis (M), and extracellular structure (W) genes appeared consistent among the three aspergilli, *A. oryzae* has an increased number of some kinases playing important roles on the regulation of cellular function. The aspergilli possessed more sensor histidine kinases (13–15) than *S. cerevisiae* (1) and *Schizosaccharomyces pombe* (3), while histidine-containing phosphotransfer factors and response regulators were found in similar numbers as in the yeasts. Of the nine histidine kinase families (HK1–9) *Aspergillus* possesses, HK8 is specific to *Aspergillus* and absent in *N. crassa* and the plant pathogens sequenced to date. Further, *A. oryzae* had two additional homologs of HK6 family as compared to the other two *Aspergillus* species, *N. crassa* and the plant pathogens, which possess a single HK6 gene. Since Nik-1, the HK6 homolog in *N. crassa*, is essential for growth in a high osmotic pressure environment, the additional homologs may play an important role to adapt *A. oryzae* to high-osmolarity conditions in making koji. Another example is mitogen-activated protein kinases (MAPK) and related kinases. While there are three MAPKK and MAPKKK in all three *Aspergillus* and *N. crassa* genomes, there are four MAPK in *A. nidulans* and *A. fumigatus* and five in *A. oryzae* but only three in *N. crassa*. This suggests that *A. oryzae* may have the most complex signal transduction cascade among the four filamentous fungi.
6.5 Notable Characteristics of the Aspergillus oryzae Genome

Genomes of the three aspergilli, A. oryzae, A. fumigatus, and A. nidulans share overall conserved synteny with each other.\(^{22,26}\) To find synteny, a particular A. oryzae gene was randomly picked and the ortholog on the genome to be compared was assigned. If the next ortholog of the adjacent gene on the A. oryzae genome could be found within 10 kb on the genome to be compared, the two genes are defined syntenic. From the analysis, the A. oryzae genome has significantly more synteny breaks than between A. fumigatus and A. nidulans. As a result, A. oryzae genome has a mosaic structure consisting of the syntenic and the nonsyntenic blocks, which are almost equivalent to the blocks common to all the three aspergilli and those specific to A. oryzae, respectively. Interestingly, the secondary metabolism genes are highly enriched on the nonsyntenic blocks.\(^{22}\) It is noted that the secondary metabolism genes are enriched with statistical significance in the regions lacking synteny with either A. fumigatus or A. nidulans (\(P = 9.8 \times 10^{-32}\)). Further, the EST analyses indicated that gene expression was considerably lower in these nonsyntenic regions (\(P = 4.1 \times 10^{-134}\)).\(^{22}\)

The A. oryzae-specific regions contained 1.7 times lower density of genes homologous to those in other eukaryotes except A. fumigatus and A. nidulans than did the common regions. Since no syntenic regions have been observed inside the A. oryzae genome, the A. oryzae specific DNA segments are supposed to emerge by horizontal transfer from foreign organisms rather than by gene duplication.\(^{22}\) Interestingly, the genes related to secondary metabolism and the genes with strong similarity to bacterial ones are highly enriched in the A. oryzae specific blocks. Mapping of ESTs\(^{22}\) and the preliminary DNA microarray experiments (unpublished data) have revealed that the expression of most of these genes are clearly weaker than the genes on the common regions under ordinal growth conditions including solid-state cultivation. Most of the extra homologs were located on the nonsyntenic blocks and their expression was not detected by ESTs.

There are two major possible mechanisms that could have made difference in genome size between A. oryzae the other two Aspergillus species (Fig. 6.3). In the first scenario, A. oryzae might have acquired extra genetic materials after A. oryzae branched off. Alternatively, in the second scenario, the ancestor might have bigger genome size as A. oryzae and A. fumigatus, and A. nidulans might have lost genetic materials to make their genome size smaller. The synteny analysis showed that the genome organization of A. fumigatus and A. nidulans is very close, indicating that the genetic materials that the two species lost are mostly common. On the other hand, the phylogenetic analysis of the three species based on the genome sequences by Galagan et al. demonstrated that A. nidulans branched off before the speciation of A. oryzae and A. fumigatus.\(^{26}\) This means that A. fumigatus and A. nidulans have independently lost the common genetic materials that reached approximately 30% of the entire genome. Since the possibility of such event is unlikely to happen, the first scenario seems more probable. No existence of detectable

![FIGURE 6.3](https://example.com/figure63.png)

**FIGURE 6.3** Possible mechanism for the expansion of A. oryzae genome. **Note:** Two possible scenarios for the expansion of A. oryzae genome are schematically drawn. The phylogenetic relationship of the three aspergilli is indicated according to the result of Galagan et al. The darker and lighter gray boxes represent the genetic blocks acquired by the A. oryzae genome and lost from the other two aspergilli, respectively. (From Galagan et al., Nature, 438, 2005.)
synteny inside the *A. oryzae* genome indicates low possibility of a large-scale gene duplication event after the speciation, and thus, supports the first scenario as well.

### 6.6 Aiming for Application of the Genome Sequence to Industries

Genome sequence of *A. oryzae* and extensive analyses based on the sequence have led to the results explaining why *A. oryzae* has been widely used in fermentation industries. Further, the sequence provides information to effectively analyze gene and cellular functions that are important for the industrial applications including those to improve productivities and to develop novel bioprocesses. Most effective approach for this purpose is to analyze transcriptional regulation, protein expression, metabolic regulation, and so on. The cDNA-based DNA microarray has been successfully applied to the development of novel application of *A. oryzae* to recycling biodegradable plastics. Recently, the DNA microarray consisting of 12,000 oligonucleotide probes has been developed by the collaborative research between Kanazawa Institute of Technology and National Institute of Advanced Industrial Science and Technology. The analysis of transcriptional regulation of the genes consisting of metabolic pathways is indeed one of the most important targets for the research and development of *A. oryzae*. Proteomic analysis is also underway in Kanazawa Institute of Technology and National Institute of Technology and Evaluation.

The compounds responsible for the flavor of fermented food such as sake (Japanese alcohol) and soy sauce are important targets for traditional fermentation industries. The analyses have been extensively performed mainly by GC/MS and LC/MS for many years. The analyses were mainly for the compounds in the product but this is now being extended to the intracellular compounds. Combination of various information including transcriptome, proteome, and metabolome will facilitate not only basic research but also research and development of fermentation technologies. The Japanese traditional fermentation was established more than 1000 years ago by accumulating knowledge obtained from daily life. Even after extensive research and development by modern biotechnology, skillful craftsmen play essential roles in the production of high quality of fermented products, *Ginjo-shu*, the premium alcohol, for example. However, recent decrease in number of the craftsmen can be a problem to maintain artistic technologies developed in a long history. Harmony of traditional fermentation technology and genomics could be a key to shed light on the unsolved mechanism for the artistic production.

Very recently, genome sequencing of *A. flavus* was commenced by the collaborative work of North Carolina State University, USDA/ARS South Regional Research Center and The Institute of Genomic Research. An extensive analysis of the close relatives, *A. oryzae* and *A. flavus*, will reveal the detailed genomic difference between these species and perhaps provide insight into the *A. oryzae* genome changes brought about by centuries of domestic cultivation. Extensive analyses of the function and regulation of *A. oryzae* genes will contribute a wealth of information important for the development of *A. oryzae* and other *Aspergillus*-based biotechnology applications.

### References


*Genome Sequence of Aspergillus oryzae* 83
II

Basic Biology of the Aspergilli
7

Signal Transduction in Aspergilli

Jae-Hyuk Yu and Christophe d’Enfert

CONTENTS
7.1 Introduction ............................................................................................................. 87
7.2 Primary Components of Heterotrimeric G Protein Signaling .................................................... 88
  7.2.1 G-Protein-Coupled Receptors (GPCRs) ................................................................. 88
  7.2.2 G Protein \(\alpha\) Subunits .............................................................................................. 89
    7.2.2.1 FadA and GpaA .............................................................................................. 90
    7.2.2.2 GanB and GpaB ............................................................................................. 91
  7.2.3 G Protein \(\beta\) Subunit ................................................................................................ 91
  7.2.4 G Protein \(\gamma\) Subunit ............................................................................................ 92
  7.2.5 Phosducin-Like Proteins (PhLPs) .................................................................................. 92
7.3 Downstream Signaling Branches ............................................................................................ 92
  7.3.1 Adenylate Cyclase and cAMP-Dependent Protein Kinases (PKAs) ............................. 92
  7.3.2 Mitogen Activated Protein (MAP) Kinases ................................................................ 94
  7.3.3 Protein Kinase C (PKC) ................................................................................................. 95
7.4 Negative Regulators of G Protein Signaling: RGSs .................................................................... 96
  7.4.1 FlbA and AfFlbA ........................................................................................................ 96
  7.4.2 RgsA ...................................................................................................................... 97
  7.4.3 Other RGSs ............................................................................................................. 97
7.5 Conclusions and Prospects ................................................................................................ .......... 98
Acknowledgments ................................................................................................................................ 99
References ............................................................................................................................................ 99

7.1 Introduction

Signal transduction is vital for the biology of all living cells, contributing to the integration of environmental cues into appropriate physiological and biochemical responses. The heterotrimeric G protein (G protein) system is conserved in all eukaryotes and is the most commonly used signal transducing system in eukaryotic cells. Basic elements of G protein signaling include a G protein-coupled receptor (GPCR), a G protein composed of \(\alpha\), \(\beta\), and \(\gamma\) subunits, and a variety of effector proteins [1–3]. In fungi, G protein signaling controls a diverse range of biological processes including growth, cell division, mating, cell–cell fusion, morphogenesis, chemotaxis, virulence, pathogenesis, and secondary metabolite production [4–9].

Sequential sensitization and activation of G protein elements translates external signals into gene expression changes, which leads to appropriate cellular behaviors. Binding of ligands to GPCRs (sensitization) induces the physical interactions between GPCR and inactive heterotrimeric GDP-G\(\alpha\); G\(\beta\)\(\gamma\), which cause GDP-GTP exchange of G\(\alpha\), resulting in the dissociation of GTP-G\(\alpha\) from the G\(\beta\)\(\gamma\) heterodimer (Fig. 7.1) [1–3]. Once separated, GTP-G\(\alpha\), G\(\beta\)\(\gamma\) or both can propagate signals through activities of (various) effector proteins. In general, G protein mediated signaling is transmitted via one or more
of the following pathways: (1) Mitogen-Activated Protein (MAP) kinases; (2) adenylyl cyclase and cAMP-dependent protein kinases (PKA); and (3) phospholipase C (PLC) and IP$_3$-[Ca$^{2+}$]-DAG (diacyl-glycerol)-dependent protein kinase C (PKC; see Fig. 7.1) [1–3]. The major portion of this chapter focuses on the characteristics and functions of the primary G protein components and downstream signaling branches in Aspergillus species, particularly in the model (Aspergillus nidulans) and pathogenic (Aspergillus fumigatus) aspergilli.

The later part of the chapter describes how G protein signaling is tightly controlled in aspergilli. Proper control of the specificity and duration of G protein signaling is necessary for the precise translation of signals into a relevant cellular response. The signal is turned off when GTP is hydrolyzed to GDP by the intrinsic GTPase activity of G$\alpha$. Thus, the rate of GTP hydrolysis of the G$\alpha$ subunit determines the strength of the signal [3,10]. Amongst several modulators, regulators of G protein signaling (RGS proteins) are major players that tightly control GPCR-G protein-mediated signaling [10]. We will discuss briefly the roles of RGS proteins in aspergilli.

### 7.2 Primary Components of Heterotrimeric G Protein Signaling

#### 7.2.1 G-Protein-Coupled Receptors (GPCRs)

The G-protein-coupled receptor (GPCR) family represents the largest and most varied collection of membrane-embedded proteins. A canonical GPCR contains a conserved structure of seven transmembrane (7-TM) spanning domains. This feature has led to the identification of greater than 16 putative GPCRs in the genomes of the 3 sequenced aspergilli (A. nidulans, A. fumigatus, and A. oryzae; see Fig. 7.2). These GPCRs have been assigned to nine classes on the basis of phylogenetic studies [11–14]. Classes I and II define GprA (PreB) and GprB (PreA), respectively, which are similar to the yeast pheromone receptors [15]. Class III includes GprC, GprD, and GprE (GprE is only found in A. nidulans) receptors that might be involved in carbon-source sensing based on their high similarity to the *Saccharomyces cerevisiae* Gpr1 protein [16,17]. Class IV is defined by GprF and GprG that are similar to the *Schizosaccharomyces pombe* Stm1 receptor involved in nitrogen sensing [18]. Class V includes GprH, GprI (GprI is not present in A. oryzae), and GprL (GprL is only present in A. fumigatus), which are similar to the *Dictyostelium*
Signal Transduction in Aspergilli

discoideum cAMP receptor cAR1 and as such have been proposed to play a role in cAMP sensing [12,14]. In addition, GprJ (class IV), GprK (class VI), GprM and GprN (class VII; GprN is only present in A. nidulans), GprO, GprP, and GprQ (class VIII), and NopA (class IX) are identified (Fig. 7.2) [14]. It is important to note that, PalH, which has not been included in the previous genome analyses, is another 7-TM protein that functions as a putative pH sensor in A. nidulans [19].

Functions of the GprA, GprB, and GprD GPCRs have been further studied in A. nidulans [11,15]. Deletion of gprD results in highly restricted hyphal growth, delayed conidial germination, and enhanced sexual development resulting in a petite colony covered by sexual fruiting bodies called cleistothecia [11]. Genetic or environmental changes obstructing sexual development rescue both growth and developmental abnormalities caused by deletion of gprD, leading to the hypothesis that the primary role of GprD is to negatively regulate sexual development, which might be needed for proper vegetative growth of A. nidulans [11]. A later study characterized the gprA and gprB genes encoding putative GPCRs similar to the yeast pheromone receptors Ste2p and Ste3p, respectively [15]. Deletion of gprA or gprB results in the formation of reduced number of cleistothecia, which are smaller than those of wild type and carry few viable ascospores (sexual spores). Supporting the potential roles of these GPCRs in sexual fruiting body development, the gprA gprB double deletion mutant is unable to produce any cleistothecia in homothallic (self-fertilizing) conditions. Perhaps somewhat unexpectedly, neither the gprA or gprB null mutation affects Hülle cell (specialized cell for supporting the development of cleistothecia) formation or cleistothecia development in outcrosses, leading to a conclusion that GprA and GprB are distinctively required for self-fertilization in homothallic conditions. Corroborating the idea that the primary role of GprD is to negatively control sexual development, and that GprA/B function downstream of GprD, diminished (or the absence of) sexual development caused by deletion of gprA and/or gprB suppresses growth defects caused by the absence of gprD [15]. A hypothesized model for GPCR-mediated signaling pathways in A. nidulans is presented in Figure 7.3.

7.2.2 G Protein α Subunits

The heterotrimeric G protein α subunit functions as the on–off switch that controls the duration of signal transduction by GPCRs. Once dissociated, the α subunit and/or the βγ dimer activate distinct downstream effectors as described in the introductory section (see Fig. 7.1). As in other filamentous fungi, three Gα
The Aspergilli proteins have been identified in all aspergilli whose genome has been sequenced except for A. oryzae, which has four. While FadA/GpaA and GanB/GpaB have been characterized in detail, little is known about the GanA Gα protein and the unusual A. oryzae GaoC protein.

### 7.2.2.1 FadA and GpaA

FadA and GpaA are highly conserved 353 amino acid-length proteins mediating vegetative growth signaling while inhibiting development in A. nidulans (FadA) and A. fumigatus (GpaA). The A. nidulans FadA (fluffy autolytic dominant) protein was identified by investigating a dominant activating mutation (d+: G42R) that caused enhanced accumulation of hyphae coupled with the lack of development followed by hyphal disintegration, which is known as the “fluffy autolytic” phenotype [20]. Constitutively active FadA mutant proteins are predicted to have reduced/absent intrinsic GTPase activity, which leads to the extended maintenance of activated FadA-GTP. The additional FadA+ mutant alleles including R178L, G183S, R178C, and Q204L also cause the fluffy-autolytic phenotype and the absence of the mycotoxin sterigmatocystin (ST) production [21–23]. On the contrary, the dominant interfering (d−) fadA G203R mutant allele causes reduced vegetative growth, hyperactive asexual sporulation, and precocious ST production [20,23]. Taken together, it is concluded that activated GTP-FadA mediates signaling that promotes vegetative growth, which in turn inhibits both asexual and sexual development as well as ST.
production (Fig. 7.3) [20–23]. Genetic studies have revealed that FadA-mediated signaling is in part transduced via cyclic AMP (cAMP)-dependent protein kinase A (PKA; see later) [24].

GpaA (G protein alpha A) is the A. fumigatus FadA homolog (97% amino acid level identity) [13,14, 25–27]. The introduction of the constitutively active gpaAQ204L allele in a wild type strain causes elevated hyphal proliferation and reduced sporulation in a dominant manner, but not autolysis [27]. Furthermore, somewhat similar to what has been observed in A. nidulans, the ectopic integration of the gpaAG203R allele results in reduced colony radial growth with normal conidiation levels. Importantly, the introduction of the gpaAG203R mutant allele into a mutant defective for a Regulator of G protein Signaling (RGS) protein restores conidiation in both air-exposed and liquid-submerged culture conditions (see later) [27]. Collectively, it has been proposed that the FadA homolog GpaA mediates signaling that stimulates hyphal growth while inhibiting asexual sporulation in A. fumigatus [26,27].

### 7.2.2.2 GanB and GpaB

In A. nidulans two additional Gα subunits (GanA and GanB; GAN stands for G protein alpha subunit in A. nidulans) have been identified, where only GanB (356 aa) has been functionally characterized [28,29]. The ganB deletion and dominant interfering (G207R) mutants abundantly produce conidiophores in liquid submerged cultures, indicating that GanB plays a role in down-regulating asexual development. Somewhat unpredictably, constitutively active GanB mutant alleles (Q208L and R182L) cause reduced hyphal growth and severely defected asexual development. Moreover, whereas the null or dominant interfering ganBG207R mutants exhibit reduced germination rates, the constitutively active ganBQ208L mutant promotes not only precocious conidial germination but also germination of conidia without any external carbon source [29]. Taken together, it is proposed that GanB negatively controls asexual development, but positively regulates conidial germination likely via sensing external carbon sources (Fig. 7.3) [29]. In fact, a later study has demonstrated that GanB mediates a rapid and temporary activation of cAMP synthesis in response to glucose during the early period of spore germination [30]. It has also been shown that GanB and SfaD-GpgA (Gβγ subunit, see later) constitute a functional heterotrimer controlling the response to glucose and consequently conidial germination, where GanB is a primary signaler and SfaD-GpgA aids proper activation of GanB signaling (Fig. 7.3) [30]. The function of GanA remains to be uncovered.

The A. fumigatus GanB homolog is GpaB (356 aa) showing 96% identity with GanB [13,26,31]. GpaB-mediated signaling is (mostly) transduced via PkaC1, a PKA catalytic subunit (see later), which is required for proper vegetative growth and normal asexual development. Deletion of gpaB and acyA (encoding an adenylate cyclase, see later) eliminate PKA activity, and supplementation of cAMP restores PKA activity in crude extracts of both the gpaB and acyA deletion strains [31]. In a low-dose murine inhalation model, the conidia of both the pkaC1 and gpaB deletion mutants are almost nonvirulent [31]. Interestingly, the expression of pksP encoding a polyketide synthase contributing to the pathogenicity of A. fumigatus is reduced in the gpaB deletion mutant. Moreover, the conidia of both the acyA and gpaB deletion mutants are much more susceptible to killing by human monocyte-derived macrophages than the wild-type conidia. Collectively, these findings indicate that the GpaB→AcyA→PkaC1 cascade constitutes a functional signaling branch that controls mechanisms by which the fungus is protected against attack by host immune effector cells [31]. A potential role of GpaB in sensing carbon sources and conidial germination remains to be investigated.

### 7.2.3 G Protein β Subunit

The Gβ subunit (SfaD) of A. nidulans is composed of 352 amino acids sharing 60% identity with mammalian Gβ subunits [28,32]. SfaD exhibits a conserved Trp-Asp sequence referred to as the “WD-40” motif [32]. Deletion of sfaD causes elaboration of conidiophores in liquid submerged culture, highly limited hyphal branching, delayed germination, restricted vegetative growth, the lack of ST production, and severe defects in sexual fruiting body formation, indicating that SfaD is required for normal hyphal growth, branching, sexual development, ST production, and proper regulation of asexual sporulation [30,32,33]. However, deletion of sfaD cannot suppress the fluffy-autolytic phenotype caused
by the FadA<sup>Δ</sup> (R178C and Q204L) alleles, indicating that constitutive activation of FadA-mediated signaling is sufficient to trigger vegetative growth signaling and that FadA might be the main signaler for hyphal proliferation [28,32]. Elimination of FadA or SfaD cannot bypass the need for FluG (an early developmental activator) [34] in asexual development, suggesting that the two (vegetative growth and asexual development) signaling pathways are separate and independent [23,32]. The requirement of SfaD for ST production is shown to be via transcriptional activation of <i>aflR</i>, which encodes a fungus-specific Zn(II)<sub>2</sub>Cys<sub>6</sub> transcriptional activator [33,35,36]. Overexpression of <i>aflR</i> under the inducible promoter <i>alcA(p)</i> restores ST production in the <i>sfaD</i> deletion mutant [33]. These results indicate that individual G protein components may play differential (or opposite) roles in controlling ST production and one of the end results of SfaD signaling may include transcriptional activation of <i>aflR</i>. Functions of the <i>A. fumigatus</i> SfaD homologue (387 aa; EAL91392; 97% identity) [13,14,26] remain to be studied.

### 7.2.4 G Protein γ Subunit

The <i>A. nidulans</i> Gγ subunit GpgA consists of 90 amino acids and exhibits 72% similarity with the yeast Ste18p [37]. It contains a characteristic coiled-coil (or GGL; G gamma-like) domain at the N-terminal region, which is required for the interaction of a Gγ with the cognate Gβ to form a heterodimer [3,37]. The <i>gpgA</i> null mutant displays delayed germination, restricted vegetative growth, and reduced/delayed asexual development [30,37]. Furthermore, deletion of <i>gpgA</i> causes severely impaired sexual fruiting body formation in self-fertilization and outcrosses, suggesting that the SfaD-GpgA heterodimer is the principal signaler for sexual development in <i>A. nidulans</i>. Deletion of <i>gpgA</i> cannot bypass the requirement for FluG in asexual development, and GpgA is also found to be required for the production of ST [33]. It appears that only one of Gβ and Gγ subunit exists in the <i>A. nidulans</i> genome. Functions of the probable <i>A. fumigatus</i> GpgA homolog (90 aa; DQ677630; 99% aa level identity) [13,14,26] have not been studied.

### 7.2.5 Phosducin-Like Proteins (PhLPs)

Phosducin or phosducin-like proteins (PhLPs) are a group of evolutionarily conserved proteins that positively regulate Gβγ signaling. PhLPs act as molecular chaperones for Gβγ assembly and are needed for proper levels of both Gβ and Gγ proteins [38–40]. The <i>A. nidulans</i> genome contains three potential PhLPs (PhnA, PhnB, and PhnC) [33] and functions of PhnA (281 aa), similar to Bdm-1 [38] a known fungal Gβγ activator, have been studied. The absence of <i>phnA</i> results in phenotypes (almost) identical to those caused by deletion of <i>sfaD</i> (see earlier), but different from those of the <i>gpgA</i> deletion mutant, suggesting that PhnA is essential for SfaD functionality. Similar to SfaD and GpgA, PhnA is necessary for sexual fruiting body formation in a dominant manner [33]. Taken together, it has been proposed that the SfaD-GpgA heterodimer is the primary signaler for sexual development, and PhnA is required for the activity of SfaD. No PhLPs in other aspergilli have been studied.

### 7.3 Downstream Signaling Branches

#### 7.3.1 Adenylate Cyclase and cAMP-Dependent Protein Kinases (PKAs)

Cyclic AMP (cAMP) produced by adenylate cyclase and cAMP-dependent protein kinase (PKA) play a central role in regulating morphology, growth, development, stress response, and virulence in a number of fungi [5–8]. In the absence of cAMP, the PKA holoenzyme exists as an inactive hetero-tetramer composed of a homo-dimeric regulatory subunit (PkaR) and two associated catalytic subunits (PKA) [41,42]. The regulatory subunit also prevents the inactive PKA holoenzyme from entering the nucleus [43,44]. The cooperative binding of two cAMP molecules to each regulatory subunit of the enzyme causes the dissociation of the active PKAs from the regulatory subunits. These active PKAs can phosphorylate downstream target proteins at serine or threonine residues [41,42].
Production of cAMP by adenylate cyclase represents a pivotal outcome of the sensitization of GPCRs and subsequent activation of heterotrimeric G proteins (see Fig. 7.3). The CyaA and AcyA adenylate cyclases have been characterized in *A. nidulans* and *A. fumigatus*, respectively [25,45]. Both proteins are large (>2000 aa) and show a modular architecture similar to that of other fungal adenylate cyclases with a leucine-rich repeat domain, a protein phosphatase 2C domain, and a highly conserved catalytic domain. Inactivation of the cyaA or acyA genes in their cognate *Aspergillus* species results in defects in conidial germination, hyphal elongation, and conidiogenesis, indicative of the role that cAMP plays at the different stages of *A. nidulans* development. Interestingly, deletion of the cyaA gene is not lethal in *A. nidulans*, whereas the simultaneous mutational inactivation of the two PKA catalytic subunits is lethal (see later) [46]. This suggests that even in the absence of cAMP produced by adenylate cyclase, PKA activity may remain above a threshold necessary for germination, hyphal growth, and conidiation.

Activation of adenylate cyclase in *S. cerevisiae* is mediated by the Ras1 and Ras2 proteins (the homologs of the human p21ras oncogene protein) and the Gpa2 Gα subunit. The functions of the *A. nidulans* RasA protein have been studied using dominant activating (G17V) and dominant interfering (S22N) mutations [47,48]. RasA has been proposed to control the sensing of the carbon source at the onset of conidial germination and the subsequent switch from isotropic to polar growth during conidial germination. Yet, the latter function of RasA is not mediated by activation of adenylate cyclase suggesting that RasA regulates *A. nidulans* development via another signaling pathway, possibly a MAP kinase pathway [45]. In contrast and as outlined earlier, evidence has now been provided that adenylate cyclase is in part controlled by the GanB Gα subunit in *A. nidulans* (Fig. 7.3) [30]. Indeed, a transient accumulation of cAMP is observed at the onset of spore germination and this phenomenon is GanB-dependent [30]. It is likely that the *A. fumigatus* GpaB Gα protein also regulates adenylate cyclase activity since addition of di-butyryl cAMP can suppress the phenotypes caused by deletion of *gpaB* and *acyA* in *A. fumigatus* [25]. Whether FadA/GpaA also regulate adenylate cyclase activity remains to be explored since a link of these Gα subunits with the cAMP signaling pathway has only been established at the level of the cAMP-dependent protein kinase [24].

In almost all cases, genomes of filamentous fungi contain two distantly related PKAs, where only one PKA is found to play a predominant role. In *A. nidulans*, PkaA and PkaB constitute the essential PKA catalytic subunits that play overlapping and opposite roles in diverse biological processes (Fig. 7.4) [24,46]. PkaA is the principal PKA and transduces the aforementioned FadA-mediated vegetative growth signaling [24]. In addition, the fact that GanB, SfaD-GpgA, and PkaA are required for proper germination of conidia [29,30] indicates that both GanB- and FadA-mediated signals are transduced via PkaA. As PkaA is a key downstream element in FadA-mediated signaling, the absence of *pkaA* function results in restricted vegetative growth and hyperactive conidiation [24]. Moreover,

**FIGURE 7.4** Overlapping and opposite roles of PkaA and PkaB in *A. nidulans.*
deletion of pkaA suppresses the fluffy-autolytic phenotype caused by the dominant activating fadA^{G42R} allele [24]. In addition, overexpression of pkaA leads to enhanced hyphal accumulation coupled with reduced sporulation and ST production [24]. Taken together, it has been proposed that the cAMP and PkaA signaling cascade plays a major role in activation of vegetative growth, repression of conidiation, and conidial germination in A. nidulans (Fig. 7.4) [46].

The secondary PKA catalytic subunit PkaB functions as a backup unit for hyphal growth and spore germination [46]. Although deletion of pkaB alone does not cause distinct phenotypic changes, the absence of both pkaB and pkaA is lethal, indicating that PkaB and PkaA are essential for viability of A. nidulans. Overexpression of pkaB enhances hyphal proliferation and rescues the growth defects caused by deletion of pkaA, indicating that PkaB plays a role in stimulating vegetative growth. However, deletion of pkaB does not suppress the fluffy-autolytic phenotype resulting from deletion of fadA, implying that PkaB is not a key signaling component for hyphal growth. While up-regulation of pkaB rescues the defects of spore germination resulting from the absence of pkaA in the presence of glucose, overexpression of pkaB delays spore germination. Furthermore, up-regulation of pkaB completely blocks spore germination on medium lacking added carbon sources. In addition, up-regulation of pkaB enhances the level of submerged sporulation caused by deletion of pkaA and reduces hyphal tolerance to oxidative stress. In summary, PkaB is the secondary PKA that has a synthetic lethal interaction with PkaA, and plays overlapping roles in vegetative growth and spore germination in the presence of glucose, but opposite roles in regulating asexual sporulation, germination in the absence of external carbon sources, and oxidative stress responses in A. nidulans (Fig. 7.4) [46].

As found in A. nidulans, two PKA catalytic subunits are present in the A. fumigatus genome [13,14], where PkaC1 (86% identical to A. nidulans PkaA) plays a principal role in regulating vegetative growth and development [26,31]. Deletion of pkaC1 results in restricted growth and delayed germination, indicating that, as in A. nidulans, PkaC1 is necessary for proper vegetative growth and germination in A. fumigatus. However, distinct from A. nidulans, deletion of pkaC1 causes reduced sporulation. As mentioned, GpaB, AcyA, and PkaC1 are proposed to constitute a major signaling cascade controlling vegetative growth, development, and virulence [25,31]. Importantly, the cAMP-PKA signaling pathways are required for proper expression of pksP encoding a polyketide synthase involved in the biosynthesis of the conidial pigment 1,8-dihydroxynaphthalene-like pentaketide melanin, which confers resistance to phagocytic cell destruction in the host. Thus, deletion of pkaC1 causes dramatically lowered expression of pksP, which contributes to the reduced virulence of the mutant [25,31]. The potential role of PkaC1 in the GpaA vegetative signaling cascade remains to be studied.

Both A. nidulans and A. fumigatus have a unique gene encoding the regulatory subunit of PKA and these have been designated pkaR in both species [49–51]. Inactivation of the pkaR gene in A. fumigatus results in reduced germination and growth rates as well as a defect in conidiogenesis, an unexpected result in light of the defective sporulation of the A. fumigatus pkaC1 mutant [50]. Importantly, inactivation of the pkaR gene increases the sensitivity of A. fumigatus to oxidative damage and, as a probable consequence, reduces the virulence of the fungus [50]. In A. nidulans, phenotypes associated with inactivation of the pkaR gene inversely mirror those resulting from deletion of the pkaA gene [51]. Indeed, deletion of pkaR causes conidial germination in the absence of a carbon source and a severe defect in sporulation. These phenotypes are suppressed by deletion of the pkaA gene but are unaffected by deletion of the cyaA gene, confirming that PkaR is a negative regulator of PkaA activity that acts downstream of adenylate cyclase [51].

### 7.3.2 Mitogen Activated Protein (MAP) Kinases

The universally conserved MAP kinase (MAPK) cascade (see Fig. 7.1) is one of the most ubiquitous signal transduction systems. This pathway is activated by a variety of stimuli. Upon activation, MAPK cascades regulate numerous physiological processes, including growth, differentiation, and high-osmolarity responses [52]. Signals are transduced by sequential phosphorylation and activation of the MAPK components specific to an individual signaling branch. In S. cerevisiae, five MAPK modules controlling mating, filamentous growth, high-osmolarity responses, cell wall remodeling, and sporulation have been studied [53,54]. This section only discusses the terminal protein kinases MAPKs in two aspergilli.
The genomes of *A. nidulans* and *A. fumigatus* contain four genes encoding MAP kinases, *mpkA*, *mpkB*, *mpkc*, and *sakA/hogA* (Fig. 7.1) [12,13,55–59]. In *A. nidulans*, the functions of *mpkA* and *sakA* have been characterized, whereas *mpkB* and *mpkc* are yet to be studied. *Mpka* (418 aa) is similar to the *S. cerevisiae* MAPK Slt2p involved in regulating the maintenance of cell wall integrity and progression through the cell cycle [59]. Deletion of *mpkA* causes impairment in conidial germination and hyphal tip growth, of which defects are partially suppressed by growing the *mpkA* deletion mutant on high-osmolarity medium [59]. The swollen hyphal tips of the *mpkA* deletion mutant indicate that MkA may function in cell wall biosynthesis and polarized growth of *A. nidulans*.

Deletion of *sakA/hogA* results in reduced (up to 60%) hyphal extension rates in the presence of high salt, enhanced branching of hyphal tips, uneven accumulation of nuclei, and the absence of septa [55]. It is proposed that SakA/HogA functions to maintain turgor pressure, which is required for proper cell expansion. A later study showed that SakA is activated in response to osmotic and oxidative stress in both *S. pombe* and *A. nidulans* [56]. Furthermore, the *sakA* deletion mutant displays premature sexual development, and produces asexual spores that are highly sensitive to oxidative and heat shock stress coupled with reduced viability upon storage. The *sakA* gene is transiently activated shortly after induction of asexual development. Taken together, it is proposed that SakA/HogA is involved in stress signal transduction and repression of sexual development, and is required for spore stress resistance and survival.

Among the four *A. fumigatus* MAPKs, SakA and Mpkc share 68% amino acid level identity, and Mpka and Mpkb share 56% identity [57,58,61]. Functions of SakA and Mpkc have been studied in *A. fumigatus*. While asexual spores of the *sakA* deletion mutant germinate and grow in the presence of osmotic pressure, germlings of the deletion mutant arrest growth in response to hypertonic stress, which is a similar phenotype to that observed in *S. cerevisiae* [57]. These results suggest that the SakA signaling pathway may be inactive in metabolically dormant spores (or used for other purposes), but actively involved in cellular responses to hypertonic stress in vigorously growing hyphae [57,58]. The inability to reinitiate vegetative growth following a hypertonic shock is a characteristic feature of mutants in this MAPK [58,60]. Somewhat unexpectedly, SakA also regulates conidial germination in response to the nitrogen source in the medium, and the mRNA of *sakA* accumulates in response to nitrogen or carbon starvation in *A. fumigatus* [57]. These results indicate that the conserved SakA MAP kinase pathway negatively regulates conidial germination and is activated in response to starvation for nitrogen or carbon sources.

A recent study characterized the functions of *mpkC* in *A. fumigatus* [61]. The *mpkC* deletion mutant is viable and exhibits normal conidial germination and hyphal growth on minimal or complete medium. Moreover, *mpkC* is dispensable for the tolerance of the fungus to osmotic, oxidative, and thermal stresses, indicating that the SakA and Mpkc signaling pathways are separate and minimally overlapping in response to these environmental signals. Importantly, the *mpkC* deletion mutant is unable to grow on minimal medium containing polyalcohol sugars, for example, sorbitol or mannitol, as a sole carbon source. This result implicates the Mpkc signaling pathway functions in carbon source sensing and utilization [61].

### 7.3.3 Protein Kinase C (PKC)

Certain GPCR-mediated signals are transduced through phospholipid signaling pathways, which involve the hydrolysis of the membrane phosphoinositide phosphatidylinositol 4,5-biphosphate (PIP$_2$) by phospholipase C (PLC), yielding two essential second messengers, inositol 1,4,5-trisphosphate (IP$_3$), and diacylglycerol (DAG; see Fig. 7.1) [62]. In mammalian cells, IP$_3$ binds to specific receptors and induces the release of calcium from storage organelles, whereas DAG activates protein kinase C [63]. Protein kinase C (PKC) comprises a superfamily of isoenzymes, which must first undergo a series of phosphorylations in order to activate downstream target proteins [63].

A recent study characterized a PKC in *A. nidulans* [64]. The *A. nidulans* genome appears to contain two putative PKCs [12,64]. PkcA shows all the architectural features of fungal PKCs, whereas PkcB apparently lacks some of the characteristic features of PKCs. Due to the evident essentiality of PkcA for the viability the fungus, its functions have been examined by expressing *pkcA* antisense RNA under the controllable promoter *alcA(p)* [64]. Production of *pkcA* antisense RNA in *A. nidulans* results in reduced growth and asexual development in *Aspergillus* minimal medium, while in fermentation medium it leads
The Aspergilli

to reduction in penicillin production, and predominant localization of AnBH1 (a negative regulator of penicillin biosynthesis) in the cytoplasm. Taken together, it has been proposed that PkcA is involved in penicillin biosynthesis via regulation of the nuclear localization of the transcription factor AnBH1 [64].

### 7.4 Negative Regulators of G Protein Signaling: RGSs

RGS proteins are negative controllers of G protein mediated signaling. In general, a RGS protein interacts with a GTP-Gα subunit and increases its intrinsic GTPase activity, leading to rapid inactivation of GPCR-mediated signaling pathways [3,10,65]. Cells can properly convert diverse incoming signals into fine-tuned cellular responses via the activities of various RGS proteins. In addition, RGS proteins can enhance G protein activation, act as effector antagonists, and serve as scaffold proteins to congregate receptors, G proteins, effectors and other regulatory molecules [10]. In Aspergilli, Gα subunits and RGS proteins govern upstream regulation of vegetative growth, development and mycotoxin/pigment production [8,9,28].

#### 7.4.1 FlbA and AfFlbA

The *A. nidulans* RGS protein FlbA (719 aa) [66] is similar to *S. cerevisiae* Sst2p, carrying one RGS and two DEP (disheveled/Egl-10/pleckstrin) domains [28,67]. The presence of repeated DEP domains is apparently fungus-specific. The DEP domain might be associated with targeting RGS proteins to the golgi and plasma membranes and inducing the expression of a group of genes containing stress response elements (STRE) in the promoter regions [68].

The FadA and FlbA pair is the first studied Gα−RGS duo in filamentous fungi, and is responsible for upstream regulation of hyphal growth, asexual and sexual development as well as biosynthesis of ST and penicillin [8,9,20,23,66,69]. As mentioned earlier in this chapter, both FadA and the SfaD-GpgA heterodimer stimulate vegetative growth in part through PkaA (see Fig. 7.3). FlbA is a specific RGS protein controlling FadA-mediated proliferation signaling, likely by enhancing the intrinsic GTPase activity

![FIGURE 7.5 Summary of G protein-RGS controlled signaling pathways in *A. nidulans*. Two independent Gα−RGS signaling pathways coordinately control various biological processes. FlbA-FadA primarily governs vegetative growth vs. development [20,66], and RgsA-GanB controls stress response (pigmentation), sensing carbon sources and spore germination [29,30,67]. Asexual development occurs through activation of brlA, encoding a key transcription factor required for conidiophore development, which requires multiple upstream genes including fluG and flbE, flbD, flbD and flbC [26,34,71]. FluG-dependent conidiation occurs via removing repressive effects imposed by the potential transcription factor SfgA with the Zn(II)2Cys6 motif [26,72,73]. Both GanB and SfaD-GpgA have been proposed to function in inhibition of asexual development (conidiation) [28,29,32,33,37]. A potential involvement of SakA (HogA) in GanB-mediated signaling is indicated [35,56,67].](image)
of FadA (Fig. 7.5) [20,66]. Loss of flbA function results in the fluffy-autolytic phenotype (almost) identical to that caused by the constitutively active FadA* mutant alleles [20–22,66]. As FadA is the primary target of FlbA function, the deletion and dominant interfering (G203R and R205H) FadA mutations suppress the fluffy-autolytic phenotype caused by deletion of flbA and restore asexual development and ST production [20,22,23]. Likewise, deletion of sfaD, gpgA, or phnA bypasses the need for FlbA in asexual development [32,33,37].

The A. fumigatus AfFlbA protein is composed of 712 amino acids and shows 79% identity with the A. nidulans FlbA protein [13,26,27]. Functional characterization of the AfflbA gene reveals that AfFlbA down-regulates hyphal proliferation, which in turn stimulates development [32,33,37]. However, distinct from A. nidulans, loss of AfflbA function does not eliminate asexual sporulation or lead to hyphal disintegration (autolysis) in A. fumigatus, suggesting that multiple mechanisms activate development in A. fumigatus and might bypass the need for AfFlbA in sporulation and allow AfflbA mutants to produce spores, thereby avoiding hyphal disintegration [26,27]. A series of genetic tests have confirmed that the FadA homolog GpaA is the primary target of AfFlbA [27]. First, similar to the effects caused by deletion of AfflbA, the introduction of the gpaA^Q203R allele in wild type causes elevated hyphal proliferation coupled with reduced sporulation in a dominant manner. Second, the ectopic integration of the gpaA^Q203R allele restores sporulation in an AfflbA loss of function mutant to the wild-type level in both air-exposed and liquid-submerged culture conditions [27]. These results indicate that inactivation of the GpaA signaling pathway bypasses the requirement of AfFlbA for developmental progression. In summary, GpaA and AfFlbA constitute a Gtr-RGS pair, which coordinates vegetative growth and development in A. fumigatus. Thus, the primary roles of FadA/GpaA and FlbA/AfFlbA are conserved in the two aspergilli [26,27].

7.4.2 RgsA

The second Aspergillus RGS protein studied is RgsA [67], displaying 28% identity and 43% similarity to S. cerevisiae Rgs2p [70]. Dissimilar from constitutively expressed flbA [66], rgsA mRNA (~2.0 kb) accumulates at high levels during early vegetative growth phase, decreases during asexual and sexual development, and increases in ascospores, indicating that rgsA is subjected to complex transcriptional control in A. nidulans [67].

RgsA is a specific RGS protein that negatively regulates GanB signaling (see Figs. 7.3 and 7.5) [28,30,67]. As deletion of rgsA would result in prolonged activation of GTP-bound GanB, it causes phenotypic changes highly similar to those resulting from the constitutive active ganB^Q208L allele [29], that is, reduced vegetative growth, germination of conidia in the absence of external carbon sources, and enhanced accumulation of brown mycelial pigments [67]. Supporting the primary role of RgsA in regulating GanB activity, only ganB deletion suppresses morphological, physiological, and metabolic alterations caused by deletion of rgsA [67]. Moreover, overexpression of rgsA causes abundant formation of conidiophores in liquid-submerged culture as observed in the ganB null and ganB^Q208L mutants [29,67].

RgsA is also involved in regulation of the cAMP/PKA pathway and conidial germination via attenuation of GanB signaling (Fig. 7.3) [30]. Importantly, the fact that deletion of rgsA causes elevated mycelial and conidial pigmentation levels, and enhanced oxidative-/thermo-tolerance [67] suggests that GanB signaling is associated with activation of stress responses and RgsA is required for tight regulation of this potentially energy draining process. In summary, GanB activates the AcyA-cAMP-PkaA signaling pathway, which in turn induces various stress responses in A. nidulans, and RgsA negatively regulates the GanB–PkaA pathway (Figs. 7.3 and 7.5). This model is opposite to the stress response mechanism in S. cerevisiae where deletion of Rgs2p reduced thermal tolerance, while overexpression of Rgs2p caused significant elevation in heat resistance [70]. Functional characterization of the A. fumigatus RgsA homolog and its interaction with GpaB remain to be investigated.

7.4.3 Other RGSs

Functions of the RgsB, RgsC, and GprK proteins are currently being characterized in A. nidulans and only limited information is available [12,14,28,67]. The rgsB gene encodes a 2.5 kb transcript, which is present at relatively constant levels throughout the life cycle of A. nidulans [67], and encodes a protein
similar to yeast Rax1p [74]. Rax1p has been implicated in bipolar budding in *S. cerevisiae*. Both RgsA and RgsB-type RGS proteins are found in fungi only [28,67].

The RgsC protein contains the RGS box in the center, and PhoX-associated (PXA) and PhoX (PX) domains at the N- and C-termini, respectively [28,67,75]. The PX domain might act as a sorting signal to allow proteins to reach their appropriate location by binding to phosphoinositides [76]. The C-terminus (780th aa ~) of RgsC including the PX domain is similar to *S. cerevisiae* Mdm1p (443 aa), which is known to be required for transmission of nuclei and mitochondria to daughter cells [77,78]. It is speculated that RgsC might function in coordinating heterotrimeric G-protein signaling, hyphal extension, nuclear positioning, and organelle transport (vesicular trafficking) [14,28,67].

GprK is unique in that it contains both 7-TM and RGS domains [12,14,28] and is similar to the *Arabidopsis thaliana* RGS protein AtRGS1 that has been shown to negatively regulate the Gpa1 Gα subunit affecting cellular propagation in *A. thaliana* [79]. The presence of GprK-like proteins (class VI GPCRs) in filamentous fungi suggests that the dual function signaling GPCRs may play crucial roles in other eukaryotes.

Homologs of the three aforementioned RGS proteins are identified in *A. fumigatus* and *A. oryzae* [12–14]. Interestingly, these two species differ from *A. nidulans* by the presence of the fifth RGS protein RgsD (see Fig. 7.1), which is similar to RgsA [12–14]. Moreover, while the genomic regions that carry the *rgsA* gene are highly syntenic among the three species, those with *rgsD* are not syntenic to themselves or to the *rgsA* region, and are rich in duplicated genes, suggesting that *rgsD* might have emerged through an ancestral duplication or horizontal transfer [14]. Similarity between RgsA and RgsD implies that RgsD might also modulate the activity of one of the three Gα subunits identified in *A. fumigatus* and *A. oryzae*.

### 7.5 Conclusions and Prospects

Cells are constantly exposed to a variety of signals and must respond to external and/or internal signals and elicit appropriate cellular changes. Our knowledge of signal transduction and its regulation in the model fungus *A. nidulans* has greatly increased in the last decade, and this knowledge is getting expanded to the pathogenic fungus *A. fumigatus*. The near-complete identification and characterization of both positive (GPCRs, G proteins, PhLPs, and effectors) and negative (RGS proteins) controllers of G protein signaling in aspergilli, especially in *A. nidulans* and *A. fumigatus*, will help us to better understand the mechanisms underlying morphogenesis, pathogenicity, and toxigenesis in this important genus.

One of many important points is that filamentous fungi have a remarkable number of putative GPCRs. Greater than 16 GPCR candidates are found in the three Aspergillus species whose genome sequence has been reported. Moreover, the *A. nidulans* genome may have at least 25 genes encoding PTH11-like membrane-spanning proteins. The *Magnaporthe grisea* PTH11 protein is proposed to form a new GPCR family based on its structure and its role in the regulation of appressorium development [80–82]. The diversity of GPCRs in filamentous fungi may reflect the ability of filamentous fungi to prosper in various ecological niches as well as the complexity of their life cycles involving differentiation of multicellular asexual and sexual reproductive structures. In any case, it should be noted that in most instances these putative GPCRs have only been characterized *in silico*. Moreover, even for those three putative GPCRs (GprA/B/D) characterized [11,15] it is not known through which heterotrimeric G protein(s) these GPCRs transmit signals. Thus, much remains to be studied in order to evaluate their biological contributions to signal transduction pathways as well as to identify their cognate heterotrimeric G proteins.

Another important point is the presence of three Gα subunits in filamentous fungi, whereas hemiascomycetous yeasts have only two Gα subunits [14]. Within the genus Aspergillus, *A. oryzae* harbors the fourth Gα subunit GaoC, although it is likely nonfunctional [12–14]. Moreover, the three sequenced aspergilli all have at least five proteins with the RGS box, where an additional protein with an RGS domain, RgsD, is found in *A. fumigatus* and *A. oryzae* [12–14]. The presence of an ample number of Gα and RGS proteins in filamentous fungi may also reflect their ability to thrive in a variety of environmental niches and to undergo complex multicellular developmental processes.

As discussed, we now have an extensive description of the components that may constitute signaling pathways in aspergilli and have elucidated some functional blocks. However, we have limited insights
into the detailed organization of signaling pathways in aspergilli and their contribution to various aspects of filamentous fungal biology. Moreover, we know relatively little about the links between the upstream and downstream components that form individual signaling pathways. The MAPK and PKC pathways, although identified, have not been linked to any GPCR. The only case where a link between a G protein and the cAMP-PKA signaling pathway has been formally established is for the *A. nidulans* GanB/CyaA/PkaA cascade during germination [30,45,51] and yet no upstream GPCR regulating GanB activity has been identified. Furthermore, molecules (if any) sensed by the multiple GPCRs predicted by genome mining remain to be identified. The same conclusions are true for other filamentous fungi such as *Neurospora crassa* or *Cryphonectria parasitica* where signaling pathways have been subject to intense investigation. Providing a detailed characterization of the signaling pipelines and their interplay thus remains a challenge for future research in filamentous fungi. In this context, aspergilli will undoubtedly continue to be key organisms to study because of the ability to employ both classical and molecular genetic approaches in the model species *A. nidulans* and because of the occurrence in the genus of medically and industrially important species such as *A. fumigatus* and *A. oryzae*.

**Acknowledgments**

We would like to express sincere thanks to Drs Jeong-Ah Seo, Sabine Fillinger, Kap-Hoon Han, Jae-Hyung Mah, Anne Lafon, Min Ni, and other fungal biologists for their great contributions toward a full understanding of signal transduction mechanisms in aspergilli and other filamentous fungi. This work was supported by National Science Foundation grant MCB-0421863 to JHY.

**References**


Signal Transduction in Aspergilli

49. Oliver, B.G. et al., Cloning and expression of pkaC and pkaR, the genes encoding the cAMP-dependent protein kinase of *Aspergillus fumigatus*, *Mycoptology*, 154, 85, 2002.


8
Gene Regulation

Mark X. Caddick and Chris Dobson

CONTENTS
8.1 Introduction ............................................................................................................. 103
8.2 Levels of Regulation ............................................................................................. 104
8.3 Transcription Factors ............................................................................................ 104
8.4 Regulatory Signals ................................................................................................. 110
8.5 Identification of Regulatory Motifs in Silico ......................................................... 111
8.6 Transcriptomics and Proteomics ............................................................................... 113
8.7 Additional Regulatory Mechanisms ......................................................................... 114
References .................................................................................................................. 115

8.1 Introduction

In the postgenomic era the importance of gene expression and regulation is becoming ever more apparent. Expression is ultimately related to function and its characterization inevitably helps in identifying the biological role of genes and proteins. Putative functional links between genes can be formed on the basis of observed coordination of expression and this in turn provides valuable information about the components of specific biological processes. Gene expression is a fundamental aspect of gene function that can be efficiently monitored globally, through transcriptomics, providing a wealth of additional information that can be mapped back onto the genome. The manipulation of biological processes will often require altered expression of genes and proteins; consequently, an understanding of how this can be achieved is invaluable. Proteins with central functions in regulation are also potentially very good targets for antifungals, where their subversion can potentially disrupt key processes involved in pathogenicity. The regulatory systems are intrinsically important biological systems in themselves, justifying an increased awareness and understanding of the components involved, their molecular interactions, and mechanisms of function. Finally, the development and evolution of these processes will inform us about key processes involved in speciation and adaptation.

Filamentous fungi, in particular *Aspergillus nidulans* and *Neurospora crassa*, have played a significant role in developing our understanding of the processes and mechanisms underlying gene regulation. Fundamental concepts relating to regulatory genes and promoters, and the coordination of gene expression in response to diverse signals, have been elegantly elucidated using these fungi as model genetic systems. The legacy is a detailed understanding of various processes and networks such as fungal development,1 the response to ambient pH2 and the regulation of metabolism.3,4 In many instances, this has been translated to an understanding of the equivalent systems in related organisms of medical and biotechnological importance. Much of this work has been extensively reviewed elsewhere,5,6 but it is our intention to focus on new developments and in particular the prospects that have arisen as a consequence of the genome sequences from a variety of aspergilli and other filamentous fungi becoming available. The emphasis is primarily on transcriptional regulation but we briefly discuss other regulatory mechanisms that have been identified in the aspergilli.
8.2 Levels of Regulation

Much of the information present within the genome is not currently understood. We cannot look at sequences and determine the level of gene expression or its regulation. We even find it difficult to accurately predict the location of genes within a DNA sequence. In particular, the transcription start site and termination sequence are problematic. This is because the genome, far from being just a linear DNA sequence encoding proteins and functional RNA molecules, is multidimensional. The DNA sequence defines the binding sites for regulatory proteins, the relative proximity and arrangement of which combine with the DNA’s physical properties, such as flexibility and melting, to determine the function of a given motif. The genomes’ functional form is the same as chromatin where the level of condensation at specific locations is a dynamic and regulated feature which significantly affects accessibility to the transcriptional machinery. The location of nucleosomes is partly dependent on the primary sequence, which affects flexibility. Other DNA-binding proteins will also affect their distribution and some of these are directly involved in regulating the level of chromatin condensation through modification of the DNA and histones. DNA is subject to chemical modifications, which will alter various properties. However, although there is some evidence for methylation in *Aspergillus* species, the levels of methylation appear to be very low.\(^7,8\) This complexity is further amplified by downstream events, such as RNA maturation and splicing, RNA stability and movement within the cell, translational efficiency, and posttranslational modification and processing. All are aspects of gene regulation and are ultimately defined within the genome. Each process has its own language that will be reflected in the genome’s sequence; consequently, the genome will have all these diverse instructions superimposed and melded together.

There are numerous examples of heterologous fungal genes being expressed in filamentous fungi. Generally, the machinery involved responds appropriately, even to the extent that genes are properly regulated. Consequently, we can assume that much of the information is well conserved. Effectively *A. nidulans* can recognize an *A. niger* or an *N. crassa* promoter and translate the transcript correctly. Similarly a wide range of transcription factors function heterologously,\(^9,10\) although they have often diverged significantly outside the DNA-binding domain. This level of conservation has been a useful tool in characterizing function and facilitating heterologous gene expression. It also provides a valuable tool for genome analysis of regulatory processes through comparative genomics, as critical sequence elements and features involved in gene expression will be conserved where you have functional conservation.

8.3 Transcription Factors

Generally the initiation of transcription is seen as the most important point at which gene expression is regulated. Significant effort has been focused on characterizing the regulatory proteins that interact at specific promoters to either activate or repress transcription. About 45 transcription factors have been identified and characterized, primarily in *A. nidulans*, and for many the respective DNA-binding motifs have also been identified (Table 8.1). However, this probably represents only about 5% of the transcription factors encoded in the genome. Considering the other known components of the transcription, RNA processing and translation machinery, including the general transcription factors, RNA polymerase components, coactivators, histone acetylases, and so on, a significant proportion of the genome is dedicated to gene expression and regulation.

The availability of the genome sequence has allowed us to identify a large number of additional genes that putatively encode DNA-binding proteins. A large proportion of these genes are likely to be involved in gene regulation. Using PFAM domains as a means of identifying DNA-binding proteins encoded within the sequenced *Aspergillus* genomes, we have identified 86 distinct classes of domain (Table 8.2). The function of these proteins will include transcriptional regulation as well as a range of activities such as DNA repair and replication, chromatin binding, and telomere integrity. Comparing the distribution of these domains across four *Aspergillus* species (*A. niger, A. nidulans, A. fumigatus, A. oryzae*) and as examples of distinct and distantly related ascomycetes *N. crassa* and *S. cerevisiae*, the relative distribution is generally well conserved. However, with respect to a few specific classes this is not the case.
TABLE 8.1
Known Regulatory Proteins in *A. nidulans*

<table>
<thead>
<tr>
<th>Factor</th>
<th>Function</th>
<th>Binding Domain</th>
<th>Recognition Motif</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbaA</td>
<td>Asexual development</td>
<td>TEA domain</td>
<td>CATTCY</td>
<td>60, 61</td>
</tr>
<tr>
<td>AflR</td>
<td>Secondary metabolism</td>
<td>Zn binuclear cluster</td>
<td>TCGN, CGA</td>
<td>62, 63</td>
</tr>
<tr>
<td>AlkR</td>
<td>Alcohol metabolism</td>
<td>Zn binuclear cluster</td>
<td>RNGCGG–AT rich</td>
<td>16, 64</td>
</tr>
<tr>
<td>AmdA</td>
<td>Acetamide utilization</td>
<td>C2H2 zinc finger × 2</td>
<td>GMGGGG</td>
<td>65, 66</td>
</tr>
<tr>
<td>AmdR</td>
<td>Induction of specific N metabolism</td>
<td>Zn binuclear cluster</td>
<td>TTCGGCGWN, SCAAT</td>
<td>67, 68</td>
</tr>
<tr>
<td>AmdX</td>
<td>Acetamide utilization</td>
<td>C2H2 zinc finger × 2</td>
<td>CAGGG</td>
<td>35</td>
</tr>
<tr>
<td>AnBH1</td>
<td>Secondary metabolism</td>
<td>bHLH</td>
<td>TCACNNG</td>
<td>69</td>
</tr>
<tr>
<td>AnCF</td>
<td>General transcription factor</td>
<td>CBF-B/NF-YA</td>
<td>CCAAT</td>
<td>70, 71</td>
</tr>
<tr>
<td>ArcA</td>
<td>Arginine catabolism</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>72</td>
</tr>
<tr>
<td>AreA</td>
<td>Nitrogen metabolism</td>
<td>GATA C4 Zn finger</td>
<td>HGATAR</td>
<td>32, 37</td>
</tr>
<tr>
<td>AreB</td>
<td>Nitrogen metabolism</td>
<td>GATA C4 Zn finger</td>
<td>Not known</td>
<td>51</td>
</tr>
<tr>
<td>AtfA</td>
<td>Oxidative stress</td>
<td>bZIP</td>
<td>Not known</td>
<td>73</td>
</tr>
<tr>
<td>BrlA</td>
<td>Asexual development</td>
<td>C2H2 zinc finger × 2</td>
<td>MRAGGGR</td>
<td>74</td>
</tr>
<tr>
<td>CpcA</td>
<td>Amino acid biosynthesis</td>
<td>bZIP</td>
<td>TTGASTCWG</td>
<td>31</td>
</tr>
<tr>
<td>CreA</td>
<td>Carbon catabolite repression</td>
<td>C2H2 zinc finger × 2</td>
<td>SYGGRG</td>
<td>75, 76</td>
</tr>
<tr>
<td>DevR</td>
<td>Development</td>
<td>bHLH</td>
<td>Not known</td>
<td>77</td>
</tr>
<tr>
<td>FacB</td>
<td>Acetate utilization</td>
<td>Zn binuclear cluster</td>
<td>TCSN₄₋₅SGA</td>
<td>18</td>
</tr>
<tr>
<td>FarA</td>
<td>Fatty acid metabolism</td>
<td>Zn binuclear cluster</td>
<td>CCTCGG</td>
<td>29</td>
</tr>
<tr>
<td>FarB</td>
<td>Fatty acid metabolism</td>
<td>Zn binuclear cluster</td>
<td>CCTCGG</td>
<td>29</td>
</tr>
<tr>
<td>FlhB</td>
<td>Asexual development</td>
<td>bZIP</td>
<td>Not known</td>
<td>1</td>
</tr>
<tr>
<td>FlhC</td>
<td>Asexual development</td>
<td>C2H2 zinc finger × 2</td>
<td>Not known</td>
<td>1</td>
</tr>
<tr>
<td>FlhD</td>
<td>Asexual development</td>
<td>myb like</td>
<td>Not known</td>
<td>78</td>
</tr>
<tr>
<td>HacA</td>
<td>Unfolded protein response</td>
<td>b-ZIP</td>
<td>CAGCGTG</td>
<td>79, 80</td>
</tr>
<tr>
<td>JibA</td>
<td>Amino acid biosynthesis</td>
<td>bZIP</td>
<td>Not known</td>
<td>81</td>
</tr>
<tr>
<td>LacA</td>
<td>Secondary metabolism and differentiation</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>21, 82</td>
</tr>
<tr>
<td>MeaB</td>
<td>Nitrogen stress</td>
<td>b-ZIP</td>
<td>Not known</td>
<td>83</td>
</tr>
<tr>
<td>MetR</td>
<td>Sulfur metabolism</td>
<td>bZIP</td>
<td>Not known</td>
<td>84</td>
</tr>
<tr>
<td>NirA</td>
<td>Nitrate induction</td>
<td>Zn binuclear cluster</td>
<td>CTCCGHGG</td>
<td>36, 85</td>
</tr>
<tr>
<td>NosA</td>
<td>Fruiting body formation</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>86</td>
</tr>
<tr>
<td>NsdD</td>
<td>Sexual development</td>
<td>C4 GATA zinc finger</td>
<td>Not known</td>
<td>34</td>
</tr>
<tr>
<td>PacC</td>
<td>pH response</td>
<td>C2H2 zinc fingers × 3</td>
<td>GCCARG</td>
<td>87</td>
</tr>
<tr>
<td>PalC</td>
<td>Phosphate metabolism and cell cycle</td>
<td>HLH</td>
<td>Not known</td>
<td>88</td>
</tr>
<tr>
<td>PrnA</td>
<td>Proline induction</td>
<td>Zn binuclear cluster</td>
<td>CCGGNCCGG</td>
<td>17, 89</td>
</tr>
<tr>
<td>QutA</td>
<td>Quinate utilization</td>
<td>Zn binuclear cluster</td>
<td>CCGGN₄₋₅CCGG</td>
<td></td>
</tr>
<tr>
<td>RlmA</td>
<td>Cell wall modification</td>
<td>MADS box</td>
<td>Not known</td>
<td>92</td>
</tr>
<tr>
<td>(A. niger)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RosA</td>
<td>Sexual development—repressor</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>93</td>
</tr>
<tr>
<td>ScfA</td>
<td>Fatty acid metabolism</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>29</td>
</tr>
<tr>
<td>SfgA</td>
<td>Asexual development</td>
<td>Zn binuclear cluster</td>
<td>Not known</td>
<td>94</td>
</tr>
<tr>
<td>SreA</td>
<td>Iron uptake and utilization</td>
<td>C4 GATA zinc finger × 2</td>
<td>GATA</td>
<td>33</td>
</tr>
<tr>
<td>SteA</td>
<td>Sexual development</td>
<td>Homeodomain and C2H2 zinc finger</td>
<td>Not known</td>
<td>95</td>
</tr>
</tbody>
</table>

*continued*
### TABLE 8.1 (continued)  
**Known Regulatory Proteins in A. nidulans**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Function</th>
<th>Binding Domain</th>
<th>Recognition Motif</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>StuA</td>
<td>Sexual and asexual development</td>
<td>bHLH</td>
<td>WCGCGWNM</td>
<td>96</td>
</tr>
<tr>
<td>StzA</td>
<td>Abiotic stress</td>
<td>C2H2 zinc finger</td>
<td>Not known</td>
<td>97</td>
</tr>
<tr>
<td>UaY</td>
<td>Purine metabolism</td>
<td>Zn binuclear cluster</td>
<td>TCGGNCGGA</td>
<td>98</td>
</tr>
<tr>
<td>VeA</td>
<td>Light response, development and secondary metabolism</td>
<td>DNA directed RNA pol domain</td>
<td>Not known</td>
<td>99, 100</td>
</tr>
<tr>
<td>WetA</td>
<td>Conidial development</td>
<td>Novel motif</td>
<td>Not known</td>
<td>101</td>
</tr>
<tr>
<td>XlnR (A. niger)</td>
<td>Xylan degradation</td>
<td>Zn binuclear cluster</td>
<td>GGCTAAA</td>
<td>102</td>
</tr>
</tbody>
</table>

*Note:* DNA-binding regulatory proteins formally characterized in A. nidulans, their regulatory role, class of DNA-binding domain, recognition motif (where known) are listed. Two A. niger proteins have also been included. Ambiguous bases in the consensus sequences are given as R = A or G, Y = C or T, H = A, C or T, K = G or T, M = A or C, S = G or C, and W = A or T, N = any base.

The zinc binuclear cluster (PF00172) proteins are arguably the most interesting group with respect to their apparent expansion and divergence within the aspergilli. The distribution is significantly different when compared with the aspergilli to other fungi, and within the aspergilli there are striking differences in frequency, with a significantly higher number occurring in A. niger (Mortimer and Caddick, unpublished data). The zinc binuclear cluster motif is generally regarded as being fungal specific and is, therefore, likely to have evolved after the fungi diverged from other eukaryotes. Proteins bearing this domain are known to be involved in regulating a wide range of biological processes including primary and secondary metabolism, development, and drug resistance (Table 8.1), the latter probably reflecting altered expression of transporters. Generally they have been shown to be DNA-binding proteins, although this may not always be the case. One interesting exception is TamA, which like in the yeast ortholog, has a well-conserved zinc binuclear cluster domain. However, this domain is dispensable for known TamA function and there is no evidence that it binds DNA directly. TamA appears to operate as a coactivator. Of the zinc binuclear cluster proteins that bind DNA directly, examples include both regulatory proteins that activate and/or repress transcription, but in at least one yeast example the protein has a distinct nonregulatory role.

The divergence in frequency between species for the zinc binuclear cluster proteins suggest that utilization of this domain and rapid amplification and divergence of the respective genes have played a significant role in the evolution of fungi, but particularly the aspergilli. Compared with the 296 proteins containing PF00172 in A. niger, there are only 66 putative proteins bearing the next most prevalent DNA-binding domain, the C2H2 type Zinc finger (PF00096) (Table 8.2). A second domain that shows similar distribution is PF04082. This was initially identified on the basis of weak homology among zinc binuclear cluster proteins. In S. cerevisiae, this domain is only associated with zinc binuclear cluster proteins but in the aspergilli it has been observed in a number of additional proteins, and these do not appear to have known DNA-binding domains associated with them. It has been suggested that this domain assists in defining the specificity of DNA binding.

The predominance of the zinc binuclear cluster is true of the other fungi analyzed. This leads to the question, why has this domain been utilized so extensively? One notable attribute is the versatility of the domain, which in different protein functions as a dimer or monomer. It can bind inverted or direct repeats as well as nonrepeated elements in some instances. This flexibility is exemplified by two specific proteins; PrnA where the recognition motifs can vary significantly with respect to spacing and FacB where distinct sequence are bound. Perhaps the emergence of this DNA-binding element after many of the key cellular processes had evolved and were appropriately regulated resulted in the duplication and divergence of this domain having less cost with respect to fitness. It will be interesting if this paradigm, that newly evolved regulatory domains are more amenable to rapid evolution, is found to be the case...
<table>
<thead>
<tr>
<th>Accession Name</th>
<th>A. nidulans</th>
<th>A. niger</th>
<th>A. fumigatus</th>
<th>A. oryzae</th>
<th>N. crassa</th>
<th>S. cerevisiae</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF0172+04082</td>
<td>Fungal specif. transcription factor domain</td>
<td>118</td>
<td>146</td>
<td>83</td>
<td>86</td>
<td>44</td>
</tr>
<tr>
<td>PF00308</td>
<td>Fungal Zn(2)-Cys(6) binuclear cluster domain</td>
<td>179</td>
<td>226</td>
<td>124</td>
<td>182</td>
<td>61</td>
</tr>
<tr>
<td>PF00096</td>
<td>Zinc finger, C2H2 type</td>
<td>253</td>
<td>296</td>
<td>188</td>
<td>176</td>
<td>89</td>
</tr>
<tr>
<td>PF00642</td>
<td>Zinc finger C-x8-C-x5-C-x3-H type (and similar)</td>
<td>69</td>
<td>66</td>
<td>52</td>
<td>42</td>
<td>54</td>
</tr>
<tr>
<td>PF00170</td>
<td>bZIP transcription factor</td>
<td>13</td>
<td>14</td>
<td>12</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>PF00249</td>
<td>Myb-like DNA-binding domain</td>
<td>24</td>
<td>13</td>
<td>16</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>PF00098</td>
<td>Zinc knuckle</td>
<td>25</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>PF00533</td>
<td>BRCA1 C Terminus (BRCT) domain</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>PF02178</td>
<td>AT hook motif</td>
<td>12</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>PF00010</td>
<td>Helix-loop-helix DNA-binding domain</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>PF01336</td>
<td>OB-fold nucleic acid binding domain</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>PF02037</td>
<td>SAP domain</td>
<td>8</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>PF00046</td>
<td>Homeobox domain</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>PF00498</td>
<td>FHA domain</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>PF05225</td>
<td>helix-turn-helix, Psq domain</td>
<td>19</td>
<td>8</td>
<td>30</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PF00320</td>
<td>GATA zinc finger</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>PF00439</td>
<td>Bromodomain</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>PF00505</td>
<td>HMG (high-mobility group) box</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>PF02463</td>
<td>RecF/RecNSMC N terminal domain</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>PF00488</td>
<td>Mus domain V</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>PF00493</td>
<td>MCM2/3/5 family</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>PF01529</td>
<td>DHHC zinc finger domain</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>PF00250</td>
<td>Fork head domain</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>PF01753</td>
<td>MYND finger</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>PF02146</td>
<td>transcriptional regulator, Sir2 family</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>PF00136</td>
<td>DNA polymerase family B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PF00730</td>
<td>HhH-GPD superfamily base excision DNA repair protein</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>PF01096</td>
<td>Transcription factor S-II (THIS)</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PF01624</td>
<td>Mus domain I</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PF01920</td>
<td>KE2 family protein</td>
<td>20</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>PF02292</td>
<td>APSIES domain</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Accession</td>
<td>Name</td>
<td>A. nidulans</td>
<td>A. niger</td>
<td>A. fumigatus</td>
<td>A. oryzae</td>
<td>N. crassa</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>PF02854</td>
<td>MIF4G domain</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PF00447</td>
<td>HSF-type DNA-binding domain</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PF03178</td>
<td>CPSF A subunit region</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PF04675</td>
<td>DNA ligase N-terminus</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>PF05190</td>
<td>MatS family domain IV</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PF05224</td>
<td>NDT80/PhoG like DNA-binding family</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>PF00319</td>
<td>MADS box/SRF-type transcription factor</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PF00565</td>
<td>Staphylococcal nuclease homolog</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PF00633</td>
<td>Helix-hairpin-helix motif</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PF01388</td>
<td>ARID/BRIGHT DNA binding domain</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>PF01422</td>
<td>NF-X1 type zinc finger</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF01424</td>
<td>R3H domain</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PF01997</td>
<td>Translin family</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PF02735</td>
<td>Ku70/Ku80 beta-barrel domain</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PF02891</td>
<td>MIZ zinc finger</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PF07529</td>
<td>HAS</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PF08265</td>
<td>YL1 nuclear protein C-terminal domain</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PF00165</td>
<td>transcriptional regulator, araC family</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>PF00352</td>
<td>Transcription factor THID (or TATA-binding protein, TBP)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF00436</td>
<td>Single-strand binding protein family</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF00649</td>
<td>Copper fist DNA-binding domain</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PF01035</td>
<td>6-O-methylguanine DNA methyltransferase, DNA binding domain</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PF01285</td>
<td>TEA domain</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PF01381</td>
<td>Helix-turn-helix</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF01426</td>
<td>BAH domain</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>PF01833</td>
<td>IPT/TIG domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF01984</td>
<td>Double-stranded DNA-binding domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PF02045</td>
<td>CCA AT-binding transcription factor (CBF-B/NF-YA) subunit B</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF02257</td>
<td>RFX DNA-binding domain</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF02671</td>
<td>Paired amphipathic helix repeat</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>PF02755</td>
<td>RPEL repeat</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Domain ID</td>
<td>Domain Description</td>
<td>Proteins</td>
<td>DNA Binding</td>
<td>Topoisomerase</td>
<td>RNA Polymerase</td>
<td>Transcriptional Initiation</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>---------------</td>
<td>----------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>PF02765</td>
<td>Telomere-binding protein alpha subunit, central domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF02791</td>
<td>DDT domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF02805</td>
<td>Metal-binding domain of Ada</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>PF02919</td>
<td>Eukaryotic DNA topoisomerase I, DNA-binding fragment</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF03126</td>
<td>Plus-3 domain</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF03871</td>
<td>RNA polymerase Rpb5, N-terminal domain</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PF04152</td>
<td>Mre11 DNA-binding presumed domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF04406</td>
<td>Type IIB DNA topoisomerase</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF04769</td>
<td>Mating-type protein MAT alpha 1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF04921</td>
<td>XAP5 protein</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF05181</td>
<td>XPA protein C-terminus</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF05764</td>
<td>YLI nuclear protein</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF06420</td>
<td>Mitochondrial genome maintenance MGM101</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF06831</td>
<td>Formamidopyrimidine-DNA glycosylase H2TH domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF08221</td>
<td>RNA polymerase III subunit RPC82 helix-turn-helix domain</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF00196</td>
<td>Bacterial regulatory proteins, luxR family</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF01498</td>
<td>Transposase</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF02186</td>
<td>TFIIIE beta subunit core domain</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF02229</td>
<td>Transcriptional Coactivator p15 (PC4)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF02892</td>
<td>BED zinc finger</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF04090</td>
<td>RNA polymerase I specific initiation factor</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF04684</td>
<td>BAF1/ABF1 chromatin reorganizing factor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PF06839</td>
<td>GRF zinc finger</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>PF07453</td>
<td>NUMOD1 domain</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Listed are the PFAM domains putatively involved in DNA binding that have been identified in at least one of the six species. The score given is the number of proteins found bearing at least one of these domains. In some cases a protein may have multiple domains. The only domain not specifically associated with DNA binding is a fungal specific transcription factor domain (PF04082) often found in association with PF00172. As such it represents a second independent assessment of this type of transcription factor. The frequency for these two domains occurring together in the same protein is also given. These data are based on the annotations available. From initial analysis it is obvious that miss-annotation is a problem and the numbers of domains identified are likely to be underestimates in the Aspergillus species.
within other genera. Finally, as can be seen in Table 8.1 many of these proteins are involved in regulating metabolism, being responsible for monitoring the presence of specific metabolites and regulating small groups of genes appropriately. The metabolic versatility of fungi, and in particular of the *Aspergillus* species, is well documented. From genome analysis, it is apparent that a large proportion of genes are involved in the uptake and metabolism of a diverse range of compounds. Appropriate expression of these genes is fundamental to fitness, and it appears that the zinc binuclear cluster proteins are ideally suited to evolving toward regulating relatively small groups of genes.

Of the remaining PFAM domains identified, three that showed very aberrant distribution across the six species (PF05225, PF00196, PF01498) are likely to be associated with transposons, which possibly explains their anomalous distribution.

Potentially these proteins represent a significant biotechnological resource, with respect to the development of biosensors and novel gene switches. Already one such protein, AlcR of *A. nidulans*, has been successfully utilized as the basis of a gene switch in plant systems.\(^\text{19}\) In this and a number of other characterized examples the protein is directly responsible for monitoring the level of specific effector molecules within the cell, the presence of which leads to altered gene expression at specific promoters. Considering the wide variety of substrates being monitored by these metabolically versatile organisms, the characterization and subsequent exploitation of these regulatory proteins should be a key goal of the postgenomic analysis of these organisms.

### 8.4 Regulatory Signals

The two key areas in bioinformatics and genome analysis are the identification of genes and reliable prediction of their function. By definition the genome sequence has all the information required for appropriate gene regulation, the problem for the biologist is identifying and interpreting the key features. A major source of underutilized information lies in the promoters and other regulatory elements, as they define which parts of the genome are transcribed, the level of expression, and its regulation. Consequently, interpreting the “transcriptional regulatory code”\(^\text{20}\) represents a major current challenge, as it is fundamental to understanding the genome, essential to precise manipulation and prediction of gene expression and provides a valuable source of information relating to gene function.

The available genomic resources from multiple-sequenced *Aspergillus* species provide us with valuable tools in the analysis of gene expression. A primary goal should be to identify key regulatory elements within the genome allowing the prediction of gene expression. This will include the identification of specific DNA motifs bound by transcription factors but additionally the characterization of the various additional features within the sequence that influence their function. The ability to interpret this information will provide new insights into the integration of different biological processes, providing information about the probable role of novel genes, clarifying the likely role of genes from specific classes that are difficult to predict on the basis of homology (e.g., transporters), providing valuable information for gene annotation in the aspergilli. Underpinning all these objectives is the involvement of bioinformaticians, in order to develop and optimize computational algorithms and applying them to the identification of significant motifs and promoter signatures.

Analysis of regulatory motifs across fungal species will also illuminate the evolutionary history of specific regulatory processes and associated factors, which played an important role in the divergence in this important group of organisms. Processes established before species divergence are likely to retain regulatory components: in particular the DNA-binding domains of regulatory proteins and their cognate sequence motifs are likely to be conserved. Examples include the GATA factors regulating nitrogen metabolism\(^\text{3}\) and the PacC orthologs regulating the response to ambient pH\(^\text{2}\), which appear to be conserved across all ascomycetes examined so far and extend into other fungal phyla. Newly evolved regulatory systems will be specific to a subgroup of species, in some cases defining novel functions that have recently evolved. Finally, specific regulatory systems will have been lost along with the biological features they regulate, as is the case for a number of regulatory genes that, for example, have orthologs across the ascomycetes but which are not present in *Saccharomyces cerevisiae*.\(^\text{21}\) Thus, the phylogenetic
footprint of a given regulatory motif will reflect its origin and subsequent adoption, modification, and loss. However, there needs to be a note of caution as apparent similarity may not in fact reflect actual conservation of the mechanism of function. A good example of this is PrnA, which in *A. nidulans* is responsible for regulating the gene cluster involved in proline metabolism. This is a zinc binuclear cluster protein that binds to the recognition motif specifically on induction by proline, leading to activation of the respective genes.\(^{17}\) The ortholog in *S. cerevisiae*, Put3p, acts very differently, being bound permanently to the promoter\(^{22}\) where it is activated by phosphorylation in response to the inducer.\(^{23}\)

### 8.5 Identification of Regulatory Motifs in Silico

There are well-established *in vitro* and *in vivo* methods to determine whether a specific transcription factor binds to a given DNA sequence. These different approaches have been applied to a range of *Aspergillus* transcription factors, leading to the identification of specific motifs (Table 8.1). The diversity of approaches makes any careful comparison of motifs difficult and their reliability uncertain. In many cases, specific motifs have been identified within the promoters of the one or two known target genes. Subsequently, these have been tested using *in vitro* DNA-binding assays. Inevitably this will not provide a robust and reliable consensus, as additional work is needed to test relative binding affinities, identify alternative binding motifs and whether the specific elements are functional *in vivo*. There are various examples where motifs appear not to be functionally significant when analyzed *in vivo*.\(^{24-26}\) The development of chromatin immunoprecipitation assays (ChIP)\(^{27}\) has recently been successfully achieved for *A. nidulans* (Joseph Strauss, personal communication) and if this could be combined with the development of intergenic arrays to undertake ChIP on CHIP experiments\(^{28}\) an invaluable global picture of transcription factor function would be possible in the aspergilli. However, conventional analysis of specific transcription factors can combine very well with genome data in the assessment of putative function, the identification of genes subject to their regulation, and establishment of robust consensus sequences. One recent example is the analysis of FarA and FarB, which are zinc binuclear cluster proteins involved in fatty acid metabolism.\(^{29}\) Genome-wide analysis of the distribution of the identified consensus binding motif was undertaken and it was found to be associated with most genes predicted to be involved in fatty acid metabolism. Furthermore, a similar association between the motif and specific structural genes extended through a range of species. The same analysis revealed that the FacB recognition motif had an overlapping but distinct distribution, being present specifically upstream of genes required for acetate utilization.

Effective computational approaches to predict potential binding sites and gene expression are highly desirable. Taking all the sequences upstream and downstream from the predicted genes, and comparing their distribution in *A. nidulans*, *A. fumigatus*, and *A. oryzae*, lead to the identification of conserved sequence elements adjacent to orthologous genes.\(^{30}\) In certain cases, these short elements are related to motifs recognized by known DNA- or RNA-binding proteins and/or the associated genes had related function or cellular location. For example, the CpcA/Gcn4p element was identified as a conserved sequence upstream of a number of genes involved in amino acid transport and metabolism, consistent with the known function of CpcA.\(^{31}\)

This type of approach does, however, have limitations. For example, an element related to the GATA sequence, which is known to be bound by the AreA transcription factor, involved in nitrogen regulation was found associated with a number of genes but there was no apparent common functional role for the associated genes. Prior to this work it was known that there are multiple GATA factors in the aspergilli, and although those tested bind very similar sequences *in vitro*, they have radically different functions including the regulation of nitrogen metabolism,\(^{32}\) iron sequestration,\(^{33}\) sexual development,\(^{34}\) light and circadian responses (H. Haas, personal communication). It is, therefore, important to be aware that these motifs are context dependent and regulatory proteins can differentiate between them *in vivo* if not *in vitro*. Additionally, it can be seen from Table 8.1 that a number of regulatory proteins bind the same or related sequences. This is a fundamental aspect of gene regulation, where different signals are coordinated via specific regulatory elements; the competition for or cooperative binding at a given sequence resulting in the appropriate regulatory response being achieved.\(^{25,35}\)
We undertook analysis of upstream regulatory elements from specific groups of genes, comparing these sequences from orthologous genes in *A. nidulans, A. fumigatus*, and *A. oryzae* (Dobson and Caddick, unpublished data). This analysis led to the following interesting conclusions:

1. The sequences immediately upstream of genes, putatively including the promoter and other regulatory elements, are not well conserved between the three species. The number, position, and orientation of known regulatory elements are also poorly conserved. Thus, the availability of closely related genomes in future analysis will be important, and this should be a factor in choosing which additional genomes are to be sequenced.

2. Key regulatory elements can be identified by examining promoters from orthologous genes. Using the publicly available software package MEME (http://meme.sdsc.edu/meme/meme.html), we conducted pattern searches using promoter sequences, such as the well-characterized intergenic regions of *niaD–niiA* and *prnD–PrnB*. This led to the identification of conserved sequence motifs close to those known to bind the pathway-specific regulators NirA and PrnA, respectively. Thus, it should be similarly possible to identify putative regulatory elements in genes of unknown function, laying the foundations for experimental exploration.

3. Computational analysis often extends the known consensus sequences. For example, NirA-binding sites, which have been defined experimentally as CTCCHG\textsuperscript{36} were identified by MEME as having the consensus WWYTCCKHG\textsuperscript{37}GV. This sequence occurs less frequently in the *A. fumigatus* genome than in the previously defined sequence but at a higher frequency in the regions upstream of genes known to be regulated by NirA. This would suggest that this sequence is a better consensus for functional NirA-binding sites, which represents a testable hypothesis.

4. Analysis of promoters of orthologous genes was unsuccessful for the shorter motif (HGATAR) bound by the global regulator AreA. However, taking promoter elements from 14 *A. nidulans* genes subject to AreA regulation and their orthologs in *A. fumigatus* and *A. oryzae*, the AreA recognition sequence was identified and again extended and refined (WGATAA\textsuperscript{37}GR).

5. Motifs of related transcription factors are distinguishable. AreA and SreA are two GATA-class transcription factors with similar affinity for DNA *in vitro* but with distinct functions; AreA is involved in regulating nitrogen metabolism\textsuperscript{38} while SreA is involved in regulating iron uptake and utilization.\textsuperscript{39} From analysis of upstream sequences of iron sequestration genes—an acyltransferase/hydrolase, a ferrochrome peptide synthetase, and a siderophore transporter—a GATA-like motif for SreA binding with consensus sequence ATCWGATWAGAT was derived. It contains a flanking GAT inverted repeat (underlined) as well as the internal GATA-like element (bold).

6. The presence of conserved motifs is partly predictive of regulation. We searched the genome of *A. fumigatus* for all potential NirA\textsuperscript{36} and PrnA\textsuperscript{37} binding sites upstream of genes and then examined the orthologs in *A. nidulans* and *A. oryzae* to determine whether these sequences were conserved. Based on the refined NirA-binding site defined by comparative analysis, no orthologous group had more than one putative binding site within 1 kb of the start codon except for the four genes known to be subject to NirA regulation. However, expression of five genes in *A. nidulans* that had at least one putative NirA binding site, were monitored. One was not expressed under any conditions and two showed significant induction by nitrate, which was NirA dependent. With respect to PrnA, eight orthologous groups were identified. Of these one gene *gdhB*, which encodes a NAD-linked glutamate dehydrogenase, could logically be expected to be regulated by proline (the role of PrnA), as proline is metabolized to glutamate, which is the substrate for the *gdhB* product. Real-time PCR analysis of *gdhB* expression successfully revealed *prnA* regulation of *gdhB*.

7. The presence of regulatory elements can be indicative of broad functional class. In *A. nidulans*, *prnB* encodes the proline permease\textsuperscript{39} and is part of a gene cluster (*prnA, X, D, B, and C*),\textsuperscript{40} which is involved in proline utilization. In both *A. fumigatus* and *A. oryzae*, the *prnB* ortholog
BLAST identifies three prnB homologs in *A. fumigatus* and seven in *A. oryzae*. However, in both species only one of the respective genes contains PrnA-binding motifs within the promoter, suggesting that these are involved in proline metabolism. Consistent with the hypothesis, these two genes are the most closely related to prnB. The function of the putative proline permeases in *A. fumigatus* and *A. oryzae* will be tested as part of this work.

From these early attempts to utilize the available *Aspergillus* genome sequences it is clear that important observations can be made and testable hypotheses derived. A key goal of research in this area should, therefore, be to optimize this type of analysis by making best use of the range of genome sequences that are now available.

### 8.6 Transcriptomics and Proteomics

Transcriptomics provides a powerful source of data allowing us to cluster genes on the basis of expression profiles, linking them to specific growth conditions or mutant backgrounds. There have been a variety of arrays produced for different *Aspergillus* species, ranging from partial cDNA arrays to Affymetrix-like arrays. Recently, 70mer oligo arrays have been produced for both *A. fumigatus* and *A. nidulans*. These are freely available through PFGR (http://pfgrc.tigr.org/slide_html/microarray_descriptions.shtml). This fantastic resource will hopefully be effectively used by the community. Up to now the number of publications featuring data from the genome-wide arrays is very limited. These have included analysis of the response to temperature and the antifungal voriconazole in *A. fumigatus*, the response of *A. nidulans* to changes in carbon source, and the role of the regulatory genes *creA* and the role of *afl R* in *A. paraciticus*.

Grouping genes on the basis of common expression profiles of itself does not tell us how genes are being regulated. Such experimental approaches also have significant limitations; indirect effects of the growth conditions or mutant phenotype may lead to the parallel regulation of genes while the complexity of cellular regulatory networks can result in similar responses being achieved using functionally distinct mechanisms. Additionally, often only a subset of the genes regulated by a specific transcription factor will be identified due to repression by a second factor or a specific requirement for an additional signal. Finally, resources limit the number of experiments that can be undertaken and the organisms that can be examined in this way. Interpreting transcriptomics data should, therefore, be accompanied by detailed analysis of the noncoding sequences associated with the genes leading to the identification of motifs and other features that define the expression profile. The identification of putative regulatory elements will provide useful data that can be mapped back onto the genomes of related organisms, providing clues as to the likely expression profiles of associated genes.

Faced with the plethora of novel genes and large gene families, an indication of function and role is very quickly gained by observing gene expression. With the availability of transcriptomics we are able to monitor global gene expression profiles and look for clusters of functionally related genes. The value of these data will increase dramatically with the ability to access with their accumulated data. This is now possible for a variety of organisms, including *Arabidopsis thaliana* and *S. cerevisiae*. It will be important for our very diverse community, working on a range of related species, that this data is easily accessible in the future, and maintained in well-curated databases.

Proteomics is also now being utilized to monitor gene expression but inevitably this is not truly global analysis. Generally, a subproteome is chosen (e.g., secreted, soluble intracellular, membrane or mitochondrial associated proteins) and the analysis is further restricted by levels of expression required for identification, and properties of each specific protein such as solubility, posttranslational modification, and so on, which will determine the ability to both isolate and subsequently characterize the protein using mass-spectrometry. Combining both proteomic and transcriptomics analysis to investigate the regulatory role of AreA in nitrogen metabolisms (Morozov, Jones, and Caddick, unpublished data), it has proved possible to identify genes/proteins that show distinct differential regulation. However, there is very little correlation between the two, with only a minority of the differentially regulated proteins identified being similarly regulated at the transcriptional level. This is not surprising in the light of observations...
in other organisms but it provides a cautionary note. Transcriptional regulation is not the only significant regulatory mechanism and up- or down-regulation at the transcriptional level may be countered at a later stage by additional regulatory steps in the gene expression pathway.

### 8.7 Additional Regulatory Mechanisms

In this review we have focused on regulation at the initiation of transcription. However, both conventional analysis and genomics reveal a range of mechanisms that are pertinent to regulation. As in higher eukaryotes, there are likely to be a significant number of examples of the use of different promoters, terminators, and differential splicing. These complexities have been identified for a number of regulatory genes and can result in distinct proteins being formed from a specific gene. The availability of EST sequences provides valuable information that has not yet been utilized to assess the proportion of genes which produce distinct transcripts in *Aspergillus*. However, a significant level of differential splicing and antisense transcripts has been observed in *Cryptococcus neoformans*. A. *nidulans* would appear to be a very good microbial system to investigate how differential splicing is regulated, *S. cerevisiae* being very limited in this respect.

Differential splicing and the use of alternate promoters can produce products that are subject to translational regulation. Based on genome analysis, short upstream open reading frames appear to be a common regulatory feature, with approximately 21% of the genes being subject to this type of regulation. Another relevant feature is the apparent use of noncanonical start codons GUG and CUG in place of AUG. The frequency of this across the genome is unknown and the biological consequences poorly understood but this is likely to have direct consequences on the levels of translation.

Regulation of mRNA decay also plays a fundamental role in the control of gene expression. In particular, the rate of mRNA decay limits how quickly the cell can respond to specific stimuli. Consequently, modulation of mRNA turnover is an important mechanism for achieving rapid responses to regulatory signals. Decay rates vary significantly between transcripts and the stability of individual transcripts can also vary significantly in response to specific stimuli. Comparative analysis of the *Aspergillus* genomes revealed various elements enriched in the 3′ UTR. The most prevalent amongst these were the Puf-binding element. There are predicted to be five RNA-binding Puf proteins in *A. nidulans*. Based on the distribution of these elements, and consistent with the function of at least one of these in *S. cerevisiae*, an *Aspergillus* Puf protein is probably involved in posttranscriptional regulation of genes involved in mitochondrial function. We have recently disrupted four of the *puf* genes and all disrupted strains were found to be viable with only minor morphological defects (Chooluck, Morozov, and Caddick, unpublished data). However, in one case the strain appears to be sterile, suggesting that these proteins have distinct and important roles.

One mechanism, which regulates nitrogen metabolism in *A. nidulans*, involves regulated mRNA decay. Transcripts for a range of genes involved in nitrogen metabolism, including the key regulatory gene *areA*, are subject to rapid degradation in response to high intracellular Gln levels. In direct contrast to this, nitrate specifically stabilizes certain transcripts involved in nitrate metabolism (*niaD* and *niiA*), but not the nitrate transporter structural genes (*crmA* and *nrtB*). Nitrate stabilization is predominant to Gln destabilization. This divergent regulation makes good physiological sense, as it is important that nitrate and nitrite reductase are retained by the cell until the potentially toxic metabolites (nitrate and nitrite) are removed. It will be intriguing to know how prevalent such sophisticated regulatory systems are and in particular if regulated stabilization of specific transcripts is a general mechanism utilized to protect the cell from the effects of toxic intermediates of biochemical pathways.

In these examples regulation of transcript stability occurs by promoting or inhibiting deadenylation of specific transcripts. From the genome sequence and functional analysis it has been shown that key components of this and other RNA degradation pathways are generally well conserved in the aspergilli. Amongst these are the components of the nonsense-mediated decay (NMD) pathway, which is responsible for eliminating defective transcripts prior to translation but which also has a regulatory role.
for some genes, and the RNA-silencing system.\textsuperscript{59} A functional RNA-silencing mechanism combined with the presence of antisense transcripts in fungi\textsuperscript{53} opens up another potentially major regulatory mechanism that has as yet not been characterized. Up to now there have been no reports of microRNAs in \textit{Aspergillus} species, and they would appear to be rare in fungi, but again this will be an intriguing possibility, which with the available resources is now open to investigation.

The postgenomic era is an exciting one for studies on gene expression and regulation. \textit{Aspergillus} species offer a wide range of valuable resources that can be readily utilized. Most importantly, fundamental biological questions remain to be addressed. A key aspect will be the effective utilization of the genomic data, developing systems to facilitate comparative genome analysis, and accessing data from transcriptomic and proteomic studies. The central role of gene regulation and expression makes its analysis pertinent to all aspects of biology. Genomics offers a global view, allowing an integrative approach. Consequently, it will be important to build on the resources we have in cell biology, physiology, genetics, and so on with the aim of developing a clearer understanding of key biological systems and how they are regulated. \textit{Aspergillus} has made a significant contribution to our understanding and we are well equipped to continue.

\begin{thebibliography}{99}
\end{thebibliography}
90. Levesley, I., et al., Domain structure and function within the QUTA protein of Aspergillus nidulans: Implications for the control of transcription. Microbiology, 142, 87, 1996.


Mitogen-Activated Protein Kinase Pathways in Aspergilli

Gregory S. May

CONTENTS
9.1 General Introduction to Fungal MAP Kinase Signaling in Filamentous Fungi ...................... 121
  9.1.1 MAP Kinase Signal Transduction Pathways in the Aspergilli ........................................... 122
  9.1.2 Cell Integrity Pathway and Functions of the mpkA Gene .................................................. 123
  9.1.3 MpkB Pathway and Mating ................................................................................................ 123
  9.1.4 MpkC Pathway and Carbon Utilization ............................................................................. 124
9.2 SakA/HogA Signaling Pathway ................................................................................................. 124
9.3 Where to Next? ......................................................................................................................... 126
References .................................................................................................................................... 126

Mitogen-activated protein kinases (MAPKs) were first described as protein kinases that are activated in vertebrate cells in response to growth factors [1,2]. We now know that MAPKs are found in all eukaryotes and that these protein kinases regulate a variety of cellular processes. MAPKs are the terminal protein kinase in a protein kinase cascade. At the top of the protein kinase cascade is a MAP kinase kinase kinase (MAPKKK) that phosphorylates a MAP kinase kinase (MAPKK) that then phosphorylates the MAPK leading to activation of the protein kinase. The biochemical events leading to MAPK activation are well understood [3–5]. The MAPKKK phosphorylates serine and threonine residues in the amino terminal region of the MAPKK resulting in activation of this protein kinase. The now active MAPKK in turn phosphorylates the MAPK on threonine and tyrosine residues separated by one amino acid that is in the activation loop of the conserved kinase domain, leading to an activated MAPK. Upstream of the MAPKKK are activating protein kinases that are different for various signal transduction pathways. These include p21 activated protein kinases and protein kinase C. Many of these signal transduction pathways are activated by heterotrimeric G-proteins whose activation is frequently coupled to cell surface receptors. The most extensively studied of these signal transduction pathways in fungi is the pheromone response pathway in the budding yeast *Saccharomyces cerevisiae* [4,6]. Substrates of MAPKs are transcription factors that change the gene expression pattern of the target cell in response to the stimulus. In vertebrate cells the stimulus is frequently a growth factor that stimulates the cell to proliferate, and disruption of this MAPK pathway has been linked to cancer [2]. The goal of this chapter is to review the work on MAPK pathways in the aspergilli and summarize the findings as they may relate overall to the role of these protein kinases in filamentous fungal cell biology.

9.1 General Introduction to Fungal MAP Kinase Signaling in Filamentous Fungi

Studies of MAPK gene functions in fungi have been conducted in a myriad of species and from these studies some universal conclusions can be drawn. MAPK signal transduction pathways respond to a
variety of environmental signals. The environmental signals include hypertonic and oxidative stresses, nutrient sensing of nitrogen and carbon sources, and mating [4]. Interestingly, it has also been shown, for plant and animal pathogenic fungi, that some of the MAPK signaling events are essential for pathogenesis. In the animal pathogen *Candida albicans* the transition from budding yeast growth to hyphal growth is regulated by MAPK signaling and this change in growth mode is directly linked to fungal pathogenesis [7–9]. Mutations in any number of the genes of the MAPK signaling pathway that control this dimorphic transition from yeast to the filamentous growth affect pathogenesis. In some plant pathogenic fungi, the MAPK signal transduction pathway that regulates the hypertonic stress response is required for virulence on the plant host [10]. Similarly, MAPKs that control the cell integrity in some plant pathogens is required for pathogenesis [10].

The yeast *S. cerevisiae* and *Schizosaccharomyces pombe* are excellent genetic systems that have helped to define the evolutionarily conserved MAPK signaling pathways of fungi and the key biochemical events of these pathways [4,6]. While there is considerable conservation of these MAPK signal transduction pathways among the yeast and filamentous fungi, there are significant differences that reflect the variable lifestyles of this diverse group of organisms [11,12]. Because of this diversity it is important to also investigate how MAPK signaling pathways contribute to filamentous fungal growth and biology. This is where studies of MAPK pathways in the aspergilli may contribute essential information on the role of these signal transduction pathways in filamentous fungal cell biology. Since the aspergilli include species that are human pathogens, colonizers of plants, and a model genetic species, there is the opportunity to explore considerable areas of biology that is not represented in the model yeast systems.

### 9.1.1 MAP Kinase Signal Transduction Pathways in the Aspergilli

There are four MAPK genes in the genomes of *Aspergillus nidulans*, *Aspergillus fumigatus*, and *Aspergillus oryzae* that are highly conserved across these species (Table 9.1). These genes are *mpkA*, *mpkB*, *mpkC*, and *sakA/hogA*. Interestingly, there are three genes that code for MAPKK polypeptides and three that code for MAPKK proteins. The three MAPKKs are Ste7 like, Pbs2 like, and Mkk2 like, suggesting they have possible roles in mating (Ste7), osmotic regulation (Pbs2), and cell wall integrity (Mkk2), respectively. Similarly, the MAPKKKs are orthologous to SteC/Ste11, Bck1, and Ssk2, and thus would appear to function in mating, cell-wall integrity, and osmotic regulation, respectively. In addition

<table>
<thead>
<tr>
<th>MAPKK</th>
<th>Bck1 like</th>
<th>SteC/Ste11 like</th>
<th>SteB/Ssk2 like</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN4887.3</td>
<td>AN2269.3</td>
<td>AN10153.3</td>
<td></td>
</tr>
<tr>
<td>AFu3g11080</td>
<td>Afu5g06420</td>
<td>Afu1g10940</td>
<td></td>
</tr>
<tr>
<td>AO090003000662</td>
<td>AO090009000610</td>
<td>AO090038000313</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MAPK</th>
<th>Mkk2 like</th>
<th>Ste7 like</th>
<th>Pbs2 like</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN4189.3</td>
<td>AN3422.3</td>
<td>AN0931.3</td>
<td></td>
</tr>
<tr>
<td>AFu1g05800</td>
<td>Afu3g05900</td>
<td>Afu1g15950</td>
<td></td>
</tr>
<tr>
<td>AO090009000347</td>
<td>AO090002000060</td>
<td>AO090005001093</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MAPK</th>
<th>MpkA</th>
<th>MpkC</th>
<th>SakA/HogA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN5666.3</td>
<td>AN3719.3</td>
<td>AN4668.3</td>
<td>AN1017.3</td>
</tr>
<tr>
<td>AFu4g13720</td>
<td>Afu6g12820</td>
<td>Afu5g09100</td>
<td>Afu1g12940</td>
</tr>
<tr>
<td>AO090009000199</td>
<td>AO090003000402</td>
<td>AO0900020000466</td>
<td>AO0900701000642</td>
</tr>
</tbody>
</table>

Numbers preceded by AN are for *A. nidulans* and were obtained from the Broad Institute (http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/Home.html).
Numbers preceded by AF are for *A. fumigatus* and were obtained from TIGR (http://www.tigr.org/tdb/e2k1/afu1/).
Numbers preceded by AO are for *A. oryzae* and were obtained at NITE (http://www.bio.nite.go.jp/dogan/MicroTop?GENOME_ID=ao).
to these core MAPK cascade proteins, additional upstream regulators of the MAPK cascades can also be found in the genomes of these species [13,14]. Thus it would appear that the primary MAPK functions in fungi for cell-wall integrity, mating, and osmotic regulation are conserved. It is interesting to note that for A. oryzae and A. fumigatus no sexual cycle has been described; yet they contain all the required genes for a sexual cycle including the MAPK, mpkB [15-18].

9.1.2 Cell Integrity Pathway and Functions of the mpkA Gene
The mpkA gene and its functions have been only investigated in A. nidulans, though the orthologous gene is present in the other two species [19]. The polypeptide coded for by mpkA is most similar to the MAPK Slt2p in S. cerevisiae and related protein kinases in other fungi. These MAPKs have a primary role in regulating cellular responses to loss of cell-wall integrity. In both plant and human pathogenic fungi the orthologous MAPK gene have been shown to contribute to fungal virulence [7,9,10]. Deletion of mpkA in A. nidulans results in a viable mutant that displays a variety of defects that is consistent with this MAPK functioning in the cell-wall integrity pathway. mpkA deletion mutants were identified as having aberrant colony formation and optimal growth of the deletion mutant was seen only on complex high-osmolarity media. Even on optimal growth medium the hyphal tips of mpkA mutants swell, suggesting that either there is a defect in cell-wall synthesis or an increase in turgor pressure resulting in tip swelling. In addition in the mpkA mutants, hyphal branching was abnormal and included branching at the hyphal tip. Furthermore, during conidial germination the mpkA mutants are unable to form germtubes, instead forming large spherical multinucleate cells that eventually lyse. Many of the defects observed are consistent with a defect in cell-wall formation.

In a study conducted in A. oryzae, it was found that deletion of the kexB gene resulted in a defect in the cell integrity-signaling pathway in this species [20]. The kexB deletion strain formed only small abnormal colonies that did not differentiate to form conidiophores or conidia. In addition, the kexB mutant strain formed a hyper branching mycelium that also branched at the hyphal tips. Just as in the case of the mpkA deletion mutant, the phenotypes of the kexB mutant phenotypes were remedied on high-osmolarity media containing salt or sorbitol. These authors hypothesized that the kexB mutant was activating the cell integrity-signaling pathway. By Northern analysis they found that mpkA messenger RNA levels were elevated in the kexB deletion mutant consistent with activation of the MpkA pathway. More interestingly, they determined that mpkA messenger RNA levels were reduced in 0.8 M NaCl again consistent with the kexB deletion mutation activating the pathway. Stronger evidence that the deletion of kexB activates the MpkA pathway was provided by western blot analysis in which the authors demonstrated that MpkA is constitutively phosphorylated in the kexB mutant. The authors conclude that deletion of kexB leads to an activation of the cell integrity-signaling pathway probably due to loss of secreted proteins that are needed for cell-wall formation and restructuring during normal hyphal growth and colony development.

Overall, the cell integrity-signaling pathway is conserved among the aspergilli and shares many of the same features identified in the yeast system. There is a readily identifiable ortholog of the MAPKKK Bck1 in each of the species and a Mkk2 MAPKK ortholog. These similarities coupled with the phenotype of the mpkA deletion mutant further suggest a conservation of many aspects of the cell integrity pathway. Additional studies will be needed to determine how this pathway functions in the aspergilli and where it may be similar or different from the pathway in yeast.

9.1.3 MpkB Pathway and Mating
The least is known about the MpkB pathway among all the MAPK pathways in the aspergilli. There are no reports of mpkB deletion mutants having been made, so it is difficult to provide specific phenotypes for deletion of this gene. In contrast, there have been experiments conducted with the MAPKKK gene steC that should be informative about this pathway [21]. The predicted SteC protein sequence contains not only the catalytic domain of the kinase but also contains the sterile alpha module domain, a protein–protein interaction domain. Deletion of steC resulted in strains with reduced growth and formed a brown-pigmented mycelium. In addition to these phenotypes, the steC mutant produced conidiophores of greater variability in length and 1–2% of the conidia formed were abnormally large. Finally, the steC deletion
The Aspergilli

mutant failed to form cleistothecia and crosses homozygous for the steC deletion failed to form cleistothecia. Interestingly, the levels of the steC messenger RNA were developmentally regulated. The levels of steC messenger RNA increase during conidiophore formation, and decrease to levels found in vegetative hyphae during sexual development. A translational fusion of green fluorescent protein (GFP) at amino acid residue 551 of SteC was used to assess SteC location during development. SteC-GFP was easily detected in the metulae and phialides of young conidiophores but was absent from mature conidiophores in the center of the colony. Thus, it would appear that expression of steC is transient and spatially restricted. In contrast, expression of this fusion protein was not detectable during sexual development.

Given the conservation of the mpkB MAPK signaling components in the aspergilli and interest in understanding how sexual development is regulated in these fungi, this pathway should be a fertile area for research [15–18]. While it is not yet clear that there is a sexual cycle in A. fumigatus, it certainly has all the genes that code for protein of the developmental program. Additional work on the MpkB pathway in A. nidulans and A. fumigatus could certainly begin to explore how sexual development is controlled and whether it may be possible to induce sexual development in A. fumigatus.

9.1.4 MpkC Pathway and Carbon Utilization

The MpkC MAPK is the odd man out of the pathways in the aspergilli because there is no orthologous gene or pathway in the yeast systems. Thus, mpkC and the proteins it acts on should be of special interest as there is new biology waiting to be discovered. The mpkC gene is found in all three species and one must, therefore, presume that the pathway and its regulation are conserved as well. The MpkC protein sequence is very similar to that of SakA/HogA MAPK and is thus a member of the stress-activated MAPK family [13,22]. Deletion of mpkC in either A. nidulans or A. fumigatus does not produce any visible phenotype under standard growth conditions. Because mpkC is a MAPK of the stress activated it is reasonable to assume that MpkC may have functions that overlap with SakA/HogA or that different stresses are signaled through these two MAPK pathways. Interestingly, growth of the mpkC deletion mutants in both A. nidulans and A. fumigatus is not sensitive to high-osmolarity media or hydrogen peroxide [13,22].

It is not clear what the upstream activating MAPKKK and MAPKK for MpkC, though it has been reported that overexpression of PbsB in A. nidulans can lead to phosphorylation of MpkC following hypertonic stress [13]. Determination of the components of the kinase cascade that activates MpkC will require additional work that does not involve overexpression. In A. fumigatus, we have determined that mpkC transcript levels increase only modestly in response to hypertonic stress when compared to the transcript levels of sakA/hogA [13]. In contrast, the transcript levels of mpkC show a rapid and significant increase in response to oxidative stress (hydrogen peroxide), while the levels of the sakA/hogA transcript increase only slightly [13]. These results suggest that signaling through the MpkC and SakA/HogA MAPK pathways may be very complex, requiring additional yet unknown components to achieve this degree of specificity. In this regard, it is interesting that these two MAPKs have distinct roles in nutrient sensing in A. fumigatus. The A. fumigatus mpkC deletion mutant is unable to use sorbitol or mannitol, polyalcohol sugars, as sole carbon source and [13] the sakA/HogA mutant is defective in nitrogen sensing [13].

9.2 SakA/HogA Signaling Pathway

The SakA/HogA MAPK signaling pathway is the most intensively studied of the MAPK pathways in the aspergilli [13,14,23,24]. Work has been carried out in both A. nidulans and A. fumigatus, which has been complementary and confirming. I refer to the SakA/HogA MAPK as just SakA for the protein and sakA for the gene through the remainder of this chapter for the sake of clarity. Deletion mutants for the sakA gene in A. nidulans have some modest growth defects on high-osmolarity media [14,23]. Transcript levels of some of the genes in the SakA pathway are transiently increased in response to hypertonic stress [14]. One of the genes whose messenger RNA levels increase in response to hypertonic stress is pbsA, the MAPKK of the pathway. Others include the downstream target zinc finger transcription factor msnA gene, and the gene ptpA that codes for a tyrosine phosphatase that dephosphorylates SakA and down-regulates the MAPK activity. Similar transcriptional responses are also seen for the orthologous genes in
A. fumigatus [24]. The hyphae of the sakA deletion mutant were abnormal when grown on high-osmolarity medium where they form a hyper branching mycelium [14]. Additional defects were seen in A. nidulans that included premature sexual development, and increased sensitivity of germlings to hydrogen peroxide and high temperature, 50°C [23]. In A. fumigatus additional phenotypes were observed [24]. One was that conidia of the sakA deletion mutant would germinate and grow in high-osmolarity medium. In contrast, germlings of the mutant when shifted to hypertonic medium growth arrest, while the wild-type parent would reinitiate, hyphal growth after a brief growth arrest. This suggests that the SakA pathway is not functional until after conidial germination. Finally, germination of conidia in the sakA deletion mutant was rapid regardless of the nitrogen source in the medium, suggesting a role for SakA in regulation of germination in response to the nitrogen source in the medium.

In A. nidulans, a more comprehensive study of deletion mutants for other genes that comprise the sakA pathway has revealed that regulation of the pathway is different from that in budding yeast. In S. cerevisiae the Hog1 MAPK pathway can be activated either through the Sho1, Ste11, or the Sn1 two-component signaling system regulator pathways [4,6]. There are orthologous genes for both of these osmotic sensing pathways in A. nidulans (Fig. 9.1). In A. nidulans a steC deletion mutant is not sensitive to hypertonic stress conditions; thus, it would appear that the orthologous Sho1, Ste11 pathway is not used in this fungus [21]. Because of this observation, I focus the remaining discussion on the two-component signaling system regulatory pathway. In this pathway TcsB functions as an osmosensor histidine kinase that signals to YpdA a phosphotransfer signal transducer that negatively regulates SskA a response regulator. In response to hypertonic stress, negative regulation of SskA is lost leading to activation of the MAPK cascade resulting in activation of SakA and adaptation of the fungus to hypertonic stress. Interestingly, deletion mutants of tcsB do not show sensitivity to hypertonic stress, suggesting that more than one histidine kinase may regulate this pathway or there remain other mechanisms for activation of the SakA pathway that are independent of TcsB [13,25]. As for A. nidulans, deletion of tcsB in A. fumigatus does not lead to sensitivity to hypertonic or oxidative stress [26]. In contrast, deletion mutants of the response regulator gene sskA, or genes for the kinase cascade steB, pbsA, and sakA result in sensitivity to hypertonic stress. Similarly, there is loss of SakA phosphorylation in these latter mutants in response to hypertonic or oxidative stress [13].

These experiments clearly illustrate that the SakA MAPK pathway is regulated by mechanisms that are different from those in budding yeast. The existence of multiple histidine kinases in the genomes of the aspergilli suggests that they may act as a functionally redundant network that cooperatively regulates this pathway. Alternatively, the SakA pathway may be regulated by an unknown mechanism that does not require histidine kinase activities. It is important to note that Furukawa et al. [13]. were unable to isolate

![Diagram](FIGURE 9.1) Hypertonic stress response pathway of Aspergillus nidulans. Note: In S. cerevisiae there are two pathways of activation the sensor—histidine kinase pathway, which is similar to the TcsB pathway in the figure, and the Sho1 pathway that activates the Ste11 pathway that is orthologous to SteC in A. nidulans. Activation of SakA by the SteC pathway is not supported by the experimental evidence and thus is marked by question marks.
a ypdA deletion mutant suggesting that like in S. cerevisiae deletion of this gene leads to constitutive activation of this pathway, which is a lethal event.

9.3 Where to Next?

I think it is clear that there is much to be learned from the study of the MAPK pathways in the aspergilli. The limited number of studies already conducted has led to the discovery of new and novel functions for these MAPks, which are very different for those studied in model yeast systems. For example, the SakA/HogA osmotic stress response pathway is also involved in nitrogen sensing regulating conidial germination [24]. Similarly, the MpkC pathway is involved in the utilization of polyalcohol sugars, the first time that a MAPK pathway has been linked to carbon source utilization [22]. Finally, the SakA/HogA osmotic stress response pathway is regulated by only one of the two pathways that control it in yeast [13]. There also seems to be a more robust insulation between the MAPK pathways that prevents crosstalk between the pathways. It will be interesting to determine how these pathways isolate from one another. Thus, there are considerable aspects to still be learned from the study of these MAPK pathways in the aspergilli that will result in novel findings and provide insight into the unique biology of this group of filamentous fungi and other filamentous fungi as well.

References


10

Glucogenic Carbon Metabolism

Michael J. Hynes

CONTENTS
10.1 Introduction ............................................................................................................ 129
10.2 Gluconeogenesis in Saccharomyces cerevisiae .......................................................... 130
10.3 General Pattern of Control of Catabolism in Aspergillus nidulans .............................. 131
10.4 Regulation by Induction ................................................................................................. 132
10.5 Substrates Metabolized Via the TCA Cycle ............................................................... 133
10.6 Acetate Utilization ..................................................................................................... 133
10.7 Organelle Localization of Pathways ............................................................................ 134
10.8 Peroxins ...................................................................................................................... 135
10.9 Fatty Acid Utilization ................................................................................................. 136
10.10 Regulation of Fatty Acid Catabolism ........................................................................... 137
10.11 Control of Gluconeogenesis ...................................................................................... 138
10.12 Conclusions ............................................................................................................ 139
References ....................................................................................................................... 139

10.1 Introduction

Most filamentous fungi have a saprophytic lifestyle where growth entails the utilization of environmental compounds as nutrients. These include sources of carbon, nitrogen, phosphorus, and sulfur. Of particular relevance here is the ability of fungi to metabolize a very diverse range of carbon compounds. Where a substrate is available as the sole carbon source then all cellular components must be synthesized from this compound via appropriate metabolic pathways. This, therefore, requires the organism to have the ability to rearrange the expression of gene-encoding enzymes catalyzing the appropriate steps in the pathways according to the substrates available. Commonly there is induction of enzymes specific to the breakdown of the particular compound. In addition, however, there is also a requirement for altering the central pathways of carbon metabolism so that the products of catabolism can be used to generate essential carbon-containing intermediates for biosynthesis of cellular components as well as energy and reducing power to deal with metabolic stresses. Furthermore, carbon metabolism plays a key role in various developmental stages such as asexual and sexual reproduction and for the provision of substrates for secondary metabolism. For fungal pathogens utilization of endogenous and exogenous carbon sources may have profound effects on pathogenicity [1,2].

Aspergillus nidulans has long been a favored organism for genetic studies of carbon metabolism [3]. As a saprophyte growing in the soil on decaying plant material it is capable of growing on an extremely diverse array of carbon sources. The ability to perform plate tests on carbon source utilization in the laboratory allows the assessment of growth on different carbon sources and the isolation and genetic
analysis of mutants. It should be noted, however, that this situation is rather artificial and in reality growth is usually on mixed complex substrates in the natural environment and even apparently weak sole carbon sources may contribute to survival. Now that many filamentous fungal genomes are available for comparison it has been found that a high proportion of genes involved in catabolism are conserved indicating selection in the wild for the maintenance of substrate utilization diversity.

This chapter presents an overview of what is known about carbon catabolism and its control in \textit{A. nidulans} with the emphasis on those carbon sources metabolized via the TCA cycle and requiring gluconeogenesis. The availability of the genome sequence now makes it possible to assess the complexity of the processes involved and to highlight the areas of our ignorance. Where appropriate, annotated genes from the genome sequence are given according to the Broad Institute site, http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/Home.html.

10.2 Gluconeogenesis in \textit{Saccharomyces cerevisiae}

Because carbon metabolism in \textit{Saccharomyces cerevisiae} has been extensively studied it is possible to use this knowledge as a starting point for explorations of unknown areas of carbon metabolism. However, as noted later, this organism has a very different lifestyle and differs particularly in how carbon sources are used.

Carbon metabolism has been much studied in this unicellular fungus and these investigations have continued in increasing detail following the development of whole genome analysis [e.g., 4]. However, this species is well known to be extremely specialized in its carbon metabolism with a strong preference for growth on fermentable monosaccharides-producing ethanol. When glucose is exhausted metabolism switches to a respiratory mode in which the ethanol is consumed. This diauxic shift involves an extensive rearrangement of gene expression patterns [4–6]. During growth on fermentable substrates the mitochondrial TCA cycle is not required for energy generation via oxidative phosphorylation. However, generation of 2-oxoglutarate from citrate is necessary for glutamate formation, which is essential for biosynthetic pathways leading to nitrogen-containing metabolites. Responses to a lack of mitochondrial function are controlled by the so-called retrograde response genes Rtg1, 2, and 3, which regulate expression of genes encoding enzymes leading to the synthesis of 2-oxoglutarate allowing glutamate formation in the absence of a complete functional TCA cycle [7].

In contrast to filamentous fungi, \textit{S. cerevisiae} can use (as sole carbon sources) only a limited number of substrates that result in the generation of TCA-cycle intermediates. The utilization of these compounds requires the net formation of sugars—gluconeogenesis. This is a reversal of glycolysis in which the TCA-cycle intermediate oxaloacetate is converted to hexose sugars (Fig. 10.1). Regulation of the enzymes involved is of great importance in avoiding futile cycling between degradation and biosynthesis of sugars by glycolysis opposed by gluconeogenesis. Gluconeogenic substrates are ethanol (generated by fermentation), acetate, and fatty acids—all of which result in the production of acetyl-CoA. Ethanol is converted to acetate by alcohol dehydrogenase and acetaldehyde dehydrogenase. Acetate generates acetyl-CoA by β-oxidation in peroxisomes [8]. The glyoxalate cycle, comprising the enzymes, isocitrate lyase (ICL, E.C.4.1.3.1) and malate synthase (MAS, E.C.4.1.3.2), is necessary for the net conversion of acetyl-CoA via malate to oxaloacetate which is then used for gluconeogenesis (Fig. 10.1).

The genes for metabolism of acetyl-CoA by acetyl-CoA synthetase, the glyoxalate bypass, and gluconeogenesis is controlled by the Zn2-Cys6 binuclear cluster proteins, Cat8 and Sip4 [5,9]. Some of the genes are also regulated by the Cys2His2 zinc finger protein Adr1 [10]. Therefore, growth on ethanol or acetate as sole carbon sources is dependent on the Cat8, Sip4, and Adr1 activators as well as the Snf1 kinase [4,5,7]. In the presence of glucose the Cys2His2 zinc finger Mig1 repressor, acting together with the corepressors Tup1 and Sn6, represses the expression of these genes [6]. Genes regulated by Cat8/Sip4 contain one or more cis-acting elements termed carbon source response elements (CSRE) in their 5′ regions and Cat8p and Sip4p bind to these elements [5,11].

The key enzymes for gluconeogenesis are phosphoenolpyruvate carboxykinase (PEPCK, E.C.4.1.1.32), which converts oxaloacetate to phosphoenolpyruvate and fructose-1,6-bisphosphatase (FBP, E.C.3.1.3.11)
catalyzing the final, irreversible step in hexose monophosphate formation—hydrolysis of fructose-1,6-bisphosphate to fructose-6-phosphate and phosphate (Fig. 10.1). In *S. cerevisiae* a single gene (*PCK1*) encodes PEPCK and its transcription is dependent on Cat8/Sip4 activation. PEPCK activity is also strictly controlled by glucose-induced inactivation of enzyme activity and increased mRNA instability. Similarly *FBP1* encoding FBP is regulated by the Cat8/Sip4 circuit as well as posttranscriptionally. Therefore, the overall situation in *S. cerevisiae* is that the expression of genes for both the glyoxalate cycle and gluconeogenesis is controlled by the same mechanism that responds to a lack of the fermentable carbon source glucose.

**10.3 General Pattern of Control of Catabolism in Aspergillus nidulans**

As expected from its growth as a saprophyte on decaying organic matter, *A. nidulans*, like other filamentous fungi, is capable of using complex carbohydrates and sugars as carbon sources. These will not be considered here. Glucose is the preferred carbon source and its presence results in repression of the transcription of genes involved in carbon catabolism. The CreA C2H2 finger protein, which has a similar DNA binding
domain to *S. cerevisiae* Mig1, binds to sequences with consensus SYGGRG to bring about repression [12–14]. The mechanism of repression is not at all clear. In some cases it may be due to direct competition between CreA and specific activators for binding to overlapping cis-acting DNA sequences in the 5′ region of controlled genes [13]. However, direct repression of transcription also occurs. In *S. cerevisiae* Mig1 recruits the general corepressors Tup1 and Ssn6. However, apart from the DNA binding domain, CreA has little homology to Mig1 and investigation of a possible Tup1 ortholog, RcoA, did not indicate a clear role as a corepressor [15]. Ssn6 has tetratricopeptide repeat domains and these are conserved in the N-terminal region of genes of unknown function in *A. fumigatus* (Afu2g11840—which has been proposed as an Ssn6 ortholog), *A. terreus*, and *A. oryzae* but strangely, a clear ortholog is not found in *A. nidulans*. The molecular mechanism of repression by CreA remains to be determined.

Orthologs of Adr1, the C2H2 finger protein of *S. cerevisiae*, which has a major role in controlling gluconeogenesis, are not found in filamentous fungi. The AMP-dependent protein kinase encoded by *SNF1* in *S. cerevisiae* has pleiotropic effects but in particular is required for growth on fermentable carbon sources and is required for nuclear exit of Mig1 and activates the Sip4/Cat8 activators [5,10]. Conserved orthologs are found in filamentous fungi (AN7695.3 in *A. nidulans*). Deletion of the cognate genes in some fungi has been found to affect the utilization of some carbon sources and a potential phosphorylation site has been described in orthologs of CreA for some fungi [13]. However, there is no clear effect specifically on carbon sources requiring gluconeogenesis. The CCAAT binding Hap2/3/4/5 complex is required for activation of the expression of respiratory genes in *S. cerevisiae* [7]. Orthologs of Hap2, 3, and 5 are found in filamentous fungi including *A. nidulans*. Deletion of these genes results in overall reduced growth and conidiation and in reduced expression of a diverse range of genes including ones not obviously related to carbon catabolism [16]. A role for this complex in nucleosomal positioning suggests that it is important for general activation of some genes as is the NFY complex in other eukaryotes [17]. Overall it seems that, although some of the general regulatory genes involved in carbon metabolism in yeast are conserved, the detailed mechanisms differ.

It is well known that glucose can result in posttranscriptional control of enzymes involved in gluconeogenesis in *S. cerevisiae*. This phenomenon, termed catabolite inactivation, is important in allowing rapid responses to the availability of glucose and prevents futile cycles of glycolysis and gluconeogenesis [6]. This phenomenon has been observed in *A. nidulans* for isocitrate lyase but the details have not been studied for a wide range of enzymes [18].

### 10.4 Regulation by Induction

The expression of genes involved in catabolism of particular carbon sources are usually subject to specific induction by the relevant substrate or a close derivative. In very many cases this induction results from activation by the Zn2 Cys6 binuclear cluster class of transcription factors as exemplified by the well-studied Gal4 in *S. cerevisiae* [19–21]. These factors are only found in fungi and appear to be well suited to bringing about a response to specific inducers [20]. In some cases there is good evidence for specific interaction with the inducer as a ligand changing the conformation of the protein allowing transcriptional activation. The conserved central domain (PF04082) found in many of these proteins may be crucial for this. In other cases specific corepressor proteins block activation in the absence of inducer but when inducer is present this block is released—for example, the Gal4–Gal80 and the QutA–QutR interactions in *S. cerevisiae* and *A. nidulans*, respectively [22,23]. The importance of this class of transcriptional factor in fungi is indicated by the fact that there are approximately 70 members present in *S. cerevisiae* and more than 200 present in *Aspergillus* species. Something is known about only a fraction of these. It is, therefore, a major challenge for fungal functional genomics to determine their roles and mechanisms of action. The large number of these proteins in filamentous fungi probably reflects the diversity of potential sources of nutrients present in nature. However, they are also used to regulate secondary metabolite pathways (e.g., AIIR in aflatoxin biosynthesis in *Aspergillus* spp.) as well as in drug resistance in yeast species [20,24].

The relationship between glucose repression and specific induction controls is of interest. In some cases CreA directly represses expression of the specific activator thereby resulting in glucose repression
of induction. This may be in addition to direct CreA repression of the structural genes and is illustrated by CreA control of both the alcR and alcA genes of the ethanol utilization regulon [13]. A feature of some systems may be strong control by CreA of permeases required for inducer uptake. Relief of glucose repression results in sufficient inducer uptake to allow pathway specific induction by the relevant activator.

10.5 Substrates Metabolized Via the TCA Cycle

A. nidulans can grow on carbon sources that enter the TCA cycle via a number of intermediates (Fig. 10.1). Aromatic amino acids such as quinate and benzoate are broken down to protocatechuic acid, which is then metabolized via the protocatechuic degradation pathway yielding succinate and acetyl-CoA [3,25]. The specific control of quinate catabolism has been well studied in both Neurospora crassa as well as A. nidulans [23,26]. The use of mutants affected in the metabolism of aromatic amino acids has enabled characterization of the corresponding genes involved in the classical human inborn errors of metabolism [27].

Propionate is converted to propionyl-CoA via at least two acyl-CoA synthetases [28]. One of these, the facA encoded acetyl-CoA synthetase is also used in acetate utilization (see later). Propionyl-CoA is converted via methyl-isocitrate to succinate and pyruvate by the enzymes methyl-citrate synthase and 2-methyl-citrate lyase located in the mitochondria [29,30]. This pathway bypasses the need for the anaplerotic glyoxalate cycle. Fatty acids with odd numbered chain lengths will yield propionyl-CoA and acetyl-CoA as end points of \( \beta \)-oxidation.

Many amino acids are converted to glutamate via amino-transferase enzymes. The glutamate produced is converted to 2-oxoglutarate and ammonium by NAD-dependent glutamate dehydrogenase encoded by the gdhB gene. This enzyme is required for the utilization of these amino acids both as sole sources of nitrogen, because of the need for ammonium to enable glutamine formation, and as sole carbon sources by 2-oxoglutarate formation [31,32]. The regulation by induction and by carbon and nitrogen metabolite repression of the transaminases, of gdhB and, indeed, of the amino acid permeases required for uptake require further study.

However, there are several systems that have been examined in some detail. L-proline is an extremely good source of nitrogen and/or carbon presumably reflecting the prevalence of this amino acid in plant material. The control of the proline utilization (prm) genes specifically required for proline utilization has been investigated. PrnA is a Zn2 Cys6 binuclear cluster protein necessary for induction by proline [33]. Nitrogen metabolite repression (via the AreA protein) and carbon catabolite repression (by CreA) affects in particular expression of prnB encoding a proline specific permease thereby leading to inducer exclusion under conditions of nitrogen or carbon sufficiency [34]. It is important to note that S. cerevisiae has a similar pathway but proline is only used as a (weak) sole nitrogen source and the genes involved are regulated by proline induction and nitrogen catabolite repression but not by glucose repression [35]. 4-aminobutyric acid (GABA) is also an excellent carbon or nitrogen source and is metabolized via transamination to glutamate and succinic-semialdehyde. Specific induction of the genes involved is by the AmdR regulatory protein [36]. Regulation of the GABA permease is particularly complex involving AmdR for induction, PacC for pH control, AreA for nitrogen metabolite control, and CreA for carbon catabolite repression [37].

10.6 Acetate Utilization

The isolation of fac mutants resistant to fluoroacetate [38] as well as a collection of acu mutants unable to use acetate as a sole carbon source isolated in the 1970s has provided a rich resource for studies in A. nidulans [39]. These include mutants specifically affected in the utilization of carbon sources metabolized via acetyl-CoA—ethanol, acetate, and fatty acids and includes the glyoxalate cycle genes acuD (ICL) and acuE (MAS). Mutations in the facB gene result in fluoroacetate resistance and an inability to grow on two-carbon compounds metabolized via acetyl-CoA and FacB activates acetate inducible
expression of *acuD* and *acuE* as well as *facA* (acetyl-CoA synthase), which are required for growth on acetate as a sole source of carbon. FacB is a Zn\(_2\) Cys\(_6\) protein with similarity to Cat8 and Sip4 of *S. cerevisiae* and binding sites for this activator have been found in the 5' region of acetate induced genes [40,41] and these are similar to the Cat8/Sip4 binding sites [5,9,11]. There is evidence that acetate control is at the level of *facB* expression—overexpression of FacB results in increased noninduced levels of expression of target genes; *facB* is subject to CreA-dependent glucose repression and *facB* is induced by acetate [42,43]. However, the activation capacity of FacB may also be increased by acetate [44].

Acetyl-CoA enters the mitochondria for metabolism via the TCA cycle as the acetyl-carnitine derivative produced by the cytoplasmic carnitine-acetyl transferase encoded by the *facC* gene specifically required for acetate utilization and regulated by FacB [45]. The *acuH* product is an acyl-carnitine mitochondrial transporter required for entry [46]. Resynthesis of acetyl-CoA from acetyl-carnitine in the mitochondria is carried out by the carnitine-acetyl-transferase product of *acuJ*, which is regulated by FacB-dependent acetate induction (see Fig. 10.2) [45,47].

Mutations in *facA*, *facC*, and the regulatory gene, *facB*, all lead to loss of growth on acetate (and ethanol) but do not affect growth on short chain (e.g., butyrate—C4; hexanoate—C6) or medium/long chain (C12-C22) fatty acids [39,47]. However, mutations in *acuD*, *acuE* (glyoxalate cycle) and *acuJ* and *acuH* result in loss of utilization of both acetate and fatty acids [39,47]. This, therefore, raises the question of fatty acid β-oxidation in peroxisomes and more generally the localization of metabolic pathways in the cell (Fig. 10.2).

### 10.7 Organellar Localization of Pathways

A variety of programs are available that predict mitochondrial targeting of proteins via the N-terminal signal peptide [e.g., see 48]. These may be confirmed by labeling with tags such as fluorescent proteins and shown to colocalize with reagents such as mitotracker [49,50]. The carnitine acyl transferase encoded
by acuJ is predicted to be localized to the mitochondria and this has been confirmed [51]. In contrast the facC encoded enzyme is predicted to be cytoplasmic.

Eukaryotes contain single-membrane bound organelles generally called microbodies containing specialized enzymes involved in a wide range of metabolic activities [52,53]. It is common for these enzymes to include oxidases that generate reactive oxygen species and the microbodies contain catalase and glutathione peroxidases to deal with these. When this occurs the microbodies are called peroxisomes. When microbodies contain enzymes of the glyoxylate cycle but not catalase they have been termed “glyoxysomes” but it is unlikely that there is any real difference in structure or biogenesis. There are two major classes of signals for targeting of proteins to the peroxisomal matrix (PTSs). The most common class (PTS-1) comprises the three C-terminal amino acids of the form S/A R/K L/M. The second, less common class, (PTS-2) is located near the N-terminus and has the consensus R/K L/V/I X5 H/Q L/A F/I where X is any amino acid. There are well-documented exceptions to these consensus sequences with proteins containing cryptic signaling sequences, context-dependent targeting, and the potential for “piggy-backing” via association with proteins containing PTS sequences [53]. Therefore, predictions of peroxisomal localization of annotated proteins in genome sequences must be treated with caution. However, the availability of many genome sequences now makes it possible to compare orthologs of proteins and see much closer fits to consensus sequences in some of these, thereby increasing confidence in predictions for proteins with ambiguous sequences. In the A. nidulans genome an analysis of predicted proteins revealed 91 proteins with the most common PTS1 signals (S/A K/R L) revealing a potential large minimal number of peroxisomal proteins [54].

Both specific enzymes of the glyoxylate cycle malate synthase (acuE) and isocitrate lyase (acuD) are peroxisomal [49,50]. However, while AcuE has a clear PTS-1 (SKL), AcuD lacks obvious peroxisomal targeting signals [55]. Like the corresponding S. cerevisiae CAT2 gene encoding carnitine acetyl transferase, AcuJ has both a mitochondrial targeting sequence and a PTS-1 and is localized to both organelles [54,56,57]. It has been shown that the NADP-dependent isocitrate dehydrogenase encoded by idpA has mitochondrial, cytoplasmic, and peroxisomal locations with the protein containing both mitochondrial and peroxisomal (PTS-1) targeting signals. Localization is determined by differential transcription start points and is in contrast to the situation in S. cerevisiae where three different genes encode the differentially localized proteins [50]. Similarly three genes encode citrate synthase in S. cerevisiae while a single gene is found in A. nidulans [58]. The protein has a mitochondrial targeting sequence and also a PTS-1 but the predicted peroxisomal localization has not been demonstrated [58].

The complex problem of transport of metabolites between cellular compartments has not been thoroughly addressed in A. nidulans. As noted earlier the shuttling of acetyl-CoA between peroxisomes, cytosol, and mitochondria is dependent on the carnitine acetyl-CoA transferases and AcuH, the acyl-carnitine transporter (Fig. 10.2). Unlike mammals, where there is an additional medium chain length carnitine fatty-acylCoA transferase, A. nidulans only has a single peroxisomal/mitochondrial short chain carnitine acylCoA transferase. A. nidulans has 37 annotated members of the mitochondrial carrier proteins (PF00153). Amongst these is an ortholog of the S. cerevisiae Sfc1 succinate—fumarate—transporter—AN7287.3. We have found that this corresponds to the acuL gene identified as being necessary for growth on acetate and fatty acids [39,59] indicating that this is required for the shuttling of succinate resulting from the glyoxylate cycle. There are clear orthologs of other characterized S. cerevisiae mitochondrial transporters—oxaloacetate (OAC1-AN0066.3, which has a PTS1 and so may also be peroxisomal), 2-oxoglutarate (ODC1-AN10172.3), dicarboxylate-phosphate exchange (DIC1-AN6254.3), and the citrate transporter (CTPI-AN3461.3). Much more analysis is required.

### 10.8 Peroxins

More than 30 proteins have been found to be involved in peroxisome biogenesis, proliferation, or function—the peroxins encoded by PEX genes. These are numbered in order of discovery—with most having been found and characterized first in S. cerevisiae. Key proteins for import of proteins into the peroxisomal matrix are PEX5, the PTS1 receptor, PEX7 the PTS2 receptor, and PEX1, 6, and 13 necessary for recycling of these receptors [60]. Single orthologs for each of these are found in A. nidulans.
and mutations or deletions of each of these have been found to affect the localization of peroxisomal proteins as predicted [61]. In *S. cerevisiae* it has recently been found that Pex3 is crucial for *de novo* formation of peroxisomes by budding from the endoplasmic reticulum [62,63]. Inactivation of the single ortholog found in *A. nidulans* is consistent with a lack of peroxisome formation [61]. It is of interest that strains containing the *pex3* mutation are viable (although there are minor abnormalities in asexual and sexual reproduction) indicating that peroxisomal functions are dispensable. A recent thorough comparison of *PEX* genes in fungal genomes has been performed [64]. In general, filamentous fungi contain the full repertoire of peroxins found in *S. cerevisiae*, plants, and humans with minor differences including additional paralogs in *S. cerevisiae*. Of particular interest is the finding in all species that there are multiple peroxin paralogs involved in proliferation of peroxisomes—the Pex11, Pex23 family. This may reflect different requirements for maintenance and expansion of peroxisome numbers in response to developmental or environmental cues. It is well known that in *S. cerevisiae* Pex11 is necessary for peroxisome proliferation in response to the presence of oleate.

### 10.9 Fatty Acid Utilization

*A. nidulans* can grow on a wide range of fatty acids, both saturated and unsaturated, and on both short and long chain lengths [39,47,49]. However, for unknown reasons, exogenously supplied C7-C11 fatty acids are inhibitory to growth [49]. There is likely to be abundant sources of fatty acids in the environment of this fungus with decaying plant materials containing waxes such as cutin (a polymer of hydroxy-oleate) and lipids. Consistent with this a variety of cutinases and lipases are found in predicted annotated genes in the genome.

β-oxidation of fatty acids entails coupling of fatty acids to CoA and then a series of reactions resulting in the chain length being shortened by two carbons generating one molecule of acetyl-CoA. In *S. cerevisiae* all steps occur entirely within peroxisomes. Fatty acids such as oleate are metabolized completely to acetyl-CoA, which is metabolized via the glyoxalate cycle and the TCA cycle [53]. Mutants lacking functional peroxisomes, β-oxidation enzymes, or the glyoxalate cycle enzymes are unable to use oleate as a carbon source. In mammals the situation is more complex with two sets of differentially localized enzymes in peroxisomes and mitochondria. Long-chain fatty acids are converted to medium-chain length fatty acids in peroxisomes and then transferred as the acyl-carnitine derivative to mitochondria for complete oxidation to acetyl-CoA. Mammals contain two acyl-carnitine transferases—one for acetyl-CoA and one for medium chain fatty acyl-CoA. As noted earlier inspection of annotated genes indicates that *A. nidulans* lacks the medium chain length enzyme.

Enzymes of β-oxidation have been demonstrated to occur in peroxisomes in *A. nidulans* [65]. The observation that a mutation in the gene encoding the peroxin Pex6 results in fatty acid toxicity presumably due to mislocalization of β-oxidation enzymes also indicates peroxisomal metabolism [47,61]. Deletion of the *foxA* gene predicted to encode a peroxisomal multifunctional enzyme (orthologous to *S. cerevisiae* Fox2) resulted in loss of growth on erucic acid (C22) and reduced growth on oleate [49]. Analysis of predicted proteins in the genome reveals a large number of potential β-oxidation enzymes indicating the likelihood of significant redundancy. This is in contrast to the situation in *S. cerevisiae* and may reflect a greater range of potential fatty acid substrates for *A. nidulans*.

The first step in the β-oxidation of fatty-acyl CoA derivatives is carried out by acyl-CoA oxidases (E.C. 1.3.3.6). While *A. nidulans* has at least two probable peroxisomal oxidase enzymes (AN6765.3, AN6752.3) *N. crassa* lacks these. This is associated with a lack of enzymes required to counteract reactive oxygen species generated in peroxisomes—peroxisomally located catalase [66] or glutathione peroxidases in *N. crassa*. In contrast *A. nidulans* has these enzymes (catalaseC—AN5918.3 [67] and glutathione peroxidase—AN5440.3). Furthermore *A. nidulans* contains some oxidases with obvious PTS1 sequences while the equivalent predicted highly conserved proteins in *N. crassa* lack this targeting sequence [68]. A very clear example is urate oxidase (AN9470.3). However, both *N. crassa* and *A. nidulans* contain genes predicted to code for acyl-CoA dehydrogenases, which contain PTS1 sequences and have high similarity to a rat liver peroxisomal acyl-CoA dehydrogenase [69]. These enzymes provide possible alternatives to the peroxisomal acyl-CoA oxidases.
The ketoacyl-thiolase enzymes illustrate the complexity of the possible β-oxidation enzymes present in *A. nidulans*. This activity is the last step in the pathway. In *S. cerevisiae* this is carried out by Pot1, which has a PTS2 targeting sequence [53]. There are three similar proteins in *A. nidulans* each with a PTS2. A second thiolase, Tes1, has been identified in *S. cerevisiae* with an imperfect PTS1 (AKF [79]) and three similar proteins are found in *A. nidulans* also with possible PTS1 sequences (AKM, ARI, ARF). In addition there are four possible enzymes that are predicted to be mitochondrial. This raises the question of a mitochondrial pathway in *A. nidulans*

This has been demonstrated [49]. A deletion of AN5916.3 (now designated echA), encoding a putative mitochondrial short-chain enoyl-CoA hydratase showed that loss of this enzyme results in an inability to grow on butyrate (C4) and hexanoate (C6) and to reduce growth on oleate (C18) and erucate (C22) and to sensitivity to these fatty acids in the presence lactose as an alternative carbon source. Deletion of AN0824.3 (now scdA) encoding a short-chain dehydrogenase also results in loss of growth on butyrate and hexanoate but does not result in fatty acid inhibition or in reduced growth on long chain fatty acids. Furthermore the scdA; echA double deletion is not inhibited by long- or short-chain fatty acids indicating that accumulation of the enoyl-CoA in echA results in inhibition [71]. The other two steps in a mitochondrial pathway—hydroxyacyl-CoA dehydrogenase, and ketoacyl-CoA thiolase—were predicted to be encoded by AN7008.3 and AN4179.3, respectively [49]. We have confirmed the role of these genes. A mutant unable to use butyrate but able to grow on oleate was found to be complemented by a sequence containing AN4179.3 and a disruption of AN7008.3 is inhibited by both long- and short-chain fatty acids [61]. This latter result indicates that accumulation of the hydroxyacyl-CoA intermediate is inhibitory. The observations that disruption of this mitochondrial pathway only affects growth on long-chain fatty acids, if accumulation of intermediates is inhibitory, indicates that it is not essential for growth on long-chain fatty acids but that fatty acyl-CoA intermediates generated by peroxisomal β-oxidation can be metabolized in the mitochondria (see Fig. 10.2). The chain length of these intermediates transferred from peroxisomes to mitochondria is not known, but as noted earlier, the lack of an ortholog of the medium-chain carnitine acyl-transferase indicates that the situation is not identical to that occurring in mammals. It is interesting to note that orthologs of the genes of the mitochondrial pathway are found in *Yarrowia lipolytica* but not in other sequenced hemi-ascomycete yeast species.

### 10.10 Regulation of Fatty Acid Catabolism

The complexity of fatty acid metabolism in *A. nidulans* is paralleled by the regulation of expression of the genes involved. In *S. cerevisiae* two Zn2 Cys6 binuclear cluster protein activators, Oaf1 and Pip2, mediate fatty acid induction of genes encoding peroxisomal and β-oxidation proteins with Oaf1 directly interacting with the inducer [72,73]. These proteins do not control expression of the glyoxalate pathway enzymes that are controlled by the Cat8/Sip4 activators (see earlier). In contrast in *A. nidulans* these enzymes, as well as the *acuJ* encoded carnitine acetyl-transferase, are subject to fatty-acid induction in addition to FacB mediated acetate induction. This is clearly shown by the ability of *facB* mutants to grow on fatty acids [39,47]. It turns out that there are three classes of genes—FacB regulated genes specific for acetate utilization (*facA* and *facC*); genes required for both acetate and fatty-acid utilization, which are controlled by FacB-mediated acetate induction, and by fatty-acid induction (*acud*, *acuE*, *acuJ*, *acuL*), and genes specifically induced by fatty acids (e.g. those encoding β-oxidation enzymes and peroxins). Regulatory genes responsible for fatty-acid induction have been identified [47].

Two genes, *farA* and *farB*, encode related Zn2 Cys6 binuclear cluster proteins that are highly conserved in filamentous ascomycetes. Mutations in *farB* result in loss of growth on short-chain fatty acids and in loss of short-chain induction of enzymes and functions involved in fatty-acid catabolism including those of the glyoxalate cycle. *farA* mutations affect growth on both long- and short-chain fatty acids and affect induction by both long- and short-chain fatty acids. Neither gene affects growth on acetate. The sequences of the DNA binding domains of these proteins are closely related to each other and, using expressed fusion proteins in electrophoretic mobility shift experiments, the core binding sequence for both FarA and FarB has been shown to be CCGAGG. The proteins turn out to be orthologous to two cutinase transcription factors previously identified in *Nectria haematococca*, which have also been found to bind to
sequences containing the CCGAGG motif [74,75]. This core sequence is found to be overrepresented in the 5′ region of a very large number of genes for proteins predicted to be involved in fatty-acid catabolism including peroxins (such as Pex11), cutinases, lipases, and glutathione peroxidase in both *A. nidulans* and other filamentous ascomycetes. This sequence is not overrepresented in genes specific for acetate utilization. Of particular interest is the finding that a single ortholog most closely related to *farA* is found in the genomes of *Candida albicans*, *Debaryomyces hansenii*, and *Yarrowia lipolytica* but not in other hemi-ascomycetes, which contain orthologs of *S. cerevisiae OAF1* and *PIP2* [47]. Consistent with this observation the CCGAGG core is overrepresented in upstream of relevant genes, thereby providing a compelling case for a conserved functional role for the FarA ortholog in these three species.

To complicate matters, mutations in a third gene, *scfA*, also encoding a Zn2 Cys6 binuclear cluster protein, eliminate short-chain induction [47]. *ScfA* is not related to *FarA* or *FarB* and, unlike these proteins, is not highly conserved in filamentous ascomycetes with clear orthologs only found in *N. crassa* and *Aspergillus spp*. The target DNA binding sequence for this protein has not been determined but, for the *acuJ* gene, the site of action is in the region bound by *FarA* and *FarB*. *ScfA* does not appear to function by regulating expression of the other two genes.

### 10.11 Control of Gluconeogenesis

Unlike *S. cerevisiae*, *A. nidulans* is capable of using amino acids such as glutamate and proline as sole carbon sources and gluconeogenesis is required for their utilization. The finding that *facB* mutants are unaffected in the utilization of these compounds clearly indicates that *FacB* is a specific regulator of two carbon metabolism and does not control gluconeogenesis [39,76,77]. Therefore, the situation is very different from that in *S. cerevisiae* where both the glyoxalate cycle and gluconeogenesis are controlled by a single circuit. The situation for *A. nidulans* is logical because during growth on carbon sources that do not generate acetyl-CoA the glyoxalate cycle is not required.

The regulation of the key genes of gluconeogenesis in *A. nidulans* is, therefore, of interest. Mutations in the *acuF* gene have been isolated by virtue of an inability to grow on acetate and found to specifically lack PEPCK activity [39]. They are also unable to grow on all carbon sources requiring gluconeogenesis. PEPCK activity is induced not only by acetate but also by glutamate, proline, and other sources of TCA cycle intermediates but is not strongly repressed by glucose [76]. Cloning of the *acuF* gene showed that it is indeed the structural gene for PEPCK and analysis of regulation of the gene via Northern blotting and *lacZ* fusions confirmed that expression is induced by sources of TCA cycle intermediates and the pattern of regulation is not consistent with direct regulation by FacB. Furthermore, deletion analysis of the 5′ region of *acuF* showed that the region responsible for induction lacks FacB binding sites [77]. *A. nidulans acuG* mutants have been isolated as unable to grow on acetate and found to specifically lack FBP activity and consistent with this they are unable to utilize any gluconeogenic carbon sources, as well as glycerol [39]. The gene is controlled by CreA-mediated carbon catabolite repression and in addition *acuG* expression is elevated under conditions where TCA-cycle intermediates accumulate consistent with an induction mechanism [78]. As for *acuF* there is no evidence for direct induction by FacB. Therefore, the question arises—what is the mechanism of induction of gluconeogenesis in *A. nidulans*?

The *acuK* and *acuM* genes were identified in the original screen for acetate mutants [39]. However, in the initial limited testing and in more extensive examination of carbon sources, it has been found that mutations in these genes do not just affect growth on acetate but also on all carbon sources requiring gluconeogenesis. Consistent with this the *acuK248* and *acuM301* mutations each result in loss of induction of the *acuF* gene by sources of TCA-cycle intermediates [78]. Similarly TCA-cycle activation of expression of the *acuG* gene has been found to be lost in these mutants [78]. A direct role in transcriptional activation for the products of *acuK* and *acuM* has been shown by the cloning of these genes indicating that the gene products contain related Zn2-Cys6 binuclear cluster DNA binding domains [78]. Apparent orthologs of these genes are present in the many ascomycete filamentous fungal genomes now available indicating that this control circuit is conserved in filamentous fungi. The previous suggestion that *acuK* and *acuM* encode NADP-malic enzyme [79] is likely to be explained by these genes regulating expression of the gene for this enzyme that may play a crucial role in providing pyruvate from malate...
Gluconeogenic Carbon Metabolism

1. During growth on amino acids metabolized via glutamate, therefore, it is proposed that AcuK and AcuM activate expression of genes for gluconeogenesis in response to a TCA-cycle intermediate—probably malate or oxaloacetate. An overview of the differences between the regulation of the pathways discussed between A. nidulans and S. cerevisiae is presented in Figure 10.3.

10.12 Conclusions

The aforementioned discussion has not attempted a comprehensive coverage of all aspects of gluconeogenic carbon catabolism but rather highlights the fact that the availability of comparative genomic data coupled with previous and ongoing molecular genetic studies will be productive in the elaboration of the complexities of the regulatory circuits. As always the availability of genetic analysis in A. nidulans provides a powerful approach to these problems that is not available in most pathogenic or industrial species. For example, the isolation and genetic analysis of mutants was crucial in the discovery of the fatty-acid regulatory genes [47]. The importance of orthologs in other species for pathogenicity, development, and secondary metabolism can now be determined by generating appropriate deletion mutants. In the future genome-wide studies employing microarrays, proteomics, and particularly metabolomics will be important in the description of all the complexities of responses to carbon nutrients. However, studies on individual genes and their cognate regulators must continue.

References


11

Amino Acid Supply of Aspergillus

Oliver W. Draht, Silke Busch, Kay Hofmann, Susanna Braus-Stromeyer, Kerstin Helmstaedt, Gustavo H. Goldman, and Gerhard H. Braus

CONTENTS
11.1 Introduction ............................................................................................................... 143
11.2 Uptake of Amino Acids in Aspergillus ...................................................................... 145
   11.2.1 Aspergillus Amino Acid Uptake in Comparison to Other Fungi ...................... 145
   11.2.2 Aspergillus Amino Acid Uptake Systems in Comparison to Mammalian Counterparts .................................................................................................................. 147
11.3 Biosynthesis of Amino Acids in Aspergillus .............................................................. 148
   11.3.1 Sensing of the Intracellular Amino Acid Pool: Sensor Kinase CpcC and the TOR Pathway .............................................................................................................. 148
   11.3.2 Cross-Pathway Control (CPC) System of Aspergillus ................................... 151
      11.3.2.1 Global Transcription Factor CpcA ......................................................... 151
      11.3.2.2 Control of the CpcA-mRNA Amount ...................................................... 151
      11.3.2.3 Transport of the Transcription Factor into the Nucleus ......................... 153
      11.3.2.4 Ribosomal CpcB Component Represses Amino Acid Biosynthesis .............. 153
   11.3.3 Examples for the Synthesis of Amino Acids and Derivatives ............................ 153
      11.3.3.1 Histidine Biosynthesis ........................................................................ 154
      11.3.3.2 Lysine and Penicillin Biosynthesis of Aspergilli .................................. 154
      11.3.3.3 Aromatic Amino Acid and Terrequinone A Biosynthesis ...................... 155
11.4 Amino Acids Obtained by Protein Degradation in Aspergillus ................................. 156
   11.4.1 Prerequisites for Protein Degradation ............................................................ 157
   11.4.2 Ubiquitylation of Phosphorylated Substrates .................................................... 160
   11.4.3 Ubiquitin Ligases ............................................................................................ 160
   11.4.4 SCF Activity Is Controlled by Alternating Neddylation Status ....................... 162
   11.4.5 PCI Complexes ............................................................................................... 164
      11.4.5.1 PCI and MPN Domain Proteins ............................................................ 164
      11.4.5.2 COP9 Signalosome (CSN) ................................................................. 166
      11.4.5.3 Translation Initiation Factor 3 (eIF3) .................................................. 166
11.5 Conclusion ................................................................................................................. 166
Acknowledgments .................................................................................................................. 168
References ........................................................................................................................... 169

11.1 Introduction

The aspergilli comprise a divergent and highly versatile group of filamentous fungi [1]. Among the over 185 aspergilli are several species with impact on human health, including 20 human pathogens. In addition, several economically, medically, and agriculturally important fungal species are part of the
Aspergillus family [1]. Bioactive molecules such as aflatoxins are secreted by Aspergillus fumigatus and Aspergillus flavus [2–4]. Additionally, A. fumigatus is an important human pathogen causing invasive aspergillosis in immunocompromised patients [5]. Aspergillus oryzae and Aspergillus niger are of high importance to produce sake, miso, soy sauce, and citric acid in industrial standards [6]. Aspergillus nidulans constitutes a representative of this fungal genus that is capable of diverse and complex biosyntheses and differentiation processes. The most complicated developmental process includes the well-characterized sexual differentiation process where after mating with a compatible partner or “selfing,” closed fruitbodies, which are called cleistothecia, are formed which contain octades of ascospores [7,8]. During the last century, molecular methods were developed to easily investigate and manipulate these eukaryotic model organisms. Therefore, Aspergillus species are particularly suited for in-depth studies on regulatory networks and cross-connections between environmental stimuli, metabolism, and development and have steadily advanced our understanding of eukaryotic physiology. The aim of this chapter is to give an overview of the metabolic potential Aspergillus species have developed to acquire amino acids. By comprehensive genome analysis regarding uptake systems, the general control/cross-pathway control (gc/cpc) of amino acid biosynthesis and the COP9 signalosome (CSN) of A. nidulans, A. oryzae, A. niger, and A. fumigatus, we describe three concepts to obtain amino acid homeostasis in an Aspergillus cell in detail: (1) uptake of free amino acids, (2) energy consuming de novo biosynthesis, and (3) controlled recycling of used amino acids. A diagram of this concept is given in Figure 11.1, respective sources for all investigated sequences are given in Table 11.1 through 11.5. Fungi, plants, and prokaryotes are able to synthesize all amino acids, whereas mammals have to take up aromatic amino acids, which they are unable to produce [9]. For efficient bio-synthesis, a well-characterized gc system, cpc, evolved, which is responsible for the regulation of amino acid biosynthesis. A similar control mechanism is known from yeast, where it is called general control (gc). Both systems have a central, global activator of transcription in common: CpcA for Aspergillus and Gcn4p for S. cerevisiae. This transcriptional activator is conserved from yeast to man, where ATF4 plays a role similar to CpcA and Gcn4p [10]. However, even fungi prefer to take up amino acids from their diet, which is simply less energy consuming than amino acid de novo biosynthesis.

Under limiting conditions or during developmental processes when specific proteins are no longer required, protein degradation results in an additional supply of amino acids [11]. A controlled timely expression and destruction of proteins is a major regulatory mechanism in cellular processes [12,13].

![Figure 11.1](image_url) Possible ways for a fungal cell to gain amino acids. Note: A living, developing fungal cell has three major ways to replenish its internal amino acid pool in favor of protein biosynthesis. The uptake of free peptides, amino acids, or amino acid precursors from the surrounding medium, energy-consuming amino acid biosynthesis controlled by the cross-pathway control or recycling of amino acids by degradation of dispensable proteins. Uptake of amino acids and internal amino acid sensing combined with regulated biosynthesis of amino acid is essential for the cell to replenish the internal amino acid pool.
Amino Acid Supply of Aspergillus

The main eukaryotic degradation machinery, the 26S proteasome, specifically degrades ubiquitylated proteins [14]. The ubiquitin-proteasome pathway acts as a very complex machinery regulating protein degradation. Unused proteins are usually phosphorylated and subsequently ubiquitylated by Cullin-RING ligases (CRL). Cullin-RING ubiquitin ligases consist of a typical cullin subunit and a RING-finger domain protein [15]. These two proteins interact with substrate specificity proteins that recruit specific substrates. The bound substrates are poly-tagged by the small protein modifier ubiquitin and thus are marked for degradation by the 26S proteasome. The activity of CRLs is regulated by yet another small protein modifier, NEDD8, which is conjugated in a similar mechanism like ubiquitin to the cullin subunit and removed by the CSN, which acts as a regulator in protein degradation and development [16]. Without any nutrition, specific protein degradation can be exchanged for unspecific bulk degradation of large compartments of cells in the vacuole by a process called autophagy [17].

11.2 Uptake of Amino Acids in Aspergillus

In fungi, amino acid uptake is carefully controlled and requires transporters for a broad range of amino acids as well as specific transporters. Amino acids are essential precursors to ribosomal biosynthesis of proteins, for nonribosomal products of secondary metabolism such as β-lactam antibiotics synthesis [18,19], but also serve as poor carbon or nitrogen sources in times of nitrogen or carbon starvation [20–22]. When available in the environment amino acid uptake is more economical for the cell due to several energy-consuming steps that are required for amino acid biosynthesis. Amino acid uptake is regulated by their availability in the environment, which is perceived by sensors that might act as transporters themselves [23]. Most fungi are also able to secrete proteases into the medium to break down extracellular proteins and take up the released amino acids, although these actions are carefully regulated and mainly happen under stress conditions such as nitrogen starvation and the presence of extracellular protein [24–26]. Fungi are able to use a variety of nitrogen sources such as proteins, amino acids, nitrate, purines, and even acetamide. The preferred nitrogen sources of aspergilli are ammonium, glutamine, or glutamate. All nitrogen-related processes pass at some stage through the amino acids glutamine and glutamate where ammonium as nitrogen donor is easily transferred between these two amino acids and other keto-group containing molecules [21,27]. Specific permeases exist for different more or less preferred nitrogen sources [21,28]. The biosyntheses of these permeases and enzymes required for the related catabolism often have to be induced according to the available nitrogen sources in the environment. Table 11.1A describes the Aspergillus genes that presumably encode amino acid transporters and their putative function (Table 11.1B).

11.2.1 Aspergillus Amino Acid Uptake in Comparison to Other Fungi

The first step before the uptake of extracellular amino acids is the sensing of amino acids outside of the cell [29]. A subsequent signal cascade is needed to pass on the signal for availability into the cell and convert the signal into appropriate actions—to take up the amino acids as nitrogen or carbon sources or for protein biosynthesis or to pass on the signal so that proteins that are no longer required can be degraded and recycled [29]. The SPS system (Ssy1p–Ptr3p–Ssy5p) is a well-characterized system in the plasma membrane that senses the presence of extracellular amino acids and is common to S. cerevisiae and Schizosaccharomyces pombe. The SPS cleaves an NLS-masking domain from the heterodimeric transcription factor Stp1p when amino acids are present in the medium, sending it to the nucleus to enhance transcription of amino acid transporters [30].

In the presence of amino acids, the system initiates the signals that ultimately lead to the expression of amino acid transporters and permeases [29]. Ssy1p is a protein consisting of 12 transmembrane helices and strongly resembles an amino acid transporter, but in contrast to these has an unusually elongated N-terminus that is required for its activity and most likely protein–protein interactions [31]. Transmembrane proteins can be found throughout the four sequenced Aspergillus genomes but none exhibits the elongated N-terminus, nor can there be found a protein with high similarity, indicating that an Ssy1p homolog is possibly not present. A region homologous to the elongated N-terminus cannot be found in any of the four
The Aspergilli

TABLE 11.1A
Genes for Proteins for Amino Acid Uptake in Aspergillus (the First Column States the Proposed Names of the Respective Proteins in A. nidulans, which were Derived from the Names of the Characterized Homologs)

<table>
<thead>
<tr>
<th>Transporter</th>
<th>Family of Mammalian Transporter</th>
<th>Putative Function Derived from Function of Homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td>GapA</td>
<td>SLC7 (System L*)</td>
<td>General amino acid permease for L-amino acids, some D-amino acids</td>
</tr>
<tr>
<td></td>
<td>LAT1</td>
<td>Trp/Tyr transporter [184]</td>
</tr>
<tr>
<td>PutD</td>
<td>SLC38 (System A)</td>
<td>Nitrogen regulated proline transporter [185]</td>
</tr>
<tr>
<td>SIC</td>
<td>SLC3 (System XAG)</td>
<td>Sodium independent neutral amino acid transporter (Q, N, I, L, Y), putative uptake of amino acids from lysosome [51,53]</td>
</tr>
<tr>
<td>SIC</td>
<td>SAT1</td>
<td>Cl- dependent GABA, betaine and taurine transporter, Na+ and Cl- dependent high-affinity glycine transporter, glutamate transporter, not found in A. fumigatus [40]</td>
</tr>
<tr>
<td>SLC</td>
<td>SLC36 (System imino)</td>
<td>Sodium independent neutral amino acid transporter (Q, N, I, L, Y), putative uptake of amino acids from lysosome [51,53]</td>
</tr>
<tr>
<td></td>
<td>PAT1</td>
<td>Sodium independent neutral amino acid transporter (Q, N, I, L, Y), putative uptake of amino acids from lysosome [51,53]</td>
</tr>
</tbody>
</table>

TABLE 11.1B
Overview of Putative Amino Acid Transporter Functions in Aspergillus

<table>
<thead>
<tr>
<th>Transporter</th>
<th>Putative Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>GapA</td>
<td>General amino acid permease for L-amino acids, some D-amino acids</td>
</tr>
<tr>
<td>PutD</td>
<td>Nitrogen regulated proline transporter</td>
</tr>
<tr>
<td>SIC</td>
<td>Sodium independent neutral amino acid transporter (Q, N, I, L, Y), putative uptake of amino acids from lysosome</td>
</tr>
<tr>
<td>SLC</td>
<td>Cl- dependent GABA, betaine and taurine transporter, Na+ and Cl- dependent high-affinity glycine transporter, glutamate transporter, not found in A. fumigatus</td>
</tr>
<tr>
<td>SLC</td>
<td>Sodium independent neutral amino acid transporter (Q, N, I, L, Y), putative uptake of amino acids from lysosome</td>
</tr>
</tbody>
</table>

genomes either. Similarly, neither Ptr3p and Ssy5p nor the transcription factor involved is present in the aspergilli’s genomes. Thus, an SPS-like system does not exist in Aspergillus species. Remarkably, the SPS system is generally not found in higher eukaryotes either.

Another prominent example of a well-characterized amino acid transporter in yeast is Gap1p (general amino acid permease). This protein constitutes a general high-capacity amino acid permease, which is known to transport all naturally occurring L-amino acids and also various D-amino acids [32] and also functions as amino acid sensor [33]. Equally well known are the Trp/Tyr transporters Tat1p and Tat2p...
All four Aspergillus species display an amino acid transport system that we called GapA, which is putatively homologous to both the Gap1p and Tat1/2p and shows over 50% similarities to the yeast counterparts. These Aspergillus transporters do not have equivalents in higher eukaryotes. A putative homolog of the proline-specific permease Put4p called PutD was also found in the aspergilli with approximately 40% similarities to the yeasts and Neurospora proteins, but is not found in higher eukaryotes. In addition, GATA transcription factors such as yeast Gln3p, Aspergillus AreA, or NIT-2 in N. crassa are activated to increase the production of enzymes needed for uptake and utilization if the fungus senses poor nitrogen sources such as amino acids [21,28,35,36]. This nitrogen discrimination pathway (NDP), which is modulated by the Tor proteins (see later), is repressed if rich nitrogen sources such as ammonium are present [37–39].

11.2.2 Aspergillus Amino Acid Uptake Systems in Comparison to Mammalian Counterparts

In mammalian systems, nutrients, especially amino acids, play an important role in the regulation of physiological processes. This is of special importance since mammalian systems are unable to synthesize several amino acids de novo making efficient uptake systems inevitable. In mammalian cells, an SPS-like system is not present and it is not entirely clear how external signals are forwarded to the nucleus to enable the cell to react. Different mechanisms have been proposed where amino acid transporters may have an impact on signal transduction. The first one would be that an amino acid transporter acts as a substrate specific receptor and passes on a signal during transport of its substrate [40]. A second impact of amino acid transporters could be that during import of the amino acid substrate other molecules are symported or antiported, which has an effect on intracellular physiology (pH, change in membrane potential, cellular volume change) [40]. Furthermore, amino acid transporters could exert influence by importing substrates and thereby activating intracellular sensors such as mGCN2 [41]. In fact, mammalian cells have a broad range of mechanisms for the transmembrane transport of amino acids [40]. The transporters have been classified into distinct “systems” defining substrate specificity, transport mechanism, and regulatory properties [40]. The systems can be divided into groups of sodium dependent and independent transporters and groups depending on the charge of the transported amino acids, but the variety of transporters is also often characterized simply by gene families that cluster certain types of transporters [40].

A thorough search for all mammalian amino acid transporters known so far [40] was carried out in the Aspergillus genomes. Our findings indicate six Aspergillus transporters SlcA-G (Table 11.1B) that resemble those of mammalian amino acid transporter groups or families, indicating simplified amino acid transport mechanisms compared to the multitude of transporters in mammals. The putative Aspergillus proteins SlcA-G, which show similarity to mammalian amino acid transporters, were also found in the genomes of Neurospora crassa or Magnaporthe grisea with identities of over 60%.

Blast searches in the genomes carried out with members of the human SLC7 or LAT family revealed proteins in all Aspergillus genomes, which we named SlcA. They display a 30% similarity to the human LAT1 transporter and 22% identity to a so far uncharacterized amino acid transporter of Drosophila. Similarities to any yeast protein are below 20%. LAT1 is supposed to be a ubiquitously expressed transporter acting as an environmental amino acid sensor [42,43]. It transports mainly large hydrophobic amino acids such as H, M, L, I, V, F, Y, W, and Q [40]. The mammalian transporters of the SLC7 family require the presence of glycoproteins to form a functional holotransporter [42].

The SlcB protein was only found in A. nidulans, no gene coding for a similar protein was found throughout the genomes of the other aspergilli, nor could any similar proteins be found in the investigated yeasts. The identities to Drosophila proteins was below 20%. The SlcB protein shows 33% identities to the human EAAT1 protein of the human SLC1 family [40]. Members of the SLC1 family K+ antiport glutamic acid and aspartic acid into the cells [40]. They are present in mammalian astrocytes and play an important role in astrocyte development, as was seen in the case of diseases such as lissencephaly [44–46]. Astrocytes have a star-shaped appearance that resembles the mycelia of filamentous fungi and form the brain-blood barrier in the brain [47]. They also support the exchange of nutrients from blood to the nervous system and the termination of neurotransmission by removing glutamate from the synaptic cleft.
The derived function of the SlcB protein might be involvement in (polar) growth in *A. nidulans* besides amino acid uptake [45,46].

The SlcC protein shows high similarities to the posttranscriptional regulator of nitrogen permeases Npr2p of *S. cerevisiae* and to the mammalian SLC38 family of transporters [40,48]. The SLC38 family of transporters is also described as System A transporters (SAT), which were the first mammalian membrane proteins described to be both sensor and transporter. This could be taken as a first indication of SlcC being a transporter upstream of a yet unknown signaling mechanism regulating the adaptive response to amino acids.

Only *A. nidulans* and *A. oryzae* contain SlcD proteins which resemble the SLC6 or BETA family of mammalian amino acid transporters. SLC6 transporters are sodium dependent and mainly expressed in the brain or neuronal tissues specifically in the neuronal gaps and transport GABA, glycine, and glutamate [40,49]. The flux of these GABA transporters is regulated by binding of diverse proteins to their N-terminus [50]. A search for System BETA transporter revealed homologs in human (32% identities), fly (29% identities), and *N. crassa* but not in yeast. The Imino or SLC36 family of mammalian amino acid transporters is essential for transport of small neutral amino acids (Q, N, I, L, and Y) from the lysosome after bulk degradation of proteins [51] and for uptake of nutrients from the gut [52]. The search for proteins similar to mammalian imino transporters (SLC36) resulted in two proteins for each *Aspergillus* species, namely SlcF and SlcG. Each protein shows approximately 33% identities to the human proteins of the imino group and 50%/45% identities to the yeast Avt2/3 proteins, respectively. These transporters were shown to effect or modulate growth through the TOR pathway in *Drosophila* in a yet not well-characterized way [53].

In summary, we found six different amino acid transporters (Table 11.1), which are all present in *A. nidulans*, whereas *A. oryzae* lacks SlcB and *A. fumigatus* and *A. niger* lack SlcB and SlcD. Phylogenetic analyses of the three aspergilli and *Neurospora* show that these fungi originate from a common ancestor and that *A. nidulans* and *N. crassa* separated first from the ancestor, *A. fumigatus* and *A. oryzae* separated later from each other [1]. The lack of both SlcB and SlcD in *A. niger* and in *A. fumigatus* further supports that these two aspergilli separated after the separation from *A. oryzae*.

### 11.3 Biosynthesis of Amino Acids in *Aspergillus*

When the uptake systems for amino acids and other nitrogen sources do not result in sufficient supply to fulfill all needs for growth and cell division, the fungal cell has to respond appropriately. Amino acids are not only precursors of translation but also educts for several products of the secondary metabolism. Inside the cell, several sensors have to monitor amino acid (and nitrogen) pools. One regulatory system that has been discovered in fungi and that is induced by amino acid starvation is cpc or gc system with a central master transcriptional regulator, CpcA (Gcn4). This protein regulates in the nucleus a multitude of target genes directly or indirectly involved in amino acid biosynthesis. Table 11.2 summarizes the genes for proteins involved in intracellular amino acid sensing and in amino acid biosynthesis.

#### 11.3.1 Sensing of the Intracellular Amino Acid Pool: Sensor Kinase CpcC and the TOR Pathway

During starvation on amino acids the amount of uncharged tRNAs increases—which in yeast is sensed by the sensor kinase Gcn2p/CpcC and results in CpcC phosphorylating eIF-2α—thus lowering overall translation. The sensor kinase CpcC is located at the ribosome and is present in the four investigated *Aspergillus* species with a protein sequence that is highly similar in all organisms [54]. The aforementioned effect is dependent on another ribosome-associated complex, which is Gcn1p/Gcn20p [55], both of which can be found in all three genomes. Their orthologs were named CpcD for Gcn1p and CpcE for Gcn20p. Interestingly, when compared to *A. thaliana*, where a similar regulation to the cpc is not yet described, no protein similar to CpcE can be found, whereas the ortholog of CpcD is almost identical.
<table>
<thead>
<tr>
<th>an</th>
<th>af</th>
<th>ao</th>
<th>ani</th>
<th>hs</th>
<th>sp</th>
<th>xc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpcA</td>
<td>AN3675</td>
<td>Afu4g12470</td>
<td>AO070194000006</td>
<td>ASN46539</td>
<td>ENSG0000128272</td>
<td>YEL009C</td>
</tr>
<tr>
<td>CpcB</td>
<td>AN4163</td>
<td>Afu4g13170</td>
<td>AO070342000193</td>
<td>ASN55338</td>
<td>ENSG00000204628</td>
<td>YMR116C</td>
</tr>
<tr>
<td>CpcC</td>
<td>AN2246</td>
<td>Afu5g06750</td>
<td>AO070326000127</td>
<td>ASN57263</td>
<td>ENSG0000128829</td>
<td>YSCGCN2</td>
</tr>
<tr>
<td>CpcD</td>
<td>AN5840</td>
<td>Afu2g07960</td>
<td>AO070340000004</td>
<td>ASN56731</td>
<td>ENSG0000089154</td>
<td>YGL195W</td>
</tr>
<tr>
<td>CpcE</td>
<td>AN4315</td>
<td>Afu4g06070</td>
<td>AO070230000003</td>
<td>ASN56318</td>
<td>ENSG00000161204</td>
<td>YFR009W</td>
</tr>
<tr>
<td>TorA</td>
<td>AN5982</td>
<td>Afu2g10270</td>
<td>AO070340000299</td>
<td>ASN118654</td>
<td>ENSG0000198793</td>
<td>YIR066W</td>
</tr>
<tr>
<td>TapA</td>
<td>AN0120</td>
<td>Afu5g11780</td>
<td>BA52890</td>
<td>ASN123042</td>
<td>ENSG0000089289</td>
<td>YMR028W</td>
</tr>
<tr>
<td>eIF2α</td>
<td>AN3156</td>
<td>Afu3g13480</td>
<td>AO070236000036</td>
<td>ASN56411</td>
<td>ENSG00000134001</td>
<td>YKRO26C</td>
</tr>
<tr>
<td>eIF2β</td>
<td>AN2992</td>
<td>Afu3g08600</td>
<td>AO070337000126</td>
<td>ASN52431</td>
<td>ENSG00000125977</td>
<td>YPL237W</td>
</tr>
<tr>
<td>eIF2γ</td>
<td>AN4470</td>
<td>Afu4g07580</td>
<td>AO070305000129</td>
<td>ASN57072</td>
<td>ENSG0000130741</td>
<td>YER025W</td>
</tr>
<tr>
<td>eIF2-Bα</td>
<td>AN0167</td>
<td>Afu5g11340</td>
<td>AO070321000079</td>
<td>ASN56808</td>
<td>ENSG0000111361</td>
<td>YKRO26C</td>
</tr>
<tr>
<td>eIF2-Bβ</td>
<td>AN1344</td>
<td>Afu1g09450</td>
<td>AO070247000020</td>
<td>ASN38610</td>
<td>ENSG0000119718</td>
<td>YLR291C</td>
</tr>
<tr>
<td>eIF2-Bγ</td>
<td>AN0978</td>
<td>Afu1g16660</td>
<td>AO070341000131</td>
<td>ASN131653</td>
<td>ENSG0000070785</td>
<td>YOR260W</td>
</tr>
<tr>
<td>eIF2-Bδ</td>
<td>AN6864</td>
<td>Afu5g13040</td>
<td>AO070314000084</td>
<td>ASN201740</td>
<td>ENSG0000115211</td>
<td>YGR083C</td>
</tr>
<tr>
<td>SprA</td>
<td>AN2142</td>
<td>Afu2g16090</td>
<td>AO070343000188</td>
<td>ASN183047</td>
<td>ENSG0000025800</td>
<td>YOR190W</td>
</tr>
<tr>
<td>KapA</td>
<td>AN0906</td>
<td>Afu1g15720</td>
<td>AO070320000082</td>
<td>ASN52336</td>
<td>ENSG0000108424</td>
<td>YLR347C</td>
</tr>
<tr>
<td>HisA</td>
<td>AN3748</td>
<td>Afu7g04500</td>
<td>AO070309000072</td>
<td>ASN125862</td>
<td>ENSG0000125862</td>
<td>YER055C</td>
</tr>
<tr>
<td>HisB</td>
<td>AN5636</td>
<td>Afu6g04700</td>
<td>AO070270000048</td>
<td>ASN48612</td>
<td>ENSG0000128829</td>
<td>YOR202W</td>
</tr>
<tr>
<td>HisC</td>
<td>AN0797</td>
<td>Afu1g14570</td>
<td>AO070239000020</td>
<td>ASN196027</td>
<td>ENSG0000145191</td>
<td>YOR202W</td>
</tr>
<tr>
<td>HisD</td>
<td>AN7044</td>
<td>Afu4g04030</td>
<td>AO070282000059</td>
<td>ASN185086</td>
<td>ENSG0000145191</td>
<td>YLR347C</td>
</tr>
<tr>
<td>HisE</td>
<td>AN0717</td>
<td>Afu1g14090</td>
<td>AO070343000449</td>
<td>ASN124008</td>
<td>ENSG0000112408</td>
<td>YLR291C</td>
</tr>
<tr>
<td>HisHF</td>
<td>AN7430</td>
<td>Afu2g06230</td>
<td>AO070229000001</td>
<td>ASN52498</td>
<td>ENSG0000128272</td>
<td>YBR248C</td>
</tr>
<tr>
<td>LysA</td>
<td>AN2873</td>
<td>Afu3g11710</td>
<td>AO070338000140</td>
<td>ASN197502</td>
<td>ENSG0000128272</td>
<td>YIR034C</td>
</tr>
<tr>
<td>LysB</td>
<td>AN5610</td>
<td>Afu4g11240</td>
<td>AO070301000010</td>
<td>ASN57361</td>
<td>ENSG0000128829</td>
<td>YBR115C</td>
</tr>
<tr>
<td>LysD</td>
<td>AN8519</td>
<td>Afu2g13630</td>
<td>AO070236000016</td>
<td>ASN175419</td>
<td>ENSG0000128829</td>
<td>YGL202W</td>
</tr>
<tr>
<td>LysE</td>
<td>AN5206</td>
<td>Afu6g07390</td>
<td>AO070199000006</td>
<td>ASN56319</td>
<td>ENSG0000128829</td>
<td>YOR202W</td>
</tr>
<tr>
<td>LysF</td>
<td>AN6521</td>
<td>Afu5g08890</td>
<td>AO070270000019</td>
<td>ASN35302</td>
<td>ENSG0000128829</td>
<td>YIR034C</td>
</tr>
<tr>
<td>LysG</td>
<td>AN5601</td>
<td>Afu4g11340</td>
<td>AO070328000003</td>
<td>ASN123525</td>
<td>ENSG0000128829</td>
<td>YLR291C</td>
</tr>
<tr>
<td>AcvA</td>
<td>P27742</td>
<td>—</td>
<td>AO070332000129</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>IpA</td>
<td>AN2622</td>
<td>—</td>
<td>AO070332000130</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

continued
<table>
<thead>
<tr>
<th>Gene</th>
<th>Anno</th>
<th>Accession</th>
<th>Homo sapiens</th>
<th>Aspergillus nidulans</th>
<th>Schizosaccharomyces pombe</th>
<th>S. cerevisiae</th>
</tr>
</thead>
<tbody>
<tr>
<td>AatA</td>
<td>AN2623</td>
<td>AO070332000131</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TrpA</td>
<td>AN3695</td>
<td>Afu6g12580</td>
<td>AO070341000319</td>
<td>ASN207673</td>
<td>—</td>
<td>SPCC1442.09</td>
</tr>
<tr>
<td>TrpB</td>
<td>AN6231</td>
<td>Afu2g13250</td>
<td>AO070304000084</td>
<td>ASN120463</td>
<td>—</td>
<td>SPAC19A8.15</td>
</tr>
<tr>
<td>TrpC</td>
<td>P06531</td>
<td>Afu1g13090</td>
<td>AO070343000580</td>
<td>ASN52667</td>
<td>—</td>
<td>SPBC1539.09c</td>
</tr>
<tr>
<td>TrpD</td>
<td>AN3634</td>
<td>Afu4g11980</td>
<td>AO070328000078</td>
<td>ASN51886</td>
<td>—</td>
<td>SPBC16G5.08</td>
</tr>
<tr>
<td>AroC</td>
<td>AN6866</td>
<td>Afu5g13130</td>
<td>AO070314000070</td>
<td>ASN49433</td>
<td>—</td>
<td>SPAC16E8.04c</td>
</tr>
<tr>
<td>AroH</td>
<td>AN6338</td>
<td>Afu2g13630</td>
<td>AO070236000016</td>
<td>ASN52238</td>
<td>—</td>
<td>SPAC56E4.03</td>
</tr>
<tr>
<td>TyrA</td>
<td>AN5959</td>
<td>Afu2g10450</td>
<td>AO070340000312</td>
<td>ASN173921</td>
<td>—</td>
<td>SPCC1494.04c</td>
</tr>
<tr>
<td>PhaA</td>
<td>AN11135</td>
<td>Afu5g05690</td>
<td>AO070271000023</td>
<td>ASN57302</td>
<td>—</td>
<td>SPBC30D10.16</td>
</tr>
<tr>
<td>LacA</td>
<td>AN08072</td>
<td>Afu1g14660</td>
<td>AO090003000489</td>
<td>ASN36075</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
The TOR pathway is known to be another sensor of amino acid availability. Whereas in *S. cerevisiae* the two redundant TOR kinases, Tor1p and Tor2p, can be found, in each of the investigated aspergilli only one gene for a TOR kinase, TorA, is present. The gene product resembles *S. cerevisiae* TOR proteins in the database with 48%, to *N. crassa* TOR 58% similarity, 54% similarities to the TOR proteins of *S. pombe* and *D. melanogaster*, 44% similarity to the single human TOR, and 42% similarity to the *A. thaliana* protein. The TOR kinase is known to interfere with gc/cpc in different ways. Inactivation of the TOR pathway in yeast results (among other processes) in a global translation repression. In addition, nitrogen starvation negatively affects the stability of high-affinity amino acid transporters in the cell membrane [56]. If nitrogen is plentiful, the TOR kinase phosphorylates Gcn2p and thus inhibits the protein kinase domain and binding of tRNA to Gcn2p [56]. Tap42p is a mediator of the TOR pathway that interacts with TOR and phosphatases and can also be found in each of the aspergilli. Tap42 is complexed with protein phosphatase 2A(-like) holoenzymes in various organisms [57] and is known to dephosphorylate Gcn2p thus inducing Gcn4p transcription when the TOR kinase is inactive [58,59]. The Tap42p ortholog TapA can be found in all three aspergilli; remarkably the similarity to the orthologs of higher eukaryotes lies under 30% whereas the fungal orthologs display similarities of 43.5% for *N. crassa*, 37% for *S. pombe*, and 34% for *S. cerevisiae*.

11.3.2 Cross-Pathway Control (CPC) System of Aspergillus

11.3.2.1 Global Transcription Factor CpcA

The transcription factor CpcA is controlling many genes involved in the biosynthesis of amino acids. In addition, CpcA was shown to play an important role in *A. nidulans* sexual development. If CpcA is overexpressed or under amino acid limitation, sexual development is arrested at the microcleistothecial stage before meiosis [60]. In *A. fumigatus* CpcA contributes significantly to the virulence of the fungal pathogen in mice. CpcA deletion strains of *A. fumigatus* display attenuated virulence in a neutropenic murine model of pulmonary aspergillosis [61]. The amount of CpcA within the cell is generally controlled at the level of CpcA biosynthesis as well as at the level of protein stability (compare Chapter 3). Due to the special structure of the promoter of the *cpcA* gene of *Aspergillus*, the translation of the *cpcA*-mRNA is increased under starvation conditions resulting in more CpcA protein (see Fig. 11.2) [62]. CpcA is able to positively auto-regulate its own transcription, as well as the transcription of amino acid biosynthesis genes through CpcA recognition elements (CPRCEs) under starvation conditions [62]. The orthologs of CpcA are well characterized [61–63]. Similarities are mainly found in the C-terminal leucine zipper region of CpcA, which is responsible for DNA binding. The basic leucine zipper motif is conserved throughout all organisms from fungus to man and can be found in the human activating transcription factor 2 (ATF2) and ATF4, which play similar roles as CpcA and Gcn4p [10,64]. The basic leucine zipper is also present in the c-Jun, JunD, and Fos-family of proteins; thus these proteins are counted as the c-Jun-like family. The leucine zippers, though consisting of 1–5 leucine residues, are fully interchangeable and restore functionality in complementation experiments [65].

11.3.2.2 Control of the CpcA-mRNA Amount

The *cpc* control of intracellular amino acid pools is reflected by the amount of CpcA within the cell. Part of this regulation includes the control of translation initiation of the mRNA for the central transcription factor CpcA [62]. For initiation of translation the eukaryotic ribosomal preinitiation complex scans mature mRNA toward the 3’-end. On arrival at the AUG Start codon the 80s ribosomes finally assemble [66]. The translation machinery needs phosphorylated eIF-2γ for initiation of translation and assembly of the ribosome. eIF-2γ is subsequently dephosphorylated during translation and it needs to be recycled by the guanine nucleotide exchange factor GEF (eIF-2B) to regain the competence for translation initiation [66]. By phosphorylation to further activate eIF-2γ the translation is repressed due to stoichiometrically decreasing GEF and a subsequent lack of the initiation factor [66–68]. Each of the components of the initiation factor eIF-2 share high similarities to the components of the initiation factors of other organisms. Among the aspergilli the components of the involved factors can be found easily, indicating a similar
The Aspergilli

mechanism of translation control as in yeast. The alpha, beta, and delta subunits of the GEF eIF-2B are well conserved between all compared organisms. The epsilon subunit is better conserved in the investigated fungi than in man, fly, and plant, with >42% similarities to its fungal counterparts and < 30% to the other eukaryotes. The gamma subunit of eIF-2B is not too well conserved at all, though orthologs can be found in all compared organisms, the highest similarity is shared with *N. crassa* followed by *H. sapiens*.

The eukaryotic translation initiation factor eIF3 functions by interacting with eIF2 and stabilizing the interaction between the ternary complex (composed of eIF2-GTP-Met-tRNA) and the 40S ribosomal subunit, thereby forming the 43S ribosomal complex [69]. Twelve well-conserved subunits of eIF3, eIF-3(a-l) can be found in the three aspergilli, whereas only six are present in the yeast *S. cerevisiae* and *A. nidulans* [89]. The respective *S. cerevisiae* homologs to the *Aspergillus* proteins are shown in brackets [178]. (From Hinnebusch, A.G. et al., Ann. N. Y. Acad. Sci., 1038, 60, 2004.)

**FIGURE 11.2** Translational control of the gc/cpc regulator mRNA CpcA. *Note:* Amino acids are either available or fungal cells starve for amino acids. Under nonstarvation conditions CpcC is not active. During amino acid limitation, uncharged tRNA molecules are sensed by CpcC in cooperation with GcnA/GcnB. CpcC then phosphorylates the subunit of the eukaryotic initiation factor eIF2. Phosphorylated eIF2 inhibits the eIF-2B complex, which acts as a guanine nucleotide exchange factor for eIF2. Inhibition of eIF-2B results in a global down-regulation of translation due to lower amounts of charged tRNAs and GTP. The low amount of ternary complexes under starvation conditions leads to delayed reinitiation of the reassembled ribosome. This in turn leads to increased translation of the CpcA ORF instead of the uORFs under starvation conditions. The translation factor CpcA binds cis-elements in the promoters of gc/cpc controlled genes (CPREs), enhances their transcription, and exhibits a transcriptional auto-regulation [62]. As an example for secondary metabolism it was shown that amino acid starvation favors amino acid biosynthesis and decreases penicillin biosynthesis in *A. nidulans* [89].
only nine in *S. pombe*. Nearly no conservation can be found for subunit eIF3j to higher eukaryotes, but this subunit is highly conserved to its counterpart of *N. crassa*. The eIF3, the lid of the proteasome, and the Cop9 signalosome (see Chapter 3) are regulatory multiprotein complexes whose components can be characterized through the specific PCI (proteasome, Cop9, eIF3) or MPN (Mpr1p, Pad1 N-terminal) protein domains [70–72].

### 11.3.2.3 Transport of the Transcription Factor into the Nucleus

Nuclear import is essential for CpcA/Gcn4p to induce transcription into its target genes in response to amino acid starvation, glucose starvation, and other stresses [73–76]. For nuclear import of Gcn4p of *S. cerevisiae* some key players are known [77], whereas the transport of CpcA as counterpart of *A. nidulans*, which also needs to be transported to the nucleus, is yet unexplored. It was shown by Pries et al. (2004) that the α-importin Srp1p and the β-importin Kap95p act as a heterodimer to channel Gcn4p into the yeast nucleus via the nuclear pores. Blast searches in the now available *Aspergillus* and *N. crassa* genomes revealed that proteins with high identities to Srp1p (over 58%) and Kap95p (over 39%) can be found to be encoded in the available *Aspergillus* and *N. crassa* genomes. The assigned ORF for the Srp1p homolog SrpA was annotated as AN2142.1 and the homolog to the Kap95p homolog KapA was annotated as AN0906.1 during the automated and manual annotation of the *A. nidulans* genome. Blast searches in the available *A. nidulans* genome revealed that there is only 1 α-importin-like protein and 12 members of the importin beta superfamily present, which suggests that the *Aspergillus* genomes harbor one less β importin member than yeast.

### 11.3.2.4 Ribosomal CpcB Component Represses Amino Acid Biosynthesis

CpcB of *A. nidulans* is a Gβ-like protein homologous to the mammalian RACK1 repressing the transcription of CpcA under nonstarvation conditions [78]. The protein seems to be constitutively expressed in *A. nidulans*, whereas it is transcriptionally up-regulated in yeast when glucose is present as the sole carbon source [78,79]. The yeast homolog Cpc2p was shown to interact directly with the 40S subunit of the ribosome providing a platform for other ribosome-bound proteins during translation with their propeller-like WD40 repeats at the mRNA exit site of the ribosome [80–82]. RACK1 and its homologs are found to be highly conserved in all organisms investigated. Interestingly the third intron of *A. nidulans* *cpcB* is conserved in the *S. cerevisiae* CPC2 and *N. crassa* cpc-2 genes and harbors the U24 small nucleolar RNA (snoRNA) coding region [78,83]. The U24 snoRNA is required for site specific 2′-O-methylation of 25S rRNA [84]. Though the coding region can be found in the ascomycete *A. nidulans*, its deuteromycete family members do not harbor the coding region of the U24 snoRNA anywhere near *cpcB*.

### 11.3.3 Examples for the Synthesis of Amino Acids and Derivatives

Since all three investigated aspergilli are able to grow without amino acid supplementation they should be able to synthesize all amino acids in contrast to mammals where numerous amino acids are essential. The biosyntheses of 19 amino acids can be divided into five groups according to the substrates of primary metabolism that are used: the glutamate group, the aspartate group, the pyruvate group, the serine/glycine group, and the aromatic amino acid group. In addition, histidine, which is closest to the RNA world, is a derivative of the pentose-phosphate pathway (Fig. 11.3). Depending on the organism, lysine biosynthesis can be accomplished in the following two different ways: the diaminopimelate way or the α-aminoadipate way. Fungi use the α-aminoadipate way to produce lysine. The biosynthesis of aromatic amino acids is the most energy intensive of all amino acid biosyntheses. We had a closer look at the biosynthesis of five amino acids and some additional derivatives of these amino acids: the biosynthesis of histidine, the biosynthesis of lysine and penicillin as derivatives, and the formation of the aromatic amino acids tryptophan and the derivative terrequinone A, tyrosine, and phenylalanine. For the ease of reading, the respective names of the *S. cerevisiae* proteins have been added in brackets to their respective *Aspergillus* homologs. Each of these biosynthetic pathways is regulated by CpcC/Gcn2p mediated control by the transcription factor CpcA/Gcn4p.
11.3.3.1 Histidine Biosynthesis

Histidine is the only amino acid, which is not synthesized in an amino acid family but is connected to the synthesis of nucleotides. Histidine is produced by the following six biosynthetic enzymes: HisA (His1p), HisB (His3p), HisC (His4p), HisD (His2p), HisE (His5p), and HisHF (His7p) starting from PRPP. HisA is an ATP phosphoribosyltransferase forming Phosphoribosyl–ATP. The next two steps are accomplished by HisC (trifunctional histidinol dehydrogenase) forming first phosphoribosyl–AMP, then phosphoribosylformimino AICAR-P, the branch point to the purine metabolism. HisHF forms imidazolglycerol-3-phosphate [85,86]. The next step is taken over by HisB forming imidazolacetol-phosphate [87]. HisE then forms L-histidinol-phosphate, which is converted to L-histidinol by HisD. The last two steps are accomplished by HisC forming first L-histidinal, then L-histidine. The genes for histidine biosynthesis were found to be highly conserved amongst the four Aspergillus species.

11.3.3.2 Lysine and Penicillin Biosynthesis of Aspergilli

The amino acid lysine is produced from aspartate through the diaminopimelate (DAP) pathway in most bacteria and higher plants. In fungi, in the thermophilic bacterium Thermus thermophilus, and in several
archaeal species, lysine is synthesized by the α-aminoadipate pathway and is part of the glutamate group as derivative of α-ketoglutarate (α-KG) of the TCA cycle [88]. In *A. nidulans*, the penicillin biosynthesis pathway branches from the lysine pathway. It was suggested that upon amino acid starvation, the cross-pathway control overrules penicillin biosynthesis and favors lysine production [89].

Lysine biosynthesis starts with homocitrate, which is converted by LysF (Lys4p) to homo-isocitrate, which is then turned over by LysE (Lys12p) to 2-oxoadipate. LysD then forms L-2-aminoadipate from 2-oxoadiapate. L-2-aminoadipate-6-semialdehyde is then formed by LysB (Lys2p) from L-2-aminoadipate. The saccharopine dehydrogenase LysG (Lys9p) then forms L-saccharopine, which is then converted to L-lysine by LysA (Lys1p). All three aspergilli have the potential to express all necessary enzymes. *A. oryzae* shows expansions of genes belonging to gene families predicted to play roles in metabolism [90]. BAT1 and BAT2, which contribute to hydrophobic amino acids lysine and serin, are overrepresented compared in a cluster of orthologous group (COG) classifications to other aspergilli and *S. cerevisiae*.

The enzymes of the penicillin biosynthesis, which require L-2-aminoadipate as one of three substrates for secondary metabolism, are only present in *A. nidulans* and *A. oryzae*. Penicillin is synthesized in a three-step reaction starting from L-2-aminoadipate mediated by the enzymes AcvA, IpnA, and AatA, leading to the assumption that *A. oryzae* is able to produce penicillin, as well as *A. nidulans*, whereas *A. fumigatus* and *A. niger* lack the necessary enzymes.

11.3.3.3 Aromatic Amino Acid and Terrequinone A Biosynthesis

The biosynthetic cascade resulting in the aromatic amino acids L-phenylalanine, L-tyrosine, and L-tryptophan links carbohydrate metabolism to biosynthesis of aromatic compounds. Whereas animals are only able to form tyrosine by hydroxylation of phenylalanine and, therefore, require this amino acid together with tryptophan in their diet, bacteria, plants, and fungi are competent to synthesize all three aromatic amino acids *de novo* [91]. The shikimate pathway leads to the formation of chorismate, the last common intermediate of the three pathways [92]. The chorismate branch point divides into the tryptophan biosynthetic branch and the biosynthesis of prephenate, which then divides into the biosynthetic branch of phenylalanine and tyrosine production. All four investigated *Aspergillus* genomes encode the enzymes necessary to produce the latter amino acids.

11.3.3.3.1 Tryptophan

The intermediate chorismate is transformed in a two-step reaction to anthranilate by the anthranilate synthase TrpA (Trp2p) and the trifunctional glutamine amidotransferase/N-(5′-phosphoribosyl)anthranilate isomerase/indole-3-glycerol transferase [93]. This situation is different from the yeast *S. cerevisiae* where the isomerase domain of the trifunctional enzyme is encoded by a separate gene (TRP1) that might be the result of a rearrangement event [94]. TrpD (Trp4p), the phosphoribosyl transferase, then forms N-(5-phospho-b-D-ribosyl)-anthranilate, which is then subsequently transformed to 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose-5-phosphate and (3-indolyl)-glycerol phosphate by TrpC (Trp3p) [95]. The final step, the formation of tryptophan, is taken over by TrpB (Trp5p) the phosphoribosyl transferase [96].

11.3.3.3.2 Tyrosine and Phenylalanine

After the formation of prephenate from chorismate by AroC (Aro7p) [97], the chorismate mutase, the phenylalanine/tyrosine-specific branch of the aromatic amino acid biosynthesis branches into two alternative routes to form the end products [98]. One proceeds via the formation of phenylpyruvate by PhaA (Pha1p), a dehydratase followed by transamination by AroH (Aro8/9p) to phenylalanine [99]. On the other hand 4-hydroxyphenylpyruvate is formed by TyrA (Tyr1p), which is subsequently transaminated by AroH to tyrosine [99].

11.3.3.3.3 Tryptophan as Precursor of Terrequinone A

Aspergilli and many other fungal species produce secondary metabolites, which are often bioactive, as amino acids usually of low molecular weight and often are produced as families of related compounds.
These compounds are only produced at restricted parts of the life cycle [100]. Secondary metabolite synthetic genes are usually clustered in fungal genomes whereas genes involved in primary metabolism are scattered throughout the fungal genome [101]. This fact can be used for genome mining and the identification of transcriptionally active gene clusters in Aspergillus. LaeA is a nuclear protein involved in global regulation of secondary metabolite gene clusters in this genus. LaeA is well conserved in all investigated aspergilli. The comparison of a LaeA mutant and the wild type in a microarray assay allows the identification of the penicillin antibiotic gene cluster (see earlier) as well as of the toxic sterigmatocystin cluster [102]. The same assay revealed the terrequinone A biosynthetic gene cluster and predicted a possible biosynthetic pathway [100]. Terrequinone A is a fungal bisindolylquinone with inhibitory properties on tumor cell lines [103], which was unknown to be produced by aspergilli. The terrequinone A biosynthetic gene cluster (tdi cluster) comprises five open reading frames, transcriptionally regulated by LaeA. Matching the chemical structure of terrequinone A to the tdi cluster explains the absence of a condensation domain within the TdiA enzyme, as no amide bond has to be closed, and implicates a speculative, yet plausible, order for the key biosynthetic events: (1) deamination of L-tryptophan to indolepyruvic acid by the transaminase TdiD; (2) activation to AMP-indolepyruvic acid by TdiA (adenylation domain), whose nonribosomal code points to an arylc acid rather than to amino acid activating function [104]; (3) dimerization of two activated indolepyruvic acid monomers to the core quinone structure, which might be accomplished by the TdiA thioesterase domain, analogous to the cyclization activity of the tyrocidine thioesterase domain [100]; and, finally, (4) the possibility of oxidoreductase TdiC playing a role in reducing the keto groups of the quinone core, perhaps to prepare it for the prenyl transfer [100]. The full metabolic pathway remains elusive and will be subject to further genetic and biochemical investigations. The tdi cluster is only present in A. nidulans. None of the corresponding genes (tdiA–tdiE) is present in any other investigated member of the genus Aspergillus.

LaeA is involved in global regulation of secondary metabolite gene clusters, but it does not extend regulation to nutrient utilization or spoC1 sporulation [102]. Although many amino acids are precursors to a multitude of secondary metabolites such as penicillin and terrequinone A, it is not completely clear whether cross-pathway control is directly involved in the regulation of secondary metabolite gene clusters. However, during amino acid starvation, cross-pathway control overrules the production of secondary metabolites such as penicillin in the favor of amino acids [89].

11.4 Amino Acids Obtained by Protein Degradation in Aspergillus

Besides uptake and biosynthesis of amino acids there is yet a third way for a living developing cell to gain amino acids. A cell can degrade unused proteins and thus recycle their amino acid components to fill up the internal amino acid pool and put these building blocks to further use. In organisms that undergo developmental differentiation, development needs to be tightly regulated spatially, as well as temporarily. A good example for differentiation is A. nidulans, which is able to produce asexual and sexual structures of a certain complexity. Another example is the degradation of the master regulator of the cpc, CpcA, which blocks sexual development under amino acid limitation and increases amino acid biosynthesis. If the stock of external amino acids is rising again, the cell needs to switch its metabolic program toward uptake of amino acids and degrade CpcA to accomplish this task.

Different ways of degrading proteins or cell compartments have been described: degradation of proteins in the vacuole or lysosomes where foreign matter or cellular compartments are bulk degraded and the degradation of cellular proteins through the ubiquitin–proteasome system where defined proteins are targeted for degradation. In general, bulk protein degradation for the sake of amino acid production has to be regarded different from the degradation of specific proteins such as CpcA to terminate its function in the cell. On the other hand even the destruction of the CpcA protein leads to replenishment of the internal amino acid pool in the end and suits as a well-understood example of protein degradation.

In this part of the chapter we focus on the possible variations of ubiquitin–ligase complexes and ubiquitin-like modifiers present in the Aspergillus genomes; the main machinery used for controlled destruction of proteins is the ubiquitin–proteasome system. Its presence was shown in a variety of eukaryotic organisms with a variety of purposes, though with the final goal to mediate destruction of...
Amino Acid Supply of Aspergillus

11.4.1 Prerequisites for Protein Degradation

The yeast transcription factor Gcn4p is a highly unstable protein with a half-life of approximately five minutes under nonstarvation conditions. A well-conserved PEST region [108] and ubiquitylation sites are responsible for the instability of this protein and its subsequent ubiquitylation by the SCF complex and degradation in the 26S proteasome [109,110]. Nothing is known about the half-life of its fungal counterparts so far. But the phosphorylation machinery that initially marks proteins for degradation is well conserved in Aspergillus. The cyclin-dependent kinases Pho85p and Srb10p are responsible for destabilization of the transcription factor in yeast [110,111]. The homologs PhoA and SrbA were found to be highly conserved in all investigated organisms including the aspergilli with the exception that the homolog of
158

The Aspergilli

Table 11.3

Genes for Protein Degradation in Aspergillus (the First Column States the Proposed Names of the Respective Proteins in A. nidulans, Which Were Derived from the Names of the Characterized Homologs)

<table>
<thead>
<tr>
<th>Name</th>
<th>an</th>
<th>af</th>
<th>ao</th>
<th>ani</th>
<th>hs</th>
<th>sp</th>
<th>sc</th>
</tr>
</thead>
<tbody>
<tr>
<td>SrbA</td>
<td>AN2489</td>
<td>Afu3g13990</td>
<td>AO070312000131</td>
<td>ASN134471</td>
<td>ENSG0000105511</td>
<td>SPAC23H4.17c</td>
<td>YPL042C</td>
</tr>
<tr>
<td>PhoA</td>
<td>AN8261</td>
<td>Afu5g04130</td>
<td>AO070310000040</td>
<td>ASN189016</td>
<td>ENSG0000123374</td>
<td>SPCC16C4.11</td>
<td>YPL031C</td>
</tr>
<tr>
<td>PhoB</td>
<td>AN4310</td>
<td>Afu4g06020</td>
<td>AO09023000995</td>
<td>ASN203823</td>
<td>—</td>
<td>YGR233C</td>
<td></td>
</tr>
<tr>
<td>SrbB</td>
<td>AN2172</td>
<td>Afu2g15790</td>
<td>AO070343000243</td>
<td>ASN56354</td>
<td>—</td>
<td>YNL025C</td>
<td></td>
</tr>
<tr>
<td>PclA</td>
<td>AN0453</td>
<td>Afu1g04750</td>
<td>AO090009000492</td>
<td>ASN55461</td>
<td>—</td>
<td>YNL289W</td>
<td></td>
</tr>
<tr>
<td>PclE</td>
<td>AN9500</td>
<td>Afu6g02020</td>
<td>AO070205000004</td>
<td>ASN53931</td>
<td>—</td>
<td>YHR071W</td>
<td></td>
</tr>
<tr>
<td>PclF</td>
<td>AN3755</td>
<td>Afu7g04640</td>
<td>AO070309000051</td>
<td>ASN57211</td>
<td>—</td>
<td>YER059W</td>
<td></td>
</tr>
<tr>
<td>PclH</td>
<td>AN5825</td>
<td>Afu2g07660</td>
<td>AO070260000023</td>
<td>ASN126937</td>
<td>—</td>
<td>YPL219W</td>
<td></td>
</tr>
<tr>
<td>PclK</td>
<td>AN4984</td>
<td>Afu3g10040</td>
<td>AO070288000052</td>
<td>ASN183571</td>
<td>—</td>
<td>YGL215W</td>
<td></td>
</tr>
<tr>
<td>SkpA</td>
<td>AN2302</td>
<td>Afu5g06060</td>
<td>AO070271000063</td>
<td>ASN214072</td>
<td>ENSG0000135558</td>
<td>SPBC409.05</td>
<td>YDR328C</td>
</tr>
<tr>
<td>HrtA</td>
<td>AN8844</td>
<td>Afu5g05790</td>
<td>AO070271000032</td>
<td>ASN192658</td>
<td>ENSG0000100387</td>
<td>SPAC23H4.18c</td>
<td>YOL133W</td>
</tr>
<tr>
<td>CdcD</td>
<td>AN5517</td>
<td>Afu6g13030</td>
<td>AO070341000037</td>
<td>ASN47323</td>
<td>ENSG0000109670</td>
<td>SPAC4D7.03</td>
<td>YFL009W</td>
</tr>
<tr>
<td>CulA</td>
<td>AN1019</td>
<td>Afu1g12960</td>
<td>AO070343000596</td>
<td>ASN55773</td>
<td>ENSG0000055130</td>
<td>SPAC17G6.12</td>
<td>YDL132W</td>
</tr>
<tr>
<td>CulC</td>
<td>AN3939</td>
<td>Afu6g08220</td>
<td>AO070341000051</td>
<td>ASN252885</td>
<td>ENSG0000036257</td>
<td>SPAC24H6.03</td>
<td>YGR003W</td>
</tr>
<tr>
<td>CulD</td>
<td>AN0037</td>
<td>Afu5g12680</td>
<td>AO070314000115</td>
<td>ASN56629</td>
<td>ENSG0000158290</td>
<td>SPAC3A11.08</td>
<td>—</td>
</tr>
<tr>
<td>CanA</td>
<td>AN2458</td>
<td>Afu6g10338</td>
<td>AO070264000026</td>
<td>ASN179224</td>
<td>ENSG0000111530</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>UbaA</td>
<td>AN2174</td>
<td>Afu2g15760</td>
<td>AO070343000247</td>
<td>ASN210285</td>
<td>ENSG0000130985</td>
<td>SPBC1604.21c</td>
<td>YKL210W</td>
</tr>
<tr>
<td>UbaB</td>
<td>AN2450</td>
<td>Afu6g10510</td>
<td>AO070264000026</td>
<td>ASN178271</td>
<td>ENSG0000125731</td>
<td>SPBC16H5.03c</td>
<td>YDR390C</td>
</tr>
<tr>
<td>UbaC</td>
<td>AN2416</td>
<td>Afu2g13730</td>
<td>AO070228000028</td>
<td>ASN174348</td>
<td>ENSG0000144744</td>
<td>SPAC24H6.12c</td>
<td>YPR066W</td>
</tr>
<tr>
<td>AosA</td>
<td>AN2298</td>
<td>Afu6g01600</td>
<td>AO070295000079</td>
<td>ASN134428</td>
<td>ENSG0000142230</td>
<td>SPAC4C5.04</td>
<td>YPR180W</td>
</tr>
<tr>
<td>UlaA</td>
<td>AN2441</td>
<td>Afu6g10600</td>
<td>AO070264000038</td>
<td>ASN25857</td>
<td>ENSG0000159593</td>
<td>SPAC323.06c</td>
<td>YPL038W</td>
</tr>
<tr>
<td>UbcL</td>
<td>AN10324</td>
<td>Afu3g14430</td>
<td>AO070286000092</td>
<td>ASN50129</td>
<td>ENSG0000130725</td>
<td>SPCC777.10c</td>
<td>YLR306W</td>
</tr>
</tbody>
</table>

SrbA cannot be found in A. thaliana. PhoA shows 81% identity to the N. crassa protein and 72% and 67% to the proteins of S. pombe and S. cerevisiae, respectively. The identity compared to the proteins of higher eukaryotes is higher than 55%. SrbA was found to be 51% identical to N. crassa and 45% and 49% identical to the respective proteins of S. pombe and S. cerevisiae. The respective proteins of Drosophila and human showed 36% and 42% identity, respectively. We could find a corresponding cyclin Srb11p, named SrbB in the aspergilli, but like in all investigated cyclins the identity to other proteins was generally low (<40%); only the proteins of N. crassa and S. cerevisiae showed identities of 47% and 42%, respectively. Under nonstarvation conditions Pcl5p is required for Pho85p-mediated Gcn4p degradation in yeast [112]. The cyclin-dependent kinase Pho85p was shown to phosphorylate Gcn4p at T165 and thus mark it for ubiquitylation and further degradation by the 26S proteasome [109]. Ten different cyclins are known to interact with the cyclin-dependent kinase Pho85p. A search throughout the genomes of the aspergilli revealed the PclA protein with relatively low homologies to the yeast cyclins Pcl1p and Pcl2p. PclL displayed 46% identity to Pho85p and 71% identities to Nuc-1 of N. crassa. PclE displayed a low identity of only 34% to Pcl5p and even lower identities of 34% and 28% to the proteins of Arabidopsis and S. pombe. The identity to the Neurospora protein is 45%. PclF resembles Pcl6p and Pcl7p with higher identity to Pcl6p (38%) and high identity to the Neurospora protein (60%). PclH shows less than 20% identity to Pcl10p, but 28% identity to Pcl8p. The Neurospora protein is 60% identical to PclH. PclL is 49% identical to Pcl9p and the respective Neurospora protein, whereas PclK shows only 26% identity to Clg1p, and 36% identity to its other fungal counterparts. Generally it has to be remarked that the cyclins are relatively weakly conserved among the investigated organisms. Comparing the Aspergillus spp. CpcAs and Gcn4p of S. cerevisiae one finds the phosphorylation site T165 well conserved through the aspergilli.
### TABLE 11.4
Putative F-Box Proteins of *A. nidulans*, Their Respective *S. cerevisiae* Orthologs Are Displayed in Brackets

<table>
<thead>
<tr>
<th>A. High-scoring F-box proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-box 1 AN6086 (Hrt3p)</td>
</tr>
<tr>
<td>F-box 2 AN6816</td>
</tr>
<tr>
<td>F-box 3 AN6183</td>
</tr>
<tr>
<td>F-box 4 AN6552</td>
</tr>
<tr>
<td>F-box 7 AN9249</td>
</tr>
<tr>
<td>F-box 8 AN9353</td>
</tr>
<tr>
<td>F-box 9 AN10061 (Rcy1p)</td>
</tr>
<tr>
<td>F-box 10 AN9460</td>
</tr>
<tr>
<td>F-box 11 AN9557 (Ufo1p)</td>
</tr>
<tr>
<td>F-box 13 AN2348</td>
</tr>
<tr>
<td>F-box 14 AN2353</td>
</tr>
<tr>
<td>F-box 15 AN2505</td>
</tr>
<tr>
<td>F-box 16 AN2636</td>
</tr>
<tr>
<td>F-box 17 AN2861</td>
</tr>
<tr>
<td>F-box 18 AN4488</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Low scoring F-box proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-box 63 AN2540</td>
</tr>
<tr>
<td>F-box 64 AN2808</td>
</tr>
<tr>
<td>F-box 65 AN3371</td>
</tr>
<tr>
<td>F-box 66 AN5034</td>
</tr>
<tr>
<td>F-box 67 AN5576</td>
</tr>
<tr>
<td>F-box 68 AN2183</td>
</tr>
<tr>
<td>F-box 69 AN6634</td>
</tr>
<tr>
<td>F-box 70 AN4548 (Ydr306c)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Other homology group</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-box 5 AN7956</td>
</tr>
<tr>
<td>F-box 6 AN9187</td>
</tr>
<tr>
<td>F-box 29 AN4200</td>
</tr>
<tr>
<td>F-box 30 AN9059</td>
</tr>
<tr>
<td>F-box 60 AN6887</td>
</tr>
<tr>
<td>F-box 61 AN7964</td>
</tr>
<tr>
<td>F-box 62 AN8128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D. Fungal-specific homology group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-box 12 AN2029</td>
</tr>
<tr>
<td>F-box 21 AN5509</td>
</tr>
<tr>
<td>F-box 32 AN2837</td>
</tr>
<tr>
<td>F-box 53 AN5263</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E. Fungal-specific homology group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-box 54 AN1705</td>
</tr>
<tr>
<td>F-box 55 AN5941</td>
</tr>
<tr>
<td>F-box 56 AN6152</td>
</tr>
<tr>
<td>F-box 57 AN6944</td>
</tr>
<tr>
<td>F-box 58 AN8051</td>
</tr>
<tr>
<td>F-box 59 AN10347</td>
</tr>
</tbody>
</table>
It is notable that the phosphorylation site is exchanged from threonine to serine in the aspergilli compared to yeast and that the site is not conserved in *N. crassa*. Phosphorylated Gcn4p is subsequently ubiquitylated by the SCF\(^{Cdc4}\) ligase, which targets it to degradation by the 26S proteasome, whereas the SCF complex is constitutive and Gcn4p stability is subject to the phosphorylation state of Gcn4p mediated by Pho85p [108,109].

### 11.4.2 Ubiquitylation of Phosphorylated Substrates

Once a protein is phosphorylated by kinases at specific residues ubiquitin is ligated to the proteins and thus the protein is finally marked for degradation. The ubiquitylation cascade has been studied in depth in general, but nothing is known so far in *Aspergillus* [113–115]. An ubiquitin activating enzyme (E1) activates ubiquitin by a thioester linkage prior to transfer to a lysine residue of a specific protein substrate by an ubiquitin conjugating enzyme (E2). An ubiquitin ligase (E3) catalyzes substrate recognition and assists in ubiquitylation. Posttranslational covalent attachment of chains of ubiquitin or ubiquitin-like proteins (UBLs) was found to play a major regulatory role in cell life in different tasks ranging from cell cycle to protein degradation, development, and signal transduction [116]. UBL conjugation cascades are initiated by the activation of the UBL by dedicated E1 activating enzymes. First, the E1 selects the respective UBL for the pathway and binds it noncovalently by adenylation. In a second step the E1 catalytic cysteine attacks the adenylate and forms a thioester bond with the UBL C-terminus [115,117]. In a next step another UBL is loaded onto the E1 by adenylation to facilitate transfer of the thioester-bound UBL to the E2 enzyme [115]. Subsequently, E1 binds to E2 and transfers the thioester bond UBL to E2 upon which the E1-E2 protein–protein interaction diminishes [115,118]. A variety of different E3 enzymes is known and will be discussed later.

### 11.4.3 Ubiquitin Ligases

Several E3 ubiquitin ligase complexes are known, most prominent is the cullin-RING-H2-family that includes SCF (Skp1p/Hrt1p/Cdc53p/F-box) [119–121], which plays an important role in the regulated destruction of the central transcription factor of the general control of amino acid biosynthesis Gcn4p of *S. cerevisiae* [110] and as our results indicate also of CpcA (Table 11.3). The E3 ubiquitin ligase specific for Gcn4p consists of the cullin Cdc53p, Skp1p, binding the specific F-box protein (Table 11.4), the RING-H2 protein Hrt1p, and the specificity protein Cdc4p [110]. We were able to identify homologs of these proteins in all three aspergilli under the names CulA, SkpA (also known as SconC [122]), HrtA, and CdcD, respectively. SkpA or SconC was earlier characterized as a Skp1p-like protein interacting with the F-box protein SconB as negative regulator of sulfur-metabolism in *A. nidulans* [122]. The cullin and the specificity protein CdcD are highly conserved among all organisms with overall identities of >56% for the cullin and >40% identities among compared F-box proteins. We were unable to find a homolog of CdcD in plants. The *Aspergillus* genomes also revealed Rub1p/NEDD8 homolog with identities >50% to the proteins of the other compared organisms. Our results lead us to the conclusion that CpcA, like Gcn4p, is degraded through a similar mechanism as Gcn4p. First the protein is phosphorylated, then tagged by ubiquitin, and subsequently degraded in the 26S proteasome. Thus even the major transcription factor of the gc/cpc serves as an example for all degraded proteins: in times when proteins are not needed any more their basic components are channeled back into the cellular pool of amino acids by controlled destruction.

Another prominent example of ubiquitin E3 ligases is the HECT-type E3s. The HECT E3s are simple polypeptides that were originally identified by the presence of a conserved 350 amino acid C-terminal HECT domain (*Homology to E6-AP C-Terminus*) [113]. This domain is able to bind an E2 enzyme and transfer the thioester-linked ubiquitin to the respective bound target protein [123]. A third group of E3 ligases is the RING/U-box ligases defined by a domain that binds zinc in a RING-H2 or RING-HC arrangement and shows potential protein-binding domains N-terminal to the RING motif. COP1 of Arabidopsis serves as a prominent example of such ligases [124]. The APC (anaphase promoting complex) is regarded as another ubiquitin E3 ligase playing a role in the control of cyclins and checkpoint regulators during eukaryote cell cycle [125,126].
All Cullin-RING ligases (CRLs) share a common appearance and are found throughout all eukaryotes. Differently composed CRLs regulate a wide variety of cellular processes with thousands of potential targets [105,127–129]. A cullin and a RING-finger protein serve as the core of each CRL [15]. Seven cullins are so far known from man and at least two proteins, APC2 and PARC [126,130], are known that share the cullin-homology domain [15]. The cullin repeat motif consists of two short (A and B) and three long helices (C–E). Helices C–E are arranged in a helix bundle, with helices A and B adjacent to the N-terminal part of the bundle. This N-terminal region also harbors a domain that is nonconserved between the cullins and much shorter in cullin 4 [131]. It is expected to bind either directly to substrate receptors or adapter proteins for substrate receptors. The C-terminal part of the cullin consists of a globular domain. It is known as the cullin-homology region and can be found in other cullin-like proteins as APC2 and PARC, as well [131]. The RING-finger protein, which is better described as zinc-binding RING-H2-domain protein, binds to the globular domain of cullins and recruits the respective E2 enzymes [15,131]. The globular domain also harbors the neddylation site in close proximity of the RING protein [131]. Substrate receptors or their adaptors bind for receptors of those bind via a common BTB or POZ (bric-a-brac, tramtrack and broad complex transcription regulators) domain to the N-terminus of specific cullins. Members of the group of BTB domain proteins are Skp1p, ElonginC, and T1-Kv [132,133]. E3 ligases containing cullin 1 and 7 recruit their targets through F-box receptors, cullin 2 and 5 E3 ligases use ElonginC to bind SOCS target adaptors, and cullin 3 E3 ligases use BTB domain proteins as direct acceptors [15]. For cullin 4 E3 ligases the DDB1 protein (DNA damage-binding protein 1) binds through a nonspecified domain to cullin 4 to accept its targets directly or through specific adaptors [134,135]. The different compositions of possible E3 ligases is founded at least on one of the core components—the cullins. In Aspergillus we were able to identify putative orthologs to cullin 1, 3, and 4b, which might indicate that, similar to plants, the Aspergillus proteome does not contain SOCS or ElonginC containing cullin-RING complexes [15,123]. The presence of different E3 complexes in the Aspergillus genome needs to be investigated more in depth to find more information on possible compositions of alternating E3 complexes besides the SCF, which was focused here.

The variety of E3 targets due to possible interactions based on different cullins, E2s, E1s, and substrate adaptor proteins is vast. E1 activating enzymes are essential for selecting the different types of ubiquitin-like proteins, whereas E2 conjugating enzymes build the bridge to the respective E3 ligase with a relatively huge number of substrate-specific adaptors. F-box proteins confer substrate specificity by binding to the Skp1 subunit through their N-termini [131]. The C-termini of F-box proteins consist of variable protein interaction domains that bind respective target proteins to the SCF for subsequent poly-ubiquitylation. The number of F-box proteins can be huge, ranging from 14 F-box proteins in yeast to almost 700 in Arabidopsis [123]. In A. nidulans we were able to identify 70 putative F-box proteins in silico, which is a number comparable to the number of F-box proteins so far found in the human genome. Since the general conservation of F-box proteins is low, there might be an even greater number of these proteins to be found in the future (see Fig. 11.5, Table 11.4).

The final step after acquisition of the target proteins by the SCF is to add ubiquitin chains to the target for subsequent proteolysis at the 26S proteasome. Ubiquitins and ubiquitin-like proteins differ significantly in size, but are characterized through the shared Ub-fold. The C-terminus of Ub-like proteins is characterized through a glycine carboxyl-group and provides the site needed for linking to target proteins. The binding site of Ub-like proteins to a target protein is usually an ε-amino group of an accessible lysine [136]. The transfer mechanism for ubiquitin, NEDD8, and SUMO from E1 to target protein is similar to the mechanism described earlier; the transfer of other polypeptide tags is not as well understood yet or is generally different [115,136,137]. So far we have been able to identify 12 ubiquitin or ubiquitin-domain proteins in Aspergillus, which can be loosely grouped into the following groups: ubiquitin, Rub/NEDD8, SUMO, autophagy tags, HUB1-like tags, and a URM-like tag [116,136]. We were able to assign five potential monomeric or heterodimeric E1 activating enzymes with significant similarities to their orthologs in other eukaryotes. We found an Uba1p ortholog, UbaA, which is responsible for ubiquitin activation and in silico results indicate that the heterodimer UbaB/AosA (Uba2p/Aos1p) is the SUMO activating enzyme. Attachment of SUMO to target proteins leads to multiple effects from crossing of the nuclear envelope or activation of transcription factors to counteracting ubiquitylation effects [138]. Of special interest in this context is also the heterodimer formed by UbaC and UlaA; the orthologous proteins
The Aspergilli to Homo sapiens are Uba3 and APPBP1 respectively, which interact with the Ubc12 ortholog UbcL (E2) to ligate NEDD8 to the cullin 1 subunit of the Aspergillus SCF-complex. High conservation was observed for 16 E2 enzymes identified in the Aspergillus genomes, though they partly seem to exhibit high conservation to different E2 enzymes of various compared species.

11.4.4 SCF Activity Is Controlled by Alternating Neddylation Status

The activity of the SCF complexes is regulated through covalent modification of the CUL1 subunit through attachment of the ubiquitin-like peptide Nedd8/Rub1. Covalent binding of Nedd8 or Rub1 to the cullin 1 subunit of the SCF, or neddylation, occurs through a pathway very similar to the ubiquitin ligating pathway: it is catalyzed by an enzymatic cascade involving Nedd8-activating enzymes APPBP1 and Uba3 (E1) and the conjugating enzyme Ubc12 (E2) [113,139]. The neddylation pathway is essential in yeast, worm, and mouse and plays a significant role in auxin response in plant [140–142].

Removal or deneddylation of the Nedd8 peptide from the cullin subunit is carried out by the metalloprotease activity of the CSN5 subunit of the COP9 signalosome (Table 11.5); interestingly noncomplexed CSN5 subunits do not exhibit this catalytic activity [143]. The CSN was found to bind to CUL1 and Rbx1 via CSN2, CSN6, and the N-terminal domain of CSN1 [144–146] and promotes the SCF function in vivo. Paradoxically deneddylation of the cullin 1 subunits inhibits SCF activity in vitro [147–149]. A possible solution for this paradoxon is that the SCF tends to auto-ubiquitylation leading to degradation
### Table 11.5

Aspergillus COPI9 Signalosome Components in Comparison to Other “Zomes” (eIF3 and Lid of the Proteasome; the First Column States the Proposed Names of the Respective Proteins in *A. nidulans*, Which Were Derived from the Names of the Characterized Homologs)

<table>
<thead>
<tr>
<th>AN</th>
<th>AF</th>
<th>AO</th>
<th>AT</th>
<th>ENSP0000347251 SPBC15.03c</th>
<th>SPAC17B11.04c</th>
<th>SPAC112.16c</th>
<th>SPAC168.11c</th>
<th>YD126C</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSN1</td>
<td>AN1491</td>
<td>A1g56140</td>
<td>A1g26990</td>
<td>A1g42520</td>
<td>A1g42570</td>
<td>A1g12230</td>
<td>A1g12230</td>
<td></td>
</tr>
<tr>
<td>CSN2</td>
<td>AN783</td>
<td>A1g56730</td>
<td>A1g56830</td>
<td>A1g56830</td>
<td>A1g56830</td>
<td>A1g56830</td>
<td>A1g56830</td>
<td></td>
</tr>
<tr>
<td>CSN3</td>
<td>AN5798</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td></td>
</tr>
<tr>
<td>CSN4</td>
<td>AN1359</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td></td>
</tr>
<tr>
<td>CSN5</td>
<td>AN1219</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td>A1g42570</td>
<td></td>
</tr>
<tr>
<td>CSN6</td>
<td>AN233</td>
<td>A1g57240</td>
<td>A1g57240</td>
<td>A1g57240</td>
<td>A1g57240</td>
<td>A1g57240</td>
<td>A1g57240</td>
<td></td>
</tr>
<tr>
<td>CSN7</td>
<td>AN8323</td>
<td>A1g54430</td>
<td>A1g54430</td>
<td>A1g54430</td>
<td>A1g54430</td>
<td>A1g54430</td>
<td>A1g54430</td>
<td></td>
</tr>
<tr>
<td>CSN8</td>
<td>AN1028</td>
<td>A1g6110</td>
<td>A1g6110</td>
<td>A1g6110</td>
<td>A1g6110</td>
<td>A1g6110</td>
<td>A1g6110</td>
<td></td>
</tr>
<tr>
<td>RPN3</td>
<td>AN3037</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>RPN5</td>
<td>AN775</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>RPN6</td>
<td>AN520</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>RPN8</td>
<td>AN511</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP3</td>
<td>AN217</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP8</td>
<td>AN121</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP9</td>
<td>AN516</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP11</td>
<td>AN492</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP12</td>
<td>AN309</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3a</td>
<td>AN208</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3b</td>
<td>AN775</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3c</td>
<td>AN520</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3d</td>
<td>AN511</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3e</td>
<td>AN492</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3f</td>
<td>AN309</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3g</td>
<td>AN208</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3h</td>
<td>AN775</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3i</td>
<td>AN520</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3j</td>
<td>AN492</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3k</td>
<td>AN309</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3l</td>
<td>AN208</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>eIF3m?</td>
<td>AN775</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>AMSP1</td>
<td>AN3003</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
<tr>
<td>PRP8</td>
<td>AN142</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td>A1g6136</td>
<td></td>
</tr>
</tbody>
</table>
of SCF components themselves (Fig. 11.4). In wild-type cells the SCF-bound substrate is poly-ubiquitylated and further degraded at the 26S proteasome; during this time the cullin subunit remains neddylated [121,150]. The \textit{in silico} investigation of all components of the de-/neddylation pathway provides a first hint for their presence in \textit{Aspergillus} and indicates similar pathways as have been found in other organisms so far.

In HeLa cells it was found that the deneddylated cullin and RING-finger part of the SCF are bound by CAND1 (cullin associated and Nedd8 dissociated). This leads to dissociation of the Skp1/F-box part from the SCF and inactivation of the ubiquitin ligase [151]. During the dissociated state of the Skp1/F-box proteins these are recharged with new substrates for ubiquitylation. Ubc12 neddylates the cullin subunit, which in turn leads to dissociation of CAND1 and reassociation of the substrate-bound Skp1/F-box proteins. A defect in the deneddylation activity leads to degradation of the target protein, but leaves the neddylated SCF intact. The SCF is now subject to auto-ubiquitylation and subsequent degradation, which leads to accumulation of SCF substrates as is seen in many organisms [121,150,152,153]. This shows that the neddylation and deneddylation of cullins is a highly dynamic and important process. On the other hand, misregulation of substrate degradation leads to severe consequences for a living organism [154]. Accordingly, a CAND1 homolog CandA is present in \textit{Aspergillus}, which strengthens the hypothesis of similar de-/neddylation pathways to other organisms.

Interestingly, the CSN does not only show deneddylation activity but it also displays an ubiquitin isopeptidase activity. The CSN can, through the metalloprotease domain of CSN5, either depolymerize ubiquitin chains or de-ubiquitylate mono-ubiquitylated substrates, which suggests that the mechanisms of deneddylation and de-ubiquitylation are similar from fission yeast to human [155,156]. So far an ubiquitin isopeptidase activity has not been detected for \textit{Aspergillus}.

11.4.5 PCI Complexes

Within the broad field of protein synthesis and degradation, three multiprotein complexes deserve a special focus: the 26S proteasome lid, the CSN, and the eukaryotic translation initiation factor 3 (eIF3). Although they have different individual functions, they share the overall activity in the control of protein levels and thus control of the free amino acid pool within the cell. They also share a similar subunit architecture containing PCI and MPN homology domain proteins (see later), for which they are referred to as “PCI complexes” [72]. In contrast to yeasts, the \textit{A. nidulans} genome seems to encode a full set of subunits required for the formation of all three PCI complexes (Fig. 11.6a). All components are also found in the genomes of \textit{A. fumigatus} and \textit{A. oryzae} (Table 11.5).

11.4.5.1 PCI and MPN Domain Proteins

The PCI and MPN domains are protein-interaction domains with an average size of 140 amino acids and 200 amino acids, respectively [157,158]. Whereas the MPN domain is well conserved and easy to detect, the degree of PCI domain conservation is highly variable and rather structure based [159]. As known so far, PCI domain proteins fulfil a mere structural role as main building blocks of the complexes and are probably important for complex assembly [160,161]. By contrast, MPN domains can include the JAMM motif conferring metalloprotease activity, and are then termed MPN+ domain [158,162,163].

The \textit{Aspergillus} genome includes at least 18 annotated proteins containing a PCI domain. It also harbors the genetic information for at least 8 MPN domain proteins, of which three include the JAMM metalloprotease motif (Fig. 11.6a). All PCI proteins could be assigned to one of the PCI complexes, which is in accordance with their proposed structural role. Of the MPN domain proteins, two belong to each of the proposed three complexes. Additionally, AN3003 is related to the MPN+ protein AMSH [164] and AN4523 corresponds to Prp8, a splicing factor containing an ubiquitin-binding variant MPN domain [165].

11.4.5.1.1 Proteasome Lid

The 26S proteasome is the cellular dustbin that is required for the targeted degradation of ubiquitylated proteins. It is one of the most complex oligomeric protein structures in eukaryotic cells and consists of
Amino Acid Supply of Aspergillus

The 20S catalytic core particle and two 19S regulatory particles located at the exterior ends of the core [70,166,167]. The regulatory particle itself is subdivided into the base, a hexameric ring of AAA-ATPases, and the lid. The complex composition of 14 catalytic and 18 regulatory components, conserved from yeast to man, has recently also been shown for filamentous fungi by genome analysis of *N. crassa* [168].

As one of the three PCI complexes, the proteasome lid of higher eukaryotes has a characteristic 6+2 PCI/MPN subunit composition [70]. The only intrinsic enzyme activity of the lid, the deubiquitinase, resides within Rpn11 [162].

Although targeted protein degradation is an important issue, no special attention has been drawn on the proteasome and its lid in *A. nidulans* so far. The six PCI and two MPN domain proteins of the proposed

![FIGURE 11.6 PCI complexes of *A. nidulans*.](image-url)

Note: (a) The *A. nidulans* genome contains all PCI/MPN domain proteins of the three PCI complexes lid, CSN and eIF3. Percent amino acid identities to the corresponding sequences of human (hs), *A. thaliana* (at), *Schizosaccharomyces pombe* (sp) and *Saccharomyces cerevisiae* (sc) are given; sequence IDs are summarized in the supplementary material. Only CsnA, CsnD, CsnE, and CsnG/AcoB had been previously described for *A. nidulans* [186]. (b) The paralogous LID and CSN subunits group together in a ClustalW based tree analysis.
Aspergillus lid show very high amino acid identities to their counterparts in other eukaryotes (Fig. 11.6a). All PCI and MPN domains were identified, including the highest conserved MPN+ metalloprotease activity domain. Subunit composition of the LID is thus highly conserved from yeast to man.

11.4.5.2 COP9 Signalosome (CSN)
The CSN regulates the ubiquitylation activity of cullin-containing E3 ligases toward their protein substrates. This regulatory role is performed by the deneddylase activity of the CSN subunit five that detaches the small ubiquitin-like protein Nedd8/Rub1 from cullins [144,146]. This way, the CSN is involved in major regulatory pathways including embryonic development in mammals [169], oogenesis in fly [161], light-signaling in plant and N. crassa [121,170,171], a cell-cycle defect in S. pombe [172,173], and altered pheromone response in S. cerevisiae, respectively [174,175]. Like the LID, CSN of higher eukaryotes has a 6+2 PCI/MPN subunit composition. These two complexes not only share a common overall structure [71], but also consist of paralogous subunits with clear counterparts [159]. Remarkably, fungi do not necessarily contain a full eight-subunit CSN. The N. crassa complex lacks subunit 8 [176], fission yeast has only six subunits [173], and in bakers’ yeast subunit 5 is solely remarkably conserved [175].

In A. nidulans, deletion of CSN subunits results in multiple mutant phenotypes, most severe in the block of the sexual cycle resulting in production of primordia that never mature into fruit bodies [177]. The Aspergillus genome encodes genes for all eight CSN subunits as known in higher eukaryotes (Fig. 11.6a), and the assembly of the corresponding proteins into a complex has been identified recently (Busch et al., submitted). The A. nidulans CSN subunit composition thus more closely resembles that of human, plant, and Aspergillus group together impressively in a ClustalW-based evolutionary distance tree (Fig. 11.6b).

11.4.5.3 Translation Initiation Factor 3 (eIF3)
The eukaryotic translation initiation factor 3 (eIF3) promotes the formation of the preinitiation complex. It facilitates the loading of the 40S subunit onto the ternary eIF2-tRNA-Met-GTP complex and interacts with other translation factors [178]. The complex is composed of at least 12 subunits in higher eukaryotes, but several subunits seem to be lacking in the yeasts S. cerevisiae and S. pombe. The eIF3 complex of higher eukaryotes contains 5 PCI and 2 MPN proteins among its 12 components. Both proposed MPN subunits of eIF3 lack the JAMM motif [158,159]. Thus, it seems that in this complex both PCI and MPN subunits have a rather structural role and the non-PCI/MPN subunits fulfil the major function in translation. An additional and alternative component of the fission yeast eIF3 was described recently as eIF3m [179]. A protein with reasonable amino acid sequence identities to eIF3m was also described as the new eIF3 component HA17 in mammals [180]. These “new” eIF3 components both contain a PCI domain, suggesting that the eIF3 complex of higher eukaryotes ranks in the 6+2 PCI/MPN subunit composition of lid and CSN.

So far, no experimental data are available on the A. nidulans eIF3 complex. The Aspergillus genome encodes counterparts to all 12 well-recognized subunits and to the proposed new thirteenth subunit (Table 11.5). Thus, the eIF3 complex of A. nidulans more closely resembles that of higher eukaryotes than that of yeasts.

11.5 Conclusion
This partial analysis of genes involved in amino acid supply in the Aspergillus genomes should be regarded as a starting point to further our understanding of filamentous fungi in comparison to other higher and lower eukaryotes. Numerous data that are accumulated here are, though well funded by in silico experiments, only wild-card guesses for the real world within the fungal cells. These genes that show certain players of interconnecting pathways need to be examined further to provide physical proof. Many genes
and proteins that were investigated in this study are strongly related to yeasts, others are not even present in *S. cerevisiae* or *S. pombe* or provide better matches to plant or human genomes. In this context we investigated three pillars of metabolism, which are at a first glance not too closely related nor are they apparently overlapping. We could show that with the common topic of “amino acids as essential building blocks” for a living, developing organism, it was relatively easy to link amino acid uptake, amino acid biosynthesis, and regulated breakdown of cellular components. Some results were rather remarkable and could yield interesting starting points for new areas of research in *Aspergillus*. The general topic of amino acid acquisition linking amino acid uptake, biosynthesis, and recycling brought us much deeper into each field of research than we would have expected before our studies and may help to develop new ideas and eventually understand more of the broad complexity of how a cell or organism works.

We were able to *in silico* identify the central components of the cross-pathway control gc/cpc that are yet only known in detail from *S. cerevisiae*. On the other hand we found that other systems accompanying the gc/cpc partly differ from yeast, mainly on behalf of the transport and sensing of extracellular amino acids. Interestingly we found that elements of a basic system like the translation machinery resemble their orthologs of higher eukaryotes much more than the compared yeasts. It was shown by Hoffman et al. (2001) that CpcA is capable of auto-regulating its own transcription under amino acid starvation conditions. Under nonstarvation conditions this auto-regulatory effect is inhibited by CpcB in a yet unclear mechanism [60]. It could be shown that necessary genes and their respective proteins needed for translational regulation of the expression of CpcA are available in all aspergilli. We were able to identify the sensor kinase CpcC, which is presumably able to sense the availability of intracellular amino acids and phosphorylate eIF-2α, a part of the translation machinery, under amino acid starvation conditions which in turn represses translation in general, though the translation of CpcA increases. The known parts of the elongation initiation factor 2, known to be involved in gc/cpc, were identified in the aspergilli. Generally the genes for the subunits of the elongation initiation factor well conserved toward the other compared fungi, whereas the guanine nucleotide exchange factor eIF-2Be subunit is far more similar to the orthologs of higher eukaryotes. The epsilon subunit is not only the largest, but also the catalytic subunit of the complex [181]. It was shown that mutations in the gene for eIF-2Be can exhibit a decrease in complex formation following decreased GTP/GDP exchange rate resulting in altered mRNA transcription and leading to leukoencephalopathy, the vanishing of white matter (VWM), which is a severe inherited human neurodegenerative disorder in man [182]. Due to the high similarity of the respective proteins in *Aspergillus*, an easily genetically manipulated organism, genetic and biochemical research on the effect of mutations in the respective genes may help understand the manifestation of this wasting disease.

According to our data the transport into the nucleus of the central transcription factor of the gc/cpc might be similar to the mechanisms taking place in yeast; at least the necessary factors are present in the aspergilli. Nothing is so far known about the half-life of the protein in the nucleus or in the cytosol. The presence of putative proteins involved in yeast in Gcn4p targeting and degradation in the 26S proteasome in *Aspergillus* indicates similar mechanisms for CpcA degradation. An additional hint is the conservation of the phosphorylatable Thr165 residue, though in all investigated aspergilli this residue was found to be exchanged for a serine residue. Interestingly an alignment of this protein region shows that this phosphorylatable residue is not conserved in *N. crassa*, though phosphorylatable threonine residues can be found in the direct vicinity of the expected spot.

Several mechanisms have been described and proposed for different amino acid uptake systems in mammalian cells. These amino acid uptake systems seem in general not only regulatable, some of them seem to transmit signals of amino acid abundance to directly or indirectly regulate corresponding cellular responses. These mechanisms are so far not well understood. We were able to identify at least four proteins SlcB and SlcC/E/F *in silico* that may have a similar effect on development and growth as their higher eukaryotic relatives. SlcB resembles transporters of the SLC1 family. The function of EAAAT1, a member of the SLC1 family, was recently shown to have a direct effect on the morphology of astrocytes, star-like glial cells. Dysregulation of this glutamate transporter expression leads to disorganized cortex formation and altered astrocytic phenotypes, as was shown for type II lissencephaly patients and cell lines [45,46]. On the other hand, amino acid transporters directly or indirectly influence cellular growth in fly through the TOR pathway [53]. *Aspergillus* has so far proven to be a good model for amino acid dependent growth and regulation, since a dysregulation of intracellular amino acid biosynthesis leads to an
arrest in fruitbody formation [60,86,87,183]. In this context it would be interesting to find new mechanisms reacting to amino acid starvation conditions that have an influence on development and growth. The mechanisms of amino acid regulated growth and development regulation are of great therapeutic interest since there are a lot of pathological circumstances associated with dysregulation of amino acid metabolism (anthropomorphic lateral sclerosis, altered amino acid availability/transport in tumor cells, and tissue response to insulin). Nutritional or pharmaceutical intervention through such mechanisms would be of great benefit. Thus the findings of amino acid transporters similar to those of mammals in filamentous fungi (but not in other lower eukaryotes) may open the way for another field of research for these model organisms.

As examples of amino acid biosyntheses we were able to prove the existence of the necessary genes encoding for the proteins for histidine, lysine, tyrosine, tryptophan, and phenylalanine biosynthesis, demonstrating that the investigated aspergilli are able to produce these amino acids if no extracellular sources are available. However, the tryptophan derivative terrequinone A can only be produced by \textit{A. nidulans} according to the genome analysis, suggesting differences in secondary metabolism. In contrast, mammals are unable to produce all amino acids and have to rely on taking up essential amino acids (isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine) and semiessential amino acids, which can substitute for essential amino acids under certain conditions (arginine, histidine, tyrosine, cysteine), from their diet. In secondary metabolism \textit{A. nidulans} is known to produce penicillin utilizing an intermediate of the lysine biosynthesis. Strikingly the genes encoding for the proteins conducting the enzymatic reactions to produce penicillin could be found in \textit{A. oryzae}, but not in \textit{A. fumigatus}, rendering it unable to produce penicillin.

Concluding, we find that the basic regulatory cascade of regulation of amino acid biosynthesis is very similar to that of higher eukaryotes, although a little bit less complex regarding the activation of target genes. External or internal (GCN2, TOR) sensors sense amino acid abundance and react to depletion by lowering the overall translation rate through eIF2. This leads to increased expression of transcription factors such as Gcn4p, CpcA, or CHOP and increased transcription of target genes. The transcription factor of the gc/cpc, the regulation of cellular expression and probably regulation of its stability seem rather to resemble those of yeast and other fungi, than higher eukaryotes. The sensing and uptake system of amino acids at least in part is more complex than in yeast. Amino acid transporters were found, which are not present in yeast but in higher eukaryotes such as human, and a yeast-like SPS amino acid sensing system is not present. This might indicate ways of uptake and sensing that are similar to those of higher eukaryotes and might have an impact on development and growth.

A full set of proteins for each of the investigated PCI/MPN complexes, the proteasome lid, CSN and eIF3, is present in the genomes of the aspergilli; even a recently discovered alternative subunit can be found. On the other hand, the compared yeasts, more or less, lacked subunits of these complexes. This in contrast to \textit{Aspergillus} and other higher eukaryotes indicates simplified versions in yeasts, which may even lack regulatory properties. The CSN was shown to be an important regulator of SCF activity. The CSN5 deneddylation activity affects SCF activity \textit{in vivo} and with CAND1 confers stability to SCF subunits until a new target of the SCF is marked for degradation by the small protein tag ubiquitin. All proteins so far known to be involved in the basic machinery dealing with SCF\textsubscript{CLL} complexes, their de-/neddylation, ubiquitin conjugation, and stabilization are present in \textit{Aspergillus}, similar to higher organisms. One might expect that these regulatory pathways might work likewise in \textit{Aspergillus} as in higher eukaryotes. So far \textit{A. nidulans} is the only filamentous fungus reported to encode all eight subunits of the CSN, though results from \textit{Neurospora} indicate that it contains at least a partial CSN as well [121]. The fact that the CSN regulates \textit{Aspergillus} development, but on the other hand is not essential in contrast to fly, man, or plant makes research on CSN very interesting and also opens new possibilities.

Acknowledgments

We thank Özgür Bayram for critical reading of the manuscript. Funding was received from the Deutsche Forschungsgemeinschaft (DFG), the Fonds der chemischen Industrie, and the Volkswagen Stiftung.
References


80. Shor, B. et al., Cpc2/RACK1 is a ribosome-associated protein that promotes efficient translation in *Schizosaccharomyces pombe*, *J. Biol. Chem.*, 278, 49119, 2003.


158. Maytal-Kivity, V. et al., MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function, *BMC Biochem.*, 3, 28, 2002.
Amino Acid Supply of Aspergillus

167. Heinemeyer, W. et al., PRE5 and PRE6, the last missing genes encoding 20S proteasome subunits from yeast? Indication for a set of 14 different subunits in the eukaryotic proteasome core, Biochemistry (Mosc.), 33, 12229, 1994.
172. Mundt, K. E. et al., The COP9/signalosome complex is conserved in fission yeast and has a role in S phase, Curr. Biol., 9, 1427, 1999.
184. Jauniaux, J. C. and Grenson, M., GAP1, the general amino acid permease of saccharomyces cerevisiae. Nucleotide sequence, protein similarity with the other baker’s yeast amino acid permeases, and nitrogen catabolite repression, Eur. J. Biochem., 190, 39, 1990.
12

Endocytosis

Juan C. Sánchez-Ferrero and Miguel A. Peñalva

CONTENTS
12.1 Introduction ................................................................. 177
12.2 Caveat Lector ............................................................... 178
12.3 What Is the Endocytic Pathway? ......................................... 178
12.4 Genes Involved in the Endocytic Vesicle Formation and Internalization Step ......................... 180
12.4.1 Clathrin-Dependent and -Independent Endocytosis ..................................................... 180
12.4.2 Endocytosis Appears to be Dynamin Independent ......................................................... 181
12.4.3 Adaptors ................................................................. 182
12.5 Accessory Proteins and Scaffolds: Actin Dynamics ...................... 183
12.5.1 Las17p Module ........................................................ 183
12.5.2 Abp1p Module ......................................................... 184
12.5.3 Pan1p Module ........................................................ 185
12.6 Membrane Identity Across the Endocytic Pathway ...................... 185
12.6.1 Phosphoinositides .................................................... 185
12.6.2 Rab GTPases .......................................................... 186
12.6.3 Rab5 Domain .......................................................... 187
12.7 SNAREs ................................................................. 188
12.8 Multivesicular Body Pathway ............................................... 188
12.9 Functional Characterization of Endocytosis in Aspergillus ..................... 189
12.10 Endocytosis and Signaling ................................................... 190
References ................................................................. 190

12.1 Introduction

Endocytosis is the process by which eukaryotic cells take up portions of their plasma membrane with associated proteins and extracellular fluid. It plays a pivotal role in nutrient acquisition either directly, for example, by mediating the uptake of iron-supplying molecules such as microbial siderophores or mammalian transferrin or, indirectly, by regulating the steady-state at the plasma membrane (i.e., the “expression” at the plasma membrane) of transmembrane domain (TMD)-containing permeases and uptake systems, with one prototypic example in the fungal world being the yeast Gap1p general amino acid permease.1 By mediating the down-regulation of plasma membrane receptors, endocytosis crucially regulates signal transduction, as illustrated by the carcinogenic effect of mutations in metazoan receptor tyrosine kinases that interfere with their ligand-induced internalization and subsequent lysosomal degradation, thus leading to their permanent signaling2–4 or by the antiproliferative effect of mutations impairing the ligand-induced endocytic down-regulation of the G-protein coupled (GPCR) Saccharomyces cerevisiae Ste2p pheromone receptor that results in the inability of mutant yeasts to recover normally from a pheromone-induced cell cycle arrest.5 A further example of involvement of endocytosis in fungal
signal transduction is the essential role of the PalF arrestin in pH.\textsuperscript{6} When coupled to exocytosis, endocytic recycling creates polarity, as demonstrated with the \textit{S. cerevisiae} v-SNARE Snc1 in shmoo tips.\textsuperscript{7} Recycling endosomes ensure delivery of chitin synthase III, a key cell wall biosynthetic enzyme, to polarized sites of growth.\textsuperscript{8}

In cells having highly active polarized secretions such as neurons, endocytosis is required for the efficient retrieval of the excess of membrane lipids and proteins (e.g., the aforementioned v-SNARE, denoted as synaptobrevin in neurons) delivered with synaptic vesicles, from the apical plasma membrane. Filamentous fungal hyphae contain an apical subcellular structure denoted as the Spitzenkörper (Spk), which involves a markedly high concentration of vesicles,\textsuperscript{9} whose almost certain secretory origin is strongly suggested by its labeling with FM4-64\textsuperscript{10} possibly through endocytic membrane recycling. Thus, the highly active localized exocytosis in hyphae poses a similar problem to that of the presynaptic terminal.

Endocytosis plays a role in determining the lipid composition of the plasma membrane. Lipid rafts\textsuperscript{11} form in the plasma membrane of metazoa and fungi. In yeast, these rafts are ergosterol- and ceramide-rich. That ergosterol plays a role in endocytosis stems from the seminal discovery by the Riezman lab demonstrating that mutations impairing cholesterol biosynthesis specifically impair fluid-phase and receptor-mediated endocytosis, indicating that certain raft domains of the plasma membrane might be preferentially endocytosed/internalized, and suggesting a possible role for lipid rafts in the sorting of TMD proteins having a high tendency to be endocytosed.\textsuperscript{12}

Thus, the importance and variety of functions performed by endocytosis is notable, and indeed endocytosis is likely to be essential for every eukaryotic cell. Thus, the report by Torralba and Heath\textsuperscript{13} suggesting the possibility that endocytosis might not occur in \textit{Neurospora crassa} hyphae come as a major surprise. The idea that endocytosis does not occur in filamentous fungi is untenable\textsuperscript{14} in view of bioinformatic data\textsuperscript{14,15} and numerous experimental observations (e.g., see the \textit{in silico} identification of all major genes for endocytosis).\textsuperscript{16,17}

### 12.2 Caveat Lector

Readers should be aware that the set of “endocytic” genes described later may have omitted, for the sake of brevity, some additional players. We have studied in detail this set of genes in \textit{Aspergillus nidulans}, aiming to provide the community of \textit{A. nidulans} researchers with a useful list of systematic designation names, but it should be noted that we have found the corresponding homologs in \textit{Aspergillus oryzae} and \textit{Aspergillus fumigatus}. Those mining the \textit{Aspergillus} genomes are certainly aware that automatic gene calling, although extremely useful, may in some cases be inaccurate due to the intrinsic difficulties of assigning intron positions using educated software in the absence of extensive cDNA sequencing coverage. One dramatic example considered later is the Fab1 homolog AN5211, which is hardly recognizable in the automatic annotation, but there will certainly be other unnoticed examples.

### 12.3 What Is the Endocytic Pathway?

Membrane trafficking involves a variety of vesicular and tubular organelles and membranous structures that are morphologically quite variable, often interconnected amongst them by forward and backward traffic that may be vesicle-mediated (one example is retrograde traffic from the Golgi to the ER), may involve direct organelle fusion (e.g., during homotypic vacuolar fusion), or may result from “maturation” (acquisition of a set of molecular properties involving changes in protein and/or lipid composition). To add yet another layer of complexity, a compartment is often dynamically subdivided in “domains” involving specific lipids and proteins set up by membrane “organizers” (Rab GTPases are membrane organizers, see later). Thus, definition of membrane compartments is quite frequently an idealization. This is complicated further by the overlapping of the late compartments of the endocytic and the biosynthetic vacuolar protein sorting (\textit{vps}) pathways.
A description of the endocytic pathway in mammalian cells is outside the scope of this chapter, and specific mention will be made later only in the context of the interpretation of some of the genomic data for *Aspergillus*. Suffice it to say that, from yeast to mammals, endocytosis starts with an internalization step (the detaching of an endocytic vesicle from the plasma membrane) and continues with the transit of endocytosed material through early and late endosomes successively, to reach the lysosome (in higher cells) or its fungal equivalent, the vacuole.\(^1\)\(^8\),\(^1\)\(^9\)

Yeast genetics has proven very useful in the identification of genes involved in endocytosis (although conclusions have been sometimes hampered by genetic redundancy—a problem that will not be encountered by *Aspergillus* geneticists!, see later), but defining the morphology of different organelles has been far more elusive, as illustrated by the publication dates of two key papers on this subject.\(^2\)\(^0\),\(^2\)\(^1\) Figure 12.1 is an adapted version of the very useful Hugh Pelham’s scheme of the endocytic pathway in *S. cerevisiae*\(^1\)\(^9\) that should be used as a guideline to the following discussion.

In the *S. cerevisiae* paradigm, endocytic vesicles detached from the plasma membrane (Fig. 12.1, step 2) reach a compartment formed by membranes derived from the Golgi (Fig. 12.1, step 3), thus sharing features of both endosomal and Golgi membranes. Such a compartment, denoted as post-Golgi endosome (PGE) in Pelham’s nomenclature, is functionally equivalent to an early recycling endosome. An identity hallmark of the PGE is the presence of the Golgi syntaxin (t-SNARE) Tlg2p but not of the “mature” (“late”) endosomal syntaxin Pep12p. Endocytic cargoes recycling to the plasma membrane via Golgi-derived secretory vesicles (Fig. 12.1, step 1) such as the exocytic v-SNARE and synaptobrevin homolog Snc1p, reach the PGE and are transported to the Golgi following the same retrieval pathway (Fig. 12.1, step 4) as Golgi-resident proteins that leak to this PGE from the Golgi.

The PGE delivers membranes to a second, “downstream” endosomal compartment, characterized by containing the endosomal syntaxin Pep12p and the endosomal phospholipid phosphatidylinositol-3-phosphate [PtdIns(3)P] (Fig. 12.1, step 5). The Pep12p-containing endosome is denoted as the prevacuolar-endosome (PVE). Pelham\(^1\)\(^9\) discusses evidence strongly suggesting that this endosomal compartment may be more akin to a mammalian “early” endosome than to a “late” endosome (one example is its containing the early endosome Rab5 homolog Yp751p/Vps21p). The PVE is a major crossroad between the endocytic and biosynthetic pathways (and an illustrating example of the complexity of membrane trafficking), as it receives biosynthetic traffic from the Golgi [e.g., certain lumenal vacuolar proteins, a prototypic example being the protease carboxypeptidase Y (CPY)] (Fig. 12.1, step 6) and plays a key role in the sorting of misfolded TMD-containing proteins that have escaped ER-mediated degradation into the multivesicular body (MVB) pathway (Fig. 12.1, step 7). Convincing evidence\(^1\)\(^9\) additionally indicates that the PVE delivers a class of secretory vesicles to the plasma membrane. Finally, Golgi-resident proteins and sorting receptors such as the CPY receptor Vps10p, which continuously cycle between the Golgi and the PVEs, are retrieved through a vesicular pathway from this endosomal compartment to the Golgi. (Fig. 12.1, step 8). This retrieval pathway involves dedicated coats (the retromer complex) and sorting proteins (sorting nexins containing PtdIns(3)P-recognizing PX domains, see Ref. 19).

---

**FIGURE 12.1** Schematic representation of membrane trafficking pathways related to endocytosis. This scheme has been adapted to a fungal cell from that described in Ref. 19 for *S. cerevisiae*. Text should be consulted for details.
PVEs mature into multivesicular endosomes (MVEs, also denoted multivesicular bodies, MVBs), which result from inward budding of vesicles emanating from the endosomal membrane into the lumen of the organelle (Fig. 12.1, step 7). As the fate of these MVEs is their fusion with vacuoles to deliver their cargo to the vacuolar lumen, membrane in these vesicles and their associated protein cargoes are predestined for degradation by vacuolar hydrolases. Thus, sorting of proteins into these vesicles is carefully regulated and the evolutionary conserved pathway denoted as the MVB pathway plays a key role in the down-regulation of plasma membrane receptors and transporters reaching PVEs via the endocytic pathway, and in the proteolysis of misfolded TMD-containing proteins reaching PVEs via the biosynthetic pathway from the Golgi. The MVB pathway involves a series of oligomeric protein complexes located at the endosomal membrane, including the three (I, II, III) ESCRTs (endosomal sorting complexes required for transport).22 One ticket that warrants entry into the MVB pathway is monoubiquitin, which is appended to cargoes by ubiquitin ligases such as Rsp5p, a soluble HECT-domain E3 enzyme that is recruited to TMD-containing cargoes directly through Rsp5-binding PPXY motifs in their cytosolic domains or indirectly through the Bsd2p TMD adaptor.23 Another ubiquitin ligase involved in the ubiquitination of TMD cargoes is Tul1p, itself a TMD-containing protein. 24 If ubiquitination takes place in the Golgi (Tul1p, e.g., is a Golgi-resident protein), the adaptor GGA (Golgi-localizing, gamma-adaptin ear homology domain, ARF-binding) proteins have the ability to sort ubiquitinated cargoes and concentrate them in clathrin-coated vesicles dispatched to endosomes.25 Maturation of PVEs into MVEs involves the generation of phosphatidylinositol-3,5-biphosphate (PtdIns(3,5)P2) (see later). MVBs fuse with vacuoles, the terminal station in the endocytic pathway (Fig. 12.1, step 9).

12.4 Genes Involved in the Endocytic Vesicle Formation and Internalization Step

12.4.1 Clathrin-Dependent and -Independent Endocytosis

The best-characterized mechanism for endocytic internalization is based on clathrin coated vesicles, which selectively incorporate lipids and protein cargoes. Proteins are sorted into endocytic vesicles by means of “adaptors,” a set of accessory proteins linking cargoes to the clathrin lattice.26,27 Adaptors typically recognize short peptidic motifs, for example, the well-characterized YxxΦ motif (where Φ represents a hydrophobic amino acid) in mammalian EGFRs, that is recognized by the oligomeric adaptor complex AP-2. Other adaptors (see later) recognize single ubiquitin moieties.β-arrestins are another class of endocytic adaptors. The β-arrestin PalF, the first experimentally documented fungal arrestin, is described later in the context of its role in pH signal transduction.

Mammalian cells contain clathrin-independent endocytic pathways seemingly associated with lipid rafts.28 Caveolae, the structures where one such pathway is thought to occur, contain a prototypic marker, caveolin, as well as the mechanoenzyme dynamin, which is thought to mediate vesicle scission both in caveolae and in clathrin-coated pits. Another protein, flotillin-1, is the marker for a recently defined second clathrin-independent pathway.29

Like budding yeast, A. nidulans (and other aspergilli) contains single genes encoding clathrin heavy (AN4463, claHchc1) and light chains (AN2050, claLclc1) (Table 12.1) but has no caveolin homolog. Clathrin-coated vesicles have been reported in N. crassa.30 The role that clathrin coats play in Golgi-to-PVEs vesicle-mediated traffic31,32 has somehow hindered the analysis of its role in endocytosis in yeast, a role that appeared hardly disputable in view of mammalian cell studies and the likely absence of a fungal caveolar pathway. Thus, the finding that yeast strains deficient in clathrin have only a modest defect in receptor-mediated endocytosis initially came as a surprise.33–35 However, recent reports demonstrated that cortical actin patches labeled with the endocytic marker Abp1p and a series of endocytic internalization factors arise from cortical clathrin patches,36,37 making the contention that clathrin plays a major role in yeast endocytosis incontrovertible. A likely explanation to this paradox is that yeast (and, by extension, fungi) has more than one endocytic pathway, as recently reported by Walter and colleagues, who described a second pathway that, unlike the one localizing to highly motile cortical patches, occurs at static sites denoted as eisosomes.38 These look somehow similar to the FM4-64-positive cortical structures described in A. nidulans.37 The recent demonstration that epsins and Eps15-related proteins suffice, by recognizing
Endocytosis

181

Table 12.1
Some Coats and Adaptors

<table>
<thead>
<tr>
<th>S. cerevisiae Homolog(s)</th>
<th>Protein Product</th>
<th>Proposed Name A. nidulans</th>
<th>Systematic Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHC1</td>
<td>Clathrin heavy</td>
<td>clat&lt;sup&gt;CHC&lt;/sup&gt;</td>
<td>AN4463</td>
</tr>
<tr>
<td>CLC1</td>
<td>Clathrin light</td>
<td>clat&lt;sup&gt;CLC&lt;/sup&gt;</td>
<td>AN2050</td>
</tr>
<tr>
<td>APL4</td>
<td>AP-1γ</td>
<td>apaG</td>
<td>AN4207</td>
</tr>
<tr>
<td>APL2</td>
<td>AP-1β</td>
<td>apaB</td>
<td>AN3029</td>
</tr>
<tr>
<td>APM1</td>
<td>AP-1μ</td>
<td>apaM</td>
<td>AN8795</td>
</tr>
<tr>
<td>APS1</td>
<td>AP1-σ</td>
<td>apaS</td>
<td>AN7682</td>
</tr>
<tr>
<td>APL3</td>
<td>AP-2α</td>
<td>apbA</td>
<td>AN7016</td>
</tr>
<tr>
<td>APL1</td>
<td>AP-2β</td>
<td>apbB</td>
<td>AN5950</td>
</tr>
<tr>
<td>APM4</td>
<td>AP-2μ</td>
<td>apbM</td>
<td>AN7741</td>
</tr>
<tr>
<td>APS2</td>
<td>AP-2σ</td>
<td>apbS</td>
<td>AN0722</td>
</tr>
<tr>
<td>APL5</td>
<td>AP-3δ</td>
<td>apcD</td>
<td>AN8291</td>
</tr>
<tr>
<td>APL6</td>
<td>AP-3-β</td>
<td>apcB</td>
<td>AN0165</td>
</tr>
<tr>
<td>APM3</td>
<td>AP-3-µ</td>
<td>apcM</td>
<td>AN5641</td>
</tr>
<tr>
<td>APS3</td>
<td>AP-3-σ</td>
<td>apcS</td>
<td>AN5519</td>
</tr>
</tbody>
</table>
| enzymes able to mediate membrane constriction, characterized by possession
| (in sequential order from the N-terminus) of dynamin GTPase (PF0035), dynamin central (PF01031), PH
| (PF00169), and GTPase effector (GED, PF02212) domains. |
| The S. cerevisiae proteome has three dynamin domain-containing proteins lacking the PH domain that interacts with lipid bilayers, thus denoted “dynamin-like” proteins. Vps1p is required for trafficking events between the TGN and the endosomal system, for peroxisomal membrane fission, and in vacuolar membrane fusion and fission events. Dnm1p and Mgm1p are involved in mitochondrial fission and fusion events, respectively. Dnm1p may additionally be involved in endocytic trafficking after the internalization step. Thus, none of the aforementioned dynamin-like molecules appears to be involved in the internalization step. 

Orthologs for these three dynamin-like proteins are conserved in Aspergillus. MgmAMgm1 and DnmADnm1 are encoded by A. nidulans AN1093 and AN8874, respectively. The A. nidulans VPS1 homolog, vpsA (AN8023), has been shown to be involved in vacuolar biogenesis. Thus, as in budding yeast, endocytosis seems to be dynamin-independent. Remarkably, aspergilli contain four additional proteins having a dynamin GTPase domain, AN1309, AN1912, AN5327, and AN5552. These are likely homologs of mammalian Mx proteins, a class of relatively poorly characterized, initially described as interferon-induced, proteins. Mx proteins possibly have a normal cellular function, very likely related to their demonstrated ability to self-assemble and tubulate lipids in vitro. MxA, one of the Mx proteins, associates with the

12.4.2 Endocytosis Appears to be Dynamin Independent
In mammalian cells, dynamins, generally regarded as vesicle fission molecules, are involved in scission of endocytic vesicles in both the clathrin-dependent and independent pathways. “Classical” dynamins are mechanochemical enzymes able to mediate membrane constriction, characterized by possession (in sequential order from the N-terminus) of dynamin GTPase (PF0035), dynamin central (PF01031), PH (PF00169), and GTPase effector (GED, PF02212) domains.

The S. cerevisiae proteome has three dynamin domain-containing proteins lacking the PH domain that interacts with lipid bilayers, thus denoted “dynamin-like” proteins. Vps1p is required for trafficking events between the TGN and the endosomal system, for peroxisomal membrane fission, and in vacuolar membrane fusion and fission events. Dnm1p and Mgm1p are involved in mitochondrial fission and fusion events, respectively. Dnm1p may additionally be involved in endocytic trafficking after the internalization step. Thus, none of the aforementioned dynamin-like molecules appears to be involved in the internalization step. 

Orthologs for these three dynamin-like proteins are conserved in Aspergillus. MgmAMgm1 and DnmADnm1 are encoded by A. nidulans AN1093 and AN8874, respectively. The A. nidulans VPS1 homolog, vpsA (AN8023), has been shown to be involved in vacuolar biogenesis. Thus, as in budding yeast, endocytosis seems to be dynamin-independent. Remarkably, aspergilli contain four additional proteins having a dynamin GTPase domain, AN1309, AN1912, AN5327, and AN5552. These are likely homologs of mammalian Mx proteins, a class of relatively poorly characterized, initially described as interferon-induced, proteins. Mx proteins possibly have a normal cellular function, very likely related to their demonstrated ability to self-assemble and tubulate lipids in vitro. MxA, one of the Mx proteins, associates with the
smooth ER but, most important, expression of an MxA GTPase-defective mutant induces expansion of a smooth ER-like compartment and promotes formation of caveolae-like buds along the plasma membrane. As noted earlier, caveolae have not been observed in *S. cerevisiae*, but its proteome has no Mx homologs either.

### 12.4.3 Adaptors

Endocytic adaptors are defined as proteins able to interact with endocytic cargoes to sort and concentrate them into endocytic vesicles. Adaptors bind endocytic accessory proteins and phospholipids, thus contributing to a network of protein–protein and protein–lipid interactions underlying cargo selection and endocytic vesicle internalization. Adaptors typically bind clathrin lattices although, as already noted, this view has recently been challenged by the finding that Eps15 and epsins behave as adaptors recognizing ubiquitinated cargoes in clathrin-independent endocytosis.

A thoroughly studied endocytic adaptor is the heterotetrameric AP-2 complex, one member of the clathrin-dependent heterotetrameric adaptor family. This includes, in *S. cerevisiae*, two additional complexes: AP-1, involved in sorting events in the Golgi, and AP-3, involved in a direct biosynthetic pathway from the Golgi to the vacuole that bypasses PVEs. In *S. cerevisiae* triple deletion of the AP-β subunits and quintuple deletion of adaptor encoding genes (additionally including the two AP180 paralogs, see later) does not produce defects in endocytosis, suggesting that other alternative adaptors (GGAs in the TGN and epsins or Eps15-like proteins in the plasma membrane) might assume key roles in clathrin assembly and cargo selection.

Aspergilli have 12 genes potentially representing single genes for the aforementioned three heterotetrameric adaptor complexes, AP-1, AP-2, and AP-3. Their systematic names and suggested designation in *A. nidulans* are shown in Table 12.1. It is notable that Aspergillus AP-2 proteins are more similar in amino acid sequence to their human than to their *S. cerevisiae* counterparts.

In *S. cerevisiae*, eight endocytic adaptors contain a phosphoinositide binding module of the ENTH/ANTH family: Sla2p (considered later as an accessory factor), five epsins (Eps interacting proteins, for Eps15-like proteins see later) denoted as Ent1-5p, and two homologs of mammalian AP-180 (YAP1801p and YAP1802p) (Table 12.1).

Both the ENTH and the ANTH domains bind phospholipids. An amphipatic ENTH helix with ability to insert into the cytoplasmic leaflet generates membrane curvature, an ability that underlies the role that epsins play in vesicle budding. Such an amphipatic helix is absent in ANTH domains, which do not have membrane-binding ability. The ENTH/ANTH domains are able to interact with tubulin, providing a possible link between endocytic internalization (and vesicle trafficking in general) and the microtubule cytoskeleton. AP180s are monomeric adaptors binding, simultaneously, PtdIns(4,5)P2 (through the ANTH domain) and clathrin, and have the ability to nucleate clathrin lattices on membranes. They additionally have tri-peptidic NPF motifs that are recognized by EH (Eps15-homology, see later) domains, thus recruiting EH-domain-containing interactors such as Pan1 and Ede1 and contributing to the network of protein–protein and protein–lipid interactions underlying endocytic internalization. Unlike yeast, *A. nidulans* has a single AP-180 ortholog (AN5224).

As already noted, epsins [Eps15 (epidermal growth factor receptor substrate 15)-interacting] are ENTH-containing endocytic adaptors localizing to endocytic sites and playing an important role in the internalization step. In common with many other internalization proteins, epsins are modular proteins having multiple interactors. Yeast epsins Ent1p/Ent2p bind PtdIns(4,5)P2, a characteristic phospholipid of the plasma membrane, through the ENTH domain, clathrin, through a clathrin binding motif and monoubiquitin tags (a recognized endocytic signal) through their UIM (Ubiquitin Interacting Motif). In addition, their amino acid sequence contains NPF tripeptidic motifs (see earlier) bound by EH domain-containing interactors. Unlike in yeast, a single Ent1p/Ent2p ortholog (AN3696) is found in aspergilli which is likely to play an essential role (Δent1 and Δent2 mutations are synthetic lethal in *S. cerevisiae*).

*S. cerevisiae* contains three additional ENTH domain-containing proteins, Ent3-5p. The function of Ent4p is unknown. Ent3/Ent5 constitutes a functionally related pair. Ent3p and Ent5p bind the γ-ear domains in Gga2p (one monomeric adaptor in the formation of Golgi-derived vesicles, see earlier) and
AP-1 (a heterotetrameric adaptor in the same clathrin-mediated budding step), which strongly suggests that they play a role in the formation of clathrin-coated vesicles carrying cargoes between the Golgi and endosomes. In addition, Ent3 and Ent5 play a role in the multivesicular body (MVB) pathway involving the ability of their ENTH domains to bind the PVE phospholipid PtdIns(3,5)P₂, as shown by the requirement of Fab1p (the yeast PtdIns(3)P 5-kinase, see later) for cargo sorting into MVB vesicles. Its role on PtdIns(3,5)P₂-containing membranes is likely related to the ability of ENTH domains to induce membrane curvature, perhaps in inward vesicle budding from the endosomal membrane during the genesis of multivesicular endosomes. Aspergilli have a single ortholog for Ent3/Ent5p (AN3821, Table 12.1) and none for Ent4p, therefore, Aspergillus can do it with just two ENTH domain proteins (EntA and EntB), both during vegetative growth and during reproduction.

12.5 Accessory Proteins and Scaffolds: Actin Dynamics

Endocytic internalization is coupled to a burst in actin polymerization mediated by the regulated activation of the Arp2/3 complex. As noted earlier, coupling of the endocytic complex effectors/adaptors/cargoes and actin polymerization involves an intricate network of multiprotein complexes. All genes encoding the seven Arp2/3 complex polypeptides and the components of the coupling multiprotein complexes that have been thoroughly analyzed in S. cerevisiae are present in aspergilli. These include homologs for Las17p, Vrp1p, one type I myosin, one actin regulating kinase, Srv2p, Abp1p, Rvs167p, Rvs161p, Pan1p, Slalp, End3p, Slalp, and Edelp (see Table 12.2). The finding of orthologs for both Rvs167 and Rvs161 in aspergilli (as in Schyzosaccharomyces pombe) strongly suggests that these two proteins have unique rather than (or in addition to) redundant functions.

In S. cerevisiae, where the endocytic eisosomal pathway taking place at static sites has only been recently reported, the best understood endocytic pathway involves highly motile cortical patches containing actin and actin cytoskeleton proteins. These cortical “actin patches,” whose protein composition changes according to a recently defined pathway, are sites for endocytosis. The actin patch lifecycle has three different steps. Very briefly, during the first step, “early” patch proteins including adaptors, Arp2/3 activators and scaffolds assemble in a nonmotile complex; during a second step, actin, Abp1p and Arp2/3 (late patch components) are recruited; actin-dependent slow movement starts and early patch components disassemble after phosphorylation of Pan1p and Slalp by the Ark1p/Prk1p kinases. This slow movement phase of the patch likely corresponds to vesicle formation and release. Actin polymerization forces drive vesicle invagination and neck contraction preceding vesicle release from the plasma membrane; during the last step, patches containing late components enter a fast movement phase during which these components are disassembled and the endocytic vesicle is propelled into the cytoplasm. In this pathway Slalp plays a key role by coupling actin polymerization and endocytic internalization. Its absence prevents patch motility and leads to continuous actin nucleation from nonmotile actin patches. Slalp is targeted to the plasma membrane via its PtdIns(4,5)P₂-recognizing ANTH domain (see discussion on AP180 earlier).

The actin nucleating activity of the Arp2/3 complex is crucially dependent on its activation by accessory proteins. Four such activators are involved in endocytic internalization in S. cerevisiae: Las17p (also named Bee1p, as it is the homolog of mammalian WASP), Abp1p, Pan1p and the Myo3/5 pair of type I myosin paralogs. Through their multiple protein interactions, these Arp2/3 activators organize multiprotein modules which in turn crossinteract amongst them. The Goode and Rodal and Engqvist-Goldstein and Drubin reviews on which the classification that appears later is based, are quite useful to categorize this complex set of yeast endocytic accessory proteins within these modules (Table 12.2).

12.5.1 Las17p Module

In S. cerevisiae, this includes, in addition to Las17p/Bee1p, the Las17p interactor and likely regulator Vrp1p (the homolog of mammalian WIP—WASP interacting protein) and their interacting Myo3p and Myo5p type I myosins, which also bind Arp2/3 and are involved in endocytosis. Fission yeast type I myosin activates the Arp2/3 complex. Single homologs for LAS17 (AN11104) and VRP1 (AN1120)
The Aspergilli

The Aspergilli are present in A. nidulans. May and coworkers showed that A. nidulans contains a single type I myosin gene, denoted as myoA (AN1588), which is involved in endocytosis (single myoA orthologs are also present in A. fumigatus and A. oryzae).

12.5.2 Abp1p Module

This includes the Arp2/3 activator, late phase patch component Abp1p, which recruits Arp2/3 to the sides of actin filaments. Abp1p additionally serves as scaffold, recruiting actin regulating kinases and the actin monomer-binding protein Srv2p to cortical patches. Single Abp1p- and Srv2p-encoding genes are found in A. nidulans (AN8873 and AN0999, respectively). In contrast to yeast (Prk1p, Ark1p) and humans (AAK1, GAK1) aspergilli have a single acting regulating kinase (that in A. nidulans we denote as ArkA, encoded by AN10515.3).

Another Abp1p interactor and key accessory factor in endocytosis is the actin patch component Rvs167p, which possesses a characteristic PFAM BAR domain (PF03114), a membrane-binding and curvature-sensing module that, by preferentially binding membranes with high curvature, has membrane deformation properties as shown by its ability to tubulate liposomes. Yeast has a second BAR domain protein, Rvs161p. In mammalian cells, two BAR-containing proteins are involved in endocytic vesicle budding, amphiphysins and endophilins. Hicke and coworkers have reported evidence strongly indicating that Rvs167p and its interactor Slalp (see later) are yeast homologs of mammalian endophilin and CIN85 (yet another endocytic factor), respectively. If confirmed, this would suggest that Rvs161p is the S. cerevisiae homolog of mammalian amphiphysin. Aspergilli (A. nidulans) contain genes for Rvs161p (AN8831) and Rvs167 (AN2516) orthologs, which in the light of this analysis agrees with the view that these proteins have unique functions. Rvs161p is involved in correct actin localization and represents yet another link between actin cytoskeleton and endocytosis.

### TABLE 12.2

Proteins Involved in the Endocytic Internalization Step

<table>
<thead>
<tr>
<th>S. cerevisiae Homolog(s)</th>
<th>Protein Product</th>
<th>Proposed Name A. nidulans</th>
<th>Systematic Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC40</td>
<td>Arp2/3 40</td>
<td>arpD</td>
<td>AN5778</td>
</tr>
<tr>
<td>ARC35</td>
<td>Arp2/3 35</td>
<td>arpE</td>
<td>AN0306</td>
</tr>
<tr>
<td>ARC19</td>
<td>Arp2/3 19</td>
<td>arpF</td>
<td>AN8065</td>
</tr>
<tr>
<td>ARC18</td>
<td>Arp2/3 18</td>
<td>arpG</td>
<td>AN8698</td>
</tr>
<tr>
<td>ARC15</td>
<td>Arp2/3 15</td>
<td>arpH</td>
<td>AN4919</td>
</tr>
<tr>
<td>ARP2</td>
<td>Arp2</td>
<td>arpP^92</td>
<td>AN0673</td>
</tr>
<tr>
<td>ARP3</td>
<td>Arp3</td>
<td>arpP^93</td>
<td>AN0140</td>
</tr>
<tr>
<td>LAS17</td>
<td>WASP</td>
<td>beeA</td>
<td>AN11104.3</td>
</tr>
<tr>
<td>RV161</td>
<td>Amphiphysin</td>
<td>ampB</td>
<td>AN8831</td>
</tr>
<tr>
<td>RV167</td>
<td>Endophilin</td>
<td>ampA</td>
<td>AN2516</td>
</tr>
<tr>
<td>SLA1</td>
<td>Sla1p/CLN85</td>
<td>slaA</td>
<td>AN1462</td>
</tr>
<tr>
<td>SLA2</td>
<td>Sla2p/Hip1</td>
<td>slaB</td>
<td>AN2756</td>
</tr>
<tr>
<td>RSP5</td>
<td>Ubiquitin ligase</td>
<td>rsvP</td>
<td>AN1359</td>
</tr>
<tr>
<td>PAN1</td>
<td>Pan1p</td>
<td>panA</td>
<td>AN4270</td>
</tr>
<tr>
<td>END3</td>
<td>End3p</td>
<td>sagAT103</td>
<td>AN1023</td>
</tr>
<tr>
<td>EDE1</td>
<td>Eps15-like</td>
<td>edeA</td>
<td>AN0317</td>
</tr>
<tr>
<td>VRP1</td>
<td>WIP (WASP interacting)</td>
<td>wipeA</td>
<td>AN1120</td>
</tr>
<tr>
<td>PRK1 and ARK1</td>
<td>Adaptor associated kinase</td>
<td>arkA</td>
<td>AN10515.3</td>
</tr>
<tr>
<td>SRV2</td>
<td>Actin-monomer binding</td>
<td>srvA</td>
<td>AN0999</td>
</tr>
<tr>
<td>ABP1</td>
<td>Actin-binding protein</td>
<td>abpA</td>
<td>AN8873</td>
</tr>
<tr>
<td>MYO3 and MYO5</td>
<td>Type I myosins</td>
<td>myoA</td>
<td>AN1588</td>
</tr>
</tbody>
</table>

are present in A. nidulans. May and coworkers showed that A. nidulans contains a single type I myosin gene, denoted as myoA (AN1588), which is involved in endocytosis (single myoA orthologs are also present in A. fumigatus and A. oryzae).
12.5.3 Pan1p Module

Pan1p is the third amongst Arp2/3 activators able to nucleate a multiprotein complex. With End3p and Ede1p (see later), they represent the three *S. cerevisiae* EH (Eps15 homology) domain-containing proteins whose function has been characterized (the function of two additional yeast EH proteins, YKR019c and YJL083w has not yet been addressed). Pan1p plays a crucial role in the internalization step through its central position in an interacting network involving endocytic machinery proteins and actin regulators. Pan1p EH domains recognize NPF tripeptide motifs in its interacting partners, including Sla1p, an endocytic factor that interacts with Sla2p to regulate actin dynamics and which has a second role as an endocytic adaptor: Sla1p recognizes an internalization signal NPFX(1,2)D in endocytic cargoes. We noted earlier that the Pan1p EH domain additionally binds yAP180A/B, Pan1p interacts physically with End3p, another EH domain-containing protein that is required for normal Pan1p localization, and it also interacts with Ent1/2p epsins showing genetic interactions with the Rsp5p ubiquitin ligase (see later). Also included in the Pan1p module is Ede1p, the third characterized EH domain protein. Ede1p is an Ent1p interactor. *EDE1* shows genetic interactions with *PAN1* and *END3*. Single genes encoding Pan1p (AN4270), End3p (AN1023), Ede1p (AN0317), Sla1p (AN1462), and Sla2p (AN2756) orthologs are found in *A. nidulans/aspergilli*. Of these, only the *A. nidulans END3* homolog *sagA* has been reported previously.

12.6 Membrane Identity Across the Endocytic Pathway

Membrane trafficking involves integral membrane proteins such as SNAREs (*N*-ethylmaleimide sensitive factor attachment protein receptors, considered later) that play a key role in determining the specificity of fusion events between membranes (dictating to a significant extent, e.g., that Golgi-derived vesicles carrying cargoes *en route* to the vacuole fuse with endosomes). Organelle-specific integral membrane protein cargoes are sorted into vesicles that deliver them to the correct subcellular location and/or retrieve them when they escape from their “correct” membrane compartment. Proteins (e.g., endocytic adaptors) play a major role in sorting events but the emerging role of lipids shown by the segregation of certain TMD proteins in lipid microdomains at the plasma membrane (the already-discussed rafts) is certainly a mainstream topic of future research. However, membrane trafficking and, by extension, the physiological role of subcellular organelles, additionally involves a number of peripheral membrane proteins that must be targeted to subcellular localizations or even to specific membrane domains within a particular organelle in a highly specific manner. To this end, organelles have molecular “codes” that determine their identity, involving a specific subset of phosphoinositides as well as small GTPases of the Rab and Arf families. Identity codes are deciphered through protein-lipid and protein-protein interactions (readers interested in the problem of membrane identity should consult the clarifying review by Behnia and Munro). Due to space limitations, we later consider separately proteins involved in phosphoinositide modification and, very briefly, Rab GTPases and SNAREs, whose genomics in aspergilli have already been the subject of a previous study.

12.6.1 Phosphoinositides

Phosphorylated derivatives of phosphatidylinositol (PtdIns), referred to as phosphoinositides, are major contributors to membrane identity “codes,” in addition to having a direct role in signal transduction. PtdIns can be phosphorylated singly or in combination in the 3’, 4’, and 5’ hydroxyl groups. The polar moieties of phosphoinositides protrude from the lipid bilayer and are recognized by code-deciphering protein modules/domains. Examples along the endocytic pathway are the already-discussed ANTH domains in AP180 or the ENTH domains in Ent1/Ent2 recognizing the plasma membrane phosphoinositide PtdIns(4,5)P₂, the ENTH domain in Ent3/Ent5 binding PtdIns(3,5)P₂, very likely at the mature endosome membrane and the FYVE domain in Vps27 (see later) recognizing endosomal PtdIns(3)P. One feature of the phosphoinositide code is that it is spatially restricted, a feature likely imposed by its need to act in
a highly specific manner. This spatial restriction is achieved by spatial (and temporal) regulation of phosphoinositide synthesis and turnover, mediated by PtdIns kinases and phosphatases.

Phosphoinositide phosphatases terminate phosphoinositide signaling. Synaptojanin, a neuronal enzyme dephosphorylating PtdIns(4,5)P\(_2\) to PtdIns(4)P\(_2\), plays a key role during the synaptic cycle in the uncoating of clathrin-coated vesicles subsequent to internalization. Synaptojanins likely dictate vesicle uncoating by debilitating interactions between the coat multiprotein complex and the vesicle membrane, switching off adaptor-mediated lipid-binding. Metazoan synaptojanins share a characteristic domain organization, with N-terminal Sac1 (PF02383) domain and a C-terminal phosphatase domain (PF03372). S. cerevisiae has seven well-characterized phosphoinositol phosphatases. Among them, three resemble mammalian synaptojanins in their domain organization and regulate the localization of PtdIns(4,5)P\(_2\) to the plasma membrane. Functional redundancy has complicated the analysis of the role of yeast synaptojanins in membrane trafficking. A. nidulans has four Sac1 domain-containing proteins, of which only the AN8288 product has the C-terminal phosphatase domain, thus representing the sole likely synaptojanin homolog, whose precise molecular role in endocytosis is an attractive avenue for future research (Table 12.3).

PtdIns and PtdIns-phosphate kinases initiate PtdIns signaling. The finding that yeast VPS34 encoding a PtdIns 3-kinase is required for vacuolar protein sorting revealed the involvement of phosphoinositides in membrane trafficking. The PtdIns(3)P 5-kinase Fab1p\(^{94}\) terminates PtdIns(3)P signaling and initiates that of PtdIns(3,5)P\(_2\) on the membrane of the mature endosome. It is required for proper trafficking and sorting of endocytic cargo through the PVEs/MVBs. These two key lipid kinases in the endocytic pathway have single homologs in A. nidulans (AN4709 and AN5211 for Vps34p and Fab1p, respectively, Table 12.3). PtdIns(3)P signals through FYVE domains specifically recognizing this phosphoinositide. Budding yeast contains five FYVE proteins and the corresponding A. nidulans homologs are described in Table 12.3. Of these, Vps27p and Vps19p (a Rab5 effector) are discussed later in their corresponding sections. Fab1p, the PtdIns(3)P 5-kinase, also contains a FYVE domain and thus it is unique in that it down-regulates FYVE domain effectors including itself and up-regulates PtdIns(3,5)P\(_2\) effectors\(^95\) such as Ent3 (discussed earlier). Pib1p is a RING E3 ubiquitin ligase located at endosomal and vacuolar membranes. The function of Pib2p is not known. A second major class of PtdIns(3)P binding proteins, (not considered here, see Ref.19 for a brief review) are sorting nexins, a class of peripheral membrane proteins consistently involved in retrieval of proteins from endosomes, which bind this phosphoinositide through their PX domains.

### 12.6.2 Rab GTPases

With phosphoinositides, small GTPases of the Rab and Arf families are major determinants of membrane identity, through their recruitment to membranes, in their activated GTP-bound state, of Rab effector proteins\(^93\) that organize lipid/protein domains (the Rab5 example is detailed later). The Rab repertoire of A. fumigatus has been analyzed previously, using a preliminary release of the sequence.

Each of the three recently published Aspergillus genomes (nidulans, fumigatus, oryzae) has coding capacity for 10 Rabs. This number contrasts with the 11 Rab genes found in S. cerevisiae, indicating that

<table>
<thead>
<tr>
<th><strong>TABLE 12.3</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphoinositide Metabolism and Phosphoinositide Binding Proteins</td>
</tr>
<tr>
<td><strong>S. cerevisiae Homolog(s)</strong></td>
</tr>
<tr>
<td>SJL1/SJL2/SJL3</td>
</tr>
<tr>
<td>VPS34</td>
</tr>
<tr>
<td>FAB1</td>
</tr>
<tr>
<td>VPS19</td>
</tr>
<tr>
<td>VPS27</td>
</tr>
<tr>
<td>PIB1</td>
</tr>
<tr>
<td>PIB2</td>
</tr>
<tr>
<td>ENT3 and ENT5</td>
</tr>
</tbody>
</table>
the presumably higher complexity of membrane trafficking in fungi compared to yeast is not reflected in a greater complexity of Rabs. Our phylogenetic analyses (Sánchez-Ferrero and Peñalva, to be published elsewhere) strongly suggest (in agreement with Gupta et al.\textsuperscript{15}) the presence of Ypt1 (ER-Golgi), Ypt6 (Golgi-endosome), Ypt7 (homotypic vacuolar fusion), Ypt31 (secretory pathway, Golgi), and Sec4 (secretory pathway, fusion events with the plasma membrane) yeast-like Rabs (Table 12.4). The \textit{A. nidulans} Ypt7 homolog \textit{avaA} has been reported and its mutant phenotype is consistent with its involvement in vacuole fusion.\textsuperscript{96} The \textit{A. niger} Ypt1p and Sec4p homologs have also been described.\textsuperscript{97} \textit{A. niger} SrgA\textsuperscript{Sec4} is not an essential protein, suggesting the intriguing possibility that a second SrgA-independent secretory pathway exists in \textit{Aspergillus}.\textsuperscript{97}

We find three members related to the Rab5 class. One Rab5 is related to \textit{S. cerevisiae} Ypt51; the second is related to yeast Ypt52; the third is closer to hsRAB24, an atypical Rab that appears to be involved in the autophagic pathway,\textsuperscript{98} than to any of the three yeast Rab5s. We confirmed the presence of hsRAB2 and hsRAB4 homologs.\textsuperscript{15} RAB2 and RAB4 homologs are absent from yeast. RAB4 is characteristic of early recycling endosomes. Its presence in fungal genomes might suggest that efficient membrane recycling pathways are required for hyphal cell growth.

### 12.6.3 Rab5 Domain

A thoroughly studied protein/lipid domain is that organized by Rab5-GTP on endosomal membranes,\textsuperscript{95} conserved to a significant extent in \textit{S. cerevisiae}. Rab5 is involved in homo- and heterotypic endosome membrane fusions. Rab5-GTP (the membrane-bound form of the GTPase, see Ref. 88) acts in the tethering step of a donor membrane to the target organelle by recruiting a series of proteins that are referred to as Rab effectors. The priming step in the assembly of the mammalian Rab5 domain is the activation of the GDP-bound Rab by Rabex-5, the nucleotide exchange factor (GEF) of Rab5. Yeast Vps9p, the Rabex-5 ortholog, is the GEF for Ypt51p.\textsuperscript{99} Human Vps34/p150 PtdIns 3-kinase is a Rab5 effector coupling Rab activation to PtdIns(3)\textsubscript{P} synthesis.\textsuperscript{100} Its yeast Vps34p homolog\textsuperscript{101} is associated with endosomes. Vps15p, a Vps34p interactor, is a protein kinase required for Vps34p activity/function\textsuperscript{101} and is conserved in aspergilli. In a positive feedback loop, Vps34p-dependent PtdIns(3)\textsubscript{P} synthesis recruits additional Rab5 effectors like Vac1p/Vps19p, a key effector with a PtdIns(3)\textsubscript{P}-binding FYVE domain. Vps19p is the structural homolog of human EEA1 (early endosome antigen 1), a tethering factor for endosome fusion. Vps19p is a promiscuous multiple interactor that binds the PGEs and PVEs t-SNAREs Tgl2p and Pep12p, respectively\textsuperscript{89,102} as well as Vps45p, a Sec1p-like protein required for SNARE function. Homologs of Ypt51p (see earlier), Vps15p (AN0576), Vps34p (see earlier), Vps19p (Table 12.3), Vps45p (AN6531), Tlg2p and Pep12p (Table 12.4) are found in \textit{A. nidulans}.

### TABLE 12.4

<table>
<thead>
<tr>
<th>\textit{S. cerevisiae} Gene</th>
<th>Proposed Name \textit{A. nidulans}</th>
<th>Systematic Name</th>
<th>Function/Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{HSE1}</td>
<td>\textit{hseA}</td>
<td>AN2066</td>
<td>Vps27 interactor</td>
</tr>
<tr>
<td>\textit{VPS27}</td>
<td>\textit{hrsA}</td>
<td>AN2071</td>
<td>FYVE</td>
</tr>
<tr>
<td>\textit{VPS23}</td>
<td>\textit{escA}</td>
<td>AN2521</td>
<td>ESCRT-I</td>
</tr>
<tr>
<td>\textit{VPS28}</td>
<td>\textit{escB}</td>
<td>AN0945</td>
<td>ESCRT-I</td>
</tr>
<tr>
<td>\textit{VPS37}</td>
<td>none recognizable</td>
<td></td>
<td>ESCRT-I</td>
</tr>
<tr>
<td>\textit{VPS22}</td>
<td>\textit{escD}</td>
<td>AN7106</td>
<td>ESCRT-II</td>
</tr>
<tr>
<td>\textit{VPS25}</td>
<td>\textit{escE}</td>
<td>contig 1.61*</td>
<td>ESCRT-II</td>
</tr>
<tr>
<td>\textit{VPS36}</td>
<td>\textit{escF}</td>
<td>AN7037</td>
<td>ESCRT-II</td>
</tr>
<tr>
<td>\textit{VPS2}</td>
<td>\textit{escG}</td>
<td>AN6898</td>
<td>ESCRT-III</td>
</tr>
<tr>
<td>\textit{VPS20}</td>
<td>\textit{escH}</td>
<td>AN1365</td>
<td>ESCRT-III</td>
</tr>
<tr>
<td>\textit{VPS24}</td>
<td>\textit{escI}</td>
<td>AN6920</td>
<td>ESCRT-III</td>
</tr>
<tr>
<td>\textit{VPS32}</td>
<td>\textit{escJ}</td>
<td>AN4240</td>
<td>ESCRT-III</td>
</tr>
</tbody>
</table>

* No systematic name available.
12.7 SNAREs

In addition to tethering factors such as the Rab effectors that were described earlier, the basic machinery ensuring the specificity of membrane fusion involves SNAREs. SNAREs are typically single-pass membrane proteins having cytosolic domains (the SNARE domain) with ability to form alpha-helical coiled-coils arranged in four-helix bundles. Based on conserved structural features, SNARE domains have been classified in Qa, Qb, Qc, and R. Specific membrane fusion events typically involve a specific four-helix bundle formed by three Q-SNAREs and one R-SNARE. Qa SNAREs, present in acceptor membranes, are homologs of neuronal syntaxin and are collectively referred to as syntaxins. Syntaxins may be used as membrane identity markers. Acceptor membranes contribute two additional helixes (of the Qb and Qc class) to four helix bundles, with one R-SNARE in the donor membrane providing the fourth helix.

Gupta et al. analyzed *A. fumigatus* SNAREs and found five Qa-SNAREs, five Qb-SNAREs, six Qc-SNAREs and four R-SNAREs (one demonstrated Qb, Sec20p does not meet the SNARE domain consensus). Our analysis of *A. nidulans* SNAREs agrees with their conclusions and confirms that yeast duplicated gene pairs are represented by single orthologs in aspergilli.

While the precise role of each SNARE is not unequivocally predictable without experimental evidence, an educated guess based on *S. cerevisiae* reveals that with a single already noted exception, aspergilli contain all SNARE machinery acting at the PGE and PVE membranes, as well as that involved in homotypic vacuolar fusion (Table 12.4). The noted exception is the absence in aspergilli of a Vam3, the yeast Qa SNARE for homotypic vacuolar fusion, whose function is likely provided by Pep12. Tlg2 is the PGE syntaxin. Homologs of the PGE “Tlg2 complex” components Tlg2 (Qa), Vti1p (Qb) Tlg1p (Qc), and Snc1p (R) are present *A. nidulans*, *A. niger*, and *A. oryzae*. Pep12p is the yeast PVE syntaxin. *S. cerevisiae* uses several SNARE endosomal complexes containing Pep12p as Qa. These complexes contain, in addition to Pep12p, Vti1p as Qb, either Tlg1p or Syn8p as Qc (mutants show synthetic interactions) and either Snc1p/Snc2p or Ykt6p as R-SNARE. The corresponding orthologs are found in aspergilli. Vti1p (Qb), Van7p (Qc), and R-SNAREs Nyv1p and Ykt6p are required for vacuolar fusion events. Ykt6p appears to be specific to fusion events between the prevacuolar endosomes and vacuoles.

12.8 Multivesicular Body Pathway

Transmembrane protein traffic in transit to the vacuole arrives to PVEs from the Golgi, via trans Golgi network (TGN)-derived vesicles, or from the plasma membrane, via endocytosis. These proteins follow two different routes. One class of cargoes stay in the membrane of the prevacuolar endosome, from which they either reach the vacuolar membrane after fusion with this organelle (this pathway is followed by resident vacuolar membrane proteins), or recycle to the plasma membrane. The second class are sorted into inwardly-budding vesicles that are released within the endosomal lumen, giving this organelle a multivesicular appearance (hence the name of “multivesicular body”). After fusion with a vacuole, MVB vesicles are released into the vacuolar lumen, where they are exposed to vacuolar hydrolases. Therefore, proteins entering the multivesicular body (MVB) pathway are usually targeted for degradation. Thus, among other roles outside the scope of this chapter, the MVB pathway, which is conserved from yeast to humans, plays a key role in the down-regulation of plasma membrane receptors and permeases.

In yeast, loss-of-function mutations in any of 17 genes referred to as the class E vps genes lead to an enlarged endosomal compartment denoted as the class E compartment and are required for sorting into the MVB pathway. Several class E proteins are organized into three sequentially acting multiprotein complexes denoted ESCRT (endosomal sorting complex required for transport)-I, -II and -III (see Table 12.5). Single orthologs for ESCRT-I VPS23 and VPS28, but not for VPS37 are found in aspergilli (note that no VPS37 homolog is evident in humans either). We found, however, single orthologs for all components of ESCRT-II (VPS22, VPS25, and VPS36) and ESCRT-III (VPS2, VPS20, VPS24, and VPS32/SNRF7). In addition, we found a gene encoding an homolog of the known Vps32p interactor Bro1p (a PalA homolog), demonstrating that, like in yeast, two Snf7 interactors (PalA and Bro1) having a Bro1 domain are present in aspergilli.
Vps27 (Table 12.5) funnels cargo into the MVB pathway using its three functionally characterized interacting domains. Vps27p binds ubiquitin through its UIM motif, thereby recognizing ubiquitin-labeled MVB cargoes. It binds endosomal membranes via its PtdIns(3)P-recognizing FYVE domain. Finally, it recruits the ESCRT-I complex through a Vps23 binding domain. Another key MVB pathway component is the AAA ATPase Vps4p, whose ATP hydrolyzing activity is required for disassembling the ESCRT complexes from the endosomal membrane to allow subsequent cycles of cargo sorting.

12.9 Functional Characterization of Endocytosis in Aspergillus

digA, a gene encoding an A. nidulans homolog of the class C vps protein Vps18p, was cloned after complementation of a temperature-sensitive mutation that, under restrictive conditions, leads to the expected vacuolar fragmentation, but additionally results in defects in nuclear migration, mitochondrial morphology and polarized growth. vpsA, the A. nidulans homolog of yeast VPS1, encoding one of the dynamin-like proteins possibly acting at the TGN and avaB, encoding an homolog of yeast Vps39p, a nucleotide exchange factor for the Ypt7 vacuolar Rab and a Ypt7 effector tethering transport vesicle to the vacuole have been disrupted. Not surprisingly, these mutations result in fragmented vacuoles and poor growth. syn8, encoding a homolog of yeast VPS8, a potential vps8p homolog is also identified. Ypt7 is a Rab protein involved in fusogenic steps, and AvaA promotes fusion of transport vesicles with the vacuole. AvaAYpt7 results in vacuolar enlargement, providing evidence that AvaAYpt7 acts in fusion steps involving the vacuole as acceptor membrane.

TABLE 12.5
A. nidulans SNAREs

<table>
<thead>
<tr>
<th>S. cerevisiae Homolog(s)</th>
<th>SNARE Function</th>
<th>Proposed Name</th>
<th>A. nidulans Systematic Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEP12</td>
<td>Qa</td>
<td>pepA</td>
<td>AN4416</td>
</tr>
<tr>
<td>SED5</td>
<td>Qa</td>
<td>sedA</td>
<td>AN9526</td>
</tr>
<tr>
<td>SSO1/SSO2</td>
<td>Qa</td>
<td>ssoA</td>
<td>AN3416</td>
</tr>
<tr>
<td>TLG2</td>
<td>Qa</td>
<td>tlgB</td>
<td>AN2048</td>
</tr>
<tr>
<td>UFE1</td>
<td>Qa</td>
<td>ufeA</td>
<td>AN6047</td>
</tr>
<tr>
<td>VAM3</td>
<td>Qa</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>BOS1</td>
<td>Qb</td>
<td>boxA, contig 1.112</td>
<td>94,812 &gt; 95,753b</td>
</tr>
<tr>
<td>GOS1</td>
<td>Qb</td>
<td>gosA</td>
<td>AN1229</td>
</tr>
<tr>
<td>SEC9</td>
<td>Qb, Qc</td>
<td>secN</td>
<td>AN2419</td>
</tr>
<tr>
<td>SEC20</td>
<td>Qb</td>
<td>secT</td>
<td>AN2969</td>
</tr>
<tr>
<td>VTI1</td>
<td>Qb</td>
<td>vtiA</td>
<td>AN1973</td>
</tr>
<tr>
<td>BET1</td>
<td>Qc</td>
<td>betA</td>
<td>AN5127</td>
</tr>
<tr>
<td>SFT1</td>
<td>Qc</td>
<td>sftA, contig 1.67</td>
<td>242,010 &gt; 242,461b</td>
</tr>
<tr>
<td>SYN8</td>
<td>Qc</td>
<td>synE</td>
<td>AN2169</td>
</tr>
<tr>
<td>TLG1</td>
<td>Qc</td>
<td>tlgA</td>
<td>AN8171</td>
</tr>
<tr>
<td>VAM7</td>
<td>Qc</td>
<td>vamS</td>
<td>AN4551</td>
</tr>
<tr>
<td>NYV1</td>
<td>R</td>
<td>nyvA</td>
<td>AN0571</td>
</tr>
<tr>
<td>SEC22</td>
<td>R</td>
<td>secS, contig 1.10</td>
<td>95,603 &gt; 96,484p</td>
</tr>
<tr>
<td>SNC1/SNC2</td>
<td>R</td>
<td>sncA</td>
<td>AN8769</td>
</tr>
<tr>
<td>SPO20</td>
<td>Qc</td>
<td>None identified</td>
<td>—</td>
</tr>
<tr>
<td>YKT6</td>
<td>R</td>
<td>yktS</td>
<td>AN8488</td>
</tr>
</tbody>
</table>

* In AN6047, automated gene calling, although detected a gene in this region, failed to yield the correct gene structure to encode this syntaxin. Start and stop codon coordinates of our manual reconstruction are given.

b Automated gene calling detected no genes in the corresponding coding regions. Start and stop codons in our reconstructed gene models are given.
homologs (and were first to describe that one species of Aspergillus contains a mammalian RAB2 homolog, absent in S. cerevisiae). Unexpectedly, they found that disruption of srgASEC4 was viable (discussed earlier). In A. oryzae, mutations leading to mislocalization of a CPY-EGFP reporter have been isolated but the corresponding genes have not yet been reported.116,117 sagA, the likely END3 homolog in A. nidulans, was cloned after complementation of a sagA1 mutation resulting in high sensitivity to alkylating agents.87

As evidenced earlier, our understanding of endocytic (and membrane) trafficking in Aspergillus is surprisingly scarce, despite the crucial importance that this pathway may have in filamentous fungal cell physiology. One example is the swoC1 mutant,118 impaired in polarity establishment. Under restrictive conditions swoC1 leads to abnormal enlargement of swelling conidia with multiple sites of germ tube emergence.118 Notably, these mutant swelling conidia are deficient in membrane internalization as determined by uptake of the lipophylic dye FM4-64.118 This suggests an as-yet-undefined connection between endocytosis and polarity in fungi. Active endocytic recycling of apical membranes has been suggested by Fischer-Parton et al.10 to explain their labeling of the Spitzenkörper, supposedly crowded with exocytic vesicles, with FM4-64. One of us17 used FM4-64 to trace membrane internalization in A. nidulans, visualize endocytic intermediates (see later), and demonstrate that the vacuolar membrane is at the end of one branch of the endocytic pathway. In a key paper illustrating the potential interest of studying the endocytic pathway in A. nidulans, Yamashita and May reported that constitutive mutational activation of MyoA, the single A. nidulans class I myosin (class I myosins are known activators of the Arp2/Arp3 complex) results in activation of endocytosis leading to accumulation of membranes in growing hyphae.16 Finally, it has been reported that PAF, a 55-residue peptide with antifungal activity, secreted by Penicillium chrysogenum, is internalized via endocytosis in A. nidulans, which might imply the existence of a specific membrane receptor(s), by analogy to the receptor-mediated endocytosis of fungal mating pheromones.

12.10 Endocytosis and Signaling

The pH signal transduction pathway (the pal pathway) involves the endocytic pathway at two different levels. The Bro1-domain protein PalA119 and very likely PalC (also a Bro1-domain protein120) are physical interactors, with the cysteine protease PalB, of ESCRT-III components presumably located on the cytosolic side of the PVE membrane, including Vps32.110 At a second level, the 7-TMD protein PalH (almost certainly the receptor of the pathway) and its accessory factor PalI act at the plasma membrane, where the cytosolic tail of PalH interacts with PalF.6 PalF is the first demonstrated example of a β-arrestin in the fungal world.6 Arrestins typically down-regulate G-protein coupled 7-TMD receptors, leading to their “desensitization.” This paradigm has been recently challenged by the finding that, like PalF, metazoan β-arrestins may have a positive, rather than a negative, role in signal transduction.121 Because β-arrestin is a well-known endocytic adaptor,26,122 this strongly indicates that β-arrestins mediate their positive role in signaling by promoting endocytosis of their cognate receptors.122 Of note, two recent examples of for this positive role of β-arrestin involve endocytosis of the 7-TMD protein smoothened.122,123 Smoothened mediates activation of the Hedgehog/Sonic Hedgehog pathway, a pathway that is likely to share ancestry with the fungal pH signaling pathway.124 The finding that certain receptors are activated at the plasma membrane but signal from endosomes is recurrently found in the eukaryotic lineage.39,125 In the fungal pH signaling pathway, signaling from endosomes may help bringing together two spatially separated subsets of Pal proteins.

References

Endocytosis


13

RNA Silencing in the Aspergilli

Thomas M. Hammond and Nancy P. Keller

CONTENTS

13.1 Introduction .................................................................................................................. 197
13.2 Experimental RNA Silencing ....................................................................................... 198
  13.2.1 Aspergillus fumigatus ........................................................................................... 198
  13.2.2 Aspergillus flavus and Aspergillus parasiticus ................................................... 198
  13.2.3 Experimental RNA Silencing During Infection .................................................. 199
13.3 Genetic Analysis of Experimental RNA Silencing ....................................................... 199
  13.3.1 Aspergillus nidulans RNA Silencing Model ....................................................... 199
  13.3.2 RNA Silencing Proteins in Aspergillus nidulans .................................................. 200
13.4 Aspergillus RNA Silencing Gene Evolution ................................................................. 200
  13.4.1 Dicers and Argonautes ....................................................................................... 201
  13.4.2 RNA-Dependent RNA Polymerases ................................................................. 201
13.5 Possible Roles of Aspergillus RNA Silencing in Nature ............................................. 203
  13.5.1 Meiotic Silencing ............................................................................................... 203
  13.5.2 Quelling ............................................................................................................ 204
  13.5.3 RNAi-Mediated Heterochromatic Silencing ....................................................... 204
13.6 Future Directions ......................................................................................................... 205
References ............................................................................................................................... 206

13.1 Introduction

Small noncoding RNAs have numerous biological functions, mediating processes such as post-transcriptional gene silencing, heterochromatic silencing, antiviral defense, and transposable element control. The proteins involved in small RNA use and production are called RNA silencing proteins. Core RNA silencing proteins are Dicer, Argonaute, and RNA-dependent RNA polymerase (RDRP). Dicer is an RNaseIII-containing protein responsible for processing dsRNA into various small RNA species, typically 21–25 nt in length. Dicer-processed small RNAs are incorporated into Argonaute-containing effector complexes, such as RISC (RNA-induced silencing complex), which uses the incorporated small RNA to find and cleave complementary mRNA. Argonaute proteins are made up of two major domains, a PAZ domain and a Piwi domain. The PAZ domain has small RNA binding activity and the Piwi domain, at least in RISC complexes, contains a “slicer” activity that degrades target mRNAs. RDRPs are thought to participate in RNA silencing processes by forming dsRNA for Dicer processing or by directly forming small RNAs for incorporation into effector complexes.

In the last 10 years a few specific biological phenomena in fungi have been linked to RNA silencing. These include cosuppression in Neurospora crassa (quelling), meiotic silencing of unpaired DNA in N. crassa, and some types of heterochromatic silencing in Schizosaccharomyces pombe (RNAi-mediated heterochromatic silencing). Most of what is known about fungal RNA silencing stems from work on these specific phenomena. How common these processes are across the Fungal Kingdom is unknown, but analysis of available fungal genomes suggests that they cannot be fundamentally conserved.
RNA silencing gene evolution in fungi is complex. For example, the basidiomycete *Phanerochaete chrysosporium* contains three Dicer encoding genes and seven Argonaute encoding genes while the basidiomycete *Ustilago maydis* contains neither of these core RNA silencing genes. The lack of RNA silencing genes in *U. maydis* is not unusual. Dicers and Argonautes are also not found in the genome of the budding yeast *Saccharomyces cerevisiae* and Dicers are not found in the genomes of the opportunistic animal pathogens *Candida tropicalis* and *Candida albicans*. It should not be possible for these fungi to perform quelling, meiotic silencing or RNAi-mediated heterochromatic silencing, at least in the manner that these processes work in *N. crassa* or *S. pombe*. It is also unknown if these specific processes exist in the many fungi that do contain RNA silencing genes.

We have analyzed the genomes of seven *Aspergillus* species (*A. clavatus*, *A. flavus*, *A. fumigatus*, *A. nidulans*, *A. oryzae*, *A. terreus*, and *Neosartorya fischeri*) and found evidence of RNA silencing gene gain and loss in individual members of this genus [Hammond et al. manuscript submitted and 22]. The mechanisms that led to RNA silencing gene gain and/or loss in aspergilli are unknown, but our analysis of *Aspergillus* genomes suggests that this genus is well suited toward studies of fungal RNA silencing evolution.

This chapter focuses on the current status of RNA silencing related research in *Aspergillus* species, including a review of the literature, a description of the RNA silencing genes found in the *Aspergillus* genome databases, and a discussion on putative natural roles for RNA silencing in the aspergilli.

### 13.2 Experimental RNA Silencing

#### 13.2.1 *Aspergillus fumigatus*

Experimental RNA silencing has been demonstrated in several *Aspergillus* species. The first use of experimental RNA silencing was with the human opportunistic pathogen *Aspergillus fumigatus*. Two genes were targeted in this study, *alb1*, encoding a polyketide synthase required for conidial pigmentation and virulence, and *fks1*, an essential beta(1–3)glucan synthase. Inverted repeat transgenes (IRTs) were designed with approximately 500 base pair (bp) fragments of each target gene, which were placed on either side of a 250 bp GFP fragment. Additionally, a double IRT construct was created with fragments of both genes to silence both targets with a single transgene. The IRTs were driven by the *A. niger* glucoamylase promoter, which is induced in maltose medium and repressed in xylose medium. All three IRTs resulted in a range of silencing phenotypes in the *A. fumigatus* transformants. While most of the transformants revealed an intermediate phenotype, in each case a few transformants (1–5%) appeared to be completely silenced. The basis for the range of silencing was not investigated, but it was proposed that random integration of the IRT could lead to differences in silencing efficiency.

Use of the inducible/repressible promoter system demonstrated that this is a useful tool for experimental RNA silencing in an *Aspergillus* species. However, the lack of total repression of the glucoamylase promoter on xylose medium prevented a switch to a true wild-type phenotype when grown on the repressive media.

#### 13.2.2 *Aspergillus flavus* and *Aspergillus parasiticus*

Subsequent work has demonstrated that experimental RNA silencing works in two additional *Aspergillus* pathogen species. These are the mycotoxin-producing plant pathogens *A. flavus* and *A. parasiticus*. Similar to the aforementioned study with *A. fumigatus*, an IRT consisting of 670 bp fragments of *A. flavus aflR* was placed on either side of a GFP spacer fragment. This IRT was driven by the constitutive *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) promoter. Because the sequences of *A. flavus* and *A. parasiticus aflR* are nearly identical, the same IRT was used to silence *aflR* in both plant pathogens. *AflR* is a transcription factor controlling aflatoxin production, a mycotoxin notorious for causing illness or death in humans and animals that eat aflatoxin-contaminated agricultural products. Introduction of the IRT into *A. flavus* and *A. parasiticus* resulted in suppression
of the aflatoxin biosynthetic pathways in both species when a single copy of the IRT integrated into the genome.\textsuperscript{28} One unexplained finding was that a strain of \textit{Aspergillus} that had integrated multiple copies of the IRT into its genome showed less suppression of the aflatoxin biosynthetic pathway than did any of the single copy integrants.\textsuperscript{28} One possibility is that the integration of multiple copies of the IRT somehow interfered with IRT transcription, perhaps caused by an unknown transcriptional silencing pathway.

### 13.2.3 Experimental RNA Silencing During Infection

In addition to revealing that RNA silencing can be used as an experimental tool in plant pathogenic aspergilli, the aforementioned study suggests that it may be possible to use RNA silencing technology to control mycotoxin production in agriculture. This is because the \textit{A. parasiticus}-silenced strain retained a silenced phenotype during corn infection. Additionally, IRT-based RNA silencing of a \textit{Fusarium graminearum} pathogenicity gene reduced virulence during infection of wheat.\textsuperscript{28} At a minimum, both findings indicate that RNA silencing in these two plant pathogens is active during host infection. If a method of dsRNA delivery to an infecting fungus is discovered, RNA silencing could be adapted as a useful control for fungal pathogens of plants.\textsuperscript{28}

A different study suggests that \textit{A. fumigatus} RNA silencing proteins are also active during infection of its host. In a work by Tsitsigiannis et al.,\textsuperscript{32} three cyclooxygenase-like enzymes were suppressed in \textit{A. fumigatus} using an IRT construct containing fragments of three different cyclooxygenase-like genes. Similar to the \textit{A. flavus} and \textit{A. parasiticus} work described earlier, the IRT was driven by the \textit{A. nidulans} \textit{gpdA} promoter. Presence of the IRT correlated with decreased transcripts of all three target genes in culture, slight decreases in prostaglandin production in arachidonic acid supplemented cultures, and an increase in resistance to reactive oxygen species (ROS).\textsuperscript{32} The IRT also correlated with hypervirulence in an invasive pulmonary aspergillosis murine model system.\textsuperscript{32} While the mechanism of increased virulence is unknown, it is possible that decreased prostaglandin production or increased resistance to ROS was a contributing factor.\textsuperscript{32} This work shows the utility of RNA silencing for efficiently targeting gene families that may contribute to fungal pathogenicity and suggests that RNA silencing processes are active during the host-infection process. As with the agricultural \textit{Aspergillus} pathogens, a dsRNA delivery method that works during the infection process could thus lead to breakthroughs in treating fungal infections of humans and animals.

### 13.3 Genetic Analysis of Experimental RNA Silencing

The aforementioned reports demonstrate experimental RNA silencing as a tool for \textit{Aspergillus} research and/or the possibility of RNA silencing based treatment/control of \textit{Aspergillus} pathogens in medicine and agriculture. Additionally, they use Southern blotting to determine that the endogenous target genes are intact, and mRNA analysis to determine that silencing correlates with the presence of an IRT. While this evidence suggests that the IRTs are causing silencing through the well-characterized dsRNA-mediated RNA silencing pathway, genetic and molecular techniques were not presented to prove this likelihood.

#### 13.3.1 \textit{Aspergillus nidulans} RNA Silencing Model

Genetic and molecular characterization of IRT-based silencing in \textit{Aspergillus} species has been carried out with \textit{A. nidulans}.\textsuperscript{22} Originally selected for genetic work by G. Pontecorvo,\textsuperscript{33} this species is well suited for genetic and molecular work because of its well-characterized sexual cycle and numerous experimental tools. These aspects of \textit{A. nidulans} allowed for the creation of an \textit{Aspergillus} RNA silencing genetic model.\textsuperscript{22} This model is based on visual assay for RNA silencing activity. Similar to \textit{A. flavus} and \textit{A. parasiticus} aflR, \textit{A. nidulans} aflR is a transcription factor required for production of the mycotoxin sterigmatocystin, the penultimate precursor to aflatoxin.\textsuperscript{34} An intermediate in the sterigmatocystin/aflatoxin pathway...
The Aspergilli

The fungal kingdom as a whole is variable with regard to RNA silencing genes, with some fungi showing a dramatic expansion of RNA silencing genes and others showing no evidence of their presence. The genus Aspergillus is represented by seven easily accessible genome sequences, including A. clavatus, A. flavus, A. fumigatus, A. nidulans, A. oryzae, A. terreus, and N. fischeri. Analysis of Dicers, Argonautes and RDRPs in these genomes reveals that RNA silencing gene evolution in the aspergilli shows some of the complexity found in the fungal kingdom as a whole. This aspect of the pathway is norsolorinic acid (NOR) and A. nidulans strains deleted of the gene encoding a sterigmatocystin biosynthetic enzyme SteC accumulate NOR. Because NOR is a bright orange compound, it can be directly visualized under certain culture conditions and it can be easily analyzed by thin-layer chromatography. Thus, A. nidulans strains carrying an IRT consisting of aflR fragments should be inhibited in NOR production if the IRT is capable of suppressing native aflR transcripts.

To eliminate the problem of the transgene inserting into an unknown ectopic location, an A. nidulans aflR specific IRT, consisting of 1300 or 900 bp aflR fragments on either side of a gfp spacer fragment and driven by the constitutive A. nidulans gpdA promoter, was directed to the A. nidulans trpC locus with a 5' fragment of A. nidulans trpC. Transformation with aflR-specific IRTs, containing either 1300 or 900 bp aflR fragments resulted in strains that produced little or no detectable NOR. Southern and northern analysis indicated that the native aflR locus was unaltered in NOR-transformants and that mRNA transcripts were significantly suppressed in IRT-carrying strains. Additionally, analysis of low-molecular weight RNAs demonstrated the presence of aflR-specific siRNAs, a hallmark of RNA silencing. Use of aflR-specific oligonucleotides as migration controls in siRNA analysis indicated that the aflR-specific siRNAs were 25 nucleotides in length. Additionally, the majority of the detected siRNAs must have come from the IRT, not the endogenous aflR transcript, as a probe specific for sequences found in aflR, but not the aflR IRT, did not detect aflR-specific siRNAs.

13.3.2 RNA Silencing Proteins in Aspergillus nidulans

The identification of 25 nt siRNAs suggested that dsRNA-mediated RNA silencing was the mechanism of silencing by IRTs. Further support for this hypothesis was that the aflR IRT only correlated with loss of NOR when an intact Argonaute gene was present. Deletion of this Argonaute by double homologous recombination resulted in aflR IRT-carrying strains that produced qualitatively normal levels of NOR. Thus it was named rsdA, for RNA silencing deficient A. In hindsight, it could have been given a name more indicative of the reverse genetics approach used to find the gene. Orthologs of rsdA in other aspergilli are named ppdA, for Paz and Piwi domain A. In addition to RsdA, another gene known to be required for IRT-based RNA silencing in A. nidulans is the A. nidulans Dicer, dclB [Hammond et al., manuscript submitted].

In contrast to the defined role of RsdA and DclB in A. nidulans RNA silencing, efforts have yet to uncover a role for Aspergillus RDRPs. Genome analysis indicated that A. nidulans encodes two RDRPs, RrpB and RrpC. Comparative genomic analysis of A. nidulans, A. fumigatus, and A. oryzae, suggests that A. nidulans has lost a third RDRP, named rpa in the other aspergilli. Syntenic analysis between the fumigatus rpa encoding region and the analogous region in A. nidulans revealed the presence of a ~4.0 kb fragment of DNA in A. nidulans with low homology to rpa-like RDRPs in Aspergillus species and other filamentous ascomycetes at the nucleic acid level. This locus was not predicted to code for a protein in the A. nidulans annotation. Our own analysis of the locus also failed to identify even a partial ORF indicative of an RDRP. This suggested that an RDRP was present at this locus in an A. nidulans ancestor, but that unknown evolutionary forces led to its degeneration.

Deletion of the two intact A. nidulans RDRPs did not affect IRT-RNA silencing. This is peculiar because a similar process in S. pombe requires an RDRP (Rdp1). The finding that A. nidulans does not require an RDRP for IRT-based RNA silencing while S. pombe does, suggests that there are fundamental differences between A. nidulans and S. pombe RNA silencing proteins or mechanisms.

13.4 Aspergillus RNA Silencing Gene Evolution

The fungal kingdom as a whole is variable with regard to RNA silencing genes, with some fungi showing a dramatic expansion of RNA silencing genes and others showing no evidence of their presence. The genus Aspergillus is represented by seven easily accessible genome sequences, including A. clavatus, A. flavus, A. fumigatus, A. nidulans, A. oryzae, A. terreus, and N. fischeri. Analysis of Dicers, Argonautes and RDRPs in these genomes reveals that RNA silencing gene evolution in the aspergilli shows some of the complexity found in the fungal kingdom as a whole. This aspect of the
RNA Silencing in the Aspergilli

13.4.1 Dicers and Argonautes

RNA silencing genes in the filamentous ascomycetes were proposed to have evolved by duplication events from a single set of ancestral RNA silencing genes, whose diversification led to paralogous groups of RNA silencing genes. This hypothesis was based on analysis of N. crassa, A. fumigatus, and S. pombe fungal genomes. An additional hypothesis that the paralogous groups of RNA silencing genes have diversified in function so that they mediate different RNA silencing processes was also proposed. Genetic characterization of RNA silencing in N. crassa generally supports both hypotheses. Our analysis of Aspergillus genomes supports the duplication hypothesis. In general, two paralogous groups of Dicers and Argonautes exist in the aspergilli, and these groups cluster with the two groups observed in N. crassa, suggesting that they are orthologous [Hammond et al., manuscript submitted]. The majority of Aspergillus species encode two Dicers (four species, dclA and dclB) and two Argonautes (four species, ppdA/rsdA and ppdB) [Hammond et al., manuscript submitted]. However, A. oryzae and A. flavus each encode an additional Dicer and Argonaute protein (dclC and ppdC), while A. nidulans has lost a Dicer and an Argonaute to truncation events [Hammond et al., manuscript submitted]. It appears that one or two evolutionary events have led to duplication of an A. oryzae and A. flavus Dicer and Argonaute gene (orthologous to the Dicer and Argonaute of the N. crassa quelling pathway), while two separate evolutionary events have led to the truncation of an A. nidulans Dicer and Argonaute gene (orthologous to the Dicer and Argonaute of the N. crassa meiotic silencing pathway) [Hammond et al., manuscript submitted].

Dicers are characterized in part by the presence of two C-terminal RNaseIII domains. These domains have been proposed to form a single processing center by intramolecular dimerization of the two RNaseIII domains. Each domain coordinates a metal ion required for cleavage of a phosphodiester bond, together cleaving phosphodiester bonds on either side of a dsRNA substrate. Biochemical and structural analysis of the cleavage process suggests that four “strictly” conserved amino acids in each domain are responsible for metal ion coordination and subsequent phosphodiester bond cleavage. These amino acids are strictly conserved in nearly all of the ascomycete Dicers we have analyzed, (N. crassa, G. zeae, M. oryzae, seven Aspergillus species, S. pombe) (Fig. 13.1). However, in A. oryzae and A. flavus DclC, five out of eight of these amino acids are not conserved (Fig. 13.1), indicating that DclC functions differently than typical dicer-like enzymes (or is nonfunctional). A similar analysis of strictly conserved residues in PpdC has not been performed, but it would be interesting to learn if PAZ or Piwi residues have been similarly modified relative to other Argonaute proteins.

In N. fischeri, a duplication of ppdB has occurred (gi:83742757, contig: AAKE02000029). However, the duplicated sequence is adjacent to a transposon-like sequence, and when aligned to N. fischeri ppdB shows a significant increase in A:T content at variable residues (in 413 variable nucleotides, 30.8% are A:T residues in ppdB while 64.9% are A:T residues in the duplicated sequence, analyzed with MEGA 3.1), as well as four minor deletions (29 bp total) and three minor insertions (7 bp total). RIPPPING, a process that mutates duplicated sequences by C to T transition events in N. crassa, or some variety of RIPPPING, could have contributed to the increase in A:T percentage in the duplicated ppdB. Supporting this hypothesis is the thought that RIPPPING serves as a highly efficient transposon defense mechanism in N. crassa (the duplicated ppdB is found next to a transposon-like sequence) and genomic analysis has uncovered evidence of RIPPPING in other Aspergillus species. A preliminary analysis of the duplicated N. fischeri ppdB sequence suggests that it does not encode a functional Argonaute protein, due to at least six early termination codons resulting from C to T transition mutations [Hammond and Keller, unpublished].

13.4.2 RNA-Dependent RNA Polymerases

To determine if RDRPs were variable in the aspergilli, RDRP-like sequences were obtained from the A. clavatus, A. terreus, A. flavus, and N. fischeri genomes and used to construct a phylogenetic tree with
While analyzing additional filamentous ascomycetes in general (Fig. 13.2). However, published work, there appear to be three general classes of RDRPs in the aspergilli specifically and three RDRP classes (Fig. 13.2). The functions of these RDRPs are unknown.

RrpA, an ortholog to of the six possible reading frames of the degenerate locus. We now know this finding to be inaccurate.

In Hammond and Keller, we determined that unknown evolutionary forces led to the degeneration of the Aspergillus RrpA locus and subsequent loss of RrpA (see earlier). One line of evidence that supported this hypothesis was that an RDRP-specific motif (DbDGD, b is a bulky residue) was not located in any of the six possible reading frames of the degenerate locus, which thus led to its identification. However, we are still unable to identify a putative cDNA identification. Additionally, a translated search of GenBank (blastx) with the analogous genomic region in A. nidulans rrpA (3e-74 and 3e-66) (Fig. 13.2). The fact that a blastx search does not find similar high identity matches with the RrpA ortholog (gi:88184235, 6e-81), and the RrpA orthologs of the fungal RDRPs.

RDRPs identified in previously published reports (Fig. 13.2). In agreement with previously published work, there appear to be three general classes of RDRPs in the aspergilli specifically and filamentous ascomycetes in general (Fig. 13.2). However, A. fumigatus and the closely related species N. fischeri are each missing an RDRP (RrpC) (Fig. 13.2). Additionally, A. nidulans is missing RrpA, an ortholog to N. crassa QDE-1. The other aspergilli were found to encode single genes in all three RDRP classes (Fig. 13.2). The functions of these RDRPs are unknown.

In Hammond and Keller, we determined that unknown evolutionary forces led to the degeneration of the A. nidulans rrpA locus and subsequent loss of RrpA (see earlier). One line of evidence that supported this hypothesis was that an RDRP-specific motif (DbDGD, b is a bulky residue) was not located in any of the six possible reading frames of the degenerate locus. We now know this finding to be inaccurate.

While analyzing additional Aspergillus RDRPs for this work, and reanalyzing the A. nidulans degenerate rpa locus, we were able to predict the probable location of this motif within the A. nidulans degenerate rpa locus, which thus led to its identification. However, we are still unable to identify a putative cDNA using a combination of gene-prediction software (fgenesh, www.softberry.com) and manual intron/exon identification. Additionally, a translated search of GenBank (blastx) with the A. nidulans degenerate rpa locus suggests that the rpa locus has very low identity to A. fumigatus RrpA (3e-13) and A. oryzae RrpA (2e-7). Oppositely, a search of GenBank (blastx) with the analogous genomic region in A. fumigatus results in very high identity matches to Rpa like RDRPs, including A. oryzae Rpa (E = 0), a putative Coccidioides immitis Rpa ortholog (gi:90301548, E = 0), N. crassa QDE-1 (3e-96), a putative Chaetomium globosum Rpa ortholog (gi:88184235, 6e-81), and the Rpa orthologs of M. oryzae (3e-76) and G. zeae (4e-74 and 3e-66) (Fig. 13.2). The fact that a blastx search does not find similar high identity matches with the A. nidulans degenerate rpa locus is supportive of A. nidulans rpa degeneration. However, we have not completely eliminated the unlikely possibility that the A. nidulans rpa locus contains a gene with a complex splicing pattern encoding an RDRP with low homology to Rpa like RDRPs.
RNA Silencing in the Aspergilli

13.5 Possible Roles of Aspergillus RNA Silencing in Nature

The natural function(s) of Aspergillus RNA silencing is unknown. Our own work with RNA silencing mutants in A. nidulans suggests that its Dicer, Argonaute and RDRPs are not required for normal growth or reproduction and we have not observed gross morphological differences in these mutants compared to wildtype [Hammond et al., manuscript submitted]. However, analysis of the fungal RNA silencing literature suggests that there are a number of other possibilities for the Aspergillus RNA silencing function, some of which might not be needed for growth under standard laboratory conditions.

13.5.1 Meiotic Silencing

N. crassa is the most thoroughly characterized filamentous fungus with regard to RNA silencing. One of the best characterized N. crassa RNA silencing phenomena is meiotic silencing by unpaired DNA.\(^\text{18,40,52}\)
Meiotic silencing occurs during the sexual cycle and is activated by unpaired DNA between sister chromosomes. The unpaired DNA nature of meiotic silencing suggests that it is perfectly suited to defend against transposable elements. In addition to a Dicer DCL-1, an Argoanute SMS-2, and an RDRP SAD-1, meiotic silencing requires a novel protein, SAD-2. This protein is required for recruitment of SAD-1 to the perinuclear region.

Whether or not some form of meiotic silencing exists in Aspergillus species is unknown. While it is clear that orthologs of DCL-1, SMS-2 and SAD-1 exist in most of the aspergilli, the orthologs do not function in a meiosis-specific process. Furthermore, a search of GenBank (blastp) with SAD-2 did not identify an obvious homolog in the A. fumigatus (Afu5g03760, 4e-6, putative chitinase) or A. nidulans (AN9132.2, 2e-04, hypothetical protein) genomes. Additionally, results in our lab indicate that RNA silencing genes, and thus meiotic-silencing homologs, are not required for the A. nidulans sexual cycle [Hammond and Keller, unpublished data]. Thus, if Aspergillus meiotic silencing exists, there are likely to be significant differences from the mechanism observed in N. crassa.

13.5.2 Quelling

A second N. crassa RNA silencing process is quelling. Quelling occurs during the vegetative cycle and is activated by high numbers of tandemly arranged transgenes. Quelling requires QDE-2 and QDE-1, and at least one of the N. crassa Dicers, DCL-1 or DCL-2. As with meiotic-silencing orthologs, the seven Aspergillus genome databases suggest that all these species have orthologs for these quelling genes [Hammond et al. manuscript submitted], with A. nidulans as the one exception. A. nidulans carries a degenerate rrpA locus (see earlier) and RrpA is orthologous to N. crassa QDE-1. Thus, if there has been a conservation in function between N. crassa and Aspergillus quelling orthologs, one would expect quelling to exist in most Aspergillus species except A. nidulans. To our knowledge, quelling phenotypes have not yet been encountered in Aspergillus research, or at least investigated in detail. However, a recent report indicates that increasing transgene dosage does not correlate with increases in heterologous protein expression in A. nidulans. Although this phenomenon was not directly connected to RNA silencing, it suggests the possibility of a quelling-like process in A. nidulans. If a quelling-like process is responsible for this phenomenon, it is possible that A. nidulans RrpB or RrpC is involved.

13.5.3 RNAi-Mediated Heterochromatic Silencing

S. pombe encodes a single Dicer (Dcr1), Argonaute (Ago1) and RDRP (Rdp1). This fact has partially contributed to it being the most thoroughly characterized fungus with regard to biochemical aspects of RNA silencing processes. Recent work has revealed a role for Dcr1 and Ago1 in S. pombe cell cycle regulation. It is currently unknown how these RNA silencing proteins regulate the cell cycle, but the mechanism is independent of Rdp1. Most S. pombe RNA silencing research has focused on a mechanism requiring all three RNA silencing genes, the initiation and maintenance of heterochromatic silencing. Also referred to as RNAi-mediated heterochromatic silencing, this process involves a specific type of repeated DNA. This repeated DNA, containing fragments referred to as dh-dg repeats, is found at a number of locations in the S. pombe genome, including the pericentromeric outer regions of all three S. pombe centromeres. The fact that dh-dg specific siRNAs can be isolated from S. pombe, and bidirectional transcripts accumulate in RNAi mutants, suggests that Dicer processing of dsRNA resulting from bidirectional transcripts may initiate RNAi-mediated heterochromatic silencing. At least three protein complexes are required for the process. These include the RNAi-mediated initiation of transcriptional silencing complex (RITS), the RNA-directed RNA polymerase complex (RDRC), and the Rik1-Clr4 complex. The RITS complex includes three proteins, Ago1, Chp1, and Tas3. Chp1 is a chromodomain protein that helps anchor RITS to heterochromatic regions by its affinity for H3K9Me2,
a histone modification associated with heterochromatin, and Tas3 is a novel protein of unknown function.\textsuperscript{62} The RDRC complex contains three proteins, Rdp1, Hrr1, and Cid12.\textsuperscript{63} Hrr1 is an RNA helicase and Cid12 is a polyA polymerase family member. The Rik1/Clr4 complex contains, in addition to Rik1, a WD-propeller-repeat protein,\textsuperscript{65} and Clr4, a histone methyltransferase,\textsuperscript{66} an additional protein known as Dos1.\textsuperscript{64} Models concerning the way these complexes interact are continuously being refined.

Although \textit{S. pombe} Dcr1, Ago1, and Rdp1 tend to cluster between the paralogous groups of filamentous ascomycete RNA silencing proteins in our unrooted phylogenetic trees, rather than clustering with one specific paralogous group (Fig. 13.2, Hammond et al., manuscript submitted, and [22]), it is clear that the filamentous ascomycetes encode homologs of these proteins. To determine if \textit{A. nidulans} and \textit{A. fumigatus} encode obvious homologs of the other proteins required for \textit{S. pombe} RNAi-mediated heterochromatic silencing, we searched (blastp) the \textit{A. nidulans} and \textit{A. fumigatus} genomes. The best matches in each genome are as follows: Tas3 (An0461.2, 1.3e-02; Afu1g04370, 1.4e-01); Chp1 (An6200.2, 3.6e-01; Afu1g02180, 2.1e-01); Hrr1 (An4669.2, 1e-54, a possible annotation error may have resulted in a lower than appropriate score; Afu5g09090, E=0); Cid12 (An5694.2, 2e-10, Afu7g04130, 2e-09); Rik1 (An0596.2, 2e-12; Afu6g10980, 1e-12); Clr4 (An1170.2 9e-53; Afu1g11090 2e-47); Dos1 (AN8282.2, 1e-22; Afu5g04300, 6e-21). Poor matching sequences were retrieved for Tas3 and Chp1, two of the main components of the RITS complex. In fact, a putative \textit{A. nidulans} retrotransposon found at An2616.2 matched Chp1 with a better score (3e-02), possibly due to the presence of a chromodomain found at its C terminal end. While these findings do not eliminate the possibility of a Tas3-Ago1-Chp1 RITS-like complex in \textit{A. nidulans}, they are at least not suggestive of its existence. Sequences with relatively high homology to Hrr1, Cid12, Rik1, Clr4, and Dos1 were found in the \textit{A. nidulans} and \textit{A. fumigatus} genomes. However, these types of proteins are thought to have additional roles in \textit{S. pombe}, and thus their presence in \textit{A. nidulans} and \textit{A. fumigatus} is also not necessarily suggestive of RNAi-mediated heterochromatic silencing in \textit{A. nidulans} or \textit{A. fumigatus}.

Overall, it is currently unknown whether or not RNA silencing genes mediate heterochromatic silencing in the aspergilli. Studies with \textit{N. crassa} suggest that its RNA silencing machinery is not required for normal DNA methylation, H3K9 methylation, HP1 localization and thus, heterochromatin formation.\textsuperscript{67} It is possible that \textit{N. crassa} lost RNAi-mediated heterochromatic silencing after divergence from an ancestor shared with the aspergilli, however, experimental analysis will be needed to determine if RNAi-mediated heterochromatic silencing exists in \textit{Aspergillus} species.

### 13.6 Future Directions

RNA silencing could be advantageous in high-throughput silencing strategies in aspergilli and other fungi. However, recent findings that nonhomologous DNA end-joining mutants are almost completely deficient in nonhomologous recombination\textsuperscript{68} suggest that targeted gene deletions and techniques such as promoter replacement are feasible in many filamentous fungi on a high-throughput scale\textsuperscript{69} including the Aspergilli\textsuperscript{70} (chapter 30 by Osmani, Hynes, and Oakley). Nevertheless, RNA silencing techniques may prove desirable under specific situations, including when the goal is to target several genes in a single transformant, such as was demonstrated for the cyclo-oxygenase family in \textit{A. fumigatus}.\textsuperscript{32} RNA silencing could also be a useful therapy for fungal diseases of humans and animals, as well as a useful control strategy for fungal pathogens in agriculture. Work with \textit{A. fumigatus}, \textit{A. flavus}, and \textit{A. parasticus} suggests that their RNA silencing machinery is active during host infection,\textsuperscript{28,32} but targeted dsRNA or siRNA delivery systems for fungi are required before such treatment/control strategies become a reality.

Genomic analysis indicates that RNA silencing gene evolution in fungi is complex.\textsuperscript{20} Further investigating this complexity should give significant insight into fungal biology and evolution. Although the \textit{Aspergillus} genus shows some of the complexity in RNA silencing proteins that are observed in the fungal kingdom as a whole, there are currently no known biological functions of \textit{Aspergillus} RNA silencing. Elucidating natural RNA silencing phenomena in aspergilli with different RNA silencing proteins is an important first step to understanding the benefits and consequences of losing RNA silencing genes. Circumstantial evidence, as discussed earlier, suggests that the processes of quelling, meiotic
silencing, and RNAi-mediated heterochromatic gene silencing, if they exist in the aspergilli, will not be identical to those of Neurospora or Schizosaccharomyces. This is not too surprising as molecular clock estimates suggest that RNA silencing processes in the Eurotiomycetes (Aspergillus lineage), Sordariomycetes (Neurospora lineage), and Archiascomycetes (Schizosaccharomyces lineage) have had hundreds of millions of years of separate evolution.71

References


14

Hyphal Morphogenesis in Aspergillus nidulans

Steven D. Harris

14.1 Introduction

Filamentous fungi are defined by their ability to form highly polarized hyphae that enable the efficient colonization of diverse environments. Several recent reviews have summarized the cellular functions and mechanisms involved in the morphogenesis of hyphal cells [1–5]. These reviews rightfully emphasize that much of our molecular understanding of hyphal morphogenesis derives from the deep insights that have been acquired into the analogous processes that drive the establishment and maintenance of cellular polarity in the yeast Saccharomyces cerevisiae. However, it has become increasingly apparent that hyphal morphogenesis involves a more elaborate set of core processes that are regulated in ways that differ from yeast. The completion of multiple fungal genome sequences and the development of tools that make it possible to undertake high-throughput functional genomic studies in filamentous fungi (see Ref. 6) will greatly facilitate the analysis of these processes and their regulation. The filamentous fungus Aspergillus nidulans possesses many attributes that make it an ideal system for the characterization of the functions required for hyphal morphogenesis. These include the coordination of morphogenesis with spore germination and the duplication cycle, the availability of numerous morphogenetic mutants, and the ability to rapidly investigate gene function. The goals of this review are to describe the patterns of cellular morphogenesis in A. nidulans, briefly summarize the gene functions known to be involved in morphogenesis, and define some of the important issues to be addressed in the future.
14.2 Patterns of Cellular Morphogenesis in *Aspergillus nidulans*

*A. nidulans* produces dormant conidiospores that possess a single nucleus arrested in the G1 phase of the cell cycle [7]. As the spore germinates, it undergoes a period of isotropic cell surface expansion that appears to be coordinated with nutritional status (Fig. 14.1a, b). On rich glucose media, spores achieve a larger volume and complete a nuclear division prior to the emergence of a hypha, whereas spores germinated in less optimal minimal media typically polarize at a smaller volume without having undergone nuclear division [8]. Upon emergence, hyphae grow in a polarized manner such that cell surface expansion and cell wall deposition are largely confined to the tip region (Fig. 14.1c). Simultaneously, nuclear division proceeds in a parasychronous manner that seems to be uncoupled from tip extension [9]. Once a hypha grows to a sufficient volume (i.e., approximately 80 μm in rich glucose media; [10]), the next round of nuclear division triggers the formation of a septum at its base near the junction with the spore (Fig. 14.1d). Septum formation occurs via invagination of the plasma membrane and the centripetal deposition of new cell wall material [11]. The completed septum retains a small pore that presumably permits cytoplasmic exchange between compartments. Following septation, the tip hyphal compartment continues to progress through the cell cycle, with each round of mitosis coupled to the subsequent formation of septa at sites specified by mitotic spindles [10]. The new subapical compartments that are flanked by septa enter a period of cell cycle arrest that is relieved by the formation of a new hyphal tip that becomes a lateral branch (Fig. 14.1e). Meanwhile, spores usually produce a second hypha in a bipolar pattern (Fig. 14.1e), and occasionally, a third hypha depending on growth conditions. As the sequence of events described above is repeated in each hypha, it becomes clear that a single spore

![FIGURE 14.1](image)

**FIGURE 14.1** Morphogenetic patterns in *A. nidulans*. Note: (a) Dormant conidiospore. (b) During germination, the spore undergoes an initial period of isotropic expansion. (c) A switch to apical growth leads to the formation of a germ tube in which extension is confined to the tip. (d) Tip extension continues even while the septum is being deposited near the junction of the spore and the germ tube. (e) The generation of additional polarity axes permits the formation of secondary germ tubes and/or lateral branches. In each panel, grey arrows indicate active sites of cell wall deposition.
is capable of producing multiple branched hyphae that extend radially to maximize the efficiency of nutrient acquisition.

Although the focus of this review is hyphal morphogenesis, it should be noted that *A. nidulans* produces several other types of polarized cells. This is most obvious during asexual development, which leads to the formation of elaborate conidiophores that harbor multiple tiers of cells that divide by budding. Both metulae and phialides superficially resemble yeast pseudohyphal cells that divide by budding from their tips [12]. In this context, *A. nidulans* should be considered a dimorphic fungus that possesses the molecular machinery needed for growth via the two major modes of cellular morphogenesis in fungi; hyphal growth and budding.

### 14.3 Morphogenetic Paradigm

Over the past decade, a consensus has emerged around the idea that the functions involved in the polarized growth of eukaryotic cells are organized in a hierarchical manner [13,14]. Three broad levels of functions have been defined (Fig. 14.2). First, landmark proteins located on the cell surface specify the site of polarized growth. These proteins may be internal cues that generate polarity, or they may be receptors that respond to external signals to direct cell growth. In either case, the second level of functions typically consists of signal transduction pathways that are locally activated by the landmark proteins. Common elements of these pathways are Rho-related GTPases and their effector protein kinases. The activated signaling pathways direct the localized recruitment of the third level of functions, which are the elements of the cytoskeleton and vesicle trafficking pathways (collectively referred to as the morphogenetic machinery) needed to deliver the components required for cell surface expansion at the specified polarization site. A crucial group of proteins that integrate positional signals with the morphogenetic machinery are modular scaffold proteins that are locally activated by Rho-related GTPases. Notably, the signaling pathways and morphogenetic machinery involved in polarized growth are well conserved from yeast to humans. By contrast, the landmark proteins do not appear to share broad homology, with different

![FIGURE 14.2](image_url)  
**FIGURE 14.2** The morphogenetic paradigm. *Note:* The hierarchy of functions involved in polarized growth is depicted. 1. Landmarks (grey flags) generate positional information based on internal or external cues. 2. Signaling pathways (grey stippled arrow) relay positional information to the morphogenetic machinery. Rho GTPases are typically key components of these pathways. 3. The morphogenetic machinery, composed of the cytoskeleton (dark lines) and associated vesicle transport complexes (black circles), orient in response to the positional signals. Modular scaffold proteins facilitate the localized reorganization of the morphogenetic machinery in response to signals from activated Rho GTPases.
proteins seemingly employed depending on the type of polarized cell (i.e., a budding yeast cell vs. a fungal hypha vs. a migrating fibroblast).

Although the hierarchical organization of polarity functions provides an effective paradigm, recent results suggest that it might be an overly simplistic model. For example, yeast cells are capable of polarity establishment in the absence of any obvious positional signal [15,16]. In this case, polarity is generated via stochastic processes dependent upon feedback loops that amplify initially weak signals to the morphogenetic machinery. These results have two important implications. First, they suggest that polarity establishment may be an intrinsically random process, with the relevant function of landmark proteins being the selective stabilization of axes that emanate from a specific site. Second, they imply that polarity functions may be organized into a complex network composed of multiple interconnected processes [17]. Nevertheless, because our understanding of the molecular mechanisms underlying hyphal morphogenesis in *A. nidulans* and other filamentous fungi are fairly limited, the hierarchical model provides an appropriate framework that is used in each of the following sections.

### 14.4 Hyphal Tip

#### 14.4.1 Landmarks

The mechanism of bud site selection in *Saccharomyces cerevisiae* provides a well-recognized paradigm for how landmark proteins spatially regulate cellular morphogenesis in fungi. Each budding pattern (axial and bipolar) is specified by distinct landmark proteins (Axl2, Bud8, Bud9, Rax2) that share common features, which include association with the cell wall and the presence of short cytoplasmic domains thought to interact with downstream signaling modules [18,19]. Additional proteins may facilitate the function of the landmarks (Bud3 and Bud4 are predicted to promote Axl2 localization; [18]). Despite the conservation of the downstream signaling modules, the landmark proteins are poorly conserved in *A. nidulans*. In some cases, such as Bud8 and Bud9, no obvious homolog can be detected, whereas in others, such as Axl2, Bud3, and Bud4, divergent homologs that presumably share a functional domain are present. Moreover, even when these divergent homologs have been characterized via gene replacement, they clearly have no detectable role in the establishment or maintenance of hyphal polarity (H. Si and S. Harris, unpublished observations). These results suggest that the function of the yeast landmark proteins in the specification of polarity sites does not apply to *A. nidulans*, or for that matter, to other filamentous euascomycete fungi (B. Rittenour, K. Xu, A. Virag, and S. Harris, unpublished observations).

How then are sites of polarized growth specified in *A. nidulans*? Given the absence of any recognizable physical marker that could designate sites of hyphal emergence, it is reasonable to speculate that these sites could be selected at random. In particular, for a saprophytic fungus like *A. nidulans*, the growth direction of the primary hypha emerging from a spore may not matter, as long as the second hypha emerges from the opposite side to generate a bipolar germination pattern [8]. The random generation of a polarity axis in *A. nidulans* may share features with the feedback-based mechanisms characterized in yeast [17]. Moreover, microtubules could conceivably play a key role in this process. The inherent dynamic instability of microtubules may be exploited to mediate the deposition of polarity factors at random sites [20]. The accumulation of these factors in excess of the needed threshold at any given site may then trigger feedback loops that reinforce the signal and generate a polarity axis. At this time, there is scant experimental evidence to support such a model, though it has been noted that specific γ-tubulin mutations prevent spore polarization [21].

#### 14.4.2 Signals

The prevailing paradigm emphasizes the role of Rho-related GTPases in the transduction of positional signals to the morphogenetic machinery. Annotation of the *A. nidulans* genome sequence reveals the presence of six such GTPases; single homologs of Cdc42 and Rac1, plus four distinct Rho homologs. Previous studies have shown that RhoA, a homolog of yeast Rho1, is involved in cell deposition and the formation of lateral branches [22]. More recent characterization of the Cdc42 and Rac1 homologs
demonstrates that Cdc42 is the predominant GTPase that regulates the establishment and maintenance of hyphal polarity, whereas Rac1 appears to play a secondary role (A. Virag and S. Harris, manuscript in preparation). However, these GTPases share at least one common function, because multiple copies of rac1 are capable of suppressing the morphogenetic defects caused by deletion of cdc42. Apparent homologs of the GEFs (guanine nucleotide exchange factors, Cdc24) and GAPs (GTPase-activating proteins, Bem3) that regulate Cdc42 in yeast have been identified in A. nidulans [2], though their specific role in hyphal morphogenesis has not been tested.

In yeast, effectors of Cdc42 include the PAK kinases Ste20 and Cla4, as well as the paralogous proteins Gic1 and Gic2. Notably, these effectors each possess a CRIB domain that mediates interaction with Cdc42. An additional Cdc42 effector characterized in yeast is the formin Bni1, which exists in an inactive conformation that is presumably “opened” by the binding of GTP-bound Cdc42 to a GTPase-binding domain [23]. Genetic analyses show that Bni1 and the Gic proteins function in parallel pathways downstream of Cdc42 [24]. In A. nidulans, the only CRIB domain proteins identified by annotation are homologs of Ste20 (AN2067.2) and Cla4 (AN8836.2), whereas no apparent homolog of Gic1 or Gic2 can be found. Indeed, the Gic proteins appear to be unique to yeast and its close relatives within the hemiascomycetes. Because A. nidulans does possess a formin with a GTPase-binding domain (SepA; [25]), the parallel Gic pathway is presumably an adaptation that is specific to the budding mode of morphogenesis.

Bem1 is a yeast scaffold protein that mediates interactions between Cdc42, its activating GEFs, and its effector PAK kinases. More recent results show that Bem1 is a crucial component of the feedback loops that amplify localized Cdc42 signals [26]. The A. nidulans homolog of Bem1, BemA, displays severe polarity establishment defects that are reminiscent of the yeast bem1 mutant [27]. Because this defect is much more severe than that observed in A. nidulans cdc42 or rac1 mutants, BemA might conceivably serve as a scaffold for additional polarity proteins, including, for example, those that could regulate microtubule organization.

The role of lipid microdomains as signaling platforms at the tips of yeast buds and mating projections remains somewhat controversial. However, their existence at the tips of A. nidulans hyphae has been documented by the use of the sterol-binding probe filipin. Additional genetic evidence supports the idea that these domains might play a crucial role in relaying signals to the morphogenetic machinery. First, mutations in mesA prevent the stable recruitment of the formin SepA to hyphal tips, thereby blocking the formation of a stable polarity axis [28]. Sterol-rich lipid microdomains are disorganized in mesA mutants, leading to the suggestion that the proper organization of these domains is required for formin localization. Second, mutations or chemical perturbations that deplete sphingolipid pools disrupt the formation of lipid microdomains at the hyphal tip and trigger the loss of hyphal polarity [29]. Notably, A. nidulans possesses two distinct ceramide synthases that each contributes to the formation of a stable polarity axis. One of these enzymes, BarA, appears to generate a specialized sphingolipid pool that directs the recruitment of the formin SepA to hyphal tips. Based on these observations, it has been proposed that like neurons [30], A. nidulans may possess distinct lipid microdomains at the hyphal tip, each of which directs the formation of different signaling complexes required for polarized growth.

### 14.4.3 Morphogenetic Machinery

Surprisingly few components of the morphogenetic machinery have been characterized in A. nidulans. Nevertheless, those studies that have been undertaken combined with results from genome annotation support the view that the morphogenetic machinery is more complex in A. nidulans than in yeast. A survey of actin-associated proteins provides a good example of this point. Although fimbren (Sac6) is the only microfilament bundling proteins found in yeast, it is not essential for viability, and the deletion of SAC6 only causes obvious morphogenetic defects when combined with mutations in other actin-associated proteins [31]. In contrast, A. nidulans possesses two distinct microfilament bundling proteins, fimbren (AN5803.2) and alpha-actinin (AN7707.2). Furthermore, mutations in either of these genes causes significant polarity defects (B. Shaw, personal communication; A. Virag and S. Harris, unpublished results), leading to the notion that A. nidulans may possess distinct populations of bundled microfilaments that are each required for polarized growth. A similar picture emerges from the annotation of actin-severing proteins. In yeast, Aif1 and coillin act together to provide the primary source of microfilament
severing activity [32]. Whereas *A. nidulans* possesses homologs of these two severing proteins (AN7448.2 and AN2317.2, respectively), it also contains two members of the gelsolin/severin/fragmin family that are not found in yeast (AN1306.2 and AN0837.2). Members of this family are capable of severing and capping microfilaments via a Ca\(^{++}\)-dependent mechanism that is presumably distinct from that of cofilin. This observation suggests the need for tighter regulation of microfilament assembly and disassembly in hyphal cells compared to yeast.

Another key distinction between *A. nidulans* and yeast is the role of cytoplasmic microtubules in polarized growth. It has long been known that the loss of microtubules in no way compromises the ability of yeast cells to establish and maintain polarity [33]. Although microtubules are similarly dispensable for the establishment of hyphal polarity in *A. nidulans*, the failure of specific gamma-tubulin mutants to undergo spore polarization suggests that microtubules may be involved in establishing a polarity axis [21]. Furthermore, in growing hyphae, microtubules are required to sustain maximal rates of tip extension and to maintain the direction of growth [34]. Both of these functions presumably reflect the importance of kinesin-mediated anterograde transport of vesicles along microtubules to the hyphal tip. Notably, in the absence of the kinesin KipA, hyphae grow in curves due to the mislocalization of the tip-localized vesicle supply center known as the Spitzenkorper [35]. This may be caused by the failure to deliver landmark proteins associated with microtubule plus ends to the hyphal tip. Further characterization of conserved plus end-binding proteins (CLIP-170, EB1, APC) may yield additional insight into this possibility [36]. At the same time, novel mechanisms, such as signaling via the ATM kinase [37], may also regulate microtubule organization at the hyphal tip to control both the rate and direction of extension.

### 14.5 Septum Formation

#### 14.5.1 Landmarks

Two general strategies are employed by cells to determine the site of cytokinesis [38]. Some cells utilize cortical markers to determine the division plane. For example, in yeast, the septation site is specified early in the cell cycle by the bud site selection proteins [18]. Cortical specification of the division site may be a unique mechanism that reflects constraints imposed by certain cell shapes such as yeast buds. On the other hand, most cells utilize the mitotic nucleus to determine the division plane [38]. In the well-studied fission yeast *Schizosaccharomyces pombe*, the septation site is specified shortly after the start of mitosis when the export of Mid1 from the nucleus to the overlying cell cortex is completed [39]. Mid1 is thought to promote the subsequent formation and stabilization of the contractile actin ring (CAR) by recruiting type II myosin to the septation site. In most animal cells, the division plane is specified during mitosis (mid-anaphase) by the central spindle. This complex comprises a set of bundled microtubules that serve to concentrate components required for assembly of the CAR [38]. These components include the central spindle complex and the aurora kinase complex, which ultimately promote the localized activation of RhoA to trigger localized actin assembly by formins. *A. nidulans* possesses a protein that might be functionally analogous to Mid1 (AN6150.2), and also possesses likely homologs of central spindle components such as the kinesin MKLP1 (AN3721.2, which has no apparent homolog in yeast; [40]), the Rac GTPase activating protein MgcRACGAP (AN1025.2), and the Rho guanine nucleotide exchange protein ECT2 (AN4719.2). How, then, is the septation site specified in *A. nidulans*?

By exploiting conditional mutants defective in nuclear distribution (*nud* mutants) and others defective in mitosis (*nim* mutants), Wolkow et al. [10] demonstrated that nuclei specify sites of septum formation in *A. nidulans*. Further analyses revealed that persistent mitotic signals are required for the assembly of the CAR [41]. These observations are consistent with the idea that, like animal cells, a central spindle complex activates a signaling pathway that provides a spatial landmark for CAR formation. The specific Rho GTPase that could be targeted by this pathway remains to be determined. Although *A. nidulans* RhoA is the best homolog of RhoA, characterization of *rhoA* mutants did not uncover a role in septation [22]. However, *A. nidulans* does possess a conserved homolog of *N. crassa* rho-4 (AN2687.2), which was recently shown to be necessary and sufficient for CAR formation [42]. Accordingly, it is tempting to speculate that a central spindle complex could trigger local activation of the formin SepA via Rho4.
A. *nidulans* hyphal cells are multinucleate, and undergo a parasympathetic wave of mitosis that is followed by the formation of a limited number of septa at appropriately spaced intervals [43,44]. Therefore, despite their role as landmarks for septation, not all mitotic spindles trigger the formation of a CAR (otherwise hyphal cells would be uninucleate). This implies that an additional mechanism limits the ability of mitotic spindles to specify a septation site. The nature of the mechanism remains a mystery, but it could conceivably involve cortical markers. Although it has been shown that displaced tip markers can designate septation sites in some filamentous fungi (*Ashbya gossypii*; [45]), this seems unlikely in *A. nidulans* because the analysis of nuclear distribution mutants demonstrates that septa can form at almost any location in a hyphal cell [10]. Instead, cortical signals may act in a repressive manner to spatially restrict the ability of mitotic spindles to specify a septation site. For example, a specified septation site may generate local signals that block nearby spindles from activating adjacent cortical regions. The ultimate test of this model will likely require the isolation and characterization of mutants that display hyperseptation.

14.5.2 Signals

Two separate studies implicate the conserved SIN/MEN pathway in the transduction of mitotic signals to the cortical septation site [46,47]. This pathway has been extensively characterized in budding yeast, where it controls mitotic exit (i.e., mitotic exit network; MEN), and in fission yeast, where it regulates the initiation of septum formation (septation initiation network; SIN) [48]. Key features of the pathway in *S. pombe* include an upstream Ras-related GTPase (Spg1) that activates a downstream protein kinase cascade (Cdc7 to Sid1 to Sid2). Two additional proteins, Cdc14 and Mob1, are subunits associated with Sid1 and Sid2, respectively. Both polo kinase (Plo1) and cyclin-dependent kinase (Cdc2/Cdc13) regulate activation of the SIN pathway and its coordination with mitotic events. The ultimate readout of SIN pathway activation in *S. pombe* is the relocalization of the Sid2/Mob1 complex from spindle poles to the already formed CAR, where it presumably provides the signal for ring contraction and septum deposition. Possible effectors that mediate this readout include Cdc15, a conserved component of the contractile actin ring [49], and Etd1, a novel fission yeast protein [50].

Genome annotation revealed the conservation of the entire SIN pathway in *A. nidulans* (Table 14.1). SepH was originally identified by a temperature-sensitive mutation that blocked septation [51]. Bruno et al. [46] subsequently showed that SepH is a homolog of fission yeast Cdc7, and, unlike *S. pombe*, that it is required

### TABLE 14.1

Conservation of the SIN/MEN Pathway in *A. nidulans*

<table>
<thead>
<tr>
<th>Gene (Sp)</th>
<th>Gene (Sc)</th>
<th>Function</th>
<th><em>A. nidulans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>plo1</td>
<td>CDC5</td>
<td>Polo kinase</td>
<td>AN1560.2 (plkA)</td>
</tr>
<tr>
<td>??</td>
<td>LTE1</td>
<td>Ras GEF</td>
<td>AN3092.2</td>
</tr>
<tr>
<td>spg1</td>
<td>TEM1</td>
<td>GTPase</td>
<td>AN7206.2 (asgA)</td>
</tr>
<tr>
<td>cdc16</td>
<td>BUB2</td>
<td>GAP</td>
<td>AN0281.2</td>
</tr>
<tr>
<td>byr4</td>
<td>BFA1</td>
<td>GAP</td>
<td>AN9413.2</td>
</tr>
<tr>
<td>cdc7</td>
<td>CDC15</td>
<td>Kinase</td>
<td>AN4385.2 (sepH)</td>
</tr>
<tr>
<td>sid1</td>
<td>Absent</td>
<td>Kinase</td>
<td>AN8033.2</td>
</tr>
<tr>
<td>cdc14</td>
<td>Absent</td>
<td>Kinase reg</td>
<td>AN0655.2</td>
</tr>
<tr>
<td>sid2</td>
<td>DBF2</td>
<td>Kinase</td>
<td>AN8751.2 (sidB)</td>
</tr>
<tr>
<td>mob1</td>
<td>MOB1</td>
<td>Kinase reg</td>
<td>AN6288.2 (mobA)</td>
</tr>
<tr>
<td>clp1/tpl1</td>
<td>CDC14</td>
<td>Phosphatase</td>
<td>AN5057.2</td>
</tr>
<tr>
<td>Absent</td>
<td>NET1</td>
<td>Cdc14 seq</td>
<td>No hit</td>
</tr>
<tr>
<td>??</td>
<td>AMN1</td>
<td>MEN reg</td>
<td>No hit</td>
</tr>
<tr>
<td>dma1</td>
<td>Absent</td>
<td>SIN reg/SAC</td>
<td>AN6908.2</td>
</tr>
<tr>
<td>sid4</td>
<td>Absent</td>
<td>SPB linker</td>
<td>No hit</td>
</tr>
<tr>
<td>cdc11</td>
<td>NUD1</td>
<td>SPB/aMTs</td>
<td>AN2459.2</td>
</tr>
</tbody>
</table>

*key:* reg = regulation, seq = sequestration, GEF = guanine nucleotide exchange factor, SAC = spindle assembly checkpoint, aMTs = astral microtubules.
for CAR assembly. Additional studies have revealed that SepH also acts upstream of the formin SepA, which is required for CAR formation, and the septin AspB [52,53]. These results support a model whereby the SIN pathway directs CAR formation at cortical septation sites in response to mitotic signals. Note that this stands in contrast to fission yeast, where the SIN pathways function after the CAR has formed.

How does the SIN pathway promote assembly of the CAR in *A. nidulans*? Kim et al. [47] recently demonstrated that MobA behaves like its homolog Mob1 in that it relocalizes from spindle poles to the septation sites concomitant with spindle disassembly. However, unlike *S. pombe*, MobA constricts, suggesting that it is associated with the CAR. Similar results were obtained for SidB, a homolog of the Sid1 kinase that associates with Mob1 in fission yeast. Perhaps the SidB/MobA complex functions as a SIN effector that locally recruits components of the morphogenetic machinery to the septation site. Moreover, a Rho GTPase such as the rho-4 homolog described above could mediate this effect. Clearly, the identification of MobA and SidB binding partners should provide valuable insight into the connection between the SIN pathway and CAR assembly in *A. nidulans*.

### 14.5.3 Morphogenetic Machinery

Like animal cells, *A. nidulans* forms a postmitotic CAR that presumably guides deposition of the septal cell wall. Comparatively little is known about the composition of the CAR and the regulation of its assembly [11]. Nevertheless, limited insights suggest significant similarities to fission yeast and animal cells. For example, like these other assemblies, assembly of the CAR appears to be largely driven by the formin-mediated nucleation of unbranched actin filaments. Both SepA and its catalytic partner BudA localize to and constrict with the CAR [52,54]. In addition, SepA is absolutely required for CAR formation [25,52]. Additional proteins associated with the *A. nidulans* CAR include tropomyosin and alpha-actinin ([28]; A. Virag and S. Harris, unpublished observation), both of which bind solely to unbranched actin filaments. The possible roles of actin patches and the Arp2/3-dependent formation of branched actin filaments have not been tested in *A. nidulans* or any other filamentous fungus. However, recent results in fission yeast and animal cells suggest that they are not required for CAR assembly [55].

In both budding and fission yeast, the septins play a key role in septum formation [56]. Functions attributed to the septins include compartmentalization of the plasma membrane at the septation site, localized recruitment of cell wall biosynthetic complexes, and coordination of septation with mitosis. Not surprisingly, septins are involved in septum formation in *A. nidulans*. The septin AspB localizes to septation sites, where it forms a single ring that subsequently splits and envelopes the constricting CAR [53]. Note that the AspB rings themselves do not constrict, and are thus not components of the CAR. Although the specific function of AspB remains unknown, it seems likely that it helps target components of the morphogenetic machinery to the septation site. *A. nidulans* possesses several additional septins whose role in septum formation has yet to be characterized [57].

A primary component of the fungal septum is chitin [58]. In *A. nidulans*, the respective class I and class II chitin synthases, ChsC and ChsA, are involved in septum formation and localize to the septum. Both proteins appear to associate with the CAR prior to constriction, and continue to colocalize with the CAR as it constricts and the septum is deposited [59]. In addition, *A. nidulans* possesses two novel chitin synthases, CsmA and CsmB that are hybrid proteins with an N-terminal myosin motor domain. Like ChsA and ChsC, CsmA and CsmB colocalize with the CAR during constriction and disappear once septum deposition is complete [60,61]. The myosin motor domain of CsmA binds to actin filaments though it likely does not have ATPase activity [60], leading to speculation that it may provide a means for targeting the chitin synthase activity to CARs (or hyphal tips) via actin filaments. Notably, homologs of CsmA and CsmB do not exist in fission or budding yeast.

### 14.6 Patterns of Morphogenesis During Development

Although the focus of this review is hyphal morphogenesis, it should be noted that *A. nidulans* converts to a budding mode of growth during later stages of asexual development. In particular, the elastic acroetal division pattern that produces chains of conidiospores from a phialide bears a striking resemblance to
yeast budding [12]. This is notable given that elements of the yeast bud site selection machinery involved in axial budding are weakly conserved in *A. nidulans* [2], including Bud3 (AN0113.2), Bud4 (AN6150.2), and Bud10/Axl2 (AN1359.2). These predicted proteins could conceivably be involved in hyphal morphogenesis. However, preliminary observations suggest that they have a direct role in cellular morphogenesis during asexual development (H. Si and S. Harris, unpublished results). For example, deletion of AN1359.2 appears to limit philalides to the production of a single spore instead of an entire chain. A common feature of yeast Bud3, Bud4, and Bud10 is their interaction with septins and their possible role in septin organization at the mother-bud junction [62]. Expression of the *A. nidulans* septins correlates with asexual development [57], and one of them, AspB, is required for normal coniation and localizes to conidiophores [53]. Thus, the interaction of Bud3, Bud4, and Bud10 with septins may be a conserved feature of budding morphogenesis that determines growth patterns during conidiation in *A. nidulans*.

### 14.7 The Future

Our understanding of the mechanisms underlying hyphal morphogenesis in *A. nidulans* has proceeded at a steady pace over the past decade. However, the availability of a complete and fully annotated genome sequence marks a pivotal transition that will greatly accelerate these studies. We now know the extent to which the relevant landmarks, signaling modules, and components of the morphogenetic machinery are conserved with budding yeast, fission yeast, animals, and plants. Moreover, the function of many of these conserved proteins in hyphal morphogenesis will no doubt soon be known. Nevertheless, to fully leverage this information and obtain a deep understanding of the molecular processes involved in hyphal morphogenesis, the following steps should be considered. First, the network of protein interactions that underlies hyphal morphogenesis must be determined. As demonstrated in budding yeast [63], this provides necessary insight into the composition of the functional modules involved in morphogenesis. A number of complementary approaches could be employed to determine the so-called morphogenetic network, including comprehensive affinity purification, yeast two hybrid, or synthetic gene interaction screens. Second, the dynamic behavior of the functional modules involved in hyphal morphogenesis should be characterized. Labeling key components of each module with GFP and various derivatives will permit real-time imaging studies that reveal the dynamic pattern of their interactions in space and time. This would also make it possible correlate the localization of specific modules with the behavior of the Spitzenkorper at hyphal tips [3]. Finally, although a valuable guide, the morphogenetic paradigm based on a linear hierarchical pathway is probably an overly simplistic model. Recent studies in yeast and animals highlight the role of multiple feedback loops and stochastic processes in the regulation of cellular morphogenesis [17]. Similar regulatory modes are undoubtedly involved in hyphal morphogenesis as well, and may indeed account for the robust ability of *A. nidulans* to form polarized hyphae across a diverse range of environmental conditions.

### References

The Aspergilli


Cytoskeleton, Polarized Growth, and the Cell Cycle in Aspergillus nidulans

Reinhard Fischer, Norio Takeshita, and John Doonan

CONTENTS

15.1 Introduction ............................................................................................................. 223
15.2 Microtubule Cytoskeleton ........................................................................................ 224
15.2.1 Fungicide-Resistance Genes Identify Tubulins in Aspergillus nidulans .............. 225
15.2.2 Organization of the Interphase Microtubule Cytoskeleton .................................. 226
15.3 Origin of Microtubules .............................................................................................. 227
15.3.1 Microtubule Plus End .......................................................................................... 229
15.3.2 MT Lattice ........................................................................................................... 231
15.3.3 MT-Dependent Motor Proteins .......................................................................... 231
15.3.4 Cell-End Markers at the Cortex ........................................................................ 233
15.4 Cell-Cycle Controls ................................................................................................ 234
15.4.1 Genetic Analysis of the Nuclear Division Cycle .................................................. 235
15.4.2 Regulation of the G2/M Transition .................................................................... 236
15.4.3 Involvement of Calcium in the G2/M Transition ................................................ 238
15.4.4 APC and the Metaphase-Anaphase Transition ................................................... 238
15.4.5 Anaphase ............................................................................................................. 239
15.4.6 Mitotic Exit and the Septation Initiation Network ............................................. 241
15.4.7 Genome Surveillance ......................................................................................... 242
15.5 Hyphal Morphogenesis and the Cell Cycle .............................................................. 243
15.6 Branching and Cell Cycle Control .......................................................................... 244
15.6.1 Developmental Regulators Impose New Discipline on the Cell Cycle ............ 244
15.7 Actin Cytoskeleton ................................................................................................. 246
15.7.1 Organization of the Actin Cytoskeleton ............................................................ 246
15.7.2 Polarisome ......................................................................................................... 246
15.7.3 Actin-Dependent Motor Proteins ...................................................................... 247
15.8 Genes Required for the Establishment of Polarity .................................................. 247
15.9 Conclusions ............................................................................................................. 248
Acknowledgments .......................................................................................................... 248
References ......................................................................................................................... 248

15.1 Introduction

Microtubules (MTs), filamentous actin (F-actin), and their associated motor proteins, kinesin, dynein, and myosin, play important roles in all eukaryotes providing cells with a dynamic structural framework called the cytoskeleton. The cytoskeleton plays crucial roles in many processes that require reorganization of the
cytoplasm, such as growth, nuclear division, and cell division. In this chapter, we review the organization of the cytoskeleton in filamentous fungi, its role in polarized growth, mitosis and cell division. We focus on *Aspergillus nidulans* because work on this species has provided major insights in this area, most of which pertains to eukaryotes generally, but other fungal systems are mentioned and compared throughout the chapter.

Genetic, biochemical, and cell biological approaches in *A. nidulans* and other fungi have provided many important insights into MT functions over the years and continue to lead to new views of many MT-related processes. For example, there is increasing evidence that MT cables, as visualized by immunostaining or GFP-tubulin fusion proteins, consist of several MTs and their dynamics differ in fast-growing hyphal tips as compared with young germlings. Whereas the spindle-pole bodies were considered as the only, or the main, MT organizing centers (MTOCs) in filamentous fungi, additional MTOCs lying outside the nuclei are contributing to the generation of the complex MT array. In addition to new insights into the MT network and its dynamics, the roles of several kinesins have been elucidated recently and their interplay with dynein investigated. Furthermore, it has become clear that MT functions are interwoven with those of the actin cytoskeleton and that three main structures are required at the tip for polarized growth, the Spitzenkörper (vesicle supply centre), the polarisome, and cell end (tip) markers at the cortex.

Another important function for MTs is in mitosis. The spindle derives its structure from a highly organized set of MTs that is generated between two MTOCs, the spindle-pole bodies, which are embedded in the nuclear envelope. At the beginning of mitosis, the spindle-pole bodies are activated by a regulatory kinase network that allows the innermost face of the spindle pole body to act as a MTOC and kinesin proteins to provide the motive power to drive pole separation, producing a bipolar scaffold on which the chromosomes are separated.

For the purpose of the studies described here, hyphal growth starts with the germination of a conidiospore, a uninucleate haploid asexual spore produced at conidiophores. This dormant cell has a thick, resistant and highly pigmented wall that is hydrophobic and adapted to spreading across liquid surfaces. On landing in a suitable environment, the cells take up water, grow, and enter their first cell cycle (see also Fig. 15.4). The first hyphal nuclear division cycle takes about 75–120 minutes, depending on growth conditions. The nuclear division cycle can be considered as four sequential phases, Gap1 (G1), S-phase when the genomic DNA is replicated, Gap2 (G2) and mitosis when the replicated DNA is separated on the mitotic spindle. The dormant spore is arrested in G1 with a highly condensed nucleus and no detectable MTs or actin filaments, but after 4–5 hours on a suitable media, will swell to several times its original size and enter S-phase, replicating its DNA. During this time, cytoskeletal elements appear and actin accumulates at the incipient point of growth. Mitosis is estimated to last 5 minutes at 37°C; G2, 30 minutes; S-phase, 25 minutes; and G1, 15 minutes (Bergen and Morris, 1983; Bergen et al., 1984). Under different growth conditions the duration of G1 and G2 phases vary but the length of mitosis and S-phase remain constant (Bergen and Morris, 1983). In this chapter, we focus mostly on mitosis and controls associated with entry into mitosis. Mitosis is a critical part of the cell cycle and involves the dramatic and highly coordinated rearrangement and separation of nuclear components. Sister chromatid separation on a spindle of MTs, the central and essential feature of nuclear division, has many common features across all eukaryotes and studies in *A. nidulans* have revealed many useful insights into the underlying mechanisms.

### 15.2 Microtubule Cytoskeleton

Microtubules are hollow tubes composed of 13 protofilaments, each of which is made up with the heterodimer αβ-tubulin, as the building block. MTs have an inherent instability but under suitable conditions can continuously elongate at their plus end, where αβ-tubulin dimers are added. One parameter that determines the elongation rate is the concentration of tubulin dimers in the cell. Both tubulin subunits contain a bound GTP. The nucleotide-binding pocket on α-tubulin is located at the interface between the α and β-tubulin subunits and thus, this GTP is rather stable. On the other hand, GTP in the β-tubulin subunit is exposed and easily undergoes hydrolysis. Once β-tubulin contains GDP, further assembly is blocked and the MT is prone to catastrophic disassembly (Nogales and Wang, 2006).
15.2.1 Fungicide-Resistance Genes Identify Tubulins in Aspergillus nidulans

Mutations in the α and β-tubulins were amongst the first cell cycle mutations to be characterized at the molecular level in A. nidulans. Screens for fungicide resistant mutants (Davidse and Flach, 1977, 1978) produced a number of strains resistant to growth on benomyl, an antimicrotubule drug. Strains resistant and sensitive to benomyl were shown to produce altered tubulin proteins (Gambino et al., 1984) that respectively either increased or decreased the stability of MTs. In both cases the cells arrested in mitosis, demonstrating that MTs were essential for mitosis. While the mitotic block caused by fragile MTs was not surprising, the mitotic block in benA33 strains where MTs are unusually stable (Jung et al., 1998; Oakley and Morris, 1981) indicated the importance of MT turnover for mitotic progression. Strains with hyperstable MTs arrest with persistent spindles. The mitotic spindle, therefore, was shown to be a highly dynamic structure, not just a passive scaffold on which the chromosomes were separated. Subsequent work has shown that the organization of the spindle is actively monitored by checkpoint mechanisms that are intimately involved in regulating all stages of mitosis.

The benA gene encodes two of the three β-tubulin isotypes (Sheir-Neiss et al., 1976; Sheir-Neiss et al., 1978). The other β-tubulin gene (tubC) plays a specialized but nonessential role in conidiation (May et al., 1985; Weatherbee et al., 1985). A. nidulans has two γ-tubulin genes, tubA and tubB. Mutations in tubA were identified as suppressors of benA-mediated benomyl resistance (Oakley et al., 1987). Molecular disruption of the tubA gene leads to a mitotic block in vegetative cells (Doshi et al., 1991), while disruption of the other α-tubulin gene, tubB, leads to a block in meiosis (Kirk and Morris, 1991). tubA encodes the major vegetative α-tubulin protein while tubB is highly expressed during sexual development, so the most likely reason for the differences in phenotype is differential expression, rather than any major functional difference (Kirk and Morris, 1993).

Suppressor analysis of the benA33 mutation uncovered a new member of the tubulin superfamily, mipA or γ-tubulin (Weil et al., 1986), which has a crucial role in MT organization and mitosis. Biochemical analysis of MTs, from A. nidulans (Weatherbee and Morris, 1984) as well as a variety of other sources, established long ago that the basic backbone consisted of equimolar amounts of α- and β-tubulin molecules, but failed to detect this novel and crucial member of the family. Until the recent advent of highly sensitive mass-spectroscopy-based methods of protein identification, biochemical approaches were simply not sensitive enough to routinely identify unsuspected minor components in such preparations and so important regulatory proteins such as γ-tubulin were not found. γ-tubulin, clearly related to both α-tubulin and β-tubulin, was sufficiently distinct from both to define a completely new class of tubulin (Oakley and Oakley, 1989) that has since been shown to be crucially important for MT organizing centers (MTOCs) in other eukaryotes (Horio et al., 1991; Joshi et al., 1992; Liang et al., 1996; Martin et al., 1997; Stearns et al., 1991). γ-tubulin is localized at the spindle poles, where it is necessary for normal MT assembly during both interphase and mitosis (Oakley et al., 1990). Its SPB location, the phenotype of cells lacking it, and the genetic evidence that mipA interacts with β- and not α-tubulin led to a now widely accepted model whereby γ-tubulin determines both the location and polarity of MT initiation (Oakley, 1992). γ-tubulin forms the basis of a high-molecular weight complex known as the γ-tubulin ring complex (γTuRC) that provides a template for MT assembly. Some ring complexes are embedded in structures such as the spindle pole body, but others are more dispersed (see later).

γ-tubulin may also have a checkpoint function. Cells with a mutant allele of γ-tubulin were originally reported to have a similar mitotic index to that of freely cycling wild type cells, suggesting that they cannot monitor successful completion of mitosis. mipAD159, another allele, allows spindles to form, but anaphase A is delayed, and late mitotic events are defective (Prigozhina et al., 2004). However, careful reexamination of γ-tubulin deletion strains indicate that nuclei arrest with condensed chromatin for about one cell cycle (Martin et al., 1997). Although spindle assembly is completely abrogated, other aspects of mitotic entry, such as SPB phosphorylation and chromatin condensation occur normally. Interestingly, the authors report that γ-tubulin is not required for cytoplasmic MT assembly, although these are abnormal.

Coordination of the complex series of events that occur during mitosis involves checkpoint controls that monitor spindle function. Thus, defects in the spindle may lead to prolonged chromatin condensation because the checkpoint pathway can sense that mitosis is incomplete and prevents a return to interphase. Mutation screens based on this logic identified the BUB genes, originally in the budding yeast, that are
required to monitor MT function (Hoyt et al., 1991; Li and Murray, 1991). These mutants, which are super-
sensitive to antiMT drugs, fail to arrest if progress through mitosis is delayed. Ascertaining that mitosis has
been correctly and completely executed is, therefore, a critical checkpoint in the cell cycle. Under normal
circumstances antiMT drugs block the cell cycle at M, probably because a crucial checkpoint that monitors
the completion of mitotic events has not been satisfied. Similar genes have been found in A. nidulans, as
the result of a screen for synthetic lethal mutants aimed at understanding the function of cytoplasmic
dynein (Efimov and Morris, 1998). Cytoplasmic dynein is a MT motor protein involved in vesicle trans-
port, mitosis, nuclear migration, and spindle orientation, and dynein mutations impair nuclear migration.
Synthetic lethal mutations that significantly reduced growth in the absence of dynein mapped to nine
different genes. Mutations in sldA and sldB also confer hypersensitivity to the MT-destabilizing drug
benomyl and are in genes homologous to the checkpoint genes BUB1 and BUB3. sldA and B mutations are
also synthetically lethal when combined with mutations in the bimC kinesin (see later).

15.2.2 Organization of the Interphase Microtubule Cytoskeleton

Microtubules are visible in fixed cells by immunolocalization light microscopy (Bourett et al., 1998;
Czymmek et al., 1996; Fischer and Timberlake, 1995) or by electron microscopy (Jung et al., 1998) but
these methods do not reveal the highly dynamic MT behaviors that occur in living cells. In vitro studies,
using mammalian brain tubulin, have shown that MT behavior is complex and their organization can be
modified by several mechanisms. These mechanisms include treadmilling, where subunits tend to fall off
the minus end and are added to the plus end. However, many MTs have their minus ends capped by virtue
of having them embedded in a MTOC, in which case a process called dynamic instability may be more
important. Dynamic instability describes the process by which a plus end can alternate between growth
and disassembly (Nogales and Wang, 2006). MTs can also interact with each other due to the action of
MT-associated proteins that can either crosslink different MTs, or facilitate sliding of one MT relative to
another (MacRae, 1992).

Direct observation of MT behavior became possible after the discovery of the green fluorescent
protein, which was fused to tubulin and expressed in cells. In Saccharomyces cerevisiae interphase cells,
short MTs are attached to nuclei and their growth toward the cortex and subsequent shrinkage causes
short-distance movement of the nuclei. The situation changes once the yeast cell enters the division cycle.
The nuclear spindle pole body divides, and as the two daughter organelles move to opposite sides of the
nucleus, they nucleate the spindle MTs. The spindle- pole bodies span the nuclear envelope and, from
their cytoplasmic faces, also nucleate cytoplasmic MTs that in turn mediate MT-cortex interactions
(Hoepfner et al., 2000). In Schizosaccharomyces pombe, interphase cells contain several cytoplasmic
MTs, which span the entire cell. Because they serve as tracks to deliver so called cell-end markers, these
MTs determine growth directionality in this yeast (Tran et al., 2001).

In filamentous fungi, GFP-tagged MTs were first studied in A. nidulans in X. Xiang’s laboratory
(Bethesda, USA). MTs are quite inflexible structures and their orientation probably mainly depends on
the shape of the cell. Hence, the bundles of MTs are mostly aligned parallel to the growth axis and their
number ranges from 3 to 6. A. nidulans MTs extend with a speed of about 14 µm per min, reach the
cortex, pause for some time and undergo a catastrophic event. Subsequently, MTs shrink with a speed of
about 30 µm per min and they may either depolymerize all the way to the MTOC, or else rescue occurs
and they may recommence elongation (Han et al., 2001). Slightly different values were recently obtained
in the group of B. Heath (Sampson and Heath, 2005). They also observed that short MT fragments
were able to slide toward the hyphal tip. In Neurospora crassa, the MT network was first visualized by
N. Reads group in Edinburgh (Scotland) and has been analyzed recently in more detail (Freitag et al.,
2004; Mouriño-Pérez et al., 2006). From observations of the MT cytoskeleton, it is obvious that the orga-
nization is quite different in these two filamentous fungi. In N. crassa the MT cytoskeleton is far more
complex than in A. nidulans. MTs, which span the entire cell. Because they serve as tracks to deliver so
called cell-end markers, these MTs determine growth directionality in this yeast (Tran et al., 2001).

In filamentous fungi, GFP-tagged MTs were first studied in A. nidulans in X. Xiang’s laboratory
(Bethesda, USA). MTs are quite inflexible structures and their orientation probably mainly depends on
the shape of the cell. Hence, the bundles of MTs are mostly aligned parallel to the growth axis and their
number ranges from 3 to 6. A. nidulans MTs extend with a speed of about 14 µm per min, reach the
cortex, pause for some time and undergo a catastrophic event. Subsequently, MTs shrink with a speed of
about 30 µm per min and they may either depolymerize all the way to the MTOC, or else rescue occurs
and they may recommence elongation (Han et al., 2001). Slightly different values were recently obtained
in the group of B. Heath (Sampson and Heath, 2005). They also observed that short MT fragments
were able to slide toward the hyphal tip. In Neurospora crassa, the MT network was first visualized by
N. Reads group in Edinburgh (Scotland) and has been analyzed recently in more detail (Freitag et al.,
2004; Mouriño-Pérez et al., 2006). From observations of the MT cytoskeleton, it is obvious that the orga-
nization is quite different in these two filamentous fungi. In N. crassa the MT cytoskeleton is far more
complex than in A. nidulans. MTs, which span the entire cell. Because they serve as tracks to deliver so
called cell-end markers, these MTs determine growth directionality in this yeast (Tran et al., 2001).
Real-time studies of MT organization and dynamics by immunofluorescence are impossible, but fluorescently labeled tubulin has recently allowed observations in living cells (Czymmek et al., 2005; Ding et al., 1998; Fischer and Timberlake, 1995; Freitag et al., 2004; Han et al., 2001). The filamentous structures observed using both immunostaining and GFP-labeled tubulin, consist of several individual MTs with mixed orientation. There is increasing evidence for this organization coming from studies with S. pombe where it was recently shown that the orientation of neighboring MTs can be opposite within the one bundle. Moreover, a kinesin-like motor protein in combination with dynein is required for sliding of individual MTs within a bundle and for maintenance of MT polarity (Carazo-Salas et al., 2005). This suggests a mechanism whereby MT bundles can quickly increase or decrease in length. In A. nidulans, Konzack et al. reported that fluorescence intensity of a MT varies dynamically and that the regions with low intensity can recover brightness after some time. Similarly, after localized bleaching of a given MT, brightness returns quickly (Veith et al., 2005), indicating active turnover of tubulin subunits within the bundle. In addition, thin MT filaments occasionally detach from a MT for some time before they merge again to form a thick MT (Veith and Fischer, unpublished results). These observations are in agreement with a model that MT filaments consist of a bundle and individual MTs within a bundle undergo individual behavior and dynamics.

15.3 Origin of Microtubules

Microtubules cannot efficiently assemble de novo in a eukaryotic cell and require MT organizing centers (MTOC), of which γ-tubulin is a characteristic and necessary component (see earlier). In higher eukaryotes γ-tubulin forms a 2.2 MDa ring complex, the γTuRC, consisting of 12–14 (different numbers exist in the literature) γ-tubulin subunits associated with other proteins (Aldaz et al., 2005). It has been known for a long time that fungal spindle-pole bodies (SPB) are very active MTOCs (Jaspersen and Winey, 2004). The SPB is embedded into the nuclear envelope, divides prior to mitosis and, by definition, localizes at the poles of the mitotic spindle. SPBs consist, in S. cerevisiae, of an inner and an outer plaque and they are able to polymerize MTs on both sides of the nuclear envelope. The outer MTs formed during mitosis are called astral MTs, but the interphase SPBs are also active MTOCs both in S. cerevisiae and in filamentous fungi (Heath, 1981). The protein composition of the S. cerevisiae SPB has been defined by John Kilmartin’s lab (Adams and Kilmartin, 1999) using Mass Spec-based identification of peptide fragments from highly purified SPBs. The availability of complete annotated genome sequences from several species, combined with the increasing sensitivity of MS peptide identification, now makes it possible to undertake similar experiments in filamentous fungi.

It seems that the SPBs are the only places from which the yeast S. cerevisiae polymerizes MTs (see accompanying movies in Hoeftner et al., 2000). However, cytoplasmic MTs only have a minor and nonessential role in S. cerevisiae, that of positioning of the nucleus prior to mitosis (Maekawa and Schiebel, 2004). The cytoplasmic MT array is not very pronounced and usually limited to a few MTs growing out of the SPB into the cytoplasm. In contrast, filamentous fungi employ MTs for their fast, polarized growth during interphase (Horio and Oakley, 2005; Riquelme et al., 2003). Nevertheless, it was assumed for a long time that SPBs are the only place for MT initiation (Czymmek et al., 2005; Oakley, 2004; Sampson and Heath, 2005). This assumption was based on the finding that the intracellular αβ-tubulin pool is used for the assembly of spindle MTs as well as for cytoplasmic MTs. Indeed, cytoplasmic MTs are generally disassembled prior to mitosis and regenerate thereafter (Ovechkina et al., 2003; Sampson and Heath, 2005). In order to determine the origin of new MTs, regrowth of MTs was observed in S. pombe after depolymerization of MTs by drugs (Mata and Nurse, 1997). These studies revealed that MTs are generated not only from the SPB but also from other MTOCs around the nucleus and in the cytoplasm. During cell division an equatorial MTOC becomes very important (EMTOC) (Hagan, 1998; Sawin et al., 2004; Venkatram et al., 2005). The origin of MTs from the cell centre leads to an orientation with their plus ends toward the growing ends.

Recently, another tool was used to determine the origin of MTs. Using MT plus-end localizing proteins, such as homologs of the mammalian EB1, MT initiation was analyzed in the plant pathogenic basidiomycete Ustilago maydis. It was found that MT nucleation occurs at three places, at dispersed
cytoplasmic sites, at a polar MTOC and at the SPB (Straube et al., 2003). Whether MTOCs exist near the apical dome of other tip growing cells is a matter of debate as MTOCs have been observed within the apical dome of plant cells such as moss protonemata (Doonan et al., 1985).

In filamentous fungi, our knowledge of MT organization is restricted to a few species, such as the chytridiomycete *Allomyces macrogynus*, the basidiomycete *U. maydis*, and the ascomycete, *A. nidulans*, which is one of the best-studied examples. Whereas Sampson and Heath (2005) reported that MTs emanate only from SPBs, Konzack et al. (2005) demonstrated that additional MTOCs also exist. This discrepancy may be due to different methods used. In the first study, the authors observed that GFP labeled MTs and the location of nuclei was determined by the absence of cytoplasmic fluorescence. The authors of the second study used simultaneous labeling of nuclei with a red fluorescent protein and GFP labeled tubulin. In addition, a plus-end tracking protein, KipA, was used to determine the origin of MTs. MTOCs were found at the SPBs but also in the cytoplasm and at septa of *A. nidulans* (Fig. 15.1). This model recently received further support from the characterization of a novel MTOC-associated protein,

![FIGURE 15.1 MTOCs in A. nidulans.](image)

(a) Hypha with DAPI-stained nuclei and GFP-labeled spindle-pole body (SPB) associated ApsB. Nuclei are evenly spaced and at each nucleus a SPB is visible. (b) Scheme of an MTOC with γ-tubulin and other proteins described in *S. cerevisiae*. (Adapted from Oakley, *Trends Cell Biol*, 10, 2000; Pereira and Schiebel, *J Cell Sci*, 110, 1997.) (c, d) MTOCs visualized by GFP-ApsB fusion, at septa. Left, phase contrast; right, same hypha under fluorescent conditions. Inset in (c), enlargement of the septum and overlay of phase contrast and fluorescent image. (e) MTOCs are found at the nuclei, in the cytoplasm and at septa. (From Konzack et al., *Mol Biol Cell*, 16, 2005. With permission.)
ApsB (Veith et al., 2005). Here, the authors demonstrated that MTOCs at septa are important for the production of the interphase cytoplasmic MT array (Fig. 15.1). These findings are in agreement with the results obtained in S. pombe and U. maydis.

It is still an open question whether there are MTOCs at hyphal tips of filamentous fungi. Whereas γ-tubulin can be visualized at tips of A. macrogynus hyphae and thus MTs polymerize from the tip to the back (McDaniel and Roberson, 1998), γ-tubulin has not yet been detected at the tip in A. nidulans, but using the kinesin motor KipA, Konzack et al. (2005) found MTs can also polymerize from the tip. However, it has to be considered that a MT occasionally might not depolymerize upon contact with the cortex but could bend along the cortex toward the rear of the hypha. If this MT would continue growth, it could explain the observed comets from the tip to the back of the hypha. In N. crassa the situation appears to be far more complicated because of the higher number of MTs and nuclei (Freitag et al., 2004; Mourino-Pérez et al., 2006) and detailed studies of MT origin have yet to be performed.

15.3.1 Microtubule Plus End

It is well accepted that the plus end consists of a large protein complex that is involved in the regulation of MT dynamics as well as in the regulation of interactions with cortical actin, membrane proteins, or proteins associated with the kinetochore of chromosomes (Akhmanova and Hoogenraad, 2005; Hestermann et al., 2002; Schuyler and Pellman, 2001b). Given the diversity of interacting partners, it is obvious that the protein composition of the plus end complex may vary depending on the function of the MT and is likely to be a highly controlled and organized structure. There are three different ways that proteins can reach the MT plus end and remain associated with it while the MT is growing (Al-Bassam et al., 2006; Howard and Hyman, 2003).

In fungi, the MT plus ends have been best characterized in S. cerevisiae and S. pombe. MT-cortex interactions play important roles for the positioning of the mitotic spindle and nuclear migration in S. cerevisiae (Schuyler and Pellman, 2001a). Dynein is a prominent example of a MT plus-end associated protein (Fig. 15.2) that localizes to the MT tip and hitchhikes with the growing filament to the cell periphery. Once at the cortex, dynein is activated and pulls the attached MT toward the cortex. This leads to translocation of the nucleus (Maekawa et al., 2003; Maekawa and Schiebel, 2004; Schuyler and Pellman, 2001a; Sheeman et al., 2003). The kinesin motor protein, Kip2, appears to be responsible for the plus end localization of several proteins, for example, the CLIP170-like protein Bik1 (Carvalho et al., 2004). As in S. cerevisiae, the CLIP170-like protein of S. pombe, Tip1, also localized to MT plus ends. The motor responsible for this localization is Tea2 (Busch et al., 2004). However, MTs are not so important for polarized growth in yeasts in comparison to filamentous fungi. However, only some components that localize at MT plus ends have been found in filamentous fungi, but amongst these are subunits of the dynein motor complex and recently the Stu2 (Alp14)-homolog AlpA (Enke et al., 2007; Zhang et al., 2002). Interestingly, conventional kinesin, KinA, is required for dynein MT tip localization (Fig. 15.2) (Zhang et al., 2003). The CLIP170-like protein, ClipA, in A. nidulans also accumulated at MT plus ends and its localization is also dependent on the Tea2/Kip2 homolog KipA (Efimov et al., 2006).

The role that plus-end localized proteins play for polarized growth remains an open question. As mentioned earlier, MT-cortical interactions are necessary for dynein-dependent nuclear positioning prior to mitosis in S. cerevisiae (Carminati and Stearns, 1997). In A. nidulans dynein is also required for nuclear positioning and migration and recently Veith et al. showed that the interaction of MT-plus ends with the cortex contribute to the dynamics of mitotic spindles (Veith et al., 2005; Xiang et al., 1994; Xiang and Fischer, 2004). Whether interphase nuclei are moved as a result of similar MT-cortex interactions is not yet clear.

Whereas the MT-plus end protein complex is widely accepted as having a role in force generation required to translocate organelles, a role in polarized growth is less obvious. Some new ideas came from observations of MTs in growing tips of A. nidulans. Konzack et al. (2005) described how MTs merge into one point in the apex. Given that vesicles constantly travel toward the vesicle supply centre, the position of MT ends determines the vesicle supply centre location. In the kipA (tea2/kip2) mutant where MTs did
not merge into a single point, the hyphae grew in meandering curves rather than straight lines. This was explained by the lack of cell-end markers, which mediate cortical contact, and are normally transported by KipA (see later). There is good evidence for such a situation in *S. pombe*, where it was shown that the cortex protein Tea1 is transported by Tea2 (Fig. 15.3) (Browning et al., 2003; Martin and Chang, 2003; Sawin and Snaith, 2004). If either of the two genes is deleted, *S. pombe* cells appear curved or T-shaped (Browning et al., 2000; Snell and Nurse, 1994). Hence, Tea1 and other proteins were named cell polarity determinants or cell end marker proteins. However, to prove such a model in *A. nidulans*, cargoes of KipA have to be identified and characterized. Another crucial piece in the puzzle is the identification of cortex proteins. Whereas cortical contacts of MTs involved in nuclear migration require the cortical protein, ApsA, in *A. nidulans* (Num1 in *S. cerevisiae*) (Veith et al., 2005) this interaction appears not to be necessary for polarized growth (unpublished results). In *S. pombe*, the Mod5 protein acts as a membrane anchor for the polarized growth machinery (Snaith and Sawin, 2003). However, in filamentous fungi, a protein with significant sequence similarity has not yet been identified.
15.3.2 MT Lattice

MT function and dynamics are not only determined by the plus and minus ends, but also by the filament lattice, which in higher eukaryotes can be decorated with a number of different microtubule-associated proteins (MAPs), which in turn may control the activity of associated motor proteins (Baas et al., 1994; Baas and Qiang, 2005; Cassimeris and Spittle, 2001). Despite the abundance of those proteins in higher eukaryotes, it is not clear yet whether proteins like Tau exist in filamentous fungi. A very profitable approach involves the direct isolation of MAPs, based on their ability to bind to and copurify with MTs. Many of the classical mammalian MAPs were found due to their copurification with brain tubulin, which can be easily induced to assemble into MTs in vitro. Tubulin from other species does not so readily self-assemble but taxol can promote in vitro MT assembly, and has extended this approach to plants where MAPs are not readily identified due to structural divergence (Korolev et al., 2005). However, filamentous fungi are quite closely related to yeast and many other MAPs such as katanin and spastin that lead to MT severing can be recognized from conserved sequences in the A. nidulans and A. fumigatus genomes (Konzack, unpublished results). Experimental data for the role of this class of MT-associated proteins are not yet available for filamentous fungi.

15.3.3 MT-Dependent Motor Proteins

Microtubules and their dynamics are, in principle, able to create force and transport proteins, such as Tea1 in S. pombe (see earlier). However, at least two classes of motor proteins have evolved that mediate fast MT-dependent movement within the cell. These are the minus-end directed dynein and the plus-end directed kinesins (Fig. 15.2), although some kinesins can also move in a minus-end mode. Both motor classes are characterized by a motor domain in which ATP is hydrolyzed (Hirokawa, 1998). The location of the motor domain within the protein can be N- or C-terminal or even the middle region.

The mechanism by which chemical energy is converted into conformational changes and force generation is best understood in conventional kinesin. Interested readers should refer to several recent reviews (Adio et al., 2006; Schliwa and Woehlke, 2003; Woehlke and Schliwa, 2000; Yildiz and Selvin, 2005).

Whereas all fungi employ a single dynein for their transport processes, their genomes usually contain several kinesin-encoding genes. For instance, A. nidulans harbors 11 and N. crassa 10, different kinesins (Fuchs and Westermann, 2005; Rischitor et al., 2004). BimC was the first kinesin discovered in A. nidulans and defines the entire class of BimC-like kinesins (Enos and Morris, 1990). The gene was discovered in a screen for temperature-sensitive A. nidulans mutants with defects in mitosis (bim = block in mitosis).
BimC has a C-terminal motor, which forms a tetramer with two motor domains opposite to each other. Because every head domain can bind to a MT, this arrangement allows crosslinking of adjacent MTs and BimC provides the motive force to build the spindle. One of the BimC’s essential functions is to separate the SPBs. Shortly after their activation as intranuclear MTOCs, BimC is required to move the SPBs to opposite sides of the nucleus as this step fails in a temperature-sensitive mutant, bimC3 (Enos and Morris, 1990). BimC is not required for spindle assembly per se as MT elongation is not impaired in bimC3 and a “unipolar” spindle is formed where MTs running from the spindle pole bodies do not overlap. BimC defects seem to lead to a mild checkpoint induction since bimC3 mutants only arrest transiently in mitosis with a temporally elevated mitotic index. The mutant cells do exit mitosis after some time, suggesting that the checkpoint becomes attenuated. Whereas BimC was discovered in a genetic screen (Morris, 1976), four other kinesins were characterized using reverse genetic approaches (Konzack et al., 2005; Prigozhina et al., 2001; Requena et al., 2001; Rischitor et al., 2004).

A second motor with functions in mitosis is the C-terminal kinesin-like protein KlpA with similarity to S. cerevisiae Kar3 (Prigozhina et al., 2001). The gene was isolated through a PCR approach and characterized subsequently. Deletion of klpA alone did not produce any severe phenotype but suppressed a bimC mutation (O’Connell et al., 1993) suggesting that these two motors act in opposing directions during the establishment of the spindle.

Another kinesin with a function in mitosis is the Kip3 family member, KipB, with a motor domain that is localized closer to the N-terminal. Gene deletion did not cause any defect in hyphal extension or organelle movement, but chromosome segregation was defective (Rischitor et al., 2004). This was surprising, because a similar motor in S. cerevisiae, Kip3, is involved in nuclear migration (Miller et al., 1998). However, the A. nidulans KipB results are in good agreement with results for the homologous proteins in S. pombe, Klp5 and Klp6 (West et al., 2002).

Two motors with N-terminal motor domains and pronounced roles in polarized growth are conventional kinesins, KinA, and the CENP-E family kinesin KipA. Deletion of kinA resulted in slower hyphal growth, which is similar to effects in other fungi (Fig. 15.2) (Lehmler et al., 1997; Requena et al., 2001; Seiler et al., 1997; Wu et al., 1998b). It is generally accepted that this motor transports vesicles toward the extending tip and provides cell wall components to the growing tip (Seiler et al., 1999). In addition, KinA appears to be involved in other cellular processes related to polarized growth, namely mitochondrial and nuclear distribution. Whereas nuclear distribution was affected in N. crassa and A. nidulans, mitochondrial distribution was changed in Nectria haematococca (Wu et al., 1998b). This may be due to the fact that mitochondrial movement depends on the actin cytoskeleton in A. nidulans (Suelmann and Fischer, 2000) and on the MT cytoskeleton in N. crassa (Fuchs et al., 2002; Fuchs and Westermann, 2005). Whether mitochondrial distribution is also altered in N. crassa conventional kinesin mutants has not yet been studied. The mechanism by which conventional kinesin may contribute to mitochondrial or nuclear distribution is not yet clear, but the effects may be indirect. KinA is required for transportation of dynein subunits to the plus end of MTs (Zhang et al., 2003) and dynein is a crucial motor for nuclear migration. Exclusion of dynein from the MT plus ends in cells lacking KinA could cause the observed nuclear clustering (Xiang et al., 1994). In addition, it has to be considered that conventional kinesin may well be involved in delivering other components of the MT plus end complex. Lack of KinA could thus influence the dynamics of MTs as well as their cortical interaction.

KipA of A. nidulans is similar to Tea2 in S. pombe and is characterized by an N-terminal motor domain (Konzack et al., 2005). Accumulation at the plus ends is dependent on an intrinsic motor activity because mutant proteins, in which a crucial residue for ATP hydrolysis was replaced, lost the ability to accumulate at MT tips, but decorated them evenly. These findings were in agreement with studies of Tea2 in S. pombe (Browning et al., 2003). Gene deletion caused a surprising phenotype in A. nidulans (Fig. 15.2). Delta kipA strains grew as well as wild-type strains but hyphal morphology and MT behavior was changed. In wild type, MTs form a focus at the apex, but are dispersed in the mutant. This suggested that the MT foci were important for controlling the direction of growth. MT foci would direct and deliver the vesicles accurately to one place, and the Spitzenkörper and hyphae would grow straight. If the MTs could not merge into one point, vesicle delivery would be less accurate with the result that new growth occurs in arbitrarily directions and leads to a twisted morphology. The KipA protein might transport proteins that are necessary for temporal anchorage of MT at the cortex at a specific point. An example for such proteins
in fission yeast are Tea1 and Tip1 (Browning et al., 2003; Busch et al., 2004). However, MT fixation at the cortex through Tea1 was not shown. Tea1 may be evolutionarily conserved among fungi, because a similar protein has been localized to the growing hyphal tip in *A. nidulans* (Takeshita, Konzack and Fischer, unpublished results).

Deletion of any kinesin motor (besides *bimC*) does not cause severe phenotypes. Interestingly, even a strain in which KinA, KipA, and KipB were deleted, was still viable although hyphal growth and development were quite severely affected (Konzack et al., 2005). This shows that kinesins can substitute for each other to some extent, and this was recently confirmed in the case of the Unc-104 homologs, Nkin-2 and Nkin-3, from *N. crassa*. Whereas Nkin-2 associates with mitochondria and connects mitochondria with MTs, Nkin-3 was found in the cytoplasm. Surprisingly, after depletion of Nkin-2, Nkin-3 was upregulated and also bound to mitochondria and MTs (Fuchs and Westermann, 2005). Homologs of these two motors also exist in *A. nidulans* and are currently being investigated (Zekert and Fischer, unpublished results).

UncA plays an important role in hyphal tip extension, whereas UncB is likely to play a role in the nucleus and at septa (N. Zekert and R. Fischer, unpublished data).

As mentioned earlier, fungi usually contain only a single dynein protein, although in some basidiomycetes the heavy chain is encoded by two genes (Eshel et al., 1993; Martin et al., 2004; Straube et al., 2001; Xiang et al., 1994; Yamamoto and Hiraoka, 2003). Dynein has a crucial role in nuclear migration but is also implicated in vesicle transport (Seiler et al., 1999). Because dynein moves toward the MT minus end, it is difficult to imagine that it is directly involved in polarized growth, given that MTs are mainly oriented with their plus-ends to the membrane. Indeed, deletion of dynein does not cause an immediate block of hyphal extension and the impact on colony growth could partly be due to the lack of nuclei and other organelles that are translocated with the help of dynein (Xiang et al., 1994).

Besides the concerted action of the cytoskeleton and associated motor proteins to translocate organelles, cytoplasmic streaming has to be considered as another mechanism to push forward the cytoplasm and organelles. Mouriño et al. showed recently in *N. crassa* that the MT array was able to advance as a unit, as the hypha elongates. The basis for this bulk flow has not yet been defined (Mouriño-Pérez et al., 2006).

If MTs play a role in vesicle delivery to the growing hyphal tip, the question remains as to how the sites for cell extension are marked. First insights into this process came from studies in *S. pombe* with the definition of cell-end markers.

### 15.3.4 Cell-End Markers at the Cortex

As described earlier, one of the first proteins (Tea1) to label the growing end of a yeast cell was discovered in *S. pombe* during a genetic screen for polarity mutants (Mata and Nurse, 1997). It was shown recently that the main membrane anchor, which recruits proteins such as Tea1, is Mod5 (Browning et al., 2003; Snaith and Sawin, 2003). This protein is posttranslationally modified with a prenyl residue, conferring membrane association. The anchored Mod5 then recruits the formin protein, For3 (Fig. 15.3) (Bretscher, 2005; Martin et al., 2005; Martin and Chang, 2006). For3 initiates the growth of actin filaments away from the growing tip that can be used as tracks for directing the vesicles necessary for cell extension. Given that the growth machinery is largely conserved in filamentous fungi, and although a crucial component, Mod5, has not yet been identified in filamentous fungi, and it has to be considered that a similar protein exists, the question remains as to what targets Mod5 or analogous proteins to the membrane near the hyphal tip rather than along the length of the cell. This points to a key function of the membrane itself, perhaps involving sterol-rich lipid rafts that may cause asymmetric distribution of membrane-associated proteins (Grossmann et al., 2006; Hancock, 2006). There is recent evidence that these membrane domains play a role in polarized growth of filamentous fungi (Martin and Konopka, 2004) and the laboratory of S. Harris showed that a ceramide synthase is important for hyphal morphogenesis (Li et al., 2006).

Because the installation of the growth machinery at a specific place determines growth directionality, one would expect that external signals influence the architecture of proteins. Indeed, recently a kinase with such a potential was described in *A. nidulans* (Li et al., 2006). The ATM kinase has a well-characterized role in DNA damage response (see Section 15.4.7) but Li et al. (2006) found that deletion also affects the establishment of polarized growth. The reason appeared to lie in a disorganization of MTs.
in the apex similar to the defect in the kinesin mutant ΔkipA (Konzack et al., 2005). Whereas MTs form a focus within the apical dome of wild-type hyphae, they are dispersed in the atmA and the kipA mutants. In both cases the authors argue that MT-cortex interaction might be affected, but it is unclear if this is due to a direct effect of ATM kinase at the tip or an indirect effect through the DNA damage checkpoint pathway.

Two further candidates for regulation of tip growth are Pod6 and Cot1, described from N. crassa. Both proteins are distributed evenly along the hypha and their role in tip growth remains to be explained (Seiler et al., 2006). The BimG phosphatase, better known for its role in mitosis, is also found in the apical dome (see Section 15.4.5) where it might control vesicle recycling.

### 15.4 Cell-Cycle Controls

Cell growth and division require careful coordination of many processes to ensure that two healthy and viable daughter cells are formed. Visually, this is most dramatically illustrated by mitosis when many subcellular structures undergo dramatic and extensive reorganization (Fig. 15.4): cytoplasmic MTs begin to disassemble and disappear, being replaced by the mitotic spindle as the spindle pole is activated. Nuclear structure changes too, with the chromosomes becoming condensed in preparation for separation. In some groups, but not in the filamentous fungi, the nuclear envelope breaks down and the nucleolus also disperses coincident with cession of ribosome biosynthesis and reduced protein production. The consequences of mitosis are profound, leading to complete and irreversible separation of the genome into two daughter nuclei. Premature entry into mitosis is prevented by a system of “checkpoints” that ensures that the prospective mother cell is capable of producing two viable daughters—that DNA replication is complete, and the cell is big enough, to name but two important attributes that must be satisfied. Therefore, the decision to enter mitosis is one of the critical transitions in the cell cycle and is under strict control by a network of regulatory pathways known as checkpoints, because, if taken at the wrong time, this would be effective suicide. These checkpoint pathways link mitosis with other cell cycle events, such as septation,

![FIGURE 15.4 Hyphal germination and nuclear division cycle in A. nidulans. Note: Scanning electron micrographs of germinating spores are arranged above the corresponding cell-cycle phases (G1, DNA replication, G2, and mitosis) and a schematic of the nuclear division cycle showing MTs (gray lines) SPB (small empty rectangles), and DNA (gray shading where the state of DNA condensation is indicated by the intensity of gray). The lower images show transmission electronmicrographs of interphase and mitotic nuclei (from Kerry O’Donnell, J Cell Sci, 99, 1991; modified from Doonan, J Cell Sci, 103, 1992) and various features are marked as follows: spb, spindle-pole body; ne, nuclear envelope; nu, nucleolus; ch, chromosome, arrowheads indicate bundles of spindle microtubules.](image)
Cytoskeleton, Polarized Growth, and the Cell Cycle in Aspergillus nidulans

DNA replication, and cell growth. Checkpoint pathways have been genetically dissected in several model organisms, including *A. nidulans*. Some of the important features of this network in *A. nidulans* are shown in Figure 15.5.

15.4.1 Genetic Analysis of the Nuclear Division Cycle

Chromatin condensation during mitosis is particularly marked in *A. nidulans* hyphae and this provided the basis for a cytological screen for cell cycle mutations. These screens directly identified mutations with altered cell-cycle dynamics and have implicated at least thirty genes in various aspects of DNA replication and nuclear division. These include temperature-sensitive lethal mutant collections (Morris, 1976; Orr and Rosenberger, 1976a,b), aneuploid generating mutants (Upshall and Mortimore, 1984), and mutations sensitive to DNA damage (Käfer, 1986; Oza and Käfer, 1990; Shanfield and Käfer, 1969). Although not all of the corresponding genes have been isolated, those that have been cloned provided profound insights into the general mechanisms that regulate mitosis in eukaryotes.

Two of the key regulators, NimA kinase and BimEAPC/C were identified using mutants from Ron Morris’ Lab. Morris (1976) screened for mutants that were conditionally defective in nuclear division but could continue limited growth at restrictive temperature. Such hyphae contained fewer nuclei per unit length, suggesting that nuclear division was specifically impaired. The mutants were further classified on the basis of chromatin configuration (condensed and, therefore, mitotic, or noncondensed and, therefore, interphase) and presence or absence of spindles at restrictive temperature. Those mutants with interphase-like nuclei were given the acronym *nim*, never in mitosis; those with condensed nuclei were called *bim*.

![FIGURE 15.5 Regulatory network controlling the mitosis in *A. nidulans*. Note: Arrows represent positive interactions or regulation; bars indicate negative regulation. Solid lines indicate genetic or protein-protein interaction evidence for the regulation. The nuclear schematic is as described in Figure 15.4 while cell cycle events are indicated inside the rectangles.](image-url)
for blocked in mitosis. To establish where the interphase mutations had their point of action relative to S-phase, the *nim* mutants were tested using a reciprocal shift method where the arrest point of a mutant is determined relative to the reversible S-phase arrest induced by hydroxyurea (HU), an inhibitor of DNA synthesis (Bergen et al., 1984). Surprisingly, no strains were identified as blocking in G1 although several mutations produced nuclei that could not be scored due to abnormal nuclei. In addition, three mutations (in the *nimL*, *nimM*, and *nimN* genes) were irreversible conditional lethals whose position of blockage could not, therefore, be determined. However, the latter three are all supersensitive to low doses of HU, implying that they may be involved in DNA metabolism (Doonan, unpublished data). Five other mutations, *sodB1*, *nimC3*, *nimG10*, *nimK14*, and *nimQ20*, conditionally block in S-phase. Mutations in at least six genes conditionally prevent the transition from G2 to M. These include *nimA*, *nimB*, *nimE*, *nimT*, *HfaB*, and *HfaF* and, together with the genes required for mitotic progression, these genes have provided some unique insights into eukaryotic mitotic control.

### 15.4.2 Regulation of the G2/M Transition

A phosphorylation cascade, cumulating in the activation of the NimX cyclin-dependent protein kinase (cdk), plays a key role in regulating mitotic entry. This heterodimeric protein kinase is composed of two subunits, a catalytic kinase subunit encoded by cdc2 gene in *S. pombe* (the *nimX* gene in *A. nidulans*) and a regulatory cyclin subunit encoded by cdc13 in *S. pombe* (Moreno et al., 1989) (*nimE* in *A. nidulans*). The kinase is kept in an inactive state by phosphorylation on tyr15 (Fleig and Gould, 1991). Activation as a mitosis-specific histone H1 kinase occurs by dephosphorylation, which is undertaken by a tyrosine-specific protein phosphatase (reviewed by Fleig and Gould, 1991), encoded by NimT in *A. nidulans* (Osmani et al., 1991b). In all of these features, the NimX*cdc2* kinase conforms to the typical eukaryotic mitotic cdk kinases. Mutations in the genes that encode the regulatory subunits of this kinase conditionally block the cell cycle in G2 (O’Connell et al., 1992). Mutant *nimT23* strains, therefore, arrest in G2 with phosphorylated and inactive p34cdc2 kinase.

Another distinct type of mitotic kinase, NimA, was discovered in *A. nidulans* and its analysis has revealed a novel parallel control pathway that acts both alongside and on the p34 kinase as a positive regulator of mitosis (Osmani et al., 1991b). At restrictive temperature, mutations in the *nimA* gene block in G2 (Bergen et al., 1984) and return of cells to the permissive temperature results in a synchronous entry into mitosis within a few minutes (Oakley and Morris, 1983). Activation of NimA and NimX*CDC2* kinases are, in part, independent of each other: NimA is fully active when NimX*CDC2* is inactivated by mutations in the *nimT* gene and NimX*CDC1* activity is very high in *nimA5* mutations (Osmani et al., 1991a). Expression of NimA in *A. nidulans* and other cell types can induce aspects of mitotic progression, especially chromatin condensation, and this does not require Cdc2 kinase activity (Lu and Hunter, 1995; O’Connell et al., 1994).

The parallel behavior of these two kinases seemed to be confirmed when NimX p34cdc2 kinase activity was examined in a *nimA5* mutant background. Despite the *nimA5* mutant being blocked in interphase at restrictive temperature, it had a fully activated p34 H1 kinase (Osmani et al., 1991b). Moreover, the NimX kinase was dephosphorylated on tyrosine 14, one of the final steps in its activation. Therefore, an active p34 kinase is insufficient to allow *A. nidulans* cells to enter mitosis. NimA kinase is fully active as a kinase in the *nimT23* mutant, which fails to dephosphorylate and activate NimX*CDC2*.

The underlying mechanism by which mitotic entry is dependent on both NimA and NimX*CDC2* kinases involves the NimA-dependent nuclear import of NimEcyclinB (Wu et al., 1998a), and possibly other cell-cycle regulators (De Souza et al., 2000). Although NimX is dephosphorylated in *nimA5* mutants, NimEcyclinB and NimX*CDC2* are retained in the cytoplasm. NimA, therefore, acts upstream of NimX*CDC2* by controlling the subcellular localization of the *nimE* encoded cyclin regulatory subunit and the cyclin-dependent kinase, NimX*CDC2*.

To control the timed nuclear uptake of NimEcyclinB NimX*CDC2*, NimA could either module nuclear pore function in a cell-cycle dependent manner or it could directly modify the CDK, but the available evidence supports the former mechanism. NimA interacts genetically with *sonA*, a homolog of the yeast nucleocytoplasmic transporter GLE2/RAE1, and interacts with SonB, a FG-repeat nucleoporin (De Souza et al., 2003). The *sonA1* mutation suppresses defective nuclear division and NimEcyclinB localization in *nimA1* cells.
without markedly increasing NimX\textsubscript{CDC2} or NimA activity, as measured in cell homogenates. NimA activation leads to partial disassembly of nuclear pores not only at mitosis, but also when ectopically expressed in S-phase cells. These results indicate that NimA promotes the nuclear localization of the NimX\textsubscript{CDC2}/NimEcyclinB complex, by modulating nuclear pore stability. The dual action of the two kinases may have evolved as a mechanism to facilitate the entry of cytoplasmic proteins into the nucleus in the absence of mitotic nuclear envelope disassembly (De Souza et al., 2004). In organisms where the nuclear envelope does break down at mitosis, the other roles of NimA-related proteins may appear more prominent.

NimA also has additional roles and targets in \textit{A. nidulans}. A major effect of ectopic NimA expression is chromosome condensation, due to the phosphorylation of histones by NimA. Prior to mitosis, NimA is uniformly distributed within the cell but, as the cell enters mitosis, it accumulates in regions of the nucleus that are undergoing chromatin condensation precisely coinciding with the phosphorylation of histone H3 on serine-10 (De Souza et al., 2000). Histone H3 is a substrate of NimA \textit{in vitro} and its phosphorylation on serine-10 is crucial to condensation of the nucleosomes in many species. NimA is also localized to the spindle and spindle-pole bodies during later stages of mitosis, suggesting that other targets may be located on the spindle poles or microtubules. TinA, a protein that locates to the mitotic spindle poles and influences astral microtubules, interacts in yeast 2H with NimA, and deletion of \textit{tina} is synthetically lethal with bimE\textsubscript{APC2} mutations (Osmani et al., 2003). TinA might provide a structural connection between NimA and other cell-cycle regulators that show spindle localization during part or all of mitosis. These include components of the anaphase-promoting complex (Mirabito and Morris, 1993) and BimGPP1 (Fox et al., 2002), which also genetically interact with \textit{nima} to control entry into mitosis. In fission yeast, Fin1 (the NimA homolog) promotes the association of polo kinase with the spindle pole body (Grallert and Hagan, 2002), although this does not appear to be the case in \textit{A. nidulans} as POLO is found at the poles throughout the cell cycle (Bachewich et al., 2005). In animals, NimA-related kinases promote the assembly of the centrosome, an analogous structure to the spindle pole body. These data suggest an evolutionarily conserved role for NimA kinases in spindle-pole function (Hayward and Fry, 2006).

NimA activity is tightly controlled at a number of levels, leading to an increase in kinase activity during mitosis of approximately 20-fold of that observed in S-phase (Osmani et al., 1991a). First, accumulation of \textit{nima} transcript increases to a maximum at mitosis and falls precipitously as the cells return to interphase (Osmani et al., 1987). This is an important regulatory mechanism as inappropriate production of \textit{nima} mRNA (from an inducible promoter) can drive cells into premature mitotic-like state, even in the presence of a HU block (Osmani et al., 1988). Moreover, overexpression of NimA prolongs mitosis leading to the formation of elongated spindles and condensed fragmented chromatin. This suggested that NIMA must be destroyed in order for cells to exit mitosis.

The accumulation of high levels of mutant NimA5 protein in double-mutant \textit{nima5 bimE7} due to a compromised APC function might explain the rather odd mitotic phenotype of these strains. After a short delay, the double mutant strain enters mitosis with aberrant spindles and nuclear membranes whereas a single \textit{nima5} mutant does not (Osmani et al., 1988). NimA protein is relatively stable in cells that have been arrested in mitosis and a destruction motif probably resides in the C-terminus of the protein as deletion leads to protein accumulation and delayed mitotic exit (Pu and Osmani, 1995b). NimA levels during interphase also respond to checkpoint mechanisms that monitor DNA replication (Ye et al., 1996). If progression through S-phase is delayed, the checkpoints normally act to delay entry into mitosis and this involves delaying the activation of NimX\textsubscript{CDC2} and NimA kinases. Under such circumstances, NimX\textsubscript{CDC2} is kept inactive via phosphorylation of tyrosine-14. Indeed, cells unable to phosphorylate tyrosine-14 on NimX\textsubscript{CDC2}, either due to mutation of NimX (\textit{nima5} \textit{bimE}) or to deletion of the wee1 kinase that carries out the phosphorylation, are viable but sensitive to S-phase delay and will enter mitosis with partially replicated DNA, and die. This premature mitotic entry is dependent on NimA accumulation, so tyrosine phosphorylation of NimX\textsubscript{CDC2} plays a key role in controlling the timing of NimA protein accumulation, at least when DNA replication is defective. APC\textsuperscript{BimE} also plays an important role in controlling NimA activity during interphase—if DNA replication is completely blocked, double \textit{bimE7nima5} mutants enter mitosis even at very high levels of DNA synthesis inhibitors, and this is associated with accumulation of NimA. This is consistent with other data that suggests the APC is activated in S or
G2 to prevent premature activation of mitotic regulators (Lies et al., 1998; Ye et al., 1998). Another regulatory mechanism involves protein phosphorylation. NimA is a phosphoprotein (Lu et al., 1993) that becomes hyperphosphorylated during G2 (Ye et al., 1995). The final activation of hyperphosphorylated NimA depends on fully activated NimX^CDC2 and is important for the coordination of early mitotic events. Several of the temperature-sensitive nim alleles produce proteins that accumulate normally at the correct time, but fail to undergo this final NimX-dependent activation step (Pu and Osmani, 1995a). Phosphorylated NimA may be a substrate for Pin1, a peptidyl-prolyl cis/trans isomerases (PPI) that might affect its activity by altering the conformation of the protein. In human cells, a NimA protein kinase interacts with, and is negatively regulated by, Pin1 (Lu et al., 1996), which is required for normal progression through mitosis. In A. nidulans, PinA (a pin1-like protein) interacts genetically with nimA5 (Joseph et al., 2004); overexpression of pinA reduces the severity of the nimA phenotype while reduction in PinA levels increases the severity. Pin1 may be dependent on cdc2 phosphorylation as its preferred targets are proline residues proximal to phosphoS/phosphoT, the product of Cdc2 phosphorylation, and it has been suggested that it amplifies the effect of phosphorylation by inducing a really major change in protein shape on proteins that have been earmarked by the kinase. Consistent with the idea that PinA may have additional substrates, reducing PinA has diverse effects on the A. nidulans cell cycle (Joseph et al., 2004).

15.4.3 Involvement of Calcium in the G2/M Transition
Calcium has been widely implicated in cell-cycle transitions in plants and animals. The gene for calmodulin, one of the major internal cellular receptors for calcium, has been cloned and sequenced from A. nidulans (Rasmussen et al., 1990). Using site-specific gene replacement to place the calmodulin gene under the control of the inducible alcA promoter indicates that calmodulin is required for G2/M progression (Lu et al., 1992). A multifunction calcium calmodulin-dependent protein kinase (CaMK) may have an overlapping role as strains with reduced CaMK activity also seem to be impaired in G2/M progression (Dayton and Means, 1996). Two additional CaMKs have been characterized and play roles in spore germination and cell-cycle progression (Joseph and Means, 2000).

15.4.4 APC and the Metaphase-Anaphase Transition
The anaphase-promoting complex (APC), or cyclosome, is an ubiquitin ligase that assembles polyubiquitin chains on to its substrates to direct them for degradation by the 20S proteosome (Gutierrez and Ronai, 2006). The APC is required for mitotic progression, particularly for the metaphase to anaphase transition and for mitotic exit. Mutants that lack APC function tend to arrest in mitosis and in A. nidulans, produce the typical bim phenotype. Genetic and biochemical analyses of the APC in diverse organisms have revealed a huge complex involving at least a dozen subunits, and two of the defining proteins were discovered as bim mutations in A. nidulans. The bimE7 mutation causes cells to block in metaphase at restrictive temperature, and is also known as APC1 (Zacharie et al., 1996). Loss of bimE function overrides a variety of interphase blocks, driving cells into premature mitosis (Osmani et al., 1988), but this requires functional NimX^CDC2 (James et al., 1995). The bimA gene encodes the APC3 subunit, a TPR protein (O’Donnell et al., 1991). Mutations in bimA also arrest in mitosis and override interphase arrests in a similar way to bimE mutations, but the phenotype is generally much weaker (Mirabito and Morris, 1993; Ye and Osmani, 1997).

Work in yeast suggests that the only essential function of APC is the ubiquitylation of securin, the separase chaperone, and cyclins. This has lead to the suggestion that APC and CDK-cyclins have coevolved (Thornton and Toczyski, 2003). In A. nidulans, the metaphase arrest caused by loss of APC function can be overridden by mutations in at least three other genes that affect either DNA replication or chromosome structure. Reductional divisions occur in nimO18bimE7 (James et al., 1999) and nimQ20bimE7 (Ye and Osmani, 1997) double mutants. NimO is structurally similar to the regulatory subunit of the cdc7 kinase, Dbf4, which phosphorylates the MCM proteins of the prereplicative complex thereby facilitating DNA replication. The cell achieves precise control over DNA synthesis, ensuring that one complete round of genome replication occurs in each cell cycle, by the binding of the prereplicative complex to discrete regions of the chromosome known as origins of replication. NimQ is homologous to one of the MCM
proteins. Both nimO and nimQ are, therefore, required for DNA replication and mutants grow at restrictive temperature arrest with unreplicated DNA.

The other mutations that can bypass the bimE7 metaphase arrest lie in the nimU gene (Pitt et al., 2004), which encodes a Pot1-like protein. Pot1 proteins form part of the shelterin complex, which acts to protect telomeric DNA from degradation and recombination, in part by regulating telomerase activity (Price, 2006). The nimU23 mutation was originally identified as having a decreased index of interphase nuclei and was, therefore, classified as being a nim “never-in-mitosis” mutant (Morris, 1976). Reciprocal shift experiments with HU indicated that the arrest point for nimU24 was in G2 (Bergen et al., 1984) but the logic of these experiments depend on there being a discrete arrest point for a given mutation: mutants that are “leaky” or do not arrest cell cycle progression can be misclassified. Loss of nimU function reduces the mitotic index because cells progress through mitosis too fast. nimU mutants show a number of defects consistent with this, including premature spindle elongation and early mitotic exit (Pitt et al., 2004) and this produces elongated large aberrant nuclei that continue to cycle but spend a reduced time in mitosis. Double nimU24bimE7 mutants enter mitosis with nearly normal bimE7 dynamics and a significant percent of cells progress through into anaphase. This indicates that NimU is required for spindle checkpoint control at the metaphase-anaphase transition and this is independent of APC function. The spindle checkpoint can be triggered by perturbing microtubule stability, either with drugs or by mutation, and results in prolonged mitosis due to inhibition of mitotic exit (Oakley and Morris, 1981). However, mitotic exit remains APC-dependent as double mutants remain blocked in mitosis with two masses of chromatin. nimU mutants may also fail to activate PP1 at metaphase as these strains also have reduced protein phosphatase activity (unpublished results). It also remains to be determined if timing of sister chromatid separation in nimU mutants is dependent on the APC and this should now be testable since centromeres can now be marked with GFP (Yang et al., 2004). The mechanism by which a telomere component affects the mitotic checkpoints, therefore, remains unclear.

The BimEAPC complex also regulates the SPB’s activity as a cytoplasmic MTOC. Loss of bimE function leads to cells containing metaphase spindles with short or no astral MTs, indicating that APC function is required for the activation of the cytoplasmic face of the SPB as an active MTOC at the end of metaphase. This requirement for APC acts through the TinA protein (Osmani et al., 2003). TinA is a coiled-coil-containing protein that was isolated as interacting with NimA in a yeast-two-hybrid screen. TinA localizes to the SPB at G2/M in a NimA-dependent manner and may act to suppress MT nucleation from the cytoplasmic face of the SPB during early stages of mitosis.

15.4.5 Anaphase

When all the chromosomes have attached correctly, via their kinetochores, to the spindle, the APC is activated and the cell enters anaphase. APC activation leads to the degradation of securin, a protein that inhibits the protease, separase. When separase is released from inhibition it causes the release of sister chromatid cohesion by cleaving cohesin. In A. nidulans, separase function is encoded by bimB (May et al., 1992), and one of the four subunits of cohesin is encoded by the sudA3 gene (Holt and May, 1996). SudA was isolated as a cold sensitive suppressor of bimD mutants. BimD is structurally related to Spd76 of Sordia, Pds5p of budding yeast, and As3 from humans, which are required for sister chromatid cohesion, DNA damage response, and normal cell-cycle progression (van Heemst et al., 2001). In yeast, Pds5p and a cohesin subunit, Scc1, are mutually required for each other’s recruitment to the chromosomes during G1 and cleavage of cohesin releases Pds5 at the metaphase-anaphase transition. In A. nidulans, as in vertebrates, BimD dissociates from the chromatin in prophase rather than the metaphase-anaphase transition, which may reflect differences in condensed chromatin structure between yeast and other eukaryotes (van Heemst et al., 2001). BimD also affects the rate of cellcycle progression, as mutants seem to progress faster through the cell cycle and cellular morphogenesis (van Heemst et al., 2001) while overexpression blocks the cell in G1 (Denison et al., 1993). The human ortholog, As3, acts as a tumor suppressor and it has been suggested that BimD plays a second role in modulating cell-cycle progression rates under unfavorable conditions (van Heemst et al., 2001).

Protein dephosphorylation plays a key role in anaphase. During mitotic entry many proteins become phosphorylated and, both to progress through mitosis and return to interphase, one might suppose that
they need to be dephosphorylated. Using an antibody, MPM2, that recognizes phosphoproteins (Engle et al., 1988), a mutation was identified with temperature-sensitive arrest in anaphase. A mutation in the \textit{bimG} gene, which encodes a type1 protein phosphatase (PP1), led to the formation of large nuclei, which failed to complete anaphase (Doonan and Morris, 1989), and also caused reduced PP1 activity (Doonan et al., 1991). At the level of gene expression, the \textit{bimG}11 allele leads to a temperature-sensitive splicing event due to a mutation in an intron. An interfering truncated protein is produced (Hughes et al., 1996), which is recessive to the wild-type protein. Consistent with this, the \textit{bimG}11 allele is suppressed by cold sensitive mutations in the \textit{sugB} gene, which encodes a splicing factor (Assinder et al., unpublished). \textit{bimG}11 is complemented by the structurally similar mammalian phosphatase (Doonan et al., 1991) but a related PP1 from Arabidopsis is only able to support hyphal growth (Arundhati et al., 1995). Interestingly, the major difference between the mammalian and plant PP1 is a C-terminal region where a functionally important regulatory site for cdc2 phosphorylation has been identified in the yeast homolog, dis2 (Yamano et al., 1994). However, mutation of the presumptive phosphorylation site in \textit{bimG} did not seriously impair its function (Fox, unpublished). BimG-GFP fusions locate to several subcellular compartments (Fig. 15.6), including the cytoplasm, the spindle pole body, the nucleolus, the spindle, and the septation site (Fox et al., 2002), reflecting the multiple functions of BimG and the pleiotropic nature of the mutation. Notably, BimG-GFP association with the spindle pole is transiently reduced

\textbf{FIGURE 15.6}  BimG distribution during mitosis. \textit{Note:} Left panels show the distribution of BimG-GFP fusion protein in a \textit{nimA5} mutant. The top panel represents time zero (\textit{nimA5} block point) and subsequent panels show representative stages of mitosis. The right panels show schematic representations of nuclei showing BimG distribution in gray. Large gray oval, nucleolus; small circles, SPB; bar, telophase spindle with BimG; ring, incipient septum. The stages of the cell cycle are given as follows: G1/S, G2, P (prophase), M (midphase), A (anaphase/telophase), and S (septation).
during early mitosis, at a time when MPM-2 phosphoprotein staining is increased. Since MPM-2 staining intensity is dramatically enhanced by the *bimG11* mutation, it seems likely that BimG is largely responsible for limiting phosphorylation of nuclear structures in G2. BimG associates with the nucleolus until quite late in mitosis and time-lapse observations suggest that nucleolar BimG-GFP is segregated into the daughter nuclei during late anaphase/telophase by streaming along the spindle. BimG is necessary for nucleolar segregation as mutants have a highly MPM-2-stained nucleolus, whose persistence is associated with spindle defects.

As the daughter nuclei separate during telophase, at some point, the nuclear envelope surrounding the two daughters must break into two and provide each daughter with its own envelope. In higher eukaryotes, this occurs by complete disassembly of the nuclear envelope and reassembly but *A. nidulans* retains a nuclear envelope throughout mitosis. NimA inactivation may play a central role in daughter nucleus separation and envelope segregation (Davies et al., 2004). TinC preferentially interacts with inactive NimA, locates to membranous structures, and is required for the final stages of nuclear separation. Expression of an N-terminal deletion derivative of TinC leads to uncoupling of DNA separation and nuclear envelope separation. TinC belongs to a fungal-specific group of proteins that includes the HetC gene of *N. crassa*, a multiallelic gene that defines heterokaryon incompatibility classes in that fungus (Sarkar et al., 2002).

### 15.4.6 Mitotic Exit and the Septation Initiation Network

Cell division is completed with the insertion of a septum (cytokinesis or septation) between two recently divided nuclei (see review by Harris, 2001). In germinating conidia, the initial cell delayed until after the third mitosis (Fiddy and Trinci, 1976; Momany and Taylor, 2000), so that the typical hyphal segment is multinucleate. Septation is not absolutely essential for the hyphal stage of growth as mutants that conditionally lack septa can produce microcolonies, but these fail to undergo any sexual or asexual development (Morris, 1976).

Septation involves the specialization of a site on the cell periphery during mitosis and the subsequent recruitment of actin filaments to form a contractile ring, similar to that observed at the division site in animals. The ring of actin filaments then contracts toward the centre of the cell, presumably pulling the cell membrane along. New cell-wall material is secreted to produce the septum. A number of mutants that have conditionally defective septation have been isolated (Morris, 1976). These have been classified into two distinct groups, called early and late. Late mutants, including *sepA, D, G*, and *H* undergo continuous nuclear division and hyphal growth at restrictive temperature but fail to form septa, while early mutants, such as *sepB, E, I*, and *J*, undergo only three nuclear divisions before arresting as aseptate cells. The late mutants probably represent genes involved in the process of septation itself, while the early genes may be involved in the coordination of mitosis and septation.

The *sepH* gene (Bruno et al., 2001), a member of the late group, is essential for the initiation of the actin ring and specially affects localization of the septin protein AspB (Momany and Hamer, 1997; Westfall and Momany, 2002), and SepA, a formin-related protein (Harris et al., 1997), which also locates to the septation site (Sharpless and Harris, 2002). SepH belongs to the family of protein kinases required for mitotic exit as defined by cdc7 from *S. pombe* and to CDC15 from *S. cerevisiae*. The yeast proteins are required for cytokinesis and for B-cyclin degradation at the end of mitosis. SepH is not required for hyphal growth and in the early stages of colony formation; the growth rate of *sepH* mutants is actually slightly higher than wild type. SepH function depends on normal cell-cycle progression: *sepH1* mutants grown at restrictive temperature will form septa on return to permissive temperatures but not if the cells have been treated with drugs that perturb cell-cycle progression.

T. Wolkow used temperature-sensitive mutants to show that septation could occur after a single mitosis, provided the cell was large enough. Indeed, downshift of *nimA5* cells back to permissive temperature is a useful means of obtaining a population of cells undergoing synchronized cytokinesis. The position of the mitotic nucleus seems to be critical in determining the position of the septum as *nim* mutants, grown at semipermissive temperature, form a septum at the site of the first nuclear division, provided the cell is large enough. Nud mutations that affect the migration and positioning of nuclei also affect the position of septa, which tend to be positioned close to the clusters of poorly separated nuclei.
Wolkow et al. also proposed a model where septation is inhibited by the proximity of the hyphal tip leading to an asymmetric cell division. Completion of mitosis is important for cytokinesis as most nim and bim mutations conditionally block septation. One exception to this general rule is hfaB mutants, which display a conditional “cut” phenotype, whereby a septum is laid down through a nondivided nucleus (Hughes et al., 2000).

The DNA damage-checkpoint pathway inhibits septation, acting through phosphorylation of NimX<sup>cdc2</sup> on tyr15 (Harris and Kraus, 1998). sepB3 mutants are conditionally defective for septation and define one of the early sep genes (Harris and Hamer, 1995). SepB is required for efficient chromosome segregation as mutants accumulate defects that eventually arrest growth and prevent septation. DNA damage induces phosphorylation on tyr15 (Ye et al., 1997) and if this is prevented either due to mutation of the tyrosine kinase, AnkA (the wee1 ortholog), or by mutation of tyr15 in NimX<sup>cdc2</sup> then the cells will undergo septation despite having suffered DNA damage (De Souza et al., 1999; Kraus and Harris, 2001). Thus, sepB3 nimX<sup>+/-24A</sup> double mutants contain a NimX protein that cannot be phosphorylated by the tyrosine kinase AnkA, but they form septa normally. This demonstrates that the DNA damage-checkpoint pathway affects septation indirectly, through regulation of NimX<sup>cdc2</sup>. A second indication of the close involvement of NimX in septation comes from the phenotype of a suppressor of nimX mutants, snxB1 (McGuire et al., 2000), which can lead to hyperseptation.

The positioning of septa may also be influenced by the energy balance or nutrient status of the cell, as the carbon source has been reported to affect the size of the tip cell formed after cell division (Muller et al., 2000). The AnkA-wee1 kinase is implicated in delaying septation under high carbon growth conditions as conditional mutations in AnkA sepaate at a given size regardless of carbon source (Kraus and Harris, 2001) and cell-size control in response to nutrient status may act through AnkA.

The APC has also been implicated in septation as a conditional mutation in sepI actually defines a new allele of the bimA gene, bimA10 (Wolkow et al., 2000). The bimA10 allele causes a splicing defect that leads to the production of an aberrant protein with an altered C-terminus and this leads to the accumulation of replication errors that trigger the DNA damage checkpoint. bimA10 double mutants with DNA damage-checkpoint mutants such as <i>uvsB110</i> or <i>uvsD153</i> partially suppress the failure to septate normally, indicating that the primary effect of the <i>bimA10</i> allele is in DNA replication and its effects on septation are secondary through activation of the DNA-damage pathway.

The spindle-pole body seems to play an important role in septum positioning since many proteins found there seem to affect septation. SnaD, a protein found at the spindle-pole body (Liu and Morris, 2000), affects the timing of mitosis and septation. Mutations in <i>snaD</i> suppress the growth defects of <i>nudA</i> mutants, not by affecting nuclear migration directly but by delaying the timing of mitosis and septation. This has the result that tip cells have a better chance of containing a nucleus and thereby retaining viability. Conversely, the <i>snaD</i> mutants have a defect in conidiation where the timing of septation is critical for cellular development. BimG PP1, another spindle-pole component, seems to play a role in septation. GFP studies indicate that BimG locates to the site of septation just after mitosis and follows the contractile ring as it divides the cell (Fox et al., 2002). Temperature-shift experiments support the idea that BimG plays a direct role in septation, but it is difficult to exclude potential indirect effects caused by perturbation of mitosis.

### 15.4.7 Genome Surveillance

Activation of the DNA-damage repair pathway can slow or arrest the cell cycle at several points, including G1/S, S-phase, and G2/M. The G2/M transition and septation are clearly very sensitive to genome damage, perhaps because these stages involve irreversible structural rearrangement of the cell that leads to separation of the genetic material. There is the possibility of repair by somatic recombination between two recently replicated DNA while they remain in the same nucleus, but after mitosis such recombinational repair is less likely and almost impossible after septation. It is easy to image, therefore, that there has been a strong selective advantage for delaying both mitosis and septation in the event of DNA damage.

Consistent with this idea, there is a conserved signal transduction pathway that detects and responds to DNA damage. In <i>A. nidulans</i>, the <i>uvsB</i> and <i>uvsD</i> genes encode crucial components and are structurally related to the ATM/ATR kinases (De Souza et al., 1999; Hofmann and Harris, 2000). The UvsB kinase is
related to Rad3 while UvsD is similar to Rad26, both PI-3 related kinases that modulate cell-cycle progression by phosphorylating several other proteins, including downstream kinases such as Rad53/Chs2, and ultimately leads to inhibitory phosphorylation on tyr14 of NIMX\textsuperscript{cdc2}, thereby delaying or blocking mitosis and septation (Fig. 15.6). The NIMX\textsuperscript{cdc2AF} mutation cannot be so inhibited and, therefore, does not respond to DNA damage by inducing cell-cycle arrest. Loss of these checkpoint functions can have downstream effects on the DNA-replication checkpoint and allow rereplication of a genome that has not completed mitosis successfully. This leads to endoreduplication and increases in ploidy.

Genetic dissection of this pathway has revealed several additional components (reviewed by Goldman and Käfer, 2004) and these are gradually being placed into what might be better described as a genome-monitoring network. Thus, uvsB acts both in the DNA replication and the intra-S-phase checkpoints (Fagundes et al., 2004) and affects both mitosis and septation (Hofmann and Harris, 2000) but other components seem to be specific for particular branches of the network. The \textit{musN} gene functions downstream of uvsB, probably in the septation-specific branch, since mutations in \textit{musN} suppress only the septation checkpoint in \textit{uvsB} mutant backgrounds (Hofmann and Harris, 2000). NpkA, a cdc2-like kinase, was isolated as a gene that is transcriptionally upregulated in response to camptothecin, a drug that inhibits type I topoisomerase and induces replication-mediated DNA double strand breaks (Fagundes et al., 2004). Deletion of \textit{npkA} partially suppresses the intra-S-phase checkpoint defect of \textit{uvsB} mutants but not the DNA-replication checkpoint. Consistent with this, deletion of \textit{npkA} is additive with \textit{ankA}\textsuperscript{Awt} mutations, but its interaction with \textit{bimE} mutations suggests that there is additional functional redundancy to be uncovered.

The \textit{sldL} gene encodes a Rad50 homolog that interacts genetically with \textit{bimE}\textsuperscript{APC}, mutants in response to DNA damage (Malavazi et al., 2005). Rad50 is a large coiled-coil protein related to the SMC protein family and is believed to bind to double stranded breaks in damaged DNA, effectively bridging the gap and allowing DNA repair enzymes to bind. Using the ScaA protein as bait in a yeast-two-hybrid screen, Semighini et al. (2003) isolated the Mre nuclease, one of these repair enzymes (Semighini et al., 2003).

\textit{A. nidulans}, in common with many fungi, can exist as haploid or diploid strains and occasionally interconvert between the two. This interconversion is the basis of the parasexual cycle that has been so useful for classical genetic analysis and mapping of mutations. Diploid strains arise very rarely, probably as a result of accidental nuclear fusion, but can be recognized and isolated from suitably marked heterokaryotic strains. The reverse process, haploidization, can be induced by transient growth on antimicrotubule drugs that lead to chromosome loss. Although the process of ploidy control remains poorly understood, it is amenable to genetic and molecular dissection. Targeted disruption of the \textit{chpA} (cysteine- and histidine-rich-domain-[CHORD]-containing protein A) gene in haploid \textit{A. nidulans} strains gives rise to \textit{chpA} knockout haploids, which are morphology normal, and heterozygous diploids, which develop abnormal conidiophores. However, \textit{chpA} knockout diploids were impossible to isolate and attempts to disrupt the remaining \textit{chpA} gene in heterozygous diploids lead to unstable aneuploids suggesting that ChpA is required for mitosis in diploid cells (Sadanandom et al., 2004). The molecular mechanism is unclear but related proteins in plants and animals interact with SGT1, a multifunctional protein associated with protein turnover.

15.5 Hyphal Morphogenesis and the Cell Cycle

\textit{bimG11} mutants also affect hyphal morphogenesis. At restrictive temperature, germinating spores fail to switch from isotropic to polar growth with the result that \textit{bimG11} spores swell to great size. GFP fusions indicate a direct role for BimG at the hyphal tip since the protein forms a collar around the base of the hyphal dome (Fox et al., 2002). \textit{bimG11} mutant strains have reduced levels of chitin in their cell wall (Borgia, 1992), and it was proposed that the phosphatase affected the pathway for chitin biosynthesis. However, it seems unlikely that the cell-wall defect is the sole basis of the phenotype as the cell swelling phenotype is much more extreme in germinating spores than in hyphae, which stop growing rather than swelling when shifted to restrictive temperature. Thus the roles of BimG seem to change after germ tube emergence: before emergence, the primary role is to define a point of polarized growth and a minor role in growth itself but after emergence the primary role is to maintain growth and the polarity maintenance role is less important.
The Aspergilli

The mechanism by which BimG establishes polar growth is uncertain, but seems to involve NimX\textsuperscript{CDC2}, nimX\textsuperscript{Af} partially suppresses the spore germination defect of the bimG11 mutant as double mutant cells germinate at restrictive temperature (Fox, Ph.D. thesis). The hyphae produced under these circumstances are unusually wide and have reduced cell extension rates, supporting the idea that BimG has an additional role in controlling growth, perhaps in vesicle recycling in the growing tip. PPI is known to affect vesicle membrane recycling in other organisms (Peters et al., 1999) and the location of the BimG halo approximately coincides with sites of endocytosis. Alternatively, BimG may affect actin function or organization since actin is intimately associated in a highly dynamic manner with both the tip and the septum.

Although NimX\textsuperscript{CDC2} is not required for hyphal outgrowth (nimX mutants germinate normally at restrictive temperature) (James et al., 1995), regulation of either NimX\textsuperscript{CDC2} or NimA activity might be required to couple spore germination with the nuclear division cycle. APC activity is essential for hyphal outgrowth as bimE7 and bimA mutants have delayed and reduced hyphal outgrowth under nonpermissive conditions. This requirement for APC depends on nimX as mutations that reduce NimX\textsuperscript{CDC2} activity suppress the polarization defect (James et al., 1995). NimA may also play a role as overexpression inhibits outgrowth so that under high nutrient conditions emergence is delayed until after the first mitosis, but under low nutrient conditions emergence takes place after the first S-phase. Completion of S-phase is essential for emergence under both conditions as mutations and drugs that lead to S-phase arrest tend to block emergence. The only exception to this general rule so far is nimX3, which blocks the nuclear cycle in late G1 but allows emergence before DNA replication is complete (Harris, 1999).

Germ tube emergence is also under developmental control. During spore formation and maturation, the nucleus becomes highly condensed and the cells enter a deeply dormant state that is dependent on WetA, a protein required to prevent precocious germination of nascent spores while they are still attached to the parent colony (Clutterbuck, 1969). However, nothing is known about the cell cycle stage-dependency of germination in wetA mutant spores.

15.6 Branching and Cell Cycle Control

Superficially similar to germ tube emergence, branching is also important for fungal growth and morphogenesis, and seems likely to be coupled to cell-cycle progression. A ts mutation in the ahbA gene leads to reduced branching and abnormal development associated with reduced nuclear number. When the gene was cloned, the ahbA1 mutation was found to be an allele of nimX (Lin and Momany, 2004). These data support the longstanding idea that the number of nuclei in a hyphal segment influences branching frequency (Dynesen and Nielsen, 2003) but the signals coupling nuclear number to branching are poorly understood. AhbB1, another gene isolated by Lin and Momany (2004), may provide a clue. Mutation of this gene also reduces branching, and it encodes a cytochrome P-450, possibly involved in steroid and fatty acid metabolism. Other lipid-derivatized compounds may also act as signals (Cheng et al., 2001). Sphingolipids are major components of the plasma membrane and their metabolism can generate potent signaling molecules. Mutations in the aur1 gene, which encodes inositol phosphorylceramide synthase, blocks in G1, perturbs the actin cytoskeleton and inhibits polarized cell growth. However, this mutation leads to changes in multiple lipid-derived signaling compounds, reducing sphingolipids and increasing levels of ceramide. To deconvolve this complexity, Cheng et al., looked at lcbA mutants defective in serine palmitoyltransferase (SPT), the first enzyme in the sphingolipid biosynthesis pathway, and at the effect of myriocin, a specific inhibitor of SPT. Reducing SPT activity led to a defect in actin-dependent hyphal morphogenesis without affecting the cell cycle. Thus, it seems likely that lipid-derived signals could be involved in the coordination of hyphal growth and branch formation with cell-cycle progression.

15.6.1 Developmental Regulators Impose New Discipline on the Cell Cycle

Coupling of the nuclear division cycle with cell growth and cytokinesis during hyphal growth differs dramatically from that observed during later stages of development. During both the asexual and sexual cycles, cell sizes and shapes and nuclear number per cell become more or less strictly defined and are...
often crucial for the identity and function of particular cell types (reviewed by Fischer and Kües, 2006). For example, the fertile hyphae within fruiting bodies are binucleate and, during asexual spore formation, cell and nuclear division become coupled so that each spore receives only one nucleus. Genes that control morphogenesis affect how cell and nuclear division are coupled. Thus the condiospores, end products of the brlA-controlled asexual reproductive pathway, contain a single nucleus per cell whereas the aerial hyphae from which they are derived contain many nuclei per cell. One effect of brlA, therefore, is to couple cell division more tightly with nuclear division than normally occurs in hyphal growth. A brlA-dependent increase in NimXcdc2 activity combined with the requirement for correct regulation of NimXcdc2 activity suggests that there is a direct developmental regulation of the cell cycle, acting through BrlA (Ye et al., 1999). Ectopic expression of the brlA gene in hyphae leads to one-to-one coupling between nuclear and cell division in hyphae leading to ectopic spore formation at the hyphal tips (Adams et al., 1988) as well as increased levels of NimXcdc2 (Ye et al., 1999). Not only are NimXcdc2 levels increased during sporulation, but regulation of its activity via inhibitory phosphorylation on tyr15 is also crucial as nimXcdc2Δ strains conidiate poorly with morphologically abnormal conidiophores. Tyr15 dephosphorylation had previously been found essential for conidiation: nimT23 mutants could be partially complemented by an extra copy of the nimE gene, which allowed hyphal growth but development was impaired (O’Connell et al., 1992). This suggests that extra nimE (cyclin B) may increase the amount of pre-MPF (tyrosine phosphorylated cyclinB/cdc2) available for nimT23-mediated activation and, eventually, the mutant phoshatase activates enough kinase to allow entry into mitosis. The filamentous hyphal cells apparently can cope with this rather sloppy control of mitotic entry but the cells that comprise the reproductive structures cannot, and development is impaired. The transcriptional regulation of NimXcdc2 could be quite direct as BrlA is a transcription factor with two TFIIB-type zinc fingers and the nimX upstream regulatory region has seven potential BrlA-binding motifs. BrlA also modulates cyclin expression: nimE transcript size is altered by expression of brlA.

Other sporulation-specific cell cycle controls may also operate to add additional layers of regulation. For example, an additional cyclin, pclA, is induced during sporulation in a brlA-dependent manner (Schier et al., 2001) and this physically interacts with NimXcdc2 (Schier and Fischer, 2002). PclA is related to pho85 cyclins in yeast but there is no evidence that PclA interacts with PhoA, one of the two A. nidulans orthologs of the yeast pho85 cyclin-dependent kinase. Deletion of the pclA gene severely reduced sporulation, indicating an important requirement for this cyclin during asexual development while mutation of phoA promotes the sexual development pathway. It is possible that these genes affect the developmental decisions made in response to environmental conditions (Bussink and Osmani, 1998).

The septation initiation network (SIN), the regulatory pathway that activates the contractile actin ring during septation, also seems to be more stringently controlled during conidiation. The mobA gene, a homolog of the yeast SIN gene, mboA, is not required for colony formation, but mobA mutants fail to conidiate. A screen for mutations that bypass the requirement for mobA has identified a number of genes that also bypass the requirement for SEPH kinase (Kim et al., 2006).

Analysis of the KfsA function supports the idea that septation is differentially regulated during development. The kfsA (kinase for septation) gene was discovered in a reverse-genetic approach and is involved in the regulation of septation in the conidiophore (Takeshita et al., 2007). The protein displayed some similarity to Kin4 of S. cerevisiae. Kin4 appears to monitor spindle misalignments and delays septation until nuclei are correctly distributed (D’Aquino et al., 2005; Pereira and Schiebel, 2005). In A. nidulans, KfsA localized to septa after the actin ring disappeared and neither deletion nor overexpression affected overall growth or the visual appearance of the colony. However, the number of conidiophores with ectopic septa in the stalk was increased and binucleate metulae were produced when KfsA levels were perturbed. This suggests similar roles for KfsA in A. nidulans and Kin4 in S. cerevisiae, but KfsA’s role in Aspergillus is only really critical during development.

The enhanced coordination between cytokinesis and nuclear division may involve other components of the brlA regulatory network. Absence of abaA function leads to incomplete separation of spores but cell growth and nuclear division continue (Sewall et al., 1990) while ectopic expression of abaA induces aberrant compartmentalization of the hyphae (Mirabito et al., 1989). Regulation of nimX may also be important for the suppression of septation during certain stages of conidiophore development as a suppressor of nimX leads to ectopic septa in the stalk (McGuire et al., 2000) as also occurs in strains with the...
activated nimXcd2AF allele (Ye et al., 1999). The molecular characterization of the nimX suppressor genes should provide insight into the interplay between the cell cycle and developmental regulation of growth.

15.7 Actin Cytoskeleton

15.7.1 Organization of the Actin Cytoskeleton

Immunostaining of actin or visualization with phalloidin derivatives revealed a spot-like distribution of the protein along the cortex in many fungi with a high concentration at the tip. In germinating spores, these actin spots are initially distributed evenly around the swelling spore but gradually accumulate at the site of hyphal emergence. In comparison, in Ashbya gossypii, actin cables are frequently seen (Schmitz et al., 2006). Meanwhile actin has been fused to GFP, which allows in vivo studies of the dynamics of actin (S. Osmani, personal communication). Furthermore, Penalva et al. fused an actin-binding protein with GFP, producing a useful tool to study actin localization and behavior in living A. nidulans cells (M. Peñalva, Madrid, personal communication). The important role that actin plays in polarized growth becomes obvious when depolymerizing agents, such as latrunculin B, or cytochalasin, are added to growing hyphae. Sampson et al. showed that addition of latrunculin B causes a rapid cessation in hyphal extension (Sampson and Heath, 2005). Likewise, deletion of the myosin gene, myoA, is lethal (McGoldrick et al., 1995). There are two likely contributions of the actin cytoskeleton to polarized growth. On the one hand, the actin-myosin cytoskeleton is used for vesicle transportation and secretion and thus the delivery of cell-wall components. On the other hand, cortical proteins are brought into place by this system in S. cerevisiae and guarantee proper attachment of MTs to the cortex (Schuyler and Pellman, 2001b).

Because MT attachment sites required for polarized growth seem to be very defined in the apical dome (see later), it is conceivable that the actin cytoskeleton plays a role at this point as well. However, further experiments are required to unravel the exact mechanisms.

Another aspect of polarized growth that we should consider is the existence of a Ca²⁺ gradient along the hypha, with a high concentration having been demonstrated at the tip of Phyllosticta ampelicida and N. crassa (Shaw et al., 2001; Silverman-Gavrila and Lew, 2003). In the absence of the Ca²⁺ concentration gradient, hyphal polarity is affected (Schmid and Harold, 1988). Although this phenomenon has been known for a long time, a direct link to the growth machinery described earlier has not emerged yet. One explanation for the role of Ca²⁺ ions is the stimulation of vesicle fusion with the membrane. The Ca²⁺ concentration appears to be regulated through a stretch-activated phospholipase C at the tip, which catalyzes the formation of inositol (1,4,5) triphosphate (IP₃) and in turn causes the release of Ca²⁺ from special vesicles (Silverman-Gavrila and Lew, 2002).

15.7.2 Polarisome

A protein complex involved in the organization of the actin cytoskeleton is localized at the incipient bud of S. cerevisiae and was named the polarsome. This structure is involved in the organization of the actin cytoskeleton and its appearance resembles the Spitzenkörper in filamentous fungi (Sagot et al., 2002). There is evidence that this protein complex also exists in filamentous fungi as a separate structure to the Spitzenkörper (Harris and Momany, 2004). The existence of polarsome components in filamentous fungi was shown first in A. nidulans. Sharpless and Harris demonstrated that SepA—an ortholog of Bni1, a key component of the yeast polarsome—colocalizes with the Spitzenkörper (Sharpless and Harris, 2002). Similarly, in A. gossypii, a filamentous fungus very closely related to S. cerevisiae (Wendland and Walther, 2005), a homolog of the S. cerevisiae polarsome protein Spa2 was analyzed (Knecht et al., 2003) and recently also the Bni1 ortholog, AgBni1 (Schmitz et al., 2006). Whereas Spa2 is not essential in A. gossypii, it is necessary for fast polarized growth, and deletion of Agbni1 caused loss of polarization and swelling of the cells to a potato-like appearance. A Spa2 ortholog has been characterized in Candida albicans as well and its role studied during filamentous growth (Zheng et al., 2003). The protein persistently localized at hyphal tips and deletion caused defects in polarity establishment. Recently, Crampin et al. suggested that the polarsome and the Spitzenkörper are distinct structures that coexist in hyphae.
(Crampin et al., 2005; Sagot et al., 2002). Similar results for Spa2 (SpaA) were obtained in *A. nidulans*, suggesting that a polarisome or the existence of polarisome components at the growing hyphal tip could be a general theme for filamentous fungi (Virag and Harris, 2006b). According to this model, filamentous fungal cells employ both the MT and the actin cytoskeleton, while the Spitzenkörper acts as a vesicle supply centre and the polarisome functions in actin organization.

The growth machinery discussed so far describes how fungi could extend their hyphae, but this picture does not yet allow any adaptation of the process to external (e.g., nutrient gradients) or internal signals (e.g., the stage of the cell cycle or life cycle). Little is known so far about the transduction of such signals into, for example, changes of growth direction, although several regulatory proteins have been described, which influence polarized growth, probably through an interaction with the actin cytoskeleton. The principle of this possible regulation is well studied in *S. cerevisiae* (Tcheperegine et al., 2005) and some of the components appear to be conserved in filamentous fungi. Among those are members of the Rho and the Rac family, small GTPases that act as molecular switches (Boyce et al., 2001, 2003, 2005; Guest et al., 2004; Momany, 2005; Virag and Harris, 2006a). However, a detailed analysis of their exact role in polarized growth in filamentous fungi remains to be done. Hyphae respond to environmental stimuli by altering their growth rate, diameter, and branching patterns. Genes such as *phoA* and *phoB*, which encode cyclin-dependent protein kinases, may play an important role in coupling growth to nutrient status (Dou et al., 2003) as available nutrients strongly affected the phenotype of knockouts.

### 15.7.3 Actin-Dependent Motor Proteins

The function of the actin cytoskeleton depends on the activity of actin-dependent motor proteins, the myosins. Myosins serve a broad range of cellular functions and are grouped into 18 different classes. In *A. nidulans*, a class-I myosin that is required for protein secretion and polarized growth and has an essential role (McGoldrick et al., 1995) and localizes to the growing hyphal tip (Yamashita et al., 2000).

Given that myosin motors are involved in vesicle transportation toward the cell cortex and vesicle fusion with the cell membrane, it is very interesting that *A. nidulans* employs a myosin-derived motor domain for the transportation of class-V and class-VI chitin synthases, where the motor domain is directly fused to the enzyme (Horiuchi et al., 1999; Takeshita et al., 2005; Takeshita et al., 2006).

Myosin motor proteins of other classes have been described. For example, in *S. cerevisiae*, a class V myosin motor is involved in peroxisomes and other organelles inheritance (Bretscher, 2003a; Fagarasanu et al., 2006). A second class V myosin is required for RNA transportation (Bretscher, 2003b).

### 15.8 Genes Required for the Establishment of Polarity

So far we have discussed polar growth in the sense of maintaining polarized extension by recruiting the cellular growth machinery for cell-wall assembly at the tip of an existing hypha. As mentioned earlier, the question how polarity is initially established starting from round spores remains largely unanswered. Hyphal emergence must involve localized cell modification such that one area of the wall becomes differentially susceptible to incorporation of new cell wall material. The polarized cytoskeleton described earlier is necessary to control this, the actin cytoskeleton being required *per se*, and the MTs fine-tuning the direction of growth. The ability to modulate cell-wall rigidity is also required as mutants lacking fibrillar components of the wall grow in a spherical manner. The secretion apparatus is also essential for hyphal outgrowth. Whittaker et al. (1999) showed that the sodVIC gene encodes an α-COP1 related protein that is essential for polarized outgrowth (Whittaker et al., 1999). In mutants that lack SodVIC function, nuclear division occurs to produce swollen deformed cells without obvious tips. COP1 proteins are important for vesicle formation and recycling. GFP-fusions with SodVIC localize to the Golgi (Assinder et al., unpublished), consistent with a role in secretion.

In ways that are not yet clear, polarity is under the control of the cell cycle regulatory network (see Section 15.5). Although this control could be indirect and through diverse pathways, it is now clear that cdc2-related proteins interact with, and control, a large number of proteins in other organisms (Ubersax et al., 2003). Consistent with this, cdc2 proteins in other organisms are located not only in the nucleus but...
also in the cytoplasm where they associate with the microtubules (Maekawa and Schiebel, 2004) and regulate MAP function.

In addition to cell cycle-related controls on spore germination, a number of additional functions have been described that directly affect polarization of the spores. Three temperature-sensitive mutations in the swoC, swoD, and swoF genes, cause spores to swell at restrictive temperature and prevent the production of a germ tube (Momany et al., 1999). The SwoC protein displayed homology to RNA pseudouridine synthases of yeast and its role in polarized growth remained obscure. SwoF on the other hand had high identity to N-myristoyl transferases and it was speculated that a polarity determinant could be the substrate for the myristoylation (Shaw et al., 2002). This posttranslational protein modification is found in proteins that switch between membrane-bound and cytoplasmic states (e.g., G-protein α-subunits), and could be important for the localization of cell-end marker proteins as discussed earlier (Bathnagar and Gordon, 1997). Therefore, the identification of prenylated or myristoylated proteins appears to be of prime importance for understanding polarity establishment in filamentous fungi.

15.9 Conclusions

The last few years have provided many new insights into the role of and interplay between actin microfilaments, MTs, and the nuclear division cycle in polarized growth of fungi. Actin plays a major role in cell growth and septation, the microtubules have a central role during mitosis, where they form the structure of the spindle, and through checkpoints mechanisms interact with the regulatory kinases that drive the cell cycle and cell growth. The circuits that connect the cell-cycle regulators to cellular morphogenesis are yielding to genetic and cell biological dissection and it will be very interesting to understand how these become modified to produce the more complex cell shapes that arise during the life cycle. During interphase in hyphae, the main function of microtubules is to deliver and direct vesicles and cell-end markers to the single point of growth at the apex of the cell, but during asexual development, growth first becomes dispersed over a large part of the vesicle surface and then becomes restricted again but to multiple points as the metulae are initiated. This function needs special attention, since only one putative cell-end marker protein has been identified so far in *A. nidulans*. If homologs of *S. pombe* cell-end markers exist in filamentous fungi, many questions remain: what is their biochemical function? Which downstream events do they trigger to allow straight hyphal growth? And with which upstream regulatory circuits are they integrated to permit the formation of different cell types? The latter question is particularly interesting since the developmental and morphological complexity of different cell types in *Aspergillus* suggest that the regulatory circuitry will be more complex than in single-celled fungi. The publication of several fungal genome sequences (Galagan et al., 2005) along with their innate genetic tractability and continuous improvement of molecular and microscopy techniques promise a fruitful future for cytoskeletal and cell-cycle research in fungi.

Acknowledgments

The work of our own research groups was supported by the German Science Foundation (DFG), the Fonds der Chemischen Industrie, the Max-Planck-Institute for terrestrial Microbiology, the special program “Lebensmittel und Gesundheit” from the Ministry of Baden-Württemberg and by the BBSRC. N. Takeshita is a fellow of the Humboldt Society.

References


Cytoskeleton, Polarized Growth, and the Cell Cycle in Aspergillus nidulans


Cytoskeleton, Polarized Growth, and the Cell Cycle in Aspergillus nidulans


16

Nuclear Pore Complex and Transport in Aspergillus nidulans

Eduardo A. Espeso and Stephen A. Osmani

CONTENTS

16.1 Introduction ................................................................. 261
16.2 Nuclear Pore Complex (NPC) and Its Regulation During Mitosis ............................................ 262
  16.2.1 Identification of the First NPC Proteins (Nups) from Filamentous Fungi Through Genetic Analysis of Mitotic Regulation ................................................................. 262
  16.2.2 Mitotic Restructuring of the Aspergillus nidulans NPC ...................................................... 264
  16.2.3 Transmembrane Nups Remain at the Nuclear Envelope (NE) During Mitosis .................. 266
  16.2.4 Structural Core of the NPC Remains at the NE During Mitosis ........................................ 266
  16.2.5 FG Repeat Nups and Other Peripheral Nups Disperse from the Structural Core of the NPC During Mitosis .......................................................... 267
  16.2.6 Gle1, a Predicted Peripheral Nup, Does Not Disperse from the Nuclear Envelope During Mitosis .................................................... 267
  16.2.7 An-Nup2 Locates to DNA During Mitosis .......................................................... 268
  16.2.8 In Aspergillus nidulans, the RanGTP Gradient Could Shift From Across the NE to Around DNA During Mitosis ......................................................... 268
  16.2.9 Mitosis in Aspergillus nidulans Is an Evolutionary Intermediate Between the Closed Mitosis of Yeasts and Open Mitosis of Higher Eukaryotes ............. 270
16.3 Soluble Nuclear Transporters of Aspergillus nidulans ................................................................. 271
  16.3.1 Importin β-Like Super Family .................................................. 272
  16.3.2 Nuclear Export Pathway and Its Regulation During Transcriptional Control ......................... 273
  16.3.3 Ntf2 and TAP/Nxt1 Transporters ...................................................... 274
16.4 Conclusions and Future Directions for Research ................................................................. 274
Acknowledgments .............................................................................................................................. 274
References .......................................................................................................................................... 274

16.1 Introduction

One of the defining features of eukaryotes is the sequestering of their genomes within nuclei. The nucleus is separated from the cytoplasm by the nuclear envelope (NE), which consists of lipid inner and outer membranes. The outer membrane is contiguous with the endoplasmic reticulum. In higher eukaryotes the nuclear lamina, consisting of intermediate filament like proteins, resides near the inner nuclear membrane and is thought to provide rigidity and the generally oval shape of the nucleus. Given that transcription is partitioned within nuclei, whereas protein translation occurs in the cytoplasm, it is essential to transport both RNA and protein in and out of nuclei. Such transport between the nucleoplasm and cytoplasm occurs through nuclear pore complexes (NPC), massive structures that provide regulated conduits through the NE. The NPC has a conserved overall structure and is constructed from multiple copies of ~30 NPC proteins termed nucleoporins, commonly abbreviated to Nups. The central transport channel of the NPC is formed by transmembrane Nups in combination with a conserved core multiNup...
The Aspergilli

16.2 Nuclear Pore Complex (NPC) and Its Regulation During Mitosis

16.2.1 Identification of the First NPC Proteins (Nups) from Filamentous Fungi Through Genetic Analysis of Mitotic Regulation

The first Nup genes to be identified in *A. nidulans* stemmed from the pioneering work of N. Ronald Morris who identified numerous genes encoding proteins specifically required for cell-cycle progression\(^1\) including proteins necessary for entry into mitosis. Of these the *nimA* gene, which encodes the essential NimA mitotic kinase, has been the most intensely studied cell-cycle regulatory gene in filamentous fungi\(^1\). More recent work has demonstrated that this kinase, along with the ubiquitous Cdk1 mitotic kinase, regulates mitosis by modifying the structure and function of the NPC\(^5,8,10,11\).

The first inclination that regulation of NPC was key to mitotic regulation in *A. nidulans* came from extragenic suppressor screens to identify mutation in genes that could suppress the temperature sensitivity, and G2 arrest, of the *nimA5, nimA7,* and *nimA1* alleles\(^10\). The extragenic suppressor approach was first utilized by Jarvik and Botstein investigating bacteriophage P22 morphogenesis\(^20\). This approach was subsequently utilized for genetic analysis of microtubule function, and most famously, led to the identification of the prototypic gamma-tubulin in *A. nidulans* by Liz and Berl Oakley\(^21\). Although the *nimA* extragenic suppressor screens utilized three different temperature-sensitive alleles of *nimA*, only the *nimA1* allele yielded mutations in genes that could suppress the original mutation’s heat sensitivity and G2 arrest phenotypes. This is probably because the *nimA1* point mutation is in the regulatory C-terminus of NimA\(^22\) which likely compromises the regulation of this kinase at higher temperatures but not its kinase activity. Presumably the defect in *nimA1* at higher temperatures can be suppressed by mutation of genes encoding proteins that interact with NimA. Conversely, the *nimA5* and *nimA7* mutations reside within the kinase catalytic domain and inactivate kinase activity at the higher temperatures. Therefore, because *nimA* is essential\(^22\) it appears that it is not possible to generate compensatory mutations in other genes to bypass NimA functions.

Mutations in two different genes, *sonA* and *sonB* (suppressor of *nimA1 A* and *B*) were identified in the extragenic suppressor screen. Subsequent molecular cloning demonstrated that the *sonA1* mutant allele...
FIGURE 16.1 The predicted localization of the nuclear pore complex proteins (Nups) of *A. nidulans* during interphase and mitosis. *Note:* Both core Nups and peripheral Nups localize to the NPC in interphase, but during mitosis the peripheral Nups disperse. Nup2 is the only Nup that associates with chromatin during mitosis. (From Osmani et al., *Mol Biol Cell*, 2006.) See color version in included CD.
The Aspergilli

encodes a WD-repeat Nup with a single-point mutation. SonA is similar to S. cerevisiae Gle2p and Saccharomyces pombe Rae1p. Importantly, an epitope-tagged version of this protein was found to locate around the nuclear periphery in a manner consistent with SonA residing in the NPC. These findings suggested that NimA may regulate the NPC and protein location during mitosis. Supporting this, it was found the Cdk1/cyclin B complex, which is normally nuclear, localizes to the cytoplasm during a nimA1 G2 arrest. Further, this defect in nuclear localization of Cdk1 and cyclin B in nimA1 mutants was remediated by the sonA1 Nup mutation which allowed entry into mitosis. This study identified the first component of the NPC in a filamentous fungus and also laid the foundation for further studies of the mitotic regulation of the NPC in A. nidulans.

Further evidence that NimA regulates the NPC came after the sonB1 mutation was found to encode a point mutation in the Nup98 component of the NPC and, hence, was named SonBnNup98. Studies in other systems have shown that orthologs of SonA and SonBnNup98 physically interact via a specific domain within Nup98 called the Gle2 binding site, or GLEBS. Accordingly, SonA was also found to physically interact with SonB via the GLEBS domain of SonBnNup98. Importantly, the sonB1 mutation changes the sequence of the GLEBS domain and weakens the interaction between SonA and SonB. To confirm that SonBnNup98 is a NPC protein it was endogenously GFP-tagged and found to locate around the nuclear periphery, similar to SonA, and characteristic of NPC proteins.

The fact that the only mutations identified that could suppress nimA1 temperature sensitivity and G2 arrest were within interacting components of the NPC suggests the NimA kinase regulates NPC function to promote mitosis. In further support of this idea, SonBnNup98 was found to be heavily phosphorylated during entry into mitosis in a NimA and Cdk1 dependent manner.

Insights to how NimA regulates the NPC to promote mitosis came from live cell imaging of the mitotic behavior of SonA and SonB endogenously tagged with GFP. Most dramatically, both SonA-GFP and GFP-SonBnNup98 were found to be completely released from the NPC and disperse throughout the cell during mitotic entry. Upon completion of mitosis, both SonA-GFP and GFP-SonBnNup98 return to the NPC as daughter nuclei are generated. This mitotic behavior of SonA-GFP and GFP-SonBnNup98 was unexpected and suggested that perhaps the NPC could be completely disassembled, as occurs during the open mitosis of higher eukaryotes. Alternatively, their behavior could be unique as both were isolated as extragenic suppressors of nimA1 and might, therefore, have specific mitotic roles. Subsequent identification (Table 16.1) and systematic tagging of most A. nidulans Nups has revealed that in fact the NPC does undergo dramatic, but not complete, disassembly during mitosis (Fig. 16.1).

16.2.2 Mitotic Restructuring of the Aspergillus nidulans NPC

Numerous studies in several different species have led to a comprehensive understanding of the general overall structure of the NPC. The NPC is constructed from ~30 different proteins, some of which interact to form subcomplexes within the overall NPC structure (Fig. 16.1a). In order for ~30 Nups to make a structure the size of the NPC they have to be present in multiple copies. The NPC is one of the largest multiprotein complexes in the cell (40 MDa in S. cerevisiae and 60 MDa in vertebrates) and is 20–30 times larger than a ribosome. The location of Nups within the overall structure of the NPC has been defined directly using immuno-EM techniques defining numerous core structural Nups and more peripheral components that decorate the core structure (Fig. 16.1a). Other more indirect approaches have been employed to help define Nups as being structural core components or more peripheral components within the NPC structure. One such approach has been to define the residence time of Nups within the NPC. This analysis indicates that central core components of the NPC are very stable whereas the more peripheral components exhibited much shorter residence times. Other studies of higher eukaryotes, which disassemble their NPC along with the NE during mitosis, have shown that structural Nups return first during the reassembly of the NPC during mitotic exit, followed by the more peripheral Nups (see Tran and Wente, 2006 and references therein). Collectively, these types of data have been used to infer whether a particular Nup plays a structural role, a transport role, or both a structural and transport role within the NPC. Recent data from A. nidulans, as discussed later, nicely support these findings. It has been demonstrated that some Nups remain during mitosis whereas many Nups disperse from the NPC during A. nidulans mitosis (Fig. 16.1b). The A. nidulans data define Nups...
TABLE 16.1
Nuclear Pore Complex Proteins (Nups) of *A. nidulans*

<table>
<thead>
<tr>
<th><em>A. nidulans</em> Protein/Systematic Name</th>
<th><em>S. cerevisiae</em></th>
<th>Vertebrates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transmembrane</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>—</em></td>
<td>—</td>
<td><em>—</em></td>
</tr>
<tr>
<td>An-Pom152/AN3454</td>
<td>Pom34p</td>
<td><em>—</em></td>
</tr>
<tr>
<td>An-Ndc1/AN4417</td>
<td>Ndc1p</td>
<td>Pomp121</td>
</tr>
<tr>
<td><em>—</em></td>
<td></td>
<td><em>—</em></td>
</tr>
<tr>
<td>Nup84/107 Complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SonBuNup96/AN5627</td>
<td>Nup145cp</td>
<td>Nup96</td>
</tr>
<tr>
<td>An-Nup133/AN4293</td>
<td>Nup133p</td>
<td>Nup133</td>
</tr>
<tr>
<td>An-Nup120/AN1238</td>
<td>Nup120p</td>
<td>Nup160</td>
</tr>
<tr>
<td>An-Nup85/AN9109*</td>
<td>Nup85p</td>
<td>Nup75/85</td>
</tr>
<tr>
<td>An-Nup84/AN1190*</td>
<td>Nup84p</td>
<td>Nup107</td>
</tr>
<tr>
<td>An-sec1/AN4317</td>
<td>Sec13p</td>
<td>Sec13</td>
</tr>
<tr>
<td>An-Seh1/AN5889*</td>
<td>Seh1p</td>
<td>Seh1</td>
</tr>
<tr>
<td><strong>Gle1 and Nup170</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>An-Gle1/AN1157</td>
<td>Gle1p</td>
<td>Gle1</td>
</tr>
<tr>
<td>An-Nup170/AN6738</td>
<td>Nup170p</td>
<td>Nup155</td>
</tr>
<tr>
<td><strong>Nic96 Complex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>—</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>An-Nic96/AN6980</td>
<td>Nic96p</td>
<td>Nup93</td>
</tr>
<tr>
<td>An-Nup192/AN0037*</td>
<td>Nup192p</td>
<td>Nup205</td>
</tr>
<tr>
<td>An-Nup188/BLASTn</td>
<td>Nup188p</td>
<td>Nup188</td>
</tr>
<tr>
<td><em>—</em></td>
<td></td>
<td><em>—</em></td>
</tr>
<tr>
<td><em>—</em></td>
<td>Nup59p</td>
<td><em>—</em></td>
</tr>
<tr>
<td><em>—</em></td>
<td>Nup53p</td>
<td>Nup35</td>
</tr>
<tr>
<td><strong>FG Repeat Nups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>An-Nup159/AN2086</td>
<td>Nup159p</td>
<td>Nup214</td>
</tr>
<tr>
<td>An-Nup57/AN1064</td>
<td>Nup57p</td>
<td>Nup54</td>
</tr>
<tr>
<td>An-Nup49/AN2431</td>
<td>Nup49p</td>
<td>Nup58</td>
</tr>
<tr>
<td>An-Nup42/AN4595</td>
<td>Nup42p</td>
<td>Npl1</td>
</tr>
<tr>
<td>An-Nsp1/AN4499</td>
<td>Nsp1p</td>
<td>Nup62</td>
</tr>
<tr>
<td>SonBuNup1061/AN5627</td>
<td>Nup116p</td>
<td>Nup98</td>
</tr>
<tr>
<td>An-Nup2/AN5485*</td>
<td>Nup2p</td>
<td>Nup50</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>An-Nup82/AN6143</td>
<td>Nup82p</td>
<td>Nup88</td>
</tr>
<tr>
<td>An-Mlp1/AN5499</td>
<td>Mlp1/2p</td>
<td>TPR</td>
</tr>
<tr>
<td>An-Sac3/AN7726</td>
<td>Sac3p</td>
<td>Shd1</td>
</tr>
<tr>
<td>SonAtnp2/AN1379</td>
<td>Gle2p</td>
<td>Rae1</td>
</tr>
</tbody>
</table>

*Automated gene structure incorrect as noted in Osmani et al., 2006.

In *A. nidulans* the transmembrane Nups, components of the Nup84/107 subcomplex along with Gle1 and Nup170 remain at the NE during mitosis and are thus considered core structural components of the NPC. All other Nups disperse from the NPC during mitosis and are considered peripheral non-structural Nups.

as either structural components that reside at the NPC throughout mitosis (Fig. 16.1) or as more peripheral components that are dispersed from the NPC during mitosis (Fig. 16.1). Here, we review these data and attempt to integrate the *A. nidulans* findings with previous studies. Each class of Nup will be covered, from the transmembrane Nups and core Nups that are thought to provide the basic framework of the NPC, outward to the more peripheral components of the complex, which are thought to help mediate transport. Although there is excellent correlation between the inferred location and function of *A. nidulans* Nups with previous studies, some surprising, and perhaps informative, differences are discussed.
16.2.3 Transmembrane Nups Remain at the Nuclear Envelope (NE) During Mitosis

Vertebrates and *S. cerevisiae* each contain three transmembrane Nups that presumably function to physically tether the NPC to the NE (Table 16.1). A part of each transmembrane Nup is thought to interact with other core components of the NPC, while transmembrane domains pass through one lipid bilayer of the NE resulting in a part of each transmembrane Nup residing in the lumen of the NE (Fig. 16.1). This arrangement is thought to anchor the NPC within the NE membrane.26 Orthologs of two transmembrane Nups exist in *A. nidulans*, An-Pom152 and An-Ndc1 (Table 16.1). Both have the expected putative transmembrane signatures and both locate at the nuclear periphery, confirming that they are components of the NPC. During mitosis, An-Pom152 and An-Ndc1 remained associated with the NPC, consistent with both proteins being part of the structural core of the NPC. In addition to locating at the NPC during mitosis, An-Ndc1 also locates to two conspicuous and dynamic foci within the NE. These foci separate as mitosis proceeds (see Figure 5 and Supplemental Movie 8 in Osmani et al., 2006). It has been proposed that the two mitotic An-Ndc1 foci represent concentration at the spindle pole bodies (SPBs) that are separating as spindle elongation occurs. This dual location of Ndc1 in the *A. nidulans* NPC and the mitotic SPB is consistent with the known location of its orthologs in both fission (Cut11) and budding yeast.27,28 The role of An-Ndc1 at the SPB during mitosis is not known, although in *S. pombe* Cut11 plays a role in inserting the SPB into the NE.29 Deletion of either An-Pom152 or An-Ndc1 does not cause any detectible phenotypes so their functions at the NPC, and/or at the SPB are not essential. At this time it is not known if *A. nidulans* can survive without both these transmembrane Nups. It is also not known if, like vertebrates and *S. cerevisiae*, a third transmembrane Nup (Table 16.1) exists in this species. It is somewhat surprising that the transmembrane proteins in *A. nidulans*, which would be expected to tether the NPC within the NE, are not essential. This suggests that perhaps there are alternative NPC-anchoring mechanisms operative in addition to the transmembrane anchors.

16.2.4 Structural Core of the NPC Remains at the NE During Mitosis

Moving outward from the NE toward the central channel of the NPC, resides the core of the NPC comprised largely of the Nup84 subcomplex.30,31 This is the largest subcomplex of the NPC and is conserved between yeast and higher eukaryotes. In vertebrates, and some other systems, the Nup84 subcomplex is termed the Nup107–160 subcomplex.32–34 All seven core members of the *S. cerevisiae* Nup84 subcomplex have been identified in *A. nidulans* (Table 16.1). All Nup84 components have been endogenously tagged with a fluorescent protein, with the exception of An-Seh1, which, for unknown reasons, has proven resistant to C-terminal tagging. All members of the Nup84 subcomplex locate to the NPC during interphase confirming them as NPC proteins. During mitosis, the *A. nidulans* Nup84 subcomplex proteins remain at the NPC, as would be expected of components of a structural core of the NPC (Fig. 16.1b). These data indicate that the Nup84 subcomplex constitutes, in addition to the transmembrane Nups, the structural core of the *A. nidulans* NPC, similar to other systems studied.

Deletion analysis has demonstrated that five of the eight Nup84 subcomplex components are not essential in *A. nidulans*. However, three of the nonessential deletions cause both temperature sensitivity and self sterility. These findings are consistent with data from *S. cerevisiae* where deletion of the orthologous five genes (*Seh1, Nup84, Nup85, Nup120*, and *Nup133*) does not cause lethality. However, in *S. pombe* deletion of either *Nup84* or *Nup85* is lethal.34 Why the NPC structure of *S. pombe* is less tolerant of these deletions is an interesting but as yet unanswered question.

Given that each of the five nonessential Nup84 subcomplex proteins are predicted to be present within each NPC structure repeated 16 times,30,35 it is surprising that the NPC can function without them. In the yeast systems deletion of some of the Nup84 subcomplex components causes a dramatic clustering of all NPCs in the NE and, as also observed for *A. nidulans*, significant temperature sensitivity. So, it is clear that the nonessential Nup84 complex components are required for completely normal NPC function, but there must be enough redundancy within the structural framework, and Nup–Nup interactions, for the NPC structure to be able to compensate for the loss of some core components. It will be interesting to determine what effects these nonessential deletions have on the minimal mitotic core of the NPC in
A. nidulans (see later) and what specific defects are caused when the temperature-sensitive null alleles are shifted to the restrictive temperature.

### 16.2.5 FG Repeat Nups and Other Peripheral Nups Disperse from the Structural Core of the NPC During Mitosis

Many Nups contain numerous FG (phenylalanine glycine) repeats in their primary sequence. The FG repeat Nups are thought to be anchored in the core structure of the NPC but also fill the central conduit of the NPC where they prevent diffusion through the pore. The FG repeat Nups also bind to, and facilitate, transport of cargo complexes through the transport channel of the NPC. This class of Nup are, therefore, not considered part of the core structure of the NPC but are essential for regulated transport through the NPC, and provide a barrier to nonspecific diffusion. *A. nidulans*, when compared to *S. cerevisiae*, appears to lack several members of the FG-repeat family of Nups including Nup1, Nup53, Nup59, and Nup60. In addition, *A. nidulans* encodes a single FG repeat Nup98-like protein whereas *S. cerevisiae* encodes three Nup98-like proteins (Nup100, Nup145, and Nup116). Thus *A. nidulans* apparently encodes six fewer FG repeat Nups compared to *S. cerevisiae* and many of the *A. nidulans* FG repeat Nups contain fewer repeats.

As described earlier, SonBnNup98 which was isolated as an extragenic suppressor of nimA1, is selectively removed from the core NPC structure during mitosis. It has subsequently been demonstrated that all FG repeat Nups are selectively removed from the core of the NPC during *A. nidulans* mitosis (Table 16.1, Fig. 16.1). Not only are all seven FG repeat Nups dispersed from the NPC during mitosis, but so are four other Nups that would be predicted to be either mobile Nups (SonAGle2) or located at the peripheral of the NPC at the cytoplasmic (An-Nup82, An-Sac3), or nucleoplasmic (An-Mlp1) side of the NPC. From these findings it is clear that FG repeat Nups cannot provide a diffusion barrier or mediate nuclear transport during *A. nidulans* mitosis. These findings also indicate that there is a mechanism by which all of the FG repeat Nups, and other peripheral Nups, are removed from the NPC specifically during mitosis. A mechanism must also exist to promote the return of the dispersed Nups to the NPC during exit from mitosis.

### 16.2.6 Gle1, a Predicted Peripheral Nup, Does Not Disperse from the Nuclear Envelope During Mitosis

There is an excellent correlation between the behavior of the *A. nidulans* Nups during mitosis and their predicted location and function. Thus, all predicted core components, including the transmembrane Nups and all Nup84 subcomplex members, remain at the NPC during mitosis. Similarly, all Nups predicted to reside around the periphery of the NPC structure are dispersed from the NPC during mitosis. However, there is one notable exception to this correlation because the RNA export factor Gle1, which is considered a peripheral Nup that resides at the cytoplasmic face of the NPC, does not disperse from the NPC during mitosis but remains at the NE. This is a very surprising finding because *A. nidulans* homologs of Nups known to interact with Gle1 in other systems are dispersed during *A. nidulans* mitosis.

Recent work has defined the biochemical role of Gle1 in RNA transport. Gle1 binds to other Nups such that it is anchored at the cytoplasmic face of the NPC where it physically interacts with the DEAD box helicase Dbp5. Dbp5's ATPase activity and RNA binding properties are stimulated by binding to Gle1, and also by inositol hexakisphosphate. Thus Gle1 has been proposed to mediate RNA export by spatially controlling local activation of Dbp5 to facilitate RNA transport and perhaps provide energy to power this transport. Given this role for Gle1 in RNA export, it is difficult to rationalize why Gle1 remains at the NE, presumably at the NPC, during mitosis. The structural Nups that remain likely provide both a conduit across the NE during mitosis and a framework to which the dispersed Nups return when regulated nuclear transport in G1 is reestablished. But Gle1 is not considered a structural Nup. Because Gle1 plays a regulatory role to help mediate RNA transport, this may suggest that some RNA transport occurs during *A. nidulans* mitosis. This is unlikely for numerous reasons, the most important of which is the fact that An-Dbp5 does not remain at the NPC during mitosis (Liu and Osmani, unpublished). At this time
the role of Gle1 at the NE during mitosis remains a mystery but will likely be the focus of further research as the current findings suggest that An-Gle1 could play a novel, and unexpected, role during mitosis.

16.2.7 An-Nup2 Locates to DNA During Mitosis

Another surprising location for a Nup during mitosis has been described for An-Nup2 which, as expected of an FG repeat Nup, is dispersed from the NPC during mitosis. However, unlike the other 13 Nups that are released from the NPC at mitosis, An-Nup2 does not disperse throughout the cell. Instead, An-Nup2 concentrates on condensed chromatin from prophase through to telophase (Fig. 16.1b). As chromatin decondenses during mitotic exit An-Nup2 is released from DNA and locates back to the NPC as nuclear transport is reestablished in G1. Given this very distinctive dual location through the cell cycle it has been suggested that An-Nup2 may play one role at the NPC to promote nuclear transport and another role at chromatin to help regulate mitosis. In this manner An-Nup2 could coordinate the changing structure and function of the NPC with mitotic specific events such as chromosome condensation and spindle formation.8

Insights to the possible function of An-Nup2 in nuclear transport, and during mitosis, are suggested from studies of S. cerevisiae Nup2p.45–53 S. cerevisiae Nup2p is a multidomain protein consisting of an N-terminal Kap60p (importin α)-binding domain followed by a NPC-targeting domain, an FG repeat domain and a C-terminal Ran binding domain.54 An-Nup2 also contains similarly placed domains. In S. cerevisiae, Nup2p, helps facilitate nuclear transport due to its ability to accelerate the release of cargo from transport complexes on the nuclear side of the NPC.54 Given the overall structural similarity between An-Nup2 and S. cerevisiae Nup2p, it is reasonable to suggest An-Nup2 could play a similar role in nuclear transport. However, during mitosis the potential ability of An-Nup2 to accelerate the release of cargo from transport complexes in A. nidulans has been hypothesized to occur in the vicinity of mitotic chromatin.8 This would have the effect of releasing proteins required for mitosis in an active form around DNA. This concept is analogous to the proven ability of RanGTP (see next section), generated around DNA due to the action of RCC1 (Ran guanine nucleotide-exchange factor = RanGEF, Fig. 16.2a), to release mitotic promoting factors from importin β around DNA during the open mitosis of higher eukaryotes55 (discussed in the following). Taking these concepts a step further, it has also been suggested that because An-Nup2 contains a C-terminal Ran-binding domain, chromatin-bound mitotic An-Nup2 could concentrate RanGTP at chromatin to further promote mitotic events.8 Clearly this is an area ripe for further experimentation.

It is noticeable that An-Nup2 is almost twice the size of S. cerevisiae Nup2. The Nup2 orthologs of other filamentous fungi are also much larger than S. cerevisiae Nup2. This increase in size could perhaps be required for the proposed mitotic functions of An-Nup2. Further supporting a potential mitotic specific role for An-Nup2 is the fact that deletion of An-Nup2 does not prevent entry into mitosis but does lead to lethal mitotic defects.8 It will be important to determine the exact nature of these mitotic defects and to understand if the defects are due to misregulation of nuclear transport in interphase, or because of the lack of An-Nup2 function at DNA during mitosis. Unlike A. nidulans, in the yeasts S. cerevisiae and S. pombe, as well as the worm Caenorhabditis elegans, Nup2 is not essential, perhaps reflecting lack of a mitotic-specific role for Nup2 in these species.

16.2.8 In Aspergillus nidulans, the RanGTP Gradient Could Shift From Across the NE to Around DNA During Mitosis

Ran is an abundant and highly conserved small Ras-like GTPase that plays essential functions in both nuclear transport and mitotic regulation in higher eukaryotes.56 Like all small GTPase enzymes, Ran can switch between a form bound to GTP (RanGTP) and another bound with GDP (RanGDP). This transition is regulated by two Ran interacting enzymes. The first, RCC1 (regulator of chromosome condensation) has guanine nucleotide exchange factor (GEF) activity and converts RanGDP to RanGTP. Because RCC1 is attached to nuclear DNA, RanGTP is generated within nuclei during interphase. The other Ran interacting enzyme, Ran GTPase-activating protein (RanGAP) is located in the cytoplasm. RanGAP stimulates the GTPase activity of Ran converting RanGTP into RanGDP in the cytoplasm. Thus, a gradient across the
FIGURE 16.2 The soluble nuclear transport machinery. Note: (a) The Ran GTPase cycle. As described in the text, Ran alternates between its GTP- and GDP-bound states. The nuclear compartmentalization of the RanGDP exchange factor (RanGEF) and the cytoplasmic localization of the RanGTPase activating protein, RanGAP, generates a gradient of RanGTP across the nuclear envelope. (b) A simplified scheme of a nuclear import pathway. Cytoplasmic NLS-cargo binds to a soluble import transporter (Imp = importin β family member), which mediates interactions with the NPC to facilitate import to nuclei. Interaction of the imported complex with nuclear RanGTP promotes the release of the cargo within the nucleus. (c) A simplified model of the nuclear export pathway. Nuclear NES-cargo binds to export transporter (Exp = exportin β family member), which mediates interactions with the NPC to facilitate export from nuclei. On the cytoplasmic side of the NPC, the conversion of RanGTP to RanGDP causes the disassembly of the transport complex and release of cargo in the cytoplasm. (For a more inclusive review see Ref. 78.) See color version in included CD.
NE is generated whereby RanGTP concentration is high in the nucleus and low in the cytoplasm\(^1\) (Fig. 16.2a). This RanGTP gradient is crucial for nuclear transport due to the effect that RanGTP has on cargo complexes. For instance, during nuclear import, cytoplasmic cargo proteins with classic nuclear localization sequences (NLS) are bound to soluble carrier proteins (importin \(\alpha\) and \(\beta\)) and transported through the NPC into the nucleus by a poorly understood mechanism. When the cargo complex arrives within the nucleus it is exposed to high RanGTP levels. RanGTP binds to importin \(\beta\), causing the release of the cargo from the transport complex, thus delivering it to the nucleoplasm\(^5\) (Fig. 16.2b).

During higher eukaryotic mitoses, the NPC is disassembled along with the NE such that no nuclear structure exists and mitosis is said to be “open.” In this state no nuclear gradient of RanGTP can exist. However, as RCC1 remains bound to mitotic chromatin, whereas RanGAP is dispersed throughout the cell, RanGTP accumulates in the vicinity of chromosomes. This concentration of RanGTP near DNA releases proteins important for mitotic progression from the inhibitory effects of importin \(\beta\) in the vicinity of chromosomes. In higher eukaryotic cells therefore, the switch from an interphase RanGTP gradient across the NE to a gradient around DNA allows proteins to be released from importin \(\beta\) in the nucleus during interphase and around condensed DNA during mitosis. It was thought that this dual role of the Ran GTPase cycle would be specific to cell types undergoing open mitosis. However, as described later, the Ran GTPase cycle could similarly have dual roles in nuclear transport and mitosis in \textit{A. nidulans}.

Ran, RCC1, and RanGAP have been identified in \textit{A. nidulans} and all have been shown to be essential.\(^8\) Although An-Ran has not yet been successfully tagged with a fluorescent protein, both An-RCC1 and An-RanGAP have been endogenously tagged with GFP and their locations followed during mitosis. As expected, An-RCC1 remains bound to DNA throughout the cell cycle, even during mitosis when chromatin becomes visibly condensed.\(^7\) On the other hand, An-RanGAP is excluded from nuclei throughout interphase and nuclei are evident as dark shadows against the cytoplasmic signal of An-RanGAP-GFP. Upon entry into mitosis, cytoplasmic An-RanGAP-GFP enters nuclei and equilibrates between the nucleoplasm and the cytoplasm. As mitosis is completed, and daughter nuclei are formed, An-RanGAP-GFP is excluded from nuclei, which then become apparent again as dark shadows within the cytoplasmic signal of An-RanGAP-GFP.\(^7\) It is, therefore, predicted that the RanGTP gradient across the NE is greatly compromised during \textit{A. nidulans} mitosis because An-RanGAP would be able to stimulate conversion of RanGTP into RanGDP within nuclei, as well as in the cytoplasm. Therefore, as in higher eukaryotic open mitosis,\(^5\) it has been suggested that the RanGTP gradient shifts from across the NE to around mitotic chromosomes in \textit{A. nidulans}.\(^7\) In this manner, mitosis in \textit{A. nidulans} could be regulated by release of importin \(\beta\) mitotic regulators around chromosomes as occurs during open mitosis.\(^5\)
during *A. nidulans* mitosis are caused by the partial disassembly of the NPC and opening of the transport conduit to allow free diffusion. Therefore, during *A. nidulans* mitosis, proteins are predicted to locate within the cell based upon diffusion and relative binding affinities rather than by active transport through the NPC.\(^7^,\)\(^8\)

In contrast to the situation in *A. nidulans*, in the yeast systems there is no indication that the NPC is disassembled during mitosis. No global changes in the subcellular location of nuclear or cytoplasmic proteins have been reported for either *S. cerevisiae* or *S. pombe*. Similarly, no Nups have been shown to be released from the NPC at any time during the yeast cell cycles. Instead, it has been shown that, at least in *S. cerevisiae*, specific Nup-Nup interactions change during mitosis and modify specific transport pathways, which may help promote transit through mitosis.\(^6\)

In contrast to the widely held view that mitosis is either open or closed, clearly there are intermediates between these two extremes. In addition to the dramatic changes in both the structure and transport properties of the mitotic NPC in *A. nidulans*, work in *Ustilago maydis* demonstrates that this fungus dramatically modifies its nuclear envelope structure during mitosis.\(^6\)\(^1\) In fact, early cytological studies of many different species indicate there is likely to be a range of mitotic types from the closed systems of yeasts, through the partially open system of *A. nidulans* to the completely open system of higher eukaryotes.\(^6\)\(^2\) The experimental methodologies available when working with *A. nidulans* and *U. maydis*, coupled with the fact that they have intermediary types of mitoses, firmly establishes these fungi as important systems in which to study and understand the structure and function of the nucleus through the cell cycle.

### 16.3 Soluble Nuclear Transporters of *Aspergillus nidulans*

Proteins destined to be transported into the nucleus contain specific sequence motifs called nuclear localization sequences (NLS). Similarly, proteins to be exported from the nucleus contain nuclear export sequences (NES). The NLS and NES targeting motifs promote interaction with soluble transport proteins termed karyopherins, which mediate transport of these NLS/NES cargo proteins through the NPC and release to the destination site\(^6\)\(^3\) (Figs. 16.2b and 16.2c). In the case of protein import, the NLS is detected by a specific class of soluble transport proteins whose role is to facilitate transport into nuclei. The soluble transport proteins provide specificity to nuclear import by detecting and binding to the NLS sequence of their specific target cargoes. In this manner only proteins containing NLS sequences are targeted for nuclear import. In addition to selecting target cargoes, the karyopherins also physically interact with FG-repeat NUPS such that the karyopherin-NLS-cargo complex first binds to the cytoplasmic side of the NPC transport channel. The complex then moves through the transport channel of the NPC by poorly defined physical interactions between the karyopherin and the FG-repeat Nups. Upon arrival within the nucleus the karyopherin-NLS-cargo complex needs to be disengaged to release the cargo protein. This is achieved by nuclear RanGTP binding to the karyopherin, which causes a conformational shift releasing the cargo, and thus completing the nuclear import of the cargo (Fig. 16.2b, see Stewart 2007\(^5\)\(^4\) for recent detailed review and further references).

The basic concepts of nuclear export are similar to protein import. The NES sequence of target nuclear cargo is bound to a specific export karyopherin forming a cargo-NES-karyopherin complex in the nucleus. However, rather than RanGTP inhibiting this interaction, the binding of cargo-NES to the export karyopherin is promoted by RanGTP (Fig. 16.2c). The cargo-NES-karyopherin complex is transported through the NPC via interactions between the karyopherin and the FG-repeat Nups. Upon arriving at the cytoplasmic side of the NPC, RanGTP is converted to RanGDP by the action of cytoplasmic RanGAP and this causes the release of the cargo-NES from its export karyopherin (Fig. 16.2c). Therefore, RanGTP has opposite roles during NLS-mediated protein import and NES-mediated protein export. Nuclear import karyopherins bind their cargo in the absence of RanGTP but release them in the presence of RanGTP. Conversely, nuclear export karyopherins bind their cargo in the presence of RanGTP and release them in the absence of RanGTP. The import and export pathways, therefore, rely on the concentration of RanGTP being high in the nucleus but low in the cytoplasm. This is achieved through the activity of RanGEF,
which is found bound to chromatin within nuclei, and the action of RanGAP, which locates in the cytoplasm as described earlier (Fig. 16.2a).

Different karyopherins can mediate either protein import or export or both. Those karyopherins involved in the transport of macromolecules from the nucleus to the cytoplasm are generally called exportins and those specialized for protein import are termed importins. Karyopherins can be classified into different families depending on their amino acid sequence similarity.

16.3.1 Importin β-Like Super Family

The karyopherin family with the most numerous members is the importin β-like super family, composed of 15 members in *S. cerevisiae*, (including importin α), while at least 28 members are encoded in the human genome. All these nuclear transporters share a common domain organization and the authors, and Mans et al., have independently searched for members of this super family of nuclear transporters in *A. nidulans*. The results of these searches are shown in Table 16.2, indicating that the soluble transport factors are generally conserved. We identified 14 *A. nidulans* importin β-like proteins (including importin α) using either yeast or human protein sequences for BLAST searches against the *A. nidulans* predicted proteome. Of the budding yeast Kaps, only the putative Kap108p/importin 7/RanBP7 ortholog has not been found in *Aspergillus*. This is most likely because in

<table>
<thead>
<tr>
<th>TABLE 16.2</th>
<th>Soluble Nuclear Transporters of <em>A. nidulans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene/Systematic Name</strong></td>
<td><strong>S. cerevisiae</strong></td>
</tr>
<tr>
<td><strong>Ntf2 Family</strong></td>
<td></td>
</tr>
<tr>
<td>ntfA/AN4942</td>
<td>Ntf2</td>
</tr>
<tr>
<td><strong>Mex67/Mtr2-TAP/p15 Family</strong></td>
<td></td>
</tr>
<tr>
<td>mexA/AN2737</td>
<td>Mex67</td>
</tr>
<tr>
<td>ntxA/AN3864</td>
<td>—</td>
</tr>
<tr>
<td><strong>Importin β-Like Family</strong></td>
<td></td>
</tr>
<tr>
<td>kapA/AN2142</td>
<td>Cap60p(Srp1p)</td>
</tr>
<tr>
<td>kapB/AN9006</td>
<td>Kap95p (Rps11p)</td>
</tr>
<tr>
<td>kapC/AN0926</td>
<td>Kap104p</td>
</tr>
<tr>
<td>kapD/AN6006</td>
<td>Kap108p (Sxm1p)</td>
</tr>
<tr>
<td>Kap119p (Nmd5p)</td>
<td>Importin 8 (RanBP8)</td>
</tr>
<tr>
<td>kapE/AN6591</td>
<td>Kap109p (Cse1p)</td>
</tr>
<tr>
<td>kapF/AN6734</td>
<td>Kap111p (Mtr10p)</td>
</tr>
<tr>
<td>kapG/AN2164</td>
<td>Kap114p (Hc1004p)</td>
</tr>
<tr>
<td>kapH/AN4053</td>
<td>Kap120p (Lph2p)</td>
</tr>
<tr>
<td>kapI/AN5717</td>
<td>Kap121p (Pse1p)</td>
</tr>
<tr>
<td>kapJ/AN2120</td>
<td>Kap123p (Yrb4p)</td>
</tr>
<tr>
<td>kapK/AN1401</td>
<td>Kap124p (Crm1p)</td>
</tr>
<tr>
<td>kapL/AN3012</td>
<td>Kap142p (Msn5p)</td>
</tr>
<tr>
<td>kapM/AN8787</td>
<td>los1p</td>
</tr>
<tr>
<td>kapN/AN7731</td>
<td>Kap122p (Pdr6p)</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

ND = not detected.
Aspergillus, and in other filamentous fungi, \textit{S. cerevisiae} Kap108p and Kap119p are encoded by a single gene, \textit{KapD}.

We have started a systematic analysis of these karyopherins, by both tagging them with fluorescent proteins and constructing null alleles. We have sequenced the cDNAs of 10 of these transporters, which confirm in most cases the coding predictions of the automatic gene annotation (Table 16.2).

\textit{KapA} (importin-\(\alpha\)), as in other organisms studied, is an essential gene (Araujo-Bazán, Osmani and Espeso unpublished). Importin \(\alpha\) mediates one of the most important nuclear import pathways and many of its cargoes have been defined. Importantly, importin \(\alpha\) also has functions in cell cycle regulation and protein degradation. Importin \(\alpha\) recognizes NLSs, which consist of basic residues (lysines and arginines) and are classified into two classes, bipartite and monopartite, as first defined in nucleoplasmin, and Simian Virus 40 large T antigen NLSs respectively. To date, three NLS sequences have been experimentally defined in \textit{A. nidulans}, the bipartite NLS present in PacC, and the monopartite NLS on Velvet (VeA), a protein involved in light-regulated asexual sporulation, and the NLS of \textit{StuA}, which is implicated in conidiophore morphogenesis.

Importin \(\beta\), as expected, is also an essential gene in \textit{A. nidulans}. The null allele, studied using heterokaryon rescue reveals that importin \(\beta\) is required during the early stages of spore germination and growth, perhaps due to a block on cell-cycle progression. In addition to its classic roles in nuclear transport, in higher eukaryotes importin \(\beta\) has been implicated in numerous other cellular functions including mitotic spindle formation, centrosome dynamics, nuclear membrane formation and NPC reassembly. In these nontransport roles, importin \(\beta\) is thought to bind to proteins and inhibit their function. These mitotic roles of importin \(\beta\) rely on the shift of the RanGTP gradient from across the NE in interphase to around DNA during mitosis. In this way, proteins bound to importin \(\beta\) are released from negative regulation near chromatin during the same mechanisms that release import cargoes within nuclei during the interphase transport cycle. Given that the RanGTP gradient likely shifts from across the NE to around mitotic chromatin during mitosis, this provides a mechanism for rapid response to changes in the supply of nutrients. In the NirA study it is suggested that nuclear export of this transcription factor is a regulatory checkpoint for nitrate induction in \textit{A. nidulans}.

16.3.2 Nuclear Export Pathway and Its Regulation During Transcriptional Control

The best characterized nuclear export pathway is mediated by Crm1. In \textit{A. nidulans} this pathway has been characterized recently. \textit{KapK} (also published as \textit{crmA} \textit{ }), is the gene coding for Crm1. KapK nuclear export activity has been defined on the basis of a mutant form sensitive to leptomycin B (LMB). This drug binds covalently to a cysteine residue and causes the specific loss of CRM1 activity inhibiting nuclear protein export. \textit{A. nidulans} is resistant to LMB because the wild-type form of KapK has a threonine in place of the cysteine targeted by LMB (aa525). By changing the Thr to a Cys within KapK, Todd et al. and Bernreiter et al. have generated strains of \textit{A. nidulans} that are now sensitive to LMB. Using these LMB-sensitive strains, it has been shown that AreA, the transcription factor mediating nitrogen regulation, and NirA, the binaural zinc cluster mediates nitrate activation, are excluded from nuclei in a KapK-dependent manner. The AreA work indicates that the quality of nitrogen source affects AreA nuclear accumulation by regulating its nuclear export. This provides a mechanism for rapid response to changes in the supply of nutrients. In the NirA study it is suggested that nuclear export of this transcription factor is a regulatory checkpoint for nitrate induction in \textit{A. nidulans}.

Crm1 recognizes nuclear export sequences (NES), which are leucine rich. The consensus NES, L-X\(_{2,3}\)L-X\(_{2,3}\)-L-X-L, is somewhat variable in that certain leucines can be changed to other nonpolar residues like isoleucine, valine, methionine, or phenylalanine. \textit{A. nidulans} KapK recognizes similar sequences as those proposed in other organisms and the first NES defined in \textit{A. nidulans} is that present in NirA. Not much is known about the mechanisms that govern KapK dynamics and the NPC, although it has been shown to interact with the nucleoporin An-Nup42/NplA, a peripheral nucleoporin that is dispersed from the NPC during mitosis. The role of this interaction is unknown but similar interactions have been described in higher eukaryotes. Finally, as expected, we have found that KapK is an essential protein in \textit{A. nidulans} (Araujo-Bazán, Osmani, and Espeso, unpublished).
16.3.3 Ntf2 and TAP/Nxt1 Transporters

There are additional, and more specialized, nuclear transport factors such as nuclear transport factor 2 (Ntf2). Ntf2 (termed NtfA in *A. nidulans*) is involved in the transport of RanGDP into the nucleus. The strong amino sequence conservation (58% identical, 77% conserved) between Ntf2 and *A. nidulans* Ntfa (Table 16.2) predicts that a similar pathway will exist in *A. nidulans*.

mRNA export does not rely on importin β-like karyopherins or the nuclear high RanGTP gradient but rather, relies on a heterodimer consisting of Mex67:Mtr2 in *S. cerevisiae*. We and others have searched for members of this mRNA export pathway. Although a clear ortholog of Mex67 exists, we have been unable to find a gene coding for a Mtr2p-like protein. Instead, we found a p15 ortholog (Table 16.2) perhaps suggesting similarity between the mechanism of export of mRNA molecules in *A. nidulans* to that of mammals.

16.4 Conclusions and Future Directions for Research

In filamentous fungi, the study of NPC proteins, soluble transporters and nuclear transport is in its infancy. However, it is clear that the unique biology of filamentous fungi provides an interesting arena in which regulation of nucleocytoplasmic transport is likely to play critical, and perhaps novel, roles. Early signs indicate that nuclear transport might well provide a fertile ground for insights into fungal biology, nuclear transport, and the structure and function of the NPC. Several areas are of obvious and immediate interest. First, the topic of cell-cycle specific regulation of the structure and function of the NPC during *A. nidulans* mitosis is positioned nicely for further experimentation. Mitotic regulation of the NPC is critical for entry into, and exit from, mitosis, and a full understanding of how the NPC is first disassembled and then reassembled, should shed light on how mitosis is regulated both in *A. nidulans* and higher eukaryotes. Secondly, the role of regulated nuclear location of transcription factors to transcriptionally control intermediary metabolism in *A. nidulans* in response to changing growth conditions is another area of high potential. Also of great interest is the role of light in regulating the nuclear localization of developmental regulators and subsequent effects on secondary metabolism. We anticipate that this first wave of fascinating studies are but the tip of a very large and exciting iceberg.

Acknowledgments

E.A.E. would like to acknowledge the nuclear transport work done by Javier Fernández-Martínez, Lidia Araújo-Bazán, Laura Cobeño, Oier Etxebeste, and Jairo Flores in his laboratory. E.A.E. also extends thanks to Ana Calvo and Joseph Strauss for productive collaborations. Thanks also to Miguel Peñalva for helpful discussions and suggestions regarding research areas and to Elena Reoyo for technical assistance. The work of E. A. Espeso has been supported by the Spanish Ministerio de Educacion y Ciencia, through grants BMC2003-00874 and BFU2006-04185. S.A.O would like to thank all members of his lab for their tireless enthusiasm and help, with particular thanks to Colin De Souza for critical reading of the manuscript. Work in the lab of S.A.O. is supported by the National Institutes of Health (GM042564).

References


Nuclear Pore Complex and Transport in Aspergillus nidulans

17

Sexual Development in Aspergillus nidulans

Dong-Min Han, Keon-Sang Chae, and Kap-Hoon Han

CONTENTS
17.1 Introduction ................................................................. 279
17.2 Morphology of Sexual Development in Aspergillus nidulans ................................................. 280
17.3 Environmental Factors Affecting Sexual Development ........................................................... 281
17.3.1 Light and VeA ................................................................. 282
17.3.2 Phytochrome Homolog, FphA ......................................................... 283
17.3.3 Endogenous Inducer—psi Factor ............................................................. 284
17.3.4 Reactive Oxygen Species and NADPH Oxidase (NoxA) ...................................................... 284
17.4 Genes Related in the Mating Process ....................................................................................... 285
17.4.1 Mating-Type Loci (matA and matB) .............................................................................. 285
17.4.2 Putative Pheromone Receptors, GprA and GprB ......................................................... 286
17.5 Signal Transduction Components for Sexual Development .................................................... 286
17.5.1 Putative GPCR, GprD ......................................................................................... 286
17.5.2 MAPKK Kinase, SteC ......................................................................................... 287
17.5.3 MAP Kinase, SakA/HogA .................................................................................... 288
17.6 Transcription Factors and Regulators Affecting Sexual Development ...................................... 288
17.6.1 Positive Sexual Regulator, NsdD ............................................................................. 288
17.6.2 Positive Sexual Regulator, SteA ............................................................................. 291
17.6.3 Transcription Factors Containing Zn(II)2Cys6 Domain, RosA and NosA ....................... 291
17.6.4 Development and Amino Acid Availability Control with cpcA and cpcB ...................... 292
17.6.5 COP9 Signalosome (CSN) and csnD/E ......................................................................... 293
17.7 Developmental Coordinators and Modifiers ........................................................................... 293
17.7.1 Spatio-Temporal Coordinators, StuA and MedA ...................................................... 293
17.7.2 Leucine Zipper Protein, DopA ................................................................................ 294
17.7.3 WD Repeat Protein, RcoA ................................................................................... 294
17.7.4 Genomics and Sexual Development ........................................................................ 294

References ................................................................................................................................. 295

17.1 Introduction

Many ascomycetes produce more than one type of spore derived mitotically or meiotically. The genus aspergilli consists of at least 186 species, and 72 species among them are known to produce ascospores as well as conidia, indicating that the genus aspergilli includes perfect fungi (Pontecorvo et al., 1953; Samson, 1994). Aspergillus nidulans is one of the genetic model organisms and is representative of the perfect Aspergillus species. The name Aspergillus came from aspergillum, which is an apparatus for sprinkling holy water, by observing the shape of the conidiophore. The species name nidulans means “nest-like,” which represents a closed fruiting body, cleistotheca. Taken together, the name Aspergillus nidulans means a fungus having asexual conidiophores and sexual cleistotheca. In addition, since it is a perfect fungus, A. nidulans has a sexual stage name, that is, teleomorph,
The Aspergilli

which is *Emericella nidulans*. Unlike many other heterothallic filamentous fungi including *Neurospora crassa*, which require a sexual partner having an opposite mating type, *A. nidulans* is homothallic, which means it does not need a different mating-type partner to produce fruiting bodies. In other words, a single conidium or ascospore of *A. nidulans* can complete both the asexual and sexual life cycle in normal conditions that are governed by a genetically programmed regulation system (Pontecorvo et al., 1954; Braus et al., 2002).

In this chapter, physiological and morphological studies on sexual development, which were reviewed previously in detail (Champe et al., 1994; Braus et al., 2002), and the relationship between fungal development and secondary metabolism, which also has been reviewed elsewhere (Calvo et al., 2002), are briefly described. In addition, several transcription factors as well as components of signal transduction pathways, are more concentrated on for presenting recent advances in the field of sexual development.

### 17.2 Morphology of Sexual Development in *Aspergillus nidulans*

*A. nidulans* undergoes sexual reproduction for generating eight sexual spores formed in one ascus that develops in a cleistothecium. This process requires sophisticated genetic regulation mechanisms, specifically to develop complex morphological structures as well as sexual spores. The sexual cycle of *A. nidulans* starts with fusion of ascogonial coils for making dikaryon, which normally occurred 50 h after germination (Fig. 17.1; Champe et al., 1994). In heterothallic fungi, such as *N. crassa* or *Podospora anserina*, this fusion event is usually tightly regulated by mating-type genes (Raju and Perkins, 1994; Coppin et al., 1997). However, the molecular event of dikaryotic hyphae formation in the homothallic fungus *A. nidulans* remains to be elucidated. This process follows the formation of primordium from ascogenous hyphae in the nest-like structure made of a number of thick-walled globose Hülle cells, which are formed from the tip of hyphae and regarded as nurse cells (Fig. 17.1; Braus et al., 2002; Hermann et al., 1983; Scherer and Fischer, 1998). As shown in Figure 17.1, the surrounding mycelia differentiate to form a network that is glued by an uncharacterized substance, cleistin, and finally matures to the cleistothecial wall (Champe et al., 1994). The primordium matures into a cleistothecium in which nuclear fusion and meiosis take place (Fig. 17.1; Sohn and Yoon, 2002). After two nuclei are fused in a crozier, forming a zygote, meiosis and the following two rounds of mitosis, result in producing an ascus that contains eight nuclei. Each nucleus undergoes an additional mitotic division, resulting in binucleate ascospores. A mature cleistothecium accumulates a pinkish red pigment called asperthecin and encloses up to 10,000 ascospores (Fig. 17.1).

Although the regulation mechanism of dikaryotic hyphae formation in *A. nidulans* is largely unknown, recent studies of genome research revealed that *A. nidulans* also contains putative mating-type genes and pheromone genes as well as pheromone receptors, indicating that there is a conserved mating process even in homothallic fungus. Similar observation was reported in *Sordaria macrospora* (Pöggeler and Kück, 2000; Pöggeler and Kück, 2001). The mating-type genes from *S. macrospora* can complement corresponding mutants of *P. anserina*, which is a heterothallic fungus (Jacobsen et al., 2002). Moreover, manipulation of the mating-type loci makes it possible to shift homothallic to heterothallic, or vice versa, in other filamentous fungi (Yun et al., 1999; Yun et al., 2000; Lee et al., 2003), implying that *A. nidulans* can possibly be converted to a heterothallic fungus by the manipulation of mating-type genes. However, currently the conversion of *A. nidulans* to heterothallic fungus has not been achieved successfully.

### 17.3 Environmental Factors Affecting Sexual Development

Environmental conditions and adaptation from the environment are important for most fungi to survive in nature. In *A. nidulans*, development as well as growth are largely affected by the surrounding environment. Han and his coworkers reported that various environmental factors mainly affect sexual development (Han et al., 2003a). The environmental factors that may affect growth of mycelia, such as nutritional status, culture conditions and several stresses, are also responsible for the developmental decision between sexual differentiation and asexual reproduction (Han et al., 2003a). Generally,
Sexual Development in Aspergillus nidulans

well-nourished conditions without any environmental stress favor sexual development, while stresses such as starvation of a carbon or nitrogen, oxidative stress, high osmolarity or intense visible light can inhibit fruiting body formation, promoting asexual development exclusively (Champe et al., 1994; Han et al., 1994a; Han et al., 2003a).

For example, sexual development of *A. nidulans* is greatly affected by amount and type of a carbon source (Table 17.1). When higher concentration of glucose is supplied, more cleistothecia are formed. On fermentable carbon sources including lactose, galactose, and glycerol, cleistothecia are developed more favorably than conidiophores, whereas on acetate, which is utilized only via aerobic respiration, no cleistothecia are developed (Han et al., 2003a). A low level of aerobic respiration caused by either plate-sealing or culture in a hypoxic chamber favors sexual development (Table 17.1; Han et al., 1990). On the other hand, either nitrogen or carbon starvation initiates asexual sporulation (Han et al., 1994a; Skromne et al., 1995). It has been suggested that the carbon source availability activates a Ras-dependent signaling mechanism with controlled rasA expression levels as the mediator of the programmed growth and the development process (Som and Kolaparthi, 1994; Osherov and May, 2000). RasA genetically interacts with DopA, which is a transcription factor for developmental regulation, mediating proper expression of several developmental transcriptions factors including *brlA, abaA*, and *steA* (Pascon and Miller, 2000).

In addition, high concentration of salt inhibits sexual development but enhances asexual development (Table 17.2). Not only *A. nidulans*, but also *A. oryzae* has higher asexual sporulation in the presence of salt (Song et al., 2001). Aeration directs the developmental program toward asexual reproduction, whereas high

---

**FIGURE 17.1** Ultrastructural morphology of sexual organs in *A. nidulans*. 1. Two separate hyphal branches participating in formation of an initial. Note the upper round core part and a loose intertwining of hypha. 2. DAPI staining of the same cell as that of Figure 1.1. Arrows indicate two hyphal partners. 3. Antheridia-like exterior electron dense hyphae and core cell (early ascogenous system) originated from the same parental hyphae (A) and (B). 4. Cleistothecial initial covered by a loosely woven hyphal net. Note that 2 or 3 hyphae from the same parental hypha joined for the formation of an initial. 9. Young cleistothecium with a few mature Hülle cells nearby. 14. Scanning electron micrograph showing an enlarged cleistothecium with that of Hülle cells. 28. Thick section of a near mature cleistothecium (170 μm in diameter) completely stripped of Hülle cells. Ascosarp contained many asci at near maturity (circled). 29. Mature ascus (ascocarp, 170 μm in diameter) containing ascospores suspended in remnant cytoplasm. Note that only two cell thick peridium with its intercellular space heavily impregnated with electron-dense material (arrows). CI, cleistothecial initial; PH, parental hypha; HC, Hülle cell; C, cleistothecium; As, ascospore; WL, peridial layer. Scale bars: 1, 2 = 10 μm; 3 = 2 μm; 4 = 2.5 μm; 9 = 10 μm; 14 = 50 μm; 28 = 30 μm; 29 = 2 μm. (From Shon and Yoon, *Mycobiology*, 30, 117, 2002.)
TABLE 17.1
Effect of Nutritional and Environmental Sources and Light on Development in FGSC4 (veA+)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Sources</th>
<th>Dark</th>
<th>Light (3000–3500 Lux)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AS^a</td>
<td>S^a</td>
</tr>
<tr>
<td>Carbon</td>
<td>Glucose (0.5%)</td>
<td>++</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Glucose (1%)</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>Glucose (3%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Lactose (2%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Acetate (2%)</td>
<td>+++</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Lactose + Acetate</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>Nitrogen sources</td>
<td>Sodium nitrate (0.2%)</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Ammonium tartrate (0.2%)</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Glutamine (0.1%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Casein hydrolysate (0.2%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Yeast extract (0.2%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td></td>
<td>Glycine (10 mM)</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>Plate sealing</td>
<td>–</td>
<td>+++</td>
</tr>
</tbody>
</table>

^Asexual development. The amount of conidia within a circled area of 1 cm diameter: –, <10^4; +, 10^4–10^5; ++, 10^5–5×10^6; ++++, >5×10^6.

^Sexual development. The amount of cleistothecia within cm^2 area: –, <1; +, 1–10; ++, 1–50; ++++, 50–100.

^Not determined.

Source: Adapted and modified from Han et al., 2003b.

CO₂ tension favors the sexual cycle (Axelrod et al., 1973). Even simple short-chain primary amines, such as propylamine, can promote asexual development but inhibit sexual development (Song et al., 2002).

17.3.1 Light and VeA

Illumination is an important environmental factor that controls induction of development (Mooney and Yager, 1990). Generally, A. nidulans predominantly produces conidiophores in the light and cleistothecia

TABLE 17.2
Effect of Environmental Factors on Development in FGSC4 (veA+) or VAJ1 (veAl)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Sources</th>
<th>veA+</th>
<th>veAl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AS^a</td>
<td>S^a</td>
</tr>
<tr>
<td>Inhibition of aerobic respiration</td>
<td>Plate sealing</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td>Sodium azide (1.0 mM)</td>
<td>2,4-dinitro-phenol (0.5 mM)</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Sodium oxalate (50 mM)</td>
<td>KCl (1%)</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td>NaCl (1%)</td>
<td>MgCl₂ (1%)</td>
<td>+++</td>
<td>–</td>
</tr>
<tr>
<td>MgSO₄ (1%)</td>
<td>Sorbitol (1%)</td>
<td>+++</td>
<td>+</td>
</tr>
</tbody>
</table>

^Asexual development. The amount of conidia within a circled area of 1 cm diameter: –, <10^4; +, 10^4–10^5; ++, 10^5–5×10^6; ++++, >5×10^6.

^Sexual development. The amount of cleistothecia within cm^2 area: –, <1; +, 1–10; ++, 1–50; ++++, 50–100.

Source: Adapted and modified from Han et al., 2003a.
Sexual Development in Aspergillus nidulans

The veA gene product is largely involved in the regulation of light-dependent developmental decision. However, most laboratory strains carry the mutated veAI allele, causing light-independent induction of conidiomata with concomitant increased levels of the asexual-specific transcription factor BrlA (Mooney and Yager, 1990). This is assumed to be caused by low functionality of the veAI product at 37°C, which is supported by the finding that the veAI mutation seems to be temperature-sensitive for cleistothecium formation (Champe et al., 1991). It means that the veAI mutant forms cleistothecia normally at 30°C, while it does not at 42°C. Such temperature-sensitivity has been explained by the finding that a veA-null mutant does not form cleistothecia even at 30°C (unpublished result). Furthermore, the finding that the mutation of G in the translation initiation codon of the wild type veA gene to T in the veAI mutant, resulting in the use of the Met codon at the 37th codon as a new translation initiation codon, and in the truncation of 36 amino acid at the N-terminus of the VeA protein is consistent with the temperature-sensitivity of the veAI mutation (Kim et al., 2002). VeA is thus hypothesized to repress initiation of asexual development and promotes sexual development (Mooney and Yager, 1990; Kim et al., 2002).

The veA gene has been known to control sexual development positively. The veAI mutation delays and reduces the development of sexual organs, which eventually results in the preferential development of asexual spores. The asexual development of veAI mutant is much less affected by various environmental factors, including nutrients, light, and temperature (Käfer, 1965; Han et al., 1994a).

The veA gene has an open reading frame (ORF) of a 573 amino acid polypeptide, which matched some hypothetical proteins whose functions were not clearly assigned yet. The veA transcript was present in the conidia and in mycelia cultured for up to 14 h and was expressed almost constitutively at an increased level throughout the asexual and sexual developmental processes, suggesting that it may act from a relatively early developmental stage and throughout developmental processes. Null mutants of the gene never formed sexual structures, even under conditions where sexual development preferentially occurs in wild types. Overexpressors of the gene formed larger numbers of sexual structures with a much reduced number of conidial heads than a control strain (a veAI mutant), even under conditions where wild-type strains form a few sexual structures but form conidial heads very well, such as in the presence of a salt at high concentrations. Furthermore, overexpressors could form Hülle cells and cleistothecia, even in a liquid culture. These results indicated that the veA gene is a positive regulator of sexual development and simultaneously, a negative one of asexual development. Later, it was investigated that the veA gene regulates several genes involved in production of sterigmatocystin (ST) and penicillin (Kato et al., 2003). The veA gene is necessary for the expression of the acvA gene for the transcription factor AcvA, which activates the gene cluster involved in the production of ST and the acvA gene, the key gene in the first step of penicillin biosynthesis. The veA gene also represses the transcription of the ipnA gene encoding isopenicillin synthetase, and regulates the brlA expression by modulating the ratio of αβ transcripts. All of these results clearly imply that the veA gene regulates sexual development positively, and asexual development and secondary metabolism, including ST and penicillin production, negatively.

17.3.2 Phytochrome Homolog, FphA

Phytochrome is a photoreceptor normally found in photosynthetic organisms such as plants and cyanobacteria. However, recent studies revealed that phytochromes are also found in heterotrophic bacteria and fungi (Kehoe and Grossman, 1996; Yeh et al., 1997). Many fungal genomes, as well as A. nidulans, include homologs of bacterial phytochrome, which have multifunctional domains such as the phytochrome domain, histidine kinase domain, and response-regulator domain. Fischer and his colleagues reported that A. nidulans phytochrome FphA (fungal phytochrome A) binds a biliverdin chromophore and represses sexual development under red-light conditions (Blumenstein et al., 2005). The ORF of the fpha gene contains one 56 bp intron and deduced polypeptide consists of 1280 amino acids with a calculated molecular mass of 140 kDa. An in vivo assembly experiment with coexpressed FphA and a heme oxygenase from Synechocystis sp. for providing biliverdin to make holo-FphA in E. coli resulted in a typical red, far-red photoreversible phytochrome signature (Gambetta and Lagarias, 2001).
Deletion of the fphA gene affects on sexual development under red light conditions. In the wild-type strain FGSC4, sexual development is repressed in the red light condition, whereas asexual development is highly repressed, and sexual development prefers dark conditions. However, in a ΔfphA mutant, the developmental pattern in the dark and red light conditions was not changed, whereas the wild-type strain largely underwent asexual development, indicating that the blocking of sexual development by red light (670 nm) in a ΔfphA mutant was largely eliminated. This phenotype of derepression was only detectable in veA+ background. If the fphA mutation was analyzed in the veA1 background, derepression of sexual development was not taken, suggesting that VeA acts downstream of FphA or the VeA and FphA place in parallel pathways (Blumenstein et al., 2005).

17.3.3 Endogenous Inducer—psi Factor

A precocious sexual inducer, psi, is an endogenous mixture of hydroxylinoleic acid moieties, which are structurally related to vertebrate eicosanoid hormones. These factors mediate balance of the asexual and sexual spore ratio in A. nidulans. Biochemical analysis showed that the A. nidulans psi factor is an endogenous hormone-like oxylipins which is composed of a mixture of hydroxylated oleic (18:1), linoleic (18:2), and linolenic (18:3) acid molecules termed psiβ, psiα, and psiγ, respectively (Calvo et al., 2001). The position and number of hydroxylations of the fatty acid backbone further identify the psi compounds as psiB (8-hydroxylinoleic acid), psiC (5,8-dihydroxylinoleic acid), and psiA (which contains a lactone ring of psiC at the 5’ position). The biosynthetic pathways of three psi factors are mediated by PpoA, PpoB, and PpoC (psi producing oxygenase; Tsitsigiannis et al., 2004a; Tsitsigiannis et al., 2005; Tsitsigiannis et al., 2004b).

The psiC is the most active compound that can be converted into its cyclic lactone, psiA. The 8-hydroxylinoleic acid psiB is presumably an intermediate (Champe and el-Zayat, 1989; Calvo et al., 1999; Calvo et al., 2001). The psiC factor acts as an inducer of sexual development (Champe et al., 1987). The psiC function also suggests that the activity in modification of membrane properties might contribute to the fusion of specialized hyphae during sexual development (Champe and el-Zayat, 1989).

17.3.4 Reactive Oxygen Species and NADPH Oxidase (NoxA)

Generation of reactive oxygen species (ROS) is an inevitable process in all aerobic organisms. Previously, it was known that ROS injures cellular components including DNA, protein, and lipids, and is also involved in the pathological process. The production of ROS, which is usually governed by NADPH oxidase (Nox) is important in response to pathogenic infection and to kill pathogens. Recently, new roles for Nox-generated ROS in eukaryotes, such as regulation of cell growth, oxygen sensing, growth factor signaling, and fertilization, have been discovered (Lambeth, 2004). In A. nidulans, the noxA gene encoding a novel microbial NADPH oxidase homologous to mammalian gp91phox has been reported as a regulator of sexual development (Lara-Ortiz et al., 2003). The noxA gene has three introns and the predicted polypeptide contains 550-amino acid residues with 64 kDa molecular mass. The amino acid sequence of this enzyme is well conserved in most filamentous fungi but not in Saccharomyces cerevisiae and Schizosaccharomyces pombe, suggesting that some fungi lost this gene during the evolutionary process (Lara-Ortiz et al., 2003).

Comparison of phenotypes in the wild type and in ΔnoxA strains showed no differences in hyphal growth and asexual development. However, sexual development in the ΔnoxA mutant is blocked, although Hülle cells are produced with a 12 h delay. Extensive dissection of a ΔnoxA mutant in brlA1 background revealed that the mutant produces primordia, but failed to generate mature cleistothecia, implying that the mutant can initiate sexual development but is blocked at the primordial stage. Kawasaki et al. (2002) reported that a mitogen-activated protein kinase (MAPK), named SakA in A. nidulans, regulates sexual development negatively. Deletion of the sakA gene resulted in overproduction of cleistothecia (see later; Kawasaki et al., 2002). But the double mutant of ΔsakA ΔnoxA was unable to develop cleistothecia, indicating that the noxA gene is required for normal cleistothecia formation.
Detection of ROS formation in the wild type and a ΔnoxA mutant indicated that young primordia and Hülle cells produce superoxide, H$_2$O$_2$, and other ROS, which are essential for cleistothecium and ascospore formation in the wild-type strain but not in the ΔnoxA mutant. This work contributed to point out that NoxA-dependent ROS generation can be a regulated and self-inflicted oxidative stress that is essential to regulate differentiation in A. nidulans (Lara-Ortiz et al., 2003).

17.4 Genes Related in the Mating Process

17.4.1 Mating-Type Loci (matA and matB)

Mating-type loci are responsible for initiating sexual development and mating in heterothallic fungi. However, in homothallic fungi, studies for the mating-type loci have not been intensively analyzed. Recent studies with genome analyses revealed that many homothallic fungi contain functional mating-type genes and are important for maintaining homothallic characteristics. In Cocchliobulous sp. and Fusarium sp., deletion of either the mat1 or mat2 gene made the strains heterothallic, producing fruiting bodies exclusively by outcrosses (Yun et al., 1999; Lee et al., 2003). In A. nidulans, the genome possesses both mating-type genes (Varga, 2003). Miller et al. (2005) have reported that two mating-type loci, namely matA and matB, are conserved in the A. nidulans genome. The matA and matB genes encode an HMG box protein and an alpha box protein, which are homologs of the products of the MAT-2 and general MAT-1-1 genes, respectively (Miller et al., 2005). Unlike other fungal mating-type loci, these two mating-type genes are not located on the same chromosome and not genetically linked to each other, suggesting that the homothallism of A. nidulans may not be caused by fusion of mating-type genes from its heterothallic ancestor. Deletion of the matA, matB, or matA/B gene(s) did not affect mycelial growth and asexual development. However, a matA deletion resulted in severely delayed and diminished production of fruiting bodies and a matB deletion strain could not undergo meiosis, resulting in cleistothecium lysis, although the matB deletion strain produced normal cleistothecia (Miller et al., 2005). These results indicate that, even in homothallic fungus, mating-type genes are required for normal sexual differentiation. Furthermore, A. fumigatus, which is an opportunistic human pathogen and has no sexual cycle, also possesses mating-type loci and other genetic components required for sexual development, suggesting that A. fumigatus may have the genetic machinery to mate and to develop sexual fruiting bodies (Paoletti et al., 2005).

17.4.2 Putative Pheromone Receptors, GprA and GprB

Genome analysis revealed that two putative pheromone receptors similar to yeast pheromone receptors Ste2p and Ste3p, which are G protein coupled receptors (GPCRs), exist in the A. nidulans genome. The mating pheromone signaling pathway is one of the well-known G protein signaling processes in budding yeast. In S. cerevisiae, Ste2p or Ste3p is bound by the opposite mating type’s pheromones such as α- or α-factor, respectively. The pheromone-bound GPCR stimulates a subsequent Gα subunit and Gβ/γ subunits, Gpa1p and Ste4p/Ste18p, respectively, for activating the Fus3p/Kss1p MAPK cascade and eventually up-regulates the Ste12p transcription factor which is responsible for regulating the mating process-related gene expression. Ste11p and Ste12p homologs in A. nidulans, SteC, and SteA, respectively, have been characterized and it has been shown that the SteC and SteA are required for sexual development (Vallim et al., 2000; Wei et al., 2003; see later).

Although sexual development of A. nidulans does not necessarily require the involvement of heterothallic mating processes, it has been shown that pheromone-like molecules including the psi factor partly affect the decision and balance of sexual development (Champe and el-Zayat, 1989; Tsitsigiannis et al., 2004a). Yu and his colleagues cloned and characterized yeast STE2 and STE3 homologs, gprA and gprB, respectively (Seo et al., 2004). Dyer’s group also deposited such identical genes as preB and preA, respectively (Dyer et al., 2003). Both genes produced two different length transcripts and the smaller transcripts are accumulated during sexual developmental process. Deletion of the gprA, gprB, or gprA/B gene(s) affected self-fertilized fruiting body formation but not outcrossing or heterothallic sexual development,
indicating that GprA and GprB play an important role in self-fertilization. This sterile phenotype could be partially rescued by an increment of nsdD expression. Overexpression of the nsdD gene in the ΔgprA/B background resulted in the production of immature and fragile cleistothecia, containing almost no viable ascospores. This result indicated that NsdD may function downstream of (or parallel to) GprA/B and the activation of NsdD is necessary but not sufficient to cleistothecia maturation and ascospore generation (Seo et al., 2004).

17.5 Signal Transduction Components for Sexual Development

G protein signaling in eukaryotic cell systems plays a pivotal role in fundamental cellular processes including growth and development. In A. nidulans Gα subunit FadA and its antagonistic regulator of G protein signaling (RGS) protein, FlbA, of the heterotrimeric G protein system, govern filamentous growth and asexual development. Three Gα proteins (FadA, GanA, and GanB), one Gβ (SfaD), and one Gγ (GpgA) were found in A. nidulans genome and, recently, these genes have been characterized (Yu et al., 1996; Rosen et al., 1999; Chang et al., 2004; Han et al., 2004; Seo et al., 2005; Lafon et al., 2005; Yu, 2006). Besides G proteins, not many genes related to the signal transduction pathway for sexual development are characterized. In this section, a putative GPCR, and MAPK components are briefly discussed. Figure 17.2 shows a proposed genetic model for a signaling pathway and transcription factors on the sexual development process.

17.5.1 Putative GPCR, GprD

GPCRs are frontiers of the heterotrimeric G protein system that accept extracellular signals and transfer to the appropriate internal response system. Han et al. (2004) reported nine putative GPCRs found in the genome of A. nidulans. Besides pheromone receptors, they were novel GPCRs in true filamentous fungi. Most of these putative GPCRs were conserved in various fungi, and recent studies identified that there are...
homologs of these GPCRs in *N. crassa* and *Cryptococcus neoformans* (Han et al., 2004; Li and Borkovich, 2006; Krystofova and Borkovich, 2006; Xue et al., 2006). Among them, GprD showed a close similarity with yeast Gpr1p, which is related to glucose sensing. The *gprD* gene has an ORF encoding 427 amino acids polypeptide with seven transmembrane domains. Deletion of the *gprD* gene caused a pleiotropic phenotype. Hyphal growth was extremely restricted in a Δ*gprD* mutant with undifferentiated aerial hyphae. GprD is also required for germination of conidia. Deletion of the *gprD* gene resulted in, at least, 3 h delay of conidial germination, suggesting that the *gprD* gene is important for hyphal growth as well as spore germination (Han et al., 2004). Furthermore, within two weeks of incubation at 37°C, the colony of a Δ*gprD* mutant was fully covered with Hülle cells and mature cleistothecia without any conidiophore. Taken together, these results indicate that GprD is necessary for activating growth and inhibiting sexual development. Inactivation of GprD resulted in derepression of *nsdD* expression, suggesting that the derepressed *nsdD* expression is important for uncontrolled sexual development of the Δ*gprD* mutant. Loss-of-function mutation in the *nsdD* or *veA* gene in the Δ*gprD* background suppresses the extreme phenotype caused by *gprD* mutation, supporting the hypothesis that the primary role of the *gprD* gene is to negatively control sexual development (Han et al., 2004).

### 17.5.2 MAPKK Kinase, SteC

The MAPK cascade is one of the central signal transduction pathways conserved in all eukaryotic systems. The kinases in the MAPK pathway are sequentially activated by phosphorylation after induction of extracellular mitogens. In yeast, Ste11p is a MAP kinase kinase kinase (MAPKKK or MAPKK kinase), which is involved in mating, pseudohyphal growth, and osmoregulation processes and enables to activate a downstream MAP kinase kinase (MAPKK or MAPK kinase) by phosphorylation. Wei et al. (2003) identified *A. nidulans* homolog of yeast Ste11p, SteC. Since the *steA* and *steB* genes were already reported, they used the name of yeast *STE11* homolog as *steC* (Vallim et al., 2000; Han and Prade, 2002). Although *steB* was also reported as a *STE11* homolog, it was not clear whether the *steB* and *steC* are identical because *steB* was identified by an expressed sequence tag (EST) sequence and the full-length DNA sequence was not available (Han and Prade, 2002). The *steC* gene contains an ORF of 886 amino acids with conserved a sterile alpha domain (SAM), which is a protein interaction domain, and a catalytic domain in its N-terminus and C-terminus, respectively. A complementation test revealed that the SAM domain is indispensable for normal function.

Deletion of *steC* resulted in reduced hyphal growth with curled and branched hyphae (Wei et al., 2003). It could be expected that the deletion mutants affect sexual development because yeast homolog *STE11* is involved in the mating process. Not only sexual development but also conidiophore development of the deletion mutant was altered. Very large conidia and secondary conidiophores, which came out from the vesicle of primary conidiophore, were observed with the rare frequency (2%) in the Δ*steC* mutant. Furthermore, a Δ*steC* mutant showed a sterile phenotype, that there is no cleistothecium in the mutant, and it could not be crossed with the conventional hyphal fusion method. Rather, the protoplast fusion method provided heterokaryotic mycelia. Heterokaryon from a Δ*steC* mutant and strains having a *steC* wild-type copy showed the typical heterokaryotic appearances with normal conidiophore and cleistothecium development. However, although small nest-like structures and Hülle cells were observed, the heterokaryon with two *steC* deletion mutants was sterile, indicating that the *steC* gene is required for heterokaryon formation and cleistothecium production (Wei et al., 2003).

Despite the important role of the *steC* gene in sexual development, the *steC* transcript was more abundant during the asexual developmental process than sexual developmental process. In addition, the GFP-tagged *steC* expression analysis also revealed that the GFP fluorescence is detected in metulae, phialides, and young conidiophores but not in stalks of conidiophores and sexual organs (Wei et al., 2003). A western blot analysis with antiphosphoantibodies of p44/42, SAPK/JNK, and p38 showed that the *steC* gene can activate at least two downstream MAPKs, a p44/42 homolog and a SAPK/JNK homolog, in *A. nidulans*. A deletion strain of the *steA* gene showed no activation or the decreased amount of the activated MAPK level during asexual development or oxidative stress condition. Although direct evidence about downstream MAPK(s) regulated by SteC has not been provided yet, it is quite clear that the MAPK cascade is involved in sexual development as well as normal conidiophore formation.
17.5.3 MAP Kinase, SakA/HogA

The A. nidulans sakA gene encodes a member of the stress MAPK family and is identical to the hogA gene that was identified as a homolog of yeast HOG1. Han and Prade (2002) reported that the hogA gene in A. nidulans plays a crucial role for regulating the osmotic stress response. However, the relationship between osmoregulation and sexual development is not included yet. Later, Kawasaki et al. (2002) provided important information of the role of the hogA/sakA gene in oxidative stress and sexual development. The sakA gene was also identified with screening of the cDNA database with yeast HOG1/SPC1 sequences, and to encode a 379 amino acid protein of which the amino acid sequence is similar to those of the stress-activated MAPK family (SAPK), having the conserved TGY phosphorylation site. Immunoblotting with antiphosphoantibody against p38 revealed that SakA is phosphorylated immediately when the fungus was subjected to an oxidative stress as well as an osmotic stress, indicating that SakA is a functional SAPK activated by various external stresses.

Deletion of the sakA gene resulted in accumulation of a reddish pigment into the medium, but did not affect vegetative growth. It also did not show phenotypic differences in hyperosmolarity conditions at 37°C. A similar observation was reported by Han and Prade (2002), while the growth of hogA deletion mutants was severely restricted when they were incubated at 30°C (Han and Prade, 2002). In a ΔsakA mutant, more Hülle cells were developed when compared to the wild type. After sexual development induction by oxygen limitation, cleistothecial development was accelerated, at least, by 24 h, suggesting that SakA represses sexual development or that other signaling pathways are derepressed by inactivation of the sakA gene. This process might be related with the steA gene function because the ΔsakA ΔsteA double mutant was not able to produce cleistothecia. In addition to the effect on sexual development, the conidia of a ΔsakA mutant lost their viability faster than a wild type, indicating that the sakA gene plays an important role in the spore viability as well as the stress resistance (Kawasaki et al., 2002).

17.6 Transcription Factors and Regulators Affecting Sexual Development

There are a few transcription factors that are known as regulators of sexual development. The GATA type transcription factor NsdD was isolated by the classical mutagenesis and the complementation analysis. Another important transcription factor SteA has a homeodomain as well as a C2H2 zinc finger. In addition, the RosA and the NosA contain a fungal-specific Zn(II)$_2$Cys$_6$ binuclear cluster, which plays an important and specific role in developmental process (Han et al., 2001; Vallim et al., 2000; Vienken et al., 2005; Vienken and Fischer, 2006). The summary of the information of the transcription factors as well as other genes involved in sexual development is listed in Table 17.3.

17.6.1 Positive Sexual Regulator, NsdD

Previously, research of sexual development had many difficulties because asexual spores were predominantly formed, and sexual organs, including primordia and cleistothecia, which are usually formed under conidiophores, were hard to observe. One breakthrough of sexual development research was that Han et al. (1994b) identified several mutants with forward genetics using the developmental characteristics of a wild-type strain in response to environmental stresses especially in the restriction of aeration or in a hypoxic stress. To isolate genes that positively control sexual development, mutants that failed to produce any sexual reproductive organs even in the hypoxic condition were screened. After isolation of several NSD (never in sexual development) mutants, Han and coworkers identified four complementation groups, nsdA, nsdB, nsdC, and nsdD (Han et al., 1994b; Han et al., 1998). Among them, the nsdD gene was isolated and identified as a positive regulator of sexual development (Chae et al., 1995; Han et al., 2001).

The isolated nsdD gene encodes a GATA-type transcription factor carrying the type IVb zinc finger DNA-binding domain at its C-terminus. The nsdD gene was expressed during the vegetative growth, and the expression level increased as sexual development proceeded. Deletion of the nsdD gene resulted in no cleistothecia formation, even under the conditions that preferentially promote sexual development, indicating that the nsdD gene is necessary for sexual development. In contrast, when the nsdD gene was
TABLE 17.3
Comprehensive Genetic and Genomic Information of the Genes Involved in Sexual Development

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of AA</th>
<th>Accession No.</th>
<th>Locus Tag</th>
<th>Domain(s)</th>
<th>Yeast Homolog</th>
<th>A. fumigatus Homolog</th>
<th>A. oryzae Homolog</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>veA</td>
<td>573</td>
<td>AAD42946</td>
<td>AN1052.3</td>
<td>?</td>
<td>?</td>
<td>Afu1g12490</td>
<td>AO090001000237</td>
<td>Light response, velvet phenotype, positive regulator of sexual development</td>
</tr>
<tr>
<td>fphA</td>
<td>1280</td>
<td>CAI30283</td>
<td>AN9008.3</td>
<td>P2, GAF, PHY, HKD, RRD</td>
<td>?</td>
<td>Afu4g02900</td>
<td>AO090001000178</td>
<td>Fungal phytochrome, repressor of sexual development under red light</td>
</tr>
<tr>
<td>ppoA</td>
<td>1081</td>
<td>AAR88626</td>
<td>AN1967.3</td>
<td>Haem peroxidase, cytochrome P450</td>
<td>?</td>
<td>Afu4g10770</td>
<td>AO0900030001138</td>
<td>psi factor production, balancing asexual and sexual development</td>
</tr>
<tr>
<td>ppoB</td>
<td>1019</td>
<td>AAX35769</td>
<td>AN6320.3</td>
<td>Haem peroxidase, cytochrome P450</td>
<td>?</td>
<td>Afu4g00180</td>
<td>AO090010000662</td>
<td>psi factor production, balancing asexual and sexual development</td>
</tr>
<tr>
<td>ppoC</td>
<td>1117</td>
<td>AAT36614</td>
<td>AN5028.3</td>
<td>Haem peroxidase, cytochrome P450</td>
<td>?</td>
<td>Afu3g12120</td>
<td>AO090003000772</td>
<td>psi factor production, balancing asexual and sexual development</td>
</tr>
<tr>
<td>noxA</td>
<td>550</td>
<td>AAN75017</td>
<td>AN5457.3</td>
<td>NADPH oxidase</td>
<td>?</td>
<td>Afu6g13350</td>
<td>AO090003000460</td>
<td>ROS generation, essential for sexual development</td>
</tr>
<tr>
<td>matA/mat2</td>
<td>318</td>
<td>AAP92161</td>
<td>AN4734.3</td>
<td>HMG-box</td>
<td>MAT A-2</td>
<td>Afu3g06170</td>
<td>AO090003001130</td>
<td>Mating type locus with HMG-box</td>
</tr>
<tr>
<td>matB/mat1</td>
<td>361</td>
<td>AAO01665</td>
<td>AN2755.3</td>
<td>Alpha box</td>
<td>MAT alpha-1</td>
<td>AAX83123</td>
<td>AO090020000089</td>
<td>Mating type locus with alpha-box</td>
</tr>
<tr>
<td>gprA/preB</td>
<td>377</td>
<td>DAA01796</td>
<td>AN2520.3</td>
<td>7 transmembrane</td>
<td>STE2</td>
<td>Afu3g14330</td>
<td>AO0900701000605</td>
<td>Similar to alpha-factor pheromone receptor</td>
</tr>
<tr>
<td>gprB/preA</td>
<td>349</td>
<td>DAA01795</td>
<td>AN7743.3</td>
<td>7 transmembrane</td>
<td>STE3</td>
<td>Afu5g07880</td>
<td>AO0900701000699</td>
<td>Similar to a-factor pheromone receptor</td>
</tr>
<tr>
<td>gprD</td>
<td>427</td>
<td>EAA63555</td>
<td>AN3387.3</td>
<td>7 transmembrane</td>
<td>GPR1</td>
<td>Afu2g12640</td>
<td>AO090026000360</td>
<td>Putative GPCR, repressor of sexual development</td>
</tr>
<tr>
<td>steC</td>
<td>886</td>
<td>CAD44493</td>
<td>AN2269.3</td>
<td>SAM/ST protein kinase</td>
<td>STE11</td>
<td>Afu5g06420</td>
<td>AO090009000610</td>
<td>MAPKKK, positive regulator of sexual development</td>
</tr>
<tr>
<td>sakA/hogA</td>
<td>379</td>
<td>AAF97243</td>
<td>AN1017.3</td>
<td>ST protein kinase</td>
<td>HOG1</td>
<td>Afu1g12940</td>
<td>AA97243</td>
<td>MAPK, sexual development repressor</td>
</tr>
<tr>
<td>nsdD</td>
<td>461</td>
<td>AAB16914</td>
<td>AN3152.3</td>
<td>GATA type</td>
<td>?</td>
<td>Afu3g13870</td>
<td>AO090012000768</td>
<td>GATA type transcription factor, positive regulator of sexual development</td>
</tr>
<tr>
<td>steA</td>
<td>692</td>
<td>AAC31206</td>
<td>AN2290.3</td>
<td>Homeodomain/ C2H2 Zn-finger</td>
<td>STE12</td>
<td>Afu5g06190</td>
<td>AO090009000638</td>
<td>Homeodomain-C2H2 transcription factor, required for sexual development</td>
</tr>
<tr>
<td>rosA</td>
<td>713</td>
<td>CAD58393</td>
<td>AN5170.3</td>
<td>Zn(II)2Cys6</td>
<td>?</td>
<td>Afu6g07010</td>
<td>AO090003001259</td>
<td>Transcription factor, repressor of sexual development</td>
</tr>
</tbody>
</table>

continued
### TABLE 17.3 (continued)
Comprehensive Genetic and Genomic Information of the Genes Involved in Sexual Development

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of AA</th>
<th>Accession No.</th>
<th>Locus Tag</th>
<th>Domain(s)</th>
<th>Yeast Homolog</th>
<th><em>A. fumigatus</em> Homolog</th>
<th><em>A. oryzae</em> Homolog</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>nosA</td>
<td>675</td>
<td>CAJ76908</td>
<td>AN1848.3</td>
<td>Zn(II)2Cys6</td>
<td>?</td>
<td>Afu4g09710</td>
<td>AO090003001259</td>
<td>Transcription factor, repressor of sexual development</td>
</tr>
<tr>
<td>cpcA</td>
<td>265</td>
<td>AAL09315</td>
<td>AN3675.3</td>
<td>bZIP</td>
<td>GCN4</td>
<td>Afu4g12470</td>
<td>AO090009000459</td>
<td>c-Jun homolog, regulate a control point for sexual development</td>
</tr>
<tr>
<td>cpcB</td>
<td>316</td>
<td>AAF98065</td>
<td>AN4163.3</td>
<td>WD40</td>
<td>CPC2</td>
<td>Afu4g13170</td>
<td>AO090009000264</td>
<td>RACK1 homolog, regulate a control point for sexual development</td>
</tr>
<tr>
<td>csnD</td>
<td>408</td>
<td>AAK14055</td>
<td>AN1539.3</td>
<td>PCI</td>
<td>CSN4</td>
<td>Afu8g05500</td>
<td>AO090005000595</td>
<td>Component of COP9 signalosome, positive regulator of sexual development</td>
</tr>
<tr>
<td>csnE</td>
<td>335</td>
<td>AAM95164</td>
<td>AN2129.3</td>
<td>MPN</td>
<td>CSN5</td>
<td>Afu2g16250</td>
<td>AO090102000238</td>
<td>Component of COP9 signalosome, positive regulator of sexual development</td>
</tr>
<tr>
<td>stuA</td>
<td>622</td>
<td>AA33325</td>
<td>AN5836.3</td>
<td>APSES</td>
<td>PHD1</td>
<td>Afu2g07900</td>
<td>AO090011000905</td>
<td>Transcription factor, coordination of sexual and asexual development</td>
</tr>
<tr>
<td>medA</td>
<td>658</td>
<td>AAC31205</td>
<td>AN6230.3</td>
<td>?</td>
<td>?</td>
<td>Afu2g13260</td>
<td>AO090026000285</td>
<td>Transcription factor, developmental modifier</td>
</tr>
<tr>
<td>dopA</td>
<td>1858</td>
<td>AAD28280</td>
<td>AN2094.3</td>
<td>Leucine zipper</td>
<td>DOP1</td>
<td>Afu2g05020</td>
<td>AO090003000304</td>
<td>Leucine zipper protein, control cellular morphogenesis</td>
</tr>
<tr>
<td>rcoA</td>
<td>619</td>
<td>AAG28504</td>
<td>AN6505.3</td>
<td>WD repeat</td>
<td>TUP1</td>
<td>Afu6g05150</td>
<td>AO090701000021</td>
<td>Pleiotropic effect on growth, asexual and sexual development</td>
</tr>
</tbody>
</table>
overexpressed, the number of cleistothecia was dramatically increased on a solid medium and also a sexual-specific organ (Hülle cells) was formed even in a submerged culture, where sexual development is completely blocked in wild types. These results indicated that the \textit{nsdD} gene functions in activating sexual development of \textit{A. nidulans} (Han et al., 2001). In several allelic mutants of the \textit{nsdD} gene that resulted in the early chain termination and lacked the zinc finger motif, the accumulation of the \textit{nsdD} transcript was greatly increased. And when the \textit{nsdD} gene was overexpressed by the \textit{niiA} promoter, the transcription under its own promoter was reduced. The mRNA levels in the strains having multiple copies of the \textit{nsdD} gene were not increased although the cleistothecial formation was dramatically increased (Han et al., 2003b). These results suggest that the expression of the \textit{nsdD} gene is negatively autoregulated and the NsdD protein in the cell is maintained within a certain level. The self-regulation or autoregulation of \textit{nsdD} expression is probably carried out by binding its own promoter of the \textit{nsdD} gene product, NsdD (Han et al., 2003b). Indeed, there are GATA-binding domains at the promoter region of the \textit{nsdD} gene, supporting the hypothesis that NsdD can bind the promoter region of the \textit{nsdD} gene. However, currently no direct evidence for the interaction is reported.

When the \textit{nsdD} gene was overexpressed, cleistothecia were formed in an excess amount even in the presence of 0.6 M KCl that inhibited sexual development in a wild type. A Northern blot analysis revealed that the expression of the \textit{nsdD} gene was repressed by 0.6 M KCl. These results strongly suggest that the inhibition of sexual development by salts was carried out via the \textit{nsdD} gene-mediated regulatory network (Han et al., 2003b).

### 17.6.2 Positive Sexual Regulator, SteA

The \textit{A. nidulans} SteA is a homolog of \textit{S. cerevisiae} Ste12p, which is a homeodomain protein governed by the MAPK signal transduction pathway. In \textit{S. cerevisiae}, Ste12p plays an important role in regulating cellular morphogenesis and the mating process, especially in the pseudohyphal growth and the karyogamy. Vallim et al. (2000) identified a yeast \textit{STE12} homolog, \textit{steA}, by the degenerate PCR method. The \textit{steA} gene contains an ORF, which encodes a 692 amino acid polypeptide, containing a conserved homeodomain in its N-terminus. However, unlike yeast \textit{Ste12p}, \textit{A. nidulans} SteA contains two tandem C2H2 Zn-finger domains on its C-terminus, which is very similar to the Ste12p of \textit{C. neoformans} and the StlA in \textit{Penicillium marneffei} (Yue et al., 1999; Borneman et al., 2001).

Deletion of the \textit{steA} gene resulted in normal vegetative hyphal growth, radial growth rate, and conidiophore morphology. On the contrary, although the \textit{steA} strains were subjected to induce sexual development by restriction of oxygen supply, they were sterile, without any cleistothecia and ascospores, while Hülle cells were still observed after four days of the induction. This phenotypic defect was complemented with the \textit{P. marneffei} \textit{stlA} gene, although \textit{P. marneffei} does not undergo sexual development. Overexpression of the \textit{steA} gene under the \textit{alcA} promoter showed that the mutant formed conidiophores with a 2–3 day delay and generated an irregular and abnormal morphology similar to the ascogenous tissue. Taken together, these results indicated that the \textit{steA} gene is a positive regulator of sexual development (Vallim et al., 2000).

The \textit{brlA} gene expression is required for regulating \textit{steA} expression because \textit{steA} expression was derepressed in \textit{brlA} mutant strains. However, \textit{steA} expression was not affected by the absence of either the \textit{stuA} or \textit{medA} gene, indicating that the \textit{steA} gene is located in the independent sexual process pathway or upstream of the \textit{stuA} and \textit{medA} genes. The \textit{medA} expression is derepressed in a \textit{steA} mutant but \textit{stuA} expression is not changed in the mutant, suggesting that the \textit{steA} gene is genetically located upstream of the \textit{medA} gene but is separated from the \textit{stuA} gene in regulating sexual development (Vallim et al., 2000).

### 17.6.3 Transcription Factors Containing Zn(II)$_2$Cys$_6$ Domain, RosA and NosA

RosA is a homolog of Pro1 of \textit{S. macrospora}, which is a Zn(II)$_2$Cys$_6$ transcription factor controlling perithecia development (Masloff et al., 1999). Fischer and coworkers identified the \textit{rosA} (repressor of sexual development) gene by screening the \textit{A. nidulans} genome database and is characterized (Vienken et al., 2005). The \textit{rosA} gene encodes a Zn(II)$_2$Cys$_6$ domain-contained 713 amino acid polypeptide in the ORF
interrupted by two short introns. Unlike many conserved transcription factors throughout filamentous fungi, a RosA ortholog was not found in yeasts and some basidiomycetes.

Deletion of the nosA gene in the veA+ background caused an increment of cleistothecia production in the dark conditions, suggesting that the nosA gene at least partially represses sexual development in A. nidulans in a veA-dependent manner. Moreover, in the low concentration of glucose or the presence of 0.6 M KCl, where sexual development is highly repressed, development of cleistothecia was allowed, especially in the dark condition. The deletion strain also produced Hülle cells even in a submerged culture. These phenotypes are very similar to overexpression of the positive sexual regulator, nsdD or veA (Han et al., 2001; Kim et al., 2002). A Northern blot analysis showed that deletion of the nosA gene up-regulated the nsdD, veA, and even stuA genes, in a submerged culture (Vienken et al., 2005). On the other hand, when the nosA gene was overexpressed, massive aerial hyphae were generated without any sexual or asexual development. Although this phenotype is quite similar to the fadA-dominant activating mutant, a genetic analysis revealed that the FadA-signaling is independent of the RosA. The expression of the nosA gene is culminated at 12 h of asexual induction transferred from a submerged culture to solid medium. However, when the fungus meets carbon starvation, the nosA gene is highly expressed after 3 h of the starvation.

On the other hand, a NosA (number of sexual spores) was also isolated by ortholog screening with the Prol from S. macrospora (Masloff et al., 1999; Vienken and Fischer, 2006). The predicted NosA protein is composed of 675 amino acids, having 44% of the sequence identity with the Prol of S. macrospora. The similarity between NosA and RosA is about 43%. A genome analysis revealed that only aspergilli, including Aspergillus oryzae and A. fumigatus, contain two Prol homologs while other fungi have just one. The expression of the nosA gene is up-regulated at the late stage of asexual development (16–26 h). Also, similar to the rosA gene, glucose starvation caused accumulation of nosA mRNA within 3 h. However, very low steady-state level of the nosA transcript was detected during the sexual-induction stage (Vienken and Fischer, 2006).

When the NosA activity was inactivated by deletion of the nosA ORF in the veA1 background, growth and asexual development were normal but no sexual development was proceeded both in the normal and favored condition, such as at the increased CO2 concentration and in a dark incubation. In the veA+ background, unlike the wild-type FGSC4, a ΔnosA mutant could not undergo sexual development. Rather, sexual development of the mutant was blocked at the primordial stage, indicating that the nosA gene is necessary for completion of sexual development but not for initiation.

Genetic analysis with the nosA and nsdD gene revealed that the nosA deletion strain in the constitutively induced nsdD background could not complete sexual development and blocked at the primordia stage, suggesting that the nosA gene is in the downstream of the nsdD gene in the same pathway (or parallel to nsdD). Genetic interaction analysis between the nosA and rosA gene was also performed. Northern blot analysis showed that nosA expression in a ΔrosA strain was up-regulated when compared to a wild type, indicating that the rosA gene represses the expression of the nosA gene (Vienken and Fischer, 2006).

In general, the Zn(II)$_2$Cys$_6$ binuclear cluster transcription factor is fungal specific. Although the genome sequence of A. nidulans revealed that 123 potential Zn(II)$_2$Cys$_6$ binuclear cluster proteins exist in the genome, only few genes have been characterized so far (Vienken et al., 2005). RosA and NosA revealed that these are important transcription factors involved in repression and activation of sexual development in A. nidulans. Lee et al. (2005) also reported development-related Zn(II)$_2$Cys$_6$ protein OefC by screening of the overexpression library, which is constructed by insertion of genomic DNA fragments into the niiA(p)-contained autonomously replicated overexpression vector (AMA-niiA(p); Lee et al., 2005). Further analysis of the functions of fungal specific development-related transcription factors will bring us more important information for understanding A. nidulans development.

### 17.6.4 Development and Amino Acid Availability Control with cpcA and cpcB

The crosspathway control is a general amino acid control in filamentous fungi. The CpcA in A. nidulans is an ortholog of a yeast transcription activator Gcn4p and plays a pivotal role in response to amino acid starvation. In A. nidulans, amino acid starvation does not affect growth or asexual development. However, mutants having a defect in their amino acid biosynthetic gene, such as the trpC or argB gene, are unable to produce cleistothecia (Eckert et al., 1999; Selupi-Crescenzi et al., 1983), suggesting that the arrest of
Sexual development can be caused by activating crosspathway control via amino acid starvation signals (Hoffmann et al., 2000).

Braus and his colleagues characterized the relationship between amino acid starvation and sexual development at the molecular level (Hoffmann et al., 2000). Amino acid starvation could be induced by adding the histidine analog 3-amino-1,2,4-triazole (3-AT). When *A. nidulans* was subjected on medium containing 3-AT, sexual development was blocked at the primordial stage. Furthermore, transferring of the primordia which were grown under amino acid starvation to the normal conditions resulted in the completion of sexual development (Hoffmann et al., 2000), indicating that the activation of the crosspathway control is responsible for the blocking of sexual development. Indeed, with overexpression of the *cpcA* gene under the inducible promoter, *alcA*(p), blocked sexual development at the primordial stage. This blockage was reversible, and turning off the activation of the *cpcA* gene led to the release of the blocking.

On the contrary, a RACK1 homolog in yeast, *CPC2*, is required to repress the crosspathway control in the presence of amino acids (Hoffman et al., 1999). The *cpcB* gene is an *A. nidulans* homolog of the yeast *CPC2* gene. Deletion of the *cpcB* gene resulted in the slight up-regulation of the crosspathway control-related genes, which is consistent with the yeast *CPC2* gene. The *cpcB* deletion caused similar blocking of sexual development to that in a *cpcA* overexpression strain or in the amino acid starvation conditions, but this blockage was not reversible. These results clearly showed that the *cpcA* and *cpcB* genes are responsible for the crosspathway control and the activation of the crosspathway control resulted in an impairment of sexual development in *A. nidulans* (Hoffman et al., 2000).

### 17.6.5 COP9 Signalosome (CSN) and *csnD/E*

Appropriate protein degradation is very important for regulating growth and development. Ubiquitylation is required for the targeted degradation and E3 ubiquitin ligase complex is a part of the enzymatic cascade. CSN, which is the constitutive photomorphogenesis complex 9 (COP9) signalosome, directly interacts with E3 ubiquitin ligases. It has been known that CSN is an important regulator of development. Busch et al. (2003) identified two components of the COP9 signalosome in *A. nidulans*, which can serve as a novel regulator of sexual development (Busch et al., 2003). One of the components was *csnD*, which encodes a PCI domain protein similar to the fourth subunit of CSN. Deletion of *csnD* resulted in a sterile phenotype of which sexual development was blocked at the primordial stage. Besides sexual development, Δ*csnD* also showed reduced radial growth and accumulation of red pigment. However, conidio- phore and conidia formation was not changed by the deletion of Δ*csnD*, indicating that the asexual development is not affected by the function of COP9 signalosome.

The fifth CSN subunit is *csnE*, containing conserved MPN domain. Deletion of *csnE* also gave almost identical phenotype of Δ*csnD*, which is blocked in sexual development, slow growth and red hyphae formation. So, the absence of either *csnD* or *csnE* generated identical phenotypes, indicating that both *csnD* and *csnE* are involved in the same function including several physiological and developmental processes (Busch et al., 2003). CSN regulation also affected light response in *A. nidulans* but *veA* expression was not altered by the CSN mutations. *veA* background or overexpression of *veA* could not overcome the sexual blockage caused by *csnD* mutation. Han et al. (1990) reported several mutants named BSDs, which are blocked in sexual development, but the corresponding genes have not been characterized so far. So, the CSN gene could be the one of the corresponding genes of the BSD mutation. This is the first report of molecular characterization of the essential player in the regulatory process that is involved in maturation of primordia (Busch et al., 2003).

### 17.7 Developmental Coordinators and Modifiers

#### 17.7.1 Spatio-Temporal Coordinators, StuA and MedA

Many mutations that were originally identified as causing defects in conidiophore development also block sexual development (Yager, 1992). Several regulatory mechanisms seem to coordinate the molecular control of sexual development as well as asexual development. Two transcription factors, the stunted protein (StuA)
and the medusa protein (MedA), have been termed as developmental modifiers of both sexual and asexual pathways. StuA is APSES (Asm1, Phd1, Sok2, Efg1, and StuA) family fungal protein, which is required for the correct spatial distribution of BrlA and AbaA (Miller et al., 1992). On the other hand, MedA is responsible for proper temporal expression of brlA transcripts and also functions as coactivator of abaA expression (Busby et al., 1996). The stuA mutant was unable to differentiate cleistothecia or Hülle cells, while medA mutant was unable to produce cleistothecia but do form Hülle cells (Clutterbuck, 1969; Wu and Miller, 1997; Dutton et al., 1997; Vallim et al., 2000). StuA also regulates some developmental specific gene expression such as cpeA, which is a Hülle cell specific catalase-peroxidase in A. nidulans (Scherer et al., 2002). Therefore, these genes may play a regulatory role in both asexual and sexual sporulation.

17.7.2 Leucine Zipper Protein, DopA
Miller and coworkers recently identified the dopey gene (dopA), which has been known previously as aco586 (Pascon and Miller, 2000; Axelrod et al., 1976). The dopA gene encodes an 1858 amino acids polypeptide with a molecular weight of 207 kDa. This putative protein has similarity from yeast Dop1 to human DopA. DopA has three leucine zipper domains in both N- and C-terminal and has a transcription activation domain of C/EBP (CAAT/enhancer binding protein) family at its C-terminus, suggesting that the DopA possesses transcriptional activation activity (Pascon and Miller, 2000).

A dopA disruption mutant shows 96% reduction of conidia production, which means the mutant is almost aconidial. Not only asexual sporulation but also sexual development is impaired and incapable of initiating sexual development in the ΔdopA mutant. DopA affects some important transcription factors of asexual or sexual development. Northern blot analysis with wild-type and ΔdopA showed that, in ΔdopA, expression of the brlA gene is delayed and diminished, and expression of the steA gene is up-regulated. Although the exact mechanism through which the DopA protein affects cell development is unknown, DopA plays an important role for normal developmental process as well as morphogenesis (Pascon and Miller, 2000).

17.7.3 WD Repeat Protein, RcoA
RcoA is a member of the WD repeat family and a homolog of Tup1p of S. cerevisiae and RCO1 of N. crassa (Hicks et al., 2001; Kelehler et al., 1992; Yamashiro et al., 1996). The functional analysis by deletion of the rcoA gene resulted in poor growth on minimal medium with small amount of conidia and irregular colony formation. The rcoA gene is also required for normal expression of the brlA, aflR and steU genes and production of a mycotoxin, steigmatocystin (Hicks et al., 2001).

Recently, it was found that rcoA is required for sexual development (Todd et al., 2006). Genetic analysis with a ΔrcoA mutant and a mutant having blA1 (blue ascus) mutation revealed that the rcoA gene is necessary for nuclear contribution to the cleistothecial walls (Todd et al., 2006). Furthermore, multicyclop rcoA strain in veAI background undergoes sexual development just like veA+ strains, suggesting that additional copy or increased expression of rcoA suppresses the veAI phenotype and promotes sexual development. When compared to ΔrcoA with veAI and veA+ background strains, the growth retardation caused by rcoA deletion was more severe in veA+ backgrounds, suggesting that veAI partially suppresses growth limitation of ΔrcoA. Although veA+ promotes sexual development (Kim et al., 2002), ΔrcoA veA+ fails to form cleistothecia. Moreover, overexpression of veA under the niiA promoter in ΔrcoA background cannot rescue the failure of sexual development of the ΔrcoA mutant, indicating that veA lies upstream of rcoA in the sexual development pathway and rcoA acts as a key regulator of sexual development as well as growth and asexual development (Todd et al., 2006). All of these developmental modifiers, including stuA, medA, dopA, and rcoA, demonstrate that a sophisticated crosstalk must exist between the two developmental pathways.

17.7.4 Genomics and Sexual Development
Sexual development and fruiting body formation of filamentous fungi is very complicated and a genetically regulated progress. Despite the importance of sexual development, the research on it has not been
intensively performed. Recent achievement of genome sequencing in various filamentous fungi including *A. nidulans*, *A. fumigatus* and *A. oryzae* (Galagan et al., 2005) provided valuable information for understanding fungal growth and development.

Before the genome sequencing was accomplished, many efforts for obtaining massive information on development of *A. nidulans*, including EST analysis, cDNA microarray, and a subtraction analysis, had been made (Lee et al., 1996; Sims et al., 2004; Ray et al., 2004). Chae and coworkers established a stage-specific cDNA library and obtained EST sequences from the early sexual developmental stage (ESD) and the late sexual developmental stage (LSD) (Lee et al., 1996). Analysis of the LSD and ESD sequences provided a lot of information about differentially expressed ribosomal genes and housekeeping genes during sexual development as well as characterization of genes specifically expressed at the late sexual developmental stage (Jeong et al., 2000; Jeong et al., 2001; Lee et al., 2001).

The whole genome sequence and gene annotation databases provided a powerful and convenient tool for transcription profiling, microarray. A cDNA microarray chip based on EST sequence, which covers about 4000 genes, was reported (Sims et al., 2004). Moreover, an oligo-DNA (70-mer) microarray chip has been constructed and is available at the pathogen fungal genomics resource center (PFGRC). These important genomic tools and proteomic analysis will be very helpful for identifying and characterizing gene sets expressed in a stage-specific manner including in the sexual developmental stage (Archer and Dyer, 2004).

Recent advances toward understanding sexual development in aspergilli suggest hidden sexual developmental ability in some species, that is, *A. fumigatus*, having only asexual cycle. Dyer and colleagues identified mating-type loci as well as the pheromone receptors in the *A. fumigatus* genome (Paoletti et al., 2005; Dyer et al., 2003). Comparative transcription profiling of *A. fumigatus* has shown that some important genes for sexual development are up-regulated during or after the hypoxic condition but some are not. These results suggest that although *A. fumigatus* may have full genetic components for sexual development, the regulation process is not maintained to undergo sexual development in *A. fumigatus*.

References


Busby, T.M., Miller, K.Y., and Miller, B.L., Suppression and enhancement of the *Aspergillus nidulans* medusa mutation by altered dosage of the bristle and stunted genes, *Genetics*, 143, 155, 1996.


Dutton, J.R., Johns, S., and Miller, B.L., StuAp is a sequence-specific transcription factor that regulates developmental complexity in *Aspergillus nidulans*, *EMBO J.*, 16, 5710, 1997.


Han, D.M. et al., Isolation and characterization of NSD mutants in *Aspergillus nidulans*, *Kor. J. Genet.*, 15, 1, 1994b.

Han, K.H. et al., Characterization of several NSD mutants that never undergo sexual development, *Kor. J. Genet.*, 20, 257, 1998.


Han, K.H. et al., Regulation of nsdD expression in *Aspergillus nidulans*, *J. Microbiol.*, 41, 259, 2003b.


Lee, B.Y. et al., Screening of growth- or development-related genes by using genomic library with inducible promoter in *Aspergillus nidulans*, *J. Microbiol.*., 43, 523, 2005.


Sexual Development in Aspergillus nidulans


Yue, C. et al., The STE12alpha homolog is required for haploid filamentation but largely dispensable for mating and virulence in *Cryptococcus neoformans*, *Genetics*, 153, 1601, 1999.


18

Aspergillus Transporters

George Diallinas

CONTENTS

18.1 Aspergillus nidulans: A Champion of Transporters ............................................................... 301
18.2 Methodological Approaches to Study Aspergillus nidulans Transporters ................................ 303
  18.2.1 Growth Tests ................................................................................................................... 303
  18.2.2 Uptake Measurements ..................................................................................................... 303
  18.2.3 Fluorescent Microscopic Imaging ..................................................................................... 304
  18.2.4 Epitope Tagging and Immunodetection .......................................................................... 304
  18.2.5 Mutations: Random, Directed, Cys-Scanning ................................................................. 305
  18.2.6 Chimeras ....................................................................................................................... 306
18.3 Integration of Multiple Physiological Signals at the Level of Transcription of Transporters ................................................................. 307
18.4 Sensing the Growth Milieu ....................................................................................................... 308
18.5 Transporters on the “Air”: Not Just Food Suppliers? ................................................................. 309
18.6 Transport in Sex ...................................................................................................................... 310
18.7 An Emerging Role of Regulated Trafficking and Endocytosis of Transporters ......................... 310
18.8 Paradigms of Transporter Structure-Function Analysis in Aspergillus ........................................ 311
18.9 Aspergillus as a Novel System for Studying Transporters from Complex Organisms ............ 313
18.10 Epilogue ................................................................................................................................. 314
Acknowledgments .......................................................................................................................... 315
References ....................................................................................................................................... 315

18.1 Aspergillus nidulans: A Champion of Transporters

The plasma membrane of all cells and internal membranes of eukaryotes contain a wide variety of proteins that ensure transmembrane solute transport. Mechanistically, they can be classified as transporters (or permeases) and channels (http://www.tcdb.org/). Transporters are classified as primary active transporters, secondary active transporters, and facilitators, depending on their energetic requirements. Active transporters catalyse the transport of metabolites up an electrochemical gradient using either ATP hydrolysis (primary transport) or the movement of another species, most commonly an ion (H⁺, Na⁺, K⁺), down an electrochemical gradient (secondary transport). Secondary active transporters can be symporters (transport of a substrate and ions in the same direction) or antiporters (transport of a substrate and ions in the opposite direction). Facilitators (uniporters) are energy-independent or passive, transporters mediating the movement of a solute across the plasma membrane along its concentration gradient. Transporters catalyse the uptake or efflux of most metabolites (amino acids, nucleobases, nucleosides, sugars, nitrogenous solutes, vitamins, etc.). In contrast, channels mediate passive transport of ions by forming an aqueous diffusion pore. Despite their structural similarity, consisting of a highly modular structure with, usually, 10–14 repeated (polytopic) hydrophobic or amphipathic α-helices, two properties distinguish channels from transporters: (1) ion flow is extremely fast, and (2) ion channels are gated, their opening frequency being regulated by changes in membrane potential, by binding of a specific ligand, or by mechanical constraints such as membrane stretching. In contrast, transporters undergo
reversible conformational changes that expose their solute-binding site alternately on each side of the membrane. However, transporter-like channels or channel-like transporters have been described (Wadiche and Kavanaugh, 1998; Boyd et al., 2003; and references therein).

Since the first molecular characterization, 22 years ago, of a specific L-proline permease (PrnB; Sophianopoulou and Scaccioro, 1989), several transport/channel systems have been characterized in Aspergillus nidulans. In fact, this model fungus possesses examples of all categories of carriers (http://membranetransport.org; http://www.broad.mit.edu/annotation/fungi/aspergillus/). It is estimated that more than 700 genes, i.e. at least 7.1% of the total A. nidulans genome, encode proteins catalyzing the transport of solutes and ions across membranes. This makes A. nidulans come second in the list of eukaryotes (Cryptococcus neoformans 7.3%). In bacteria transporters can be up to 13.8% and most often exceed 11%. Similar percentages exist for Archaea. In other eukaryotes this can be from 2% (Dictyostelium discoideum, protozoa) to 3.4–4.1% (Caenorhabditis elegans, Drosophila melanogaster, plants, and mammals). Data concerning comparative genomics of transporters/channels and transporter families can be obtained from the excellent sites http://www.membranetransport.org/ and http://www.tcdb.org/.

The 703 putative transporter/channels of A. nidulans are 81.5% secondary active transporters (573), 11.8% primary-active transporters (83) and 4.4% channels (31), the remaining 1% being unclassified. Nearly 2/3 of the secondary active transporters classify within the Major Facilitator Superfamily or MFS (358). Other large families are the APC (amino acid, GABA, choline transporters; 55 proteins), ABC (multidrug resistance or fatty acid efflux proteins; 47), P-ATPases (ion or phospholipids transporters; 23), MC (mitochondrial carriers; 35), AAAP (amino acid; 14), NCS1 (nucleobase transporters; 11), and DMT (metabolite/drug transporters; 11). The rest make families of 1–9 members. Interestingly, Aspergillus fumigatus and Aspergillus oryzae have different numbers of transporters. A. fumigatus has 79 transporters less than A. nidulans and some families are differentially represented. For example the MFS of A. fumigatus has 275 members, 83 less than A. nidulans, but a large family of CytB ferric reductase channels (17 members compared to the single-membered family of A. nidulans and A. oryzae). A. oryzae has 245 transporters more than A. nidulans, with major differences coming from the 507 MFS, 76 APC, 72 ABC, 24 AAAP, and 16 POT (oligopeptide transporters) families. DEFINE All aspergilli have far more transporters than Neurospora crassa (364 proteins). It is interesting to compare the MFS number of N. crassa (141) with that of A. oryzae (507), as well as the differences in APC and AAAP, the two amino acid transporters families (19 versus 96). Saccharomyces cerevisiae and Schizosaccharomyces pombe have 323 and 211 transporters, respectively. Considering the number of genes and genome size of S. cerevisiae, the total number is similar to the number in filamentous fungi. However, there are important differences in the number of family members, which can vary significantly as in the case of MFS (85 versus 358–507), and in some families, can be absent, as for example the NCS2/NAT purine transporters missing from S. cerevisiae. S. pombe has clearly less transporters, with, for example, only 58 MFS, 17 APC, 9 ABC.

It is also interesting to compare Aspergillus transporters and their numbers with other nonfungal systems. E. coli and B. subtilis have 532 and 423 transporters respectively, of which very few are channels (<2.8%). Several prokaryotic-specific systems, including phosphotransferases, exist in these model bacteria. D. discoideum, with a genome size similar to aspergilli, has only 267 transporters, with great increase in primary-active transport (35.6%) and even greater reduction in the MFS (32 members). The relative percentage of channels, primary- and secondary-active transporters in yeast, filamentous fungi, and D. discoideum remain rather similar. Among protozoa, Trypanosoma brucei has 334 transporters, with very different composition and size of families. Only 20 MFS, but 82 AAAP, and no NCS1 or NCS2 families, are present in this parasite. Leishmania major is similar while Plasmodium falciparum has only 105 transporters with reduced percentage for secondary-active transporters (46.7%) and increased for primary active-transporters (41.9%). Interestingly, the ENT (Equilibrative Nucleoside Transporter) family is very variable in these three protozoa, with 21, 5 or 2 members in T. brucei, L. major, P. falciparum, respectively. Aspergilli and N. crassa have only 1 ENT carrier. C. elegans, with a genome three times bigger in size compared to that of aspergilli and at least 5000 genes more, has 666 transporters, of which the secondary-active transporters diminish to 52.4%, and channels increase to 34.5% of the total number. D. melanogaster is similar to C. elegans, with 632 transporters (55% secondary-active transporters and
Aspergillus Transporters

28.3% channels). Humans have 934 transporters, with an increased number and type of channels (37.5%, only 81 MFS) and reduced numbers of secondary-active transporters (36%). The increased number of channels apparently reflects their role in the development of a nervous system. Arabidopsis thaliana has many transporters (990), with several plant-specific transporters and very dissimilar numbers within different families, as for example, the reduced number in MFS (90) and the increased number in ABC (108) and DMT (121) families. It is noticeable that there are families with well-conserved numbers in several organisms, such as the mitochondrial transporters (MC) (34–45 in fungi, D. discoideum, C. elegans, D. melanogaster, protozoa, and mammals). The Most divergent numbers in MC transporters are present in S. pombe (22) and in A. thaliana (52).

The medical importance of transporters is directly apparent from their biological role. Several human genetic disorders such as cystic fibrosis, X-linked adrenoleukodystrophy, and diastrophic dysplasia are caused by alterations in membrane transport proteins that have retained high similarity to Aspergillus proteins (http://www.membranetransport.org/ and http://www.tcdb.org/). The human ascorbate transporters, proteins essential for life, are homologous to the A. nidulans uric acid-xanthine transporters (Koukaki et al., 2005). Ammonium transporters of Aspergillus and other fungi are homologous to the human and mouse Rhesus proteins, which may also function as ammonium transporters (Avent et al., 2006). The role of transporters in modern pharmacology and agriculture is of increasing interest. Today, most drugs, including antifungals, are the products of massive random screens and their discovery depends less on knowledge of metabolism and enzyme-substrate interactions. Most drugs need to enter a cell, as the plasma membrane is practically an impermeable barrier. Transporters might serve as specific gateways to deliver drugs selectively to target cells (microbes, pathogens, cancer cells) and thus avoid side effects from uptake in other host tissues (Kraupp and Marz, 1995; De Koning et al., 2005; Blagini et al., 2005; Ho and Kim, 2005). In addition, the efflux of drugs and xenobiotics, leading to pleiotropic or multidrug drug resistance (PDR/MDR), through ABC and MFS transporters (Sipos et al., 2006 and references therein) is of primary importance in various areas of biological research, from treating cancer to fighting pathogenic fungi. Understanding how transporters function, how they obtain their final topology and how they are regulated, or what determines substrate specificity, binding and transport, constitute not only an essential part of the basic understanding of cell function, but is expected to contribute to better pharmacological therapies and play a significant role for novel targeted phytopharmacological protocols, so much needed today.

18.2 Methodological Approaches to Study Aspergillus nidulans Transporters

18.2.1 Growth Tests

Several transport mutants have a visible growth phenotype. When the transported metabolite is needed for growth (as a source of carbon, nitrogen sulfur, phosphorus, vitamin, etc.) lack of uptake can lead to a range of visible phenotypes. Leakiness also directly reflects the existence of more than one uptake system for a particular metabolite. Even in cases where lack of uptake leads to no visible phenotype (e.g., lack of uracil uptake), the use of toxic analogs (see introduction) or specific mutations resulting in other metabolic defects can be used to obtain a discernable phenotype. For example, lack of uracil uptake can either be seen as resistance to 5-fluorouracil, a highly toxic uracil analog, of by selecting uracil uptake mutations in strains partially deficient in uracil biosynthesis (Pulmer et al., 1975; Amillis et al., 2006).

18.2.2 Uptake Measurements

Solute uptake measurements in filamentous fungi can present some technical problems related to mycelium mat. In A. nidulans this problem can be overcome by performing uptake studies at a stage before the emergence of the germ tube. Studies concerning the expression of several genes have shown that at this stage germinating conidiospores are physiological equivalent to young mycelium (Tazebay et al., 1997; Amillis et al., 2004). Another convenient aspect for performing uptake studies using
germinated conidia is the fact that all transporters analyzed so far are developmentally expressed earlier, during the isotropic growth phase of germination, eliminating the need to induce transporter expression by physiological signals, usually by substrate induction. This also eliminates the time-consuming and laborious washing of conidia in order to remove excess “cold” substrate before performing the uptake assays. A final advantage of using individual conidia for performing transport kinetics is that data concerning transport rates can be given per viable conidia.

18.2.3 Fluorescent Microscopic Imaging

Tagging transporters with GFP has proved an excellent tool for following the in vivo dynamic fate of *Aspergillus* transporters (Valdez-Taubas et al., 2000; Tavoularis et al., 2001; Koukaki et al., 2005; Forment et al., 2006; Pantazopoulou et al., 2006). In addition, detecting transporter topology is a prerequisite for classifying transporter mutations to classes affecting structure/topology and those affecting kinetic characteristics such as $V_m$ or $K_m$ per se. Both epifluorescence and confocal microscopy can be used, the former being usually sufficient to draw conclusions and easier to perform. Several transporters tagged, at their C-termini, proved fully or sufficiently functional to perform cytological studies. In some cases, GFP tagging was entirely “silent” in respect to transport function (as for purine transporters UapA, UapC, AzgA; Koukaki et al., 2005; Pantazopoulou et al., 2007), while in other cases it can partially affect the kinetic parameters of the transporter (PnB proline permease; Tavoularis et al., 2001). In the latter case, the problem can be solved or diminished by employing different amino acid linkers between GFP and the transporter. The 2–4 amino acid-linkers consisting of Gly residues were the best for PnB, but this should be examined for each transporter. An alternative to GFP is the use of mRFP1, which was employed to detect the UapA purine transporter (Koukaki et al., 2005; Lemu and Diallinas, unpublished).

These studies have shown that transporters are localized principally in the plasma membrane, in the area of the septum and in the vacuoles, which appear as cortical punctuate or larger granules/organelles (Fig. 18.1). Localization and degradation in the vacuoles reflect the final step in a transporter life cycle but also directly depends on the physiological state of the cell and on several stress responses (temperature, ageing, overloading with excess substrate, preference of an alternative metabolite due a sudden shift in media composition, etc.; Tavoularis et al., 2001; Valdez-Taubas et al., 2004; Pantazopoulou et al., 2007). The identity of vacuoles can easily be traced using specific chemical probes, such as CMAC or CDCFDA. Normally, transporters cannot be seen en route from the ER, the Golgi or other secretory compartments, such as the late or early endosomes. GFP- or RFP-tagged transporters can be used as tools for studying systematically, endocytosis and trafficking from ER to the Golgi, the endosomes or the vacuole (discussed later; also see Fig. 18.1). UapC-GFP has already been used to this aim (Higuchi et al., 2006). The use of GFP- or RFP-specific antibodies provides an extra advantage for using the corresponding tags in immunological studies concerning the proteomics of transporters (Kinghorn et al., 2005; Pantazopoulou et al., 2007).

18.2.4 Epitope Tagging and Immunodetection

Immunodetection with transporter-specific antibodies has proved a difficult task. High titre affinity-purified antibodies, obtained against various UapA or UapC transporter domains, failed to recognize any specific polypeptide in Western blots (Valdez-Taubas et al., 2000; Sophianopoulou, unpublished). Anti-NtrA antibodies have been more successful in the detection of the nitrate transporter expressed from its native promoter, but still detection was laborious (Kinghorn et al., 2005). Tagging with immunologically detected epitopes provides a more efficient alternative tool. Standard epitopes such as V5 or His$_6$ have been used for the nitrate (NrtA; Unkles et al., 2004) and uric acid/xanthine (UapA) or purine (AzgA) transporters (Pantazopoulou et al., 2007), respectively. These tags were cloned C-terminally and transporter-specific bands could be efficiently detected in single-copy transformants, after expression from native promoters, by using monoclonal or polyclonal antibodies. Importantly, both of these tags did not affect the kinetic characteristics of the transporters studies. An alternative tag used successfully for the
study of the proline transporter PrnB (Kafasla et al., 2007) was that of the Biotin-Acceptor-Domain (BAD) (Consler et al., 1993) followed by a dodecapeptide epitope corresponding to the C-tail of the E. coli lactose permease (Carrasco et al., 1984). This "silent" tag permits the purification of any transporter, through affinity chromatography using avidin columns, and detection via either anti-BAD or anti-C-tail polyclonal antibodies.

18.2.5 Mutations: Random, Directed, Cys-Scanning

Transporters and channels are very difficult to study using classical approaches of structural biology. Crystals are very difficult to obtain and NMR cannot be applied for such long hydrophobic proteins, at least not yet. Most of these problems arise from the hydrophobic nature of these transmembrane proteins. Thus, it is not surprising that among more than 30,000 entries of solved structures in the Protein Data Base only around 100 concern membrane proteins and far less polytopic transporters and channels (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Randomly selected and oligonucleotide-directed mutations are classical complementary approaches to study structure-function analyses in transporters, proteins very difficult to be studied by standard biochemical and structural approaches. The paradigmatical work of R. Kaback on the lactose permease has illuminated the way (Kaback et al., 2005). Through selecting, constructing, and combining hundreds of mutations, Kaback and colleagues have proposed a model on how lactose permease folds and binds its substrates, which proved very close to the structural model solved recently after X-ray crystallography. A. nidulans provides an excellent system to select

FIGURE 18.1 Cellular expression of A. nidulans transporters. Note: (a) Wild-type and mutant forms (Q85T, H86A, H86N, H86D; Pantazopoulou and Diallinas, 2006) of UapA expressed in undifferentiated hyphae. Arrows highlight septa, the ER membrane (as a perinuclear ring), and vacuoles. (b) A truncated form of UapA-mRFP1, missing transmembrane-helix 12, is retained in the ER, which is simultaneously detected with ShrA-GFP (see text) (Vlanti et al., 2006). (c) Expression of AzgA during sexual differentiation (Pantazopoulou and Diallinas, unpublished). Upper panel shows a young cleistothecium and hülle cells (arrows). Notice the fluorescence of hülle cells and of interconnecting hyphae on the surface of the cleistothecium. Lower panel shows an ascogenous hyphae (AH). Arrow indicates a septum and an open circle highlights a swelling area, corresponding to the site of initiation of the development of a cleistothecium. (d) UapA expression in metulae and PrnB expression in conidiospores (Pantazopoulou and Diallinas, 2006).
conditional, kinetic, or specificity mutations in several solute transporters. Thermo- or cryosensitive mutations can be affected in either topogenesis or the kinetic parameters of a specific transporter. The use of strains in which the transporter gene to be studied is tagged with gfp and integrated in the endogenous locus, should allow classification of mutants into topological or kinetic/functional mutations. Affinity mutants can be selected as mutants lacking uptake, and thus unable to grow on appropriate media at low concentration of substrate, but permitting growth at higher concentrations of the substrate. Specificity mutations can be even easier to select as mutants permitting growth on a given substrate but not on another, or as mutants leading to resistance to a toxic analog while still allowing growth on physiological substrates. Intragenic second-site suppressors can be selected and employed to understand interdomain interactions. In the case of the UapA uric acid/xanthine transporter, for example, allele-specificity between two sites in the transporter provided strong evidence for the functional and possibly physical, interaction of the proposed substrate binding site and a distant substrate selectivity filter (Vlanti et al., 2006). Random mutation and primary amino acid alignments should also direct the rational design of directed mutations. Directed mutations usually concern conserved or semiconserved residues with particular physicochemical properties, such as charged amino acids or prolines and glycines within α-helical transmembrane domains. All these genetic approaches are complementary and have led to important conclusion on the structure-function relationships of the nitrate transporter NtrA, the proline transporter PrnB, and mostly on the UapA uric acid/xanthine transporter (see section on structure-function studies later).

Cys-scanning mutagenesis combined with site-directed sulphydryl labeling and various biochemical and spectroscopic techniques is a powerful approach to study structural and dynamic aspects of membrane protein structure and function (Frillingos et al., 1997, Miklos Sahin-Toth et al. 2000; and references therein). The power of this approach in the aforementioned aspects was first demonstrated with the LacY permease. In a first step all Cys residues should be removed from a given transporter. Apart from investigating the essentiality and involvement of cysteine residues per se, this technique leads to the generation of Cys-less transport proteins, which, if functional, will form the matrix on which single Cys residues can be introduced at any position and studied for their accessibility to chemical modification. Studies addressing the effect of substrate presence on chemical modifications of single-Cys versions of transporters can then be used to further investigate the proximity of the residues studied to the substrate binding and translocation site. Two functional Cys-less versions of A. nidulans carriers, the NtrA nitrate transporter (Unkles et al., 2005) and the PrnB proline transporter (Kafasla, Frillingos, and Sophianopoulou, in preparation), have been recently made and evaluated as functional (30–50% $V_m$ values compared to wild-type alleles and wild-type $K_m$s). It would appear therefore, that none of cysteines, in both transporters, is involved directly in substrate transport or topogenesis. Having established the experimental conditions for sulphydryl (NEM, etc.) treatment of conidiospores, the Cys-less versions of these transporters are currently analyzed by systematic Cys-scanning mutagenesis. The regions studied were selected on the basis of pilot classical mutagenesis studies, which indicated the importance of specific amino acids in structure–function relationships.

18.2.6 Chimeras

The construction and analysis of chimeric proteins composed of parts of similar transporters provides an excellent a priori approach to identify domains important for topology, function, and specificity. An example concerned the similar, but kinetically distinguishable, UapA and UapC uric acid/xanthine transporters of A. nidulans. Using a rapid, in vivo, approach, based on recombination in E. coli, several UapA-UapC and UapC-UapA functional chimeras were made and showed that a 69 amino acid residue-long region, including an amphipathic loop and one putative transmembrane segment, was necessary for substrate binding and transport (Diallinas et al., 1998). Further mutational analysis has provided strong evidence that this region includes at least part of the substrate-binding site (Meintanis et al., 2000; Koukaki et al., 2005). Combined with more recently developed techniques, such as the use of fluorescent tags (see earlier) and PCR-based joining of DNA segments (Yang et al., 2004; Yu et al., 2004), chimeras should provide a powerful tool for transporter analysis.
18.3 Integration of Multiple Physiological Signals at the Level of Transcription of Transporters

Expression of solute transporters is a primary target of several metabolic control transcriptional circuits. Most transporters related to catabolism are regulated by substrate-induction and general metabolite repression (Wiame et al., 1985; Davis et al., 1993; Scorzocchio 1992; Wilson et al., 1998). In A. nidulans known examples include transporters specific for nitrogen sources, such as amino acids, GABA, purines, urea, nitrate, ammonium, and carbon sources, such as glucose, ethanol, fructose, monocarboxylic acids, etc. Substrate-induction of a specific transporter is a prerequisite for efficient induction of all genes encoding enzymes involved in the utilization of this substrate. Repression of transporter synthesis on the other hand, is a very efficient mechanism, present in many fungi, for repressing the synthesis of enzymes by inducer-exclusion, and thus eliminating the need for individual repressor target sites in the promoter of genes encoding the enzymes for a particular catabolic pathway. While in some cases fungal genes are repressed only by inducer-exclusion (proline utilization; Cubero et al., 2000), in some other cases, inducer-exclusion operates hand-in-hand with direct repression of individual genes (nitrate and purine utilization; Oestreicher and Scorzocchio 1993; Glatigny and Scorzocchio 1995; Punt et al., 1995), providing a more efficient and versatile double-lock mechanism for transcriptional shut-off.

Some metabolites can serve very different functions in catabolic and anabolic pathways. Amino acids can be used as nutrients, but also directly channeled to protein synthesis. Nonoxidized purines can serve as nitrogen sources, but also channeled to nucleotide and nucleic acids synthesis. Uric acid and proline, besides being nutrients or constituents of macromolecules, might also serve protective roles for the cell. Most solutes can serve to a variety of cellular responses. This poses an apparent problem to the cell that can be solved by two strategies. One is to use different transporters for the same substrate, depending on the particular role of the solute taken up in a given physiological condition or a given developmental stage. In A. nidulans, very recent evidence suggests that while purines, when abundant, are taken up very efficiently by the purine-inducible AzgA transporter (Cecchetto et al., 2004), the same purines are continuously scavenged, probably for anabolic purposes, by a high-affinity, very low-capacity constitutive transporter, tentatively called FcyB (Vlanti and Diallinas, unpublished). AzgA- and FcyB-close homologs exist in all aspergilli with known genomes. The alternative solution to the need for the uptake of solute that can be used in different pathways, is to have alternative mechanisms for the regulation of transporter synthesis. The example of regulation of the prnB, the gene encoding the major proline transporter, is an excellent one. Proline can serve as nitrogen and carbon source, but can also be channeled to protein synthesis or it might serve as an antidrought molecule (see Section 18.5). It has been shown that prnB transcription is regulated by no less than six independent or partially independent mechanisms: proline induction, nitrogen metabolite repression, carbon catabolite repression, the general control system regulating amino acid pools, an independent mechanism in response to germination, and an unknown conidiospore-specific mechanism, Cubero and Scorzocchio 1994; Tazebay et al., 1995; 1997; Gonzalez et al., 1997; Cubero et al., 2000; Pantazopoulou and Diallinas, unpublished). Similarly, transcription of uapA, the gene encoding the xanthine/uric acid transporter, is regulated by purine-induction, nitrogen metabolite repression, an independent mechanism in response to germination, and a developmental, metulae-specific mechanism (Gorfrinkiel et al., 1993; Diallinas et al., 1995; Amillis et al., 2004; Pantazopoulou et al., 2007). Some of these aspects are also discussed in more detail in later sections.

Work on the transcriptional regulation of transporters has revealed some very interesting aspects on A. nidulans molecular biology. Studies on the regulation genes encoding the UapA and UapC purine transporters, the PrnB L-proline transporter or the GABA transporter, have confirmed the crosstalking of pathway-specific and the general transcription factors AreA and CreA, mediating nitrogen catabolite repression (NCR) and carbon metabolite repression (CMR), respectively. Such interactions have been previously proposed on the basis of genetic evidence. In particular, the molecular and functional analysis of a number of areA (the gene encoding the general transcription GATA-like factor AreA, necessary for the transcription of more than 100 genes implicated in the utilization of nitrogen sources; Arst and Cove, 1973) DNA-binding specificity mutations and cis-acting regulatory mutations in the promoter regions of
uapA and uapC have proved valuable tools in identifying the binding sites of both AreA and UaY (the positive-acting, pathway-specific, regulatory protein, which in the presence of uric acid mediates the induction of most genes involved in purine utilization) (Ravagnani et al., 1997; Oestreicher, Diallinas, Gomez, Gordon, de Queiroz, and Scanzocchio, unpublished) Gel shifts, in vitro footprinting, and interference assays (Ravagnani et al., 1997), and in vivo nucleosome positioning (Kagias and Strauss, unpublished) have established important aspects on the molecular interactions at these binding sites. It seems that UaY binding to its site(s) plays a critical role in directing the binding of the general factor AreA and thus initiating chromatin remodeling and gene expression activation. Similarly, derepressed mutations, called ppr^t, mapping in the prnB (proline transporter) promoter (Arst and Cove, 1973; Arst and MacDonald, 1975; Sophianopoulou et al., 1993) defined the binding site for the CreA, which was subsequently confirmed by in vitro and in vivo studies to be so (Cubero and Scanzocchio, 1994). Similar studies have led to the identification of the binding sites for the PrnA (the pathway-specific regulator responsible for induction by proline (Cubero et al., 2000; Gomez et al., 2002) and AreA (Gonzalez et al., 1997; Gomez et al., 2003) in the prnB promoter region. It was further shown that under glucose/ammonium-repressing conditions, partial nucleosomal positioning depends on the CreA repressor’s binding to two specific cis-acting sites (Garcia et al., 2004). AreA was not involved in nucleosome positioning, which contrasted with its role in another promoter (niiA-niad; nitrate utilization). Interestingly, prnB transcription induced by amino acid starvation, possibly through the action of the general regulator CpcA/Gcn4, leads to a different chromatin rearrangement. Default nucleosome positioning and partial positioning under induced-repressed conditions seem to depend on deacetylated histones (Garcia et al., 2004).

Analogous studies with the GABA transporter gene have led to significant conclusions concerning interactions of the gaba promoter with several trans-acting regulators, such as IntA, CreA, AreA, and PacC. Early genetic analysis has shown that gaba expression is subject to carbon catabolite and nitrogen metabolite repression (Bailey et al., 1979), mediated by CreA and AreA (see earlier), and ω-amino acid (such as β-alanine or GABA) induction (Arst, 1976; Bailey et al., 1979) mediated by the zinc binuclear cluster protein IntA/AmdR (Andrianopoulos and Hynes, 1990). In addition, gaba is regulated in response to the pH of the growth medium. The physiological target sites for all these regulators have been identified in the gaba promoter. It was shown that a double PacC binding site overlaps the binding site for the transcriptional activator IntA and competes for DNA binding. Thus, PacC was shown to act as a genuine repressor for an acid-expressed gene through preventing the binding of a positively acting transcription factor. Since it was also shown that PacC acts also as a transcriptional activator for other alkali-expressed genes (Penalva and Arst, 2004), a dual role of PacC as activator and repressor in pH regulation was established. Interestingly, close homologs of the A. nidulans GabA can be missing from some Aspergillus species, such as A. oryzae.

All genetic and molecular evidence has established that AreA is an activator, inactivated by ammonium, while CreA is a repressor, activated by glucose. However, the recent work with ammonium transporter (AMT/MEP) genes and a low-affinity glucose transporter (MtsE) revealed new roles for AreA and CreA. In particular, the expression of ammonium transporter genes under all nitrogen conditions, including ammonium-rich media, was shown to be dependent on AreA (Monahan et al., 2006). On the other hand, mstE expression was induced at the transcriptional level in the presence of glucose and other repressing carbon sources and this induction was CreA-dependent (Forment et al., 2006). It remains, however, to be determined whether these phenomena are due to a direct effect mediated by AreA- and CreA-binding sites or rather a cryptic, indirect, consequence. In any case, the studies presented here, show that transporter genes have evolved alternative uses for these transcription factors.

### 18.4 Sensing the Growth Milieu

Conidial germination seems to involve the transcriptional activation of not only house-keeping genes, such as the actin-encoding gene actnA (Tazebay et al., 1997), but also of genes encoding transporters. Two old physiological studies have shown that nitrate and ammonium transport activities in A. nidulans are very low in resting conidiospores, increase dramatically during germination to reach a maximum associated with germ tube emergence, and drop to basic levels in mycelium (Cook and Anthony, 1978;
Brownlee and Arst, 1983). Later, it was shown that PrnB, the proline transporter of Aspergillus nidulans, is not expressed in resting conidiospores but is transcriptionally activated during the isotropic growth phase of germination (Tazebay et al., 1995). It was shown that prnB transcriptional activation, although it partially responds to physiological signals such as amino acid starvation or proline induction, is independent of the known transcription regulators PrnA and CpcA/Gcn4, which operate fully only after polarity establishment (Tazebay et al., 1997). Recently, growing evidence has accumulated to show that the effect seen with prnB is much more general, as several other transporters, such as those specific for aspartate/glutamate (Apostolaki and Scazzocchio, unpublished), allantoin (Hamari, Amillis, Diallinas, Scazzocchio, in preparation), nucleoside (Hamari, Amillis, Diallinas, and Scazzocchio, in preparation), glucose (Forment et al., 2006), and lactic acid (Diallinas, unpublished), as well as several homologs of the uracil transporter family (Amillis et al., 2006; Hamari, Amillis, Diallinas, and Scazzocchio, in preparation), are also transcriptionally activated during the isotropic growth phase of germination. Again, this activation was independent of both the physiological conditions of germination and the pathway-specific transcription factors. More detailed analysis was performed for the purine transporters UapA, UapC, and AzgA (Amillis et al., 2004). Transcriptional activation of the uapA, uapC, and azgA genes occurs during the isotropic growth phase, prior to the first nuclear division, and leads to the appearance of the corresponding purine transport activities with a small time delay (30–60 min). Similarly to prnB, uapA, uapC, and azgA transcriptional activation was independent of the major, pathway-specific, transcription factor known as UaY, as both loss-of-function (uaY) and constitutive (uaY) mutations had no effect on transcription during germination. In fact, the only requirement for this transcriptional activation was the presence of a carbon source (glucose or fructose) in the germination medium. Interestingly, moreover, the de novo transcription of all three purine transporter genes is activated even in the absence of any carbon source. The lack of repression in the presence of ammonium or glutamine was in line with the observation that this novel transcriptional activation mechanism is also independent of AreA, the general GATA factor, mediating nitrogen catabolite repression. Thus it seems that a novel control operates very early during germination specifically for transporters, and not for metabolically relevant enzymes. What makes unique the promoter of a transporter gene or how such an idiosyncrasy of transporter promoters has evolved, is not known. This novel mechanism of regulation should serve as a transient system for sensing various solutes, and accordingly regulate the expression of the corresponding metabolic pathways. Thus, unlike S. cerevisiae, where highly specific sensor proteins can activate true transporters (Iraqui et al., 1999; Forsberg and Ljungdahl, 2001), Aspergillus nidulans might use its transporters per se for both sensing the environment and for the bulk transport of solutes. The protagonists controlling this novel regulation system remain unknown.

### 18.5 Transporters on the “Air”: Not Just Food Suppliers?

In our lab, strains expressing functional versions of transporter genes fused with gfp were used for studying transporter expression in the asexual conidial apparatus. UapA-GFP, UapC-GFP, AzgA-GFP, and PrnB-GFP have been studied and led to surprises (Pantazopoulou et al., 2007). Figure 18.1 shows that UapA-GFP was not expressed in the conidiophore stalk and the vesicle, but was highly expressed in the periphery of the metulae. The same result was obtained when the strain was grown under inducing (uric acid) or noninducing (urea) conditions. UapC was conditionally expressed in metulae, in samples grown only in the presence of uric acid. AzgA-GFP was not expressed in any of the asexual structures of A. nidulans under any condition used (urea or hypoxanthine). PrnB-GFP (urea or proline) was expressed specifically and intensively in conidiospores, and much less in phialides, but not at all in the metulae, the vesicle, or the conidiophore (Fig. 18.1). Interestingly, GFP studies showed that uricase (UaZ), the first enzyme involved in uric acid oxidation, is also expressed in the metulae, the phialidies and the conidiospores, suggesting that the machinery for uric acid catabolism is operating during asexual reproduction (Langousis and Diallinas, unpublished). Why this is so in the absence of uric acid from the growth medium seems paradoxical. A speculation might be that low levels of uric acid are continuously synthesized by oxidation of purines and its further oxidation to ureides or urea is necessary for detoxication of Aspergillus. Uric acid can be a very strong antioxidant, but also a pro-oxidant in metazoa and
plants (Ames et al., 1981; Motchnik et al., 1994), and thus mechanisms to control its cellular pools should be necessary.

In contrast to evidence for an ongoing uric acid oxidation in the conidiophore, proline catabolism does not seem to operate in the vesicle, the metulae, the phialides or the conidiospores. This is evidenced by the lack of expression of GFP-tagged proline oxidase (PrnD) and Δ1-pyrroline-5-carboxylate dehydrogenase (PrnC), the two basic proline catabolic enzymes, in these structures (Pantazopoulou, Demais, Scazzocchio, and Diallinas, unpublished). This means that strong PrnB expression in the conidiospores should serve in the accumulation of proline rather than its use as a carbon or nitrogen source. Proline is known for its role as an antistress, particularly antidrought molecule (Hoekstra et al., 2001; Kempf and Bremer, 1998). Does this mean that PrnB-mediated accumulation of proline in the conidiospores operates for protecting proteins from desiccation? Another possibility that should not be dismissed is that these transporters serve functions other than transport. Recently a *D. melanogaster* amino acid carrier was shown to be necessary for insect development through a function other than its capacity for amino acid transport.

### 18.6 Transport in Sex

*A. nidulans* accumulates cell-wall components during vegetative growth and breaks them down during sexual development. *hxtA*, a gene encoding a putative hexose transporter, was isolated from a differential library, which was enriched for sexual-specific genes. Deletion of *hxtA* does not impair growth on a variety of carbon sources nor does it inhibit sexual development, suggesting redundant sugar uptake systems. *hxtA* is repressed under high glucose conditions and expressed in vegetative hyphae upon carbon starvation and during sexual development. Using GFP fusions, it was shown that HxtA is expressed in developing cleistothecia, specifically in ascogenous hyphae. Based on this, it was proposed that HxtA is a high-affinity glucose transporter involved in sugar metabolism during sexual development (Wie et al., 2004). In a recent work, two purine transporters (UapA and AzgA) were also found to be expressed in ascogenous hyphae, Hüll cells, and interconnecting hyphae of the latter (Pantazopoulou et al., 2007; see also Fig. 18.1). Specific transporter-gene expression during sexual development might also prove a key for assigning function in several orphan putative transporters genes in *A. nidulans*.

### 18.7 An Emerging Role of Regulated Trafficking and Endocytosis of Transporters

Recent evidence shows that the need of tight and rapid control of transporter expression is not only operating at the transcriptional or posttranscriptional levels but also at the level of protein ontogeny (Dupre et al., 2004). Transporters are made on ribosomes and directly channeled to the ER membrane through the translocae complex (Meacock et al., 2000; Dalbey and Chen, 2004; Perry and Lithgow 2005). Aberrant transporter folding or transporters foreign to the host are usually retained in the ER and undergo ER-associated degradation (ERAD) (Meusser et al., 2005). This is also associated with an unfolded protein response reaction (Zhang and Kaufman 2004 and references therein). Efficient further transport from the ER is not a default process. ER-exit is dependent not only on proper transporter folding and cis-acting motifs (Nishimura and Balch, 1997; Dominguez et al., 1998) but also trans-acting protein factors (Fromme and Scheekman, 2005). Trans-acting factors include several *sec* genes (e.g., *SEC18* in yeast; Bisson, 1988; Riballo et al., 1995; Beck et al., 1999) but also family-specific ER-resident chaperones. In *S. cerevisiae*, for example, ER-exit of amino acid transporters, hexose transporters or phosphate transporters, requires the function of proteins known as Shr3p, Gfs2p, and Pho86p, respectively (Ljungdahl et al., 1992; Kuehn et al., 1996; Kota and Ljungdahl, 2005). Such specialized transmembrane chaperones were proposed to prevent aggregation of transporter in the ER (Kota and Ljungdahl, 2005). Shr3p functional homologs have been described in *S. pombe* (Martinez and Ljungdahl, 2000), *C. albicans* (Martinez and Ljungdahl, 2004) and also in *A. nidulans* (Eerpenpazoglou et al., 2005). ΔshrA mutants of *Aspergillus* show leaky phenotypes only on proline, glutamic, and aspartic acid as nitrogen sources. GFP studies have shown that in a ΔshrA strain, the proline (PrnB) and the glutamic/aspartic acid (AgpA) transporters have reduced, but
Aspergillus Transporters

still significant, expression in the plasma membrane. A functional ShrA-GFP chimeric molecule confirmed that ShrA is indeed an ER membrane-resident protein (Erpapazoglou et al., 2006) and is now used as a marker for transporter trafficking studies (Vlanti et al., 2006). Post-ER transporter trafficking takes place in vesicles, is directed to the Golgi and then to the vacuole, the peroxisomes, or the plasma membrane. Specific proteins and cofactors that control distinct steps including vesicle budding, transport, docking, and fusion with target membranes regulate vesicle biogenesis. Budding requires an assembly of a coat protein complex on the membrane, membrane deformation, and subsequent cleavage of the nascent vesicle from the donor membrane. Sec proteins bind to other factors and form COPI and COPII complexes (Duden, 2003). This dynamic control of transporters has been, and is, extensively studied in yeast and mammalians cells. In yeast, down-regulation of Gap1p, the general amino acid permease, in the presence of ammonium, and of Fur4p, the uracil permease, in the presence of excess uracil, are excellent examples of regulated endocytosis in fungi. In A. nidulans, similar down-regulation of UapC and UapA by ammonium-induced endocytosis has also been detected recently (Valdez-Taubas et al., 2004; Pantazopoulou et al., 2007). The specialized growth pattern of filamentous fungi likely imposes additional needs on membrane trafficking as compared to yeasts (Harris and Momany, 2004). Interestingly, early genetic work in A. nidulans has led to the identification of mutants with pleiotropically reduced (leaky) capacity for growth on amino acids, purines, or other nitrogen sources (Scazzocchio, unpublished). Some of them, particularly uapB70 and aauZ102, were compatible with compromised uptake of these nutrients (Scazzocchio, Vlanti, Diallinas, unpublished). Such mutants might be affected in proteins involved in transporter trafficking and recycling, and thus it will be interesting to clone them, now that the genome sequence of A. nidulans is available.

18.8 Paradigms of Transporter Structure-Function Analysis in Aspergillus

Genetic, biochemical, and biophysical approaches have tremendously contributed to our knowledge concerning the mechanism of substrate recognition and transport, and on the role of ion coupling in the function of secondary transporters (Kaback, 2005). A. nidulans can provide a unique system, equivalent to other microbial model organisms such as E. coli and S. cerevisiae, to approach the structure and/or function relationships of transporters.

The most studied A. nidulans transporter in respect to the molecular basis that determines aspects of its function is the UapA xanthine-uric acid acid transporter. All aspergilli have close homologs (>70% identity) of UapA and preliminary evidence shows that the A. fumigatus protein has a function very similar to UapA (Goudela, Reichard, and Diallinas, in preparation). UapA historically defined the NAT/NCS2 family (Diallinas et al., 1995) and NAT members are ubiquitous in filamentous fungi. Among the yeasts, there is one gene in C. albicans and S. pombe, but none in S. cerevisiae. Bacterial members (PyrP from B. subtilis or UraA from E. coli) are known to transport uracil (De Koning and Diallinas, 2000).

Work from our laboratory has led to significant conclusions concerning substrate binding and transport, as well as important topological determinants. An approach using chimeric UapA-UapC transporters (as discussed earlier, UapC is a paralog of UapA with distinguishable kinetics) has initially identified a relatively short segment, including two putative transmembrane domains and their connecting loops, which determines the kinetics and, possibly, the specificity of these purine transporters (Diallinas et al., 1998). In the original article, UapA was proposed to have 14 TMS but today, using more sophisticated algorithms and multiple alignment topological programs, it is rather believed that it most probably has 12 TMS, with an extra topologically ambiguous amphipathic α-helix between TMS8 and TMS9 (Koukaki et al., 2005). In that model, the result from chimeric analysis suggests that the region determining UapA (or UapC) kinetics corresponds to a region starting from the loop downstream of TMS8, the following topologically ambiguous amphipathic α-helix, the next loop, and TMS9. Randomly selected or in vitro-constructed mutations within the loop downstream of TMS8 either inactivate UapA or modify its specificity (Diallinas et al., 1998). More interestingly, in the following short loop connecting the ambiguous amphipathic α-helix and TMS9, a highly conserved sequence ([Q/E/P]408-N-X-G-X-X-X-T-[R/K/G])^{17} (the NAT signature motif; Diallinas et al., 1998) was shown to include residues involved in substrate
were characterized and molecular motives important for the movement of nitrate across the membrane (Scazzocchio and Diallinas, unpublished observations). The L-proline transporter, which has no close homologs in any other species of aspergilli (Tavoularis et al., 2004; Kinghorn et al., 2005), proline (PrnB; Tavoularis et al., 2004), and ammonium (MeaA; Monahan et al., 2005) transporters, have close homologs (75–85% identities) in all aspergilli. An impressive exception, however, is the PrnB protein, the highly specific L-proline transporter, which has no close homologs in any other species of aspergilli of known genome (Scazzocchio and Diallinas, unpublished observations).

NtrA and MeaA are homologous to transporters with solved structures (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html), and thus conclusions from mutational analyses can be combined with 3-D data. In a classical genetic screen selecting for chloride resistance, several interesting ntrA mutants were characterized and molecular motives important for the movement of nitrate across the membrane were identified. These include the highly conserved nitrate signature motif (residues 166–172) in TMS5, the absolutely conserved charged residues R87 (TMS2) and R368 (TMS8), as well as the highly conserved aromatic residue F47 (TMS1). The polar sides of TMS5 and TMS8 have been shown to be parts of the substrate translocation pathway in other MFS proteins, and thus R368 might be directly involved in substrate binding and transport (Kaback, 2005). This assumption is in agreement with the fact that replacement R368K is tolerated, but increases the \( K_m \) for nitrate influx from \( \mu \)M to mM values. A similar

Mutational studies have also been performed with transporters specific for nitrate (NtrA; Unkles et al., 2000; Amillis et al., 2001), proline (PrnB; Tavoularis et al., 2004), and ammonium (MeaA; Monahan et al., 2002). NtrA and MeaA, as most studied A. nidulans transporters, have close homologs (75–85% identities) in all aspergilli. An impressive exception, however, is the PrnB protein, the highly specific L-proline transporter, which has no close homologs in any other species of aspergilli of known genome (Scazzocchio and Diallinas, unpublished observations).

The Aspergilli
kinetic change was also seen with R87K, but given the position of TMS2 in the 3-D NrtA structure (by homology threading, not shown), the role of R87 on transport is probably indirect. A second-site suppressor of R87T, that restored the ability to grow on nitrate, was mutation N459K, present in the second copy of the nitrate signature in TMS11. This result is fully reasonable given that in the 3-D NrtA structure, TMS2 and TMS11 are next to each other, and thus, K459 probably bypasses the need for R87. Finally, F47 and other aromatic residues lie on the same side in the TMS1 of NtrA and thus, may either close the translocation pore following binding of substrate, or affect the flexibility of the translocation pore.

In the studies concerning PrnB and MeaA, several of isolated or constructed missense mutations affecting function were mapped in TMSs and the borders of cytoplasmic loops with TMSs. In PrnB, mutations were classified to those affecting function per se, and those affecting topogenesis, based on the topology of PrnB-GFP versions of mutations. Despite failure to obtain mutations altering the specificity of PrnB, an important role of helix TMS6 for proline binding and transport was proposed based on the kinetic profiles of mutations K245L and F248L. Results on PrnB were in line with those from a limited number of analogous studies on yeast amino acid transporters. Several mutations concerning MeaA mapped in a motif (161-GAVAERGR-168) connecting TMS3 and TMS4, which may be important for the translocation of ammonium, and in the conserved P186 (TMS4). Homology threading (not shown) supports these results. Interestingly, mutation G447D, in the C-tail of MeaA, trans-inhibited the activity of not only the endogenous MeaA, but also of the other ammonium transporter, MepA. These results suggest that MeaA may interact with itself and with MepA, although any heterointeraction is not required for ammonium transport function.

18.9 *Aspergillus* as a Novel System for Studying Transporters from Complex Organisms

Heterologous expression systems, such as *Xenopus* oocytes, tissue-culture cells, insect cells, and yeast cells have been used to characterize proteins from complex organisms. Cloning or/and functional analysis of eukaryotic transporter genes by functional complementation in yeast is the most used system (Frommer et al., 1993; Hsu et al., 1993; Chiu and Bush, 1996; Eide et al., 1996; Gillissen et al., 2000; Mäser et al., 2001; Vickers et al., 2001). We have shown that *A. nidulans* can also be used for transporter functional complementation. The function of Leaf Permease1 (LPE1), a protein that is necessary for proper chloroplast development in maize, was characterized by functional expression in an *A. nidulans* mutant lacking all endogenous purine and pyrimidine transport activities (Schultes et al., 1996). In that case, the choice of *A. nidulans*, instead of *S. cerevisiae*, for studying Lpe1 was dictated by the particular genetic and physiological features of purine transport and metabolism in *A. nidulans*, but also by the fact that, unlike the *A. nidulans* UapA and UapC proteins, *S. cerevisiae* has no homologs of Lpe1. In some cases *A. nidulans* transporters are more similar than those of *S. cerevisiae*, in sequence, to plant or metazoan transporters (e.g., the sarco/endoplasmic reticulum Ca\(^{2+}\)-ATPase SERCA2b homolog, responsible for the Darier-White Disease in humans). Other efforts to express several *Arabidopsis*, *Drosophila*, and human NAT homologs, purine transporters from *Trypanosoma* or the human sodium/iodide symporter, are proving problematic (Gournas, Pitis, Kafasla, Erpapazoglou, Billini, Maurino, Tazeby, de Koning, Sophianopoulou, and Diallinas, unpublished). Very recent evidence using GFP chimeras shows that human or *Drosophila* NAT transporters are stably translated but are retained in the ER and other secretory compartments (Gournas, Sophianopoulou, and Diallinas, unpublished). ER retention is further supported by induction of the unfolded protein response, evident at the level of mRNA accumulation of the Hsp70 (Pantazopoulou, Gournas, and Diallinas, unpublished). Aberrant transporter trafficking is the major problem for expressing heterologous transporters in yeast as well (Wieczorke et al., 2003; Flegelova et al., 2006). In general, plant transporters are much more easily expressed in fungi than metazoan transporters. One should not be pessimistic, however. Fungi are unique genetic tools where everything may become possible. Functionality can be achieved by employing several approaches such as modifications of codon usage, truncation of C-terminal regulatory sequences, expression in mutant strains, coexpression with trafficking partners, construction of protein chimeras, various growth conditions, or chemical
cannot bind and activate the transcription of suppressor of (Scazzocchio and Arst, 1975; Arst and Scazzocchio, 1978). This mutation, papers in Nature describing a mutation leading to strong constitutivity for the uric acid-xanthine permease function mutations and the isolation of transport genes. In the 1970s Scazzocchio and Arst published two becoming prototype paradigms for transporter studies. Classical genetics did not only lead to loss-of-1999), including those encoding the first known nitrate and uric acid transporters, both of which are regulation and showed that they were linked to permease with enlarged specificity for other purines (Diallinas et al., 1995). Bailey and Arst (1979) and Arst and MacDonald (1975) isolated mutations defining promoters of permeases were selected. D. Gorton, C. Scazzocchio, and H. Arst (unpublished), using an allele that codes for a version of the GATA-like transcription factor AreA that cannot bind and activate the transcription of uapA. uapA100, which also resulted in constitutivity and an uppromoter effect, was genetically found to be tightly linked to the uapA structural gene whose expression it controlled in the cis configuration. Several cis-acting regulatory mutations for uapA have then been selected as suppressors of areA102 for utilization of uric acid or xanthine. It was proposed that these mutations define the promoter of uapA, including the binding site of two positive transcription factors, the pathway-specific UaY and AreA. This was in fact one of the first formal genetic proofs that positive regulation operates in eukaryotes. At more or less the same time, several other cis-acting regulatory mutations defining promoters of permeases were selected. D. Gorton, C. Scazzocchio, and H. Arst (unpublished), using an areA102 uapA- allele, selected cis-acting suppressors able to grow on uric acid and showed that they were linked to uapC, a gene shown in the 1990s to encode a uric acid-xanthine permease with enlarged specificity for other purines (Diallinas et al., 1995). Bailey and Arst (1979) and Arst and MacDonald (1975) isolated gabI and prnB mutations as suppressors of areA mutations for GABA or L-proline utilization, respectively, and showed that these mutations map tightly linked to the gabA and prnB structural gene encoding GABA or L-proline permeases. The prnB phenotype suggested that these mutations define the binding site of CreA, the general transcription factor mediating carbon catabolite repression. Finally, isolation of the sBo-90 (Lukszkiewicz and Paszewski, 1976) defined the sulfate permease sB promoter. When genetic transformation became available, cis-acting regulatory mutations controlling uapA, uapC, and prnB expression were cloned and shown to define physiological DNA-binding sites for transcription factors UaY, AreA, and CreA. Moreover, the molecular identity of these mutations established the promoter-specific recognition profile of the UaY, AreA, and CreA. The aforementioned highlights from the classical period of transporter genetics do not only present historical importance and a personal taste. They show that microbial genetics is still a powerful approach to understand regulatory circuits, trafficking and structure–function relationships. Combined with reverse genetics, improved methods to perform uptake studies and modern cell microscopy techniques provide a unique tool for investigating the biochemical and physiological role of transporters.
Aspergillus Transporters

The functional expression of any foreign transporter gene in A. nidulans converts this gene to fungal and permits assays unique to simple microbial systems. Good knowledge of A. nidulans transporters and proteins involved in their regulation of expression, trafficking, or function will certainly assist in achieving functional expression of transporter genes from animals or plants. S. cerevisiae has shown the way but is not a panacea for expression and structure-function analyses of transport proteins. The metabolism of filamentous fungi reflects better than the one of “domesticated” S. cerevisiae a generalized eukaryotic metabolism. Some A. nidulans proteins are significantly more similar than S. cerevisiae to transporters of medical importance. The Darier-White disease is due to a defect in the sarco/endoplasmic reticulum Ca$^{2+}$-ATPase (SERCA2b), encoded by the ATPLA2 gene (Vangheluwe et al., 2005). A. nidulans has a unique homolog which is 52% identical to the human SERCA2b. In S. cerevisiae, the closest homolog to SERCA2b shows only 30% identity. Moreover, some A. nidulans transporters are not present at all in S. cerevisiae. An example constitutes the Aspergillus UapA and UapC purine transporters, which are homologous to the vitamin C transporters of mammals (Liang et al., 2001), purine transporters of plants, and other proteins of unknown function in model animals such as D. melanogaster and C. elegans (De Koning and Diallinas, 2000). UapA structure-function mutational analysis has given hints for the molecular determinants that might be critical for purine rather the vitamin C recognition (Koukaki et al., 2005). The Aspergillus NtrA nitrate transporter is homologous to plant nitrate transporters, to a Drosophila protein (malvolio) needed for normal taste behavior, and to the mouse NRAMP-1 protein expressed in macrophages and the nervous system (Rodrigues, Cheah, Ray, Chia, 1995). Thus, A. nidulans transporters are not only important for understanding the fungus itself, but also for basic research or medical and economic applications.

Acknowledgments

I would like to thank all previous and present members of my lab, Sotiris Amillis, Marina Koukaki, Sophia Goudela, Areti Pantazopoulos, Anna Vlanti, Ioannis Papageorgiou, Njimoh Deudonnë Lemuh, Harris Gioule, Stella Tournaviti, Konstantinos Kagiás, Lena Tsiliki, Katerina Roumelioti, Christos Gournas, and Gerasimos Langousis. I would also like to thank Dr. V. Sophianopoulou and members of her lab, Dr. George Zacharioudakis, Dr. Dimitris Hatzinikolaou, Ass. Prof. S. Frillingos, Ass. Prof. Harry De Koning, and Dr. Utz Reichard. I also especially thank A. Pantazopoulou, A. Vlanti, and Prof. Claudio Scaccio for suggestions on this review article. Finally, I am grateful to C. Scaccio for his ideas and constructive criticism in all these years of a unique collaboration.

References


Aspergillus Transports


Aspergillus Transporters


Rodrigues V, Cheah PY, Ray K, Chia W, Malvolio, the Drosophila homologue of mouse NRAMP-1 (Bcg), is expressed in macrophages and in the nervous system and is required for normal taste behaviour. EMBO J. 1995 Jul 3;14(13): 3007–3020.


In all eukaryotes, with the possibly lonely exception of dinoflagelates (1), nuclear DNA is compacted in a structure called chromatin. The first universal level of organization is the nucleosome (see later). Other, higher levels of organization have been described. The next order of complexity is proposed to be the 30 nM chromatin fiber. In this fiber, six nucleosomes would be wound in a solenoidal structure, which would result in a second-level compaction. It has been proposed that the linker histone H1 is instrumental to organize the solenoidal structure. Other levels, like a series of Russian dolls, would lead to the compact structure called the chromosome. A description of different levels of organization can be found in the monographs by Wolfe (2) and Richmond and Widom (3).

Work on model organisms has lead to the discovery and underlining of the similarities found among organisms separated by over a thousand million years of evolution such as fungi, metazoans, and plants. We are, however, convinced that the universality of some obvious features of chromatin hides profound divergences of structure and function at the root of the eukaryotic tree, and that these divergences extend to the different phyla of the fungal kingdom and even within the phylum ascomycetes. There are fundamental differences in chromatin function and organization between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. The heterochromatin of *S. pombe* is not that different from that of *Drosophila melanogaster*; that of *S. cerevisiae* it is a world of its own (4). Perhaps other worlds are there to be discovered. Unfortunately, the experimental work on chromatin of the aspergilli pales when compared with that carried out in other model ascomycetes, *S. cerevisiae*, *S. pombe*, and...
Neurospora crassa, the latter limited to some very specific aspects connected with DNA methylation and heterochromatin (5 and references therein). *S. pombe* could represent an ancestral stage of chromatin organization, evoking perhaps the common ancestor of the different eukaryotic taxa. The similarities found between plant and metazoans may indicate that neither kingdom has diverged very much from this ancestral organization, while other eukaryotic taxa might have diverged more substantially. Many branchings of the fungi have diverged substantially from the *ur-fungus*, the last common ancestor of all fungi, possibly embodied today by *S. pombe*. The substantial number of ascomycete genomes now available (6) allows one to address how this phylum differs from other eukaryotes and may even result in uncovering genus or species differences. A tour of the genomes of the aspergilli has convinced us that some challenging differences are extant even within this genus. It would be impossible, in the limits of space and time of this chapter, to draw an exhaustive picture of the chromatin of the genus *Aspergillus*. While the experimental work may be paltry, three complete different genomes have been published, three more are available on the databases, and an additional two are near completion (7–9). Thus, a search and comparison of all proteins involved in chromatin organization would grow into an unwieldy monograph. Therefore, no completeness and even balance will be attempted. This review is limited to the basic organization of chromatin and to the experimental work extant. The histone modifications (except for some specific aspects relevant to Section 19.3) and heterochromatin are not discussed—the latter will be discussed in a separate publication (Scazzocchio and Ramón, in preparation).

The study of chromatin of the aspergilli was initiated by Ron Morris and his colleagues, who demonstrated that the nuclei of *Aspergillus nidulans* contain the full complement of histones, including the linker histone H1 (10). He demonstrated by micrococcal DNase digestion of chromatin prepared from isolated nuclei, that the nucleosomal repeat is about 154±7 base pairs (bp) compared with 198 bp in rat liver (11). This meant that the stretches of DNA between 146 and 147 nucleosomal cores are rather short. Subsequently, the genes coding for the H2A, H2B, H3 proteins and two isoforms of histone H4 (H4.1 and H4.2) were cloned and sequenced (12,13). No further functional studies were carried out on the chromatin of *A. nidulans*, let alone other aspergilli, until Ramón Gonzalez, then a post doctoral fellow in our laboratory, developed a rapid method of chromatin isolation. We estimated the repeat size to be 159±7 nucleotides, which agrees nicely with the earlier estimate (14).

## 19.1 Basic Nucleosome Structure: Core Histones, Unexpected Occurrences

The nucleosome is an almost universal structure of eukaryotic cells. Around 146–147 bp of DNA are wrapped around an octamer containing two molecules of each of histones H2A, H2B, H3, and H4. The H3/H4 tetramer is the scaffold to which are attached the two dimers of HA2/H2B. All these histones are evolutionarily related and contain a common domain, called the histone fold. Proteins with this domain are present in Archea (15,16). Notwithstanding the strict conservation of the nucleosome, eukaryotes contain a number of variants of the core histones. For both the H3/H4 scaffolds and the H2A/H2B dimer, one molecule, namely, H4 and H2B, respectively, shows striking phylogenetic conservation, while the other is less conserved. Moreover, within the same organism, all nucleosomes contain identical H4 and H2B molecules while H3 and H2A show a number of variants. Within one organism these variants may be non-uniformly distributed in different sections of the genome, and/or may show in multicellular organisms, a cellular or developmental-specific distribution.

### 19.1.1 Histones of the Nucleosomal Core: The Conserved H4

In the aspergilli, as in all ascomycetes, there are two extremely similar H4 histone genes present in the genome. The variation between the two homologs concerns only three amino acids. One of the genes is always transcribed divergently from the canonical H3 gene, as reported previously for *A. nidulans* (13). The second isogene is unlinked. The first variable amino acid (residue 2) is serine in both the linked and unlinked copies in *A. nidulans*, *A. flavus*, and *A. oryzae*, and threonine in both copies of *A. terreus*, *A. niger*, and *A. fumigatus*. Conversely, the other two variable residues are specific for the divergently
transcribed and the unlinked copies in the six species (relevant sequences TFLEG for the divergently transcribed copy, SPLES for the unlinked copy).

In contig 4 in chromosome VIII of *A. nidulans*, between autocalled genes AN0177.3, a protein highly conserved in the ascomycetes, and AN0178.3, encoding a dynein heavy chain, there is a short open reading frame with some clear identities with histone H4, including some completely conserved blocks. There is no evidence as to whether this open reading frame corresponds to a translated protein.

### 19.1.2 Variable H3

In almost all eukaryotes, including basidiomycetes such as *Ustilago maydis* and *Criptococcus neoformans* (16 and our observations), there are two variants of the H3 histone, a canonical form, and the H3.3 variant. In the ascomycetes there is only one form, more similar to the H3.3 than to the canonical H3 (16). The H3 translated sequences are 100% identical in all the sequenced variant. In the ascomycetes there is only one form, more similar to the H3.3 than to the canonical H3 (16).

The second distinguishing motif is in the canonical H3 has, in the amino-terminus, an APTAG sequence. H3.3 has APSTG, which is the sequence found in the H3 of the ascomycetes. The second distinguishing motif is in the amino-terminal residues that can be modified post-translationally (acetylated, methylated, or phosphorylated, 16;) black arrows indicate those that are not conserved in the H3.3 sequence with the H3-like protein from *A. oryzae* (AO090023000662). Above the sequence, gray bars indicate the amino-terminal (αN) and histone fold (α1-α3) conserved α-helical domains. The arrows above the sequence indicate amino-terminal residues that can be modified post-translationally (acetylated, methylated, or phosphorylated, 16;) black arrows indicate those that are not conserved in the *A. terreus* sequence, blank arrows indicate modifiable residues that are conserved. In both panels, the diagnostic sequences for histone H3.3, and the sequences aligned with them are boxed. A clustalW (http://www.ebi.ac.uk/clustalw/) alignment, visualized with boxshade (http://www.ch.embnet.org/software/BOX_form.html) was used.

**FIGURE 19.1** Canonical and aberrant H3 histones of the aspergilli. Top panel, comparison of the canonical H3.3 *Aspergillus* sequence, conserved in all *Aspergillus* genomes (called H3) with the aberrant protein from *A. oryzae* (AO090023000662). Bottom panel, comparison of the canonical *Aspergillus* H3 sequence with the H3-like protein from *A. terreus* (ATEG_04922.1). Above the sequence, gray bars indicate the amino-terminal (αN) and histone fold (α1-α3) conserved α-helical domains. The arrows above the sequence indicate amino-terminal residues that can be modified post-translationally (acetylated, methylated, or phosphorylated, 16;) black arrows indicate those that are not conserved in the *A. terreus* sequence, blank arrows indicate modifiable residues that are conserved. In both panels, the diagnostic sequences for histone H3.3, and the sequences aligned with them are boxed. A clustalW (http://www.ebi.ac.uk/clustalw/) alignment, visualized with boxshade (http://www.ch.embnet.org/software/BOX_form.html) was used.
A second H3 gene is present in both *A. oryzae* and *A. terreus*. The *A. oryzae* translated sequence shows an almost total identity from K58 onward with the canonical H3, but has a completely aberrant amino-terminus (Fig. 19.1). It is obviously the result of a duplication of a section of the H3 gene, fused with an unrelated amino-terminus. This protein is absent from *A. flavus*, which suggests the duplication to be a relatively recent event, as *A. oryzae* is supposed to be a domesticated strain of *A. flavus* (see below). The *A. terreus* additional H3 histone is, on other hand, unique. In this case the amino-terminus is conserved, but different conservative and non-conservative substitutions occur throughout the protein. Figure 19.1 shows a comparison of the common *Aspergillus* H3 sequence with the anomalous H3 histones from *A. oryzae* and *A. terreus*. This novel *A. terreus* protein is neither a typical H3 nor a typical H3.3. The α2 diagnostic motif is more similar to the ascomycetes H3.3 than to the canonical H3. Some residues involved in crucially important modifications are not conserved. Strikingly, this includes Lys 9 and 27, which when methylated are recognized by the chromodomains proteins of the HP1 or polycomb families respectively, leading to the formation of heterochromatin. The substitution of Lys 27 may not be that important, as in spite of its universal conservation in both canonical H3 and H3.3 histones, no methylation of this lysine has been described in the fungi. This is surely not the case for lysine 9, as methylation of this lysine is important in the establishment of heterochromatin in *S. pombe* and *N. crassa* (4,5) and tri-methylation of this residue has been shown experimentally in *A. nidulans* (Reyes, Y. and Strauss, J., unpublished data). No histone similar to the novel *A. terreus* H3 was found in the databases. The presence in only one species, of a protein clearly related to histone H3, but carrying substitutions in many of the residues where modifications occur, is rather extraordinary and calls for in vivo to follow the in silico research.

### 19.1.3 Centromere H3 Histone Variant

In the centromere core, including in the minimal centromeres present in *S. cerevisiae* (see below), the H3 histone is substituted by a specialized protein (16 for review). A gene coding for this protein is present, not surprisingly, in all the aspergilli. The sequence LPFQR in the α1 domain typical of
sequences of the centromeric H3 histone of eight aspergilli. A diagnostic motif of the centromeric histone is boxed by a black line in Figure 19.3. The first loop of the histone fold domain is always longer than in a canonical H3 histone, can be quite divergent, and has been shown to underlie the specificity of the H3 centromeric-histone in Drosophila species (17). The amino-terminus of the CenH3 variants of the aspergilli show a clear divergence, which is consistent with their molecular phylogeny. The proteins of A. flavus and A. oryzae are identical, confirming their very close parentage. Similarly, the proteins of A. fumigatus and Neosartorya fischeri show only four amino acid substitutions throughout the sequence (Fig. 19.3).

Henikoff and coworkers have proposed an evolutionary rationale for the rapid evolutionary divergence of the CenH3 histone. This divergence presents an apparent paradox, as the role of this histone is to interact with proteins of the kinetochore, which is a universally eukaryotic conserved structure (18). Henikoff and coworkers propose that the evolution of centromeres and CenH3 histone resembles a host/parasite coevolution pattern and have proposed the term “centromere drive” to account for it (19). Centromeres of all eukaryotes, with the exception of S. cerevisiae and its relatives, are composed of highly repetitive DNA, which impede assembling of sequences and make the whole centromeric region a “black hole” (20). Pericentromeric and centromeric sequences may be quite different even within one genus. The fact that centromeres are not represented in any of the Aspergillus genomes, the frustrating attempts to walk cosmid libraries towards the centromere of chromosome IV (21) of A. nidulans, the estimated gaps

![](image_url)

FIGURE 19.3 Centromeric histones of the aspergilli. Note: ClustalW (as in Fig. 19.1) alignment of the translated sequences of the centromeric H3 histone of eight aspergilli. A diagnostic motif of the centromeric histone is boxed by a light gray rectangle. The arrows indicate the few residues where Aspergillus fumigatus differ from Neosartorya fischeri. The black line is above the sequences corresponding to loop 1, which is highly variable between different Drosophila species (24).
in the chromosome assembly of *A. fumigatus* (8; S1) and the pericentromeric regions available for the chromosome VI of *Aspergillus fumigatus* (Nieman, W. and Fedorova, N., personal communication), all imply that the centromere complexity typical of most eukaryotes is extant in the aspergilli. The rapid evolution of centromere sequences is shown by the appearance of functionally active neocentromeres in both *Drosophila* and human cells, which bear no sequence similarity to physiological centromeres (22,23). What is being proposed is that the evolution of CenH3 histones is driven by the evolution of centromere DNA. This “centromere drive” is proposed in the context of organisms, which like metazoans and plants, show an asymmetric female meiosis. In these organisms, only one of the four products of meiosis in females will give rise to a fertile gamete. An expansion of the CenH3-binding sequence in a chromosome will attract additional CenH3 molecules, which in turn will attract more kinetochore proteins and microtubules, and this will eventually lead to a meiotic drive resulting in the “expanded” chromosome being preferentially transmitted in the unequal female meiosis. This process will provoke a clear disadvantage at male meiosis, most clearly seen for the X/Y chromosome pair—at male meiosis, the Y chromosome, which cannot be subject to meiotic drive, will be impaired in its ability to bind microtubules, leading to abortive meiosis and male sterility. Any CenH3 allele that restores parity of chromosome segregation will thus be selected. This process will lead to rapid divergence of the sections of the CenH3 histone where mutation will change DNA-binding properties without impairing its basic nucleosome scaffolding function. This has been proposed as a mechanism of speciation, as centromeric DNA and CenH3 histones will be subject to independent and different evolutionary races in isolated populations, leading to cross-sterility (18,19). This scenario makes a clear prediction: in fungi where meiosis is symmetric and all ascospores are equivalent, there should be no rapid divergence of CenH3 histones. Qualitatively, the alignment shown in Figure 19.3 shows a clear divergence of the amino-terminus, while the sequences in the histone fold loop 1 and sequences amino-terminal to it within the histone fold, found to be subject to positive selection in either or both *Drosophila* species (17,24) and in the *Brassicaceae* (25), show a much higher degree of conservation. Indeed, some of the positions subject to positive selection in the *Brassicaceae* (25) are invariant in the aspergilli. In the absence of quantitative substitution data, a very preliminary conclusion would be that positive selection in the histone fold could well be driven by the asymmetry of female meiosis, but the divergence of the amino-terminal tails is not. It must be stressed that while structural considerations permit to make good guesses at the function of residues in the histone fold, the function of the long divergent amino-terminal tail of the CenH3 is unknown. *S. cerevisiae* strains deleted for the CenH3 amino-terminus are not viable, but overexpression of the tailless histone fold domain is sufficient to ensure viability and correct chromosome segregation. It can be concluded that the amino-terminal divergent tail is necessary as a CenH3 assembly factor and becomes redundant once the histone is incorporated into the nucleosome. Overexpression of the histone fold domain would bypass this process simply by mass action (26). *S. cerevisiae* centromeres are about 125 bp, and CenH3 may be compacted in as little as one nucleosome. Thus, it would be unwise to extrapolate this finding to other organisms. However, the aforementioned would imply an interaction of the amino-terminal domain of CenH3 with other nuclear and/or centromeric proteins. The extreme variability of the amino-terminus of the CenH3 suggests a coevolution scenario similar to that proposed to occur between centromeric DNA and the histone fold domains of CenH3 (see above). The *Aspergillus* genomes can provide a useful tool to explore this possibility.

### 19.1.4 Histone H2B

All *Aspergillus* genomes contain one highly conserved H2B homolog.

### 19.1.5 Histone H2A

The H2A sequence found in the ascomycetes (16), including the aspergilli, belongs to what in other eukaryotes, is the H2A.X variant, rather than the canonical H2A. This is characterized by the carboxy-terminus extension, which contains in the ascomycetes, a conserved SQEL sequence (published consensus for the H2A.X variant S Q (E/D) (I/F/L/Y) (16,27; Fig. 19.4). This is particularly interesting,
because phylogenetic analysis shows that H2A.X variants are not monophyletic, having arisen separately from canonical H2s repeatedly in evolution (16,28). Basidiomycetes have both a canonical H2 and an H2A.X variant (our unpublished observations), the loss of H2 seems to be an ascomycete taxonomic marker.

In addition to the H2A.X histones in all Aspergillus genomes, we find a histone H2A variant, the histone H2A.Z. This was not detected either by the biochemical or the cloning procedures employed before the genomic era. Figure 19.4 shows an alignment of the H2A.X and H2A.Z histones of A. nidulans, A. fumigatus, and A. oryzae. Besides the three clusters of difference described by Malik and Henikoff (16), other highly specific differences are extant in the amino-terminus of the proteins. As in other organisms, a cluster of different residues appears in the carboxy-terminal docking domain, which interacts with the H3-H4 tetramer. This region is essential for H2A.Z function in Drosophila melanogaster and cannot be replaced by the cognate domain of the canonical H3 histone (29). Structural studies (30) have shown subtle differences between the docking domains of the canonical H2A histones and the H2A.Z histones. However, these specific differences cannot be extrapolated from vertebrates to ascomycetes. While a crucial glutamine-to-hydrophobic amino acid substitution (marked with a star in Fig. 19.4) is extant in the H2A.Z sequences of the aspergilli, the two histidines that chelate an Mn++ ion in the crystal (30), are not present in the H2A.Z sequences of the aspergilli. Indeed, one of the histidines is present and conserved in the H2A.X sequence (positions also marked with stars). Very little is known about the function and genomic distribution of the variant H2A.Z histone in any ascomycete other than S. cerevisiae. In the latter organism, it tends to be associated with repressed promoters. Genes that depend on this variant histone for their transcription cluster near telomeres, where H2A.Z protects these genes from Sir-mediated silencing. Thus, this histone variant would limit the spread of heterochromatin to subtelomeric genes (31,32). S. cerevisiae silences its telomeres employing a complete different set of proteins than most eukaryotes, from S. pombe to mammals (4). Thus it would be unwise, in the absence of additional experimental evidence, to extrapolate this function of H2A.Z from S. cerevisiae to other, more orthodox, organisms. In fact, in mammals, H2A.Z is reported to be associated with heterochromatin and more specifically, with HP1 (heterochromatin protein 1; 33). Neither in S. cerevisiae or S. pombe is

![FIGURE 19.4](image-url)  
**FIGURE 19.4** Histone H2A variants. *Note:* ClustalW (as in Fig. 19.1) alignment of the H2A and H2AZ histones of A. oryzae, A. fumigatus, and A. nidulans. The H2A histone of the aspergilli is more similar to the H2AX variant than to the canonical H2A. The H2AX carboxy-terminal diagnostic motif is boxed by a light gray rectangle. The arrows indicate the residues where H2A and H2AZ variants differ throughout evolution (16). All these replacements are found in the aspergilli, except for the proline residue indicated with a blank arrow. The carboxy-terminal docking domain is underlined (30). The stars mark three positions where differences between H2AZ and H2A result in altered interactions in the docking domain of the vertebrate histones. (From Suto, R. K. et al., Nat. Struct. Biol. 7, 2000.)
the H2A.Z protein essential, nor, in the former organism, can it substitute the canonical H2A histone (34, 35). The nonessentiality of H2A.Z in the yeasts, is particularly interesting as H2A.Z is present in all eukaryotes where it has been searched for, and is actually more highly conserved than the canonical H2A or the H2A.X variant. The H2A.Z histone is essential in metazoans and in the ciliate *Tetrahymena thermophila* (36).

### 19.2 Linker Histone H1: A Challenging Mystery

The linker H1 histone is almost universally present in eukaryotes. Linker histones do not belong to the same family of proteins as the core histones. They do not show a histone fold. Instead, they are composed of amino- and carboxy-terminal basic, unstructured domains, which bracket a globular domain. The structure of a number of linker histone globular domains has been determined and it is a conserved winged helix motif, different from the histone fold found in the core histones and related to archael proteins. It has been proposed that linker histones have evolved from bacterial, rather than archael proteins (37).

The core histones are wrapped by a 146–147 bp DNA sequence constituting the core nucleosomes, which are separated by a linker DNA of variable length. This length does not only vary among organisms but also between different tissues or different developmental stages of the same organism. H1 histone and its variants are supposed to bind to the DNA between the nucleosome cores (2). Both plants and metazoans have typically more than one linker histone. For a long time it was supposed that both model yeasts, *S. cerevisiae* and *S. pombe* did not posses an H1 histone. This is surely the case for *S. pombe*, an observation that has became even more puzzling in the wake of extensive work on this species, which shows that the chromatin of this organism has a striking resemblance to that of metazoans. However, in *S. cerevisiae*, the failure to find a linker histone was due to the aberrant biochemical properties of the protein (38). The sequence of the genome indicated that a protein, which was likely to be a linker histone was present. This protein shows two contiguous, rather than one globular, domains and it lacks the amino- and carboxy-terminal basic domains. The first domain has a typical winged H1 globular domain structure, while the second domain is unstructured under physiological conditions (39–41). Deletion of the gene results in only marginal apparent phenotypes (38, 42 and references therein). However, it has been shown that the H1 histone of *S. cerevisiae* is inhibitory for the DNA homologous recombination repair pathway, including the recombination pathway of telomere maintenance (43).

Deletion of the linker histone of the ciliate *Tetrahymena termophila* results also in nondramatic phenotypes (44, 45). However, this protein is also aberrant. It contains the two basic tails, but it lacks the central globular domain. It is a moot point whether, as it is commonly accepted, the aberrant linker histones of some protists are phylogenetically related to the “mainstream” linker histones.

We have purified the histone H1 from *A. nidulans*, determined its terminal amino acid sequence and proceeded to clone its cognate gene. The protein shows a canonical domain structure, with a typical globular domain. The sequence of its globular domain is very similar to that of the globular domains of the H1 of *S. cerevisiae*, which strongly suggests that the aberrant H1 present in this organism arose from a duplication of the ascomycete ancestral domain (46). The globular domain of the *A. nidulans* H1 can be easily modeled on, and superimposed to, the structure of the chicken H5 histone (Scazzocchio, C., unpublished data). While some features of the fungal linker histone are metazoan-like, the first intron of the *A. nidulans* sequence (and of *N. crassa*, 47; and other aspergilli, but not of *Ascobolus immersus*, which has lost the first two introns, 48), is in an identical position of a number of plant H1 histones, a feature that is not found in any metazoan H1 gene (46).

We presented strong evidence that there is only one H1 gene in *A. nidulans* (*hhoA*) (46). The gene was mapped by CHEF in chromosome VI, using suitable translocated strains (A. Ramón and C. Scazzocchio, unpublished data). Both these data were confirmed by the genomic sequence. A deletion of the gene has no observable phenotype. The deletion has no effect on growth, conidiation, conidial viability or the sexual cycle, in spite of an H1-GFP fusion protein being clearly visible at mitosis in the *A. nidulans* chromosomes. The deletion strain has the same nucleosomal repeat as the wild type, and chromatin structure is maintained under both nonexpression and expression conditions for the *acnA* (actin) promoter, the
niiA-niaD bidirectional promoter (46), and for the prnD-prnB bidirectional promoter (Ramón, A., Gonzalez, R., and Scaccizocchio, C., unpublished data; see Section 19.3.3 for the chromatin structure of this promoter). Gross overexpression of the gene using the alcA promoter does not result in any obvious observable phenotype, nor does it alter the size of the nucleosomal repeat (Ramón, A. and Scaccizocchio, C., unpublished data). We have also shown that the expression of the H1 gene follows the same pattern as that of the H3 gene (Fig. 19.2). The H1 protein is present in resting conidia as shown by an H1-GFP fusion, and while a low level of the cognate mRNA is seen in conidia, transcription starts, coordinately with that of H3 histone and actin, between 90 minutes and two hours, thus, during the phase of isotropic growth and before the first mitosis (Fig. 19.5). A hhoA-deleted strain is as sensitive as a wt strain to UV irradiation, which suggest that at variance with S. cerevisiae H1(43), it does not have a role in DNA repair mechanisms in A. nidulans. Recently, we had shown that the deletion does not affect either map distances nor the frequency of gene conversion in recombination experiments involving closely linked markers (Hamari, Z. and Scaccizocchio, C., unpublished data). Thus the role (if any!) of the canonical linker histone of A. nidulans remains a mystery.

In the wake of this work, the genes coding for the H1 genes of N. crassa (47) and Ascobolus immersus (48) were cloned and inactivated. In the former, subtle phenotypes are observed, including derepression of the gene coding for pyruvate carboxylase. In the latter, the inactivation of the H1 coding gene affects the accessibility of the chromatin to micrococcal DNase digestion, and results in hypermethylation and in a shortened life span (48,49). The latter is particularly interesting, strains of A. immersus carrying an inactivated H1 gene, show a sudden growth arrest between 6 and 13 days after ascospore germination. An arrest of growth is what is seen in yeast strains lacking telomerase, and thus unable to restore DNA telomeric repeats (50). It would be worthwhile to investigate if in A. immersus H1 is necessary for telomerase recruiting or activity.

All the aspergilli sequenced to date have one typical H1 coding gene. There are several problems worth investigating. Is the distribution of H1 uniform in the genome of the aspergilli? The early observation of Felden et al. (10) that H1 occurs in stoichiometric amounts with other histones is particularly challenging, in view of the absence of any phenotype resulting from the deletion of the gene. Is H1 involved in the differential expression of specific genes or sections of the genome? Obviously, both ChIP and transcriptomic studies should be carried out. What is the evolutionary rationale for the conservation of an apparently useless histone? The results presented suggest that both at the root of the eukaryotic tree and even within the ascomycetes, there has been quite a divergence in the function of the linker histone. Studies with fungi belonging to other phyla than the ascomycetes would be most desirable.

![Graph](image-url)

**FIGURE 19.5** Both histones H1 and H3 genes are transcribed during the phase of isotropic growth. *Note:* Spores of a pabaA1 strain were harvested from 48 hours plates, filtered and inoculated to liquid minimal media (as in Fig. 19.1), and mycelia grown in shaking culture for the indicated times at 37°C. RNAs were isolated and analyzed by Northern blots. Hybridization with a probe corresponding to the 18s ribosomal RNA was used to monitor the loading of the gel. Hybridization to acnA (α-actin) is included to verify the germination state of the conidiospores (94).
19.3 Transcription Factors and Chromatin Structure

There is a substantial body of work on control of gene expression and transcription factors in the aspergilli, mainly in the model species *A. nidulans*. A number of pathways have been thoroughly studied. These include nitrate, purine, proline, ethanol, and acetamide utilization and pH-mediated regulation. Work in organisms ranging from *S. cerevisiae* to mammals have established detailed mechanistic models that connect the control of transcription and the restructuring of chromatin in promoters. It would not be possible to review such work here. On the whole, expression of regulated genes is associated with loss of nucleosome positioning in promoters, while repression is associated with establishment or reestablishment of nucleosome positioning. There are two basic, nonexclusive, mechanisms involved. One is the acetylation-deacetylation of histones H3 and H4, the former carried out by large complexes such as ADA and SAGA, the latter by complexes containing deacetylases. Acetylation is associated with activated states, deacetylation with repressed states (51). The second mechanism is the recruitment by transcription factors of chromatin remodeling complexes, which always contain one protein with ATPase activity. These complexes can mediate a number of nucleosome rearrangements (52). The work in *A. nidulans* has not gone beyond the phenomenological level and is briefly described here. A survey of the transcription factors involved in different regulatory processes in the aspergilli, can be found in the chapter by Mark Caddick in this book (53).

In our laboratory, we have described the chromatin rearrangements extant in three systems, one inducible and nitrogen-metabolite repressible, a second inducible and carbon-catabolite repressible, and a third one that is inducible and repressed synergistically by both the carbon and nitrogen repression systems.

19.3.1 Nitrate Assimilation Gene Cluster

In order to be utilized as a nitrogen source, nitrate must be taken up by the cell by specific transporters (54) and reduced to ammonium by the successive activities of nitrate and nitrite reductases (55,56). We have focused in the 1200 bp bidirectional promoter driving the genes coding for nitrite and nitrate reductases. This promoter is shown in Figure 19.6. Transcription driven from this promoter is induced by nitrate and repressed by ammonium and glutamine (Fig. 19.6). Expression demands the synergic action of two transcription factors, NirA, which is pathway-specific, and the GATA factor AreA, which is inactivated by ammonium and glutamine through a number of concurrent mechanisms (53). Expression is accompanied by a drastic chromatin restructuring of the promoter. Six nucleosomes are positioned in the wild type under nonexpression conditions (Fig. 19.6) and they lose their positioning under expression conditions (57). Four GATA sites are situated in a nucleosome-free region. Of these, site 5 is the transcriptionally most important GATA site, and the only one that can be revealed by *in vivo* methylation protection (57). Site 2 is the most important NirA binding site (58), it is also the only NirA site that can be revealed *in vivo* by methylation protection and it maps in the nucleosome-free region in the boundary of nucleosome-1 (Fig. 19.6) (59,60).

The loss of positioning is independent from transcription and depends strictly on the GATA factor AreA, but not under the physiological conditions tested by us, on the specific transcription factor NirA. However, the latter may not be strictly true under all conditions. In a recent article, and using different physiological conditions and a different assay for nucleosomal positioning, J. Strauss and his coworkers have shown that loss of positioning of nucleosome-1 (Fig. 19.6) depends on NirA (60). Moreover, a NirA-constitutive mutation leads to loss of nucleosome positioning in the absence of the inducer (61).

Work from the laboratory of J. Strauss has also shown that overexpression of NirA can bypass the requirement for AreA for transcriptional activation and thus, presumably also for chromatin restructuring. Interestingly, work from the laboratory of Michel Hynes has shown that nitrogen starvation results in AreA accumulation in the nucleus (62). The early experimental protocol used in our studies involved a transfer from neutral conditions (urea as a nitrogen source) to nitrate as “inducing conditions.” Under these conditions, a mutant in the nirA gene will be nitrogen starved, and would possibly accumulate AreA in the nucleus. Unpublished work from the lab of J. Strauss using ChIP and nuclease accessibility assays,
has indeed found that the two hours nitrogen starvation conditions used in our early studies, lead to loss of nucleosome positioning at the promoter, increased acetylation of histone H3 in the region of nucleosome-1 and to accumulation of DNA-bound AreA, but not NirA, to sites in the nucleosome-free region. Neither of these long-term effects are dependent on NirA or transcription (Berger, H., Böck, S., and Strauss, J. unpublished results).

Being that as it may, the role of AreA in chromatin rearrangements in the \( niiA-niaD \) promoter is firmly established. Mutation of the four AreA sites in the nucleosome-free region results in a strong diminution of transcription of both \( niiA \) and \( niaD \), but does not prevent chromatin remodeling, which implies that other AreA binding sites are competent for the latter (57). A far-fetched alternative is that AreA does not need to be bound to DNA to promote chromatin remodeling, but could be recruited by other proteins, such as NirA. The interaction of the NirA protein with an AreA fragment comprising the DNA binding domain has been shown in vitro (63).

An \( areA \) mutation that results in both constitutive expression of \( niiA-niaD \) and constitutive chromatin rearrangements has been described. The chromatin rearrangements are NirA independent, while the constitutivity is only partially so (63). This mutation maps in the basic carboxy-terminus of the AreA DNA binding domain, and behaves for most of the genes controlled by AreA as a partial, or even complete (as, e.g., for \( uapA \), 63; encoding the main urate-xanthine transporter, see chapter by G. Diallinas, 64) loss-of-function mutation. In fact, it results in strongly diminished in vitro affinity for all the 10 AreA binding sites of the \( niiA-niaD \) promoter, most noticeably for the four binding sites located in the nucleosome-free region, which contribute to about 80% of the transcriptional competence of the promoter (57,63). We have identified mutations in the \( nirA \) gene that are completely analogous; these result in constitutivity of both expression and chromatin rearrangements, which are now AreA independent (65). A bypass of the role of AreA in both the chromatin rearrangement and activation requires at least two mutations in NirA, a constitutive mutation such as that described recently by Benreiter et al. (61), and a second mutation, in a carboxy-terminal basic domain (66 for the sequence of one such mutation, Ramón, A. Strauss, J. and Scaccazzocchio C, unpublished results). The data indicate that NirA requires interacting AreA for in vivo DNA binding and that transcriptional activation needs the synergistic action of both transcription factors. In addition, NirA seems to play a crucial role to initiate transcription-associated

![FIGURE 19.6](image-url)
processes of chromatin remodeling whereas AreA is required for a transcription-independent maintenance of an open chromatin structure associated with increased histone acetylation (59–61,63). Strikingly, very subtle mutations in either partner can result in a bypass for the requirement of the other. This implies that these mutations can result in AreA of NirA proteins than can recruit all that is necessary for both chromatin rearrangements and transcriptional activation. These mutations are priceless in the identification of the factors interacting with NirA and/or AreA, work that is being carried out in the laboratory of J. Strauss.

The niaA-niaD intergenic region is a genuine bidirectional promoter, rather than two juxtaposed promoters driving genes transcribed in opposite directions. This is shown by the fact that mutations in NirA or AreA binding sites affect the transcription of both genes (57,58,63). In A. nidulans the clustering of the nitrate assimilation genes comprises the nitrate and nitrite reductase genes and the gene encoding one of the nitrate transporters. A second nitrate transporter and the nirA regulatory gene are not linked to the cluster. In other filamentous ascomycetes, such as N. crassa, Magnaporthe grisea, and Fusarium graminearum the nitrate assimilation genes, highly homologous to the A. nidulans ones, are scattered in the genome. In the aspergilli the clustering is maintained (58 and our unpublished results). It is particularly interesting that clustering of the nitrate assimilation genes seems to have occurred several times independently in evolution. In Hansenula polymorpha, a yeast able to assimilate nitrate, the homologs of niaA and niaD, a gene encoding a nitrate transporter and two genes encoding transcription factors not homologous to nirA, are tightly clustered, with extremely short intergenic regions (67). Clustering of several genes of the nitrate assimilation pathway also occurs in the alga Chlamydomonas reinhardtii (68).

19.3.2 alc Gene Cluster

Following the unpublished work of Mary Page and David Cove, John Pateman and his colleagues carried out a physiological and genetical analysis of ethanol utilization in A. nidulans, (69). Robin Lockington took this system to the molecular era, cloning the alcA and alcR clustered genes and the aldA unlinked gene (70). The group of Betty Felenbok made it into one of the best-studied systems in fungi, with their work including the determination of a unique mode of binding to DNA of the transcription factor AlcR (71,72). Two enzymes, alcohol dehydrogenase (encoded by alcA) and aldehyde dehydrogenase (encoded by aldA) are necessary for the conversion of ethanol to acetic acid, which can then be used as a carbon source after incorporation into acetyl-CoA. The alcA gene maps within a large cluster of genes in chromosome VI. This cluster comprises alcA, the gene coding for the specific transcription factor alcR, and four other genes, alcO, alcM, alcS, and alcU. These four genes are inducible by ethanol (through its conversion to acetaldehyde, which is the molecule that activates AlcR), but they are not necessary for the utilization of ethanol as carbon source (73,74). The aldA gene maps in chromosome VIII outside the alc gene cluster.

Only those features of the regulation of alcR, alcA, and aldA that are necessary as a background to the understanding of the chromatin restructuring in their cognate promoters are described here. Briefly, the reader is referred to a review and recent articles of Felenbok and coworkers for further details (71,74,75). The gene coding for the positive-acting transcription factor AlcR shows a low basic level of transcription, which is ethanol-inducible and carbon catabolite (glucose)-repressible. The alcA gene is strongly inducible by ethanol and carbon catabolite-repressible. Carbon catabolite-repression acts by a double lock mechanism on alcA; it represses alcR, which encodes the transcription factor essential for alcA transcription, and it directly represses alcA. The aldA gene, on the other hand, is only subject to AlcR regulation, glucose repression occurs only through the repression of alcR transcription (71). The phenomenology of alc regulation is shown in Figure 19.7.

The negative transcription factor, CreA, which mediates carbon catabolite repression, has been identified by purely genetic means by Herb Arst and his colleagues in the 1970s (76,77) and characterized by Kelly and coworkers (78–80). Mutations in creA are derepressed for the expression of the alcR and alcA genes, and indirectly for aldA. The binding sites for AlcR and CreA have been identified in the alcR and alcA (and for AlcR also for aldA) promoters in vitro and in vivo by directed mutagenesis (71).

In the absence of an inducer, irrespective of whether a neutral (0.1 fructose, 3% lactose) or repressing (1% glucose) carbon source is present, an array of positioned nucleosomes can be detected in both the alcA and alcR promoters. The aldA promoter does not show any positioned nucleosomes. Both the alcA
and alcR promoters undergo a drastic chromatin restructuring upon induction, in which nucleosomes lose their positioning. This rearrangement is independent from transcription and strictly dependent on the AlcR transcription factor. Of the AlcR 821 residues, at most the first 241 are required for transcriptional activation and chromatin rearrangements. Upon glucose repression in the presence of an inducer, we see a pattern of partial nucleosome repositioning. This is partial in two different ways, some nucleosomes being positioned while others are not, and the positioning is, for some individual nucleosomes, partial (81). The MNase digestion pattern obtained in the positions occupied by these nucleosomes is exactly what would be obtained by superimposing the nonpositioned with the positioned patterns. This has been interpreted as a metastable positioning of each nucleosome in each nucleus rather than as an heterogeneity of nuclei with and without positioned nucleosomes (81). Figure 19.7 shows the patterns of rearrangements extant in the alcR promoter.

19.3.3 prn Gene Cluster

A. nidulans can utilize a number of amino acids as sole carbon and nitrogen sources. The more thoroughly studied pathway of this class is that of proline utilization. The group of Herb Arst described in detail the genetics and physiology of this pathway (82). In A. nidulans, a cluster in chromosome VII, shown in Figure 19.8, comprises all the genes involved in proline utilization. prnA encodes the pathway specific transcriptional activator, prnX encodes a gene of unknown function, which, however, shares the same control system with the other genes of the cluster, prnD encodes proline oxidase, prnB encodes the specific proline transporter and prnC encodes the Δ1′-pyrroline-5-carboxylate dehydrogenase. The pathway is identical in S. cerevisiae and indeed in every organism where it has been studied, however, the cognate genes are scattered in S. cerevisiae and there are substantial differences in the pattern of regulation between the two model ascomycetes. The pattern of clustering is variable in the aspergilli and will be discussed elsewhere (83, Demais, S. and Scaggzocchio, C., unpublished results).
The prnD, prnB, prnC, and prnX genes are subject to proline induction, mediated by PrnA. They are also subject to metabolite repression. However, efficient repression is only achieved when a repressing carbon and a repressing nitrogen source are present simultaneously (Fig. 19.8). Metabolite repression acts directly only on the prnB gene, other genes are repressed due to inducer exclusion, resulting from the repression of the transporter gene prnB (84). Strains carrying a prnB deletion show a residual uptake of proline (85,86). In these strains, the prnD and prnC genes are still fully inducible. This implies that other amino acid transporters, such as GAP (general amino acid permease) can take proline and must also be repressed and/or inhibited drastically to account for the efficient inducer exclusion seen under fully repressing conditions (84). In fact, one such amino acid transporter, which we have studied recently, is exquisitely sensitive to nitrogen metabolite repression (86).

The prnD and prnB genes are transcribed divergently from an intergenic region of 1.7 kb. This region is quite complex. It acts as a bidirectional promoter in relation to proline induction mediated by the PrnA transcription factor. Mutating both the PrnA-binding sites (Fig. 19.9), which can be revealed by in vivo methylation protection (87), virtually abolish proline induction for both the prnD and prnB transcripts (García, I., Gómez, D. and Scaccoziochi, C., unpublished results). The GATA factor AreA (53,76,88) is necessary for expression only in the presence of glucose, when CreA, the protein responsive to carbon catabolite repression, is bound to both the sites shown in Figure 19.9 (89–92). Of the 15 GATA-binding sites present in the prnD-prnB intergenic region, only sites 13 and 14, shown in Figure 19.8, are important for prnB transcription (92). Figure 19.9 shows the structure of the prnD-prnB intergenic region, with all the relevant binding sites and the different chromatin structures described in the text.

Under noninduced conditions, in the absence of proline, whether in the presence of repressing carbon and nitrogen sources or not, the intergenic region shows a closed chromatin structure with eight clearly positioned nucleosomes. Upon induction, under nonrepressing conditions, a massive restructuring occurs, in which the eight nucleosomes lose their positioning and a new nucleosome is placed in between the

![Figure 19.8](image-url)  
**FIGURE 19.8** Regulation in the *prn* gene cluster. *Note:* Upper panel, the proline gene cluster and the proline utilization pathway. SPC, Δ¹-pyrroline-5-carboxylate, GSA glutamic semialdehyde. These two compounds can be converted nonenzymatically to each other, and they presumably are in equilibrium within the cell. Lower panel, the phenomenology of *prnD* and *prnB* expression. NI, absence of induction in the presence of neutral nitrogen and carbon sources (0.1% fructose, urea), I, induction by proline in the presence of neutral nitrogen and carbon sources, IG induction in the presence of glucose, IN, induction in the presence of ammonium, IR, induction in the presence of both glucose and ammonium. (From González, R. et al., *EMBO J.*, 16, 1997 and García, I. et al., *Eukaryot. Cell*, 3, 2004.)
positions previously occupied by nucleosomes +1 and +2. The PrnA binding sites are in a short nucleosome free sequence, and the restructuring relieves the occlusion of the CreA and AreA physiologically relevant binding sites and of the TATA box (Fig. 19.9) (93).

Upon repression, which only occurs when both a repressing nitrogen source and a repressing carbon source are present (e.g., ammonium and glucose, see bottom panel of Fig. 19.8), a pattern of partial positioning is seen (Fig. 19.9). We have determined that nucleosome loss of position is independent from transcription and strictly dependent on the specific transcription factor PrnA. The rearrangement seen under conditions of repression in the presence of proline, is, as described earlier for the _alcA_ and _alcR_ promoters, strictly dependent on CreA (93).

While the _prnD_ gene is specifically induced by proline, the _prnB_ gene is subject to two additional alternative controls. Its transcription is activated, in common with a number of other transporter genes, during the isotropic phase of conidial germination, independently of the specific induction system. It is, moreover, activated by amino acid starvation (94). Activation of transcription by amino acid starvation results in a chromatin- restructuring pattern radically different from that resulting from proline induction. Under conditions of amino acid starvation, only nucleosomes +3 and +4, those proximal to the _prnB_ start of transcription, lose their positioning (Fig. 19.9). Induction by amino acid starvation depends most probably on the homolog of GCN4, CpcA, since a mutation in a putative GCN4 (CpcA)-binding site in the _prnD-prnB_ region prevents this induction. Thus, two different transcription factors, acting independently at different sites in the same promoter, result in radically different chromatin restructuring.

![Diagram showing the chromatin structure of the _prnD-prnB_ bidirectional promoter under different conditions.](image)

**FIGURE 19.9** Chromatin structure of the _prnD-prnB_ bidirectional promoter under different conditions. (I) The default structure of chromatin in this region, in the absence of induction, irrespectively of whether nitrogen and/or carbon sources are neutral (NI, noninduced) or repressing (R). (II) Chromatin structure seen upon induction, in the presence of neutral nitrogen and carbon sources (I), in the presence of only a repressing carbon source (IG), in the presence of only a repressing nitrogen source (IN). It is also found under IR conditions in the presence of trycostatin A (TSA). (III) Chromatin structure seen in induced cultures in the presence of repressing nitrogen (ammonium) and carbon (glucose) sources (IR). (IV) Chromatin structure upon induction of _prnB_ by amino acid starvation. (V) Chromatin structure seen under noninducing conditions (as in NI) in the presence of trychostatin A. Blank ovals, fully positioned nucleosomes, gray shaded ovals, partially positioned nucleosomes. Only physiologically relevant transcription factor binding sites are shown. ( Adapted and redrawn from García, I. et al., *Eukaryot. Cell*, 3, 2004.)
We have also studied the chromatin structure of the prnD-prnB intergenic region under different conditions in the presence of trichostatin A, an inhibitor of histone deacetylases. The inhibitor does not affect induction. In the presence of the drug and the absence of induction, chromatin is in a partially open configuration, but no transcription occurs. Under induced conditions, there is no difference in chromatin structure in the presence or absence of the drug. The most striking difference is found under repressed conditions in the presence of an inducer, where any restructuring fails to occur, and the chromatin structure is identical to that seen under induced conditions (Fig. 19.9). The drug leads to partial derepression of the prn genes. These experiments strongly suggest that histone acetylation is involved in the opening of chromatin mediated by PrnA and deacetylation in the restructuring mediated by CreA. Results obtained recently by ChIP (chromatin immunoprecipitation) indicate that acetylation of histone H3 is correlated with nucleosome loss of positioning in the prnD-prnB promoter (Reyes, Y., Narendja, F., Berger, H., Gallmetzer, A., Fernández-Martín, R., García, I., Scazzocchio, C., and Strauss, J., unpublished results).

The three systems described earlier show different degrees of gene clustering, a feature that is not universal for the same pathways across the ascomycetes phylum. The nature of the selective pressures that result in dispersal or clustering of homologous genes involved in any single pathway is an old and unsolved problem. It is befitting to this chapter to propose that the chromatin structure at a level above higher order structure has a role in the prn gene cluster. Some old data of Herb Arst (95) indicate an action at a distance of the prnD-prnB region on the expression of the distally located prnC gene. We have recently found that this is exactly the case, the PrnA binding sites of the prnD-prnB region are necessary for optimal expression of prnC. Moreover, prnX is proline inducible, but its promoter does not include any PrnA binding sites. Again, this role is carried out by PrnA binding sites in the prnD-prnB intergenic region, which is thus not only a bidirectional promoter but also an enhancer for prnC and prnX (83 and Demais, S., Gómez, D., and Scazzocchio, C., unpublished).

### 19.3.4 Mechanism of Chromatin Rearrangements: Toward a Genomic Approach

The availability of complete genomes and of several thoroughly studied promoter regions permit to investigate the mechanism(s) by which transcription factors elicit chromatin rearrangements. This can be investigated by deletion or conditional expression of presumed actors belonging to the acetylation and deacetylation complexes and of the nucleosome rearrangements ATP-dependent complexes (see previous section). It can be coupled to ChIP techniques recently adapted to *A. nidulans* (61 and Fernández-Martín, R., Gallmetzer, A., Reyes Domínguez, Y., García, I., Scazzocchio, C., and Strauss, J., unpublished results). Similar ChIP work is been carried out in the laboratory of N. Keller (Bok, J-W., Shwab, E. K., and Keller, N.P., unpublished work).

The *A. nidulans* deacetylases have been studied by the group of S. Graessle. Histone deacetylases have been classified in a number of phylogenetically related classes (96–98). Three classes are present in fungi and metazoans, while a fourth class is exclusive of plants. Classes I and II are mechanistically and phylogenetically related, while class III, comprising the Sir2 protein of *S. cerevisiae* are NAD-dependent deacetylases and are phylogenetically unrelated to the other two classes. Graessle and coworkers have defined two class I and two class II deacetylases in the genome of *A. nidulans* (99,100). They showed that the putative ortholog of the atypical class II deacetylase HOS3, HosB, possesses deacetylase activity *in vitro*, and that this activity is trichostatin A resistant. As the positioning of nucleosomes in the prnB-prnD bidirectional promoter (see earlier) is trichostatin A sensitive (93), we can assume that HosB is not the major actor involved in this process. The major histone deacetylation activity in *A. nidulans* is carried out by HdaA, a typical class II protein (100). The only phenotype described for hdaA deletion strains is a greatly increased sensitivity to oxidative stress, implying that the one or more enzymes involved in resistance to oxidative stress are down-regulated in the hdaA deletion. This was found to be the case for *catB*, encoding a stress inducible mycelial catalase. It would be most interesting to investigate the deacetylase-deleted strains for the chromatin rearrangements in the promoters described earlier in this section.
Some work involving the deletion of actors in chromatin rearrangements has been carried out in collaboration with Joseph Strauss. Deletion of the only A. nidulans homolog of ADA2, adaB, encoding a conserved partner of GCN5 in acetylation complexes such as ADA and SAGA (101,102), results in a gross morphological alteration of the fungus, including specific modifications of the conidiophore. We have studied the effect of this deletion on the expression of a number of promoters. Most noticeably, the histone H3 acetylation of the region covered by nucleosome 2 in the prnD-prnB bidirectional promoter (see Section 19.3.3 and Figs. 19.8 and 19.9) is strongly decreased. In spite of this, the inducibility of this promoter is not affected, and paradoxically, we observed a partial derepression when the mutant strain is grown under induced-repressed conditions (as in Fig. 19.8). This is coupled with an inability to reposition nucleosomes upon repression, in the deleted strain. We have observed a similar pattern of events in the alcA and alcR promoters. The deletion of genE, the ortholog of GCN5, results in very similar phenotypes, including the partial derepression of some carbon catabolite repressible promoters and the same defects in the conidiation process (Reyes-Domínguez, Y., Neredja, F., Berger, H., Gallmetzer, A., Fernández-Martín, R., García, I, Scazzocchio, C., and Strauss, J., unpublished results). The results summarized here imply that, as in organisms from S. cerevisiae to mammals, the association of ADA2 and GCN5 is conserved in A. nidulans, but the function of the complex is not necessarily maintained within the ascomycetes.

In S. cerevisiae, several repression processes, including carbon catabolite, are mediated by the TUP1/SSN6 complex (103). This complex acts by recruiting deacetylase HdaI (a member of the class II group of deacetylases, see earlier), and in turn, deacetylated H3/H4 histone would recruit additional molecules of the complex, resulting in occlusion of the TATA box, and direct repressive interactions with components of PolII (103). As carbon catabolite repression has been thoroughly studied in A. nidulans and shown to involve chromatin restructuring (see Sections 19.3.2 and 19.3.3), it is interesting to investigate the role of the TUP1/SSN6 complex in this organism. The TUP1 homolog of A. nidulans (rcoA) has been deleted. The deleted strains show a drastically altered morphology. The RcoA protein plays a crucial role on asexual development and secondary metabolite synthesis, and is necessary for the sexual cycle (104,105). The protein does not seem necessary for carbon catabolite repression (104). We have confirmed the latter results for the alcA, alcR, and prnD-prnB promoters. Interestingly, the deletion of the rcoA gene alters drastically the nucleosome-positioning pattern in the alcR (but not in the alcA) and prnD-prnB promoters. In the deleted strain some nucleosomes are not positioned even in the absence of induction, and repositioning upon repression is partial for the prnD-prnB promoter, and does not occur for the alcR promoter. Thus the rcoA deletion uncouples carbon catabolite repression from its concomitant nucleosome repositioning (García, I., Mathieu, M., Felenbok, B., and Scazzocchio, C., unpublished results). Close homologs of rcoA are present in all the sequenced Aspergillus genomes, including the genomes of A. clavatus and Neosartorya fischeri. We then wanted to address whether the Ssn6p partner has a similar role to that of RcoA. We identified an SSN6 homolog in the genome of A. nidulans and extremely conserved orthologs in all the aspergilli. Differently from S cerevisiae, a deletion of the A. nidulans SSN6 homolog (to be called ssnF) is lethal. This essential function has been shown very recently also for the S. pombe homolog (106). This implies that whether or not the A. nidulans RcoA acts in a complex with SsnF, the latter has specific functions not shared by RcoA (Mathieu, M., Nikolaev, I., Felenbok, B., and Scazzocchio, C., unpublished results).

These scattered observations imply that when investigating the specific functions of chromatin modifying or interacting proteins, it is unwise to rely on only one model, and that further systematic experimental work is needed to determine the function of these proteins in the aspergilli. The availability of the genomes, of novel techniques of gene substitution that can be used to obtain conditional expression of essential genes (107–110), and the development of ChIP (61), make this systematic approach feasible.

### 19.4 Conclusion

We hope that the work presented in this review will result in genomic experimental approaches to a number of chromatin-related problems. The role of different histone variants, the presence of surprising H3 paralogs, the fact that functions of some chromatin-associated proteins cannot be predicted from the
roles found in other model ascomycetes, the availability of many genomes of the same genus, should stimulate a systematic approach to chromatin structure and function within the genus *Aspergillus*.

### 19.5 Databases for the *Aspergillus* Genomes

- **A. nidulans**: http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/Home.html
- **A. fumigatus**: http://www.tigr.org/tdb/e2k1/afu1/
- **A. oryzae**: http://www.bio.nite.go.jp/dogan/MicroTop?GENOME_ID=ao
- **A. terreus**: http://www.broad.mit.edu/annotation/genome/aspergillus_terreus/Home.html
- **A. flavus**: http://www.aspergillusflavus.org/genomics/
- **A. niger** (DoE JGI sequence): http://genome.jgi-psf.org/Aspni1/Aspni1.home.html

Many sequences from *A. clavatus* and *Neosartorya fischeri* have been entered in the NCBI site and can be recovered by appropriate blasts: http://www.ncbi.nlm.nih.gov.gate1.inist.fr/BLAST/

### Acknowledgments

We thank Christine Drevet for preliminary work on histone genomics, including the detection of the *A. oryzae* H3 aberrant histone. We thank all colleagues who allowed us to quote unpublished data, and Joseph Strauss for critical reading of a part of the manuscript. Claudio Scazzocchio thanks the Institut Pasteur, Montevideo and the Department of Molecular Microbiology and Infection, the Flowers Building, Imperial College London for office and computer facilities.

### References


20

Transposable Elements and Repeat-Induced Point Mutation in Aspergillus nidulans, Aspergillus fumigatus, and Aspergillus oryzae

A. John Clutterbuck, Vladimir V. Kapitonov, and Jerzy Jurka

CONTENTS

20.1 Introduction ................................................................. 343
20.2 Spectrum of Transposable Elements in the Three Species ........................................ 345
20.3 Repeat-Induced Point Mutation ........................................ 347
20.4 A+T Content ................................................................. 348
20.5 Clustering and Fragmentation ........................................... 349
20.6 Transposon-Related Genes in Aspergillus nidulans ........................................... 351
20.7 Transcription from TEs and Transposition ........................................ 352
20.8 Implications for Genome Expansion, Sexuality, and Chromosomal Rearrangement .... 352
20.9 Conclusions ................................................................. 353
Acknowledgment ...................................................................... 354
References ............................................................................ 354

20.1 Introduction

The release of genome sequences for three Aspergillus species provides an opportunity to examine transposable elements (TEs—here used interchangeably with “transposons” and “retrotransposons”) in three related organisms, treating these genomes as historical records of successive waves of TE proliferation and subsequent decay. Only a small proportion of these relatively compact genomes are made up of TEs, but despite this, all three include a wide spectrum of TE types. There is strong evidence, in all three fungi, of a destructive process specific for repeated sequences, namely repeat-induced point mutation (RIP), now widely reported in filamentous ascomycetes.
20.1.1 TE Detection

Genome sequences of *Aspergillus nidulans*, *Aspergillus fumigatus*, and *Aspergillus oryzae* [1–3], were obtained from the respective sequencing centers: The *A. nidulans* sequence is available from the Aspergillus Sequencing Project Broad Institute of MIT and Harvard (http://www.broad.mit.edu/annotation/fungi/aspergillus/). Initial analyses of *A. fumigatus* and *A. oryzae* genomes were done on preliminary releases containing ~98% of the later versions. Data in Table 20.1 and supplementary Table 4 have subsequently been updated to correspond with currently available versions, obtainable from the Aspergillus web site, University of Manchester (http://www.aspergillus.org.uk/indexhome.htm) and the Japanese National Institute of Technology and Evaluation (http://www.bio.nite.go.jp/dogan/Top). All three genomes are also deposited with GenBank/NCBI.

TEs were identified in all three genomes using standard methods [4]; briefly, these involved screening genomes for known transposable element (TE) sequences from Repbase Update [5] using Censor; recognized sequences were then extended to identify and compile consensus sequences for specific elements, in some cases encoding typical TE-specific transcripts or displaying TE features such as the long terminal repeats (LTRs) of appropriate retrotransposons, inverted terminal repeats of some DNA transposons, and target site duplications. More elements were identified by further rounds of BLAST searching using consensus sequences as queries, and by examining interruptions in identified elements and gaps between them. TE nomenclature followed the standardized system outlined in Repbase [5]. It should be noted that these search procedures are not exhaustive: unusual elements, and families represented by only a few copies, may be missed. While establishment of a consensus sequence for a young family is not difficult, older degraded elements may be lumped together although they were derived from different families. Alternatively, some old degraded fragments may be assigned to a similar, but younger family. Consensus sequences of TE families have been deposited in Repbase Reports and are listed, along with TE locations in the three genomes, in supplementary Tables 1–4.

### TABLE 20.1

<table>
<thead>
<tr>
<th>Genome</th>
<th>TE Superfamily</th>
<th>Number of Families</th>
<th>Total Copies</th>
<th>Intact Copies</th>
<th>Total Kb</th>
<th>% of Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Copia</td>
<td>4</td>
<td>64 (+200(^a))</td>
<td>3</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>31 Mb</td>
<td>Gypsy</td>
<td>3</td>
<td>73 (+78(^a))</td>
<td>15</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td></td>
<td>solo LTR(^c)</td>
<td>6</td>
<td>227</td>
<td>—</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I (LINE)</td>
<td>1</td>
<td>84</td>
<td>13</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mariner</td>
<td>10</td>
<td>183</td>
<td>52</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hAT</td>
<td>3</td>
<td>211</td>
<td>52</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DNA (MUDR?)</td>
<td>4</td>
<td>70</td>
<td>14</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helitron</td>
<td>2</td>
<td>28</td>
<td>2</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>1218</td>
<td>151</td>
<td>795</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Copia</td>
<td>2</td>
<td>29 (+35(^a))</td>
<td>2</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>29 Mb</td>
<td>Gypsy</td>
<td>7</td>
<td>190 (+115(^a))</td>
<td>35</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I (LINE)</td>
<td>1</td>
<td>31</td>
<td>16</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mariner</td>
<td>8</td>
<td>195</td>
<td>77</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>595</td>
<td>130</td>
<td>821</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>Gypsy</td>
<td>3</td>
<td>91 (+108(^a))</td>
<td>8</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>37 Mb</td>
<td>Solo LTR(^c)</td>
<td>2</td>
<td>78</td>
<td>—</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I (LINE)</td>
<td>6</td>
<td>121</td>
<td>14</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SINE</td>
<td>2</td>
<td>28</td>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mariner</td>
<td>11</td>
<td>399</td>
<td>54</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>857</td>
<td>81</td>
<td>453</td>
<td>1.2</td>
</tr>
</tbody>
</table>

\(^a\) ≥ 95% of consensus length.

\(^c\) Solo LTRs of identified elements. Attached LTRs are counted along with their internal portions.

\(^c\) Solo LTRs of unknown elements.
20.2 Spectrum of Transposable Elements in the Three Species

Table 20.1 and Figure 20.1 summarize the spectrum of TEs found; see supplementary files TE-S1–4 for TE sequences and genomic sites for each fungus.

The first important point to make is that while all three species contain a wide variety of elements from many of the known TE families, there are no identical elements in common among the three species, i.e., although these fungi are assigned to the same morphological genus, there appears to have been a complete turnover of TE content during the considerable time [1] since their divergence. It is, of course, possible that unrecognized remains of common elements may exist among the degraded fragments present in all species. New TE families may have arisen by mutation within each organism, or by evolution elsewhere and horizontal transfer [6].

*A. nidulans* contains the largest number and greatest diversity of elements, but a similar proportion of the genome of *A. fumigatus* is composed of TEs since it includes more intact large elements. *A. oryzae* has fewer recognizable TEs and many of these are highly degraded, resulting in a considerably smaller proportion of the genome identified as TE-derived. The precise number of elements or fragments of elements found depends on the degree to which sequencing has extended into telomeric and centromeric regions that contain large numbers of TE fragments.

### 20.2.1 TEs in Aspergillus nidulans

The most prominent TE families in this genome, both in terms of element size and number of copies, are *Gypsy-I_AN* and *I-1_AN* retrotransposons (LTR and nonLTR, respectively). Both include many intact copies, some of which have evidently been subjected to repeat-induced point mutation (RIP; see later), the results of which are seen as reduced C+G content in Figure 20.1a. Other notable TE components are Helitrons [7], and a considerable number of smaller elements, not included in Figure 20.1a. These include 505 solo LTRs, the result of excision of the body of retrotransposons, such as *Gypsy* or *copia*, by recombination between LTRs at each end. Some of these can be identified with full-length elements, while others are orphans. There are also two families of short nonautonomous *hAT* DNA transposons (*hAT-N1_AN* and *hAT-N2_AN*), totaling 187 elements in all. Only two full-length copies and some fragments, all degenerated but including traces of *hAT* transposase, remain, of an autonomous *hAT* element. This element might have been responsible for transposition of the nonautonomous elements but does not resemble them closely enough to suggest that they were derived from it. As with the other two species, the *A. nidulans* genome includes a variety of *Mariner* DNA transposon families, most of which include both young (i.e., undergraded) copies and copies affected by mutation, including RIP. In addition, though unlike the other species, *A. nidulans* also has a number of elements in four unidentified DNA transposon families, possibly belonging to the *MuDR* superfamily. One of these, *DNA-3_AN*, has also been reported as the *MATE* element, fragments of which promote plasmid replication [8].

Two other elements have been reported earlier: cosmid SW06E08, sequenced by Kupfer et al. [9] included a *fot/pot*-like element, including intact transposase ORF F2P08. This element agrees exactly with that of the *Mariner-6_AN* consensus reported here. However, the genomic version of this particular copy, in contig 1.14, differs from its consensus by four single base insertions that disrupt the ORF, although the remainder of the 38.8 kb cosmid sequence is identical with the genomic version (S.A. Osmani, personal communication). This points to a need to be wary of genome compilation errors for repeat sequences, although single base insertion differences from the consensus do not appear to be particularly widespread in other elements.

Another report, by Nielsen et al. [10], described a single *Dane* element in two cosmids representing a repeated chromosomal segment. The repeated region turns out to be subtelomeric, part of a complex of segmental repeats, fragments of which are found on a number of chromosomes. No other copies of the *Dane* element are found elsewhere in the genome. Subtelomeric regions are underrepresented in the Broad genomic sequence (see Chapter 5 by John Clutterbuck and Mark Farman), which contains no sequence restriction pattern corresponding to that of the second cosmid analyzed by Nielsen et al. *Dane* is distantly similar to the internal portion of *Gypsy-I_AN* and encodes remnants of the *Gypsy* polyprotein.
Despite its apparent age, it is bounded by LTRs with matching target site duplications, which are also found as 23 solo copies, here designated \textit{Gypsy-3\_AN\_LTR}.

### 20.2.2 TEs in \textit{Aspergillus fumigatus}

\textit{Afut1\_AF} and \textit{Afut2\_AF} are previously described families of \textit{Gypsy} LTR retrotransposons [11,12] that make up a considerable proportion of the TE population of this fungus. Both families are evidently ancient, most copies diverging considerably from their respective consensus sequences. Two \textit{copia}, five other \textit{Gypsy} families and one “\textit{I}” nonLTR retrotransposon family are similarly diverse. In contrast to this, out of nine \textit{Mariner} DNA transposon families, four consist almost entirely of intact copies differing little from their respective consensus sequences, suggesting recent proliferation. However, each of these families also has assigned to it a number of short segments (Fig. 20.1b), mostly with reduced similarity to the consensus. It is likely that these are evidence of a much older generation of \textit{Mariner} transposons, now seen only as degraded fragments.

A recently reported member of the \textit{Mariner} superfamily, \textit{Taf1} [13], is here annotated as \textit{Mariner-4\_AF}. Both \textit{A. fumigatus} and \textit{A. oryzae} genomes include small numbers of long \textit{Mariner-L} elements, bearing one or more ORFs in addition to the normal transposase.

### 20.2.3 TEs in \textit{Aspergillus oryzae}

The \textit{A. oryzae} genome contains a relatively small proportion of DNA recognizable as transposon-derived. Furthermore, many of those elements identified are highly fragmented and depleted of G+C content, suggesting extensive RIP (Fig. 20.1c). Exceptions to this picture are a few families (\textit{Gypsy-2\_AO}, \textit{I-4\_AO}, \textit{I-5\_AO}, \textit{Mariner-3\_AO}, and \textit{Mariner-L1\_AO}), all of whose full-length copies agree closely with their respective consensus sequences.

\textit{A. oryzae} is also the only one of the three fungi carrying a substantial number of \textit{SINE3} retrotransposons employing pol III promoter derived from 5S rRNA. These form two families: \textit{SINE3-1\_AO} is short (206 nt), while \textit{SINE3-2\_AO} is mainly represented by overlapping fragments conforming to a 3294 nt consensus. \textit{A. nidulans} and \textit{A. fumigatus} also contain only small numbers of \textit{SINE3} fragments, but these are insufficient to compile consensus sequences. Previously \textit{SINE3} elements have been found only in the genomes of zebrafish and red flour beetle (V.V. Kapitonov, unpublished).

\textit{Tao1}, belonging to the \textit{Mariner} superfamily has previously been published under GenBank accession number AB021710, and RIP has been studied in transposase sequences corresponding to this element [14].

### 20.3 Repeat-Induced Point Mutation

RIP was discovered in \textit{Neurospora crassa} [15], but has now been demonstrated or inferred in other filamentous ascomycetes, for example [14,16,17]. It results in multiple C$\rightarrow$T transitions (and consequently G$\rightarrow$A transitions on the opposite strand) in repeated DNA. In \textit{N. crassa}, RIP occurs to repeated sequences longer than 400 nt at the premeiotic dikaryon stage of sexual reproduction, and is dependent on a specific DNA methylase encoded by the \textit{rid} gene [18], homologs of which, labeled \textit{dmtA}, are found in all three \textit{Aspergillus} species considered here [17]. RIP has not been experimentally demonstrated in \textit{Aspergillus}, but multiple C$\rightarrow$T transition mutations in repeat elements have been taken as evidence for its occurrence [8,11].

RIP-specific methylated DNA has not been detected in \textit{N. crassa}, suggesting that the Dmt enzyme is also a deaminase, but it should be noted that in \textit{N. crassa} RIP-affected sequences can subsequently be methylated in vegetative mycelium by a different enzyme encoded by \textit{dim-2} [19,20]. No methylated DNA has been found in \textit{A. nidulans} [21], but vegetative mycelium of \textit{A. oryzae} has been reported to contain very low levels of 5-methylcytosine [22].

\textit{In silico} detection of the effects of RIP depends on examination of families of repeats consisting of RIP-affected copies, along with enough unmethylated or lightly mutated copies to provide a reliable consensus
for the original sequence. Such families are abundant in *A. nidulans* and *A. fumigatus* but scarce in *A. oryzae*, where, as noted earlier, most TE families consist either of unmutated copies or heavily degraded ones, with few intermediates. The presence of both RIP-affected and unaffected TE copies implies that RIP is sporadic, at least for *A. nidulans*, where the sexual cycle is active, i.e., only a proportion of repeated elements are mutated in any one passage through the cycle. This is in contrast to *N. crassa* where no RIP-free TEs have been found in the sequenced genome [17]. A distribution of RIP-affected and unaffected elements in *A. fumigatus* similar to that in *A. nidulans* suggests that both monitor their genomes for repeats with low efficiency, while *A. oryzae*, with few RIP-free elements, may be more like *N. crassa* in its degree of vigilance. However, firm conclusions, in the absence of experimental evidence, depend on assumptions about the frequency of sexual reproduction in the history of these species. This can be assessed for the sexually active *A. nidulans* [23], but is unknown in the other two apparently asexual species, as discussed further on.

RIP frequencies are affected by the environment of the affected cytosine, in particular, the following base [8,15]. Table 20.2 shows transition frequencies for one TE family typical of each *Aspergillus* species, and, for comparison, for the ζ-η sequence of *N. crassa*. As previously observed [8,11], it can be seen that in *A. nidulans* and *A. fumigatus* similar to that in *A. nidulans* suggests that both monitor their genomes for repeats with low efficiency, while *A. oryzae*, with few RIP-free elements, may be more like *N. crassa* in its degree of vigilance. However, firm conclusions, in the absence of experimental evidence, depend on assumptions about the frequency of sexual reproduction in the history of these species. This can be assessed for the sexually active *A. nidulans* [23], but is unknown in the other two apparently asexual species, as discussed further on.

RIP frequencies are affected by the environment of the affected cytosine, in particular, the following base [8,15]. Table 20.2 shows transition frequencies for one TE family typical of each *Aspergillus* species, and, for comparison, for the ζ-η sequence of *N. crassa*. As previously observed [8,11], it can be seen that in *A. nidulans* and *A. fumigatus* similar to that in *A. nidulans* suggests that both monitor their genomes for repeats with low efficiency, while *A. oryzae*, with few RIP-free elements, may be more like *N. crassa* in its degree of vigilance. However, firm conclusions, in the absence of experimental evidence, depend on assumptions about the frequency of sexual reproduction in the history of these species. This can be assessed for the sexually active *A. nidulans* [23], but is unknown in the other two apparently asexual species, as discussed further on.

### 20.4 A+T Content

*A. nidulans* and *A. fumigatus* have been described as exhibiting a mild version of the RIP mechanism [1,17], comparable to that found in *Magnaporthe grisea*, but it now appears that RIP in *A. oryzae* has been both more widespread and more severe. Many of the best TE consensus sequences that can be reconstructed for this species are themselves AT-rich and contain only disrupted ORFs. An early finding was that in BLAST searches for matches to such consensus sequences, both strands were frequently recognized by the same query; and on examination these genomic regions were seen to consist of pure A+T mixtures.

Machida et al. [3] observed that DNA stretches of 50 or more nucleotides with composition >90% A+T were six to nine times commoner in *A. oryzae* than in the other two species. Since RIP in *A. fumigatus* and *A. nidulans* gives rise to less extreme nucleotide bias, we have expanded that observation by scanning the genomes for sequences of more than 75% A+T (Table 20.3). The most striking species difference, in agreement with Machida et al., is for >85% A+T, where *A. oryzae* has approximately four times as much as *A. nidulans* and *A. fumigatus*. In the range 75–84% A+T, *A. nidulans* has approximately 60% of the figure for the other two fungi. A further observation is that in *A. fumigatus* >75% A+T runs are significantly shorter than in the other two fungi and only *A. oryzae* has an appreciable number of runs of more than 500 nt.

<table>
<thead>
<tr>
<th>Elements Sampled</th>
<th>RIP-Free</th>
<th>RIP-Affected</th>
<th>Total Transitions</th>
<th>C→T Transitions as % of Available Doublets</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. nidulans</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mariner-6</td>
<td>18</td>
<td>6</td>
<td>437</td>
<td>14 0.5 26 0.9</td>
</tr>
<tr>
<td><em>A. fumigatus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mariner-4, -4b</td>
<td>11</td>
<td>4</td>
<td>654</td>
<td>69 2 42 3</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mariner-2</td>
<td>13</td>
<td>3</td>
<td>375</td>
<td>41 13 8 8</td>
</tr>
<tr>
<td><em>N. crassa</em></td>
<td>ζ-η</td>
<td></td>
<td>268</td>
<td>64 5 13 18</td>
</tr>
</tbody>
</table>

aData from reference 15.
It was expected that a proportion of A+T-rich DNA would overlap with RIP-affected TEs; in this respect *A. nidulans* and *A. oryzae* were similar, but *A. fumigatus* showed a much greater overlap (Table 20.3, row 5), implying that detection of RIP-affected elements has been most successful in this species.

It can be concluded from this that *A. oryzae* has long stretches of A+T-rich DNA, much of which is unrecognizable as TE-derived. A+T-rich sequences in *A. fumigatus* are shorter than in the other two species, but there is still a considerable quantity of moderately A+T-rich DNA, approximately half of which has also been identified by the transposon search. *A. nidulans* has the least A+T-rich DNA.

### 20.5 Clustering and Fragmentation

In all three species a large proportion of TEs are clustered, here defined as within 1 kb of another element (Table 20.4). Such clusters occur most prominently at both telomeric and centromeric ends of supercontigs, but also in other positions. Table 20.4 shows that in all three species clustered elements are more affected by RIP and are more fragmented than scattered ones, implying that they are generally older.

#### 20.5.1 *Aspergillus fumigatus*

Of the three species studied, *A. fumigatus* has the fewest recognized TEs, but in spite of this, it has a total quantity of transposon-derived DNA nearly as great as that of *A. nidulans* (Table 20.1), i.e., it has more

### TABLE 20.4

Properties of Scattered and Clustered TEs in Three *Aspergillus* Species

<table>
<thead>
<tr>
<th></th>
<th>A. nidulans</th>
<th>A. fumigatus</th>
<th>A. oryzae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scattered</td>
<td>332</td>
<td>200</td>
<td>331</td>
</tr>
<tr>
<td>Clustered</td>
<td>799</td>
<td>252</td>
<td>438</td>
</tr>
<tr>
<td>RIP Index</td>
<td>0.30</td>
<td>0.47</td>
<td>0.48</td>
</tr>
<tr>
<td>% Length</td>
<td>62</td>
<td>61</td>
<td>28</td>
</tr>
</tbody>
</table>

In all three species RIP index and % length values for scattered and clustered elements are significantly different ($p < .05$ or lower).

---

*Genomes were scanned for stretches of DNA consisting of the specified percentage of A+T in a 50 nt window, step 25 nt.

*Significantly shorter than the for the other two fungi: $p < .01$. 
intact elements and large fragments than the other two species. For this reason, clustering and fragmentation of TEs has been examined most closely in this species. Despite the overall impression of a younger set of TEs in *A. fumigatus* than in the other two species, Figure 20.1b shows that all *A. fumigatus* TE families include both intact copies and small, C+G-deficient fragments, and it is apparent from Table 20.4 that clustered elements have suffered a greater degree of RIP than in either of the other two species (although, as noted above, it is probable that the most heavily RIP-affected sequences in *A. oryzae* are no longer recognizable as TE). Many members of the *Afut1_AF* and *Afut2_AF* families, in particular, are clearly ancient and degraded. Figure 20.2 demonstrates that both type I retrotransposons and type II DNA elements fall into two groups with relatively few intermediates: young full-length elements, many of which have high similarity to their consensus sequences, and fragments, most of which have lower consensus identity. This suggests a history of two periods of proliferation, one early, leaving mainly degraded fragments, and one recent, resulting in relatively intact insertions.

A proportion of clusters consist of “nested” elements, arising through invasion of a preexisting element by a new one. Nested clusters are most conspicuous in *A. fumigatus*, comprising 40% of clustered elements, compared to 9% in *A. nidulans* and 5% in *A. oryzae*. Table 20.5 gives the properties of “host” and “invader” components of such nested sets for *A. fumigatus*: as expected, invaders generally have the properties of younger elements, while host elements show evidence that they are older: they have a multiple differences from their consensus, and are usually more affected by RIP. A preference for RIP-affected insertion sites can be explained in two ways: firstly, transposons will only be observed if their insertion did no serious harm to the host, therefore, insertion into preexisting transposons will be common, and such older transposons will have had the opportunity to become affected by RIP. Secondly, transposons are likely to evolve to prefer nonfunctional target sites, and this may lead to preferential invasion of A+T-enriched DNA.

Other sets of adjacent elements could be the result of splitting of TE nests by chromosomal rearrangement, resulting in pairs, or more complex sets of apparently unrelated fragments. Many such clustered elements appear to have been subjected to multiple destructive processes, including truncation and internal deletion; most of them are short, and heavily affected by RIP (Table 20.5). Some groups of related fragments suggest tandem duplication, with or without inversion. Other cases of apparent fragmentation are due to recognition in BLAST searches of only short stretches of a diverged sequence; and in some of these instances, the gaps can be recognized as A+T-rich DNA.

Given the degenerate nature of most clustered elements other than nested invaders, it is not surprising that the majority of clustered elements are from the older families of LTR-retrotransposons; while 82%
of LTR retrotransposons are clustered, 74% of Mariner elements are solitary. *I-1_AF* elements are a special case in that 10/16 intact *I-1_AF* elements are inserted into the same target site in *Afut2_LTR*s.

### 20.5.2 *Aspergillus nidulans*

In contrast to *A. fumigatus*, Figure 20.1a suggests a much more continuous history of proliferation of both type I and type II transposons, leaving a number of intact elements, and many more small degraded fragments, but also many of apparently intermediate age. The spread of transposon families across clustering categories analyzed for *A. fumigatus* in Table 20.5 is also more uniform (data not shown). It is seen that 71% of *A. nidulans* TEs are clustered, and of these only 25% are full-length, compared to 39% of solitary elements.

As in *A. fumigatus*, RIP-affected elements appear to be particularly prone to further transposon invasion, e.g., of 22 full-length *Mariner-6_AN* elements, 6 show evidence of substantial RIP and four of these are hosts to other elements. The same is true of *DNA-3_AN/MATE* elements: of the five full-length copies, only the two RIP-affected ones harbor other elements [8].

### 20.5.3 *Aspergillus oryzae*

It is clear from Figure 20.1c that the majority of elements are fragmentary and considerably diverged from any consensus. There are some intact elements of both types I and II but, as in *A. fumigatus*, fewer TEs of apparently intermediate age. It is seen that 57% of TEs are clustered. Only 15% of clustered elements and 18% of solitary elements are full-length.

### 20.6 Transposon-Related Genes in *Aspergillus nidulans*

A total of 260 autocalled genes from Broad Institute *A. nidulans* genome gene prediction, version 2, wholly or partially overlap our TE predictions. Of these, 77 consist of ORFs within TEs, 73% of these being recognizable as encoding transposition functions. There are 71 cases of TEs, mostly solo LTRs or other fragments, inside autocalled genes; of these, 16 are recognized as being, at least in part, transposition related. Only 19 TEs are wholly within predicted introns. The 84 remaining overlaps are partial, 49 of the relevant ORFs showing some transposition-related homology. Many of the predicted overlapping genes do not look as if they have typical fungal intron-exon structures, so they are likely candidates for revision.

Only two experimentally verified genes overlap predicted transposons: *Helitron-N1_AN* is represented by a single full-length copy in the *A. nidulans* genome, and this overlaps exon 1 and part of intron 1 of the

### TABLE 20.5

<table>
<thead>
<tr>
<th>Elements</th>
<th>Number</th>
<th>Mean RIP Index</th>
<th>Mean % Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nested: hosts</td>
<td>37</td>
<td>0.83</td>
<td>69</td>
</tr>
<tr>
<td>Nested: invaders</td>
<td>39</td>
<td>0.63</td>
<td>81</td>
</tr>
<tr>
<td>Other clustered</td>
<td>150</td>
<td>0.84</td>
<td>35</td>
</tr>
</tbody>
</table>

*14 insertions of *–1_AF* into *Afut2-LTR_AF* not included. Invading elements that are themselves hosts to further invaders, are most similar to other hosts and are counted as such.

*Calculated from doublet frequencies (TpA − CpG)/(TpA + CpG).

*Significantly lower (*p* < .01 or lower) than values for the other two classes.

*Lengths of fragments are combined for host elements split by invaders.*
20.7 Transcription from TEs and Transposition

Transcription of a number of transposases has been demonstrated in different species, for example, the "Fot1-like" Mariner-6_An element [26]. The same paper reported active transposition of an introduced Fusarium oxysporum Fot1 element, which was evidently not repressed by the resident homologs. On the other hand, active transposition of these resident elements has not been shown, and surveys of spontaneous mutants at a number of loci have uncovered no inserted TEs [26], but this could be a sign of preferential insertion into noncoding sequences. Wild strains of A. nidulans have been shown to be polymorphic for sites of both DNA-3_AN/MATE [8] and Mariner-6_AN (Fot1-like) [26], confirming the occurrence of transposition at the population level.

20.8 Implications for Genome Expansion, Sexuality, and Chromosomal Rearrangement

20.8.1 Genome Expansion

Galagan and Selker [17] postulated that a paucity of expanded gene families in N. crassa correlates with the efficiency of destruction of multicopy sequences by RIP in this fungus. It is, therefore, ironic that of the three fungi discussed here, A. oryzae is celebrated for its diversity of expanded metabolic gene families [3] but, unlike the other two Aspergillus species, has a relaxed doublet preference for RIP comparable to that in N. crassa, and shows the highest level of RIP-degradation of TEs. Possible explanations for gene family expansion despite efficient RIP are that many of the additional genes have been derived by horizontal gene transfer and are not seen as duplicates of older genes [3]. Even if blocks of genes were introduced in this way, as postulated by Machida et al. [3], duplicates could have escaped RIP if sufficiently diverged from each other before transfer.

RIP, where it has been demonstrated experimentally, is strictly associated with sexual reproduction, so another possible explanation for successful gene duplication is that it may have occurred since the loss of sexual reproduction in A. oryzae (see later).

20.8.2 Sexuality

Galagan et al. [1] reported that the sequenced genomes of all three Aspergillus species include a full set of genes associated with sexual reproduction, but while the sexual fertile, homothallic A. nidulans has genes determining both mating types, A. oryzae and A. fumigatus have single mating-type genes, implying that both are heterothallic. Sexual reproduction is a regular feature of A. nidulans biology, both in the laboratory and in the wild [23], but has never been observed in the other two species, despite the existence in the wild of both mating types of A. fumigatus [27]. Since RIP in N. crassa is strictly associated with sexual reproduction, a prolonged asexual existence should be reflected in the accumulation of RIP-free transposons.

This seems to be the case for A. oryzae (Fig. 20.1c) where most families consist of a few full-length copies, showing no signs of RIP, plus many small fragments, all depleted in C+G content. The Mariner-3_AO family has 12 intact members, all RIP-free, while 44 very A+T-rich fragments were sufficiently different for standard procedures to lead to the construction of a separate consensus Mariner-4_AO. This suggests a history of extensive and efficient RIP, followed by a significant period when RIP was inactive.
Only the Mariner-2_AO family has a mixture of 3 full-length but RIP-affected members and 13 unaffected ones: a pattern familiar in A. nidulans, where it is ascribed to the sporadic occurrence of RIP.

The picture for A. fumigatus is less clear; the majority of TE families look like those in A. nidulans (Fig. 20.1a,b) in comprising a mixture of RIP-affected and RIP-free members of various lengths. Two families are of interest, however: the Mariner-4_AF family has 15 full-length copies, and like Mariner-3-4_AO, two of these are so heavily affected by RIP as to generate a separate consensus, Mariner-4b_AF. They may well be evidence of their proliferation earlier than that of the remainder of the Mariner-4 family. The 13 other Mariner-4 members between them differ from their consensus by 36 transitions, 3 transversions, and 2 single base deletions, suggesting some degree of aging. However, only half of the transition mutations are C→T in a CpR doublet context, suggesting that most, if not all, of these mutations are due to standard mutation rather than RIP. This suggests a considerable period during which this family has decayed by simple mutation, in the absence of RIP. On the other hand, the Mariner-1_AF family has 20 intact copies, of which two exactly fit the consensus and the remainder show evidence of limited RIP, with an average of 4 C→T transitions typical of RIP, and less than one other mutation per copy. This suggests the recent occurrence of mild RIP.

20.8.3 Chromosomal Rearrangement

Galagan et al. [1] also noted that while the number of mutational differences between the three species for orthologous genes suggests similar evolutionary branch lengths since divergence, chromosomal rearrangements have occurred approximately one third as often in the branch leading to A. fumigatus as in the branches leading to A. nidulans and A. oryzae. Since both aberrant transposition and ectopic recombination between similar TEs at different sites are potent sources of chromosomal rearrangement [28], it is of interest to see if any correlation can be discerned between TE populations and chromosomal breakage in the three species.

It might be argued that the postulated interval between earlier and more recent TE proliferations in A. fumigatus left that fungus with relatively few TEs for a long period. On the other hand, since A. oryzae appears to have the most active nonhomologous recombination machinery. This could also account for the large proportion of intact TEs in this fungus (Table 20.1). There is little evidence for or against this hypothesis in the literature; Strømnes and Garber [29] were unable to find mitotic recombinants between spore-color mutants, but this is now predictable since these genes are tightly linked in this species [30]. More relevant, but contrary to our hypothesis, may be the finding that electroporation of heterologous plasmids induced numerous chromosomal rearrangements [31]. On the other hand, if the ratio of solo to attachedLTRs is taken as a measure of recombination, this is low (1.0) in A. fumigatus, but much higher in A. nidulans (6.5) and A. oryzae (5.8). While this ratio must depend on the age of these elements in each species, Figure 20.1a through 20.1c do not suggest that LTR retrotransposons are younger in A. fumigatus than in the other two species.

20.9 Conclusions

At first glance the TE populations in these three fungi are quite similar, making up 1.3% to 2.7% of the genome, but closer examination reveals differences. There has evidently been a complete turnover of TE content since divergence of the three species, and each fungus has its own history of TE proliferation and mutational decay. Prominent among the mechanisms of decay is RIP, evidence for which is strong in all three species. A. nidulans and A. fumigatus have a restricted doublet preference for RIP-determined C→T transitions, not shared by A. oryzae, in which evidence suggests that RIP has been both widespread and intensive, to the point where much TE-derived DNA is now seen as amorphous A+T-rich sequence. A. fumigatus contains the largest number of full-length elements, possibly reflecting a reduced propensity
for genome rearrangement in this species. Examination of clustering and fragmentation in this species suggests that “nesting,” resulting from invasion of preexisting elements by new arrivals, is a primary cause of both phenomena. The fact that gene-family expansion has occurred to a greater extent in A. oryzae than in the other two species, despite more intensive RIP, supports the hypothesis that much of this expansion is due to horizontal transfer rather than gene duplication. The presence of apparently undamaged elements alongside heavily RIP-affected ones in A. oryzae also supports the idea that RIP has been inactive in this fungus for some time, due to the absence of sexual reproduction. Similar evidence for A. fumigatus is equivocal.

Acknowledgment

AJC is grateful to Derek Gatherer for initiation into Perl scripting.

References

III

Medically Important Aspects of the Genus
Clinical Aspects of the Genus Aspergillus

William J. Steinbach

CONTENTS

21.1 Clinical Presentation ................................................................. 360
  21.1.1 Epidemiology and Risk Factors ........................................... 360
    21.1.1.1 Pulmonary Aspergillosis ................................................. 361
    21.1.1.2 Invasive Aspergillus Sinusitis ......................................... 361
    21.1.1.3 Cerebral Aspergillosis .................................................. 362
    21.1.1.4 Cutaneous Aspergillosis ............................................... 362
    21.1.1.5 Chronic Aspergillosis .................................................. 362
    21.1.1.6 Aspergilloma .............................................................. 363
    21.1.1.7 Allergic Bronchopulmonary Aspergillosis ....................... 363
  21.2 Diagnosis ........................................................................... 363
    21.2.1 Cultures ................................................................. 364
    21.2.2 Radiology .................................................................. 364
      21.2.2.1 Bronchoalveolar Lavage .............................................. 365
      21.2.2.2 Galactomannan Antigen ............................................. 365
      21.2.2.3 (1,3)-β-D-Glucan ....................................................... 366
      21.2.2.4 Polymerase Chain Reaction (PCR) ......................... 366
  21.3 Treatment ........................................................................ 367
    21.3.1 Immunomodulatory Therapy ............................................ 368
  21.4 Prognosis ........................................................................ 369
  References .................................................................................. 370

Aspergillosis refers to infection with any of the approximately 185 recognized species of the genus Aspergillus, of which only 20 are known to cause human disease. Most human disease is primarily caused by Aspergillus fumigatus, Aspergillus flavus, Aspergillus niger, Aspergillus terreus, and Aspergillus nidulans [1]. Aspergillus fumigatus causes approximately 90% of cases of invasive aspergillosis [2], and most pulmonary disease is caused by A. fumigatus, while most isolated sinus disease is caused by A. niger and A. flavus [3]. A review of Aspergillus cultures found that amphotericin B-resistant A. terreus was seen in only 3% of isolates in cases of invasive aspergillosis and found exclusively in cases of invasive disease, and not in patients with colonization [3].

Although yeasts such as Candida species cause the most common fungal infections, the incidence of Aspergillus infections is increasing and carries a dismal mortality [4–6]. Aspergilli are the most common cause of mortality due to invasive mycoses, likely due to the success with prophylactic regimens and easier diagnostic examinations for Candida infection. A complete review of the entire clinical diagnostic spectrum and therapeutic armamentarium, including all the many nuances that are so critical to effective patient care, is simply impossible in a single chapter. Instead, the general clinical presentations of the varied forms of
disease caused by the genus *Aspergillus* are presented, along with the summary of some important concepts on patient management targeted for the non-clinician.

### 21.1 Clinical Presentation

*Aspergillus* species produce a range of disease, including allergic bronchopulmonary aspergillosis (ABPA), aspergilloma, chronic necrotizing aspergillosis, and various forms of invasive aspergillosis [1]. The most common forms of invasive aspergillosis are acute pulmonary aspergillosis, acute invasive rhinosinusitis, cerebral aspergillosis, and disseminated disease. *Aspergillus* aerosolizes conidia readily and while immunocompetent people breathe and clear conidia everyday [1,3], immunocompromised patients are at risk for the development of invasive aspergillosis. Since the route of infection appears to be pulmonary, the first line of defense is formed by alveolar macrophages. *In vitro* studies with murine cells have suggested that resident pulmonary macrophages are responsible for digesting inhaled *Aspergillus* conidia [7,8]. If conidia escape and germinate into hyphae, then the hyphae become susceptible to neutrophil killing through the release of toxic oxygen radicals. Thus, disease risk is associated with neutropenia, challenge with overwhelming microbial doses, and/or corticosteroid suppression of macrophage conidiacidal activity [9]. The mechanism by which T cells function to protect against invasive aspergillosis is not clear, but they may enhance phagocyte killing of conidia [10].

#### 21.1.1 Epidemiology and Risk Factors

Invasive aspergillosis is a leading cause of infectious death in hematopoietic stem cell transplant (HSCT) recipients and one study showed that 36% of all confirmed nosocomial pneumonia in these patients was caused by *Aspergillus* infection, yielding a crude mortality rate of 95% [11]. The incidence of *Aspergillus* infection in HSCT recipients has ranged from 3% to 7% [12,13], but the true incidence is dependent on the follow-up duration of individual studies. A review of patients from 1990 to 1998 found the yearly invasive aspergillosis incidence increased in both allogeneic transplants (5% increasing to 12%) and autologous transplants (1% increasing to 5%). This difference in incidence in the two types of transplants makes clinical sense, as in allogeneic transplants the recipients receive stem cells from another individual and there is greater risk for rejection, while patients with an autologous transplant receive their own stem cells back after appropriate conditioning. Importantly, the incidence of non-*A. fumigatus* species as cause of invasive pulmonary disease also dramatically increased after 1995 (18% increasing to 34%) [4].

There is a well-characterized bimodal distribution of aspergillosis in HSCT recipients which correlates with pre-engraftment neutropenia with a median of 16 days and the peak of graft-versus-host disease (GVHD) with a median of 96 days [14]. GVHD is the immunologic situation where the patient’s own immune system (the host) battles the incoming graft and wages an inflammatory and immunologic war. These risk periods likely relate to the two major mechanisms of protection against invasive aspergillosis, alveolar macrophages and granulocytes. Most patients (86%) with autologous transplants were diagnosed with invasive aspergillosis while neutropenic, while patients with allogeneic transplants were at greatest risk after engraftment or during impairment of cell-mediated immunity due to cytomegalovirus (CMV) infection or GVHD [14]. As early posttransplant management and survival improves, the peak of invasive aspergillosis appears to be shifting to the outpatient setting.

Neutropenia is the time-honored risk factor for invasive mold infections and the risk of invasive aspergillosis is calculated to increase from 1% per day after the first three weeks of neutropenia to 4–5% per day after five weeks [15]. The incidence of invasive aspergillosis can be as high as 70% if neutropenia exceeds 34 days [16]. Repeated cycles of neutropenia may be an added risk factor. Corticosteroids suppress the ability of monocytes/macrophages to kill conidia through inhibition of nonoxidative processes and impairment of lysosomal activity, and also inhibit polymorphonuclear neutrophils in their chemotaxis, oxidative burst, and antifungal activity against hyphae [17]. The results of one *in vitro* study suggest that corticosteroids may actually accelerate the growth of *A. fumigatus* [18].
21.1.1.1 Pulmonary Aspergillosis

Separate clinical manifestations in different patient populations have been reviewed [2], but disease in most immunocompromised patients is often diffuse pulmonary infection. The most common presentation is unremitting fever [6], but high fever may be absent in those patients receiving steroid therapy [19]. Other early symptoms of pulmonary disease include a dry cough and possibly chest pain. Dyspnea (pain on breathing) is more common in patients with diffuse disease, and the presentation in some patients is similar to a pulmonary embolism. Hemoptysis (coughing up blood) can occur and can be fatal with the first presenting episode [13], while in neutropenic patients a pneumothorax is also an occasional presenting feature [2].

Invasive pulmonary aspergillosis is the leading cause of mortality in patients with chronic granulomatous disease (CGD), and may be the first manifestation of CGD. Invasive aspergillosis in a patient with CGD usually presents within the first 20 years of life, and invasive aspergillosis in a child or adult without a known predisposing risk factor should prompt an evaluation for CGD. Diagnosis in a patient with CGD does not often contain typical clinical symptoms (including a completely asymptomatic patient), and may consist of only an elevated erythrocyte sedimentation rate (ESR) as a general marker of inflammation in the setting of no fever. In a review of invasive aspergillosis in 23 CGD patients, only one-third (8/23) were symptomatic at diagnosis, only one-fifth (4/23) were febrile at diagnosis, white blood count was <10,000 cells/μl in 13/23 cases, and ESR <40 mm/h in 9/20 cases [20].

In early disease in a CGD patient there is an acute neutrophilic response where the neutrophils surround hyphae. However, in this patient the hyphae remain intact due to impaired neutrophil-mediated killing of hyphae. In this setting pulmonary aspergillosis is a chronic progressive infection, which may spread locally to involve pleura, vertebrae, and the chest wall. In contrast to patients with neutropenia, hyphal angioinvasion is not a feature of disease in patients with CGD. The halo sign (angioinvasion with surrounding tissue ischemia), cavitated lesions, and pulmonary infarcts are not typical in CGD. There are areas of tissue destruction secondary to reactive acute and granulomatous inflammatory process rather than directly due to growth of the hyphae [21]. Infection with *A. nidulans* is also more common in patients with CGD, with pulmonary disease more likely to involve adjacent bone, more likely to cause disseminated disease, generally refractory to intensive antifungal therapy, and more likely to require surgery than non-CGD patients [22].

21.1.1.2 Invasive Aspergillus Sinusitis

Invasive *Aspergillus* sinusitis is underdiagnosed because of the lack of detailed examination, but patients can present with ear pain or discharge, facial pain or swelling, localized pallor of the nasal septum or turbinate mucosa, epistaxis (nose bleeds), orbital swelling, or headache [2,19]. The maxillary sinus is most commonly involved, followed by the ethmoid, sphenoid, and frontal sinuses [23]. A careful rhinoscopic examination is needed to look for insensitive areas with decreased blood flow, frank crusting or ulceration, or blackened necrotic foci. One review found 11 patients with invasive fungal sinusitis after bone marrow transplantation, including 8 patients with *Aspergillus*. The mean interval from bone marrow transplantation to diagnosis of fungal sinusitis was 22.5 days and all patients had maxillary sinus involvement, half of the patients had ethmoid sinusitis, and the majority of patients showed extension into the orbits, bone, or brain [23]. One major difficulty is the critically important diagnostic distinction between zygomycosis and aspergillosis, due to the divergent therapeutic approaches. In one study evaluating pulmonary aspergillosis and pulmonary zygomycosis, concomitant sinusitis was significantly associated with zygomycosis and not aspergillosis in patients with radiographic pulmonary disease [24].

A high index of suspicion is necessary in immunocompromised patients. These infections are characterized by mucosal invasion with infarction and spread of infection in centrifugal fashion to contiguous structures. Early diagnosis is imperative, and the onset of new local symptoms, such as epistaxis, nasal-orbital pain, a positive nasal swab culture in a febrile, susceptible host, or an abnormal sinus radiographic finding should lead to immediate otorhinolaryngologic evaluation, including careful inspection of the nasal turbinates. Although surveillance nasal cultures are of questionable value, baseline sinus radiographs or limited computed tomography (CT) should be considered in these high-risk patients. *T*<sub>2</sub>-weighted magnetic resonance imaging (MRI) images may show decreased signal intensity compared to those of bacterial sinusitis, which show increased signal intensity [2].
21.1.1.3 Cerebral Aspergillosis

Cerebral involvement has been noted in up to 40% of patients with invasive aspergillosis [6,12]. The pathogenesis is thought to be due to hematogenous dissemination from an extracranial focus, most commonly the lung, or by direct extension through the sinuses. *Aspergillus* hyphae are angioinvasive, and thrombose arteries create hemorrhagic infarcts that are then converted to abscesses. Cerebral aspergillosis is the most common brain abscess in HSCT recipients; in one study 58% of brain abscesses were caused by *Aspergillus* and 87% of those patients had concomitant pulmonary infection [25]. In another report, cerebral involvement was seen in 10/18 patients, including three patients who presented with neurologic signs and no pulmonary symptoms [26]. The classic features of abscesses such as headache, nausea, and vomiting can be present in <10% of cases, with more prevalent features, including altered mental status, confusion, hemiparesis, and cranial nerve palsies. Multiple lesions in the corticomedullary junction are consistent with infarct due to *Aspergillus* vasculopathy, with dilated cortical vessels located in the central portion of the lesions in the corticomedullary junction often a distinctive sign in diagnosing cerebral aspergillosis [27]. Definitive diagnosis requires a brain biopsy, but these patients are often too coagulopathic for the diagnostic operation.

A frequent target of disseminated disease is the central nervous system (CNS), where hematogenous spread results in occlusion of intracranial vessels and infarction. This may manifest as the characteristic single or multiple cerebral abscesses, or meningitis, an epidural abscess, or a subarachnoid hemorrhage. Cerebral aspergillosis has been noted in 25–40% of patients with invasive pulmonary disease [4,6,12,26]. The classical presenting features of abscesses such as headache, nausea, and vomiting are rare (<10% of cases). More frequently, presenting signs and symptoms include altered mental status, confusion, hemiparesis (paralysis on one side of the body), and cranial nerve palsies. CT of the head often reveals one or multiple hypodense, well-demarcated lesions. Hemorrhage and mass effect are unusual, but for patients with adequate peripheral white blood cell counts, ring enhancement and surrounding edema are frequent [2].

The cerebrospinal fluid (CSF) glucose level is normal, and cultures of the CSF are negative. Biopsy of these lesions is warranted, if feasible, to differentiate *Aspergillus* infections from those caused by other fungi, such as *Pseudallescheria*, dematiaceous fungi, Mucorales or *Fusarium*, which may alter one’s choice of antifungal therapy. A surgical approach leads to laboratory characterization of the causative agent together with removal of nonviable tissue, which may not be well-penetrated by systemic antifungals.

21.1.1.4 Cutaneous Aspergillosis

Cutaneous aspergillosis can be either primary, usually from skin injury or traumatic inoculation, or secondary from contiguous extension of hematogenous dissemination. In general, burn victims, neonates, and solid-organ transplant recipients develop cutaneous inoculation after prolonged local skin injury. HSCT recipients often develop secondary disease from contiguous extension of infected structures under the skin or from hematogenously disseminated embolic lesions [28]. A review of 15 cases of cutaneous aspergillosis from over 4000 patients with malignancy found an incidence of 4% of patients with documented *Aspergillus* infection [29], while another review found embolic lesions present in approximately 11% of patients with disseminated aspergillosis [30].

Cutaneous aspergillosis often begins as an area of raised erythema that progresses to include pain, and skin involvement can be the first presenting sign of invasive aspergillosis (IA) [31]. The center of the lesion changes from red to purple and then to black and may ulcerate [2]. Infections arising at the site of an intravenous catheter puncture typically begin with erythema and induration and progress to necrosis that extends radially [32]. Patients with primary cutaneous aspergillosis appear to present with significantly less necrosis and systemic toxicity than wound zygomycosis [28].

21.1.1.5 Chronic Aspergillosis

Acute invasive aspergillosis almost uniformly occurs in patients with profound immunosuppression and displays rapid progression with poor outcome. However, there are patients who clearly have a more chronic characteristic to their *Aspergillus* infection. Chronic aspergillosis has been previously known by several different terms, most notably semiinvasive aspergillosis, chronic invasive pulmonary aspergillosis,
Clinical Aspects of the Genus Aspergillus

and symptomatic pulmonary aspergilloma. The blurred entity of chronic aspergillosis has been recently reclassified [33] to include chronic cavitary pulmonary aspergillosis to note radiologic formation and expansion of multiple cavities with some containing fungus balls, and chronic fibrosing pulmonary aspergillosis where this progresses to marked and extensive pulmonary fibrosis. The final category is called chronic necrotizing pulmonary aspergillosis, or subacute invasive pulmonary aspergillosis. These are patients with a mild or moderate defect in immune function, unlike the other two chronic classifications, where there is a slow and progressive enlargement of an Aspergillus-containing cavity. Surgery appears to play a smaller role in managing chronic aspergillosis, and the mainstay of treatment is long-term antifungal therapy to halt the progression of disease. While there have been no randomized clinical trials demonstrating the benefit of antifungal therapy in chronic aspergillosis, anecdotal evidence suggests slow improvement [33].

21.1.1.6 Aspergilloma

Aspergilloma is considered a “saprophytic,” or noninvasive, form of infection in which Aspergillus may colonize preexisting pulmonary cavities due to tuberculosis, sarcoidosis, bullous emphysema, bronchiectasis, or other etiologies. Aspergilloma will develop in cavitating lung disease from tuberculosis in approximately 15–25% of patients [34]. Aspergillomas can be divided into simple and complex aspergillomas based on radiographic criteria. The simple aspergilloma can be differentiated from the complex aspergilloma by the absence of constitutional symptoms, para-cystic lung opacities, cyst expansion, or progressive pleural thickening [35]. In one study, chest radiographs showed a “fungus ball” in the cavities of 67% (42/61) of cases of pulmonary aspergilloma and thickening of the cavity wall in 26% (16/32) of the cases [36]. Patients with aspergilloma may be asymptomatic, but many have persistent and productive cough, hemoptysis, and weight loss. Surgical management to completely eradicate the aspergilloma is the preferred treatment. Although the postoperative morbidity rate is higher in complex aspergilloma, in one series of 88 patients surgical management led to nearly 80% survival rates in both patients with simple or complex aspergilloma [37]. Systemic antifungal therapy is often unsuccessful as penetration of the antifungals into the cavity is poor. Percutaneous intracavitary instillation of antifungals designed to fill the cavity and create an anaerobic environment for the Aspergillus has led to some success [38], but should be reserved for inoperable patients [39].

21.1.1.7 Allergic Bronchopulmonary Aspergillosis

Allergic bronchopulmonary aspergillosis is a hypersensitivity lung disease resulting from sensitization to environmental exposure to allergens of A. fumigatus. The A. fumigatus grows saprophytically to colonize the bronchial lumen and results in persistent bronchial inflammation. Conidia trigger an IgE-mediated allergic inflammatory response in the bronchial airways, leading to bronchial obstruction. ABPA is primarily a disease occurring in patients with asthma (1–2%) or cystic fibrosis (1–15%). The manifestations are due to the immunologic responses to the A. fumigatus antigens, including wheezing, pulmonary infiltrates, bronchiectasis, and fibrosis. Immunologic manifestations include peripheral blood eosinophilia, immediate cutaneous reactivity to A. fumigatus antigen, elevated total levels of serum IgE, presence of precipitating antibody to A. fumigatus, and elevated specific serum IgE and IgG antibodies to A. fumigatus [40].

The mainstay of therapy for ABPA is corticosteroids for attenuation of the inflammation and immunologic activity, while antifungals are used to decrease the fungal burden and therefore the antigen load. Itraconazole was found in a randomized, double-blind study to be superior to placebo for treatment of ABPA [41]. While there is insufficient data to recommend itraconazole for initial therapy for ABPA exacerbation, it should be added to therapy if there is a slow or poor response to corticosteroids [40] as part of a two-armed attack of corticosteroids and antifungals to optimize therapy.

21.2 Diagnosis

As with most invasive mold infections, the clinical signs and symptoms are very nonspecific. Unfortunately the most immunocompromised patients are those least likely to have symptoms and progress most
The Aspergilli rapidly, whereas less immunocompromised patients (e.g., patients with diabetes mellitus) usually have indolent symptomatic presentations [2]. In severe disease, an aggressive, invasive approach, as well as a tissue diagnosis early in the illness, appears to be a key to survival. In the appropriate clinical setting, such as an immunocompromised host with fever and a pulmonary infiltrate on X-ray, repeated isolation of the same species in culture, and particularly a bronchoalveolar lavage (BAL) or other endobronchial culture, correlates with invasive disease; sometimes even a single sputum culture (especially with heavy growth) may have to be the stimulus for therapy if invasive procedures cannot be carried out.

Diagnosis is difficult because aspergilli frequently are contaminants in sputum and even in other cultures during handling. Despite many efforts in developing new and exciting detection tools such as PCR assays, the diagnosis of invasive aspergillosis still remains very difficult. Several reasons are responsible for these limitations. First, invasive aspergillosis often shows nonspecific and variable clinical signs, and the manifestations are subtle and occur late in the course of disease. Second, invasive aspergillosis occurs in many different patient cohorts, those at risk for a short period of time or for years. Because of residual defects and tissue infarcts, the disease has a potential to reactivate, mainly during prolonged or continuous immunosuppression and it may occur as a subacute or chronic infection. Third, no unique universally applicable test with sufficient sensitivity and specificity exists, and in consequence, invasive aspergillosis is often diagnosed late leading to a delayed initiation of antifungal therapy resulting in a fatal outcome. While culturing *Candida* species is still the hallmark of diagnosis for that invasive fungal infection, a culture is often neither easy to obtain nor interpret in patients with potential invasive aspergillosis.

### 21.2.1 Cultures

The “gold standard” of tissue biopsy for culture is often considered too invasive and complicated by bleeding or secondary infection in HSCT patients. In tissues, aspergilli may be seen as dichotomously branched (resembling the divergence of fingers from one another) septate hyphae, and may produce their characteristic conidia in tissues or artificial media. If the septation can be seen, they can be differentiated from the zygomycetes, but aspergilli may be confused with *Pseudallescheria boydii* unless the characteristic terminal spores of the latter are seen.

A positive culture of *Aspergillus* from an otherwise sterile site provides proof of the disease. However, culture may have a reduced ability to detect *Aspergillus* at an advanced stage of the disease, owing to necrosis occupying a large portion of the lesion. Aspergillosis diagnosed by blood culture is rare. There has been speculation that only 32 cases of true *Aspergillus* fungemia have been correctly documented in patients with hematological disease [42]. In a study of 1477 separate positive cultures there were more than a dozen positive blood cultures, but most were associated with pseudofungemia or terminal events noted at autopsy [3]. In general, the *Aspergillus* hyphal mass that develops in the lumen during angioinvasion remains in place until the force of blood flow causes hyphal breakage, which then allows the mass to circulate. The likelihood of a blood culture capturing these irregularly and infrequently discharged units is small. This difficulty in detection of *A. fumigatus* in blood culture stands in contrast to other angioinvasive filamentous fungi (e.g., *Fusarium* species, *Paecilomyces lilacinus*, *Scedosporium prolificans*, *Acremonium* species) that have the ability to discharge a steady series of unicellular spores into the bloodstream, which are more likely to be captured in a blood sample. This ability to sporulate in tissue and blood has been termed adventitious sporulation [43]. As *A. terreus* also displays adventitious sporulation, histopathology and KOH examination of these spores also can allow rapid, presumptive identification of *A. terreus*. Therefore, a positive blood culture with *A. terreus* or another fungi that demonstrate adventitious sporulation should not be ignored.

### 21.2.2 Radiology

Different radiological tools such as chest radiography, ultrasonography, CT, and MRI are available. The appearance of invasive aspergillosis on chest radiographs is extremely heterogenous. The most
distinctive appearances are cavitations and pleural-based, wedge-shaped lesions. In addition, nodular shadows with and without cavitation and thin- or thick-walled cavities (especially in patients with AIDS) are typical signs of invasive aspergillosis. However, pulmonary invasive aspergillosis often results in false-negative chest radiographs. Therefore, high-resolution CT scans often play an important role in the detection of invasive aspergillosis. Invasive pulmonary aspergillosis characteristically manifests on radiographs as multiple, ill-defined, 1–3 cm peripheral nodules that gradually coalesce into larger masses or areas of subsegmental and segmental consolidation. Lobar or diffuse pulmonary consolidation are common findings [44]. Chest radiographs can be abnormal, but in one series they were actually normal in approximately 30% of patients in the week preceding death [45].

There are two classic radiologic signs of invasive pulmonary aspergillosis. The “halo sign” occurs in neutropenic patients with a hemorrhagic nodule due to angioinvasion. An early CT finding of the halo sign is a rim of ground-glass opacity surrounding the nodule. In one study, the halo sign was seen in all patients with biopsy-proven invasive pulmonary aspergillosis, but it is so nonspecific it was also seen in patients with zygomycosis, organizing pneumonia, or pulmonary hemorrhage [46]. These early lesions subsequently change into a cavitory lesion or lesion with an “air crescent sign” 2–3 weeks later when neutropenia recovers [44,47]. In one study this was seen in 48% of patients 3–10 days after recovery of neutropenia [26]. Cavitation of the nodules or masses occurs in about 40% of patients and is characterized by an intracavitary mass composed of sloughed lung and a surrounding rim of air. In one study the appearance of the air crescent sign had no relationship to duration of neutropenia, and showed a tendency to appear in large lesions such as consolidation or mass rather than small lesion-like nodules [48].

21.2.2.1 Bronchoalveolar Lavage

The BAL was analyzed in 23 consecutive patients with histologically proven invasive aspergillosis, and only 7 patients (30%) had BAL specimens diagnostic for invasive aspergillosis. In that group where it was diagnostic, 71% had multiple changes on thoracic CT compared to 25% of patients with negative BAL. The diagnostic yield of BAL was not associated with clinical symptoms or duration of neutropenia. A thorough review of the diagnostic yield of BAL specimens in histologically proven invasive aspergillosis patients yielded sensitivities of approximately 40% (range 0–67%) [49], but in one study BAL had a sensitivity of only 50% even in patients with focal invasive aspergillosis [50]. The sensitivity of respiratory tract culture specimens in general has ranged from 15% to 69% [3], and in one study it increased 50–70% in high-risk invasive aspergillosis groups. Therefore, a negative BAL in a high-risk patient does not conclusively exclude the possibility of invasive aspergillosis.

21.2.2.2 Galactomannan Antigen

Galactomannan (GM) is a major cell wall component of Aspergillus and it is known that the highest concentrations of GM are released in the terminal phases of the disease [1]. An enzyme-linked immunosorbent assay (ELISA) technique was introduced using a rat anti-GM monoclonal antibody, EB-A2, which recognizes the 1→5-β-D-galactofuranoside side chains of the GM molecule [51]. The threshold of detection with ELISA improved to 5 ng/ml [51]. A sandwich ELISA technique was introduced in 1995 [52] and by using the same antibody as both a capture and detector antibody in the sandwich ELISA (Platelia® Aspergillus, Bio-Rad, France) the threshold for detection can be lowered to 1 ng/ml. The Platelia assay was approved for use in the United States in May, 2003 and is the preferred commercially available assay.

Positive GM findings are suggestive of invasive disease, but false-positives are especially high during the neutropenic period following HSCT. A three-year prospective trial showed the sensitivity of serial monitoring was 90%, specificity 98%, and negative predictive value of 98%. All 30 patients with proven IA tested positive, with no false-negatives. GM detection preceded the development of infiltrates on chest radiograph in 68% of patients. The false-positive rate was 14%, and therefore the improved sensitivity over latex agglutination is counterbalanced by the loss of specificity and greater false-positives [53].
Other studies have reported false positives at 5–8% and suggested they were due to cytotoxic agents, increased resorption of GM, or cross-reacting factors from the intestines [54]. Antigenemia can be observed from one week up to two months, depending on the type of patient and serial testing at least twice a week has been recommended [55]. In one study, an increase in value during the first week of observation was predictive of treatment failure in allogeneic HSCT patients [56]. In a large prospective study of hematology and HSCT patients with confirmed or probable invasive aspergillosis, GM was detected in 65% of patients, an average of 8.4 days before positive CT scans or cultures, and GM was detected in 40% of patients before the onset of clinical symptoms by a mean of 6.9 days. The sensitivity of GM detection in bone marrow transplant (BMT) patients was 89%, with a specificity of 98% [57]. One prospective study involving 186 consecutive patients yielded a sensitivity of 93% and specificity 95%, and in most cases antigenemia was detected a median of six days before clinical symptoms [51]. Unfortunately, the GM assay has decreased sensitivity in the setting of a patient receiving anti-Aspergillus antifungals, while the specificity for detection does not change [58]. While the GM assay has brought with it much promise for a noninvasive diagnostic tool, there remain many challenges, including the exact cut-off value used, optimal sample to test (serum versus BAL fluid), utility in pediatric versus adult patients, and many other unresolved issues.

21.2.2.3 (1,3)-β-D-Glucan
(1,3)-β-D-glucan is an integral cell wall component and, in contrast to GM, is not normally released from the fungal cell [1]. Factor G, a coagulation factor of the horseshoe crab, is a highly sensitive natural detector of (1,3)-β-D-glucan [59]. The “G test” detects (1,3)-β-D-glucan via a modified limulus endotoxin assay and detects Aspergillus, Candida, and even Cryptococcus, but does not identify the genus of the fungi detected [59]. The G-test is widely used in Japan, however, these tests yield positive results only at advanced stages of infection in a majority of patients [60]. A one-year prospective study of patients with hematological malignancy and controls found that the sensitivities of real-time PCR was 79%, GM was 58%, and G-test was 67%, with specificities of 92%, 97%, and 84%, respectively [61].

In a study comparing (1→3)-β-D-glucan and GM, the sensitivity, specificity, and positive and negative predictive values for GM and β-glucan were identical. False-positive reactions occurred at a rate of 10.3% in both tests, but the patients showing false-positive results were different in each test. Both tests anticipated the clinical diagnosis and CT abnormalities, but β-glucan tended to become positive earlier than GM. A combination of the two tests improved the specificity (to 100%) and positive predictive value (to 100%) of each individual test without affecting the sensitivity and negative predictive values [62].

Another study compared GM, PCR, and β-glucan on hematological disorders and the receiver-operating characteristic analysis showed an area under the curve was greatest for the GM assay, using two consecutive positive results. This suggests that the GM was the most sensitive at predicting the diagnosis of invasive aspergillosis in high-risk patients with hematological disorders [63]. The β-glucan assay is currently approved for use in the United States, but there are few studies yet validating its role in diagnosing invasive aspergillosis.

21.2.2.4 Polymerase Chain Reaction (PCR)
Although GM assays created a noninvasive test with improved sensitivity and specificity, recent efforts have focused on defining an optimal primer sequence for a PCR detection method. At present this diagnostic method is not commercially available, and reports can be difficult to interpret due to the lack of experimental standardization between centers. Due to the ubiquitous nature of the mold, the value of this test is its high negative predictive value. Issues remaining unresolved in the use of PCR are the best source of material (e.g., whole blood, serum, plasma, BAL specimens), the amplification protocol (e.g., real-time, sample volume, extraction methods), and primer selection (e.g., “panfungal,” 18S rRNA, 28S rRNA, mitochondrial DNA, etc.) [64]. Using PCR in BAL specimens as compared to blood samples seems less promising due to higher number of false positives. Real-time PCR assays seem to decrease the risk of false-positive results and have better reliability than conventional PCR [65].
Reported sensitivities in numerous retrospective reports of PCR are 55–100%, with negative predictive values generally around 100% [66]. The high-negative predictive value is consistent with PCR as a sensitive marker for any colonizing or infected aspergilli, and a negative PCR in a patient with a suspected invasive aspergillosis most likely does not have the organism [67]. In a two-year prospective study of 121 hematologic malignancy patients analyzed twice-weekly with screening of whole blood samples, the negative predictive value of two or three consecutive positive PCR results was 98%, with a sensitivity of 75% and a specificity of 96%. When PCR was detected only once it was never associated with disease and resolved without antifungal treatment, thus indicating only a transient Aspergillus DNAemia [68].

In another prospective PCR study in 84 allogeneic HSCT patients, all patients with proven or probable invasive aspergillosis were PCR positive. PCR was found to be the earliest indicator for patients with new onset of invasive aspergillosis, preceding the first clinical signs by a median of two days and preceding the diagnosis based on typical radiographic findings by a median of nine days. For patients without a history of invasive aspergillosis who tested positive PCR twice, sensitivity was 100% and the negative predictive value was 100% [69]. In another study, patients with invasive aspergillosis diagnosed by a positive PCR preceded radiologic signs by a median of four days. Additionally, the disappearance of fungal DNA correlated with successful therapy, and the patients who died of IA remained PCR positive [70]. In another study, PCR positivity preceded radiologic diagnosis in 11–18 patients and an ELISA did not precede a positive PCR on any patient. PCR was more sensitive than ELISA, although ELISA had no false positives [71].

21.3 Treatment

The armamentarium of antifungal choices for invasive fungal infections has increased substantially in recent years. While once choices were quite limited and often ineffective, clinicians must now consider many factors before deciding on therapy, including the organism and its epidemiology, antifungal resistance patterns, and any prior antifungal therapy. Currently licensed systemic antifungals include amphotericin B and its lipid derivates; 5-fluorocytosine; the triazoles, including, itraconazole, voriconazole, and posaconazole; and the echinocandins, including caspofungin, micafungin, and anidulafungin.

Antifungal prophylaxis of susceptible patients, such as immunocompromised hosts, using intranasal, inhaled, or systemic antifungals, is an approach to avoid disease and the need for therapy. One strategy for this would be to identify the highest risk patients, such as those identified by screening respiratory cultures as colonized, or those with HSCT and GVHD, and targeting prophylaxis to them. Reducing airborne spores, such as by filtering hospital air, keeping patients in rooms with positive pressure relative to the corridor and frequent air changes and unidirectional air flow in the room, reducing activities that increase spore counts such as room maintenance when the patient is in the room, separating patients from areas of construction, substituting sponge baths for showers and bottled water for tap water, and restricting contaminated materials (e.g., potted plants, sterilization of spices), are believed to be worthwhile efforts for patients who will be transiently immunosuppressed or neutropenic.

In invasive disease, prompt, aggressive antifungal therapy has produced superior survival statistics at some institutions, although recovery from neutropenia is a necessary accompaniment of recovery in almost every success. Therapy needs to be initiated on only a high degree of suspicion. Surgical excision has an important role in the invasion of bone, burn wounds, epidural abscesses, and vitreal disease. It may have a function in invasive pulmonary disease for which chemotherapy has failed or where disease impinges on major vascular structures, and there is a heightened risk of sudden, fatal exsanguination. In pleural disease, locally instilling antifungals may be useful. Therapy should be continued after lesions are resolving, cultures are negative, and reversible underlying predispositions have abated. Reinstating therapy in patients who have responded should be considered if immunosuppression is reinstituted or neutropenia recurs.

For years, the response rate for invasive aspergillosis was only approximately 30% using either amphotericin B or itraconazole treatment [72]. Amphotericin B was previously the “gold standard” since its approval in 1958, but currently voriconazole is the clear choice for primary therapy of invasive aspergillosis. In the pivotal study comparing voriconazole to amphotericin B deoxycholate [73], there was a statistically
superior response rate with voriconazole (53%) versus amphotericin B (32%). This response rate also translated into improved patient survival (71% versus 58%) for initial therapy with voriconazole. While some criticized this study for the use of other licensed antifungals after initial randomization to voriconazole or amphotericin B deoxycholate, a subsequent analysis revealed the strategy of initial therapy with voriconazole was also superior to liposomal amphotericin B [74]. An additional open-label study confirmed the benefit of voriconazole as primary therapy versus salvage therapy [75], highlighting the importance of using the best available antifungal therapy first.

There have been two other randomized clinical trials for invasive aspergillosis. The first study examined two different doses of liposomal amphotericin B (1 mg/kg/day versus 4 mg/kg/day) and found that the lower dose of amphotericin B tended toward an improved response. However, when using only proven or probable cases of invasive aspergillosis, the higher dose of amphotericin B appears to be more effective [76]. The second clinical trial compared amphotericin B colloidal dispersion formulation versus amphotericin B deoxycholate, and found equivalent therapeutic response but superior renal safety with the lipid formulation [77].

The echinocandin antifungals appear to act primarily on the growing hyphal tip of the Aspergillus [78]. There are no randomized studies utilizing caspofungin or micafungin as primary therapy for invasive aspergillosis. In the large open trial examining caspofungin as salvage therapy for invasive aspergillosis, there was a 45% (37/83) favorable response rate [79], suggesting an option for salvage therapy if initial antifungal therapy should fail. Micafungin has been studied as primary therapy for invasive aspergillosis with initial success, but large-scale trials need to be completed [80].

In the past, combination antifungal therapy for IA was of little consequence since there were only a handful of possible permutations available, including the use of other agents such as rifampin and fluconazole. However, the recent surge of newer antifungals has created myriad potential combinations, most of which are just beginning to be tested in the laboratory and at best used in a few anecdotal clinical case reports. An exhaustive review of all previous in vitro, animal model, and clinical reports of combination antifungal therapy for IA [81] revealed that clinical combination therapy was approximately 64% effective. Although there have been concerns of antagonism, especially with polyene-azole interaction, these concerns have generally not been clinically relevant [72,81]. While there are reports of success with combination antifungal therapy for invasive aspergillosis [82], there are varied laboratory data that suggest the entire range of outcomes from antagonism to synergy. Future clinical combination trials will be crucial to determine optimal therapy for patients failing monotherapy. Similar to combination therapy has been the debate of “sequential therapy.” There are reports of various patterns of sequential antifungal therapy which raises another issue other than concomitant therapy: the appropriate and safe sequence of agents. Confounding matters is the long half-life of AmB, so even sequential use has an element of concurrent therapy [83].

### 21.3.1 Immunomodulatory Therapy

Host defense is paramount as invasive aspergillosis generally only develops in certain subsets of severely immunocompromised patients. Few patients with persistent neutropenia and invasive aspergillosis survive, and indeed resolution of invasive aspergillosis has followed neutrophil recovery in most cases. In bone marrow transplant patients the risk for invasive aspergillosis remains even after engraftment, highlighting the fact that although the number of phagocytes is important, their ability to kill must be adequate [84]. Immunotherapy offers many therapeutic advantages through the availability of a wide range of recombinant cytokines that exert their effects indirectly through leukocyte activation rather than directly on the fungus. Immunotherapy is designed to increase the number of phagocytic cells and shorten the duration of neutropenia, modulate the kinetics or actions of those cells at the site of infection, and/or activate the fungicidal activity of phagocytes to kill fungal cells more efficiently [1,85].

Human recombinant granulocyte colony-stimulating factor (G-CSF) has been approved for clinical use since 1991 [86]. The potential of exogenously administered G-CSF therapy seems to be in maintaining the innate signal for longer production of polymorphonuclear leukocytes (PMNs) or initiating that signal earlier if endogenous production is decreased or insufficient during a specific time, for example, neutropenia after bone marrow transplantation [87]. One fear has been the unwanted side effect of increased...
inflammatory products, such as the untoward release of reactive oxygen species and lysosomal contents with G-CSF use [88]. However, in vivo and human studies have also shown G-CSF reduces production of inflammatory mediators such as IL-1, TNF-α, and IFN-γ [87].

In addition to increasing the number of mature circulating PMNs, G-CSF enhances phagocytic activity and oxidative burst metabolism. Human G-CSF only affects function of granulocytes, not macrophages, and has been shown to have a protective effect in murine models of IA. Prophylaxis with human G-CSF and AmB or itraconazole showed some additive effect in neutropenic animal models of IA but not in those immunosuppressed with cortisone, which has a greater effect against macrophages. In a neutropenic (cyclophosphamide induced) murine model, human G-CSF alone was ineffective but with AmB showed synergy in survival greater than with itraconazole and G-CSF [89]. Pretreatment of neutrophils with G-CSF and/or IFN-γ can attenuate the inhibitory effect of corticosteroids on PMN-induced hyphal damage [90]. G-CSF administered to human volunteers increased the fungicidal activity through enhanced respiratory bursts of their PMNs against Aspergillus conidia of their PMNs by fourfold [91]. However, there is no clear evidence G-CSF benefits patients with aspergillosis. One review found no significant reduction in fungal infections in acute myelocytic leukemia patients treated with G-CSF [92].

TNF-α is a proinflammatory cytokine secreted by various macrophage populations and shown to be a critical initiator in innate immunity against respiratory pathogens [93] including A. fumigatus [94]. In vitro TNF-α appears to enhance early host defense against Aspergillus invasion, with slight increases in macrophage oxygen radical production, upregulation and activation of alveolar macrophage phagocytosis, and augmented production of other cytokines such as GM-CSF. It also augments a late defense with increased PMN hyphal damage by oxygen radical production [95,96].

In vitro GM-CSF and TNF-α administration have been shown to counteract dexamethasone-induced immunodeficiency [97]. Animal model depletion of TNF-α results in increased fungal burden and mortality [94] and resistance is further impaired in IFN-γ deficient mice [98]. Treatment of mice with neutralizing antibodies to TNF-α and GM-CSF reduces the influx of PMNs into the lungs and delays fungal clearance [99]. Intratracheal administration of a TNF-α agonist resulted in survival benefits when given three days before A. fumigatus inoculation but not when given concomitantly with conidia, suggesting that pretreatment may provide macrophage priming [94]. However, excessive toxicities in doses required to have a biologically useful effect preclude safe administration in humans [84,95].

IFN-γ promotes TNF-α production [100] and enhances PMN and mononuclear cell-induced damage by increasing the oxidative burst of PMNs in response to stimuli such as nonopsonized hyphae of A. fumigatus [96]. IFN-γ and G-CSF can each enhance the oxidative bursts and fungicidal activity in vitro of human PMNs against A. fumigatus hyphae, with the combination of the two cytokines showing an additive effect [101]. IFN-γ can also restore the corticosteroid-suppressed fungicidal activity of human PMN and elutriated monocytes [97,102,103], and IFN-γ-treated human monocytes show enhanced oxygen radical production and damage to A. fumigatus hyphae [102].

Exogenous administration of IFN-γ and TNF-α has resulted in protective effects in a murine model of IA [104] by decreasing mortality and the number of organs affected by Aspergillus. Conversely, IFN-γ and TNF-α neutralization resulted in increased disease and increased expression of IL-10. Although IFN-γ is better than G-CSF or GM-CSF at enhancing PMN hyphal damage and both IFN-γ and GM-CSF demonstrate enhanced hyphal damage by PMNs in vitro [105], combination treatment does not increase damage [102,106]. In vitro IFN-γ augments PMNs of CGD patients by an undetermined mechanism [107], although previous work demonstrated a myeloperoxidase-dependent oxidative process [108]. IFN-γ has been proven to help prevent IA in CGD patients [109], and there are case reports of the successful use of antifungals and IFN-γ for treatment in CGD patients [110,111]. One recent case report details use of L-Amb and both GM-CSF and IFN-γ in successful therapy of sino-cerebral aspergillosis, with the addition of the IFN-γ temporally related to clinical resolution [112].

### 21.4 Prognosis

The outcome with invasive aspergillosis absent immune reconstitution is generally poor, highlighting the paramount importance to recover immune function. The mortality associated with untreated invasive
The Aspergilli

aspergillosis is nearly 100% in some patient groups, and the overall survival rate among patients treated with amphotericin B had been approximately 34% [113,114] but is now generally improved to approximately 50% with voriconazole [73]. Cerebral disease is rapidly and near uniformly fatal irrespective of treatment [25]. Prognosis for focal pulmonary aspergillosis is more favorable than diffuse bilateral disease, as focal disease tends to progress more slowly. However, focal disease carries an increased risk of hemoptysis, often life-threatening [2]. In one series the median duration of survival of patients diagnosed with invasive aspergillosis before death was 29 days [4]. In another series median survival after invasive aspergillosis diagnosis was 36 days, with a one-year survival rate estimated to be 22%, while mean survival for 11 patients with CNS disease was only 19 days [6].

References


Clinical Aspects of the Genus Aspergillus


22

Pathogenicity Determinants and Allergens

Sven Krappmann

CONTENTS

22.1 Aspergillus—A Saprophytic Pathogen? ......................................................... 377
22.2 Aspergillosis—Disease Forms and Defense Lines ........................................... 378
22.3 Aspergillus fumigatus and Its Pathogenic Relatives ...................................... 380
22.4 What Makes Aspergillus Virulent? ................................................................. 380
22.5 Thermophyly—A Key Feature of Fungal Pathogenicity .............................. 381
22.6 Contacting the Host ...................................................................................... 382
22.7 Sensing the Host ............................................................................................. 383
22.8 Feeding from the Host .................................................................................... 384
22.9 Damaging and Fighting the Host ................................................................... 385
22.10 Special Features ............................................................................................ 388
22.11 Sensitizing the Host ...................................................................................... 388
22.12 Conclusions and Outlook ........................................................................... 391
References ........................................................................................................... 392

Aspergilli are exceptional pathogens as they are able to harm immunocompromised individuals or to elicit allergic responses. During the last decades, aspergilloses have emerged as a major threat for a certain group of patients in specific clinical settings with Aspergillus fumigatus representing the predominant menace. A variety of cellular attributes accompanied by environmental factors contribute to the virulence potential of this filamentous fungus, and we are in the initial stage of understanding what renders Aspergillus pathogenic or allergenic under certain circumstances. In this chapter, the forms of disease and its epidemiology are briefly summarized. Cellular traits that have been characterized to influence the pathogenic capabilities of Aspergillus are described, as well as allergenic determinants, with the aim to give an up-to-date overview on the news and views related to the topic of Aspergillus pathogenicity.

22.1 Aspergillus—A Saprophytic Pathogen?

Fungal infections and the diseases emerging from them have become an increasing threat for a certain group of the population. Especially immunocompromised individuals are prone to develop several forms of mycoses, with invasive forms being the most severe and drastic ones. Among the estimated 1.5 million fungal species,1 only a few are able to cause diseases in mammalian hosts, and the identification of factors that render these species pathogenic is a rapidly developing field in modern mycological research.2 The recent years have seen a couple of excellent publications reviewing the field of Aspergillus pathogenicity and aspergillosis, covering many aspects of the topic such as epidemiology, virulence, or disease models.3–8 It is not the intention of this chapter to reiterate these data in detail but to extract and condense them to focus on attributes that support virulence or determine the allergenic potential of Aspergillus.
The predominant fungi to cause fungal infections are Candida and Aspergillus species, the latter having evolved over the last two decades as the major pathogen in distinct clinical settings such as solid organ transplantation, allogeneic bone marrow transplantation, acute leukemia, or immunosuppressive therapy. In a retrospective survey, an exponential trend in mortalities due to Aspergillus spp. was evident from 1980 to 1995 with a peak at 0.42 deaths per 100,000 population, which equals a 375% increase. Estimation of the costs yielded an average burden of $65,000 per case of invasive aspergillosis (IA) totaling to $633 million in the United States during 1996. Systemic infections with Aspergillus still result in high mortality rates of 30% to 90%, which is based on inadequate diagnostic capabilities and restricted therapeutic options.

Besides being opportunistic aggressors, aspergilli bear the potential to cause allergic responses, making them unique microbial pathogens according to the damage-response framework concept (see later). Moreover, their basic lifestyle is that of a saprophyte, raising the question whether Aspergillus represents a true pathogen at all. The answer to this may lie in the viewpoint on the interplay of the fungus and its environment: pathogenicity strictly relies on a host to be infected and damaged, therefore in the setting of Aspergillus colonizing this specific ecological niche it has to be regarded as a pathogen; in case of fungal proliferation in the absence of a host, saprophytic propagation might be used as proper description.

### 22.2 Aspergillosis—Disease Forms and Defense Lines

Aspergilli are ubiquitous molds able to colonize a broad spectrum of substrates, and this mirrors their capacities to infect different kinds of hosts, such as mammals, birds, or even sea fans. The primarily airborne, infectious propagules are the asexual uninucleate spores, which are released in vast amounts from the conidiophores into the environment. Several kinds of disease caused by Aspergillus species exist that can manifest both in immunocompetent as well as in immunocompromised individuals.

Invasive forms in which tissue is colonized and eventually invaded are generally called aspergilloses, but this term also comprises allergic disease forms. From this, the so-called aspergilloma or “fungus ball” has to be distinguished, which is the growth of solid masses of fungal mycelia inside preformed cavities of the lung.

Allergic forms may be subdivided into Aspergillus asthma, extrinsic allergic alveolitis, allergic fungal sinusitis, hypersensitivity pneumonitis, and, predominantly, allergic bronchopulmonary aspergillosis (ABPA).

Invasive and saprophytic forms of aspergillosis can be classified on the basis of pathology and pathogenesis. Among the affected organs, the lung and the lower respiratory tract are the predominant sites of infection and disease manifestation, and disease forms range from invasive pulmonary aspergillosis (IPA) with or without angioinvasion to chronic noninvasive forms. Aspergilloses of the upper respiratory tract include acute invasive sinusitis, chronic invasive or granulomatous sinusitis, and sinus aspergilloma. Other forms of the disease may affect the cardiovascular system by, for example, colonizing native or prosthetic valves, or the central nervous system; furthermore, cutaneous, ocular, osteoartic-ular, genitourinary, or gastrointestinal forms of aspergillosis have been described. The most severe form of IA is the systemic, disseminated one, which is characterized by the occurrence of IA at two or more noncontiguous sites.

The fact that aspergilli are able to cause disease in the absence as well as presence of an immune response places them in a distinct group within the damage-response framework (Fig. 22.1). This concept relates the extent of the immune response to the degree of damage or benefit that is manifested in the host infected by a microbial pathogen. In the case of an impaired immune system such as neutropenia, types of IA may develop. In contrast, allergic complications such as ABPA arise in the case of a hyper-active immune response. Several additional forms of aspergillosis can be placed along the framework’s parabola, which therefore provides a useful concept to comprehensively describe the outcome of this specific host-microorganism interaction.

Once inhaled by an immunocompetent individual, Aspergillus conidia face two lines of defense by effector cells of the innate immune system; resident alveolar macrophages and neutrophilic granulocytes,
Pathogenicity Determinants and Allergens

FIGURE 22.1 Aspergillus is a unique pathogen. Note: In the damage-response framework of microbial pathogenesis, Aspergillus spp. exhibit a unique curve progression when plotting the outcome of infection versus the host’s immune status. In case the immune response is severely impaired, serious and systemic forms of IA may develop, whereas in the case of a hyper-reactive immune response, allergic forms like ABPA can arise.

The latter being recruited to the site of inflammation (Fig. 22.2). Both cell types support mucociliary clearance to eradicate fungal spores that are inhaled continuously. Macrophages ingest and kill swollen but ungerminated conidia in the phagolysosomal compartment, whereas neutrophils contact hyphal surfaces once a germ tube has emerged from an escaped spore to result in a respiratory burst and eventually degranulation. Accordingly, prolonged neutropenia, treatment with immunosuppressive agents, highly dosed corticosteroids, or forms of leukemia may result in an impairment of these cellular defense

FIGURE 22.2 Schematic presentation of defense lines of the innate immunity against Aspergillus. Note: After inhalation, asexual spores of Aspergillus spp. are able to reach the lower respiratory tract down to the alveoli. There, two types of phagocytic cells are present in immunocompetent individuals to defeat infection and to eliminate the fungus. First, alveolar macrophages (AM) phagocytose the conidia (C) to entrap them in the cellular compartment of the phagosome. By fusion with endosomes phagolysosomes are formed, in which swollen conidia (sC) are destroyed by mainly nonoxidative mechanisms. Second, conidia that are able to escape their elimination by alveolar macrophages to germinate attract polymorphonuclear leukocytes, also called neutrophils (N); these are recruited to germ tubes (GT), attach to them, and damage them predominantly by oxidative mechanisms and eventually degranulation to result in killing of the hyphae.
lines and enable fungal germination and growth to establish one or more forms of aspergillosis.26–28 The Toll-like receptors TLR2 and TLR4 as well as the dectin-1 receptor are involved in recognition of the conidia by these effector cells via pathogen-associated molecular patterns such as cell wall components, and a crucial role for the long pentraxin PTX3 in conidia recognition during the innate antifungal immune response could be validated as well.29–32

For allergic reactions such as ABPA, the mechanism of immunopathogenesis is not fully understood. With respect to hypersensitivity classification, ABPA represents a mixed immune response of the type 1, 3, and 4. Underlying diseases are often asthma or cystic fibrosis, and exposure to Aspergillus allergens (see later) results in a Th2 response accompanied by elevated total serum IgE levels, anti-Aspergillus antibodies of the IgE and IgG type, and eosinophilia.33 Probable genetic risk factors have been described that may contribute to the skewing of Th2 responses to Aspergillus allergens.34

22.3 Aspergillus fumigatus and Its Pathogenic Relatives

The genus Aspergillus comprises about 180 accepted species, but only few of them, approximately 40, are described to be able to cause disease.35 By far the most frequent one is A. fumigatus, which accounts for about 80% of human opportunistic mycoses provoked by Aspergillus.36 Spores of A. fumigatus are ubiquitously distributed over the world but account for a small fraction of environmental Aspergillus conidia.37 There is apparently no distinct difference between environmental and clinical isolates as determined by DNA fingerprinting, which rules out the possibility that certain isolates enriched in hospital settings express specific virulence traits.38,39 In particular institutions other species of Aspergillus may predominate, and these non-fumigatus species are identified at increasing frequencies to cause disease in immunocompromised hosts.40 In correlation to their environmental occurrence,41 A. flavus, A. terreus, and A. niger are commonly reported to cause aspergilloses. Interestingly, the model ascomycete A. nidulans is most frequently identified as causative agent of aspergillosis in chronic granulomatous disease (CGD) patients, which carry a genetic defect in the respiratory burst of phagocytes.42 Furthermore, A. ustus has been isolated from cases of IPA and primary cutaneous aspergillosis.43

It is noteworthy that the model fungus A. nidulans is able to cause disease in general experimental settings using neutropenic mice, albeit at about 10-fold higher spore counts administered, which enables virulence studies employing isogenic A. nidulans mutants.44 The ability of several species to cause disease in an immunosuppressed individual implicates that under appropriate conditions any Aspergillus species can provoke different forms of aspergillosis; however, the fact that A. fumigatus is by far the most commonly identified species in pulmonary mycosis although its relative abundance among environmental Aspergillus conidia is low, is in favor of the existence of specific cellular attributes that support its growth inside the ecological niche “immunocompromised host.”

22.4 What Makes Aspergillus Virulent?

Descriptive terms to illustrate the pathogen–host relationship for an opportunistic pathogen are often confusing and unexact, and several concepts and definitions have been suggested during the history of medical microbiology to describe the microbial trait of pathogenicity comprehensively.45,46 Especially, factors that determine virulence of fungal opportunistic pathogens are hard to define, as the host’s immune status is crucial for the outcome of infection; moreover, general as well as specific cellular attributes of the fungus have a large impact on its survival inside the hostile environment of an infected individual. In this chapter, the term “virulence determinant” is used in a broad sense to describe gene products and cellular aspects of Aspergillus that were characterized to support its capacity to cause disease in an immunocompromised host. This includes common traits that account for the physiological versatility of this fungus or its saprobic lifestyle, although these features represent factors that are required for growth in general. However, by the identification of such features, deeper insights into the specific fungal
requirements during disease are gained to assist in elucidating the fungal side in the pathogen-host system of aspergillosis.

Virulence of *Aspergillus* is multifactorial, and several key features contribute to its capacity to infect and damage a susceptible host (Fig. 22.3). As a general factor, given the fact that a correlation of growth rate and virulence of *A. fumigatus* is evident, fungal physiology and metabolic flexibility have to be considered, since fitness and, therefore, growth strictly relies on this. A basic feature that supports infection is the fungal morphology: the small size of the conidia of about 2–3 µm assists in reaching the primary site of infection during IPA, the alveoli, and the hyphal growth mode of this filamentous fungus may support tissue penetration and eventually dissemination. In a recent comparative study to identify distinctive determinants of *A. fumigatus* pathogenicity, no unique gene products or exclusive pathways could be identified. Based on this, environmental and therefore host factors have to be considered to contribute significantly when the saprophyte *A. fumigatus* becomes pathogenic.

### 22.5 Thermophyly—A Key Feature of Fungal Pathogenicity

As a general feature of pathogenic fungi, the capacity to sustain elevated temperatures as they are present inside a mammalian host is evident. Most fungal species have an optimal growth temperature of 25–35°C but the basal body temperature of warm-blooded animals is above this level. Pathogenicity of aspergilli is made up by a set of attributes, and its thermophyly is a strict precondition to propagate inside the host and express them. *A. fumigatus* is able to survive at temperatures of up to 75°C, and growth can be maintained at 55°C while it is rapid at 37°C. A comprehensive profiling study has tried to elucidate the transcriptional reprogramming of *A. fumigatus* upon a shift from 30°C to 37°C or 48°C. Interestingly, except for a catalase-encoding one, no gene that has been characterized before to be involved in pathogenicity displayed a higher expression level at 37°C than at 48°C, which implies that a specific response toward the host’s body temperature does not exist. In particular, two genes of *A. fumigatus* have been characterized with respect to their contribution to thermotolerance, *thtA* and *cgrA*. The *thtA* locus has been identified by a genetic screen in search for genes required for growth at elevated temperatures, and its gene product renders the fungus tolerant to a temperature of 48°C but is not required for virulence. The *cgrA* gene product plays a role in ribosome biogenesis, and a corresponding deletion mutant displayed reduced virulence in murine model of pulmonary aspergillosis.
22.6 Contacting the Host

The contact with the host is mediated by the conidial surface or the hyphal cell wall, and both cellular structures are highly specific for fungi and essential for growth of Aspergillus. When infecting a susceptible host, conidia first encounter epithelial barriers and adhere to them. Accordingly, much attention has been drawn to this particular pathogen/host interplay, mainly focusing on conidial and their molecular interaction with extracellular matrix (glyco)proteins such as fibronectin, laminin, type I and type IV collagen, or fibrinogen.55,56 Lung epithelial cells synthesize and secrete basolaterally fibrinogen, which is actually the major plasma glycoprotein and involved in coagulation, and A. fumigatus conidia specifically bind fibrinogen in a saturable manner.57 Moreover, they are able to adhere by a specific cell wall glycoprotein to laminin, a component of basement membranes that underlie epithelia and endothelia.58 A tempting assumption is that wounded epithelia expose subepithelial membrane structures and deposit fibrinogen and laminin, to which A. fumigatus conidia can adhere to initiate colonization. Another glycoprotein that is present in extracellular matrices is fibronectin, and conidia of A. fumigatus are able to adhere to its fibrillar form via two polypeptides in an RGD-dependent manner.59 Furthermore, a specific lectin was found to be present at the surface of conidia, which probably mediates binding of terminal sialic acid residues of glycoconjugates as they are present in fibrinogen and laminin.60 However, detailed molecular knowledge on the ligands or on genes that encode fungal receptors to mediate specific interactions with extracellular matrix components is still scarce.

Hydrophobins are structural parts of the outer conidial layer, which appears as rodlets, and conidia from A. fumigatus mutants impaired in hydrophobin synthesis are described to be more sensitive toward killing by alveolar macrophages.61,62 Interestingly, mutants impaired in pigment synthesis due to a deletion of the glucan synthase activity, which is encoded by the essential chsC gene, display altered hydrophobin expression, and filamentation of the ΔchsC mutant strain, suggesting a regulatory role of the pigment on conidial surface hydrophobicity.63,64 Although mutations in the hydrophobin genes encoded by theags1, ags2 and corresponding mutant strains have been studied in detail,65-67 among the three hydrophobin genes encoded by the A. fumigatus genome, onlyagsΔ mutants displayed a reduction in virulence. Galactomannan, a component of the conidial cell wall, is highly specific for fungi and essential for growth of A. fumigatus. The polymer chitin is synthesized in moulds by six different classes of chitin synthase (Chs) enzymes, and in A. fumigatus eight genes covering each class were identified.68 Among them, disruption of chsE or chsG results in distinct phenotypes, and a chsE−; chsG− double mutation displays additive phenotypes but not synthetic lethality. Reduced virulence could be determined for chsG− and chsC−; chsG− mutant strains, which were, however, still able to cause pulmonary disease in neutropenic mice despite their severe morphological variations. In contrast, wild-type virulence capacities were determined for single chsD− or chsE− mutants, illustrating the high degree of redundancy in chitin biosynthesis of A. fumigatus.69-72

In addition to its polysaccharide carcass, the A. fumigatus cell wall is interspersed with proteins, and only a few of these have been characterized in this pathogenic fungus. An A. fumigatus mutant lacking the GPI-anchored cell wall protein Ecm33 displays a range of phenotypes related to cell wall integrity
and cell–cell adherence but cell–matrix interaction being unaffected. Strikingly, it displays increased virulence in a murine model of disseminated aspergillosis, which accentuates the significance of the fungal cell wall in *Aspergillus* pathogenesis.

### 22.7 Sensing the Host

Perceiving and coping with rapid changes in the environment, or, in case of infection, the host, is vital for fungal propagation. A variety of systems have evolved in fungi to serve this task, which can be grouped according to the components that build up a signal transduction cascade. General eukaryotic stress response sensors are kinases that phosphorylate initiation factors of translation. This, in turn, results in a general translational shut-off accompanied by forced expression of specific regulators to create a cellular response with the aim to counteract the initial stress condition. Higher eukaryotes carry four types of such eIF2α kinases, each responding to a different kind of environmental challenge like oxidative or unfolded protein stress, viral infection, or nutritional depletion. The *A. fumigatus* genome encodes two eIF2α kinases, one of them being the cross-pathway control component CpcC. This sensor kinase acts in response to nutritional stress conditions to derepress expression of the transcriptional activator CpcA. As a consequence, a global transcriptional response is executed to counteract the environmental stress situation. Preliminary data indicate that neither *A. fumigatus* eIF2α kinase is necessary for or involved in virulence.

Another major mechanism by which fungi adapt to their environment is by the action of two-component phosphorylays of the histidine kinase type. These consist of a sensor domain, which contains a conserved histidine residue that becomes phosphorylated in reaction to an environmental input, and a receiver domain, which contains a conserved aspartate residue to which the phosphate group becomes transferred to create an output. *A. fumigatus* has the capacity to express 13 of these, based on similarities of deduced orthologs, but only the *fos-1* and *tcsB*-encoded ones have been studied. The *fos-1* gene is highly expressed in *in vivo*, and in a murine model of systemic aspergillosis a *fos-1Δ* deletion strain appeared significantly attenuated in virulence. No clear data on the input signal perceived by the *Fos-1* histidine kinase were evaluated.

GTP-binding and -hydrolyzing proteins represent highly conserved molecular switches within eukaryotic signal transduction cascades. Members of the Ras subfamily of such GTPases play a well-established role in morphological processes of fungi, among them pathogenic species. For *A. fumigatus*, two genes encoding Ras proteins, *rasA* and *rasB*, were identified, and by expression of dominant negative or dominant active alleles of each gene their impact on timing and morphology of asexual development could be demonstrated. Moreover, by testing a *rasBΔ* deletion mutant, which displays pleiotropic phenotypes related to germination, growth rate, and hyphal morphology, in a murine model, the necessity for proper polarized growth in IA was deduced.

Environmental stress conditions are sensed and retorted by an additional type of signal transduction cascade, which acts essentially by serial phosphorylation of three protein kinases. These mitogen-activated protein kinase (MAPK) modules are crucial for fungal physiology in response to changing environmental conditions. Among the four MAP kinases—MpkA, MpkB, MpkC, and SakA/HogA—that can be deduced from the *A. fumigatus* genome sequence, only one has been studied in detail: the SakA MAPK is necessary for the osmotic stress response but is also involved in nutritional sensing with respect to nitrogen source-dependent conidial germination or carbon or nitrogen starvation during vegetative growth. A deletion of the encoding gene does not result in an attenuated mutant strain, which indicates signaling redundancy among MAPK modules or, alternatively, the absence of environmental stress conditions during the course of infection. Chapter 23 (Protein Kinases and Pathogenesis, by G.S. May) discusses the relationship of kinases and *Aspergillus* virulence in greater detail.

Much attention has been focused on the cAMP-mediated signaling cascade of *A. fumigatus*. Growth experiments with the second messenger cAMP have shown that the signal transduction pathways and receptors to sense different carbon sources are not uniform among the aspergilli: whereas *A. niger* displayed growth reduction in the presence of cAMP and glucose, *A. fumigatus* was refractory to these conditions. Only when grown on a catabolite nonrepressing carbon source, extracellular cAMP resulted in an increase of PKA activity and reduced growth rates. First hints that this regulatory network might
contribute to pathogenicity of this fungus stem from differential display studies on \( A. \) \( fumigatus \) cells grown in the presence or absence of endothelial cells, where the expression of the gene encoding the regulatory subunit of cAMP-dependent PKA was found to be up-regulated.\(^9\) Moreover, transcript levels of the \( pkaR \) and \( pkaC \) genes, which code for the regulatory and a catalytic subunit of protein kinase A, respectively, are increased upon cocultivation with alveolar epithelial cells.\(^9\) The two genes \( gpaA \) and \( gpaB \) encoding \( \alpha \)-subunits of heteromeric G proteins and the \( acyA \) gene encoding the adenylate cyclase were characterized in detail: ingestion of conidia by human monocyte-derived macrophages and subsequent determination of survival rates revealed reduced viability rates for \( gpaB \) and \( acyA \) deletion mutants.\(^9\) Furthermore, expression of the virulence-determining factor \( PksP \), the polyketide synthase involved in conidial pigment synthesis, was shown to be reduced in hyphae that were grown under standard conditions in a \( gpaB \Delta \) background, an effect likely based on decreased intracellular cAMP levels. These findings are complemented by studies on the \( pkaC1 \) gene, which encodes a protein kinase A catalytic subunit.\(^9\) \( pkaC1 \Delta \) deletion mutants display a reduction in growth and conidiation and, furthermore, \( PksP \) expression is decreased. Most interestingly, in a low-dose inhalation model of pulmonary aspergillosis strains carrying deletions in the \( pkaC1 \) or \( gpaB \) locus are almost avirulent. Deletion of the \( pkaA \) gene that encodes the PKA regulatory subunit resulted in strains impaired in growth and germination and that were more sensitive to oxidative damage.\(^9\) Accordingly, virulence of this mutant was greatly reduced. Taken together, these findings implicate a relevance of cAMP-mediated PKA signaling for \( A. \) \( fumigatus \) virulence, an important notion that is also reflected by studies on a variety of other pathogenic fungi as well.\(^9,10\)

Changes in ambient pH represent a severe environmental stress as pH homeostasis is crucial for all intracellular processes. A large body of research has been carried out on the pH response of \( A. \) \( nidulans \), which strictly depends on proteolytic processing of the transcription factor PacC under alkaline conditions to result in induction of a subset of alkaline expressed genes and repression of a subset of acid expressed ones.\(^10\) Using isogenic strains of \( A. \) \( nidulans \) in a murine model of pulmonary aspergillosis, Bignell et al. (2005) were able to demonstrate that PacC itself, its processing or the ambient pH signal transduction pathway is required for virulence.\(^4\) Accordingly, expression of a constitutively activated allele of \( pacC \) that mimics response to alkaline conditions resulted in enhanced virulence and extensive fungal growth in the lungs of infected animals. The high degree of conservation of this signaling cascade among aspergilli supports the conclusion that PacC as pH-responsive transcription factor is also key for \( A. \) \( fumigatus \) virulence.

A variety of other cellular stress responses and in particular calcium signaling is mediated in eukaryotes by the calcineurin pathway.\(^2\) Calcineurin is a serine-threonine-specific protein phosphatase that is activated by \( \text{Ca}^{2+} \) and calmodulin. It has a heterodimeric composition, consisting of a catalytic \( \alpha \) subunit and a regulatory, \( \text{Ca}^{2+}\)-binding \( \beta \) subunit, and is highly conserved among eukaryotes and the fungal kingdom. Targeting and deleting of the \( \alpha \) subunit-encoding gene \( calA/cnaA \) in \( A. \) \( fumigatus \) resulted in severely impaired mutant strains that displayed defects in conidiation and filamentation as well as increased branching.\(^3\) Moreover, and in line with the growth defect of the calcineurin \( \alpha \) mutant, virulence was clearly reduced in different animal models of aspergillosis, reflected by decreased invasion and fungal burdens.\(^3\)

### 22.8 Feeding from the Host

Physiology has been studied in several \( Aspergillus \) species over decades, and it is of logical consequence that the results and insights gained from scrutinizing the metabolism were assigned to study pathogenicity mechanisms of \( A. \) \( fumigatus \). In a pilot study, Purnell assessed virulence capacities of auxotrophic \( A. \) \( nidulans \) mutant strains in a systemic infection model.\(^2\) Some of these data were supported in later studies using defined deletion mutants of \( A. \) \( fumigatus \), and several biochemical pathways of primary metabolism were shown to be required for growth inside the murine lung. Among them are lysine and pyrimidine biosynthesis and synthesis of the folate precursor \( para\)-aminobenzoic acid.\(^2\) The citric acid cycle activity encoded by the \( mcsA \) gene product was shown to be required for full virulence in an alternative infection model using caterpillars of the greater wax moth \( Galleria \) \( mellonella \).\(^2\) Related to this, the \( A. \) \( fumigatus \) gene coding for the glyoxylate bypass enzyme isocitrate lyase \( acuD \) has been
Pathogenicity Determinants and Allergens

characterized in detail, as enzymes of this metabolic shunt are putative targets for antifungal therapies.\textsuperscript{110} However, the virulence of respective mutant strains has not been assessed to date.

In accordance with their natural habitat, aspergilli are able to utilize a broad range of carbon sources. As soil-borne organisms, they especially have to deal with complex substrates, and the plethora of enzymes secreted by the fungi to degrade polymeric and oligomeric substrates like (hemi)cellulose, pectin, lignin, starch, or sucrose gives reference to this fact.\textsuperscript{111} Inspection of the \textit{A. fumigatus} genome sequence has revealed a huge array of genes encoding different glycosylhydrolase activities that allow degradation of these components, which are predominantly found in plant cell walls.\textsuperscript{48} Generally, low levels of these degrading enzymes are expressed and secreted into the environment to produce small molecules by the breakdown of the substrate, which in turn act as molecular signals for increased enzyme synthesis to accelerate polymer degradation. Facilitated by specific uptake systems, sugars are transported into the cell to be metabolized by a range of different pathways.\textsuperscript{112} Utilization of hexoses like glucose and fructose has been studied in \textit{Aspergillus} in great detail, and key enzymes and genes of glycolysis, pyruvate metabolism, the pentose phosphate pathway, and others have been characterized in different species.\textsuperscript{113} Environmental changes in glucose availability are sensed by the cAMP/protein kinase A (PKA) pathway, a conserved signal transduction cascade counteracting nutritional stress situations (see above). Furthermore, aspergilli are generally able to grow on alcohols, with the exception of methanol, or in the presence of polyols like glycerol, and studies on the ethanol utilization pathway and its regulation have emerged as a prime example for carbon catabolite repression.\textsuperscript{113} This global regulatory system ensures that preferred carbon sources like glucose or xylose are preferentially utilized by the fungus. Studies on defined mutants of the PKA signal transduction pathway imply that a favored carbon source might be a limiting factor for \textit{in vivo} growth of \textit{A. fumigatus}.\textsuperscript{114} Nevertheless, defined mutant strains that are impaired in metabolizing specific carbon sources were not tested for their virulence capacities, although these metabolic features are likely to support growth and dissemination within an infected host.

The role of nitrogen metabolism in supporting \textit{Aspergillus} virulence has been addressed in several studies.\textsuperscript{115} Gene products and corresponding mutants that were tested for reduced virulence include the nitrate assimilation regulator \textit{AreA}, the Ras-related protein \textit{RhbA}, which is required for growth on poor nitrogen sources, and the regulator \textit{CpcA} of the cross-pathway control system of amino acid biosynthesis.\textsuperscript{81,116–118} Infections using \textit{areA} mutant strains of \textit{A. fumigatus}, which are accordingly unable to utilize certain sources of nitrogen, resulted in delayed development of pulmonary aspergillosis, suggesting that an \textit{AreA}-regulated metabolic pathway is required for proper utilization of nitrogen from the lung tissue. Mutants of \textit{A. fumigatus} deleted in the Rheb homolog \textit{RhbA} were severely attenuated in virulence due to limited \textit{in vivo} growth inside the lungs of immunosuppressed mice, which also supports the necessity of accurate nitrogen signaling for virulence. Fungal amino acid homeostasis is regulated in a global cellular manner via action of a conserved regulatory system termed cross-pathway control. This serves to sense nutritional stress conditions, such as amino acid starvation, by an eIF2\textalpha{} kinase, \textit{CpcC}, to result in elevated expression levels of a transcriptional activator, \textit{CpcA}, which creates a transcriptional read-out to neutralize the cellular stress.\textsuperscript{119} Mutants of \textit{A. fumigatus} deleted for the \textit{cpcA} gene appeared attenuated in a murine model for pulmonary aspergillosis, demonstrating the general requirement of the cross-pathway control regulator for full virulence. As regulation of \textit{CpcA} expression is complex,\textsuperscript{120} the necessity of a derepressed cross-pathway control system cannot be deduced from these data, leaving the question open as to what extent the host’s lung is a nutritionally poor environment.

22.9 Damaging and Fighting the Host

Specific studies on growth substrates for \textit{A. fumigatus} are incomplete and only a few aspects have been addressed particularly in this species. A survey on different carbon sources had been undertaken by Demain and coworkers with the aim to study the effect on fumagillin synthesis, a fungal antibiotic produced by \textit{A. fumigatus}.\textsuperscript{121} In total, 29 carbon sources (poly-, oligo-, and monosaccharides as well as organic acids) in chemically defined media were tested as well as mixtures containing a secondary carbon source. Studies on more complex substrates revealed that \textit{A. fumigatus} is able to utilize cellulose, bark, aromatic compounds, melanin, or even chicken feather keratin.\textsuperscript{122,123} Considering the \textit{in vivo}
situation where A. fumigatus germinates and invades the host’s lung, elastin and collagen are the predominant macromolecules that have to be metabolized by the fungus. Therefore, proteases that degrade these biopolymers are important in promoting growth in vivo and an extensive body of research has focused on these enzymatic activities as virulence-determining factors. Three main proteolytic activities are secreted by A. fumigatus in vivo, the Alp alkaline serine protease, and the Mep metalloprotease, and the Pep aspartic protease. Expression of these proteases is repressed by the presence of free amino acids or oligopeptides, whereas they are secreted extensively when protein is the sole source of nitrogen. Alp is a subtilisin-like protease with elastinolytic activity that is secreted in the germ tube and the hyphal apex. The structural gene could be cloned and inactivated by disruption to create a mutant exhibiting strongly reduced proteolytic activity. From such a mutant, Mep was isolated to be characterized as a collagen-cleaving activity. The mep gene could also be inactivated in a wild-type background and an alp-strain, and the single mutant still retained 70% of wild-type levels activity in vivo while the double mutant lacked neutral extracellular proteolytic activity completely. The Pep aspartic protease belongs to the pepsin protease family (“aspergillopepsin”) and is secreted in germ tubes and penetrating hyphae. For all three of these proteolytic activities, additional members with their corresponding genes have been identified in A. fumigatus, such as Alp2 and Exalp as alkaline serine proteases, Mep20 and MepB as metalloproteases, or Pep2 as a cell wall-associated aspartic protease. Moreover, other degradative activities have been detected like dipeptidyl-peptidase enzymes (DppIV and DppV) or phospholipases (Plb1, Pl2, and Plb3). Especially with respect to the secreted proteases, their role in pathogenicity of A. fumigatus is still questionable, although some mutant strains have been created and were subjected to virulence studies. Single alp-, mep-, or pep- mutants retained full virulence in animal models as did an alp-; mep-; pep- double knock-out strain. Based on the fact that a whole array of proteolytic activities encoded by several genes can be produced by A. fumigatus, this implies some degree of redundancy and the ability to compensate the loss of one particular activity. As a consequence, and as already suggested by Monod et al., the complete inactivation of a family of genes by subsequent gene targeting seems necessary to obtain conclusive results.

Fungi and, in particular, aspergilli display a complex and capacious secondary metabolism yielding several characteristic and biologically active compounds such as polyketides, nonribosomal peptides, terpenes, or indole alkaloids. The contributions of these molecules to virulence of Aspergillus has been studied to a limited extent, but along with the annotation of secondary-metabolite genes and clusters in the genome sequences, defined mutant strains deficient in metabolite production have become feasible. For instance, an ergot biosynthetic gene of A. fumigatus, dmaW, has been characterized and a corresponding mutant strain was devoid of all ergot alkaloids this fungus is able to produce, but its virulence capacity has not been assessed. Mycotoxins are probably the most attractive substances with respect to the pathogen/host interaction. Toxic substances associated with conidia of A. fumigatus have been described, and some of them were characterized with respect to their effect on macrophages and phagocytosis. Additionally, different toxins released by the hyphae could be identified in culture filtrates of A. fumigatus. Of all the mycotoxins produced by A. fumigatus, five immunosuppressive ones—gliotoxin, fumagillin, helvolic acid, fumitremorgin A, and Asp-hemolysin—could be identified up to now. The alkaliaromatic derivative gliotoxin shows the strongest immunosuppressive activities as to specifically inhibit NADPH oxidase and to interfere with superoxide release, migration, leukocyte cytokine release, or T-lymphocyte-mediated cytotoxicity. Moreover, gliotoxin is described as genotoxic and able to trigger apoptosis in macrophages. However, data on its in vivo production are scarce, which challenged the assumption that gliotoxin acts as a true virulence factor in terms of host damaging. By genetic analysis and generation of a strain disrupted in the nonribosomal peptide synthase-encoding gene gliP, an A. fumigatus mutant completely abolished in gliotoxin production could be tested for virulence. No significant difference of this mutant in comparison to its wild-type progenitor was observed in an inhalation model using neutropenic mice, demonstrating that gliotoxin biosynthesis is not required for virulence of A. fumigatus in this experimental setting.

Ribotoxins are enzymatic activities that inactivate ribosomes by cleavage of a phosphodiester bond in the large subunit, and they are produced by aspergilli as well. The major ribotoxin of A. fumigatus is mitogillin (syn. restrictocin), a 18 kDa protein encoded by the mitF/res gene. Yet, as deduced
from virulence studies using disruption mutants that do not produce this cytotoxin, no important role for this ribotoxin in IPA could be deduced.\textsuperscript{165,166}

An assortment of additional toxic substance is synthesized by \textit{Aspergillus} species, but experimental evidences on their involvement in pathogenesis is lacking for most of them. Here again, redundancies and synergistic effects of toxins and secondary metabolites that are secreted into the environment may hamper a clear correlation of particular molecules to virulence.

The main lines of host defense rely on effector cells of the innate immune system and ROS produced by them. Accordingly, neutralizing antioxidant activities have been in the focus in defining detoxifying systems as virulence factors.\textsuperscript{167} The \textit{A. fumigatus} genome bears the information for two pathways to counteract reactive oxidants as they are generated in phagocyte defense, one using oxidases such as catalases, peroxidases, or superoxide dismutases, the other acting via glutathione.\textsuperscript{48} For the potential catalases, three have been studied in detail: CatA is expressed in conidia, whereas Cat1/CatB and Cat2 are mycelial.\textsuperscript{166,167} Conidia from cat\textit{A}\textsuperscript{−} mutants appear more sensitive toward \textit{H}_2\textit{O}_2 but are not killed more effectively by alveolar macrophages, therefore, eliminating this enzyme as virulence determinant. A cat\textit{I}\textsuperscript{−}; cat\textit{A}\textsuperscript{−} double mutant was also more sensitive against \textit{H}_2\textit{O}_2 treatment and, moreover, attenuated in an experimental rat model; however, this mutant was as easily destroyed by polymorphonuclear neutrophils as a wild-type strain. Superoxide dismutases (SOD) have been suggested to play a more profound role in protecting hyphae against ROS. Two types of SOD were identified in \textit{A. fumigatus}, both being immunoreactive in humans: one Cu/ZnSOD, which is extracellular, and two intracellular MnSODs.\textsuperscript{170–172} Moreover, numerous genes putatively encoding components of the glutathione pathway were annotated in the \textit{A. fumigatus} genome sequence, and three genes encoding glutathione transferases, \textit{gstA-C}, have been characterized by defining enzymatic activities from recombinant proteins.\textsuperscript{173} For both neutralizing systems, SOD- or glutathione-mediated, data on virulence of corresponding mutant strains are not available, yet. Related to this, the product of the nonribosomal peptide synthase gene \textit{pesI} was recently described to confer protection against oxidative stresses in \textit{A. fumigatus}; however, besides a conidial morphology phenotype no references have been gained on its actual cellular role.\textsuperscript{174}

As this arsenal of enzymatic activities to counteract oxidative stress appears highly flexible and effective, it is likely to contribute to virulence of \textit{A. fumigatus}. However, similar systems are generally present in nonpathogenic fungi, and thus they may not be assigned as specific virulence factors. Nevertheless, this may not be case for unique aspects of \textit{A. fumigatus} physiology.

It may well be the case that aspergilli, and in particular \textit{A. fumigatus}, express factors specifically to modulate the host’s defense.\textsuperscript{5} Gliotoxin, for instance was shown to interfere with mucociliary beating, which represents a first mechanical barrier in infection.\textsuperscript{175} Proteases support epithelial damage and inflammation by activation of epithelial cells.\textsuperscript{176} The capacity to counteract the primary immune defense lines to a certain extent contribute to virulence of \textit{A. fumigatus}, as this trait is expressed when these defense lines are weakened. Particularly, intervention with phagocytosis events or with opsonization assists in establishing fungal growth. Besides the described action of toxins or antioxidant pathways, additional factors have been characterized to specifically interfere with components of these immune actions: a “conidial inhibitory factor” impedes phagocytosis and superoxide anion and hydrogen peroxide production by phagocytes in response to \textit{A. fumigatus};\textsuperscript{177–179} the \textit{A. fumigatus} diffusible product \textit{AfD} inhibits transcription of genes-encoding proinflammatory cytokines in rat alveolar macrophages;\textsuperscript{180} a diffusate released from \textit{A. fumigatus} conidia reversibly diminishes ingestion by alveolar macrophages in a time- and dose-dependent manner;\textsuperscript{181} binding of \textit{Aspergillus} conidia by C3, the key component of the alternative complement pathway, appears to be less for the pathogenic species \textit{A. fumigatus} and \textit{A. flavus}, and from hyphae of \textit{A. fumigatus} a factor could be isolated that interferes with C3b-binding and C3b-dependent phagocytosis and killing.\textsuperscript{182–184} For none of these factors, however, detailed molecular data such as sequences of the encoding genes are available, hampering investigation of mutant strains and clear correlation to virulence.

A more detailed analysis with respect to modulation of the host’s immune response by \textit{A. fumigatus} was carried out by scrutinizing the role of oxylipins and prostaglandins. In higher eukaryotes, prostaglandins are characterized to be involved in numerous cellular processes, among them regulation of inflammation or allergic responses.\textsuperscript{185} Accordingly, the observation that eukaryotic pathogens produce prostaglandin derivates upon infection suggests that these bioactive lipids play a role in the interplay between host and infecting microbe.\textsuperscript{186} Three \textit{ppo} genes could be identified in the \textit{A. fumigatus} genome that encode fatty
acid oxygenases similar to mammalian synthetases for prostaglandins. By silencing expression of these \textit{ppo} genes, mutant strains with slightly decreased prostaglandin production were generated, and in a murine model of pulmonary aspergillosis these appeared to be hypervirulent, accompanied by increased conidial resistance toward oxidative stress as induced by H$_2$O$_2$. However, a precise mode of action of \textit{A. fumigatus} prostaglandins in modulating the host’s immune response waits to be demonstrated.

### 22.10 Special Features

Besides general factors of fungal lifestyle, some cellular features of \textit{A. fumigatus} in particular could be identified that seem specifically required for survival in the host’s environment. For these, the term “virulence factor” seems more appropriate, although alternative definitions of this term emphasise host damage as an effect of a true virulence factor.

As mentioned earlier, conidia of environmental \textit{A. fumigatus} isolates are characterized by their gray-green pigmentation, which is arises from a pentaketide melanin that is similar to 1,8-dihydroxynaphthalene. Interestingly, white mutants that are unable to produce this pigment because of an impaired \textit{PksP} (\textit{Alb1}) polyketide synthase are more sensitive against oxidative stress conditions and damage by murine macrophages or human monocytes; moreover, the corresponding mutant was shown to be attenuated in virulence in an intravenous infection model of mice. However, as melanin by itself seems not sufficient to render \textit{Aspergillus} pathogenic, an additional role of the \textit{pksP} gene product might contribute to virulence of \textit{A. fumigatus}.

The ability to acquire, store, and mobilize iron is a strict prerequisite of microbial pathogenesis. Fungi have evolved highly complex systems to cope with toxicity of this element and its low abundance in the environment due to limited solubility. Aspergilli synthesize siderophores to chelate, transport, or store ferric ions, and the iron utilization system is conserved among \textit{Aspergillus} species. When scrutinizing the iron mobilization and reductive iron assimilation systems in \textit{A. fumigatus}, striking differences with respect to virulence were discovered: deletion of the high affinity iron permease-encoding gene \textit{ftrA} had no effect on virulence in a murine model of pulmonary aspergillosis, whereas elimination of hydroxamate-type siderophore biosynthesis by deletion of the \textit{sidA} gene resulted in complete avirulent strains. Especially the fact that the fungal siderophore system is absent in mammals makes it an attractive target for antifungal therapy.

An additional attribute of aspergilli is their multifaceted secondary metabolism. In a screen in search for \textit{A. nidulans} mutants impaired in the production of the carcinogen sterigmatocystin, a global regulator of the production of secondary metabolites could be identified, the \textit{laeA} gene product. Further investigations revealed that also the expression of \textit{Aspergillus} toxins depends on the presence of LaeA, making it a prime candidate for a virulence-determining factor. Expanding the research to \textit{A. fumigatus} demonstrated that LaeA is required for virulence and that conidia from corresponding mutant strains were more effectively ingested by human monocyte-derived macrophages. Moreover, hyphal killing of polymorphonuclear neutrophils (PMNs) was reduced, but the ability of PMNs to kill hyphae of a \textit{laeA} mutant was as effective as for a wild-type strain. Interestingly, expression was delayed for the conidial surface associated hydrophobin genes \textit{rodA} and \textit{rodB} as well as for the conidial pigmentation gene \textit{alb1}, which is identical to \textit{pksP}, accompanied by alterations in adherence and the conidial surface.

In summary, an array of factors determines the outcome of the host/fungus interplay once a susceptible individual has been infected by \textit{Aspergillus} spores (Table 22.1). Risk factors and underlying diseases that support fungal growth have been defined and categorized in clinical studies, and certain fungal attributes that support establishment and dissemination of the disease could be identified (Fig. 22.4). However, far more details on this interlaced relationship need to be evaluated to open novel and promising therapeutic perspectives.

### 22.11 Sensitizing the Host

\textit{Aspergillus} is one of the most common fungal genus involved in allergic responses. Particularly \textit{A. fumigatus} is characterized as one of the major sources for fungal allergens, and its repertoire on
TABLE 22.1
List of Genes Described to be Involved in *Aspergillus* Pathogenicity

<table>
<thead>
<tr>
<th>Category</th>
<th>Gene</th>
<th>Function</th>
<th>Accession No.</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Metabolism</strong></td>
<td>lycF</td>
<td>Homoaconitase</td>
<td>CAC48042</td>
<td>[108]</td>
</tr>
<tr>
<td>Citric acid cycle</td>
<td>mcsA</td>
<td>Methylcitrate synthase</td>
<td>CAI61947</td>
<td>[109]</td>
</tr>
<tr>
<td>Folate biosynthesis</td>
<td>pabaA</td>
<td>PABA synthetase</td>
<td>AAD31929</td>
<td>[106]</td>
</tr>
<tr>
<td>Pyrimidine biosynthesis</td>
<td>pyrG</td>
<td>OMP decarboxylase</td>
<td>CAA72161</td>
<td>[107]</td>
</tr>
<tr>
<td><strong>Secondary Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pigment synthesis</td>
<td>pksP/alc1</td>
<td>Polyketide synthetase</td>
<td>AAC39471/CAA76740 (190,193)</td>
<td></td>
</tr>
<tr>
<td>Oxilipin biosynthesis</td>
<td>ppo</td>
<td>Fatty acid oxygenases</td>
<td>XM_746657, XM_741345</td>
<td>(187)</td>
</tr>
<tr>
<td>Siderophore biosynthesis</td>
<td>sidA</td>
<td>L-ornithine N(^2)-monooxygenase</td>
<td>AAT84594</td>
<td>[197]</td>
</tr>
<tr>
<td><strong>Cell Wall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chitin biosynthesis</td>
<td>chsG</td>
<td>Chitin synthetase</td>
<td>CAA63928</td>
<td>[75]</td>
</tr>
<tr>
<td>Cell wall integrity</td>
<td>ecm33</td>
<td>GPI-anchored cell wall protein</td>
<td>Q4WNS8</td>
<td>[78]</td>
</tr>
<tr>
<td><strong>Stress Resistance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxidative stress</td>
<td>cat1, cat2</td>
<td>Catalases</td>
<td>AAB71223, AAM95780</td>
<td>[169]</td>
</tr>
<tr>
<td><strong>Signal Transduction and Regulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrogen metabolism</td>
<td>areA</td>
<td>Transcription factor</td>
<td>EAL85842</td>
<td>[116]</td>
</tr>
<tr>
<td>Ca(^2+) signaling</td>
<td>calA/cnaA</td>
<td>Calcineurin subunit</td>
<td>XP_753703</td>
<td>[104]</td>
</tr>
<tr>
<td>Cross-pathway control</td>
<td>cpcA</td>
<td>Transcriptional factor</td>
<td>AAQ14858</td>
<td>[81]</td>
</tr>
<tr>
<td>Secondary metabolism</td>
<td>laeA</td>
<td>Methyltransferase</td>
<td>AAR01218</td>
<td>[64]</td>
</tr>
<tr>
<td>Nutrient sensing</td>
<td>rhaA</td>
<td>Ras/Rheb GTPase</td>
<td>AAN17787</td>
<td>[117]</td>
</tr>
<tr>
<td>Morphology</td>
<td>rasB</td>
<td>Ras GTPase</td>
<td>AAP94030</td>
<td>[89]</td>
</tr>
<tr>
<td>cAMP signaling</td>
<td>pkaCl</td>
<td>PKA catalytic subunit</td>
<td>CAC82611</td>
<td>[97]</td>
</tr>
<tr>
<td>cAMP signaling</td>
<td>gpaB</td>
<td>G(_3) protein subunit</td>
<td>CAC81805</td>
<td>[96]</td>
</tr>
<tr>
<td>cAMP signaling</td>
<td>guvA</td>
<td>adenylate cyclase</td>
<td>CAC81748</td>
<td>[96]</td>
</tr>
<tr>
<td>pH signaling</td>
<td>pacC</td>
<td>Transcription factor</td>
<td>CAA87390</td>
<td>[44]</td>
</tr>
<tr>
<td>?</td>
<td>fos-1</td>
<td>Histidine kinase</td>
<td>AAK27436</td>
<td>[84]</td>
</tr>
</tbody>
</table>

molecules sensitizing the host’s immune system is large and complex. Historically, galactomannan as essential polysaccharidic part of the *Aspergillus* cell wall was the first antigen to be detected in patients with IA, and its detection in body fluids of patients like serum, urine or bronchoalveolar lavage is still a valuable and useful diagnostic criterion. In a phage surface display-based screen, at least 81 cDNAs encoding IgE-binding gene products could be identified, and more than 20 of them have been cloned and characterized for allergenicity (Table 22.2). Comparative analysis of these sequences indicates that orthologs exist for almost all *A. fumigatus* allergens in the less allergenic species like *A. nidulans* or *Aspergillus oryzae*. A crude classification of these antigens divides them into highly conserved proteins and secreted glycosidases, chemically they may be distinguished as proteins, polysaccharides, or glycoproteins. On a functional level, *A. fumigatus* allergens split into enzymatic activities (proteases, endonucleases, or oxidases) and regulatory proteins (heat shock factors).

Among the list of allergens, Asp f 1, Asp f 5, and Asp f 10 are exceptional as they are considered to represent species-specific allergens, based on the assumption that they seem restricted to the genus *Aspergillus*. Asp f 1 was the first fungal allergen cloned and to be tested in recombinant form in humans. It is identical to the ribonucleolytic enzyme mitogillin and belongs to the standardized diagnostic antigen (SDA) group of *Aspergillus* antigens relevant for immunodiagnosis. It is one of the major polypeptides secreted by *A. fumigatus* under standard in vitro conditions. Asp f 5 is identical to the 42 kDa Mep metalloprotease, and Asp f 10 equals the aspartic protease aspergillopepsin F. Asp f 6, a MnSOD, is an example for a pan-allergen as it shows cross-reactivity with its human homologous protein at cellular and humoral level; human MnSOD binds IgE from patients sensitized to *A. fumigatus* MnSOD and elicits specific skin test reactions. Other such pan-allergens are represented...
FIGURE 22.4 Overview on key factors determining the host-pathogen system of invasive aspergillosis. Note: Several predisposing factors make a putative host susceptible to develop forms invasive aspergillosis: neutrophils represent a major defense line in the innate immune defense against *Aspergillus* conidia; accordingly, depletion of these immune cells is a high risk factor, and hematological disorders or malignancies may result in such prolonged neutropenia. Immunosuppression, as applied in the course of solid organ transplantations, or long-term corticosteroid therapy interfere with cellular mechanisms of the innate host defense on several levels, therefore supporting disease. For the fungus, a variety of attributes contribute to its pathogenic potential: adaptation to the ecological niche "host" relies on signal transduction cascades resulting in altered levels of gene expression; growth and expansion is supported by primary metabolism and specific secondary metabolites; overcoming defense lines of the innate immune system requires stress tolerance mechanisms and means to modulate effector cells; enzymatic activities able to degrade polymeric substrates assist in tissue penetration for invasion and dissemination.

### TABLE 22.2

*A. fumigatus* Allergens Cloned by IgE Binding

<table>
<thead>
<tr>
<th>Allergen</th>
<th>Size (kDa)</th>
<th>Function/Similarity</th>
<th>Accession No./Gene Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp f 1</td>
<td>16.9</td>
<td>Ribotoxin</td>
<td>S889330/Afu5g02330</td>
</tr>
<tr>
<td>Asp f 2</td>
<td>37.0</td>
<td>b-Glucanase</td>
<td>U56938/Afu4g09580</td>
</tr>
<tr>
<td>Asp f 3</td>
<td>18.5</td>
<td>Peroxiosomal protein</td>
<td>U58050/Afu6g02280</td>
</tr>
<tr>
<td>Asp f 4</td>
<td>30.0</td>
<td>Glucosidase</td>
<td>AJ001732/Afu2g03830</td>
</tr>
<tr>
<td>Asp f 5</td>
<td>42.1</td>
<td>Metalloprotease</td>
<td>Z30424/Afu8g07080</td>
</tr>
<tr>
<td>Asp f 6</td>
<td>23.0</td>
<td>MnSOD</td>
<td>U55661/Afu1g14550</td>
</tr>
<tr>
<td>Asp f 7</td>
<td>11.6</td>
<td>Glucosidase</td>
<td>AJ223315/Afu4g06670</td>
</tr>
<tr>
<td>Asp f 8</td>
<td>11.1</td>
<td>P2 ribosomal protein</td>
<td>AJ224333/Afu2g10100</td>
</tr>
<tr>
<td>Asp f 9</td>
<td>32.3</td>
<td>Cell wall glucanase</td>
<td>AJ223327/Afu1g16190</td>
</tr>
<tr>
<td>Asp f 10</td>
<td>34.0</td>
<td>Aspartic protease (aspergillopepsin F)</td>
<td>X85092/Afu5g13300</td>
</tr>
<tr>
<td>Asp f 11</td>
<td>18.8</td>
<td>Cyclophilin</td>
<td>AJ006689/Afu2g03720</td>
</tr>
<tr>
<td>Asp f 12</td>
<td>65.0</td>
<td>Heat shock protein</td>
<td>U924665/Afu5g04170</td>
</tr>
<tr>
<td>Asp f 13</td>
<td>34.0</td>
<td>Alkaline serine protease</td>
<td>Z11580/Afu2g12630</td>
</tr>
<tr>
<td>Asp f 15</td>
<td>19.5</td>
<td>Serine protease?</td>
<td>AJ002026/Afu2g12620</td>
</tr>
<tr>
<td>Asp f 16</td>
<td>43.0</td>
<td>b-Glucanase</td>
<td>G3643813/Afu1g16190</td>
</tr>
<tr>
<td>Asp f 17</td>
<td>19.4</td>
<td>Galactomannan protein MP1</td>
<td>AJ224865/Afu4g03240</td>
</tr>
<tr>
<td>Asp f 18</td>
<td>34.0</td>
<td>Vacuolar serine protease</td>
<td>Y13338/Afu5g09210</td>
</tr>
<tr>
<td>Asp f 22</td>
<td>46.0</td>
<td>Enolase</td>
<td>AF284645/Afu6g00770</td>
</tr>
<tr>
<td>Asp f 23</td>
<td>44.0</td>
<td>L3 ribosomal protein</td>
<td>AF464911/Afu2g11850</td>
</tr>
<tr>
<td>Asp f 27</td>
<td>18.0</td>
<td>Cyclophilin</td>
<td>AJ937743/Afu3g07430</td>
</tr>
<tr>
<td>Asp f 28</td>
<td>12.0</td>
<td>Thioredoxin</td>
<td>AJ937744/Afu6g10300</td>
</tr>
<tr>
<td>Asp f 29</td>
<td>12.0</td>
<td>Hioredoxin</td>
<td>AJ937745/Afu5g11320</td>
</tr>
</tbody>
</table>

by Asp f 8, the cyclophilins Asp f 11 and Asp f 27, and the thioredoxins Asp f 28 and Asp f 29.\textsuperscript{204} Synergistic activation of the immune system could be demonstrated for Asp f 13 and Asp f 2, which enhance lung inflammation when being present simultaneously.\textsuperscript{214} Some \textit{A. fumigatus} allergens—Asp f 2, 4, 7, 8, 9, 16, and 17—are similar to secreted glycosidases like to glucanases or cellulases and may therefore play a role in tissue adhesion. Moreover, additional allergen sequences could be extracted from the genome of \textit{A. fumigatus} by sequence comparisons with other established fungal allergens.\textsuperscript{51}

Although allergenicity is determined by other factors than the primary sequence, the knowledge on the identities of \textit{A. fumigatus} allergens assists in diagnostic matters and dangles immunotherapy or vaccination. Producing recombinant forms of allergens in suitable expression systems has become the method of choice to obtain pure and reliable antigens. Solving crystal structures may help in defining surface structures of allergens relevant for antibody binding. Studies using recombinant \textit{A. fumigatus} proteins have identified disease-specific allergens to discriminate between allergic asthma and ABPA: at least four allergens (Asp f 2, 4, 6, and 8) were shown to be highly specific for sera of ABPA patients, allowing discrimination between ABPA and \textit{A. fumigatus} sensitization at 100% specificity and 90% sensitivity.\textsuperscript{204,215,216}

### 22.12 Conclusions and Outlook

In the last 15 years medical mycology has faced the emergence of fungal pathogens, and in particular \textit{Aspergillus}, in distinct clinical settings to cause increasingly severe complications. On the other hand, this was accompanied by expanding knowledge and advanced techniques to study this genus on all kinds of aspects. This development was capped by the final determination of several genomic sequences from different \textit{Aspergillus} species, among them the major pathogen \textit{A. fumigatus}.\textsuperscript{51,205,217–219} As a consequence, profiling studies to monitor expression levels comprehensively have become feasible,\textsuperscript{220} and this approach is promising to shed light on the intimate host/pathogen system of \textit{A. fumigatus} infecting an immunocompromised or allergic individual. Global transcriptional profiling studies of immune effector cells phagocytosing conidia have been carried out to elucidate the host side: regulation of innate host defense molecules in response to \textit{A. fumigatus} was comprehensively described for the first time by transcriptional profiling data derived from human monocytes challenged with conidia.\textsuperscript{221} These data provide valuable insights into the initial immune response and elucidate the intimate interaction of the fungal microorganism and the effector cells. Thorough evaluation of the fungal transcriptome under \textit{in vitro} and, more importantly, \textit{in vivo} conditions is just a logical next step on this experimental line.\textsuperscript{222} In parallel and to correlate transcription profiles with expression levels, proteomic analyses need to validate the mass of data generated in this postgenome era of \textit{Aspergillus} research.\textsuperscript{223–225}

Most approaches to identify virulence-determining factors involved in \textit{Aspergillus} pathogenicity were based on gene identification and targeting to assess virulence of corresponding mutant strains in animal models of disease. As pathogenicity of \textit{Aspergillus} is most likely a polygenic trait, targeting single genes is not promising except for regulatory factors influencing the expression of a subset of genes related to a functional cellular category. Likewise, targeting several genes belonging to a gene family that encodes redundant activities makes generation of complete null mutant backgrounds possible. Testing such multiple deletion mutants enables the general evaluation of a cellular attribute and its contribution to pathogenicity. The molecular toolbox to manipulate \textit{Aspergillus} strains has expanded enormously over the last decade and allows such gene targeting strategies with reasonable effort.\textsuperscript{226,227} However, even by comprehensive evaluation of mutant strains and their phenotypes, specific virulence factors for \textit{Aspergillus} may be hard to define and comparative genome studies support that notion.\textsuperscript{48} Therefore, it remains to be shown whether \textit{Aspergillus} and in particular \textit{A. fumigatus} is a true pathogen at all or a saprophyte that just got lost in the wrong ecological niche—more molecular as well as epidemiological data need to be evaluated to find an answer to this provocative question.

Gaining more knowledge on the host side and the preconditions supporting the saprophyte \textit{Aspergillus} to become a pathogen will surely assist in evaluating the overall pathogenic potential of this fungal genus. Detailed knowledge on cellular attributes that support infection, germination, invasion, or dissemination is a strict prerequisite to define promising targets with the aim to allow rational design of antifungal substances, which defines the fundamental goal to overturn an unfavorable fungus/host relationship.
The Aspergilli

References

42. Segal, B.H., DeCarlo, E.S., Kwon-Chung, K.J. et al., Aspergillus nidulans infection in chronic granulomatous disease, Medicine, 77, 435, 1998.
The Aspergilli


78. Romano, J., Nimrod, G., Ben-Tal, N. et al., Disruption of the Aspergillus fumigatus ECM33 homologues results in rapid conidial germination, antifungal resistance and hypervirulence, Microbiology, 152, 1919, 2006.


95. Oliver, B.G., Panepinto, J.C., Fortwendel, J.R. et al., Cloning and expression of *pkaC* and *pkaR*, the genes encoding the cAMP-dependent protein kinase of *Aspergillus fumigatus*, *Mycopathologia*, 154, 85, 2002.


208. Moser, M., Crameri, R., Menz, G. et al., Cloning and expression of recombinant *Aspergillus fumigatus* allergen Ia (rAsp f Ia) with IgE binding and type I skin test activity, *J. Immunol.*, 149, 454, 1992.


The hallmark of translational research is the ability to transfer discoveries from the laboratory to the patient, and the crucial link in this marriage is the animal models that successfully mimic the human condition studied. Invasive aspergillosis (IA) is a complicated process requiring more than a single approach to accurately decipher and understand the pathobiology of both the fungus and the host. While there have been great advances in understanding genomics and proteomics in the last decade, the true application in medical mycology progresses through effective animal model testing.

Since testing hypotheses in humans afflicted with IA is not a practical option, mammalian models are a critical component of understanding because these models represent the best approximation possible of human disease. Animal models are designed to help answer widely varied hypotheses, and thus are tailored to the individual class of experiments. No single model will be ideal for testing all possible parameters. Major classes of experiments that can be addressed by animal models include the following: (1) evaluation of key fungal virulence factors identified in vitro, (2) exploration of host factors involved in disease susceptibility, and (3) evaluation of antifungal agents and/or immunomodulatory agents. Depending on the hypothesis to be tested, the investigator must decide on several critical factors, each of which will yield a slightly different representation of IA.

Since Aspergillus infection is relatively unique amongst microorganisms in that it afflicts patients in a variety of clinical presentations, extending the spectrum from deadly invasive disease in immunocompromised
patients to highly allergic disease in atopic patients, the investigator should first determine which aspect of aspergillosis is to be examined. There are numerous other nuances to consider, including which animal type or species, which method or dose of immunosuppression (if any), and what route of infection is appropriate. Further considerations also include fungal inoculation dose and identification of key experimental endpoints.

This chapter will address the types of mammalian models available, focusing on testing IA and not the many allergic forms of Aspergillus disease, outlining advantages and disadvantages to all options as well as highlighting key elements of the decision process of which model to choose. The development of standardized animal models is highly desirable, not only for ease of comparison of studies, but also for accurate and reproducible testing of new diagnostics and new therapeutic strategies.

### 23.1 Choice of Animal

#### 23.1.1 Murine Models

Murine models are the most common animal model used for the study of IA for several practical reasons. As with the study of many disease states, murine models afford the researcher not only with an inexpensive and easy system, but also the advantage of relative similarity between murine and human immunology and genetics. Furthermore, the availability of transgenic or targeted mutation ("knock-out" or "knock-in") mice as well as inbred murine strains allows the investigator to pursue candidate host susceptibility genes in a focused manner. This approach could be quite helpful for both testing disease in specific genetic backgrounds to mimic specific patient populations (i.e., chronic granulomatous disease) as well as testing a purposefully varied population such as would be represented with outbred mice. Murine genetics has undergone a recent revolution [1], and the availability of easily searchable murine genetic databases provides novel avenues for understanding disease.

Use of murine models to study IA is subject to the same scientific rigor that one would use in designing any animal model experiment. Some of the most common pitfalls in experimental design are the inadequate use of an appropriate number of mice to establish statistical significance of the findings, inadequate number of replicate experiments to validate data collected, and the appropriate use of uninfected controls. Simply understood, the smaller a difference in outcome expected, the greater number of animals required to demonstrate a statistically significant finding. There are a myriad of examples in the published literature where the investigators used an inadequate number of mice in each experimental arm and, therefore, failed to robustly support their potential conclusions. Inherent in this issue is the necessity to devise a statistical analysis plan prior to beginning the experiments. The assistance of an investigator adequately trained in statistical methodology will greatly aid in the effective completion of animal experimentation.

Most mammalian models of IA involve a particular immunosuppressive regimen, including appropriate control animals which undergo immune suppression and sham saline inoculation to document that the endpoints attributed to IA are indeed related to disease. Severely immunocompromised mice are typically given antibacterial prophylaxis during the time course of an experiment to prevent succumbing to bacterial infection during the periods of intense immunosuppression. Such prophylaxis historically has included supplemented water, including acidified water [2] as well as water supplemented with tetracycline 1 mg/ml [3], or supplemented with triple antibiotics (vancomycin, gentamicin, and clindamycin) [4]. Others have chosen to administer systemic antibacterials via daily subcutaneous injection (ceftazidime 50 mg/kg) [5] or an intravenous injection [6]. Choice of antibacterial prophylaxis regimen depends on the severity of immune compromise as well as concern of interactions during antifungal testing studies.

The choice of murine strain is ultimately dependent on the nature of the hypothesis to be tested. Transgenic mice are best utilized for host candidate gene/pathway studies. For example, such mice have been utilized effectively to characterize the role of innate immune pathway (Toll-like receptors 2 and 4 [7–11], Myd88 [12], natural killer (NK) cells [13], MIP-1α [14]) as well as adaptive immune pathways (IL-10) in the pathogenesis of IA. Additionally, transgenic mice can be utilized to mirror specific human disease states that confer increased susceptibility to IA due to an inherited genetic defect, typified by the
use of the p47\textsuperscript{phox}\textsuperscript{-/-} mice as a model for chronic granulomatous disease [15,16]. Appropriate control mice for these studies are wild-type litter mates. Comprehensive lists of commercially available transgenic mice can be found from major laboratory mouse suppliers such as The Jackson Laboratories (Bar Harbour, ME; www.jaxmice.org) and Charles River Laboratories (Wilmington, MA; www.criver.org). While an advantage of these murine strains is the very specific characterized genetic background to help control host variables, these mice are often very expensive.

Outbred mice are typically referred to as stocks [17]. The international standardized nomenclature for outbred stocks of both mice and rats [18] takes the following form: the company or laboratory designation followed by a colon, the stock designation, and then a hyphen followed by the mutation designation (if a mutation is known to be present). Like inbred mouse strains, outbred mouse stocks have an official definition: “a closed population (for at least four generations) of genetically variable animals that is bred to maintain maximum heterozygosity” [19]. As such, mice from a given outbred strain are genetically similar, but not identical, to one another. Outbred mice are often chosen for experiments because of low cost, or in the case of toxicology experiments, to represent a degree of genotypic variance across a population.

Important considerations when using outbred mice include the calculation of sample size, as the increased phenotypic standard deviation in outbred mice tends to obscure the effects of treatments, leading to low-powered experiments [17]. Thus, the use of several inbred strains, as opposed to outbred strains, can provide genetic variation (interstrain) as well as homogeneity (intrastrain), without the additional unknown variables inherent in outbred mice [20]. Additionally, outbred mouse stocks from different suppliers have been subject to different selective pressures over time, and thus may provide varying results in the same experiment [17]. Given the increasing recognition of the role of genetics in the outcome of both simple and complex traits, the general opinion of experts in murine model systems is that outbred mice should be used with caution and for highly specific reasons [17,20].

Inbred mice are commonly used in all types of biomedical research and offer the advantage of genetic homogeneity within each strain, such that each mouse is isogenic with another mouse of the same strain. As discussed earlier, such inbred strains offer greater control of phenotypic variability within an experiment. The obvious corollary to this fact is that what is found phenotypically for one strain of inbred mice may not be generalizable to all strains of inbred mice. Thus, to obtain a true sampling of population variation, experts recommend studying multiple inbred strains to offer a greater degree of control over population variability [17,20]. Inbred mice also allow the investigator to evaluate the genetic basis of phenotypic response, if this is the goal. A leap forward in the quest for resolution of the genomic basis for inherited traits was made in 2002 when mouse genomic sequences first became available [21,22], enabling interstrain sequence variation to be observed across the genome for the first time [23]. Definitions of commonly used categories of laboratory mice are shown in Table 23.1.

### 23.1.2 Nonmurine Models

While murine models often allow an easy and inexpensive route to pathogenesis modeling, there are circumstances where a nonmurine model system is preferred. For instance, due to the small size of a mouse there is a low limit on the amount of blood that can be obtained. Tail vein or retroorbital sampling on a mouse will generally only yield \( \leq 50 \) µl of blood, while even a cardiac puncture on a euthanized mouse will yield \( \leq 1 \) ml of total whole blood. For experiments where there is a need for serial monitoring, if using a murine model there must therefore be staggered animals used where different animals are sampled at different timepoints since one animal will not be able to be sampled at multiple points. The investigation of bronchoalveolar lavage (BAL) fluid is also complicated in a murine system. While the BAL can be performed on mice, the procedure is a fatal one, eliminating the possibility of serial BAL sampling on the same mouse over a time course.

The two most common nonmurine mammalian models used to study IA are rabbits and guinea pigs. Rabbits are generally 2–4 kg in size and have several unique advantages due to their larger size. Permanent vascular access can be obtained through the placement of a catheter for infusing both immunosuppression as well as antibiotic prophylaxis. Additionally, this catheter can be used for easy serial blood sampling, diagnostic surrogate markers, or antifungal pharmacokinetic studies. In the rabbit model it is possible
Guinea pigs, generally approximately 300–500 g in size, do not offer the same increased size as rabbits do over mice, and have been largely employed not for their larger size but for their preferred metabolism of certain antifungals. Many investigators studying the antifungal voriconazole determined that it is difficult to administer voriconazole to mice on a twice daily dosing schedule due to the rapid clearance of the agent. The guinea pig was found to possess different clearance times and could allow twice daily dosing of this particular important antifungal against IA [26]. Unfortunately, while larger than mice, the guinea pigs do not offer all the conveniences of a larger rabbit—including the permanent vascular access and the ability to image the lungs by CT.

### 23.2 Immune Suppression

Like humans, immunocompetent mice are fairly resistant to the development of IA. Establishment of IA in human hosts involves disruption of either alveolar macrophage and/or neutrophil-based lines of defense, and there is also an increasing recognition of the role of NK cells [13] and the adaptive immune system [8,9,27,28]. As such, investigators using mouse models have employed various methodologies for disrupting the macrophage- and neutrophil-based immune responses to *Aspergillus* infection. Importantly, infection parameters such as inoculum size needed or overall outcome may vary based on choice of immunosuppressive agent and immune suppressive schedule [29], and this should be taken into consideration when comparing studies using different forms of immune suppression.
Choice of agent and overall dosing regimen should be made with careful attention to the hypothesis of the study; that is, the use of a neutrophil-depleting regimen with additional corticosteroids may be most representative of hematopoietic stem cell transplant recipients during the period of graft-vs-host-disease prophylaxis [3,30], whereas a pure neutrophil-depleting monoclonal antibody (MAb) may be most useful in a study designed to evaluate specific roles of immune system components. Depending on dosing schema, immune suppressive agents can also be used to create a state of persistent or transient immune compromise. Here again, the methodology should reflect the desired effect. Key elements to utilization of any immune suppressive regimen are documentation of cell count kinetics in control animals to validate the establishment of neutropenia.

Not surprisingly, choice of immune suppressive agent(s) can significantly affect not only the establishment of disease, but also affect outcome. This is evidenced by a comparison of corticosteroid-based immune suppression with chemotherapeutic (vinblastine)-induced immune suppression. Despite identical inocula of fungi, there were important differences in disease pathophysiology due to the type of immune suppressive regimen chosen [31]. Pulmonary IA pathogenesis involved predominantly fungal development in mice treated by chemotherapy but an adverse host response in mice treated with a corticosteroid. Differences have also been noted in granulocytopenia versus corticosteroid and cyclosporine A induced immunosuppression, with scant mononuclear inflammatory infiltrate in the granulocytopenic model [32]. These differences should be taken into account in evaluations of the pathogenesis of IPA in animal models.

23.2.1 Neutrophil-Based Immunosuppression

Neutrophils are the main line of defense against Aspergillus hyphae, and as such, neutropenia remains an important clinical risk factor for the development of IA in at-risk patients. Establishment of neutropenia in murine models can be accomplished by the use of traditional chemotherapeutic agents or selective antibodies developed against neutrophils. Importantly, just as outcome can be influenced by the class of immune suppressive agents used, outcome can be influenced by neutrophil-depletion strategy [29]. In one particular study, various immune suppressive regimens were compared, and outcome was assessed following inhalation or intratracheal instillation of A. fumigatus conidia in both Balb/C and C57Bl/6 mice. Neutrophil-depleting regimens consisted of either a MAb RB6 or cyclophosphamide. Investigators evaluated the kinetics of neutrophil, lymphocyte, and splenocyte suppression using these agents alone or in combination. While the MAb is theoretically specific for neutrophils, depletion of CD8 cells and monocytes occurred. Coadministration of 25 µg of MAb RB6 and 150 mg of cyclophosphamide per kg on days 0, 3, 6, and 9 after infection resulted in significantly more severe neutropenia on day 3 than did administration of MAb RB6 or cyclophosphamide alone. As expected, administration of cyclophosphamide depleted a broad range of host cells, including neutrophils, and repeated administration allowed more sustained effects. While further details of the regimens are less important, the key concept illustrated by this investigation was that outcome following experimental IA can be extensively influenced by dosing regimen and choice of immune suppressive agents and that “selective” immune suppression may have additional unmeasured effects.

23.2.2 Macrophage-Based Immunosuppression

Alveolar macrophages ingest conidia and prohibit germination of conidia into invasive hyphae. Corticosteroids are toxic to alveolar macrophages, thus breaking down the first barrier to infection [33,34]. Additional effects of corticosteroids may be to suppress cytokine or chemokine production, specifically IL-1a, TNFa, and MIP-1a [35]. This blunting of alveolar macrophage cytokine production by corticosteroids may affect secondary pulmonary defenses against IA [36]. While not all murine models of IA utilize a corticosteroid-induced blunting of macrophage response, use of such agents is common to many murine models. Some models [37] utilize only a corticosteroid-based immunosuppression. In these systems, the doses of corticosteroids are often greater, as it is designed to overwhelm the immune system in order to establish disease. The individual corticosteroid utilized can also be important and issues such as half-life of the agent should be considered to obtain the desired effect. Often corticosteroids are administered
subcutaneously to allow more of a “depot” release of the agent, as opposed to the intraperitoneal or intravenous delivery of chemotherapeutic agents for establishment of neutropenia.

### 23.3 Fungal Inoculum Size

Fungal inoculum dose is again dependent on the model chosen for inoculation as well as the desired outcome. The effect of inoculum size is illustrated by one study comparing the effects of $2 \times 10^7$, $2 \times 10^8$ (10-fold increase), and $2 \times 10^9$ (100-fold increase) conidia/ml delivered in 30 µl droplets to the nares of anesthetized mice [38]. Mice were immunocompetent or injected with cortisone acetate or cyclophosphamide in various dosage regimens. In this particular model, outcome was affected both by fungal dose and immune suppressive regimen. Immunosuppressive treatment and $2 \times 10^7$ conidia/ml inoculum induced approximately 50% mortality. In contrast, mortality followed a fungus dose response in mice receiving immunosuppression with either cortisone acetate or cyclophosphamide. Different inoculum sizes are also required in different delivery methodologies, as intravenous delivery will require far less conidia/ml than an inhalational approach due to the more direct route of delivery.

### 23.4 Choice of Inoculation Route

Many *Aspergillus* inoculation routes have been utilized in murine models, some of which are more representative of human disease than others. Intranasal, intratracheal, and inhalational models each introduce conidia into the lung, mimicking the route of human disease acquisition, albeit some directly introduced and some allowing the animal to inhale the conidia in a more natural fashion.

#### 23.4.1 Intranasal

The intranasal delivery of conidia is probably the easiest method for establishing a pulmonary infection. Conidia are directly instilled down the nares of an anesthesized mouse, followed generally by a period where the mouse is held upright to allow the conidia to gravitationally fall toward the lungs. This methodology, while simple and reproducible as a very specific amount of conidia are instilled each time, has some potential drawbacks. While the infection is not isolated to the pulmonary system, as invariably there are signs of disseminated infection as conidia likely also move retrograde up the nasopharynx and establish cerebral infection, the instillation delivery method does not always establish a homogenous infection throughout all lung fields. Quantitative fungal PCR has shown that an intranasal instillation methodology does not yield statistically similar burden in all areas of the lung, which could lead to potential problems when only certain segments of the lungs are sampled for various measurements (i.e., histology, PCR, colony count, RNA extraction, etc.) [3].

#### 23.4.2 Intratracheal

The intratracheal route of infection was developed to circumvent the problem of conidia lodging in the nares during intranasal infection. This methodology allows for precise delivery of a conidial suspension directly into the pulmonary tree. Drawbacks include the need for a “survival surgery” wherein the animals are anesthetized, a catheter inserted into the trachea and conidial suspension instilled as described later. This methodology is not only time-consuming, but requires skilled laboratory personnel. Other tracheal instillation models involve a neck incision to facilitate direct visualization of and access to the trachea. This approach has the obvious downside of introducing a portal for infection in severely immunocompromised animals. Finally, it is unclear if the histopathology introduced by intratracheal infection is a true mimic of human disease [3] and how much this may impact study results.

A typical protocol for intratracheal inoculation of conidia involves general anesthesia achieved with a mixture of ketamine (40 mg/kg) and xylazine (8 mg/kg) administered via the intraperitoneal route.
For performing inoculation without a neck incision, a catheter (diameter, 0.86 mm) is inserted into the trachea via the oropharynx. Proper insertion of the catheter into the trachea (as opposed to the esophagus) is verified by checking the formation of mist due to expiration on a mirror placed in front of the external end. A 50-µl conidial suspension is introduced into the lungs using a pipette with a sterile gel loading tip placed into the internal end of the catheter. Following instillation of the conidial suspension into the lung, mice are immediately held upright in order to facilitate inhalation of conidia and maintained in an upright position until normal breathing resumes. Investigators using this protocol have demonstrated highly reproducible infection of the lungs with 10 times more inoculum reaching the lungs via this route than via the intranasal route [31].

### 23.4.3 Inhalational

Inhalational models of aspergillosis are attractive as they likely best mimic disease acquisition in the human host. Once developed, these systems can infect large numbers of mice at one time (20 mice in a standard Hinner’s chamber; 100 mice in a Madison chamber). As such, these models may be best suited to standardization across multiple laboratories [5]. Older inhalational models utilized a modified Erlenmeyer flask with side-arms where a lawn of conidia was grown on the bottom. After sufficient conidia growth, unanesthetized mice would be placed in the side-arms with noses pointing into the flask. At the top of the flask was a stopper attached to a plunger where investigators would compress the plunger, resulting in aerosolization of conidia and thus murine exposure [39]. While this model would establish aspergillosis, it was clearly hampered by the inevitable technical inconsistencies between experiments and operators for aerosol generation.

Inhalational models became more advanced with the development of nebulized delivery systems for delivery of conidial suspensions. This delivery system overcomes the inherent variation in conidial aerosolization from the side-arm flask method and can deliver a consistent amount of conidia on repeated experiments [5]. Inhalational models have been demonstrated to induce homogenous whole-lung pulmonary infection [3,40] as determined by quantitative Aspergillus DNA PCR. While there are many advantages to an inhalational system, including the homogeneity of delivery in a manner most akin to human disease acquisition, it does require the use of an aerosol chamber. While technically easier than intratracheal instillation, this method is more demanding than the simpler intranasal instillation.

Currently an NIH consortium grant has been awarded to standardize inhalational models of aspergillosis in an effort to lend some consistency to the field [41]. This consortium is working on several methods, including an apparatus developed by this program known as a Madison Chamber that is designed to hold up to 70 mice per inhalation. Standard operating procedure for this apparatus is found on the following website http://www.sacmm.org/sop.html.

### 23.4.4 Intravenous

While intravenous inoculation with conidia results in rapid and consistent disease acquisition [42–44], it does not mimic human disease either in route of acquisition or areas of heaviest dissemination (kidney) [42–44]. However, while the intravenous route of acquisition does not completely mimic human disease, it does create a disseminated disease and is a useful model for certain purposes. This model allows for highly standardized doses of fungus to be administered to immunocompetent mice, which can be a valuable system for initial evaluation of antifungal agents. A typical intravenous model involves delivery of a conidial suspension through the lateral tail vein, with observation of mice for survival and determination of fungal burden in the kidney as endpoints. Fungal inoculum size can be varied to induce different degrees of mortality, with inoculum sizes of $5.5 \times 10^7$ to $7.0 \times 10^7$ conidia/ml producing nearly 100% mortality in immunocompetent CD-1 mice [45].

Overall advantages to this model system include ready standardization—the fungal inoculum is delivered directly into the tail vein. Tail vein injection requires technical skill, and thus practice is necessary to perfect the technique. An additional advantage is the lack of immunosuppression due to the direct intravenous delivery which apparently overwhelms the animal sufficiently that there is no necessity to
first establish an immunocompromised state. This is not an insignificant fact, as there is a careful and
delicate balance to attain and maintain with the immunosuppression of animals. Lastly, this model effec-
tively creates a disseminated infection, while the inhalational model generally creates only a pulmonary
infection. This serves both as an advantage and a disadvantage. While this allows recovery of the
Aspergillus from the kidneys, akin to Candida infection murine models, patients with IA rarely have
kidney involvement and generally instead develop pulmonary disease.

23.4.5 Cerebral

The central nervous system (CNS) is the most common site of disseminated aspergillosis, and until
recently this form of animal model had been neglected. Establishment of pulmonary infection in a murine
model rarely results in dissemination, and systemic infection through an intravenous infection does not
consistently establish CNS disease. To facilitate study of CNS aspergillosis, a difficult clinical problem
for treatment due to the frustration in achieving adequate levels of antifungals in the brain, a murine
model of CNS aspergillosis was developed. This model is suitable both for study of host response to CNS
aspergillosis as well as the evaluation of therapeutic regimens. Goals of the model were to induce 100% mortality,
so that no animal would have insufficient progression of disease and resolution. However, in a
model system, rapid mortality is not desirable because of the need for sufficient time for evaluation of
potentially useful therapeutic interventions. Given these factors, the model was designed such that 100% mortality
would occur over the course of several days [46] using both outbred (CD-1) and inbred (C57Bl6/J
and DBA/2N) mice.

Investigators discovered several important experimental points regarding murine CNS aspergillosis
during the development of this model. First, immunocompetent mice were either resistant to a low-dose
infection (5 × 106 conidia inoculated intracerebrally), or experienced 100% mortality within two days
of a high-dose inoculation (5 × 107 conidia inoculated intracerebrally). As such, a cyclophosphamide-
based regimen was evaluated. A regimen of 200 mg/kg IP on day −2 and every five days after inocula-
tion of conidia produced both reliable neutropenia as measured by cell count and differential, as well
as acceptable and reproducible 100% mortality by eight days after infection when using an inoculum
size of 5 × 106 conidia.

23.4.6 Noninvasive to Invasive Disease

An interesting model has been developed for the study of the development of IA following established
noninvasive disease. In this model, mice are inoculated intratracheally with a 20 μl suspension of
Aspergillus laden agarose beads (10^8 conidia/ml). Following inoculation, all mice rested for
two weeks. Then, a subset of mice were given cortisone acetate 125 mg/kg SQ daily for seven days and
four week mortality was observed. Immune competent mice cleared the Aspergillus burden, whereas
the immune compromised mice developed pulmonary aspergillosis with neutrophilic and lymphocytic
invasion, with hyphal growth and tissue necrosis. This model is valuable for studying a unique form of
pulmonary aspergillosis [47].

23.5 Measurement of Outcome

Equally important to the accurate establishment of the correct model through determining the optimal use
of animal type, immunosuppression, and inoculation route, is the assessment of valid and reproducible
endpoints. Outcomes measured related to IA pathogenesis include survival, tissue fungal burden, and
cytokine/chemokine response levels in either lung lavage fluid or serum. Choices of endpoints are based
both on hypothesis and goals of the particular study. Overall survival is an important primary endpoint in
all experiments and needs to be assessed with consistency. Ideally, the animals are euthanized before they
succumb to disease, but in order to do this there needs to be uniform standards used for determining
which animals meet criteria for euthanasia. Generally animals with imminent deadly infection will have
Mammalian Models of Aspergillosis

little movement, ruffled fur, and breathe very rapidly. While some investigators also measure weight daily as an indication of disease, this is complicated in immunosuppressed mice as the agents themselves will induce a weight loss up to 15% of body mass. Once survival is assessed the data need to be presented correctly, and for this a Kaplan-Meier curve is the standard used.

After the primary endpoint of survival, there are many other secondary endpoints that can be evaluated. Recently, there is great interest in methodology for measurement of tissue fungal burden, as this represents both an internal control for equivalence of fungal dose as well as a representation of host control of disease or efficacy of a particular antifungal regimen. The most commonly used methodologies for determination of tissue fungal burden include colony count, chitin assay, and quantitative Aspergillus DNA PCR. Other potential measurements include surrogate markers such as the galactomannan assay or the β-glucan assay to quantify serum or BAL fluid fungal burden.

In Candida infection models, it is easy to obtain accurate colony counts because the organism can be retrieved as a simple cellular yeast. However, while homogenization, serial dilution and quantifying infected tissue for colony counts is an accurate and easily performed means of determining organism burden for bacterial and yeast infections [48,49], this methodology can be problematic when dealing with filamentous fungi such as Aspergillus. The complication arises because the Aspergillus exists as spores which are not easily confined and when germinate they grow as successive generations of hyphal expansions, very quickly blurring the distinction of how many colonies originally existed. Comparisons of fungal burden as determined by colony counts versus quantitative PCR have noted colony counts to consistently underestimate the fungal burden [50,51], possibly due to shearing of filamentous organisms during homogenization or due to clumping of fungal elements. Conversely, spuriously high estimates of fungal burden have resulted when using colony count to quantify fungal burden in echinocandin antifungal based studies due to hyphal tip breakage introduced by the antifungal agent [52,53].

The chitin assay is another tool and this assay measures the fungal cell wall component chitin. This method is performed on tissue homogenates, with fungal burden measured as glucosamine equivalents [54]. While likely more quantitative than colony counts, the chitin assay is more technically demanding than quantitative PCR [55].

Quantitative Aspergillus DNA PCR has been validated as an outstanding method for measuring fungal burden in animal models. This methodology measures the Aspergillus 18s ribosomal RNA gene, which is relatively easy to perform and is well-standardized. This is the preferred method for the measurement of fungal burden by the IAAM consortium [5]. One disadvantage of this methodology is its expense, as the most cited and standardized methodology is a Taqman-based assay [50,51].

Histologic analysis of infected tissue is useful for determining the extent of cellular injury. Typically, histologic analysis evaluates inflammation, fungal burden, necrosis, and hemorrhage. Methodologically, infected lung tissue is harvested and stained with hematoxylin and eosin to characterize inflammation and Gomori’s methenamine silver stain to document fungal invasion. Scoring systems have been developed to quantify infection according to a five-point pulmonary infarct score that incorporated necrosis, hemorrhage, edema, and hyphal presence [3].

23.6 Conclusion

Mammalian animal models play a major role in elucidating IA pathogenesis as well as investigating exciting new diagnostic and therapeutic strategies. Mammalian models thrive because they can mimic the human condition and disease can be evaluated via numerous endpoints. Crucial to selecting the correct mammalian model is a firm understanding of the hypotheses to test and the eventual application of the results. Important considerations include the type of animal and species or genetic background, the degree and nature of immunosuppression, the infection route and amount, and the type of endpoints to evaluate. It is crucial to understand that no single mammalian model will serve all lines of inquiry, but with a proper understanding of the benefits and differences in each model system, valuable data can be obtained to advance the work against aspergillosis.
References


Emerging Role of Mini-Host Models in the Study of Aspergillosis

Georgios Chamilos and Dimitrios P. Kontoyiannis

CONTENTS

24.1 Introduction ......................................................................................................................... 413
24.2 What Is a Suitable In Vivo Model of IA Pathogenesis? ...................................................... 414
24.3 Evolutionarily Conserved Innate Immunity Pathways Against Aspergillus in Insects .......... 415
24.4 Humoral Responses Against Aspergillus in Insects ........................................................... 415
24.5 Toll Pathway .......................................................................................................................... 415
24.6 Cellular Immunity Against Aspergillus in Insects ............................................................... 416
24.7 Invertebrate Models of IA ................................................................................................... 416
24.8 Methods of Studying IA Using a Fly Model .......................................................................... 417
24.9 Infection Models of Aspergillosis in Drosophila ................................................................. 417
24.9.1 Methods of Screening of Candidate Compounds with Anti-Aspergillus Activity in Drosophila ................................................................................................................................................................................. 420
24.9.2 Pilot Studies of Antifungal Drug Efficacy in Drosophila Model of IA ......................... 421
24.9.3 Virulence Testing of Aspergillus in Drosophila ............................................................... 421
24.9.4 Galleria mellonella ........................................................................................................... 422
24.10 Virulence Studies in G. Mellonella and Other Invertebrate Models of Aspergillosis ............ 423
24.11 Future Perspectives .............................................................................................................. 423
24.12 Concluding Remarks .......................................................................................................... 424
24.13 Conflicts of Interest ............................................................................................................ 424

24.1 Introduction

In recent years, opportunistic fungi have emerged as leading causes of morbidity and mortality in immunocompromised individuals [1,2]. The epidemiology of invasive fungal infections has evolved over the past two decades, partially because of the widespread use of antifungal agents with good activity against Candida, such as fluconazole. Since the early 1990s, ubiquitous airborne saprophytic molds have become the leading fungal pathogens in severely immunocompromised patients [1]. Aspergillus is the most common of these molds, and invasive aspergillosis (IA) has emerged as the major problem of modern mycology [2].

Despite the recent introduction of new antifungal agents with promising anti-Aspergillus activity, the mortality rate associated with IA remains high, approaching 80–90% in high-risk leukemia patients and allogeneic bone marrow transplant recipients [1,2]. Aspergillus fumigatus is the species that causes most cases of IA; however, non-fumigatus species that are less susceptible to antifungal agents, such as...
Aspergillus terreus, have been increasingly reported in severely immunocompromised individuals [2]. Of even more concern, however, is that infections caused by multidrug-resistant non-fumigatus Aspergillus species have begun to emerge [3].

There is a dire need to develop novel therapeutic approaches to the treatment of IA; thus, expanding our knowledge of the immunopathogenesis of Aspergillus infections is vital [4]. Protection of the immunocompetent host against Aspergillus infection is mediated by a highly coordinated response that involves both innate and adaptive immunity [4–6]. However, despite significant improvements, the immunopathologic mechanisms of aspergillosis are not completely understood.

Specifically, uncertainty remains regarding the attributes that mediate Aspergillus pathogenicity [5]. Several putative virulence factors in A. fumigatus have been identified, including various proteolytic enzymes (elastases, collagenases, and trypsin), phospholipases, ribotoxin, hemolysin, and toxins [5,6]. Gliotoxin in particular has received substantial attention, as it has been shown to inhibit the phagocytosis of macrophages, promote the apoptosis of macrophages, and block B- and T-cell activation [7]. In addition, the melanin pigment and rigid protein coat layer (composed of rodlet fascicles) that are present on the Aspergillus conidial surface may confer resistance to phagocytosis [6]. However, molecular studies have yet to identify a single target that meets Koch’s postulates for virulence in IA. On the other hand, recent evidence from a comparative genomic analysis across filamentous fungi suggests that A. fumigatus virulence results from either the failure of host immune responses or the genetic susceptibility of the host rather than from specific, unique fungal determinants [5].

Overall, the versatility and plasticity of cellular functions of this important opportunistic mold emphasize the importance of understanding the nature of host–pathogen interactions in robust experimental systems. Over the past decade, invertebrate mini-host models with well-characterized genetics and simple immunity have been used effectively to explore several aspects of both fungal pathogenicity and host immune responses [8–11]. Several factors sparked the development of these models. First, traditional animal models remain logistic barriers to large-scale studies. In addition, the innate immune mechanisms are evolutionarily conserved between invertebrates and mammals, and several common virulence factors are involved in fungal pathogenesis in phylogenetically disparate hosts, such as fruit flies, nematodes, and mammals [12,13]. Furthermore, invertebrate organisms have been increasingly used as alternative in vivo assays to perform antifungal drug efficacy studies because of their low cost and simplicity [14–16]. Last, recent studies have revealed the feasibility of in vivo high-throughput screens for antimicrobial compounds in invertebrates [17–19].

24.2 What Is a Suitable In Vivo Model of IA Pathogenesis?

Ideally, an animal model of fungal pathogenesis should simulate the pathophysiologic characteristics of the corresponding human infection, including colonization and invasion from a specific route of entry, and stimulate key host immune defenses. In addition, the tempo of experimental infection should be sufficiently protracted to effectively account for the virulent attributes of the fungal pathogen.

Mammalian animal models have been invaluable tools for the elucidation of the molecular and cellular pathogenesis of fungal infections [20]. Apart from the ethical dilemmas associated with experimentation in humans, the principal advantage of modeling infections in animals is that both the host and its environment can be precisely controlled, allowing for a comprehensive analysis of host–pathogen interactions. Thus, small mammals such as rats, mice, and rabbits are the gold standard for pathogenesis studies because of their relative anatomic, physiologic, and immunologic similarities to humans.

However, several problems remain with the use of these experimental systems. Although mammalian models are amenable to reverse genetics through the generation of knockout mutants, the identification of genes by large-scale forward genetic screening is challenging. In addition, dissecting the molecular attributes of host immune response against fungi is difficult in these pathosystems because of the complexity of mammalian immunity. Furthermore, the use of large numbers of mammals is not feasible for logistic, economic, and ethical reasons. This is a particularly timely issue, as the genome sequences of medically important fungi such as Aspergillus, Candida, and Cryptococcus have recently been completed [21,22]. This surge in genetic information and the rapid development of molecular toolsets to study gene
Emerging Role of Mini-Host Models in the Study of Aspergillosis

function in various fungi has created an increasing need for simple, yet innovative, ways to screen for virulence mechanisms and assess the contribution of individual genes to fungal pathogenesis.

Pioneering studies over the past decade have demonstrated that a variety of pathogenic fungi can invade and cause fatal infection in a variety of simple invertebrate hosts, such as the fruit fly \textit{Drosophila melanogaster}, the roundworm \textit{Caenorhabditis elegans}, and the greater wax moth \textit{Galleria mellonella} \cite{8–10,14–16,22}. Also, research has shown that important aspects of innate immunity have been evolutionarily conserved across phylogeny. Hence, comparative genomic studies illustrate that a high percentage of human protein homologs, particularly those involved in pathogen recognition, signal transduction, and innate immune responses, exist in various invertebrates, such as \textit{D. melanogaster} (60\%) and \textit{C. elegans} (55\%) (http://www.ncbi.nlm.nih.gov/sutils/taxik2.cgi). Thus, because of their simple immunity and because both the host and pathogen are amenable to genetic analysis and high-throughput screening in each of these pathosystems, the use of invertebrate models has accelerated studies of microbial virulence and host immunity.

24.3  Evolutionarily Conserved Innate Immunity Pathways Against \textit{Aspergillus} in Insects

Although they lack adaptive immunity, invertebrates are capable of mounting efficient responses to an array of pathogens in their natural habitats. Their self-defense system is composed of epithelial responses, a primitive phagocytic response, and the release of natural defensins through stimulation of innate immunity. First, because epithelial surfaces are where potentially invading pathogens come into contact with these hosts, physical barriers such as chitin-rich rigid membranes and a low pH constitute the first line of defense and prevent contact and colonization by invaders \cite{11,12,22,23}. When these physical barriers are breached, the invasion of pathogenic microorganisms within the insect body induces a strong and highly coordinated immune response that has both cellular and humoral components and is strikingly similar to the mammalian immune response to pathogens \cite{11,12,22,23}.

24.4  Humoral Responses Against \textit{Aspergillus} in Insects

In insects, the hallmark of an innate immune response against pathogenic microorganisms is the induction of a battery of antimicrobial peptides, which are secreted by the fat body (equivalent to the mammalian liver) into the hemolymph \cite{23}. Despite the broad spectrum of antimicrobial peptides, specificity exists upon their induction against various microbial pathogens. For example, in \textit{Drosophila}, fungi (including \textit{Aspergillus}) and gram-positive bacteria mainly induce drosomycin and metchnikowin via the \textit{Toll} pathway, whereas gram-negative microbes induce diptercin, attacin, and cecropin through the \textit{Imd} pathway, which are analogous to mammalian \textit{Toll}/interleukin-1 receptor and tumor necrosis factor (TNF) signaling in humans, respectively \cite{11,12}.

The specificity of innate immunity is conferred through pattern recognition receptors (PRRs). PRRs are soluble or transmembrane proteins common to various mammals and insects that recognize essential molecules present exclusively in microbes, such as lipopolysaccharide, lipoteichoic acids, and peptidoglycan; the so-called pathogen-associated molecular patterns. The results of recent studies in \textit{Drosophila} indicate that peptidoglycan recognition proteins and gram-negative bacteria-binding proteins comprise the main PRRs. Recently, the results of a mutation analysis in \textit{Drosophila} implicated that gram-negative bacteria-binding protein 3 is a candidate PRR for fungal pathogens, including \textit{Aspergillus} \cite{24}.

24.5  \textit{Toll} Pathway

In both insects and mammals, the interaction of invading fungi, including \textit{Aspergillus}, with specific PRRs leads to activation of intracellular phosphorylation cascades, with subsequent upregulation of
The Aspergilli

antimicrobial peptide-encoding genes through the translocation of nuclear factor κB-like transcriptional factors to the nucleus [11–13]. Specifically, in Drosophila, gram-negative bacteria-binding protein 3 activates the serpin Persephone, which is implicated in the cleavage of another serpin, Spatzle. Spatzle subsequently activates the transmembrane receptor Toll, its downstream adaptor proteins MyD88 and Tube, and the threonine-serine kinase Pelle, which are homologs of human MyD88, Mal (functional equivalent), and IRAK, respectively [11–13]. This proteolytic cascade ultimately leads to degradation of the 1xK inhibitor Cactus and nuclear translocation of the nuclear factor κB-like transcriptional factors Dorsal and Dif, which induce the expression of antimicrobial peptide-related genes [11–13]. The predominant role of the Toll pathway in Drosophila immunity against Aspergillus was first demonstrated in a landmark study by Lemaitre et al. [8], who found that Toll mutant flies, in contrast to wild-type flies, were extremely susceptible to Aspergillus infection.

24.6 Cellular Immunity Against Aspergillus in Insects

Cellular responses against Aspergillus and other pathogenic fungi are less well characterized in Drosophila. Nonetheless, the cellular immune responses of insects against Aspergillus have been better characterized in G. mellonella, in which at least six types of hemocytes (prohemocytes, coagulocytes, spherulocytes, oenocytoids, plasmacytes, and granulocytes) participate in specialized processes, including phagocytosis, nodulation, and melanization [23,25]. Of these cells, plasmatocytes and granulocytes are the predominant phagocytic cell types. Importantly, the mechanisms of killing efficacy of phagocytic cells in G. mellonella seem to be analogous to those in mammals. Thus, insect phagocytic cells are also capable of generating an oxidative burst of oxygen radical intermediates in a manner analogous to that in mammals [23].

Of interest, infection of G. mellonella larvae with nonpathogenic fungi (e.g., S. cerevisiae) results in high hemocyte concentrations in the hemolymph, whereas Aspergillus and other pathogenic fungi (e.g., Caenorhabditis albicans) induce a significant reduction in the number of hemocytes. These findings suggest that hemocyte concentrations could serve as a marker of the pathogenicity of fungal invaders in the Galleria model [25]. In addition, the role of cellular immune responses against fungi has been increasingly appreciated in Drosophila. For example, we recently showed that S2 embryonic phagocytic cells phagocytosed Aspergillus conidia and caused significant damage within 1 h after exposure to Aspergillus hyphae [26].

24.7 Invertebrate Models of IA

Drosophila melanogaster. The fruit fly D. melanogaster (=3 mm long) is larger than roundworms but significantly smaller than caterpillars. The pathogen of interest is typically injected into the dorsal thorax; of fruit flies; however, other more physiologic methods of infection (e.g., rolling or feeding) have also been used [14]. Female flies are typically used in infection experiments because of their larger size and relative resistance to injection injury when compared with male flies. Because wild-type Drosophila is resistant to Aspergillus and most other pathogenic fungi, fly mutants deficient in various components of the Toll cascade are used to model infections. In most cases, crossing different loss-of-function alleles is required to generate homozygous Toll mutant flies. The genetic tractability and well-characterized immune system of Drosophila is a major advantage. Hence, Drosophila is amenable to both forward and reverse genetics, and large collections of Drosophila mutants and transgenic cell lines are commercially available (http://flybase.net/). Also, the Drosophila genome sequence was one of the first to be completed and is probably one of the most fully annotated eukaryotic genomes found in a database (http://flybase.net/annot/). As a result, double-stranded RNA has been synthesized for each of the Drosophila genes (http://www.flyrna.org). The application of RNA interference technology in the Drosophila S2 phagocytic cell line confers an additional advantage, enabling a functional genome-wide analysis of host–pathogen interaction at the cellular level [12,27].
However, there are limitations to the study of fungal pathogenesis in this elegant mini-host model. Thermotolerance is a universal virulence trait of pathogenic fungi [5]; thus, the fact that the infection and maintenance of Drosophila as well as most invertebrate hosts takes place at a low temperature is problematic for the study of temperature-sensitive Aspergillus mutants. For example, a recent study found that a gene that regulates the expression of the nucleolar protein CgrA plays a pivotal role in Aspergillus thermotolerance and that a ΔCgrA mutant displayed attenuated virulence in mice at 37°C [28]. However, the ΔCgrA mutant was fully virulent in Toll-mutant flies infected and maintained at 25°C.

24.8 Methods of Studying IA Using a Fly Model

Maintenance of fly stocks. Flies are maintained in vials containing standard Caltech (cornmeal) medium (http://flybase.net), which can be stored for a month at 4°C until use. For optimal fly hatching, sets of 20–30 flies are placed in each vial. Flies are typically transferred into new vials every three to four days to continuously renew the stock. A critical component of fly hatching is optimal humidity (60%), which requires the use of specific fly humidifiers. Optimal temperature (25°C) and a day/night cycle of 12 h are also needed. Alternatively, flies can be maintained at ambient temperature, which requires supplementation of vials with distilled water every three to five days to retain humidity. Manipulated and nonmanipulated flies are kept in separate containers except when genetic crosses are to be made. All containers and material wastes are adequately sealed before disposal and deposited in clearly labeled bins prior to being decontaminated by approved methods. All containers that contain or have contained infected flies are decontaminated with 100% ethanol solution.

Identification of males versus females and virgin versus nonvirgin adults. In our model of IA, we used Drosophila female adult (two to four days old) transheterozygous mutant flies (Tl<sup>632</sup>/Tl<sup>RXA</sup>). These flies are generated by crossing virgin flies carrying a thermosensitive allele of Toll (Tl<sup>632</sup>) with flies carrying a null allele of Toll (Tl<sup>RXA</sup>) [14]. Tl<sup>632</sup>/TM6B and Tl<sup>RXA</sup>/TM6B Drosophila alleles can be phenotypically distinguished by their color. Specifically, the former are light brown, whereas the latter are dark gray. Oregon wild-type flies are also dark gray in color (resembling Tl<sup>RXA</sup>/TM6B flies) but have red eyes. The identification of male and female flies is on the basis of their genitalia (Fig. 24.1a, c). Virgin female flies can be distinguished by a dark mark on the ventral side of the abdomen (Fig. 24.1a) that corresponds to an embryonic residue that is excreted from their gastrointestinal tract upon matura- tion. Typically, at their optimal temperature of 25°C, female flies are considered virgins during the first 6–8 h of their lives.

Crossing of Toll Drosophila alleles and identification of transheterozygous Tl<sup>632</sup>/Tl<sup>RXA</sup> mutants. Importantly, each Drosophila allele carries a balancer (marker) chromosome, which is used to provide flies with unique phenotypic characteristics (e.g., eye color or wing or bristle pattern) that can be used to distinguish different fly crossings phenotypically. In our Toll mutant fly model, the balancer is called TM6B; flies with this balancer have a “hairy-type” bristle in their upper lateral thorax/torso (flies lacking TM6B balancer have a “double hair-type” bristle at the same location) (Fig. 24.1b, d). To obtain the Tl<sup>632</sup>/Tl<sup>RXA</sup> Drosophila mutants used in the Aspergillus infection models, we need to either cross virgin Tl<sup>632</sup>/TM6B females with Tl<sup>RXA</sup>/TM6B males or cross virgin Tl<sup>RXA</sup>/TM6B females with Tl<sup>632</sup>/TM6B males at a ratio of 2:3. Tl<sup>632</sup>/Tl<sup>RXA</sup> Drosophila mutant flies have a light brown color and do not possess the balancer TM6B; as a result, they have the double hair-type of bristle (Fig. 24.1b).

24.9 Infection Models of Aspergillosis in Drosophila

Injection assay. Few fungi that naturally infect invertebrates have been identified. Therefore, most fungal pathogens are introduced directly into the body cavity by pricking the insect’s cuticle with a sharp needle or microinjecting a precise dose of fungal cells into the body cavity. However, this procedure bypasses the natural routes of infection, including mucosal colonization and the initial interaction of the pathogen with
The Aspergilli

specific epithelial receptors. It also induces the expression of a wide spectrum of antimicrobial peptides, some of which are nonspecific, as they might be induced by the physical injury.

For the injection assay, 20–30 female Tlr632/Tlr-RXA flies are placed on a fly pad anesthetized with CO2, and subsequently injected into the dorsal side of the thorax by a sterile metal needle (0.1 μm) previously dipped in a concentrated solution of A. fumigatus conidia (Fig. 24.2). As a control, another group of 20–30 female Tlr632/Tlr-RXA flies are injected with a sterilized needle (septic injury control). Following infection, flies are transferred in fresh vials and are observed for the next 3 h. Flies that die within 3 h after infection are considered to have succumbed as a result of the procedure and these flies are excluded from survival analysis. Flies are housed at 29°C (the temperature that results in optimal susceptibility to fungal infection) and transferred to new vials daily when drugs are tested or every two days when only crude mortality without drug protection is assessed. Survival is assessed daily until day 8 of the infection. The injection assay is technically a more standardized and reproducible (almost semiquantitative) method of infection and allows us to estimate fungal inoculum by serial dilutions of Aspergillus conidia introduced by a needle [14]. Nonetheless, parenteral inoculation results in a more overwhelming infection, which may not be suitable for pathogenesis studies. For example, we recently reported [14] that the alb1 A. fumigatus mutant, which is hypovirulent in mice, exhibited attenuated virulence in Toll-deficient flies when introduced using the rolling and ingestion methods; however, this effect was not observed when it was introduced by injection.

Rolling and ingestion assay. A more physiologic method of infection is typically achieved by feeding insects in a lawn of yeast or molds or rolling insects over a fresh carpet of fungal spores (Fig. 24.2). In the latter case, the fungal pathogen penetrates the insect’s exoskeleton. This method results in a more
protracted model of infection than that of injection and thus may allow for more reliable assessment of certain attributes of fungal pathogenicity. However, standardization of the infecting inocula is difficult with natural infection methods such as ingestion.

For the rolling assay, flies are placed on the fly pad and anesthetized for 3–4 min. Groups of 30–50 flies are transferred to YAG plates containing a fresh carpet of *Aspergillus* conidia that has been grown for three days. The flies are then rolled on the plate surface for 2 min to allow *Aspergillus* conidia to cover their surface (Fig. 24.2). Importantly, flies should be transferred temporarily (e.g., for 1–2 h) to new vials to avoid exposure to a substantial amount of conidia. After this 1–2 h period, flies can then be placed in new vials at 29°C. Again, flies that die within 3 h of rolling (typically <1%) are considered to have died as a result of the procedure and were not included in the survival analysis.

For the ingestion assay, special fly vials are prepared that contain YAG medium. *Aspergillus* conidia are put on the surface of the YAG medium (100–200 μl of a 10⁸ conidia/ml solution) for three days until a conidial carpet has formed. Groups of 30–50 flies are left for 6–8 h to feed on the conidia inside the vials (Fig. 24.2). It should be noted that feeding times longer than 6–8 h can result in death from dehydration and starvation, perhaps because *Aspergillus* conidia is not an optimal nutritional medium for flies. Flies are transferred again (as described previously for the rolling assay) in temporary vials.
for 1–2 h so that a substantial amount of conidia gets out of their cuticles. They are then transferred to new vials at 29°C.

24.9.1 Methods of Screening of Candidate Compounds with Anti-Aspergillus Activity in *Drosophila*

In addition to study of fungal virulence, the capacity of a new animal model in assessing the efficacy of therapeutic agents adds value. In *Drosophila*, the easiest way to expose a large number of flies to an antifungal compound is by mixing the drug with fly food, which is typically achieved by dissolving it into a yeast-sucrose medium. However, this technique is not suitable for screening antifungal compounds that have low solubility in aqueous solvents. In addition, precisely estimating the amount of the drug ingested with this method is difficult. To indirectly determine whether adult flies eat the compound, a red water-soluble dye plus the antifungal agent of choice (e.g., voriconazole) is added to the regular fly food. Flies are subsequently fed in the vials for 24 h. The presence of red dye in the fly’s ventral abdominal surface is an indirect method of determining whether they have eaten the food that contains the drug (Fig. 24.3).

---

**FIGURE 24.3** Outline of the drug protection procedure. Vial containing fly food and a red food dye plus 1 mg/mL of voriconazole (a). Colorimetric assay for drug absorption (b). A fly was left to feed in the vial; red dye was visible 24 h later, indicating ingestion of voriconazole.
Methods of drug delivery in fruit flies. Horizontal and vertical abrasions (2–3 mm in depth) are made on the surface of the fly food using a sterile spatula. A calculated concentration of the antifungal agent of interest (e.g., voriconazole) is added to the surface of the fly food. Next, yeast particles are soaked in the surface of the vial using a 1-ml pipette tip, and the vials are left to dry for 24 h. An outline of the drug protection procedure is shown in Figure 24.3. In brief, flies are put in empty vials for 6–8 h to starve and become dehydrated. After this starvation period, flies are transferred to the drug-containing vial for 24 h to ensure adequate drug levels prior to infection. Overall, feeding is regarded as the method of choice when long-term treatment is required.

24.9.2 Pilot Studies of Antifungal Drug Efficacy in Drosophila Model of IA

Despite the limitations of the fly model of IA mentioned earlier, our group has recently shown [14] that Drosophila is a reliable model for testing orally absorbed antifungals with anti-Aspergillus activity. Specifically, Toll-mutant flies fed in vials containing voriconazole and infected with A. fumigatus had significantly better survival rates than did control (untreated) Toll-mutant flies infected and maintained in regular vials without drugs. In addition, we found that the fungal burden was significantly decreased in voriconazole-treated flies compared with control flies, as assessed using quantitative real-time polymerase chain reaction and histopathologic analysis. Furthermore, the combination of voriconazole and terbinafine, two drugs that block sequential steps in the ergosterol pathway and show synergy in vitro against Aspergillus, was synergistic in the Drosophila model of IA [14].

For most experimental invertebrate models, the preferred method of administration for the study of the pharmacologic characteristics of agents is injection. Injection is relatively easy to perform in large insects such as caterpillars [15,16], but it is time-consuming and requires technical sophistication and specialized equipment in fruit flies [29].

Overall, a significant limitation of antifungal drug efficacy studies in Drosophila and other mini-host models is that they do not measure drug levels or include a pharmacokinetic analysis. Although high-performance liquid chromatography analysis and bioassay methods are feasible in invertebrates, such studies are more cumbersome, imprecise, and technically challenging in these models than in mammals. In addition, the injection of multiple doses of antifungal agents for long periods of time, a requirement for pharmacokinetic studies, is difficult in invertebrates because it results in increased mortality rates associated with repeat injury. Finally, little is known regarding the metabolism and elimination pathways of drugs and the potential for drug–drug interactions in mini-host models.

24.9.3 Virulence Testing of Aspergillus in Drosophila

Only a few entomopathogenic fungi are able to infect fruit flies in nature, via penetration of fly exoskeleton. Even when fungal pathogens are experimentally introduced directly into the fly hemolymph wild-type flies are still capable of effectively combating infection. For example, an injection of 10^6 conidia of A. fumigatus in wild-type D. melanogaster resulted in a survival rate of almost 100% [9,14]. In the mid-1990s, Lemaire et al. [9] were the first to show that A. fumigatus was able to infect and kill flies carrying mutations in various aspects of the Toll pathway. The usefulness of the fly model in studying virulence mechanisms in A. fumigatus is also suggested by the fact that the alb1 A. fumigatus mutant [30], which lacks the ability to produce melanin and exhibits attenuated virulence in a mouse model of IA, displayed a hypovirulent phenotype in Toll-mutant flies infected by ingestion or rolling (Table 24.1) [14]. Nevertheless, similar to recent findings with the ΔCgrA mutant [28], putative virulent factors of A. fumigatus with a role in thermotolerance may not be encountered in Drosophila or other invertebrate models because infection in these mini-hosts takes place at temperatures much lower (25°C) than the mammalian physiologic temperature (37°C).

High throughput screens for A. fumigatus mutants with attenuated virulence in fruit flies. The recent completion of the sequencing of the A. fumigatus genome and the development of molecular toolsets to study the biology of A. fumigatus is expected to lead to the generation of multiple Aspergillus mutants and creates a need for high-throughput strategies capable of assessing the contribution of individual genes to Aspergillus virulence.
We recently screened for the concordance in virulence between mice and the *Toll* fly model of aspergillosis. Specifically, we tested the virulence of 22 *A. fumigatus* strains in *Tl* flies in triplicate, in a blinded fashion. Three of 22 strains submitted for testing were hypovirulent mutants in mice (ΔcpcA, Krappmann et al., [31]; ΔGlip, Spikes et al., [32]; H515, Brown et al., [33]). We infected flies by injecting a standardized inoculum of *A. fumigatus* (~800 conidia/fly) into the thorax and monitoring survival daily for eight days after inoculation. *A. fumigatus* isolates with attenuated virulence in the fly model were selected if they satisfied the following criteria: infection of flies with the *A. fumigatus* mutant resulted in statistically significant differences in the mortality rate (P < 0.05) compared with the wild-type strain, as assessed by a Kaplan Meier analysis of the mortality curves; the complemented *A. fumigatus* strain exhibited restored virulence, similar to that of the wild-type strain in flies; and the insertion mutation in the gene did not result in severely impaired (e.g., 50%) *in vitro* growth of the mutant compared with the isogenic wild-type strain.

Similar to its effect in mice, ΔcpcA (which is impaired in the stress response to amino acid starvation), was hypovirulent in *Tl* flies (median survival, 3.5 days) compared with the wild-type D141 strain (median survival, two days, P = 0.02). Also, ΔGlip (which is deleted in the gene that encodes gliotoxin production), displayed attenuated virulence in *Tl* flies (median survival, four days) compared with the wild-type *Af293* flies (median survival, two days, P = 0.001), whereas the PABA auxotroph H515 was found to be completely avirulent in both *Tl* flies and mice. Finally, all 19 *A. fumigatus* strains that were virulent in mice retained their virulence in flies.

# 24.9.4 *Galleria mellonella*

Larvae of *Lepidoptera* insects such as the greater wax moth *G. mellonella* and the silkworm *Bombyx mori* have been successfully used as models of fungal pathogenesis because of their relatively large size (~2 cm and 5 cm long, respectively), which allows for the injection of standardized fungal inocula and studies of drug pharmacodynamics. In addition, *in vivo* studies of phagocytic cell function are feasible with these invertebrates (Table 24.2). Importantly, although their optimal temperature of growth and maintenance is 29°C, *G. mellonella* and *B. mori* larvae are able to survive at the mammalian physiologic temperature (37°C), which may allow for the expression of certain temperature-regulated virulence factors of fungal pathogens in these models. Nevertheless, the effect of increased temperatures on the *G. mellonella* immune response has not been studied in detail. In fact, some studies have shown that *G. mellonella* exhibits increased susceptibility to pathogenic fungi at mammalian physiologic temperatures [15,22]. In contrast, in all other invertebrate models, infection experiments are typically performed at temperatures of 22°C–30°C [10,12,22]. The major disadvantage to using *Lepidoptera* insects is the lack of tools for genetic analysis and the lack of genome sequencing. However, the *B. mori* genome sequencing project is nearly complete, and the number of genetic tools and techniques for lepidopteran genetic analysis are quickly accumulating (http://www.ab.a.u-tokyo.ac.jp/lep-genome).

### Table 24.1 Overview of Published Virulence Studies of *A. fumigatus* in Mini-Host Models

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Mini-Host Model</th>
<th>Phenotype</th>
<th>Comments</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔcgrA</td>
<td><em>D. melanogaster</em></td>
<td>CGRA encodes a nucleolar protein implicated in thermotolerance</td>
<td>Hypovirulent in mice, virulent in <em>D. melanogaster</em></td>
<td>28</td>
</tr>
<tr>
<td>Δalb1</td>
<td><em>D. melanogaster</em></td>
<td>Defective in melanin production</td>
<td>Hypovirulent in mice and <em>D. melanogaster</em> (feeding assay)</td>
<td>30</td>
</tr>
<tr>
<td>ΔcnaA</td>
<td><em>G. mellonella</em></td>
<td>CnaA encodes for the catalytic subunit in the calcineurin pathway</td>
<td>Almost avirulent in mice and <em>G. mellonella</em></td>
<td>36</td>
</tr>
<tr>
<td>Δpes1</td>
<td><em>G. mellonella</em></td>
<td>Pes1 encodes for a nonribosomal peptide synthetase</td>
<td>Reduced virulence in <em>G. mellonella</em></td>
<td>37</td>
</tr>
</tbody>
</table>
24.10 Virulence Studies in *G. Mellonella* and Other Invertebrate Models of Aspergillosis

The only other mini-host model in which *A. fumigatus* virulence has been studied of is *G. mellonella* (Table 24.1). *G. mellonella*, like the other invertebrates, is enormously resistant to infection by *A. fumigatus*. Importantly, it was recently demonstrated that the stage of conidial germination of *A. fumigatus* has a significant effect on the virulence of *A. fumigatus* in *G. mellonella* larvae because of the associated differences in the rate of phagocytosis [34]. Specifically, while resting conidia of *A. fumigatus* were avirulent in the larvae of *G. mellonella*, even when injected in high inocula (up to 10⁷ conidia per insect), swollen (more than 3 μm in size) or germinating conidia were highly virulent and were associated with significantly reduced rates of phagocytosis by hemocytes.

In addition, the same group of investigators recently reported that *G. mellonella* was extremely susceptible to the *A. fumigatus* strain ATCC 26933 [35]. This *A. fumigatus* isolate was shown to produce gliotoxin. Gliotoxin is an *Aspergillus* metabolite that exhibits immunosuppressive and apoptotic activity against immune effector cells *in vitro*, implying that it is a virulent factor in the *Galleria* model of IA. However, comparative studies of gliotoxin gene-deletion *A. fumigatus* mutants with isogenic controls are needed to provide definitive answers regarding the role of gliotoxin in *Aspergillus* virulence. Recently, an *A. fumigatus* mutant deleted in calcineurin A (ΔCnaA), the catalytic subunit of the calcineurin pathway, exhibited significant defects in conidial cell wall structure and lateral filamentation and was shown to be hypovirulent, both in the mouse model of aspergillosis and in *G. mellonella* [35]. Similarly, another *A. fumigatus* mutant (ΔPes1), deleted for a nonribosomal peptide synthetase with a potential role in tolerance against oxidative stress, was found to be hypovirulent in *G. mellonella* [36].

### TABLE 24.2
Comparison of Different Mini-Host Models in the Study of Aspergillosis

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Fruit Fly (<em>D. melanogaster</em>)</th>
<th>Wax Moth (<em>G. mellonella</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic tractability</td>
<td>++</td>
<td>−</td>
</tr>
<tr>
<td>Sequenced genome and annotation</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Adaptive immunity</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Survival at 37°C</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Phagocytic cell studies</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Suitable for antifungal studies</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Correlation of <em>Aspergillus</em> virulence factors in mammalian models of IA</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Potential for high-throughput screening</td>
<td>++</td>
<td>±</td>
</tr>
</tbody>
</table>

24.11 Future Perspectives

The prospects of mass screening of candidate antifungal compounds in invertebrates. Despite their limitations, invertebrates are attractive models for mass screening candidate antifungal compounds that will require subsequent validation in mammalian systems. Such approaches have been used successfully in *Drosophila* to select life-extending compounds [19] and in *C. elegans* to screen for novel anthelmintic microbial molecules [18]. Recently, investigators performed the first *in vivo* high-throughput screening of antimicrobial compound libraries by developing a *C. elegans–Enterococcus faecalis* model in liquid medium using standard 96-well microtiter plates [18]. They identified 16 of 6000 synthetic compounds and 9 of 1136 natural product extracts that promoted nematode survival. Importantly, many of the compounds and extracts had little or no effect on *in vitro* bacterial growth but significantly promoted the survival of nematodes *in vivo*. The results of this pioneer study indicate the significant advantages of *in vivo* mass screening, including simultaneous toxicity testing and identification of not only traditional...
antibiotics but also prodrugs and compounds that target bacterial virulence or stimulate host defense. Similarly, another group of investigators combined the ex vivo culture of Drosophila cells with a reporter gene that reflects the heat shock response and demonstrated that the resulting system is capable of screening compounds that act specifically on innate immunity, including mammalian innate immune responses [19]. Using this ex vivo system, they identified novel compounds that selectively activated the Imd pathway.

Dissecting the molecular attributes of host immune response against Aspergillus in invertebrates. Recently, we used a genomic approach to explore the molecular aspects of immunopathogenesis of A. fumigatus (nonpathogenic) and Zygomycetes (pathogenic) infection in fruit flies. We infected wild-type flies with Rhizopus or A. fumigatus and performed a whole-genome microarray analysis to identify genes that were differentially expressed within 12 h of infection [26]. In addition, we used the S2 D. melanogaster phagocytic cell line to determine fly cellular immunity to Zygomycetes compared with A. fumigatus by assessing phagocytosis and hyphal damage. Compared with Zygomycetes, A. fumigatus infection up-regulated 36 genes that encode for pathogen-recognition (2), immune-defense (5), stress response (3), detoxification (2), steroid metabolism (1), tissue reconstruction (13), and unknown functions (10). Furthermore, and similar to findings with human phagocytic cells [38], A. fumigatus hyphae were much more susceptible to killing by S2 cells than were Zygomycetes hyphae [26].

24.12 Concluding Remarks

In recent years, mini-host models have been increasingly appreciated as important in determining the molecular mechanisms of Aspergillus pathogenesis and host innate immunity. A variety of genetically amenable invertebrate hosts are expected to become key components of genomic strategies to scan the entire genomes of medically important fungi, such as A. fumigatus, for pathogenicity-related genes. Also, these elegant pathosystems may be adapted as high-throughput assays to screen for new types of antifungal compounds that target specific virulence attributes of pathogenic fungi. However, significant questions concerning the future role of mini-host models continue to develop. Importantly, the pathophysiology of mini-host death caused by Aspergillus and other opportunistic fungi has not been fully explored. In addition, the immune mechanisms that mediate resistance to Aspergillus in invertebrate epithelia are less well characterized. Furthermore, the essential molecular structures of Aspergillus that mediate recognition by the Toll pathway in Drosophila and other insects have not been elucidated. For example, it is unclear whether the recognition of different classes of opportunistic fungi is induced by the same pathogen-associated molecular patterns in Drosophila.

Finally, despite the considerable similarities in innate immune mechanisms, invertebrate models are not directly comparable with mammalian models. Thus, it is reasonable to speculate that some of the virulence attributes of Aspergillus and other pathogenic fungi that affect mammals might not be important in invertebrate mini-host models. Each of the existing mini-host models has advantages and disadvantages, which highlights the need for several models to help us understand the mechanisms of Aspergillus pathogenicity.

24.13 Conflicts of Interest

DPK has received research support and honoraria from Merck & Co., Inc.; Fujisawa, Inc.; Enzon Pharmaceuticals; and Schering-Plough. GC has no conflicts of interest.

References


25

Food Products Fermented by Aspergillus oryzae

Keietsu Abe and Katsuya Gomi

CONTENTS

25.1 Introduction .................................................................................................................................. 429
25.2 Soy Sauce Fermentation ............................................................................................................... 430
  25.2.1 Background ................................................................................................................................ 430
  25.2.2 History ........................................................................................................................................ 430
  25.2.3 Outline of the Manufacturing Process of Soy Sauce .............................................................. 431
  25.2.4 Raw Materials ............................................................................................................................ 431
  25.2.5 Treatment of Raw Materials ..................................................................................................... 431
25.3 Soybean Paste (MISO) Fermentation (Rice-Miso and Barley-Miso) ............................................. 433
25.4 Sake (Japanese Rice Wine) Fermentation ....................................................................................... 434
  25.4.1 Outlines of the Manufacturing Process of Sake .................................................................... 434
25.5 Shochu (Japanese Spirits) Fermentation ......................................................................................... 436
25.6 Conclusion ...................................................................................................................................... 437
References .............................................................................................................................................. 438

25.1 Introduction

Traditional fermented foodstuffs may be divided into six categories as follows: (1) alcoholic beverages fermented by yeasts; (2) vinegars fermented with *Acetobacter* or *Gluconobacter*; (3) fermented milk products treated with lactic acid bacteria; (4) pickles fermented with lactic acid bacteria in the presence of salt; (5) fermented fish or meat treated with enzymes together with lactic acid bacteria in the presence of salts; and (6) fermented plant proteins treated with molds, with or without lactic acid bacteria and yeasts in the presence of salt [1]. Since ancient era, people all over the world have developed traditional processes for the production of alcoholic beverages. The beverages are prepared principally by converting sugars present in the raw materials into alcohol by the action of yeasts. Vinegars have been made from almost all of these alcoholic beverages by converting the alcohol into acetic acid by the action of *Acetobacter* or *Gluconobacter*. Fruit wines have been made from sweet fruits such as grapes, apples, and oranges, and the corresponding vinegars made from these wines.

To prepare wines from starchy raw materials such as wheat, barley, corn, or potatoes, these raw materials must first be degraded into sugars, mainly glucose, to allow fermentation by yeasts; there are significant differences between the saccharification process of Western countries and that of the Orient. The amylolytic enzymes used for the saccharification in Western countries have been derived from sprouted barley or malt, but in the Orient, *Aspergillus* or *Rhizopus* molds have been utilized as the source of amylolytic enzymes. Accordingly, in the preparation of beer, whisky, gin, and vodka, the starch in the corresponding raw materials is first saccharified (enzymatically hydrolyzed) using malt, while in the preparation of alcoholic beverages from rice, wheat or sweet potatoes in the Orient, *Aspergillus* and *Rhizopus* molds are cultured on part of these solid raw materials to produce amylolytic enzymes. These cultured materials are called "chu" in China and "koji" in Japan; the koji is then mixed with the remainder
of the raw material and water to make a mash. The mash is concurrently subjected to enzymatic saccharification, lactic acid fermentation, and yeast fermentation.

Producers for enzymatically hydrolyzing certain protein foods into amino acids and lower peptides to make them more attractive and nourishing have been known for a long time. In the Western world, the flavor of cheese has been enriched by fermenting it with some Penicillium molds; in the Orient, the flavors of fish, poultry, meat, pulses, and some cereals have been improved by fermenting them with proteolytic and amylolytic enzymes produced by Aspergillus. Rhizopus, or Rhizomucor molds or by Bacillus, sometimes accompanied by fermentation with lactic acid bacteria and yeasts in the presence of high salt concentrations [1]. These foods, formerly called chang and shi in China, hishio or sho, and kaki in Japan, can still be found in every Asian country including Japan. They are believed to be the forerunners of foods such as shoyu (soy sauce) and miso which are now in use; their records date back 3000 years in China and they are believed to have been introduced from China to Japan 1300 years ago or more.

Nowadays the typical fermented foods produced in Japan are shoyu (soy sauce), miso (fermented soybean paste), sake (Japanese rice wine), shochu (Japanese spirits), rice vinegar, natto (soybeans fermented by Bacillus), and pickles. With the exception of natto and pickles, there is a common method for production of the fermented foods. The first step is cultivation of molds belonging to the genus Aspergillus, such as Aspergillus oryzae, Aspergillus sojae or black aspergilli (Aspergillus awamori and Aspergillus kawachii) on part or all of the raw material to produce plant-tissue degrading enzymes such as amylases, proteases, lipases, cellulases, and pectinases [2]. The mold-cultured material is called koji in Japanese, and the koji is used to degrade raw materials with water or sodium chloride solution. According to old records, Rhizopus or Mucor were used in China and are still used in other Oriental countries, for example, Taiwan and Indonesia, but only Aspergillus molds have been used in Japan. In Japan, annual production volume of koji is over one million ton and most of the koji is produced by large-scale solid phase culture systems of A. oryzae, A. sojae, A. awamori, and A. kawachii. Because koji fermentation in Japan is the largest industrial solid-phase culture system of Aspergillus in the world, the fermentation system represents Oriental fungal industry. This article describes fermentation process of Japanese food products made by koji molds such as A. oryzae, A. sojae, A. awamori, and A. kawachii.

25.2 Soy Sauce Fermentation

25.2.1 Background

Soy sauce produced in Japan is classified into three different types on the basis of three production methods [1,3]. The Japan Agricultural Standard (JAS) recognizes three production methods as follows: (1) genuinely fermented; (2) semichemical in which fermented soy sauce mash or soy sauce is mixed with a chemical or enzymatic hydrolyzate of plant protein, in amounts of 50% and 30%, respectively, on nitrogen basis, and fermented and aged more than one month; and (3) amino acid solution mixed in which (1) or (2) is mixed with a chemical or enzymatic hydrolyzate of plant protein in amounts of less than 50% or 30%, respectively, on nitrogen basis. According to JAS, in the production of genuine fermented soy sauce, heat-treated raw materials, soybeans and wheat should be inoculated with koji molds (A. oryzae or A. sojae) and cultured to make koji, and then the koji is mixed with salt water to make a mash (so called “moromi” in Japanese). The moromi-mash is fermented with the halophilic lactic acid bacterium, Tetragenococcus halophilus [4–6], and the halo-tolerant yeast Zygosaccharomyces rouxii [7,8]. The annual production volume of soy sauce is about one million Kl of which approximately 80% is genuine soy sauce [9].

25.2.2 History

The origin of soy sauce (shoyu) is thought to originate from a Chinese food, called “sho” in Japanese. The first description of sho was found in “Shurai” (in Japanese), which was originally written in China 3000 years ago. According to the book, sho was made by aging a mixture of mold culture, foxtail millet, dried meat, and liquor in a bottle for 100 days. The final product was in the form of a mash or “miso” (fermented soybean paste) which is one of a member of typical fermented foods in Japan (a semisolid
Food Products Fermented by Aspergillus oryzae

slurry food made from soybeans, rice or barley, and salt by fermentation). Soybeans are not described in the book as one of the raw materials of sho, however, it is quite possible that soybeans were used, because it is known that soybeans were widely cultivated in China 4000 years ago. The first documentation of the clear liquid part of sho is the Chinese book “Chin-Min-Yao-Shu” (Saimin-Yojutsu in Japanese) written in AD sixth century. Sho or hishio was also made from fish and salt at the beginning of Japanese history. Sho made from soybeans is presumed to have been introduced from China into Japan along with other foods by the Chinese Buddhist priest Ganjin (AD 1254). The first record of the name of “shoyu” (soy sauce) appeared in AD 1595, and it is likely that the basic manufacturing process of today’s Japanese-type soy sauce was derived by the early seventeenth century. In the Edo era (1603–1867) of Japan, the technology for preparing soy sauce developed dramatically and the scale of production increased. However, remarkable improvements in the processing of soy sauce have taken place since 1950, with dramatic advance in both biochemistry and technology.

Soy sauce was exported to the Netherlands from Japan for the first time back in 1668, and at present it is exported to more than 90 countries as a seasoning of world-wide appeal [1,3,9,10].

25.2.3 Outline of the Manufacturing Process of Soy Sauce

We describe the process for manufacturing of “koikuchi” soy sauce (deep reddish brown color soy sauce, which is most popular all over the world now) as a representative of several different types of soy sauce [1,3,9,10]. The process consists of five major steps as follows: (1) treatment of raw materials; (2) koji making; (3) mash fermentation and aging; (4) pressing and refining; and (5) pasteurization. The outline of the process is shown in Figure 25.1.

25.2.4 Raw Materials

The raw materials of soy sauce and soybean paste are soybeans, wheat, rice, barley, rye, salt, and water. Soybeans and defatted soybeans: The protein content of soybeans and the amino acid composition of that protein greatly influence the quality and yield of soy sauce, which is usually calculated on the basis of nitrogen content. Whole and defatted soybeans contain approximately 33% and 45% (w/w) protein, respectively. The key component for soy sauce production is protein, because about three-quarters of the total nitrogen in soy sauce originates from soybeans. Generally, soybeans contain about 20% crude oil, of which 94% to 97% is composed of glycerides of higher fatty acids (linoleic 10%; linolenic 55%; oleic 20%; stearic 10%; and palmitic 12%) and 2% of phospholipids. Soybean oil is degraded into higher fatty acids and glycerol by the action of lipase derived from koji molds mainly in the salty mash. Soybean oil or degraded fatty acids are present at the upper layer of soy sauce and thus can be separated from soy sauce that is derived from mash by filter-press as described later and only glycerol remains in soy sauce. Wheat: The total nitrogen contained in soy sauce is derived from soybeans (75%) and the remainder from wheat kernels. The ratio of soybeans and wheat as raw materials for “koikuchi” soy sauce ranges from 6:4 to 4:6. The glutamic acid contents of soybeans and wheat are 20% and 30% of the total amino acids, respectively. Proteins present in wheat kernels are good sources of glutamic acid, which is an important taste ingredient of soy sauce. Salt and water: Salt of purity greater than 95% is generally used; water low in Fe, Mn, and Cu is preferred because these metals promote oxidative browning of the color of soy sauce.

25.2.5 Treatment of Raw Materials

Whole soybeans, or more commonly defatted soybean grits are moistened and cooked with steam under pressure. The process greatly influences the digestibility of soybean protein. Wheat kernels are roasted at 160–180°C for less than one minute, then coarsely crushed.

Koji production: The main purpose of this process is to produce enzymes required for hydrolysis of the raw materials. Many nutrients for lactic acid bacteria and yeasts are supplied by enzymatic hydrolysis of the raw materials in the next step. Some flavor components that influence the quality of soy sauce are also produced in this process.
The Aspergilli

**Procedure:** Nearly equal amounts of cooked soybeans and roasted crushed wheat (6:4–4:6) are mixed and inoculated with a pure starter culture of *A. oryzae* or *A. sojae*—so-called koji starter or seed mold. The inoculated mixture is then transferred to the equipment for koji cultivation, and spread onto large perforated stainless steel trays (3 [wide] × 12 [long] × 0.6 [high] m) to a depth 30–40 cm, and then incubated in a room at 25–30ºC for two to three days. During this period, the temperature, moisture, and aeration are controlled to allow the seed mold to grow on the mixture, and to promote the production of enzymes. A temperature above 35ºC leads to a death of the koji mold. The temperature of the koji is controlled by stirring. The first stirring is performed at about 20 h, and the second at about 25 h, after inoculation. The resulting end product (clear yellow to yellowish green color) is koji in which the surfaces of the raw materials are covered by grown fungal cells. The aforementioned method is a batch system. In recent years, an advanced system for continuous koji cultivation has been developed and industrially employed [1].

**Koji molds:** The koji molds used for soy sauce are *A. oryzae* and *A. sojae* [1]. These species differ not only in their conidial, morphology, but also in several physiological characteristics of soy sauce-manufacturing importance. Generally, *A. oryzae* is characterized by a productivity of α-amylase, and...
A. sojae is characterized by a high productivity of protease. A. oryzae is used not only for soy sauce production but also for the other Japanese fermented foods such as miso and sake. However, the use of A. sojae is limited to soy sauce manufacture [1,2,10].

The breeding of koji molds has been performed, using mutation, crossing, and protoplast fusion, in an attempt to increase their enzyme productivities. At the present time, several koji molds with a high enzyme productivity of protease and amylase are used to ensure a high productivity of the whole process.

Mash production and aging: In mash production, the koji is mixed with 120 to 130 volume percent water containing 23% to 25% salt. The mash that is called moromi in Japanese is transferred to deep tanks (50–300 kl). The mash is kept for four to eight months, depending upon its temperature, with occasional agitation with compressed air to mix the soluble components uniformly and to promote microbial growth. During the fermentation period, the enzymes from the koji mold hydrolyze most of the protein into amino acids and low molecular weight peptides. Around 20% of the starch is consumed by the mold during koji cultivation, but almost all of the remaining starch is converted into simple sugars; more than half of these are fermented to lactic acid and alcohol by lactic acid bacteria and yeasts. The pH drops from an initial value of 6.5–7.0 to 4.7–4.9. The lactic acid fermentation is gradually replaced by yeast fermentations.

Cultures of T. halophilus and Z. rouxii are added to the mash. The salt concentration of the mash limits the growth to a few desirable microorganisms.

Mash pressing: A matured mash is filtered at high pressure through cloth and liquid part is squeezed out under hydraulic pressure, which sometimes reach 100 kg cm$^{-2}$, for one to three days. The difficulty of pressing soy sauce mash is due to the viscosity of more than 300 cps.

The liquid part of the mash is called raw soy sauce or “nama-shoyu” in Japanese. The residue from the pressing is called soy sauce cake, or “shoyu-kasu” in Japanese. This can be used as an additive in animal feed. The final moisture content of soy sauce cake is less than 25%.

Refining: The liquid part of the mash obtained is stored in a tank and separated in three layers; sediment in the bottom, a clear middle layer, and an oily layer floating on top. The oil layer is removed by decanting. The middle layer is heated at 113–120ºC for a few seconds in a heat exchanger in order to pasteurize any microbial cells, denature enzymes, coagulate proteins, develop the reddish brown color, and generate aroma. The resultant clear soy sauce is then filtered, bottled, and marketed.

### 25.3 Soybean Paste (MISO) Fermentation (Rice-Miso and Barley-Miso)

Whole yellow soybeans are used for the preparation of ordinary miso. Dehulled soybeans or soybean grits are sometimes employed for the production of white or pale yellow rice-soybean paste. The outline of manufacturing of soybean paste is described as follows (Fig. 25.2) [1,10]. Soybeans are soaked in water until saturation and then cooked for 30–60 min at normal pressure, or cooked in four volumes of water for 20–30 min at a pressure of 0.5–0.7 kg cm$^{-2}$, or steamed for 20 min at a pressure of 0.7 kg cm$^{-2}$ (115ºC), either batch wise or continuously. Cooked soybean granules are preferably pressed using less than 0.5 kg cm$^{-2}$ pressure. Milled rice or barley or rye is soaked in water and then steamed batch wise in an open cooker for 40 min or continuously on a net conveyor in a closed autoclave for 30–60 min. The koji cultivation on rice or barley or rye with A. oryzae is conducted at 35ºC to 38ºC, sometimes with an increase in temperature up to almost 40ºC in the final stage, for 40–48 h. The finished koji is mixed with salt to stop further mold growth and to minimize the inactivation of enzymes. The amount of salt used is about 30% by weight of the koji. Finished koji is sometimes cooled instead of adding salt. Various types of koji fermenter are employed nowadays. Cooked soybeans are mixed with salted rice- or barley-koji, a small amount of water and inoculums of cultured lactic acid bacteria (T. halophilus) and yeasts (Z. rouxii), if necessary. It is important to mix these materials uniformly so that the variation in salt concentration in the mash is less than 0.5%. The mixture is packed in a fermentation tank; it is moved from one tank to another at least twice during the fermentation to mix the contents and to provide aerobic conditions suitable for growth of yeasts. Fermentation is carried out at around 30ºC for one to three months depending on the type of soybean paste. Well-ripened soybean paste is then blended and mashed if necessary and pasteurized using a tube heater. About 2% alcohol is added to the product to terminate the growth of yeasts.
25.4 Sake (Japanese Rice Wine) Fermentation

Sake is the traditional Japanese alcohol beverage, as beer is the German one and wine is the French. Old Japanese historical books tell us that Japanese people enjoyed drinking sake at the court and at home about 2000 years ago as well as do Chinese history book that describes Japanese people’s life. In ancient times, people used to drink turbid sake (*doburoku*, sake without pressing or filtration), but later they invented a way to filtrate the fermented main mash (*moromi*) easily. Filtrated sake (*seishu*, which means clean sake) since then has become popular.

The main process of sake fermentation is outlined in Figure 25.3 [11,12] and is somewhat complicated, compared with that of beer or wine, because two major kinds of microorganisms are used to make sake; one is koji mold (*A. oryzae*) for saccharification of starch in rice, the other is sake yeast (*Saccharomyces cerevisiae*) for ethanol fermentation. In addition to *A. oryzae* and *S. cerevisiae*, lactic acid bacteria such as *Leuconostoc mesenteroides* and *Lactobacillus sake* are involved in traditional procedures for sake fermentation to reduce pH of seed mash (*moto*) into acidic condition and prevent from contamination of wild yeasts and deleterious bacteria. Nowadays, lactic acid is commonly added in the process of seed mash simultaneously with addition of sake yeast instead of lactic acid fermentation.

25.4.1 Outlines of the Manufacturing Process of Sake

Raw materials used in sake fermentation are rice and water. Since water comprises as much as 80% of the final product, fine and clean water is of importance in sake fermentation. In particular, iron should not be contained in water used for sake fermentation, because a kind of siderophore called deferriferrichrysin produced by *A. oryzae* during koji preparation chelates a ferric ion to yield ferrichrysin with red color, which deteriorates the quality of final product [13]. Moreover, rice used for sake making is also pivotal for good quality of sake. Rice suitable for sake fermentation is as follows: (1) rice grain is relatively large,
(2) white opaque portion (called shinpaku in Japanese) resulted from small gaps between starch granules appears inside of rice grain, and (3) content of protein in rice is relatively low.

**Rice milling**: The peripheral layers of brown rice contain lipids, proteins, and minerals, excess of whose content is undesirable for the quality of sake, and thus brown rice is polished by milling to remove these materials, with a polishing ratio (means the left over weight after milling) below 70%. Recently, the polishing ratio of white rice used for making premium sake is 50% or less.

**Rice steaming**: The polished rice is washed to remove the bran left on the surface of the grains, and then soaked in water to attain about 30% of water content. When rice has been polished below 50% of polishing ratio, it absorbs water fast so much so that soaking period should be shorter than usual (less than 30 min). After soaking, the rice is drained off and steamed for about 1 h, and then cooled prior to use. Out of rice used for sake fermentation, about 20% is used for koji preparation, and the rest is directly used for main mash (moromi).

**Koji preparation**: The steamed rice is cooled to about 35°C and taken to the koji-making room (koji-muro) where higher temperature (about 30°C) and humidity (about 80–90%) are maintained. Then conidiospores of *A. oryzae* (tane-koji, seed mold or koji starter) are sprinkled on steamed rice, then mixed thoroughly, gathered to heap and covered with cloths. After 20 h of inoculation, mycelia of *A. oryzae* develop over the rice grain and can be seen by the naked eye, then the temperature of the molded rice rises, which thus is controlled to not exceed over 40°C to 42°C for the next 20 to 24 h. In contrast to koji preparation in soy sauce fermentation where the temperature should not be over 35°C, the temperature of koji-making for sake fermentation should be over 35°C, because production of amylolytic enzymes is suitable above 35°C while that of proteolytic enzymes prefers below 35°C. The final product of koji is white in color and smells faintly of sweet chestnuts.

Koji starters are generally provided to sake factories by their manufacturers, which preserve hundreds of proprietary strains of koji molds. In sake making, *A. oryzae* is used in soy sauce as well as soybean paste production, but the properties of strains used are different with respect to enzyme production.
profiles. For example, *A. oryzae* strains used for sake making have an ability of production of abundant amylolytic enzymes (α-amylase and glucoamylase), while those for soy sauce making can produce much proteolytic enzymes. In addition, strains used for sake making are desired to produce a little peptidase and less tyrosinase. As stated earlier, because deferriferrychrysin produced by *A. oryzae* binds a ferric ion to form a red colored compound, ferrychrysin, strains with less production of it are preferable.

*Seed mash (moto) preparation:* Seed mash or yeast starter, called “moto,” is first prepared prior to the main fermentation. Seed mash plays an important role in sake fermentation by providing a large amount of pure and healthy living cells of sake yeast. For seed mash preparation approximately 7% of the total rice is used including that for koji. Finished koji and freshly prepared steamed rice are mixed with water and lactic acid, followed by addition of pure cultured sake yeast. General seed mash is ready for use as a starter for main mash (*moromi*) in two weeks, whereas it takes more than four weeks for traditional type of seed mash where lactic acid fermentation is carried out before addition of pure yeast.

*Main mash (*moromi)*: Main mash is prepared by adding finished koji, steamed rice, and water to seed mash in three successive stages over four days, roughly doubling the volume of the batch each time. Briefly, on the first day in main mash preparation, seed mash is transferred into the large main fermentation tank, followed by addition of koji, steamed rice, and water. The temperature of the main mash is generally 15ºC on the first day and then kept on the second day to allow sake yeast to proliferate rigorously. On the third (10ºC) and fourth (7–8ºC) day, additional koji, steamed rice, and water are added to the main mash tank. This main mash is maintained at about 10–15ºC over the next two to four weeks. During the fermentation of main mash, the starch in rice is liquefied and saccharified by the action of amylolytic enzymes from koji, such as α-amylase, glucoamylase, and α-glucosidase, and the sugar thus formed is fermented into ethanol by sake yeast. Thus, in sake fermentation, both processes of saccharification and ethanol fermentation take place simultaneously in a well-balanced manner, which is a very unique and complex method, called “multiple parallel fermentation,” contributes ethanol content up to 20%, higher than any other naturally fermented beverages.

*Pressing and filtration:* The fermented main mash is pressed through cloth or canvas-like bags at high pressure to separate the clear sake from the unfermented solids, called *sake-kasu* (sake cake). At first the clear sake is slightly turbid due to the presence of fine lees containing yeast cells and undissolved steamed rice, which are settled out and removed through filtration.

*Pasteurization and aging:* The fresh sake is heated at 60–65ºC by passage through a heat exchanger in order to inactivate the remaining enzymes and to kill the deleterious microorganisms such as *hiochi* bacteria, a kind of lactic acid bacteria (*Lactobacillus fructivorans*) whose growth requires *hiochic* acid (mevalonic acid) produced mainly by koji mold. This pasteurization procedure had been already employed about 500 years ago in sake making, while Louis Pasteur discovered it in mid-nineteenth century in wine making. The pasteurized sake is left to age for about three to six months, whereby the taste and flavor becomes to smooth and well balanced. Finally, the sake is bottled and shipped.

### 25.5 Shochu (Japanese Spirits) Fermentation

In southern parts of Japan, particularly in Kyushu and Okinawa Islands, shochu (Japanese spirits) is generally produced from various raw materials. There are various kinds of shochu depending on the raw materials and areas where shochu is produced. For example, shochu made from sweet potato (*imo-shochu*), from rice (*kome-shochu*), that from buckwheat (*soba-shochu*), and that from barley (*mugi-shochu*) are popular in Kagoshima, in Kumamoto, in Miyazaki, and in Oita, Nagasaki and Fukuoka, respectively. In addition, shochu made in Okinawa is commonly called “*awamori*” which is distilled from the main mash fermented of long-grain rice (indica subspecies of *Oryzae sativa*), whereas short-grain rice (japonica subspecies of *Oryzae sativa*) is used for shochu making in other areas. Each raw material brings a different unique flavor and aroma to the final product.

The procedure for shochu making is similar to that of sake making, except for distillation of the main mash (Fig. 25.4). In shochu making, black koji molds (black aspergilli) such as *A. awamori* and *A. kawachii* are commonly used to saccharify starchy materials. Since black koji molds produce much citric acid, koji made by these molds reduces the pH of the main mash to low (approx. pH 3–3.5), which
is very useful to prevent from infection by harmful bacteria in southern warmer districts. *A. kawachii* was isolated as an albino mutant from *A. awamori* in 1918, and has been used mainly in Kyushu Island.

In shochu making, distillation of the main mash is carried out by single distilling procedure, such as in whisky and brandy. Traditional distilling method, where the fermented mash is boiled by passage of steam up to 90–100°C in a pot still, is commonly employed. However, recently decompression distillation method has been developed to drop the boiling point of the fermented mash by reducing the atmosphere pressure to one-tenth of it through generating a vacuum in the pot still. Because the mash is distilled at low temperature (approx. 50°C), final product of shochu made by decompression distillation is mild and light.

### 25.6 Conclusion

The most distinguished characteristic of the processes for fermented food manufactures in the Orient is the solid fermentation using filamentous fungi such as *Aspergillus*, *Rhizopus*, *Rhizomucor*, etc. In Japan, the solid fermentation process is called “koji-making,” where *Aspergillus* molds are grown on steamed rice, wheat, or soybean. The koji is used as a source of various hydrolytic enzymes required for the production of fermented foods. Solid fermentation (koji-making) is an efficient process for enzyme production, and resulted enzymes are useful for food processing. Commonly, *Aspergillus* fungi can produce more hydrolytic enzymes in solid fermentation than in liquid fermentation. In particular, glucoamylase (GlaB) and aspartic protease (PepA) of *A. oryzae* and acid-stable α-amylase of *A. kawachii,
The Aspergilli

which play important roles in sake and shochu fermentation, are only slightly produced in liquid fermentation, but are produced to a significant level in solid fermentation. Recent research on their gene expression revealed that the genes encoding those enzymes are expressed at high levels in solid fermentation and at very low levels in liquid fermentation [14–16]. Although the molecular mechanism that controls the expression of those genes remains to be elucidated, gene expression profiles of Aspergillus during solid and liquid fermentation seem to be quite different [17,18].

In addition, solid fermentation has not only been used for fermented food manufacture but also employed for industrial enzyme production. The first commercial microbial enzyme in the world was “Takadiastase,” production process for which was developed by Jokichi Takamine in 1894 based on solid fermentation of A. oryzae grown on wheat bran. The same process is still now employed industrially for production of various fungal enzymes in Japan. Thus, solid fermentation process using Aspergillus molds originally developed in the production of traditional fermented foods has been further exploited for fungal biotechnology including commercial enzyme production.

References


26

Aspergillus as a Cell Factory for Protein Production: Controlling Protease Activity in Fungal Production

Machtelt Braaksma and Peter J. Punt

CONTENTS
26.1 Introduction .............................................................................................................. 441
26.2 Strain Development ................................................................................................. 442
  26.2.1 Classical Methods to Screen for Protease Mutants ......................................... 442
  26.2.2 Molecular Genetic Methods to Construct Protease Mutants .............................. 443
    26.2.2.1 Protease Genes ..................................................................................... 443
    26.2.2.2 Protease Regulators .............................................................................. 444
  26.2.3 A Novel and Efficient Method for Isolation of Protease-Deficient Fungi ............ 446
26.3 Fermentation Conditions ......................................................................................... 446
  26.3.1 pH Regulation ................................................................................................. 447
  26.3.2 Carbon Catabolite Control ............................................................................ 448
  26.3.3 Nitrogen Metabolite Control ........................................................................ 448
  26.3.4 Sulfur and Phosphorus Metabolite Repression .............................................. 449
  26.3.5 Induction of Protease by Protein .................................................................. 449
  26.3.6 Bioprocess Engineering ............................................................................... 450
  26.3.7 System Biology Approach ........................................................................... 450
26.4 Conclusion ............................................................................................................. 451
Acknowledgments ........................................................................................................ 451
References ..................................................................................................................... 452

26.1 Introduction

Since ancient times microorganisms have been used in a variety of traditional food processes (e.g., the production of alcoholic beverages, cheese, and bread). Fungi are applied in cheese-making and in traditionally oriental food such as soy sauce, tempeh, and sake. However, the presence and role of these microorganisms was for most processes only identified in recent times. Fungi, like Aspergillus oryzae in the production of sake, were discovered to play a key role in the production of this product by the excretion of enzymes. In 1894, the first microbial enzyme that was commercially produced appeared on the market, called “takadiastase”; it was in fact fungal amylase produced by Aspergillus oryzae [1]. Nowadays, a large number of fungal enzymes are commercially available and their application extends well beyond their traditional use in food processes. Glucoamylase, α-amylases, cellulase, lipase, and protease are only a few examples of enzymes produced by filamentous fungi that are commercially available. Aspergillus species, and particularly A. niger and A. oryzae, play a dominant role in the production of many of these enzymes [for a list of commercial enzymes see the Association of Manufacturers and Formulators of Enzyme Products (AMFEP) at www.amfep.org].

441
The Aspergilli

For the last two decades, filamentous fungi have also been explored as hosts for the production of heterologous proteins. Because of their established use as production host of homologous proteins aspergilli are the obvious expression system for heterologous proteins. The Danish company Novozymes A/S was in 1988 the first on the market with a nonnative fungal lipase (Lipolase) produced from a genetically modified A. oryzae strain (www.novozymes.com/en/mainstructure/ourscience/gene+technology) [2]. Since then several species of Aspergillus have been used to express a wide variety of foreign genes (see also the list of commercial enzymes of the AMFEP at www.amfep.org). However, the production of heterologous as well as homologous proteins is often limited by the high levels of proteases also produced by the fungal host organism. This review will focus on the role of protease activity in strain and process development. Both classical mutagenesis and gene disruption techniques have been applied to generate strains with reduced protease activity. In addition indeed production levels improved significantly when using protease deficient strains (e.g., tissue plasminogen activator (t-PA) production with a protease deficient A. niger strain [3]). Controlling the culture conditions can result in a further improvement of the heterologous protein production (e.g., green fluorescent protein (GFP) production with a protease deficient A. niger strain at controlled pH [4]). However, the production levels for heterologous proteins are in most cases one to two orders of magnitude lower than for homologous proteins.

With the availability of the complete genome sequence of several Aspergillus strains (A. oryzae [5] at www.bio.nite.go.jp/dogan/MicroTop?GENOME_ID=ao; A. nidulans [6] and A. terreus at www.broad.mit.edu/annotation; A. niger at www.jgi.doe.gov/aspergillus; A. fumigatus [7] at www.sanger.ac.uk/proj-ects/A_fumigatus; A. flavus at www.aspergillusflavus.org/genomics), homology searches for genes involved in the proteolytic systems of these organisms resulted in a number of genes encoding protease activity much higher than previously known. For example, for A. niger over 150 protease genes were found in the genome [8]. In comparison, before the genome sequence of A. niger was known, an extensive analysis of the proteolytic system of A. niger led to the identification of only eight protease genes [9]. Given this very large gene potential, actual protease production and its regulation is expected to be very complicated.

The understanding of the regulation of the proteolytic system of Aspergillus strains is still only in its infancy. The involvement of several wide domain regulatory systems (carbon catabolite repression, nitrogen metabolite repression, pH regulation [9]) and probably sulfur metabolite repression [10] in the overall regulation of protease expression in Aspergillus is suggested. In our research, we plan to use fungal transcriptomics and metabolomics [11,12] to further elucidate the proteolytic system and its regulation in these organisms in the years to come. This review gives state of the art in the protease research field and provides an outlook on new research approaches.

26.2 Strain Development

26.2.1 Classical Methods to Screen for Protease Mutants

Mutagenesis by means of X-ray or UV irradiation and chemicals mutagenesis were discovered in the first half of the past century. Hara et al. [13] describe the successful attempts of Iguchi (1955–1956) to isolate a mutant strain producing higher levels of protease compared to the parent strain. After X-ray irradiation a large number of isolates were screened in a laborious and time-consuming effort for a hyperproducing mutant. The screening procedure was greatly improved by the method developed by Sekine in 1969 which enabled the screening of a large number of isolates (see Hara et al. [13]). Around colonies grown on casein-containing medium a halo (clear zone) was formed of which the diameter has a significant correlation with the protease production (see Fig. 26.1).

These classical methods to generate and screen for mutants with altered levels of excreted protease are still successfully applied. Nowadays mutagenesis of spores is most often conducted with the less-aggressive irradiation with ultraviolet light instead of X-rays. This approach has been applied to isolate several protease-deficient mutants in different aspergilli, such as A. niger [14,15] and A. nidulans [16]. Also mutagenesis with mutagens such as nitrosoguanidine has been described [17,18]. After mutagenesis the spores are plated on milk or gelatin-casein medium. Mutants with low proteolytic activity are screened for reduced
degradation of casein which results in a reduced or no halo on those plates. In this way, Mattern et al. [14] isolated A. niger mutants with residual extracellular proteolytic activities that vary from 2% to 80% of the protease activity of the parental strain. Katz et al. [16] describe A. nidulans mutants with tenfold reduced levels of extracellular protease compared to the parental strain.

26.2.2 Molecular Genetic Methods to Construct Protease Mutants

26.2.2.1 Protease Genes

Clearly, the random mutagenesis approach results in potent production hosts, but the genetic basis of these mutants remains unknown and may have unwanted pleiotropic effects on fungal fermentation performance (e.g., gene expression, growth rate). Therefore, with the development of molecular genetic tools also a more targeted approach to obtain protease-deficient mutants became available.

The general strategy for this approach is the so-called reverse genetics. By separating proteins produced in culture medium by SDS-PAGE or chromatography and subsequently testing for protease activity (as determined, e.g., by protease activity on skim milk agarose) of the different bands or fractions several proteases can be identified. By determining the (partial) amino-acids sequence the protein oligonucleotide probes corresponding to these sequences can be designed. These oligonucleotides or PCR fragments generated by using similar oligonucleotides are subsequently used to screen genomic libraries to clone the corresponding protease genes. With the resulting clones a disruption vector for the protease gene can be constructed for actual gene disruption. The more recent availability of genome databases makes it also possible to use obtained amino acid sequences directly to clone the corresponding genes by genome mining using sequence comparison algorithms such as BLASTX. However, even with knowledge of the genome sequence, an activity screen (most preferably based on proteolytic activity against the protein one wants to produce) is still necessary to identify which of all the protease genes present in the fungal genome is actually new and most active and thus the desired target for gene disruption. Berka et al. [19,20] was the first to describe the construction of gene replacement vectors for Aspergillus, which were used to specifically delete the chromosomal DNA of the protease gene encoding the major extracellular acid protease aspergillopepsin A (PEPA) in A. awamori. Disruption of this pepA gene reduced extracellular proteolytic activity compared to the wild type. Similar results were achieved by disruption of the aspergillopepsin A gene in A. niger [14]. Probes containing part of the coding region of this pepA gene were also used to screen the genomic library of an A. nidulans strain [10]. And although A. nidulans appears to lack detectable acid protease activity, a clone which hybridized with the pepA gene was
obtained. This aspartic protease gene, which was designated *prtB*, was only expressed at a very low level. Furthermore, homologs of the *pepA* gene have been cloned from other *Aspergillus* species, such as *A. fumigatus* [21], *A. oryzae* [22], and *A. satoi* [23].

In nonacid-producing aspergilli, such as *A. nidulans*, neutral or alkaline proteases are responsible for the major part of the extracellular protease activity. Disruption of the gene coding for the dominant extracellular serine protease in *A. nidulans* strain resulted, when cultured under various medium limitations, in reduced levels of proteolytic activity under all culture conditions [10]. Controlled batch fermentations of an *A. sojae* strain with a disruption of an alkaline protease gene resulted in about 40% reduction of proteolytic activity in comparison to the wild type [24]. Shake flasks cultures with *A. oryzae* expressing the heterologous protein endoglucanase showed enhanced stability of this protein when an alkaline protease gene of the host strain was disrupted [25].

Not in all cases disruption of a protease gene results in decreased protease activity. Disruption of the serine protease gene (*sep*) in *A. flavus* led to a compensatory increase in the expression and production of metalloproteinase gene (*mep20*) [26]. Both wild type and mutant degraded elastin at the same rate. The authors concluded that the expression of the genes encoding both proteases is controlled by a common regulatory system and that the fungus has a mechanism to sense the status of the extracellular proteolytic activities.

An alternative method for reduction of expression of a particular gene is the use of antisense RNA. This approach was applied in an *A. awamori* strain used to express the heterologous protein thaumatin [27]. Even though an insertion in the *pepA* gene had resulted an inactive PEPA protein, thaumatin was still degraded. Another protease, aspergillopepsin B (previously believed to be a pepstatin-insensitive aspartyl protease, but more recently established to be a member of the newly discovered family of glutamic proteases [28]), was identified as the most likely protease responsible for this degradation. Expression of *pepB* antisense RNA improved thaumatin production with 30%. Nevertheless, thaumatin was still degraded, indicating the antisense mRNA had only a partial silencing effect on *pepB* gene expression. Disruption of the *pepB* gene resulted in a significant further increase of the thaumatin production. However, an advantage of gene silencing with respect to gene disruption is that it can be used to suppress the expression of complete gene families. Zheng et al. [1] describe that the expression of antisense RNA of the structural gene of carboxypeptidase in *A. oryzae* did not only decrease the activity of that carboxypeptidase, but also of two other carboxypeptidases [29].

Yet another approach to obtain strains with low protease levels is disruption of proteases that proteolytically activate other protease precursor proteins which require processing for their activation. Disruption of such a protease gene will have a direct effect on the protease activity of one or more other proteases, as was described for *A. niger*. Disruption of the gene of an intracellular acid protease (PEPE) in *A. niger* did not only reduce the intracellular pepstatin-inhibitable aspartyl protease activity, but also intracellular serine protease and serine carboxypeptidase activities were significantly reduced in the ΔpepE strain [30]. The transcription of these nondisrupted genes was not affected by the disruption of the single *pepE* gene. According to the authors this may indicate the presence of a cascade activation mechanism for several vacuolar proteases, triggered by the PEPE protein. A similar mechanism has been described for *Saccharomyces cerevisiae* [31].

In Table 26.1 describes disruptions of protease genes in *Aspergillus* strains and the resulting residual proteolytic activities are summarized. In this table the construction of multiple disruptants can lead to further decrease of proteolytic activities. This was shown for a ΔpepA ΔpepB ΔpepE triple disruptant in *A. niger* [30] and disruption in *A. fumigatus* of both a gene encoding an extracellular serine alkaline protease and a gene encoding an extracellular metalloprotease [32].

### 26.2.2.2 Protease Regulators

Finally, a very efficient approach to generate strains with low protease levels is through disruption of genes that influence the expression of multiple protease genes. Two groups of regulatory genes have been described so far. In the first place, genes that encode specific regulators of protease genes; second, genes that encode wide domain regulators. Interestingly, in the first group, to date, only one single gene has been identified both in fungi and yeast species. This gene is the *prtT* gene, as cloned from an UV-induced
**TABLE 26.1**
Effects on Secreted Protease Activity of Protease Gene Disruption Strains in Aspergilli

<table>
<thead>
<tr>
<th>Species</th>
<th>Name Disrupted Gene</th>
<th>Residual Extracellular Protease Activity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracellular serine protease (fam. S8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. flavus</em></td>
<td>sep</td>
<td>100%</td>
<td>[26]</td>
</tr>
<tr>
<td><em>A. fumigatus</em></td>
<td>alp</td>
<td>0–30%</td>
<td>[32,93,94]</td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>prtA</td>
<td>10–50%</td>
<td>[10]</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>alp</td>
<td>&lt;WT</td>
<td>[25]</td>
</tr>
<tr>
<td><em>A. sojae</em></td>
<td>alpA</td>
<td>60%</td>
<td>[24]</td>
</tr>
<tr>
<td>Vacular serine protease (fam. S8)</td>
<td><em>A. oryzae</em></td>
<td>pepC</td>
<td>N/A</td>
</tr>
<tr>
<td>Extracellular aspartyl protease (fam. A1)</td>
<td><em>A. awamori</em></td>
<td>pepA</td>
<td>&lt;&lt;WT</td>
</tr>
<tr>
<td><em>A. fumigatus</em></td>
<td>pep</td>
<td>&lt;&lt;WT</td>
<td>[96]</td>
</tr>
<tr>
<td><em>A. niger</em></td>
<td>pepA</td>
<td>15–20%</td>
<td>[14,30]</td>
</tr>
<tr>
<td>Vacular aspartyl protease (vacuolar) (fam. A1)</td>
<td><em>A. niger</em></td>
<td>pepE</td>
<td>−100%</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>pepE</td>
<td>N/A</td>
<td>[95]</td>
</tr>
<tr>
<td>Extracellular glutamic protease (fam. G1)</td>
<td><em>A. awamori</em></td>
<td>pepB</td>
<td>&lt;parentb</td>
</tr>
<tr>
<td><em>A. niger</em></td>
<td>pepB</td>
<td>95%</td>
<td>[30]</td>
</tr>
<tr>
<td>Extracellular metallo protease (fam. M35)</td>
<td><em>A. nidulans</em></td>
<td>pepI</td>
<td>N/A</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>npII</td>
<td>&lt;WT</td>
<td>[98]</td>
</tr>
<tr>
<td>Extracellular metallo protease (fam. M36)</td>
<td><em>A. fumigatus</em></td>
<td>mep</td>
<td>70%</td>
</tr>
<tr>
<td><em>A. niger</em></td>
<td>pepH</td>
<td>&lt;WT</td>
<td>[97]</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>npI</td>
<td>N/A</td>
<td>[98]</td>
</tr>
<tr>
<td>Multiple disruptants</td>
<td><em>A. fumigatus</em></td>
<td>alp, mep</td>
<td>&lt;&lt;WT</td>
</tr>
<tr>
<td><em>A. niger</em></td>
<td>pepA, pepB</td>
<td>10%</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>pepB, pepE</td>
<td>−ΔpepB</td>
<td>[30]</td>
</tr>
<tr>
<td></td>
<td>pepA, pepB, pepE</td>
<td>&lt;10%</td>
<td>[30]</td>
</tr>
</tbody>
</table>

*a*As determined with protease assays and expressed as percentage compared to the parent strain; N/A is data not available.

bParent strain is not the WT strain, but a classical pepA-deficient mutant.

*A. niger* mutant [33] [Punt et al., in preparation]. This mutant was suggested to be a regulatory mutant as at least two proteases, including aspergillopepsin A, were missing from the culture medium, while genetic data indicated the presence of a single semidominant mutation, not linked to the pepA gene [14]. Recent analysis has indeed shown that the prtT gene is actually a regulatory gene encoding a member of the Zn-binuclear cluster family [33] [Punt et al., in preparation]. Interestingly, this gene is unique for *Aspergillus* species but actually absent in *A. nidulans*. With the disruption of the prtT gene in *A. niger* total protease activity was reduced to 20% of the wild type [34].

Besides regulatory genes specific for proteases expression wide domain regulatory genes affect the expression of a broad spectrum of enzymes including proteases as a response to ambient pH (*pacC* gene), nitrogen source (*areA* gene) or carbon source (*creA* gene).

The *pacC* gene is expressed at alkaline pH and encodes a protein, which is able to activate the expression of other alkali-expressed genes and to prevent the expression of acid-expressed genes [35]. In
The Aspergilli

A. nidulans the expression of the major alkaline protease prtA gene is activated by PacC. However, disruption of the pacC coding region results in very poor growth, making this approach not very interesting to generate hosts for protein production [36].

The gene areA is expressed in the absence of preferred nitrogen sources such as ammonium and encodes a protein that activates transcription of genes encoding enzymes (like proteases) involved in the utilizing of other resources [37]. Disruption of the areA gene in A. oryzae resulted in increased production of the heterologous protein chymosin due to reduced protease activity [38]. Unfortunately, disruption of the areA gene in A. niger as well as A. oryzae also affected growth, even in culture medium with (low levels of) ammonium; this reduced growth was not noticed in A. nidulans [39,40].

The gene creA is expressed in the presence of preferred carbon sources like glucose. The CreA protein represses the synthesis of enzymes (like proteases) involved in the catabolism of alternative carbon sources [41]. However, attempts to disrupt the complete creA gene from A. nidulans resulted in lethal phenotypes [42] or mutants with extremely severe effects on morphology (namely reduced growth rate and reduced conidiation) [43].

Altogether, the approach of using gene disruption of wide domain regulatory genes seems unsuitable to generate protease-deficient fungal host strains for protein production due to pleiotropic growth defects of this type of mutants. Specific mutation of these regulatory genes, alleviating the severe phenotypic effects of the complete knockout mutants could be used [44]. However, this approach relies on selection of specific spontaneous mutants making this approach not generally applicable.

The wide domain regulatory mechanisms will be discussed in more detail later on in this chapter.

### 26.2.3 A Novel and Efficient Method for Isolation of Protease-Deficient Fungi

Although both the classical screening approach and the gene-based approach have resulted in improved host strains, it is clear that both approaches have their limitations. The classical approach is very labor-intensive, whereas the disruption approach is limited by the availability of gene information. Therefore, we have developed a (direct) mutant selection approach, similar as is available for a number of other traits in filamentous fungi (pyrG [45], niaD [46], sC [47]). This proprietary approach is based on a suicide substrate (SUI) to which protease mutants of fungi and yeasts are more resistant (SUI^R) than the parent strains [Punt et al., unpublished results]. The method can be used to select spontaneous mutants or mutants generated by mutagenesis by ultraviolet light irradiation. After a first round of selection the resulting mutants can be screened in a conventional milk halo screening. As shown in Table 26.2 the number of colonies resulting in a decreased halo formation is about 10% of the initial SUI^R strains even without UV-mutagenesis. In previous studies using milk halo screening after UV-mutagenesis only 0.1% of the surviving spores resulted in a reduced milk halo. With UV-mutagenesis prior to selection with the suicide substrate the efficiency of isolating protease-deficient mutants can be even further increased to over 50% [Punt et al., unpublished result].

In Table 26.3 the analysis of a number of available and newly selected protease mutant strains is shown. Interestingly, also a mutant with a deficient intracellular protease gene (pepe), which results in no significant decrease of extracellular protease activity [30], can be selected with this method. From Table 26.3 it is also clear that, as is the case with virtually every method, not every type of protease mutant can be selected in this way. For example, a mutant lacking the major protease gene (pepA) in A. niger, which results in a residual extracellular protease activity of less than 20% [14,30], had no higher resistance against the suicide substrate than the wild-type strain. Remarkably, with this approach also mutants with enhanced protease activity were selected [Punt et al., unpublished results].

### 26.3 Fermentation Conditions

Strain improvement has proven to be a very useful tool for reducing the proteolytic degradation of especially heterologous proteins produced in the Aspergillus host strain. However, the large number of (extracellular) proteases able to degrade these heterologous proteins and the varying susceptibility of the produced heterologous proteins for the different proteases [15,48] makes one single (permanent) solution
of the problem impossible. Therefore, also development of fermentation conditions repressing protease production can be an additional way to improve heterologous protein production. Although numerous empirical approaches have been followed to address the protease issue, only very few systematic studies have been performed. From these studies three environmental parameters have emerged which have been studied in somewhat more detail, that is, ambient pH, carbon catabolite control and nitrogen metabolite control.

26.3.1 pH Regulation

Ambient pH was shown to be an environmental parameter greatly influencing the expression of proteases. Controlled fermentations with \textit{A. niger} at pH 4 or pH 5 resulted in a significant decrease of protease activity at higher pH. When cultured at pH 6, protease activity was even further decreased [Braaksma et al., unpublished results]. Culture pH was also suggested to be a key player during the production of recombinant GFP by \textit{A. niger} and \textit{A. sojae} [24,49]. GFP excreted by the recombinant \textit{A. niger} strain was rapidly degraded, whereas in \textit{A. sojae} significant amounts of extracellular GFP could be detected. Acidification of the culture medium of \textit{A. niger} was suggested to be the cause for proteolytic degradation of GFP, as under identical conditions \textit{A. sojae} did not significantly acidify. Maintaining the pH at 6 during

### TABLE 26.2

Efficiency of Isolation of Protease-Deficient Mutants by Spontaneous Resistance to Suicide Substrate (SUI) Compared to UV Mutagenesis

<table>
<thead>
<tr>
<th>Strain</th>
<th>Number of Initial Spores</th>
<th>Number of Colonies SUI</th>
<th>Rescreen SUI</th>
<th>Reduced Milk Halo</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Aspergillus sp. section Nigri strain A}</td>
<td>4.10^4</td>
<td>590</td>
<td>160/590</td>
<td>45/160</td>
</tr>
<tr>
<td>\textit{Aspergillus sp. section Nigri strain B}</td>
<td>4.10^4</td>
<td>200</td>
<td>85/200</td>
<td>20/85</td>
</tr>
</tbody>
</table>

### UV mutagenesis of \textit{A. niger} [14] and \textit{A. nidulans} [16]

<table>
<thead>
<tr>
<th>Strain</th>
<th>Number of Initial Spores</th>
<th>Survival Rate After UV Mutagenesis</th>
<th>Number of Spores Screened for Reduced Milk Halo</th>
<th>Reduced Milk Halo</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{A. niger}</td>
<td>5.10^4–1.10^6</td>
<td>10–20%</td>
<td>1.10^4</td>
<td>7/1.10^4</td>
</tr>
<tr>
<td>\textit{A. nidulans}</td>
<td>2.5.10^4–2.5.10^6</td>
<td>1–10%</td>
<td>2.5.10^4</td>
<td>29/2.5.10^4</td>
</tr>
</tbody>
</table>

### TABLE 26.3

Protease Mutants Show Higher Resistance to the Suicide Substrate than WT Strains^a

<table>
<thead>
<tr>
<th>Species</th>
<th>SUI (mg/l)</th>
<th>Residual Protease Activity (Intracellular)</th>
<th>Residual Protease Activity (Extracellular)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{A. niger} WT</td>
<td>0</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>\textit{A. niger} pepA</td>
<td>100</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>\textit{A. niger} pepE</td>
<td>200</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>\textit{A. niger} prtT</td>
<td>300</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>\textit{A. niger} prtT/phmA^b</td>
<td>400</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>\textit{A. niger} prtT/phmA^c</td>
<td>500</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

^a[Punt et al., unpublished results].

^bThe \textit{A. niger} \textit{prtT/phmA} mutant is a derivative of \textit{A. niger} \textit{prtT} that does not acidify its medium.
the production of GFP with A. niger resulted in a tenfold increase of GFP levels compared to a culture controlled at pH 3 [4]. This increase was due to reduced degradation of GFP by proteases. Also production of the human cytokine interleukin 6 (II-6) in a protease deficient strain and a derivative of that strain that did not acidify resulted in improved yield and stability of II-6 in the nonacidifying host strain [50].

The genes encoding the two major extracellular proteases of A. niger, pepA and pepB, were not expressed under alkaline conditions [51]. On the other hand, the transcript levels of the major alkaline protease gene prtA produced by A. nidulans was elevated under alkaline conditions [36], although this was not confirmed by similar experiments conducted by Katz et al. [16], where nitrogen starvation appeared to override the repression of prtA by low culture pH [10]. From these results we conclude that ambient pH is a regulator of protease expression. In A. nidulans pH regulation is mediated mainly by seven genes, pacC, palA, palB, palC, palF, palH, and palI, where pacC plays the key role in the regulation of gene expression by ambient pH [30]. The products of the pal genes transduce a signal able to trigger the PacC into an active form. This active PacC is able to activate the expression of alkali-expressed genes (including prtA) and to inhibit the expression of acid-expressed genes [35]. Homologous of the pacC gene and the pal genes have been identified in other aspergilli, such as A. niger [52], A. fumigatus [53], and A. oryzae as well as all major groups of ascomycetes [35]. The involvement of pH control in extracellular protease production was further confirmed by analysis of protease expression in PacC mutants of A. nidulans and A. niger [36,44]. However, the expression of three vacuolar proteases in A. niger is not regulated by PacC, which may also be the case with intracellular proteases of other aspergilli [44].

26.3.2 Carbon Catabolite Control

Growth on glucose or other favored carbon sources prevents the synthesis of enzymes involved in the utilization of other substrates, like polysaccharides [37]. This seems to apply for fungal extracellular proteases as well. Unfortunately, literature about the effect of carbon source on protease production by aspergilli is scattered and in addition often rather outdated. However, a few examples of the repressing effect of glucose and other carbon sources on the levels of excreted proteases have been described. When mycelia from A. nidulans were transferred to a medium without carbon source, extracellular proteases were abundantly produced. When mycelia were transferred to medium with glucose, lactose, galactose, or glycerol, protease production was severely repressed [54]. Similarly, transferring experiments with A. oryzae showed a strong decrease of protease production when mycelia were transferred to medium with casein and glucose compared to medium with casein only [55].

The expression of the two extracellular proteases pepA and pepB of A. niger was studied in the presence of various carbon sources [51]. When cells were transferred to medium supplemented with glucose, expression of both protease genes was repressed. In the presence of the less favorable carbon source glycerol the pepA gene was derepressed and in medium without carbon source pepA and pepB were both strongly derepressed. Thus, protease expression is clearly affected by glucose (or carbon catabolite) repression. Repression may be caused by various other carbon sources, but glucose is suspected to be the most repressive. The repressor protein CreA plays a major role in carbon repression. CreA inhibits transcription of many target genes by binding to specific sequences in the promoter of these genes [41]. The gene encoding this protein has been identified in several Aspergillus species, like A. nidulans [56], A. oryzae [57], and A. niger [58]. With Northern blot analysis, protease expression in creA mutants of A. niger gave clear evidence for the involvement of carbon catabolite control [44]. Similarly, this was suggested by the fact that two of the isolated A. nidulans mutants, xprF and xprG, which carry a mutation in a hexokinase-like protein and an acid phosphatase, respectively, are thought to be involved in carbon catabolite repression and maybe also nitrogen, sulfur, and phosphate regulation [54,59].

26.3.3 Nitrogen Metabolite Control

Similar as for the repression by glucose, the presence of preferred nitrogen sources such as ammonium suppress the production of enzymes, such as extracellular proteases, for utilizing other nitrogen sources [37]. For example, high concentrations of the preferred nitrogen source ammonium resulted in increased concentrations of bioactive t-PA produced by A. niger, which was suspected to be due to less degradation
of this heterologous protein [3,60]. Extracellular protease levels of *A. nidulans* were significantly lower in a growth medium with ammonium compared to a nitrogen-free medium [10]. The influence of nitrogen source on the expression of the *pepA* and *pepB* gene in *A. niger* was investigated by transferring cells to medium with and without ammonium. Cells grown with ammonia showed very low levels of both protease transcripts, whereas the levels of mRNA were much higher when cells were grown without ammonia [51].

The gene *areA* has been implicated in mediating the nitrogen metabolite control regulatory mechanism and it has been extensively studied in *A. nidulans* [61]. The *areA* gene encodes a protein that activates transcription of many target genes by binding to specific sequences in the promoter of these genes. Homologous of this gene have also been identified with other *Aspergillus* species, such as *A. oryzae* [40] and *A. niger* [62].

A study with an *A. niger* wild-type strain and several different *areA* mutants (obtained by UV-mutagenesis and selection on chlorate plates) demonstrated that three intracellular protease genes were not controlled by AreA, because both wild-type and *areA* mutants showed unaltered expression of these three genes [44]. The same study showed that three extracellular proteases were apparently regulated by AreA. However, the expression of the corresponding extracellular protease genes was not modulated in the same way in the different *areA* mutants, but depended on the combination of the protease gene and the particular *areA* mutation.

### 26.3.4 Sulfur and Phosphorus Metabolite Repression

Several decades ago the first studies on the effect of phosphorus and sulfur sources on protease expression in aspergilli were reported, but hardly any articles have been published on this subject since [63–65]. Today, little is known about sulfur and phosphorus metabolite repression in aspergilli and putative involvement in protease regulation. However, more recently a strong effect of sulfur limitation on the increase of protease activity for *A. nidulans* has been described [10]. In addition, expression analyses of *prtA*, encoding the major extracellular protease in *A. nidulans*, showed a high transcript level when mycelia was transferred to sulfur-free medium [16,66].

Although the regulatory factors involved in sulfur metabolite repression are known [67,68], no information is available regarding protease gene expression. The regulatory factors involved in phosphorus metabolite repression are yet unknown. Identification of the role of these factors may help for a better understanding of the overall protease regulation.

### 26.3.5 Induction of Protease by Protein

The fact that in the presence of protein the production of proteases is stimulated has been applied for years in the production of extracellular proteases by the use of complex nitrogen and/or carbon sources [55,69–71].

However, the opposite effect has also been described. Extracellular GFP could not be detected when the *A. niger* host strain was cultured on defined medium [49]. When modified soya milk medium was used, fluorescence could be detected in the culture medium. The authors indicate that this was probably not due to a repressive effect of the soya milk protein, but due to the natural protease inhibitors that are present in the soya milk medium and the fact that the ambient pH can be maintained for longer than with defined medium at a value which limits protease induction. Another explanation is that the abundant availability of substrate for the proteases delayed the degradation of GFP.

The *A. niger pepA* and *pepB* protease genes were induced when mycelia was transferred to medium with elastin [41]. Medium containing glucose next to elastin repressed expression of both proteases. Comparable experiments by Jarai and Buxton [51] showed a somewhat different picture, as *A. niger* expressed *pepA* and *pepB* in the presence of glucose if BSA was also present. When additional ammonia or urea was supplemented both protease genes were repressed. These results suggest that induction by the presence of extracellular protein plays only a secondary role in the regulation of extracellular proteases. As for the sulfur and phosphorus regulation mechanisms little is known about the mechanism of specific induction of protease gene expression by external addition of proteins. It is also possible that protein itself is not an inducer, but that the added protein or its peptide degradation products, being a complex carbon and nitrogen source all in one, play a role in the wide domain regulation mechanisms of nitrogen metabolite and carbon catabolite control.
26.3.6 Bioprocess Engineering

Affecting protease production by the means of bioprocess engineering has also proved to be a successful means of controlling extracellular protease activity. However, again very little has been published on the subject. Immobilization of the cells of *A. niger* to materials like a metal-coated pad or Celite beads reduced secretion of extracellular protease and increased the secretion of glucoamylase [72,73]. Manipulating the morphology of *A. niger* by means of inoculum levels (concentration of spores) or inoculum type (vegetative or spores) was also shown to affect protease levels [74,75]. Growth of the mycelium in the form of (large) pellets resulted in lower specific protease activities and increased protein production compared with a filamentous morphology. Morphology clearly affects protease secretion as well as protein production, but the exact mechanism needs further investigation [76].

The effect of the bioprocess parameters agitation intensity, dissolved oxygen tension as well as initial glucose and yeast extract concentration on protease and heterologous protein production has been studied in *A. niger* [77]. However, altogether these studies should be considered as exploratory no systematic analysis was performed.

26.3.7 System Biology Approach

Strain development and optimization of fermentation conditions have improved the production of (heterologous) proteins by aspergilli to a considerable extent. However, the problem of proteases has in most cases been approached by trial-and-error, without taking the interaction between strain development and improvement of fermentation conditions into account (e.g., the best mutant may not be the best producer on the medium previously optimized for a precursor strain). Furthermore, the mechanism of induction and repression of protease production is far from completely understood. A more integrated approach is, therefore, desirable to come to a better understanding of the issue and from this to a solution that is also more generally applicable.

Recently developed techniques such as (comparative) genomics, transcriptomics, proteomics, and metabolomics will very likely play a crucial role in understanding the proteolytic system of aspergilli. In addition to these -omics approaches we would also like to consider the role of the various physiological parameters involved in the fermentation process. These “physiomics” parameters such as pH, oxygenation, viscosity, agitation and so on, add a further layer of data to be included in a full systems biology approach to study the proteolytic system of aspergilli.

The first articles reporting application of genomics techniques for research of *Aspergillus* strains have been recently published. However, only one of them has dealt specifically with fungal proteolytic processes [8].

With the complete genome sequences of several *Aspergillus* strains open to the public [5–7] and more to be expected in the near future, comparing the presence or absence of specific genes or gene clusters is a first indication of hitherto undiscovered pathways [78].

Transcriptomics is the most established of these genomics techniques. However, only a few studies on gene expression profiling for *Aspergillus* species have been reported at this time (*A. nidulans* [79–81], *A. oryzae* [82], *A. flavus* [83]). Most of these studies use microarrays based on cDNA, so only covering that part of the genome that is expressed under certain conditions. With the genome sequences available microarrays covering the complete genome can be created, although they are much more expensive. Mogensen et al. [79] constructed a microarray with probes for only one-third of the annotated genes of *A. nidulans* (as annotated by the Broad Institute), selecting only those genes with value-added annotation. They studied the gene expression of *A. nidulans* wild-type and a creA mutant during growth on glucose or ethanol. Analysis of the supplementary data associated with the article showed that from the approximately 70 hypothetical protease genes induced in the study, only three were significantly affected. A hypothetical aspartyl protease was expressed at higher levels in the wild type compared to the creA mutant when grown at glucose, while two metalloproteases were expressed at lower levels in the wild type compared to the creA mutant.

The method for the identification of all proteins in complex mixtures is proteomic analysis. Initial approaches involved studying the proteins to be separated by one-dimensional (1D) SDS-PAGE. This
approach was also used to identify proteins produced by *A. oryzae* on solid cereal substrates [84]. This resulted in about 10 proteins, which were identified by N-terminus sequencing. With the development of 2D gel electrophoresis, often coupled to mass spectrometry in order to identify the proteins, proteomic analysis has become a very powerful method for the identification of proteins in complex mixtures, like culture samples or cell extracts, and can also be used to study the alteration of protein production under different environmental conditions. At this time, few studies applying this functional genomics tool have been published for *Aspergillus*. For *A. flavus*, an extensive proteomic study was conducted to identify the secreted proteins during growth on the flavonoid rutin and nonrutin culture medium [85]. Over 100 different protein spots were found, however only 22 spots were identified. On rutin-containing medium an alkaline protease, which could be produced to provide the fungus with sufficient energy for growth, appeared to be the most abundant protein based on the intensity of the spots. For *A. fumigatus* the intracellular proteins were analyzed on 2D gels in combination with mass spectrometry [86]. This resulted in a total of 180 spots, from which 50 distinct proteins (in 65 individual spots) could be identified. Proteomic analysis for extracellular proteins from *A. oryzae* grown under submerged and solid-state culture conditions revealed that the cultivation environment greatly affects the types of secreted proteins [87].

One of the more recent functional genomic tools is metabolomics, the analysis of all intracellular and extracellular metabolites. Although already in the mid-1990s, a method to extract intermediary metabolites from *A. niger* has been described by Ruijter and Visser [88], very little has been published on metabolomics involving *Aspergillus* species since. However, since then analytical platforms for metabolite detection have gone through major developments. While Ruijter and Visser used an automated spectrophotometer to analyze the metabolites, nowadays most quantitative strategies combine a separation technique (e.g., capillary electrophoresis, liquid or gas chromatography) with MS or NMR-based detection, making detection of a large dynamic range of metabolites possible, both known and unknown [89].

All functional genomics tools are still under development, with identification of expressed genes or proteins as the major challenge for transcriptomics and proteomics, respectively. But for all genomic tools extracting relevant biological information from the overwhelming amount of data resulting from these tools is perhaps the biggest challenge. Focusing on the biggest changes in gene expression or protein or metabolite concentration does not automatically lead to the identification of the most important parameter in a biological process [90]. The choice for a data pretreatment method and a data analysis method greatly affects the outcome [91]. The final goal will be to combine the results distilled from the high-throughput functional genomics methods with information from small-scale studies focusing on particular cellular functions and systems in order to construct a biological network of all protein and genetic interactions. A comprehensive collection of experimentally observed interactions has been put together for the best-studied eukaryote, the budding yeast *Saccharomyces cerevisiae*, but it is suggested that there are probably many more interactions to be discovered [92]. For *Aspergillus*, the study of complex biological networks, among which are also the proteolytic systems, is still in its infancy and will provide the scientific community with a huge challenge on the road to a more complete understanding of this type of organism.

### 26.4 Conclusion

Based on the results described in this chapter, it is clear that understanding of the regulation of the highly complex fungal proteolytic system, aimed at improvement of the use of the fungal cell factory, will require a multidisciplinary systems biology approach linking the various layers from gene to phenotype. Current research from our Group within the framework of the Kluyver consortium is focused at this approach.

### Acknowledgments

The Kluyver Centre for Genomics of Industrial Fermentation, which is supported by the Netherlands Genomics Initiative (NROG), is gratefully acknowledged for the financial support of this research.
References

38. Christensen, T. and Hynes, M.J., Fungus wherein the *areA* gene has been modified and an *areA* gene from *Aspergillus oryzae*, US patent US6025185, 2000.
Aspergillus as a Cell Factory for Protein Production

90. van den Berg, R.A. et al., Centering, scaling, and transformations: improving the biological information content of metabolomics data, BMC Genomics, 7, 142, 2006.
94. Christensen, T. and Lehmbbeck, J., Fungus wherein the areA, pepC and/or pepE genes have been inactivated, US patent US6013452, 2000.
27

Mycotoxin Production and Prevention of Aflatoxin Contamination in Food and Feed

Jiujiang Yu, Gary A. Payne, Bruce C. Campbell, Baozhu Guo, Thomas E. Cleveland, Jane F. Robens, Nancy P. Keller, Joan W. Bennett, and William C. Nierman

CONTENTS
27.1 Introduction ........................................................................................................... 457
27.2 Significance of Mycotoxin Contamination ................................................................. 458
  27.2.1 Mycotoxins and Aflatoxins .............................................................................. 458
  27.2.2 Health Risk, Food Safety, and Economic Impact of Aflatoxin Contamination .......... 458
27.3 Aflatoxin Biosynthesis ............................................................................................. 459
  27.3.1 Aflatoxin Biosynthetic Pathways and Enzymes .............................................. 459
  27.3.2 Aflatoxin Biosynthetic Pathway Genes .......................................................... 460
  27.3.3 Aflatoxin Biosynthetic Pathway Gene Clusters .............................................. 460
  27.3.4 Regulation of Aflatoxin Biosynthesis ............................................................. 461
27.4 Factors Affecting Aflatoxin Formation ...................................................................... 461
  27.4.1 Nutritional Factors .......................................................................................... 461
  27.4.2 Environmental Factors .................................................................................... 462
  27.4.3 Developmental Factors .................................................................................... 462
  27.4.4 Stress Factors and Antioxidation .............................................................. 462
27.5 Strategies in Preventing Aflatoxin Contamination .................................................... 463
  27.5.1 Biological Control by Nontoxicogenic Aspergillus flavus Strain ...................... 463
  27.5.2 Host Resistance Through Crop Breeding and Genetic Engineering .............. 464
  27.5.3 Crop Genomics for Improving Host Resistance in Corn, Cotton, and Peanuts .... 464
  27.5.4 Gene Profiling Through Aspergillus flavus Genomics ...................................... 465
27.6 Summary and Prospects ......................................................................................... 466
References ..................................................................................................................... 466

27.1 Introduction

Aflatoxins are the most prominent groups of mycotoxins. They are known as the most toxic and most potent carcinogens naturally produced by fungal molds, mainly Aspergillus flavus and Aspergillus parasiticus. Over 40 years of research and investigations generated a wealth of publications on fungal biology, toxicology, and toxin biosynthesis. In this review, only the most pertinent aspects on aflatoxin biosynthesis and prevention are discussed. In deciphering the aflatoxin biosynthetic pathway, significant progress has been made in the last decade. A complete aflatoxin biosynthetic pathway gene cluster consisting of 29 genes has been cloned. Details of these genes and their encoded enzymes as well as the regulation of gene expression have been reported. Sterigmatocystin (ST) or dihydrosterigmatocystin...
The Aspergilli

(DHST), the penultimate precursor of aflatoxins, is produced by several species including *A. versicolor* and *A. nidulans*. ST and DHST are toxic and carcinogenic. They share strikingly similar common biochemical pathway, homologous genes, and regulatory mechanism to aflatoxin synthesis in *A. flavus*. The ST biosynthetic pathway and genes are discussed as well.

### 27.2 Significance of Mycotoxin Contamination

#### 27.2.1 Mycotoxins and Aflatoxins

Mycotoxins are commonly found in foods and feeds all over the world. It has been estimated that a quarter of the world’s crops are contaminated to some extent with mycotoxins. Mycotoxins rank as the most important noninfectious, chronic dietary risk factor, higher than synthetic contaminants, plant toxins, food additives, or pesticide residues. Mycotoxins are low molecular weight secondary metabolites produced by filamentous fungi that display various degrees of toxicity to vertebrates, invertebrates, plants, and microorganisms. Research on mycotoxins gained worldwide attention since the notorious “Turkey X disease” in 1960 near London, England, that killed approximately 100,000 turkey poults. This mysterious disease was later found to be caused by feeding peanut (groundnut) meal infested by *A. flavus* and so the toxins were named “aflatoxin” for *A. flavus* toxins.

Since aflatoxins are the most toxic and carcinogenic mycotoxins, considerable attention has been given to their study. Aflatoxins are difuranocoumarin derivatives by chemical definition. Aflatoxins B₁, B₂, G₁, and G₂ (AFB₁, AFG₁, AFB₂, and AFG₂) are the four major aflatoxins based on their blue or green fluorescence under ultraviolet light, and relative chromatographic mobility by silica gel thin layer chromatographic separation. In addition to the four major aflatoxins, about a dozen other aflatoxins (e.g., P₁, Q₁, B₂a, G₂a) were described. Aflatoxin M₁ is secreted in milk when cows metabolize aflatoxin B₁ from feed into a hydroxylated derivative. *A. flavus* produces aflatoxins B₁ and B₂, cyclopiazonic acid, kojic acid, β-nitropropionic acid, aspertoxin, aflatem, and aspergillic acid. *A. parasiticus* produces the four aflatoxins: B₁, B₂, G₁, and G₂. Some other *Aspergillus* species are identified to produce aflatoxins such as *A. nomius*, *A. pseudotamarii*, *A. bombycis*, *A. ochraceoroseus*, and *Emericella venezuelensis* (Klich, unpublished data).

#### 27.2.2 Health Risk, Food Safety, and Economic Impact of Aflatoxin Contamination

*A. flavus* can cause diseases in animals and human beings. The diseases caused by fungal invasion into animal or human hosts are collectively called “mycoses,” while the diseases or symptoms caused by the toxic fungal metabolites are collectively called “mycotoxicoses.” The diseases caused by the *Aspergillus* species including *A. flavus*, is called “aspergillosis.” *A. flavus* is the second leading cause of invasive and noninvasive aspergillosis in humans and animals next to *A. fumigatus*. Due to the spread of AIDS and organ transplantation, the incidence of aspergillosis in immunocompromised patients is rising. There is no effective antifungal drug available on the market to control fungal growth in human patients and so invasive aspergillosis is often fatal.

*A. flavus* is a weak opportunistic plant pathogen that causes diseases of many agricultural crops such as maize (corn), cotton, groundnuts (peanuts), and tree nuts. Few plant pathogenic fungi such as *A. flavus* have such a broad host range that it can attack seeds of both monocots and dicots, and to infect seeds produced both above and below the ground. It infects corn ears, cotton bolls, and peanut pods after insect or mechanical damages occurs. The postharvest aflatoxin contamination is sometime problematic if food grain storage is poorly managed.

Aflatoxin is associated with both toxicity and carcinogenicity in human and animal populations. Aflatoxicosis is poisoning resulting from ingestion of moderate to high levels of aflatoxins in contaminated food or feed. Toxicological studies demonstrate that ducklings, hamsters, rats, trout, rabbits, and a number of other vertebrates are all susceptible to aflatoxin poisoning. Acute aflatoxicosis results in rapid
progressive jaundice, edema of the limbs, pain, vomiting, necrosis, cirrhosis, or in severe cases, acute liver failure and death.\textsuperscript{27–30}

Outbreaks of acute aflatoxicosis from contaminated food in humans have been documented in Kenya, India, Malaysia, and Thailand (Council for Agriculture Science and Technology).\textsuperscript{31} However, the most widely spread outbreak of aflatoxicosis in humans occurred in more than 150 villages in western India in 1974 where 397 persons were affected and 108 persons died.\textsuperscript{32} As recent as in July 2004, an incident of aflatoxin poisoning in Kenya had occurred involving 317 cases and 125 deaths due to consumption of aflatoxin contaminated maize (corn), the largest and most severe outbreaks of acute aflatoxicosis documented worldwide.\textsuperscript{29,30}

Chronic aflatoxicosis results in cancer, immune suppression, and other “slow” pathological conditions. The liver is the primary target organ when mammalian species are fed with aflatoxins. Cytochrome P450 enzymes in the liver convert aflatoxins to the reactive 8,9-epoxide form, which is capable of binding to both DNA and proteins.\textsuperscript{26} Aflatoxin B\textsubscript{1}-DNA adducts can result in the GC to TA transversions in the \textit{p53} tumor suppressor gene at codon 249. Inactivation of the \textit{p53} tumor suppressor gene is the culprit in the development of primary liver cancer.\textsuperscript{33,34}

In developing countries, food safety is the major problem where detection and decontamination policies are impractical. In those countries where populations are facing starvation routine ingestion of aflatoxin-contaminated food may occur. Statistics indicates that worldwide liver cancer incidence rates are 2 to 10 times higher in developing countries than in developed countries.\textsuperscript{35} This is because, in the developed countries the toxic contaminants are monitored and regulated. The maximum allowable amount of aflatoxin in food and feed for human consumption and for livestocks has been mandated by laws. A guideline of 20 parts aflatoxin per billion parts of food or feed substrate (ppb) is the maximum allowable limit imposed by the U.S. Food and Drug Administration. The European Union has a maximum level of 2 ppb for aflatoxin B\textsubscript{1} and 4 ppb for total aflatoxins.\textsuperscript{36} The aflatoxin contaminated harvest is destroyed resulting yearly in billion dollar losses worldwide. Aflatoxin contamination is a chronic problem in some parts of Southern United States\textsuperscript{37} such as in Southern cotton belt, mid-South corn, and Southeast peanut-farming regions.\textsuperscript{38} Severe outbreaks of aflatoxin contamination have occurred in the Midwest U.S. cornbelt in 1977, 1980, and 1988. The resulting economic losses are enormous.\textsuperscript{38} Thus, aflatoxin contamination is not only a serious food safety concern, but has significant economic impact in agriculture worldwide.

### 27.3 Aflatoxin Biosynthesis

#### 27.3.1 Aflatoxin Biosynthetic Pathways and Enzymes

The aflatoxin pathway is one of the best-studied pathways of fungal secondary metabolism. Within the last decade, the major biochemical pathway steps have been elucidated and the chemical structures of these aflatoxin intermediates have been defined.\textsuperscript{1,5,39} At least 23 enzymatic reactions are estimated to be involved in aflatoxin formation. No less than 15 structurally defined aflatoxin intermediates have been identified in the aflatoxin/ST biosynthetic pathway.\textsuperscript{2,3,9}

Studies demonstrated that aflatoxins are synthesized from malonyl CoA, first with the formation of hexanoyl CoA, followed by formation of a decaketide anthraquinone.\textsuperscript{39,40} There are two fatty acid synthases (FAS) and a polyketide synthase (PKS) involved in the synthesis of polyketide from acetyl CoA.\textsuperscript{41} Norsolorinic acid (NOR) is the first stable aflatoxin intermediate identified in the pathway.\textsuperscript{42} Aflatoxins are formed after a series of highly organized oxidation–reduction reactions.\textsuperscript{4,40,43,44} The general accepted aflatoxin biosynthetic pathway scheme is: hexanoyl CoA precursor $\rightarrow$ norsolorinic acid, NOR $\rightarrow$ averantin, AVN $\rightarrow$ hydroxyaverantin, HAVN $\rightarrow$ Oxoaverantin, OAVN $\rightarrow$ averufin, AVF $\rightarrow$ hydroxyversicolorone, HVN $\rightarrow$ versicalonal hemiacetal acetate, VHA $\rightarrow$ versicalonal, VAL $\rightarrow$ versicolorin B, VERB $\rightarrow$ versicolorin A, VERA $\rightarrow$ demethyl-sterigmatocystin, DMST $\rightarrow$ sterigmatocystin, ST $\rightarrow$ O-methylsterigmatocystin, OMST $\rightarrow$ aflatoxin B\textsubscript{1}, AFB\textsubscript{1}, and aflatoxin G\textsubscript{1}, AFG\textsubscript{1}. After the VHA step, there is a branch point in the pathway that leads to aflatoxins B\textsubscript{2} and G\textsubscript{2} formation, AFB\textsubscript{2} and AFG\textsubscript{2}.\textsuperscript{45,46}
27.3.2 Aflatoxin Biosynthetic Pathway Genes

A total of 29 genes or open reading frame (ORF) involved in aflatoxin formation have been identified, cloned and characterized. Several key genes for aflatoxin biosynthesis are milestone discoveries. The first aflatoxin pathway gene identified was aflD (nor-1) encoding for a ketoreductase in A. parasiticus for the conversion of NOR to averantin (AVN). Disruption or deletion of the aflD (nor-1) gene leads to the accumulation of brick-red pigment and loses all aflatoxins and their intermediates in the fungus. The second important gene cloned was aflM (ver-1), encoding for a ketoreductase, which is required for the conversion of versicolorin A (VERA) to demethylsterigmatocystin (DMST) and versicolorin B (VERB) to demethylhydrosterigmatocystin (DMDHST) in A. parasiticus.

In early step of aflatoxin biosynthesis, two large genes (7.5-kb transcripts), aflB (fas-1) and aflA (fas-2), encoding beta (FASβ) and alpha-subunit (FASα) of FAS, respectively, were identified. The third important large gene in aflatoxin synthesis is the aflC (pksA) gene encoding a PKS for the synthesis of polyketide. Disruption of the aflC (pksA) gene produced no aflatoxin or aflatoxin intermediates. The predicted amino acid sequences of these PKS contains four typical conserved domains common to other known PKS proteins: \(\beta\)-ketoacyl synthase (KS), acyltransferase (AT), acyl carrier protein (ACP), and thioesterase (TE). The aflA, aflB, aflC genes are shown to be directly involved in the backbone formation for the conversion from acetate to NOR in aflatoxin synthesis. In later step of aflatoxin biosynthesis, an important gene named aflP (omtA) encoding an O-methyltransferase for the conversion of ST to OMST and DMST to dihydro-O-methylsterigmatocystin (DHOMST) was cloned by antibody screening of a cDNA expression library from A. parasiticus.

The gene and enzyme for G-group aflatoxin formation has been a long-time mystery for many molecular biologists in elucidating aflatoxin pathway until the cloning of aflQ (ordA) and aflU (cypA). There are two separate pathways leading to B-Group (AFB1 and AFB2) and G-Group (AFG1 and AFG2) aflatoxin formation. A gene, named aflQ (ordA) encoding a cytochrome P-450 monoxygenase, was demonstrated to be responsible for the conversion of O-methylsterigmatocystin (OMST) to AFB1 and AFG1, and demethylhydrosterigmatocystin (DMDHST) to AFB2 and AFG2 in A. parasiticus and in A. flavus. Expression and substrate feeding in yeast system demonstrated that an additional enzyme is required for the G-group aflatoxin (AFG1 and AFG2) formation. Functional studies demonstrated that the aflU (cypA) gene in A. parasiticus, encoding a cytochrome P450 monoxygenase, is responsible for the conversion from OMST to AFG1 and DHOMST to AFG2. A partial deletion of this gene results in loss of G-group aflatoxin production in A. flavus.

27.3.3 Aflatoxin Biosynthetic Pathway Gene Clusters

The initial evidence indicating clustering of aflatoxin pathway genes was the fact that the aflID (nor-1) and aflIM (ver-1) genes were linked with the regulatory gene aflR in a cosmid clone. The aflatoxin pathway gene cluster was established when nine cloned genes including aflID (nor-1), aflR, aflM (ver-1), and aflP (omtA), were mapped to overlapping cosmids clones in A. parasiticus and A. flavus. The completed aflatoxin pathway gene cluster was established when a 82 kb DNA sequence harboring a total of 29 aflatoxin biosynthetic pathway genes (or ORFs) and 4 sugar utilization genes was reported.

Clustering of genes allows regulatory elements to be shared. A primary advantage of gene clustering may be for the purpose of coordinated gene expression. Gene complementation experiments performed in this laboratory demonstrated that the aflatoxin pathway genes are expressed properly only when they are targeted into the gene cluster. Gene clustering may influence gene expression and regulation through modulation of localized chromatin structure.

In A. parasiticus, a partial duplicated aflatoxin gene cluster consisting of seven duplicated genes was identified and characterized. These duplicated genes are named aflR2, aflJ2, adhA2, estA2, norA2, ver1B, and omtB2, respectively with the number “2” indicating second copy. Due possibly to the chromosome location, the genes within this partial duplicated cluster, were likely nonfunctional under normal conditions though some of the gene sequences are intact.
27.3.4 Regulation of Aflatoxin Biosynthesis

There exists a positive regulatory gene, aflR (originally named afl-2 and apa-2) in both the aflatoxin and ST gene clusters, for activating toxin pathway gene transcription.\(^6^4,6^5\) Disruption of aflR gene leads to loss of aflatoxin pathway gene expression and aflatoxin production. Introduction of an additional copy of the aflR leads to the overproduction of aflatoxin biosynthetic intermediates.\(^6^6\) The fact that aflatoxin and ST biosynthetic pathway genes are tightly compacted on a single chromosome within a 75 kb DNA region in both *A. parasiticus* and *A. flavus* as well as in *A. nidulans*\(^6^1,6^7\) leads to the presumption of gene expression in concert in the genome. The aflR gene, coding for a sequence-specific zinc binuclear DNA-binding protein, is required for transcriptional activation of most, if not all, the aflatoxin pathway genes.\(^6^6,6^8–7^0\) The AflR protein has major domains typical of fungal and yeast Gal4-type transcription factors\(^6^6;\) an N-terminal cysteine-rich stretch, \((\text{Cys}_6-\text{Zn}_2)\) DNA-binding domain,\(^6^6,7^0\) an arginine-rich (RRARK) nuclear localization domain; and a transcription activation domain in the C-terminus.\(^6^6,6^9\) The aflatoxin pathway gene transcription can be activated when the AflR protein binds to the palindromic sequence \(5\prime-\text{TCGN5CGA}-3\prime\) (also called AflR-binding motif) in the promoter region of the structural genes\(^7^1–7^3\) in *A. parasiticus*, *A. flavus*, and *A. nidulans*. *A. sojae*, a nontoxigenic strain used in industrial fermentations, was found to contain a defective aflR transcription activation domain due to early termination of 62 amino acids from its C-terminus.\(^7^4,7^5\) Thus, with the absence of the functional regulatory protein, no induction of aflatoxin can occur in this food grade *Aspergillus*.

In the aflatoxin gene cluster adjacent to the aflR gene, a divergently transcribed gene, aflS (aflJ), was also found to be involved in the regulation of transcription.\(^7^5,7^7\) The AFLJ protein binds to the carboxy terminal region of AflR and may affect AflR activity.\(^7^7\) Disruption of aflS in *A. flavus* resulted in a failure to produce any aflatoxin pathway metabolites.\(^7^9\) It was also found that a transcription factor required for nitrate assimilations, AreA,\(^7^8\) bound to sites near the aflS (aflJ) transcription start site in the aflR-aflS intergenic region, suggesting that aflS (aflJ) expression could be mediated by nitrogen source via the action of AreA.

Most recently, Professor Keller’s group identified a new gene named laeA, for lack of aflR expression.\(^7^9,8^0\) Interruption of this laeA gene resulted in loss of not only the aflR gene expression for ST synthesis, but also the expression of the genes involved in penicillin biosynthesis in *A. nidulans* as well as the genes involved in gliotoxin biosynthesis in *A. fumigatus*.\(^8^1\) It is likely that the laeA gene is involved globally in the regulatory circuit of the secondary metabolites, aflatoxins, ST, penicillin, and gliotoxin in several fungal species.

27.4 Factors Affecting Aflatoxin Formation

Many biotic and abiotic factors, including nutritional and environmental factors, are known to affect aflatoxin production by toxigenic aspergilli. The molecular mechanisms for these effects are still unclear in spite of numerous studies.\(^1,8^2\) Some of these factors may affect expression of the aflatoxin regulatory gene, aflR, or alter the expression of globally acting transcription factors that respond to external signals.

27.4.1 Nutritional Factors

Nutritional factors such as carbon, nitrogen, amino acid, lipid, and trace elements have long been observed to affect aflatoxin production.\(^2,5^1,8^4\) The relationship of carbon source and aflatoxin formation has been well established.\(^1\) Simple sugars such as glucose, sucrose, maltose, but not peptone, sorbose, or lactose support aflatoxin formation. However, the mechanism of carbon source in the regulation of aflatoxin pathway gene expression is poorly understood. Nitrogen source affects aflatoxin formation in varying ways.\(^1\) Aflatoxin production levels are different when on nitrate than on nitrite medium. Amino acids could have opposite effect on aflatoxin production.\(^8^5\) Recent studies show that tryptophan inhibit aflatoxin formation while tyrosine spikes aflatoxin production in *A. flavus* (Wilkinson et al., unpublished data). Trace elements (metal ions) were also reported to affect aflatoxin pathway gene expression.\(^8^4,8^6\) Lipids have tremendous effects on
27.4.2 Environmental Factors

The external environmental factors such as temperature, pH, water activity (drought stress), and other stress factors, have been shown to affect aflatoxin production. Recent studies suggest that aflR transcription is responsive to a G-protein signaling cascade that is mediated by protein kinase A, such a signaling pathway may respond to the environmental effects on aflatoxin biosynthesis. Optimal aflatoxin production is observed at temperatures about 30°C (28°C to 35°C). As temperature increases to above 36°C, the fungus stops aflatoxin production though the aflatoxin pathway genes are expressed. Optimal aflatoxin production is closely related to pH changes. Aflatoxin biosynthesis in A. flavus occurs in acidic media, but is inhibited in alkaline media (Wilkinson et al., unpublished data). The presence of a putative PacC-binding site close to the aflR transcription start site may play some role in pH regulation on aflatoxin production and the PacC- and AreA-binding sites in the aflR-aflJ intergenic region are the potential evidence that gene expression is regulated by environmental signals (pH and nitrate).

27.4.3 Developmental Factors

Secondary metabolism is associated with fungal developmental processes such as sporulation and sclerotia formation. It was observed that the environmental conditions required for secondary metabolism and for sporulation are similar. It was also reported that the spore formation and secondary metabolite formation occur at about the same time. Mutants that are deficient in sporulation were unable to produce aflatoxins. Certain compounds in A. parasiticus that exhibit the ability to inhibit sporulation have also been shown to inhibit aflatoxin formation. Chemicals that inhibit polyamine biosynthesis in A. parasiticus and A. nidulans inhibit both sporulation and aflatoxin/ST biosynthesis. A critical advance in this regard was the recent finding that the regulation of sporulation and ST production is by a shared G-protein mediated growth pathway in A. nidulans. Mutations in A. nidulans flbA and fadA genes, early acting members of a G-protein signal transduction pathway, resulted in loss of ST gene expression, ST production, and sporulation. It has been demonstrated that this regulation is partially mediated through protein kinase A. This G-protein signaling pathway involving FadA in the regulation of aflatoxin production also exists in other aspergilli including A. parasiticus and A. flavus. Most recently, progress has been made in identifying putative ligands regulating both morphological developments with toxin formation. Reviewed in Brodhagen and Keller, a sophisticated lipid signaling program has been uncovered between host seed and pathogenic aspergilli where the fungi can induce the production of seed oxygenated fatty acids that, in turn, stimulate sporulation and regulate expression of toxin genes.

Chapter 7 by Jae-Hyuk Yu and Christophe d’Enfert discusses developmental regulation in great detail.

27.4.4 Stress Factors and Antioxidation

We now know that oxidative stress induces aflatoxin biosynthesis. Jayashree and Subramanyam were the first to report that oxidative stress induced aflatoxin biosynthesis in A. parasiticus. More recently, Kim et al. showed that treatment of A. flavus with tert-butyl hydroperoxide induced significant increases in aflatoxin production. Their research showed that aflatoxin production under standard test regimens progressively increases over a four-day period, peaking at day 5, and then declines. However, A. flavus treated with hydroperoxide clearly had >30% greater levels of aflatoxin by day 5 and by day 9 >100% greater levels compared to unstressed cohorts of A. flavus. Similar treatment of A. parasiticus also induce aflatoxin production.
Alternatively, hydrolyzable tannins significantly inhibit aflatoxin biosynthesis with the main antiaflatoxigenic constituent being gallic acid. Gallic acid reduces expression of genes within the aflatoxin biosynthetic cluster, but surprisingly not the aflatoxin pathway gene regulator, \(aflR\). From these results, it appears gallic acid disrupts signal transduction pathway(s) for aflatoxigenesis somewhere upstream of the gene cluster. When certain phenolics or other antioxidants, such as ascorbic acid, are added to oxidatively stressed \(A.\ flavus\), aflatoxin production significantly declines, with no effect on fungal growth. Microarray analysis of \(A.\ flavus\) treated with caffeic acid, another antioxidant that inhibits aflatoxigenesis, shows that the gene apparently involved in quelling the signal for aflatoxin production is \(ahpC2\), the gene currently annotated as alkyl hydroperoxide reductase (Kim et al., unpublished). Surprisingly, the caffeic acid treatment showed no notable effect on expression of \(laeA\), a gene encoding a global regulator for secondary metabolism in \(Aspergillus\). This indicates that aflatoxin production is possibly regulated independently of overall secondary metabolism or by a system other than the one involving \(laeA\). Inhibition of aflatoxin biosynthesis by \(A.\ parasiticus\) appears to be associated with activation of an \(hsf2\)-like transcription factor that triggers antioxidative enzyme production. The antiaflatoxigenic activity of antioxidants appears to attenuate the oxidative stress responses in aspergilli.

The evolutionary advantage resulting from biosynthesis of aflatoxins has never been established. It may be that the chief reason for the evolution of the aflatoxin biosynthetic pathway is to relieve oxidative stress. Indeed, the fact that oxidative stress induces its biosynthesis, and that the biosynthetic precursors of aflatoxins are phenolics, having predictable antioxidant activity, supports this conjecture.

### 27.5 Strategies in Preventing Aflatoxin Contamination

Though we have a good knowledge on aflatoxin biosynthetic pathway and pathway cluster genes, many important questions remain. Not all of the genes responsible for aflatoxin production in the pathway have been accounted for. The mechanism of aflatoxin production in response to environmental changes is poorly understood. The mechanism of global regulation that turns on \(aflR\) expression is not yet defined. To identify all of the genes responsible for aflatoxin formation and to understand the regulatory mechanisms of aflatoxin formation present a daunting task and are impossible to accomplish by traditional cloning techniques. Genomics of host crops and crop pathogens will provide vital clues for devising strategies in solving aflatoxin contamination of food and feed.

#### 27.5.1 Biological Control by Nontoxigenic \(A.\ flavus\) Strain

Applying nonaflatoxicogenic strains of \(A.\ flavus\) and \(A.\ parasiticus\) in the field is a very successful strategy to reduce aflatoxin contamination of preharvest crops. The nonaflatoxin-producing strains occupy the same niche as the natural toxigenic strains. In another words, they have the same ability to survive in the natural environment and possibly to out-compete toxin-producing fungal strains. The method of competitive exclusion has been successfully applied to cotton, peanut, and corn fields. The fungal strains selected as the competitive biological control agents must meet the two basic criteria: they must be genetically stable so as not to produce aflatoxins under all conditions; and secondly, they must be competitive in excluding the toxigenic strains in the wild. Dr. Peter Cotty studied the competitiveness of seven nontoxigenic \(A.\ flavus\) strains in cotton in greenhouse experiments by coinoculating them with representative toxigenic strains. All of the nontoxigenic strains showed a significant effect in reducing aflatoxin content in cotton seed. One of the strains, named AF36, showed highest survivability in green house and in cotton field tests and the largest reduction in aflatoxin level as a biological control strain in field applications approved by EPA. The aflatoxin level in the cottonseed from treated fields was \(<20\) ppb compared to \(>100\) ppb from the untreated field. EPA has approved treatment by the biocontrol method of up to 20,000 acres by 2002. This strategy has also been shown to be effective in peanut and corn.
27.5.2 Host Resistance Through Crop Breeding and Genetic Engineering

*Aspergillus flavus* and *A. parasiticus* occur in a wide range of commodities. These fungi can be found virtually everywhere in the world. Damage due to insects or environmental stress can enable the fungi to invade seeds where they thrive at high temperatures and drought, such as those frequently experienced in the South during the summer. Various approaches have been suggested for genetic control of preharvest aflatoxin contamination and include breeding and use of crops with resistance to insects, resistance to plant stress, and resistance based on agronomic traits. Developing resistant commercial crop varieties is the most effective and most economic strategy for reducing and eliminating aflatoxin contamination of preharvest crops. Consequently, several sources of resistance were identified and released for use by public and private breeders in corn\textsuperscript{115–117} and in peanut.\textsuperscript{118–121} Crop resistance to aflatoxin contamination can be achieved mainly by the three venues: (1) resistance to the fungal invasion, (2) inhibition of aflatoxin formation, and (3) resistance to insects. Genetic engineering has been used in the last decade to improve crop resistance to diseases and insects. A typical example is the commercialization of Bt (bacterium *Bacillus thuringiensis*) corn\textsuperscript{122,123} or Bt peanut.\textsuperscript{124} The transgenic crop produces a protein (termed a Cry protein because of its crystalline nature) that is toxic to certain insects. Transgenic Bt crops have been commercially available since the mid-1990s, and have shown reduced level of aflatoxin contamination.\textsuperscript{125–127}

27.5.3 Crop Genomics for Improving Host Resistance in Corn, Cotton, and Peanuts

Plant-host resistance is a highly desirable tactic that can be used to manage aflatoxin contamination of corn and peanut before harvest. Using the combination of genetic and genomic approaches to elucidate crop defense pathways and to understand the resistance mechanism and regulation will enhance genetic breeding for better crop cultivars and improved host resistance.\textsuperscript{128,129} In traditional crop improvement systems, breeders develop improved varieties based on genetic crosses and phenotypic assessment. In the future, molecular methods, in particular those based on the genomics revolution, have great potential to revolutionize breeding for crop improvement. It will aid in the identification and molecular cloning of high-value genes, allow integration of trait information across species boundaries, and greatly facilitate the introgression of transgenic traits into elite germplasm. Notable research progress made in peanut is the research conducted in USDA-ARS laboratories in Tifton, Georgia,\textsuperscript{130–133} which initiated an Expressed Sequence Tag (EST) sequence project to develop genomic tools and resources in order to decipher the chromosomal location and biological function of genes in the peanut genome, to understand the uniqueness of the peanut plant, and to mitigate aflatoxin contamination. Messenger RNAs (mRNAs, the coding part of the active gene) provide the opportunity to obtain significant information in a more rapid and usable form than studying the entire genome by converting the labile mRNAs into stable double-stranded (ds) DNA for cloning as complementary DNAs (cDNAs). The basic strategy for partial sequencing of the active/expressed genes (ESTs) and using as bait fishing the whole gene is a rapid, efficient method for researching for active gene sequences. Up to today, a total of 43,296 cDNA clones from 10 peanut cDNA libraries have been sequenced. Progress also has been made in short sequence repeat (SSR) marker development in peanut\textsuperscript{130,134} and a few hundred SSR sequences have been released in GenBank (accession number AY237736-237798, AY-310535-310564, AY526357-526456, AY731521-731698). A mapping population, derived from two diverse peanut cultivars, Tifrunner, a Runner type with resistance to tomato spotted wild virus (TSWV) and leaf spots, and GT-C20, a Spanish bunch type with reduced aflatoxin contamination and resistance to leaf rust and bacterial wilt, has been developed for construction of genetic linkage map and encompassing different quantitative trait locus (QTL) traits, such as resistance to TSWV, leaf spots, and aflatoxin contamination.

Application of functional genomics in elimination of aflatoxin contamination in corn and cotton is limited. The corn and cotton communities have developed long oligo microarray (http://www.maizearray.org/ and http://cottongenomemcenter.ucdavis.edu/), which will be, indeed, a valuable asset to scientists working on elimination of aflatoxin contamination in corn and cotton. Luo et al.\textsuperscript{135} have been conducting research experiments using the maize microarray to test the hypothesis that the metabolic pathways in the
developing kernels are affected differently in drought tolerant and sensitive lines under water deficit stress in order to understand the changes of gene expression in response to drought stress during the late stages of seed development and to identify the related biochemical pathways and resistant genes in two corn genotypes, Tex6 and B73. Transcriptional profiles of kernels at 25, 30, 35, 40, 45 days after pollination were compared under normal and water-deficit conditions using the 70-mer maize oligonucleotide arrays from Maize Oligonucleotide Array Project. Several inbred lines with different tolerance of drought stress and Aspergillus flavus infection have been tested for validation of gene expression detected by microarray study. Ten “cross-talking” genes have been identified.

27.5.4 Gene Profiling Through Aspergillus flavus Genomics

Genomic methods such as EST, microarray technologies, and whole genome sequencing provides robust tools for profiling genes involved in aflatoxin production and for studying the regulatory mechanisms of gene expression. A. flavus EST has been completed using wild-type strain NRRL 3357 (ATCC# 20026). Over 26,110 cDNA clones from a normalized cDNA expression library were sequenced at The Institute for Genomic Research (TIGR). After comparison and assembly of overlapping sequences, 7218 unique sequences were identified. These EST sequences have been released to the public at the NCBI GenBank Database (http://www.ncbi.nlm.nih.gov/). The Gene Ontology data were compiled to construct the A. flavus gene index which can be accessed at the TIGR web site (http://www.tigr.org/tdb/tgi/). From the EST database, not only the known aflatoxin pathway genes in the aflatoxin pathway gene cluster in A. parasiticus and A. flavus were identified, but also four new transcripts (hypB, hypC, hypD, and hypE) were identified. Other categories of genes identified could be potentially involved directly or indirectly in aflatoxin production such as in global regulation, signal transduction, pathogenicity, virulence, and fungal development. The genomic DNA amplicon microarray consisting of 5002 gene-elements was constructed at TIGR. Profiling of genes involved in aflatoxin formation using microarrays, performed at Southern Regional Research Center (SRRC) and at TIGR, identified hundreds of genes that are significantly up or down regulated. Further study on their functional involvement in aflatoxin formation is underway.

The A. flavus whole genome sequencing by a shotgun approach was carried out at TIGR led by Dr. William C. Nierman with the funding from USDA, National Research Initiative awarded to Professor Gary Payne, North Carolina State University, Raleigh, North Carolina, U.S.A. The Food and Feed Safety Research Unit of Southern Regional Research Center, USDA/ARS, provided funding for fine finishing and gene calling. Currently, the sequencing has been completed. Primary assembly indicated that the A. flavus genome consists of eight chromosomes and the genome size is about 36.3 Mb. Annotation of the A. flavus genome sequence data with the help of A. flavus EST database demonstrated that there are about 13,000 functional genes in the A. flavus genome similar to related Aspergillus species. Genes responsible for the biosynthesis of secondary metabolites such as aflatoxins are those encoding PKSs, nonribosomal peptide synthetases (NRPS), cytochrome P450 monoxygenases, FAS, carboxylases, dehydrogenases, reductases, oxidases, oxidoreductases, epoxide hydrolases, oxygenases, and methyltransferases. Primary annotation revealed that there exist over two dozens of PKSs, many NRPS, and numerous cytochrome P450 enzymes in the A. flavus genome. With the availability of the A. oryzae whole genome sequence, a close relative of A. flavus, which is used in industrial fermentation for enzyme production that produce no aflatoxins, we have compared the two genomes and identified unique gene sets in each organism. Comparative genome hybridization (CGH) has been planned to confirm the unique genes in the aflatoxin-producing strain (A. flavus NRRL 3357) and the food grade industrial non-aflatoxin-producing strain (A. oryzae RIB 40). Further comparative analysis of the unique genes from both species and further studies on their biological functions will reveal the secret of the mechanism of aflatoxin formation. A. flavus genomics is expected to provide valuable information for devising strategies in solving aflatoxin contamination of food and feed.

Chapter 2 by Gary Payne et al., Chapter 6 by Masayuki Machida, and Chapter 3 by William Nierman discuss A. flavus genomics in great detail.
27.6 Summary and Prospects

*A. flavus* is the most common cause of aflatoxin contamination in preharvest field crops and postharvest grains. The chemistry, biochemistry, or molecular biology, and synthesis of aflatoxins B1 and B2, and their transcriptional regulation have been studied in great detail. We have discovered the aflatoxin pathway gene cluster, a sugar utilization gene cluster, and a nitrogen pathway gene cluster. With the rapid progress in genomics of host crops and fungal pathogens, the mechanisms of aflatoxin formation, pathogenicity of the fungus, and crop-fungus interaction are expected to be revealed. A large research community has been formed in an effort to understand the biology of the fungus and biosynthesis of aflatoxins with the goal of developing novel control strategies. *A. flavus* genomics will contribute greatly to the accomplishment of this goal. Identification and functional elucidation of those genes that are responsible for aflatoxin formation, regulation, signal transduction, pathogenicity, and the environmental effects on aflatoxin production by the fungus could provide vital information for devising novel strategies to eliminate preharvest aflatoxin contamination resulting in a safer, economically viable food and feed.

References

45. Yabe, K., Ando, Y., and Hamasaki, T., Desaturase activity in the branching step between aflatoxins B

V

Methods: Techniques and Resources
28

Microarrays in Aspergillus Species

Andrew Breakspear and Michelle Momany

CONTENTS
28.1 Introduction to Microarrays ................................................................. 475
28.1.1 Platforms ................................................................. 475
28.1.2 RNA Isolation ................................................................. 475
28.1.3 Labeling and Hybridization ....................................................... 476
28.1.4 Image Analysis, Normalization, and Data Interpretation ............... 476
28.2 Microarray Studies in Aspergilli ....................................................... 476
28.3 Conclusion .................................................................................. 478
Resources .......................................................................................... 479
Analysis Software ................................................................................ 479
References .......................................................................................... 480

28.1 Introduction to Microarrays
Since 2003, 11 microarray studies in Aspergillus species have been published (Table 28.1). Before discussing these studies, we provide a brief overview of microarray techniques and highlight resources available. There are many excellent reviews on microarray analysis which the reader should consult for more detail (i.e., Hegde et al. [1]; Schulze and Downward [2]).

28.1.1 Platforms
There are currently three types of microarray platforms: cDNA, long oligonucleotide, and short oligonucleotide [3]. For cDNA microarrays, PCR amplicons are printed onto coated glass slides and typically hybridized with two separate cDNA probes labeled with different fluorescent dyes. Differentially expressed genes are identified by the ratio of dye intensities for each spot. cDNA microarrays can be manufactured “in house” or through organizations such as COGEME (Resources box). For long oligonucleotide microarrays, oligonucleotides of 50–80 bases are printed onto coated glass slides and hybridized with fluorescently labeled cDNA probes. Long oligonucleotide microarrays have now been made available for Aspergillus fumigatus and Aspergillus nidulans and can be obtained from the PFGRC (Resources box). For short oligonucleotide microarrays, oligonucleotides of 25–30 bases are synthesized in situ using photo-chemistry. Short oligonucleotide microarrays are currently manufactured by Affymetrix (GeneChips) and Nimblegen and are hybridized using a single biotinylated cRNA probe.

28.1.2 RNA Isolation
Obtaining high-quality RNA is essential for successful microarray analysis. While a number of commercial reagents have been used for extracting RNA from the aspergilli [4,5], TRIzol is the most popular [6–8]. Typically, frozen mycelia are ground into a fine powder under liquid nitrogen using a mortar and pestle. TRIzol is then added to the frozen powder and RNA purified. RNA purity and integrity are estimated by measuring the A260/A280 ratio and observing ribosomal RNA (rRNA) bands on a denaturing agarose gel.
### TABLE 28.1
Microarray Studies in *Aspergillus* Species

<table>
<thead>
<tr>
<th>Organisma</th>
<th>Resultsa and Validationc</th>
<th>Array Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>Identified terrequinone A biosynthetic gene cluster (N)</td>
<td>Nim (Nim) 9541 genesf</td>
<td>GenePix Pro EBarrays [8]</td>
</tr>
<tr>
<td><em>A. flavus/A. parasiticusb</em></td>
<td>Identified aflatoxin biosynthesis genes (RT)</td>
<td>cDNA (TIGR) 5002 unique ESTs</td>
<td>UCSF Spot SAS [14]</td>
</tr>
<tr>
<td><em>A. flavus/A. parasiticusb</em></td>
<td>Identified aflatoxin biosynthesis genes (PM)</td>
<td>cDNA (NCSU) 753 unique ESTs</td>
<td>Scanalyze SAS [7]</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>Obtained profiles of metabolic and industrially important genes (N)</td>
<td>cDNA (ATGC) 2070 unique ESTs</td>
<td>GenePix Pro Genomic Profiler [4]</td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>Evaluated subtraction library enriched with polysaccharide metabolism genes (N)</td>
<td>cDNA (OSU) 728 unique ESTs</td>
<td>GenePix Pro Genesis [5]</td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>Evaluated array by examining known metabolic genes (PN)</td>
<td>cDNA (COGEME) 3752 ESTs</td>
<td>GenePix Pro [13]</td>
</tr>
<tr>
<td><em>A. flavus/A. parasiticusb</em></td>
<td>Identified aflatoxin biosynthesis genes (NV)</td>
<td>cDNA (NCSU) 753 unique ESTs</td>
<td>Scanalyze [6]</td>
</tr>
<tr>
<td><strong>Pathogenesis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. fumigatus</em></td>
<td>Identified voriconazole adaptation genes (RT)</td>
<td>cDNA (TIGR) 9516 genes</td>
<td>TIGR Spotfinder MIDAS TIGR MeV [15]</td>
</tr>
<tr>
<td><em>A. fumigatus</em></td>
<td>Identified temperature regulated genes (NV)</td>
<td>cDNA (TIGR) 9516 genes</td>
<td>TIGR Spotfinder MIDAS TIGR MeV [16]</td>
</tr>
<tr>
<td><strong>Protein Secretion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>A. niger</em></td>
<td>Evaluated subtraction library enriched with dithiothreitol stress genes (SL)</td>
<td>Affy (Affy) Whole genome</td>
<td>Affymetrix Microarray Suite [18]</td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>Identified unfolded protein response genes (NV)</td>
<td>cDNA (COGEME) 5579 ESTs</td>
<td>GenePix Pro MaxdView [17]</td>
</tr>
</tbody>
</table>

*aGenome sequences are available for all species except *A. parasiticusb* (Resources box).

*bFurther details can be found in the text.

*cThe method used for validation is given in parentheses; N, Northern analysis; PN, compared to previously published Northern; PM, compared to previously published microarray; RT, Real-Time quantitative PCR; SL, compared to genes isolated from subtraction library; NV, no validation reported.

*dRepresents 96% of predicted number of genes in *A. fumigatus* genome.

*eFurther information is available to academic groups and nonprofit organizations (hans.roubos@dsm.com).

*fRepresents each predicted gene in *A. nidulans* genome database (Broad institute).

*gSecond generation arrays of *OBrian et al. [6]* with additional positive and negative control spots.

*hThe source of the microarray is given in parentheses; Affy (Affymetrix); ATGC (Asahi Techno Glass Corporation); COGEME (Consortium for the Functional Genomics of Microbial Eukaryotes); Nim (Nimblegen); OSU (the Microarray Core facility at Oklahoma State University); NCSU (the Payne lab at North Carolina State University); TIGR (The Institute for Genomic Research).

**iRepresents each predicted gene in *A. nidulans* genome database (Broad institute).*

**jA. flavus* cDNA microarray hybridized with *A. parasiticusb* probes.

**kSecond generation arrays of *OBrian et al. [6]* with additional positive and negative control spots.

**lArray contains some duplicate sequences. A further 340 PCR products were designed from sequences deposited in GenBank.

**mRepresents 96% of predicted number of genes in *A. fumigatus* genome.

**nFurther information is available to academic groups and nonprofit organizations (hans.roubos@dsm.com).

**oSecond generation arrays of *Sims et al. [13]* including extra PCR products designed from ESTs generated using negative subtraction hybridization. Also used first generation arrays from *Sims et al. [13]*.
28.1.3 Labeling and Hybridization

Fluorescent dyes (usually Cy3 and Cy5) are routinely used to label probes for hybridization to cDNA and long oligonucleotide arrays. For both direct and indirect labeling fluorescent cDNAs are made prior to hybridization by incorporation of modified nucleotides during reverse transcription. The 3DNA system (Genisphere) introduces dyes after microarray hybridization and is much less dependent on cDNA sequence or length. For comparisons of all three labeling methods see Manduchi et al. [9] and Badiee et al. [10]. After the cDNA probes have been mixed together they are hybridized to a microarray slide under a glass coverslip and scanned at wavelengths corresponding to each fluorescent dye. Detailed labeling and hybridization protocols can be found on the PFGRC website (Resources box). Protocols for the labeling and hybridization of short oligonucleotide arrays are available to registered users from Affymetrix (Resources box).

28.1.4 Image Analysis, Normalization, and Data Interpretation

Raw tiff images generated by the laser scanner are analyzed to calculate intensity values for each spot on the array. The values are then normalized to remove variation in relative intensity that is not related to biological expression. Although differentially expressed genes can be identified following normalization, a variety of statistical methods are available for further interpretation. Algorithms such as hierarchical clustering [11] and k-means clustering [12] are used to identify sets of genes with similar expression patterns across a range of experimental conditions. A variety of software packages are available for imaging, normalization, and data analysis, many are continually updated. A selection, including software used in Aspergillus studies are listed in the Analysis Software box.

28.2 Microarray Studies in Aspergilli

Microarray analysis depends on comparing gene expression in two well-defined conditions. Studies in aspergilli have so far compared gene expression under different nutritional conditions (i.e., glucose-rich versus glucose-depleted), at different stages of development (i.e., conidiation), and in the presence or absence of stressors (i.e., drugs, high temperature). These studies have addressed questions in the areas of metabolism, pathogenicity, and protein secretion (Table 28.1).

Several studies have exploited microarrays to examine the expression of specific primary metabolic genes in aspergilli. A. oryzae is used extensively for industrial fermentations. Maeda et al. [4] used an A. oryzae cDNA array to investigate the gene expression of A. oryzae grown on different media, focusing on conditions important in industry. They showed that most catabolic genes of the glycolytic pathway and the tricarboxylic acid cycle were expressed at higher levels in glucose-rich medium than in glucose-depleted medium. Maeda et al. also showed that growth on wheat bran produced the richest set of industrially important hydrolytic enzymes and validated its current industrial use. In a similar study, Sims et al. [13] evaluated an A. nidulans cDNA array by examining the expression of previously characterized metabolic genes in response to a change of carbon source. In an unusual use of the technology, Ray et al. [5] used cDNA microarrays to confirm that genes identified by negative subtraction hybridization were induced by a switch from medium containing glucose to medium containing specific polysaccharides.

The ability of Aspergillus species to produce a vast array of secondary metabolites has inspired several microarray studies. O’Brian et al. [6] aimed to identify genes with roles in aflatoxin biosynthesis in A. parasiticus. Comparison of expression profiles in cultures prior to and during aflatoxin production identified 42 differentially expressed genes. Because aflatoxin production is tied to development, some of these differentially expressed genes are likely not directly associated with aflatoxin. A further study investigated the influence of growth conditions on aflatoxin production [7]. Cluster analysis identified one particular group of differentially expressed genes predicted to constitute a putative regulon. Price et al. [14] used a different experimental approach to target a more specific set of genes. A wild-type strain was compared with a mutant (ΔaflR) missing the aflatoxin pathway regulator. The study identified 20 genes in the
The Aspergilli

478

The Aspergilli

The Aspergilli

aflatoxin biosynthetic cluster and 3 genes outside the cluster all with AflR-binding sites. A similar strategy was used by Bok et al. [8] to identify *A. nidulans* secondary metabolite gene clusters controlled by the transcriptional regulator LaeA. A comparison of an laeA deletion mutant (ΔlaeA) with wild type revealed a biosynthetic cluster in which all five genes were down-regulated in the laeA strain. High pressure liquid chromatography and mass spectroscopy confirmed the biosynthetic cluster made the antitumor compound terrequinone A, a metabolite not previously described in *A. nidulans*.

Microarray analysis has also been used to identify genes associated with pathogenicity. Ferreira et al. [15] examined the ability of the human pathogen *A. fumigatus*, to adapt to the antifungal agent voriconazole. The study identified 2271 genes which were differentially expressed over a time course following exposure to the drug. While the list of genes alone would be of little use to the investigator, cluster analysis allowed their organization into groups of genes with similar expression patterns. One particular cluster contained an ABC multidrug transporter and a glutathione S-transferase, both thought to have roles in voriconazole detoxification. Nierman et al. [16] investigated the thermotolerance of *A. fumigatus*. Almost 2000 differentially regulated genes were identified and cluster analysis was applied to distinguish those of interest. Only eleven of the 551 homologs of the *Saccharomyces cerevisiae* general stress response genes found in *A. fumigatus* are up-regulated at high temperature suggesting that the stress response machinery will be very different in this pathogen.

Microarray analysis has also analyzed the industrially important process of protein secretion. Sims et al. [17] used a cDNA microarray approach to compare a recombinant chymosin-producing and a wild-type strain of *A. nidulans*. The study identified many secretion related genes involved in the unfolded protein response (UPR). The UPR was also induced using the secretion blocker dithiothreitol (DTT). Although many differentially expressed genes were identified it was suggested that many may have been artifacts associated with DTT treatment. A similar study [18] using *Aspergillus niger* also concluded that DTT does not specifically induce UPR genes and underscores the value of comparing gene expression under multiple conditions.

28.3 Conclusion

Microarray studies of secondary metabolite biosynthesis in aspergilli have been particularly fruitful. For many metabolic pathways, a change in development or environment induces a transcriptional regulator that in turn induces dramatic changes in the expression of a whole suite of genes—exactly the sort of pattern cluster analysis is best at finding. Several early studies in aspergilli have highlighted the value of cluster analysis in organizing vast lists of differentially expressed genes into groups with similar expression profiles. It is unlikely that the putative aflatoxin regulon in *A. parasiticus* would have been identified without this powerful tool.

Insight into the evolution of several aspergilli has been gained using comparative genomics [19]. Comparative functional genomics, analyzing expression profiles of diverse *Aspergillus* species, will enhance our understanding of how this important group of fungi evolved. Comparative functional genomics, perhaps using cross-species microarray analysis between less-studied aspergilli and their better studied relatives, should also allow insight into evolutionary questions such as how gene position affects expression. The studies of aflatoxin biosynthesis pioneered by OBrian et al. [6], demonstrated the successful application of cross-species analysis. However, *A. parasiticus* and *A. flavus* share almost identical sequence in the aflatoxin biosynthetic cluster. The evolutionary limits for cross-species analysis of aspergilli have yet to be established. An investigation detailing the usefulness of current microarrays for studies on other related members of the genus would be of value to the *Aspergillus* research community. It would also be interesting to establish which microarray platform is best suited to cross species study. The redundancy associated with the short oligonucleotide platform should increase the likelihood of detecting genes from related species.

The recent completion of several *Aspergillus* genomes and advances in their annotation has paved the way for the fabrication of more “whole genome” microarrays. The next few years should see an explosion of microarray studies within the genus contributing to an understanding of gene expression and how it is regulated.
Resources

General Microarray Resources

Microarray Worlda http://www.microarrayworld.com/
TIGRb http://www.tigr.org/db/microarray/

Microarray Manufactures/Suppliers

Affymetrixc http://www.affymetrix.com/
Agilentd http://www.agilent.com
COGEMEc http://www.cogeme.man.ac.uk/
Nimblegenf http://www.nimblegen.com/
PFGRCd http://pfgrc.tigr.org/

Microarray Data Repositories

ArrayExpress http://www.ebi.ac.uk/arrayexpress/

Aspergillus Genomes

A. flavuse http://www.aspergillusflavus.org/genomics/
A. fumigatusf http://www.sanger.ac.uk/Projects/A_fumigatus/
A. nidulang http://www.broad.mit.edu/annotation/fungi/aspergillus/
A. nigerh http://genome.jgi-psf.org/Aspni1/Aspni1.home.html
A. nigerj http://www.integratedgenomics.com/index.html

Analysis Software

Free “Open Source” Software

EBarrays http://bioconductor.org/packages/1.9/bioc/html/EBarrays.html
MaxdView http://www.bioinf.man.ac.uk/microarray/maxd/index.html
TIGR MeV http://www.tm4.org/mev.html
TIGR MIDAS http://www.tm4.org/midas.html
TIGR Spotfinder http://www.tm4.org/spotfinder.html

aContains extensive links detailing protocols, instruments, hardware, software, and databases.
bThe Institute for Genomic Research; provides a variety of microarray resources including protocols, software and tutorials.
cDesign and manufacture custom microarrays.
dPathogen Functional Genomics Resource Center; supplies A. fumigatus and A. nidulans oligonucleotide microarrays in addition to providing software and protocols.
e5x coverage sequenced by TIGR; publicly available.
f10.5x coverage sequenced by The Sanger Institute and TIGR; publicly available.
g13x coverage sequenced by TIGR and Monsanto; publicly available.
h8x coverage sequenced by the Department of Energy’s Joint Genome Institute; publicly available.
i8x coverage sequenced by Gene Alliance for DSM; available to academic and nonprofit organizations.
j4–6x coverage sequenced by Integrated Genomics; available on request (scott.baker@pnl.gov).
k>9x coverage sequenced by the National Institute of Technology and Evaluation (NITE); publicly available.
Free to Academic and Nonprofit Organizations

Cluster http://rana.lbl.gov/EisenSoftware.htm
GeneCluster http://www.broad.mit.edu/cancer/software/
Genesis http://genome.tugraz.at/genesisclient/genesisclient_description.shtml
Hierarchical Clustering Explorer http://www.cs.umd.edu/hcil/hce/
ScanAlyze http://rana.lbl.gov/EisenSoftware.htm
TreeView http://rana.lbl.gov/EisenSoftware.htm
UCSF Spot http://jainlab.ucsf.edu/Downloads.html

Commercially Available

Affymetrix Microarray Suite http://www.affymetrix.com/products/software/specific/mas.affx
Expressionist http://www.genedata.com/productoverview/expressionist/index_eng.html
GCOS http://www.affymetrix.com/products/software/specific/gcos.affx
GenePix Pro http://www.moleculardevices.com/pages/software/gn_genepix_pro.html
ImaGene http://www.biodiscovery.com/index/imagene
SAS http://www.sas.com/software/sas9/
Spotfire DecisionSite http://www.spotfire.com/products/decisionsite.cfm

References

29

Chemostats and Microarrays

Manda Gent and Karin Lanthaler

CONTENTS
29.1 Introduction ............................................................................................................ 483
29.2 Comparison of Batch and Chemostat Fermentations ............................................. 483
29.3 Microarrays and Filamentous Fungi ....................................................................... 487
29.4 Design of Experiments ........................................................................................... 487
29.5 Conclusions ............................................................................................................. 489
References ..................................................................................................................... 489

29.1 Introduction

The recent availability of full genome sequences for some filamentous fungal species, for example, the human pathogen 
*Aspergillus fumigatus* (Nierman et al., 2005), the commercially important species 
*A. oryzae* (Machida et al., 2005) and *Trichoderma reesei* (http://www.genomesonline.org/), the model organism 
*A. nidulans* (Galagan et al., 2005) and the white rot fungus *Phanerochaete chrysosporium* (Martinez et al., 2004) (for a full list of all completed genomes and genomes being sequenced see http://
www.genomesonline.org/), together with the development of cutting-edge technologies has led to the possibility of investigating global protein, mRNA and metabolite profiles of a chosen organism under certain defined growth conditions, as a result of a specific treatment or in response to an engineered change in the expression of a single gene. These technologies provide the possibility of characterizing cell physiology at a molecular level, providing temporal, spatial, and even real-time information (Hoskisson and Hobbs, 2005). However, these technologies require the production of reproducible, reliable and homogeneous datasets in order to gather meaningful information. The cell populations must, therefore, be grown in defined, ideally constant, controllable, physico-chemical conditions.

Simple “batch culture” systems result in dynamic physico-chemical conditions that are difficult to reproduce and produce complex data patterns that are often difficult or impossible to interpret. Where possible, a continuous culture system is preferable where growth conditions can be constantly controlled and maintained and importantly, reproduced.

29.2 Comparison of Batch and Chemostat Fermentations

For physiological reasons it is vital for microarray experiments to compare either different strains under equivalent culture conditions or the same strain under conditions where only a single variable (e.g., carbons source, temperature, growth rate etc.) is changed. To appreciate the full impact the choice of fermentation technique has upon the physiological environment, and therefore on differential gene expression,
it is necessary to understand the underlying growth kinetics of the organism and the underlying kinetics of chemostat cultivations from a physiological point of view, before the technique can be applied in a relevant manner.

Experimental and theoretical studies of chemostat cultivations were conducted in the 1950s (Herbert et al., 1956) and the foundation for the correct theoretical treatment was described as early as 1950 by Monod (1950) and Novick and Szilard (1950). Theoretical descriptions of steady state and batch cultivations based on the Monod principle can be found in any comprehensive Microbiology text book (e.g., Madigan et al., 2000).

Chemostats are widely applied for sophisticated studies of bacteria and yeast cultures under steady-state conditions and the theoretical background is well understood and proven to be correct by numerous experimental approaches.

A schematic of a completely mixed bioreactor with fittings can be found in Figure 29.1a.

The vessel geometry, stirrer type and layout may vary greatly depending on the specific organism or process requirements. General guidelines can be found in (e.g., Pirt, 1975).

In contrast to a batch reactor, which in its simplest form consists of an agitated Erlenmeyer flask, a bioreactor run in chemostat mode is constantly supplied with fresh growth medium at a constant flow rate \( f/lh^{-1} \) and spent medium and surplus biomass are constantly removed from the growth vessel at the same rate. The working volume \( v/l \) of the vessel is kept constant throughout the experiment. It should be emphasized at this point, that for fermentations with filamentous fungi, it is essential to fill up the reactor completely so that the working volume equals the vessel volume \( (V_m) \), \( (v = V_m) \). Conventional fermentation approaches leave one-third of the vessel volume as headspace, but filamentous fungi grown in such a setup have the ability to form a “biofilm” by attaching to the top plate and filling up this headspace with mycelium. This is especially undesirable when biomass samples are used for microarray investigations as attached biomass is in stationary growth phase and, therefore, in a physiologically different state compared to the biomass in the liquid phase. Pieces of the attached biomass detach on a regular basis and so “contaminate” the steady-state biomass in liquid culture. Chemostat fermentations with filamentous fungi are, therefore, best carried out according to the method of Wiebe and Trinci (1991), as described later.

Biomass levels are kept low (ideal below 5 gl^{-1}) in order to guarantee that dissolved oxygen is not the limiting nutrient (unless this is intended). Dissolved oxygen probes normally used in fermentation technology are usually not applicable with filamentous fungi as they rapidly grow over the membrane and, therefore, do not allow for correct measurement of dissolved oxygen. With biomass levels below 5 gl^{-1}, oxygen probes can be dispensed with, once oxygen excess has been proven which thus removes a surface that otherwise would promote surface attachment.

The residence time is the time an average fungal filament remains in the bioreactor. It is coupled with the flow rate \( f \) and the working volume \( v \) via the dilution rate \( D \).

The dilution rate \( D/h^{-1} \) is defined as \( f/lh^{-1}/v/l \) and is a measure for the number of complete volume-changes per hour. The basic question we need to ask in order to determine the residence time therefore is: How long does the pump need to run before the vessel volume is replaced? The answer is expressed as \( 1/D \) [h].

In liquid medium, growth is usually described using the familiar equation for exponential growth:

\[
\left(\frac{1}{x}\right) \left(\frac{dx}{dt}\right) = \frac{d(lnx)}{dt} = \mu = \ln2 \frac{t_d}{t_d} \tag{29.1}
\]

Where \( x \) is the biomass expressed as dry weight per volume [gl^{-1}], \( t \) is the time, \( \mu \) is the specific growth rate \( [h^{-1}] \), and \( t_d \) is the doubling time [h].

\( \mu \) and \( t_d \) may be assumed to be constant if all substrates necessary for growth are present in excess.

Monod (1942) was the first to show that there is a correlation between the specific growth rate \( \mu \) and the concentration of the limiting substrate according to:

\[
\mu = \mu_{max} \left(\frac{s}{K_s + s}\right) \tag{29.2}
\]

where \( \mu_{max} \) is the maximum specific growth rate \( [h^{-1}] \), \( s \) is the concentration of the limiting substrate and \( K_s \) is the saturation constant and equals the substrate concentration at \( \frac{1}{2} \mu_{max} \).
The logical conclusion that can be drawn from Equation 29.2 is that $\mu$ can have any value between 0 and $\mu_{\text{max}}$, only requiring that the concentration of the limiting substrate be held constant at the appropriate set point value. This cannot be achieved in batch fermentations but is one of the key features of chemostat cultivations. The simplest experimental chemostat set-up uses a growth medium that contains only a single substrate which is growth limiting with all the other substrates present in excess.
The Aspergilli

(providing adequate stirring and oxygen supply), although chemostat experiments have been conducted with bacterial cultures, where multiple substrates were growth limiting (Egli, 1991; Gottschal, 1993; Zinn et al., 2004).

The variables which are under control of the experimenter will therefore be, temperature, medium composition, and flow rate for the incoming fresh growth medium.

The physiological parameters, such as the maximum specific growth rate which can be reached by a filamentous fungus on changing media or by changing the temperature, as well as the yield $Y = \text{weight of biomass formed/weight of substrate used}$, lies beyond the control of the experimenter. These are constants for a given strain when grown under the same culture conditions. It is important, therefore, to determine these carefully for each strain before undertaking fermentations of biomass intended for microarray experiments.

The flow rate must not be set to a rate higher than the maximum specific growth rate of the fungus because this would cause wash out and therefore loss of the biomass according to:

$$-\frac{dx}{dt} = Dx \quad (29.3)$$

In a bioreactor, which is run in chemostat mode, the fungus is growing at a rate described by Equation 29.1 and is simultaneously washed away by a rate given by Equation 29.3. Therefore, the rate of the change of fungal biomass in the vessel is given by the simple equation; increase in biomass = growth – output, or in mathematical terms:

$$\frac{dx}{dt} = \mu x - Dx \quad (29.4)$$

In steady state the biomass is constant and, therefore, $\frac{dx}{dt} = 0$. Solving Equation 29.4 with $\frac{dx}{dt} = 0$, it can be seen that, in steady state, $\mu = D$ and, therefore, the growth rate can be set by the applied dilution rate (which itself can be set by the flow rate at which fresh medium is pumped into the culture vessel).

Chemostat cultivations reduce the physiological background, where sampling is independent of the cultures age (within known time limits before evolved strains start to out-compete the parental strain (Wiebe et al., 1991, 1992, 1998; Swift et al., 2000). Chemostat cultivations allow the comparison of different strains under equivalent culture conditions and, therefore, allow the experimenter to assess the impact which the change of a single variable (e.g., carbon source, temperature, pH, growth rate) has upon gene expression. Furthermore, chemostats allow the study of growth rate dependent phenomena.

Table 29.1 gives a brief comparison of growth characteristics in batch fermentation compared to chemostats.

<table>
<thead>
<tr>
<th><strong>TABLE 29.1</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison of Growth Characteristics in Batch Fermentation and Chemostats</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Batch Cultivation</th>
<th>Chemostat Cultivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag phase</td>
<td>Constant growth rate which is set by the applied dilution rate</td>
</tr>
<tr>
<td>Log phase</td>
<td>Constant biomass</td>
</tr>
<tr>
<td>Deceleration phase</td>
<td>Constant nutrient supply and constant nutrient concentration</td>
</tr>
<tr>
<td>Stationary phase = variable growth rate during cultivation</td>
<td>Metabolites removed constantly during the cultivation</td>
</tr>
<tr>
<td>Increase in biomass</td>
<td>Controllable variables:</td>
</tr>
<tr>
<td>Changes in nutrient concentration</td>
<td>pH</td>
</tr>
<tr>
<td>Accumulation of metabolites during the cultivation</td>
<td>T</td>
</tr>
<tr>
<td>Controllable variables:</td>
<td>Initial medium composition</td>
</tr>
</tbody>
</table>
29.3 Microarrays and Filamentous Fungi

A microarray works by exploiting the ability of a given mRNA molecule to hybridize specifically to a DNA template probe attached to a solid support such as a glass microscope slide, silicon chip, or nylon membrane. The probe can be DNA, cDNA, or oligonucleotides representing a specific sequence from each gene in the genome. Each gene from an organism can be represented either singly or in replicates, together with relevant control sequences. cDNA prepared from mRNA extracted from biomass grown in two different conditions under study is labeled with one or other of two fluorescent dyes (usually Cy3 or Cy5) and competitively hybridized to the array. A comparison of the quantities of cDNA hybridizing to the array from the two different conditions gives a ratio which is proportional to the difference in transcription levels of each gene represented on the array. For a full description of the use of microarrays see Momany et al., (ibid) and for a more detailed discussion of microarrays and the analysis of data obtained, see the excellent text by Causton et al. (2003).

Functional genomics analyses using microarrays require highly controlled experimental conditions in order to gather relevant data from a small set of experiments. In order to obtain relevant and conclusive information, it is vital to keep the number of confounding variables affecting an experiment to an absolute minimum. Using chemostat cultures enables the experimenter to exercise control over the physiological and chemical environment during growth of the fungus and, therefore, over the number of variables that might influence the overall outcome of the experiment (Hayes et al., 2002; Wu et al., 2004). During classical batch growth experiments, the growth rate may change even in exponential growth phase and the chemical and physical environment of the culture changes constantly throughout the course of the experiment. This results in changes in the concentrations of nutrients, metabolites, and biomass (Hoskisson and Hobbs, 2005; Ter Linde et al., 1999). Moreover, the experimenter lacks complete control over the growth rate a particular fungal strain is able to reach in a given environment.

All of these variables may be removed by using chemostat cultivations (Pirt, 1975), where, during steady state, all the culture variables are held constant, including pH, temperature, nutrient concentrations, excreted metabolites, biomass level, and growth rate (Wu et al., 2004). Whereas the use of chemostats is routine in genome-wide transcriptome experiments involving the unicellular microbes such as the yeast Saccharomyces cerevisiae (Darán-Lapujade et al., 2004; Piper et al., 2002; Hayes et al., 2002; Saldanha et al., 2004, etc.) and many bacterial species (Bacon et al., 2004; Shockley et al., 2005; Silberbach et al., 2005; Silberbach and Burkovski, 2006 and references therein), to date, there have been relatively few transcriptome studies employing microarray technology in the filamentous fungi (Chambergo et al., 2002; Aign and Hoheisel, 2003; Foreman et al., 2003; Sims et al., 2004; 2005; Xie et al., 2004; Maeda et al., 2004; Babu et al., 2005; Felipe et al., 2005; Mackenzie et al., 2005). Only two of these, Sims et al., 2004 and 2005, describe using chemostats for the growth of the biomass used in the transcriptome experiments.

29.4 Design of Experiments

Microarray experiments usually generate large amounts of data which can be mined over a long period of time. It is essential, therefore, that experiments are carefully planned so that data can be added to or combined with data from other laboratories in order to extend an experimental study and extract the most out of the data. A single microarray experiment can produce data on the expression of thousands of genes but sometimes the data can be noisy and individual data points may be unreliable, particularly for genes with low abundance transcripts. A particular gene of interest may not be apparent on a microarray and conversely, a large fold change does not necessarily reflect greater biological relevance. A microarray experiment may not, therefore, necessarily be the best form of experiment to carry out. Northern analysis or real-time PCR may be better for studying the expression of a single or few specific genes, or a proteome study may be more informative for the analysis of protein abundance. There are two different formats for microarrays, glass slides and the Affymetrix silicon chips. As the only readily available arrays for filamentous fungi are the glass slide arrays from TIGR (http://pfgrc.tigr.org/resources.shtml) these alone will be referred to here.
A large number of variables can contribute to differences in transcription. Variation in batches of culture media or reagents, for example, can bring about subtle changes in gene expression and thus complicate data analysis, especially where the number of samples is small. The use of chemostat cultures, as described earlier, will minimize most sources of variability in growth conditions. However, the largest sources of variation are often due to the person conducting the experiment or environmental factors such as the humidity, temperature, and ozone levels in the laboratory environment. Replicates of experiments are very important in order to minimize the effects of these variations. Technical replicates are arrays that use the same RNA samples and also the same treatments, thus the only differences in measurements are due to the technical differences in array processing. These can, therefore, be used to assess experimental noise. Biological replicates use different RNA samples generated in the same conditions and will determine the natural variability in the system.

The design of a microarray experiment will depend largely on what it is you wish to investigate. The simplest design is the comparison of genes transcribed in one condition with those genes transcribed in another condition, or a wild-type strain compared to a mutated or otherwise altered strain. For this type of experimental design, called the “universal reference design” depicted in Figure 29.2a (Kerr and Churchill, 2001), it is preferable to use a “dye flip” labeling design where the reference and control samples are labeled separately with both Cy dyes so that gene-specific biases in labeling are removed from the data. Each sample is compared to the same reference sample. This results in more data sets for the reference sample which are of little interest, than for the experimental samples, which are of interest. The same design can be used for time course experiments as well, where each time point is referred to one time point (typically time zero). Additionally, the reference sample can be compared to itself to provide more data about the variation in the experimental process (shown in Fig. 29.2c).

An alternative to this is the “loop design” proposed by Kerr and Churchill (2001), in which labeled extracts are serially compared with each other rather than to a common reference, as shown in Figure 29.2b. In this design, each hybridization measures expression in a biologically interesting sample, allowing more relevant data to be generated. There are disadvantages with this design in that more RNA may be required than for a reference design and a single poor quality hybridization can introduce uncertainty and missing data into the analysis. Additionally, the loop design is not easily extendable, unlike the reference design.

---

**FIGURE 29.2** Experimental designs.
(For a full discussion of experimental design see Kerr and Churchill, 2001.) Figure 29.2d shows a modified loop design which can be used for six samples where it is not possible to do dye flip labeling as it still allows the collection of more data on the variations of interest but assumes that there are no gene-specific dye effects.

A description of the analysis of microarray data is beyond the scope of this chapter but it is worth emphasizing the need for normalization and rigorous statistical analysis using an established method or software package, for example, GeneSpring (Agilent Technologies), MaxdView (available free from http://bioinf.man.ac.uk/microarray/maxd/) or MIDAS (available from TIGR at http://www.tm4.org/midas.html).

29.5 Conclusions

In many cases, it will not be possible to perform microarray experiments growing the biomass in chemostat fermenters perhaps because of economical constraints, as in the case of a study involving an expensive drug treatment, (Lepak et al., 2006), or because the fungus is an obligate pathogen such as *Ustilago maydis*, which must be grown on plant material (Babu et al., 2005). Similarly, if the study in question is one of comparing growth on different materials, such as the one carried out by Maeda et al. (2004), where they compared the differential expression of genes in *Aspergillus oryzae* grown on different industrial solid-phase media, clearly a chemostat cannot be used. However, where possible, it has been demonstrated that the use of a controlled chemostat approach is preferable where the experiment sets out to compare the gene expression of one strain under different culture conditions or to compare genetically distinct strains under the same culture conditions (Sims et al., 2005). A chemostat approach allows for carefully controlled growth conditions which result in considerably lower variation of gene expression patterns throughout the experiment, thus greatly reducing the complexity of the system under study and allowing a much more focused analysis of the factors affecting gene expression (Hayes et al., 2002).

References


Advances in Gene Manipulations Using Aspergillus nidulans

Stephen A. Osmani, Hui-Lin Liu, Michael J. Hynes, and Berl R. Oakley

CONTENTS
30.1 Introduction ............................................................................................................. 493
30.2 General Considerations for Gene Manipulations ..................................................... 494
  30.2.1 Efficient Gene Targeting Requires >500 bp Homologous DNA in A. nidulans .......... 494
  30.2.2 Making Gene Targeting Constructs Using Fusion PCR ..................................... 494
  30.2.3 Recent Advances in Improvement of Gene Targeting—The Ku Story ................. 495
  30.2.4 Types of Transformation Markers for the Aspergilli ........................................... 497
    30.2.4.1 Auxotrophic Markers .............................................................................. 497
    30.2.4.2 Markers Providing Both Positive and Negative Selection ...................... 498
    30.2.4.3 Drug Resistance Markers ................................................................. 498
    30.2.4.4 Markers that can be Recycled for Multiple Gene Modifications ............ 498
    30.2.4.5 Considerations Regarding Accuracy of Gene Calling and Gene Manipulations ................................................................. 499
  30.2.5 Confirming Gene Targeting ..................................................................................... 500
  30.2.6 Nonintegrative Gene Expression Utilizing the AMA1 Sequence ........................... 501
30.3 Specific Types of Gene Manipulations ............................................................................... 502
  30.3.1 Gene Deletion and Promoter Rundown .............................................................. 502
    30.3.1.1 Gene Deletion and the Heterokaryon Rescue Technique ........................ 503
    30.3.1.2 Promoter Rundown ............................................................................ 503
  30.3.2 Fluorescent Protein Tagging .................................................................................... 505
  30.3.3 Affinity Tags for Protein Purifications and Proteomics ........................................ 505
  30.3.4 Two-Step Site-Specific Mutation ............................................................................. 506
30.4 Conclusions ............................................................................................................. 508
Acknowledgments ............................................................................................................. 508
References ................................................................................................................................ 508

30.1 Introduction

The value and utility of any model genetic organism relies on many factors, including the basic biological features of the system and the ease with which the organism can be experimentally manipulated. Aspergillus nidulans has proved to be a highly valued model fungus not only because of its natural biology (e.g., it can undergo self- and out-crosses) but also because over 50 years of development of classical and molecular genetic methodologies have provided a multitude of techniques by which this organism can be experimentally manipulated. Since the advent of gene transformation over twenty years ago gene
The Aspergilli

manipulations in *A. nidulans* have been routine because homologous recombination occurs at a reasonable frequency in this organism.\(^4\) However, recent technical advances have greatly improved gene targeting. These advancements were spurred by the availability of high-quality genome sequences and the desire to easily target all *A. nidulans* genes. In this chapter, we review the types of approach that have been developed to manipulate *A. nidulans* for global functional gene analysis.

### 30.2 General Considerations for Gene Manipulations

#### 30.2.1 Efficient Gene Targeting Requires >500 bp Homologous DNA in *A. nidulans*

One of the most useful features of *Saccharomyces cerevisiae*, which has become the most intensely used and manipulated eukaryotic species, is its ability to utilize very short DNA sequences to mediate homologous recombination into its genome. Because of this ability, which is inherent to the species, it is possible to generate linear gene targeting DNA constructs using polymerase chain reaction (PCR) and primers that directly incorporate 50–80 bp of homology. Thus, when deleting a gene in *S. cerevisiae* or *S. pombe*, it is usual to amplify a selectable marker using two primers that incorporate 50–80 bp corresponding to the sites at which the marker gene is to be landed. For gene deletions, these typically correspond to the ends of the coding region of the target gene. After transformation, the linear DNA finds its way into the nucleus and, at high frequencies, the two 50–80 bp regions of homology locate their homologous sequences, and homologous recombination occurs at both sites leading to a clean replacement of the target gene by the selectable marker.\(^5\) Unfortunately, in *A. nidulans*, and most other eukaryotes for that matter, linear DNA targeting cassettes with 50–80 bp of homology do not integrate into the genome via homologous recombination. Luckily, however, if the targeting DNA domain is expanded to >500 bp of homology, then targeting by homologous recombination is greatly increased, although there is still a significant level of nonspecific integration due to nonhomologous recombination. For linear constructs, circularization of the DNA often occurs followed by a single nonhomologous recombination event that allows the marker DNA to be expressed, but the target gene remains intact.\(^6\) Thus, many transformants are obtained but few have the target gene deleted. Because of these limitations, carrying out gene deletions and other gene targeting events, as described later, in *A. nidulans* has been slow compared to the situation in the yeasts. Designing and making gene targeting constructs with conventional cloning approaches often involved several steps and weeks of work to generate a single construct. After transforming the resulting constructs into *A. nidulans*, it was often necessary to screen through many transformants (this number being variable, but for some constructs up to 50) to identify one with the correct gene modification. However, two new technical advances have greatly improved the efficiency of accomplishing gene targeting in the aspergilli. Although targeting constructs still need to incorporate >500 bp of targeting homologous DNA, these can now be efficiently made using the technique of fusion PCR. Fusion PCR utilizes the sequence data available for many of the aspergilli, along with improved methodologies for faithful replication of DNA by the PCR, to generate targeting constructs that can specifically modify any region of the genome.\(^7,8\) In addition, the frequency of correct targeting of the resulting constructs into the genome has been greatly improved by utilizing strains in which nonhomologous recombination has been crippled such that homologous recombination is the overwhelming way in which the introduced DNA is incorporated into the genome\(^9-12\) as first discovered in *Neurospora crassa*.\(^13\) These advances, as described later, have shortened the time it takes to generate targeting constructs and perform the downstream effort to correctly identify targeted strains immensely, making efficient gene targeting possible in the aspergilli at the genome-wide level.

#### 30.2.2 Making Gene Targeting Constructs Using Fusion PCR

Several methodologies have been published that generate gene targeting constructs using fusion PCR. The purpose of this review is not to explain the procedures\(^7,8,14\) in technical detail but to give a conceptual overview of what is now possible. As a general example, we will consider gene deletions using fusion PCR–generated deletion constructs.
The essence of fusion PCR relies on the ability of DNA polymerases to accurately copy DNA from DNA primers that anneal to the target DNA. Using two primers that anneal in opposite directions on the target DNA and multiple rounds of DNA synthesis, it is, therefore, possible to make, or amplify, any region of DNA for which the sequence is known to allow appropriate primers to be designed. When designing PCR primers, additional nucleotides, which do not anneal to the target DNA being amplified, can be added. Then during PCR amplification this adds specific DNA sequence to the target DNA being made. For example, as described earlier, when gene deletions are completed in yeasts, a selectable marker is amplified adding 50–80 bp to the primers used to amplify the marker gene. These sequences are homologous to the 5' and 3' ends of the gene to be deleted (Fig. 30.1a). However, 50–80 bp of homologous DNA is not sufficient to promote homologous recombination in the aspergilli. However, it is possible to amplify two flanking domains before and after the target gene—utilizing primers that have sequence identity to the sequence that is present on either end of the first amplified DNA product (Fig. 30.1b). If the three fragments are added to a PCR reaction in which two primers are used that prime at either end of the two flanking domains, the full-length construct becomes amplified which contains the marker gene flanked on either side by the DNA that is homologous to sequences 5’ and 3’ of the target gene. Thus, instead of the marker gene having 50–80 bp of targeting sequence, it now has >500 bp. This length of homology then ensures increased levels of accurate targeting of the construct.

Several variations of this basic process have been published that are conceptually the same. One conceptually different approach is the use of a split marker and transformation of two pieces of DNA which recombine when transformed into A. nidulans to generate both a functional marker gene and the desired modification within the genome. Both of the pieces of DNA, each with a region for targeting into the genome, but one containing the 5' end of the transformation marker and the other, the 3' end of the marker, are also generated using fusion PCR. However, in this case, the two targeting constructs are made using two-way fusion PCR and, importantly, the two regions of the transformation marker overlap. The rationale behind this approach is that the only way the two fragments can generate a functional nutritional marker is via homologous recombination, and it is thought that recombination will occur only when the two fragments integrate at the target genomic site. This may not be true as transformants can be obtained at a high frequency without the desired genomic modification, such as gene deletions. The increased rates of homologous recombination using the split marker approach do not approach the frequencies possible using conventional whole marker constructs in combination with recipient strains defective in nonhomologous end joining (as discussed later).

### 30.2.3 Recent Advances in Improvement of Gene Targeting—The Ku Story

The lab of Hirokazu Inoue, while studying the DNA damage response in N. crassa, made the observation that when orthologs of N. crassa, Ku70 or Ku80, were deleted, the resulting strains had amazingly increased frequencies of correct gene targeting via homologous recombination. Subsequently, this phenomenon has been observed in numerous other filamentous fungi including A. nidulans, A. fumigatus, A. sojae, and A. oryzae. Therefore, the original findings in N. crassa are applicable to other fungi and enable the generation of recipient transformation strains lacking either Ku70 or Ku80, with rates of homologous recombination that rival what naturally occurs in S. cerevisiae. This has leveled the playing field between filamentous fungi and S. cerevisiae and S. pombe, somewhat, when it comes to ease of targeted gene manipulations. However, as first defined in N. crassa, although the Ku70- or Ku80-deleted strains integrate target constructs at near 100%, there is still a necessity for the target constructs to have >500 bp of homology to attain these high levels of correct targeting. However, as described in Section 30.2.2 with fusion PCR approaches this requirement is no longer rate limiting. In addition, recently, it has been shown that deletion of the Lig4 ortholog of N. crassa completely eliminates nonhomologous recombination and allows gene targeting with reduced lengths of homologous DNA. This effect has not been tested in the aspergilli.

A natural question arises regarding why deletion of either Ku70 or Ku80 increases the percentage of transformants in which the transformed DNA integrates via homologous recombination? These two proteins form a heterodimer and play a role in the DNA damage response by mediating nonhomologous...
DNA end-joining (NHEJ). When linear DNA is introduced during transformation, it likely promotes activation of the DNA damage response pathway which is known to be activated by DNA ends that are not capped by telomeres. In \textit{S. cerevisiae}, the “repair” process mounted against free DNA ends relies mainly on the homologous recombination (HR) pathway, hence, the linear DNA gets integrated at homologous sites via homologous recombination. However, in \textit{N. crassa}, and other filamentous fungi, both the HR pathway and the NHEJ pathway appear to try and “fix” the ends of the introduced DNA, resulting in

\textbf{FIGURE 30.1} Gene deletion strategies in yeast and \textit{A. nidulans}. Note: (a) To delete a gene in yeasts, a selectable marker gene is amplified using primers that incorporate short (~50bp) regions of homology to the flanking regions of the target gene. Upon homologous recombination between these sequences and the genome, the marker gene replaces the target gene generating a null allele. Strains with the null allele are typically confirmed using diagnostic PCR and primers, indicated as 1 and 2, which will amplify a band from DNA isolated from the null strain but not from wild-type strains. (b) It is necessary to add larger regions of homology (>500 bp) to the deletion construct because in \textit{A. nidulans} homologous recombination is not promoted by short stretches of homology. This is done using fusion PCR. The larger regions of homology promote homologous recombination between the flanking regions of the deletion construct and the genome thus generating a null allele. Diagnostic PCR using primers 3 and 4 generate a positive result in the wild type and in the gene-modified strain. However, the gene modified strain should not amplify the wild-type size band but instead amplify a predictable size band, depending on the relative size of the marker gene and the gene being deleted.
a mixture of both random integrations via the NHEJ pathway and homologous integration via the homologous recombination system. In *A. nidulans*, it has also been demonstrated that, in many instances, linear DNA molecules are first circularized, presumably by the NHEJ system, and then integrate via a region of homology in the resulting circular “plasmid” into the genome via a single homologous recombination event. Thus, both the NHEJ and HR systems can act upon transformed DNA molecules. The net result of these different fates of transformed DNA is a competition between the NHEJ and HR systems. Therefore, as previously described by Inoue, by removing either Ku70 or Ku80 this balance is shifted almost exclusively toward the HR system due to lack of NHEJ caused by deficiency of the Ku heterodimer or exclusively towards the HR system after deletion of Lig4.

It is noticeable that in *A. nidulans*, deletion of orthologs of either Ku70 or Ku80, or the double mutant, are not appreciably sensitive to DNA-damaging agents, including agents that should promote DNA double-strand breaks. This is somewhat surprising given the dramatic effects these deletions have on the fate of transformed DNA, which indicates that the NHEJ pathway is certainly deficient in these strains. Why lack of a NHEJ system does not cause sensitivity to DNA damaging agents remains an open question but suggests there could be a second NHEJ system in *A. nidulans* acting independently of the Ku heterodimer. It will be interesting to ascertain if the *A. nidulans* Lig4 mutant displays a DNA-damage phenotype. Although the Ku70 and Ku80 deletions are surprisingly without measurable phenotypes, it is wise to remove these markers by genetic crosses to avoid complications that could arise from synthetic interactions with mutant alleles created by mutagenesis or molecular genetic manipulations.

By utilizing a combination of fusion PCR techniques and specific gene-modifying templates, as described later, in combination with Ku-deficient host transformation strains, it is now possible to manipulate all the genes of *A. nidulans* with ease. For instance, recent work using these approaches have defined the phenotypes caused by deletion of 30 genes involved in nuclear transport (see Chapter 16 by Espeso and Osmani) and enabled the detailed definition of the disassembly of the nuclear pore complex during *A. nidulans* mitosis utilizing endogenously fluorescent tagged versions of these proteins. It is safe to say that without fusion PCR technologies, and the Ku-deficient recipient strains, these comprehensive studies would not have been attempted.

### 30.2.4 Types of Transformation Markers for the Aspergilli

#### 30.2.4.1 Auxotrophic Markers

To introduce DNA constructs into cells, it is necessary to have a means by which to select for the transforming DNA. In *A. nidulans*, this has historically been done by using DNA constructs that encode genes that are able to complement auxotrophic markers. By plating transformants on to media that lacks the required nutrient, only strains complemented with the nutritional marker encoded by the DNA construct are able to grow. Because of the rich historical use of *A. nidulans* for studies of intermediary metabolism, many auxotrophic mutations have been generated providing a wealth of potential genes to use as transformation markers (http://www.gla.ac.uk/Acad/IBLS/molgen/aspergillus/loci.html). Those most widely used include pyrG89, argB2, trpC3801, riboB2, and pyroA4.

The earliest published report of transformation of *A. nidulans* employed a heterologous cloned gene, *N. crassa* pyr4, to complement a mutant pyrG89 strain on media lacking uridine or uracil. No homologous DNA was present on the transforming DNA. With the development of targeted transformation strategies it has become apparent that the use of heterologous markers is beneficial because use of homologous genes provide a second landing site for transforming DNA, in addition to the target locus. It is now typical to employ genes from other Aspergillus species to transform *A. nidulans* because they tend not to have enough homology to target the auxotrophic marker site, but have the advantage that they function better than equivalent genes from more distantly related fungi such as *N. crassa*. A series of nutritional markers from other aspergilli have now been cloned and can be employed for gene targeting experiments including *A. fumigatus* pyrG, Al-trpC, Al-trpB, and Al-pyroA.
30.2.4.2 Markers Providing Both Positive and Negative Selection

It is often advantageous to use a transformation marker for which positive selection can be employed, first to select for integration of the marker into the genome, and then to select against the marker in order to remove it from the genome. For example, positive/negative selection has been used to introduce new alleles into the genome of *A. nidulans* using a two-step gene replacement protocol.\(^{23,24}\) In these examples, the *pyrG* nutritional marker was used to select for the initial transformation event. The toxicity caused by 5-fluoro-orotic acid (5-FOA) when the *pyrG* function is wild type was then used to negatively select against the introduced nutritional marker. Further details are given in Section 30.3. Other examples of markers providing both positive and negative selection include: *niaD*, auxotrophic marker with positive selection for growth without nitrate and negative selection using chlorate toxicity; *sC*, auxotrophic marker with positive selection for growth without sulfate and negative selection using selenate toxicity; *ble-HSV1 tk* gene chimera cassette with positive selection using resistance to phleomycin and negative selection using FUDR toxicity.\(^{11}\)

30.2.4.3 Drug Resistance Markers

Another useful type of transformation marker is the drug resistance marker. Such markers work by making a normal strain resistant to an added toxic compound, hence, it is not necessary to have a genetically marked recipient strain in which to introduce the DNA. This is particularly useful for species with no auxotrophic markers and those that will not undergo sexual crosses. Also on the positive side, the encoded resistance genes typically do not have sequence similarities to the host genome. On the negative side, drug resistance markers need to be engineered so that the encoded gene can be expressed. This requires addition of promoter sequences, for transcriptional activation, and 3′ termination sequences, to terminate transcription and provide sequence information for polyadenylation of the resulting transcripts. Examples of resistance markers used in the aspergilli include the pyrithiamine resistance gene *ptrA*,\(^{25}\) the phleomycin resistance gene *ble*,\(^{26}\) and the glufosinate resistance gene *bar*.\(^{9}\)

30.2.4.4 Markers that can be Recycled for Multiple Gene Modifications

It is often desirable to make more than one change in the genome of a particular strain. For instance, deletion of a member of a functionally redundant gene family will not reveal a phenotype. This can necessitate completing several sequential deletions in a single strain. In addition, it is often useful, after a gene is deleted, to define the phenotypic consequences. This analysis can be aided by completing the deletion in a strain in which marker proteins are tagged by GFP, or similar moieties. This type of study might require generating a strain with genes modified with GFP and mRFP, for example, followed by deletion of a third target gene.

In theory, multiple gene modifications can be completed using different selective gene markers. For instance, the related *phoA* and *phoB* genes encode protein kinases with overlapping functions and deletion of either does not cause lethality. In an attempt to determine the defects caused by double deletion, one was deleted using the *pyroA* marker and the other with the *pyrG* marker. Subsequent crossing of the two deleted strains revealed that the double mutant was not viable and caused a block in cell cycle progression.\(^{24}\) It is thus possible to utilize several markers for generating multiple gene modifications but it is also possible to reutilize the same marker multiple times.

Two basic approaches have been taken in order to reutilize transformation markers. Both approaches rely on utilizing a transformation marker that provides both positive and negative selection and incorporation of repeated DNA sequences either side of the transformation marker. In the first approach, the direct repeats act as targets for spontaneous recombination and do not have to have a defined sequence.\(^{27,28}\) Typically ~300 bp is sufficient. In the second approach, the direct repeat sequences are composed of 34 bp *loxP* sites which, recombine\(^{29,30}\) when acted upon by Cre recombinase. In both approaches, after recombination between the flanking repeated sequences the marker gene is excised from the genome, leaving behind a single copy of the repeated sequence and a strain in which the transformation marker can be reutilized.

For the nonspecific DNA direct repeat system to work, no additional modifications of the transformation strain are necessary. However, in the Cre/loxP-based method, it is necessary to express the
Cre recombinase gene of bacteriophage P1 in the recipient strain. This has been achieved in *A. nidulans* by placing expression of Cre under control of the *A. nidulans* xylose-inducible, glucose-repressible *xlnA* promoter and using *trpC* sequences for termination. In *A. fumigatus*, expression of Cre has been placed under control of the promoter and termination sequences of *A. nidulans* *niaD*. In addition, the *A. fumigatus* Cre-expression cassette has been placed into a nonintegrating, self-replicating plasmid (see Section 30.2.6 later) which is readily lost during mitotic divisions, enabling easy removal of the Cre-expressing gene. In the *A. nidulans* Cre system, it would be necessary to cross the transformation strain in order to remove the Cre-expressing gene. In both *A. nidulans* and *A. fumigatus* the Cre-expressing clones should be readily transferable to other aspergilli and other fungi.

In the *A. nidulans* Cre system, it has been noticed that, at a low frequency, loop-out via recombination between the *loxP* sites occurs without the necessity for expression of Cre recombinase. Because there is positive selection for the loop-out event (i.e., negative selection for the marker gene flanked by the *lox* sites—5-FOA toxicity for *pyrG* in *A. nidulans* system and FUDR toxicity for the ble-HSV1 *tk* marker in the *A. fumigatus* system) a very low level of loop-out can be selected for. Therefore, at least for *A. nidulans*, it may not be necessary to go to the extra effort to express the Cre recombinase in recipient transformation strains to remove markers flanked by *loxP* sites.

### 30.2.4.5 Considerations Regarding Accuracy of Gene Calling and Gene Manipulations

The genomic sequence of numerous members of the aspergilli have been determined, and are in the process of being sequenced (http://www.genomesonline.org/), which provide priceless resources that will revolutionize research on this most important group of fungi. All of the gene manipulations described in this chapter rely on the accurate prediction of gene structures and the protein encoding regions of the genome. To date, this has largely been done using automated gene prediction pipelines, although more recent efforts have concentrated on additional manual genome annotation (http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/News.html). However, given the number of genome sequences available, or soon to be available, it is unlikely that all aspergilli genomes will undergo intensive manual annotation. At this time, it is known that a significant percentage of predicted gene structures have errors, especially those deposited at public databases such as NCBI. More details are provided in the Chapter 3 by Fedorova et al.

It is wise to critically question the accuracy of predicted gene structures when considering any gene modifications based upon this information because of the potential for misinformation regarding gene structures. The definitive gene structure has to be derived from cDNA sequence. Often, cDNA sequence information is available and defined at NCBI, if the cDNA sequence has been published and/or submitted to NCBI. In addition, several of the databases available for specific sequences, such as at the Broad Institute site for *A. nidulans* (http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/Home.html), provide cDNA and EST sequence comparisons to the genome sequence, thus providing definitive gene structure data for the region of sequence overlap.

It is advisable to identify the latest gene structure predictions available when considering any gene manipulation experiments because gene annotation is a continuous process, constantly evolving. As an example, the annotation of *A. nidulans* has undergone three major revisions resulting in significant reorganization of predicted gene structures. Thus, the original Broad Institute data set consisted of 9541 protein-coding genes but this number has increased to 10,701 due to improved gene annotation (http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/News.html). The improved gene structures will not be available for some period at NCBI although they are available at the Broad Institute. Therefore, for each sequenced *Aspergillus* species, it is recommended that the latest version of the predicted gene structures are referred to at the respective sequencing centers.

All of the gene manipulations described in this chapter depend on the accurate prediction of either the start codon or the stop codon or both the start and the stop codon. For example, for accurate and specific gene deletions, the region of the genome from the start codon to the stop codon is replaced by a transformation marker. This requires that both the start and stop codons are defined. For endogenously tagging proteins at the C-terminus, the stop codon needs to be defined. For endogenously tagging proteins at their
N-terminus, the start codon needs to be defined. Finally, in order for the expression of any gene to be placed under control of a regulatable promoter, for overexpression or promoter rundown experiments, the start codon needs to be known as well as the limits of the nearest upstream gene. Thus, for promoter replacement, information about two gene structures is required.

It is often possible to gain convincing evidence that the start site and stop site are accurately predicted in the absence of cDNA sequence. If the predicted protein is highly conserved and BLASTP searches using the predicted protein sequence demonstrate high levels of identity with many orthologs at the N- and C-termini, this would strongly suggest that the predicted gene structure is correct. For less well-conserved proteins, this approach becomes problematic. However, by comparing predicted gene structures from related species, as is possible with the annotated genome sequences of several of the aspergilli, it is often possible to gain convincing evidence that a predicted gene structure is either wrong or, more than likely, correct. This is particularly so if one of the gene structures is based upon cDNA sequence.

If convincing evidence is lacking after predicted gene comparisons, it is possible to readily define the 5′ and 3′ limits of mRNA sequences by cDNA sequencing. This can be done using the RACE procedure (rapid amplification of cDNA ends) or by utilizing DNA from a high-quality cDNA library.

In the RACE protocol, cDNA is synthesized but not cloned. After total cDNA synthesis, a gene-specific primer based upon the predicted gene structure and a second primer based upon a constant sequence used in the RACE ready cDNA synthesis are used in PCR reactions to selectively amplify cDNA corresponding to the target gene. Different gene-specific primers are used to amplify the 5′ and 3′ cDNA ends. Typically, a specific band is amplified which is visible after agarose gel electrophoresis of the PCR products. If not, a second round of PCR is carried out using a second “nested” gene-specific primer and the first PCR products as template. When a specific cDNA band has been amplified, it is cut from the gel and cloned into a suitable vector for sequence analysis. The sequence analysis provides definitive data regarding the start and stop site of the target gene.

An alternative approach is to generate two primers based upon the predicted gene structure that will prime DNA synthesis before the predicted start codon and after the predicted stop codon. The primers can also be based upon the 5′ and 3′ RACE analysis. It is possible to use such primers and the RACE-ready cDNA to amplify a full length cDNA. However, because the RACE-ready cDNA is not cloned, it is typically in limited supply. Therefore, assuming a high-quality cloned cDNA library is available, the PCR reaction can use amplified DNA made from the cloned cDNA library as template. After amplification, the cDNA can be sequenced directly or cloned and subsequently sequenced.

If sequence analysis reveals that the expected open reading frame (ORF) terminates at a stop codon in the 3′ region of the cDNA, then this defines the stop codon. If the predicted initiation methionine is the first initiation codon, and if there are stop codons in the same reading frame 5′ to the initiation codon, it is safe to conclude that the correct initiation codon has also been defined. However, in some instances the reading frame of the predicted initiation codon is open upstream. In such instances, it is necessary to complete 5′ RACE to ensure a full length 5′ cDNA sequence has been isolated because of the possibility that the predicted initiation codon corresponds to a methionine encoded within the protein if the encoded protein is longer than that predicted.

### 30.2.5 Confirming Gene Targeting

Once a particular gene manipulation has been completed, it is important to confirm that the desired change has been achieved and that additional changes have not occurred within the genome of the modified strain. In all of the gene manipulations described in this chapter, a marker gene is incorporated in a stable manner into the genome at the site of the gene that is modified. This will normally have the effect of making that region of the genome larger or smaller than the wild type. In order to confirm that a gene-modifying construct has landed site specifically in *S. cerevisiae*, it is typical to utilize diagnostic PCR to amplify a region of the modified genomic area. Two primers are utilized: one that will prime within a unique sequence of the introduced DNA cassette, such as the selection marker and the other is designed such that it will prime upstream of the introduced cassette and, if the cassette has landed site-specifically, will amplify DNA in PCR reactions containing the other primer...
Advances in Gene Manipulations Using Aspergillus nidulans

(Fig. 30.1a, primers 1 and 2). However, if no site-specific integration has occurred, no PCR product is generated.

In theory, similar diagnostic primers can, and have been, used to determine if specific gene modifications have occurred in A. nidulans. However, it has been shown, at least in Ku+ strains, that linear DNA constructs often become circularized before integrating by a single site specific cross-over event.6 In such reactions, the genomic site is increased in size proportional to the size of the circular DNA molecule rather than generating just the desired gene replacement. What is more, using the primer pair as described earlier (Fig. 30.1a) will amplify the expected sized band even though the desired modification alone has not been achieved. It is, therefore, better to design primers for the diagnostic PCR that will prime within the genomic region at either side of the region replaced by the modifying cassette (Fig. 30.1b, primers 3 and 4). Using such primers, a positive result is obtained both in the wild type and in the gene-modified strain. However, the gene-modified strain should not amplify the wild type size band but instead amplify a predictable size band, typically larger than the wild type, although for deletions the band can be smaller.

For gene modifications involving tagging proteins with either fluorescent proteins or with small affinity tags, it is also advisable to complete Western blot analysis to confirm that protein of the correct size is being synthesized. Antibodies for most tags are commercially available.

Finally, Southern blot analysis has historically been used to confirm correct gene replacements in the aspergilli. If designed correctly, Southern blot analysis has the advantage over diagnostic PCR approaches in that not only can the correct gene replacement event be confirmed but the presence of other ectopic integrations can also be detected. Such ectopic events typically involve nonspecific integrations of plasmid constructs via nonhomologous recombination. By utilizing Ku-deleted strains in which NHEJ is suppressed, such ectopic events have been almost eliminated as revealed by Southern blot analysis of many different gene-modified strains in Ku-deleted backgrounds.9

30.2.6 Nonintegrative Gene Expression Utilizing the AMA1 Sequence

In addition to landing DNA constructs by stable integration into the host genome, it is also often useful to employ expression constructs that do not integrate. By transforming a genomic library cloned into an integrative plasmid vector, the lab of John Clutterbuck isolated a plasmid from an unstable transformant which had the ability to replicate without integration into the genome. Subsequent experiments revealed that the plasmid contained an A. nidulans sequence which consisted of an inverted repeat termed the AMA1 replicator.35–37 The AMA1 replicator has the ability to maintain plasmids as extrachromosomal elements which never integrate into the host genome. This has the effect of increasing the frequency of transformation more than 250-fold. AMA1-containing plasmids support replication and expression of introduced genes and provide the ability to readily cure strains of the plasmid because the plasmids are naturally lost during asexual spore (conidia) formation such that >50% of spores lack the plasmid.38–40 These features have made AMA1 plasmids of great use in the aspergilli as all these features are also displayed when AMA1 plasmids are introduced into other species of Aspergillus.25

One of the main uses of AMA1 plasmids stems from the extremely high rate of transformation possible. This likely indicates that one of the limiting factors during Aspergillus transformations is the integration event. This bottle-neck is circumvented on using AMA1 plasmids, as they never integrate and once present in nuclei can give rise to transformants. The lab of Greg May has produced several genomic libraries in AMA1 plasmids that provide a selection marker for transformation which have been used for gene complementation experiments in A. nidulans41 and A. fumigatus.42–44 The advantage of this system is not only the very high rates of transformation but also the ease by which the complementing plasmid can be reisolated in E. coli. Once isolated, vector-based primers can be used to sequence into the insert and candidate-complementing genes identified in the corresponding genome sequence. Another advantage is afforded because of the copy number of AMA1 plasmids, which has been estimated to be 10–15 copies per nucleus. Therefore, in addition to cloning the wild-type copy of the mutated gene that is complemented, high copy number suppressors can also be isolated.45

Another application of AMA1 plasmids is the transient expression of genes. It is sometimes beneficial to express a certain gene function for a period and then remove it. This can be achieved using regulatable
promoters to turn gene expression on and then off. An alternative approach is to put the desired gene to be expressed on an AMA1 plasmid, either with constitutive expression or using a regulatable system. The plasmid can be introduced by transformation and the gene expressed in the host strain. Then, in order to remove the plasmid, the strain is allowed to undergo asexual spore formation. During this developmental stage, AMA1 plasmids are lost from nuclei, which removes the expression of the introduced gene. For instance, in the *A. fumigatus* Cre-lox system referred to earlier (in Section 30.2.4.4) the expression of the Cre recombinase is placed under control of a regulatable promoter cloned within a selectable AMA1 plasmid. Once Cre had been induced to promote excision of the marker gene flanked by *loxP* site, conidia from the strain were streaked on media nonselective for the AMA1 plasmid. Many resulting colonies had lost the AMA1 plasmid and, therefore, were unable to express Cre. In this way, Cre was expressed to fulfill its role in the excision reaction but then removed to prevent any potential phenotypic consequences of continued cre expression in subsequent experiments. 30

### 30.3 Specific Types of Gene Manipulations

#### 30.3.1 Gene Deletion and Promoter Rundown

When studying the function of a gene, it is desirable to manipulate the gene in several different ways. One of the most common and vital modifications is deletion of the gene to ask if it is essential or if the deletion causes any phenotypes, such as developmental or conditional phenotypes, for example. The null allele is important not only regarding studies of gene function but also because it can be used as a tool to garner evidence that tagged versions of the encoded protein are functional. For example, for essential genes, if an endogenously tagged version of the gene, such as a GFP chimera, grows normally, it is safe to conclude the tagged version of the protein is functional. Because such analysis proves the GFP chimera is functional, its location within the cell is more likely to reflect the normal location of the wild type protein. Many genes are nonessential but their deletion can often cause a phenotype. For example, deletion of DNA repair genes or spindle assembly checkpoint genes does not cause lethality but does generate sensitivity to DNA-damaging agents or microtubule poisons, respectively. In such instances, it is possible to test strains with the tagged chimera for sensitivity to such drugs and if wild-type resistance is seen the tagged versions of the proteins can be inferred to be functional.

To delete a gene, it is necessary to target the gene with a deletion or disruption construct such that after homologous recombination the target gene is removed from the genome or rendered inactive. The easiest construct to make that disrupts the function of a gene is a plasmid containing a selectable marker for transformation and an internal fragment of the target gene. After a single site-specific integration event, the target gene is disrupted leaving 5′ and 3′ truncated, and hopefully nonfunctional, copies of the gene. 46 Although technically easy to accomplish, such approaches do not necessarily generate a clean null allele of the target gene. This is because of the potential for generating a poison peptide, or a truncated but functional gene, using this approach. However, targeted linear deletion cassettes, when integrated correctly, provide an unequivocal null allele (as will be discussed in Section 30.3.1.1).

One use of disrupting a gene with an internal fragment is to target plasmids to defined locations within the genome. This can be done very conveniently by targeting constructs to γA or to wA using plasmids with a heterologous selectable marker (*Af-ribob* for e.g.) and an internal fragment of the color gene. When integrated specifically at the color gene, the integration causes disruption of the color gene and correctly targeted plasmids thereby result in transformant colonies with white or yellow conidia rather than the wild type green color. This allows visual detection of the correct transformants and makes it easy to follow the construct in crosses. The approach also allows *in vitro* manipulations of a gene cloned into the vector and repeated targeting to the same defined place in the genome (Hynes and R. Genovese, and S.L. Murray, unpublished). Transformants require checking by Southern blot because integration of more than one plasmid copy at the targeted locus is common.
To generate a null allele using fusion PCR, a linear DNA deletion cassette is generated whereby a transformation marker is flanked by targeting sequences homologous to the DNA sequence upstream and downstream of the start and stop sites of the gene to be deleted (Fig. 30.1b). Any selectable marker should suffice. If a nutritional marker is used for the transformation selection, it is advisable to use a marker cloned from a different Aspergillus species from the one to be transformed, such as Af-pyrG, Af-pyroA, or Af-riboB to transform A. nidulans. As mentioned earlier, this decreases the chance that the deletion cassette will be targeted to the marker site rather than to the site of gene deletion. It is also worth noting that the expression of nutritional markers can be affected when integrated at certain target sites. Therefore, any phenotype generated after deletion using a nutritional marker should also be confirmed with addition of the nutritional supplement.

When a nonessential gene is deleted, it is easy to propagate the null allele and confirm the deletion using diagnostic PCR and/or Southern blot analysis. If the deleted gene is essential then it cannot be propagated as a haploid. One approach to get around this is to complete the deletion in a diploid strain. It is then possible to confirm the null allele and wild type allele are present in the diploid. In order to “prove” that the null allele is essential, the diploid is broken down to the haploid state. This is typically done by propagating the heterozygous null diploid on media containing microtubule poisons which promotes haploidization. If the marker gene that was used to delete the target essential gene is not recovered in the resulting haploid strains, this is taken to mean that the gene deleted is essential. This approach is not ideal because the conclusion that a gene is essential comes from negative data. In addition, it is not possible to gain any information regarding the phenotypic consequences of the gene deletion using this approach. Both of these problems can be circumvented by one of two approaches; (1) using the heterokaryon rescue technique or (2) putting the gene under control of a regulatable promoter and complete a promoter rundown.

### 30.3.1.1 Gene Deletion and the Heterokaryon Rescue Technique

As in all filamentous fungi, A. nidulans cells can maintain many nuclei within a common cytoplasm. Typically, these nuclei are genetically identical but, if appropriate selection is imposed, two genetically distinct nuclei can coexist in the same cell. Such cells are called heterokaryons. During the deletion procedure, the deletion cassette DNA is introduced to protoplasts which can be multinucleate. Upon integration of the deletion cassette, the target gene is deleted (ΔgeneX) and at the same time that nucleus becomes positive for the marker gene (e.g., pyrG+). Such a nucleus could not support growth because the essential gene is deleted. However, nontransformed nuclei in the same cytoplasm would be geneX+ but pyrG−. When selection for the transformation marker is imposed, it is, therefore, possible to select for heterokaryons where the deleted nuclei provide the pyrG+ function and the nondeleted nuclei the essential geneX+ function. Such heterokaryons readily form after deletion of essential genes in A. nidulans. The beauty of the heterokaryon rescue technique is that when A. nidulans undergoes asexual development to form spores, called conidia, the heterokaryotic state is not propagated because conidia contain a single nucleus. Instead, two types of conidia are formed from the parent heterokaryon. These are, pyrG+ geneX− spores and pyrG− geneX+ spores. Therefore, by plating the conidia from a heterokaryon on to selective media for the pyrG marker the pyrG+ geneX− spores cannot germinate or grow due to lack of pyrG function. However, the pyrG+ geneX+ spores can germinate because they are pyrG+ but will arrest growth when the essential gene function becomes limiting. For instance, deletion of cell cycle–specific functions allows short-term growth but not mitotic division and so the null allele terminal phenotype is a short germling with a single cell cycle arrested nucleus. By using the heterokaryon rescue technique it is possible to define not only that a gene is essential but also to define the specific defects caused by lack of the essential gene. For a more in-depth description of the heterokaryon rescue technique and a detailed protocol see Ref. 49.

### 30.3.1.2 Promoter Rundown

An alternative approach to generate the equivalent of a null allele of essential genes is to employ the promoter rundown technique. This is achieved by replacing the promoter region of the target gene with one that is conditionally regulatable. This has been achieved in A. nidulans using the alcA-regulatable
promoter sequences.\textsuperscript{50,51} The desired regulation can be achieved by generating a plasmid construct in which the 3' truncated coding region of the target gene is cloned downstream of the \textit{alcA} promoter. When the plasmid lands at the target site via a single homologous recombination event, the endogenous promoter will express the truncated gene which, hopefully, will be inactive. The full-length coding sequence of the target gene is under control of the \textit{alcA} promoter (Fig. 30.2a). Again, this is a relatively easy manipulation but there are two potential problems with this approach. First, the truncated version of the gene under the endogenous promoter potentially can be expressed and either function or produce an interfering protein. Second, because the plasmid sequence is flanked by direct repeats of the target gene the integration is not stable and can loop-out via a reversal of the recombination event that integrated the plasmid. This is particularly true during genetic crosses when recombination is enhanced.

An alternative approach is to use fusion PCR to generate a linear replacement cassette which, when integrated, puts the target gene under control of the regulatable promoter. In this approach, there is no chance of the endogenous promoter expressing any portion of the target gene, and there are no direct repeats generated upon integration. Therefore, the gene modification is completely stable. Suitable core constructs have been generated\textsuperscript{14,52} that include a transformation marker 5' to the \textit{alcA} promoter region.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure30.2.png}
\caption{Generation of promoter replaced strains using plasmid or linear constructs.\textit{Note}: (a) A plasmid construct is generated with the target gene promoter replaced by that of \textit{alcA} (\textit{alcA}-p). The target gene in the plasmid is also 3' truncated. After homologous recombination between the 3' truncated target gene in the plasmid with its homologous genomic region, the plasmid is integrated generating two versions of the target gene which flank the integrated plasmid sequence. The 5' version of the target gene is regulated by the endogenous promoter but is 3' truncated and thus should not be expressed. The 3' version of the target gene has its expression regulated by the \textit{alcA} promoter and expresses a full-length protein. (b) A gene-replacement cassette is generated using fusion PCR containing, from 5' to 3', a targeting domain with homology to the 5' region of the target gene—the \textit{Af-pyrG} selectable marker—the \textit{alcA} promoter—a targeting domain with homology to the target gene coding region. After homologous recombination between the targeting domains and the genome the expression of the target gene is placed under control of the \textit{alcA} promoter.}
\end{figure}
Advances in Gene Manipulations Using Aspergillus nidulans

Using fusion PCR, it is, therefore, possible to target the pyrG-alcA cassette such that the alcA promoter is placed before the coding region of the target gene and thus puts the gene under control of alcA (Fig. 30.2b). In addition to the published work, this approach has been used by our labs to control the expression of several genes which, when placed in repressive media, generate the expected phenotype of a null allele (Osmani and Oakley, unpublished). However, it should be noted that even on rich glucose-containing media the alcA promoter is not completely off, and some low level of transcription occurs. For this reason, especially for genes that typically have low levels of expression and/or encode particularly stable proteins, it is sometimes not possible to recapitulate the effects of the null allele by placing the gene under alcA and turning off the promoter.54

Another highly regulatable expression system for the aspergilli has been developed based upon the human estrogen receptor.55 The human estrogen receptor’s ability to regulate transcription depends on its binding to activating ligands, such as estrogen, which causes an allosteric change in the receptor, which promotes location of the receptor to nuclei where it binds to estrogen-responsive elements present in the promoter regions of target genes. The binding of the activated receptor to the estrogen-responsive elements turns on expression of the gene. This system has the advantage that the induction can be completed using glucose-containing media, but its utility for doing promoter rundown experiments is yet to be tested. Other potentially usable promoters include a xylP promoter-based expression system, derived from the Penicillium chrysogenum endoxylanase gene, which puts expression of target genes under induction by xylose and repression by glucose56 and the A. oryzae thiA promoter which puts gene expression under control of thiamine levels in the growth media.57

30.3.2 Fluorescent Protein Tagging

Another powerful tool commonly used to investigate the function of a protein is to tag proteins with GFP or a similar fluorescent protein. This aspect of A. nidulans gene manipulations is covered in Chapter 31 by Oakley and Xiang.

30.3.3 Affinity Tags for Protein Purifications and Proteomics

Another exceptionally useful type of gene modification is the addition of an affinity tag via homologous recombination. There are several widely used affinity tags available which are, typically, relatively short amino acid sequences that have the ability to bind reversibly and specifically to a second protein or other types of molecule. By incorporating the affinity tag at the C- or N-terminus of a target protein, using molecular genetic approaches, it is possible to subsequently biochemically purify the tagged protein in a single affinity-purification step. The process will be described for the S-tag that has proved to be particularly well suited for protein purification of proteins from A. nidulans (and Liu and Osmani unpublished).

When purifying an affinity-tagged protein it is advisable to work with an endogenously tagged strain so that artifacts associated with over- or underexpression of the tagged protein are avoided. In addition, by working with an endogenously tagged protein it is possible to determine that the tagged protein is functional (as described earlier). A core cassette has been developed and deposited at the FGSC (pAO81) which can be used to generate targeted S-Tagged gene-replacement cassettes using fusion PCR (Section 30.2.2). One nice feature of this cassette is that the same primers used to generate GFP fusions can be reutilized to complete the S-Tagging. This is because all of the core C-terminal tagging cassettes incorporate a repeated glycine-alanine linker. This has been incorporated because a flexible linker has been found to enhance the functionality of tagged chimeras.59 In addition, the cassettes that have been generated utilize the A. fumigatus pyrG gene. Thus, all of the core tagging cassettes have the same sequence at their 5’ and 3’ ends enabling primers designed to generate constructs to land GFP at the C-terminus of a gene to be used to generate constructs to introduce mRFP or the S-Tag. Future generations of cassettes that utilize other transformation markers will also be designed to incorporate the same sequences to make the primers for any C-terminal tagging event universal between all of the tags. Another advancement being developed is the incorporation of loxP sites flanking the A. fumigatus pyrG gene in these constructs (Liu and Osmani, unpublished). This will allow loop-out of the pyrG marker after a gene has been GFP-tagged to
enable subsequent tagging or gene deletions to be carried out reusing the pyrG marker (see Section 30.2.4.2).

Once the appropriately S-Tagged strain has been developed and confirmed as functional, the tagged protein can be purified using single-step affinity purification. The components required for the S-Tag affinity-purification system are commercially available (Novagen). The S-Tag consists of 15 aa which bind with high affinity to the S-Protein which can be purchased linked to Agarose beads. Protein extracts generated from the S-Tagged strain are incubated with S-Protein agarose beads. The S-Tagged protein will bind with high affinity to the S-Protein bound to the beads. Subsequent washing of the beads removes all unbound proteins, but the S-Tagged protein and its binding partners remain bound to the beads. The purified proteins can then be released from the beads by boiling in SDS sample buffer, or alternatively, the beads can be exposed to an excess of the S-Peptide to release the S-Tagged protein in an active form. This approach is suitable for maintaining the enzymatic activity of purified proteins.

As an example of the utility of the S-Tag affinity-purification approach, it is possible to purify to homogeneity μg quantities of nuclear pore complex proteins of A. nidulans from protein extracts made from 1000 ml of A. nidulans culture. The purified protein, and binding partners, can readily be identified using mass spectroscopy analysis of Coomassie-stained bands from SDS PAGE gels (Liu and Osmani, unpublished).

In addition to the S-Tag, there are many other affinity tags that should work in A. nidulans and could readily be incorporated into core fusion PCR cassettes. In addition to single-step affinity tags, in some systems two affinity tags are attached to a target protein, and two affinity purifications are completed in series. These dual-affinity tags are commonly termed TAP-Tags (tandem affinity protein). In certain instances, a TAP-Tag purification could be superior to a single-affinity purification. The lab of Gerhard Braus has generated a suitable TAP-Tag for use in the aspergilli which has been optimized with the codon usage of A. nidulans to prevent the TAP-Tag from affecting translation of the chimera protein (Gerhard Braus, personal communication).

The generation of affinity-tagging systems for A. nidulans sets the stage for large scale analysis of the proteome of this species. With the availability of high-quality sequence and gene predictions, it is possible to utilize mass spectroscopy to identify purified proteins and their binding partners. All of the tagging technologies developed should be readily transferable to other aspergilli and perhaps other filamentous fungi. In addition, purified proteins can also be identified without specific affinity purifications as recently demonstrated using mass spectroscopy to identify proteins of the A. fumigatus conidia cell wall.

30.3.4 Two-Step Site-Specific Mutation

During functional gene analysis, the need to introduce a specific mutation can arise. As an example, cyclin-dependent kinase 1 (Cdk1) is required for entry into mitosis in all eukaryotes. However, the gene encoding this kinase was not identified during a screen for temperature-sensitive mutations affecting A. nidulans cell cycle progression. The Cdk1 gene of A. nidulans (nimX) was subsequently cloned and deleted to reveal it to be essential. It was possible to introduce specific missense mutations into nimX and thus generate temperature-sensitive strains that arrested the cell cycle at the restrictive temperature because of several temperature-sensitive alleles of this kinase having been isolated in the S. pombe homolog (cdc2) and the fact that this is a highly conserved protein. In a more recent example, a missense mutation in the H2A gene has been introduced to convert a phosphorylatable serine residue to a nonphosphorylatable amino acid to demonstrate the site is essential for the DNA damage response.

Several methods have been employed to introduce point mutations into the A. nidulans genome. In the method termed a two-step gene replacement, a plasmid copy of the gene is subjected to mutation using a standard in vitro mutagenesis protocol. In addition to the genomic copy of the gene (with promoter and 3’ processing sequences), the plasmid should also contain a selectable marker for transformation. With the availability of genome sequences, it is easy to generate a cloned gene in an appropriate plasmid via PCR amplification. When the mutated gene is transformed and has integrated via homologous recombination, the plasmid sequence is integrated which is flanked by repeated copies of the target gene, one of which is wild type, the other the mutant allele (Fig. 30.3). For the next step, it is useful if the transformation marker can be selected against, although this is not completely necessary. At some frequency, if
selection for the transformation marker is removed, mitotic recombination occurs between the repeated sequences leading to a loop-out of the plasmid sequence (Fig. 30.3). This event can be scored as loss of the transformation marker. If the transformation marker provides negative selection, such as 5-FOA sensitivity for pyrG, asexual spores from colonies grown without selection can be plated on 5-FOA containing plates to select for those which have looped out the marker. If the marker does not provide negative selection the strain can be put through a self-cross during which loop-out events occur at an elevated frequency. Once a strain is cured of the plasmid, it can be tested for the presence of the mutated allele. This is straightforward if a phenotype is caused by the mutation, such as temperature or DNA-damage sensitivity, as described earlier for nimX and H2A alleles. If not, the locus can be amplified using PCR and sequenced to determine if the wild type or mutant allele exists.

Another variation on this approach has been recently described using the split marker concept. In this approach, two DNA fragments are generated using fusion PCR. Both fragments contain the target gene in which the desired mutation has been introduced. One fragment contains the 5′ region of the marker gene fused to the 3′ end of the target gene. The other fragment contains the 3′ region of the marker gene fused to the mutated genes 5′ end (see Fig. 2 in Nielsen et al., 2006). The two regions of the marker gene overlap and, in the published case, the pyr4 gene of N. crassa was utilized. When the two sequences integrate via homologous recombination, the desired outcome is a direct repeat of the mutated gene flanking

![Diagram of gene replacement process](image.png)

**FIGURE 30.3** Two-step gene replacement to generate new alleles. *Note:* A copy of the target gene in a selectable plasmid is mutated using standard *in vitro* mutagenesis methodologies. After transformation into a recipient strain, homologous recombination generates two copies of the target gene flanking the selectable marker on the plasmid. Subsequent mitotic recombination can occur between the flanking repeated genes and, depending on where the crossover occurs, the mutant allele can be left in the genome as indicated. This loop-out event can be selected for, if there is negative selection against the transformation marker on the plasmid.
the reconstituted and functional pyr4 gene. In reality, when the homologous recombination event occurs, depending on where the actual crossovers occur, only one of the resulting alleles may have the mutation.\textsuperscript{16} Upon loop-out of the pyr4 gene, as described earlier, at some frequency, the mutated allele will be left in the genome if both a wild type and mutant copy are present. If both copies are mutant then all loop-out events will leave just the mutant allele.

### 30.4 Conclusions

The ability to manipulate genes in \textit{A. nidulans} and other aspergilli has undergone a recent revolution. The most important advance that has facilitated this revolution is the availability of high-quality genome sequence of numerous aspergilli. To fully utilize the untapped wealth of information encoded within the genome sequence of these organisms, new methods for gene manipulations have been developed, in part based on the pioneering work completed in \textit{N. crassa}. The only limit to our ability to fully realize the potential of the aspergilli as allies in basic research, biotechnology, and the food industry is our imagination and resolve. Similarly, with the ability to manipulate the genomes of this most important genus at will, it is hoped that imaginative research and continued efforts will enhance our capacity to control the growth of aspergilli when they become pernicious opportunistic pathogens.

### Acknowledgments

SAO and BRO would like to acknowledge support from the National Institutes of Health and MJH support from the Australian Research Council. We would also like to thank all members of our labs, past and present, for their contributions to the development of many of the experimental approaches outlined in this chapter.

### References


The Aspergilli

53. Bussink HJ, Osmani SA. A cyclin-dependent kinase family member (PHOA) is required to link developmental fate to environmental conditions in Aspergillus nidulans. EMBO Journal 17:3990–4003, 1998.
31

Fluorescent Labels for Intracellular Structures and Organelles

Berl R. Oakley and Xin Xiang

CONTENTS
31.1 Introduction ............................................................. 513
31.2 Histological Organelle-Staining Methods and DNA-Binding Fluorochromes ........................................ 514
31.3 Vital Dyes .................................................................... 514
31.4 Immunofluorescence Microscopy ................................................................. 514
31.5 Fluorescent Protein Tagging ................................................................. 515
  31.5.1 Labeling with Multiple Fluorochromes ........................................ 516
  31.5.2 Additional Techniques Using Fluorescent Proteins .................. 517
  31.5.3 Fluorescent Protein Tagging Approaches ........................................ 518
    31.5.3.1 Plasmid-Based Approaches ....................................................... 518
    31.5.3.2 Fusion PCR-Based Approaches .................................................. 518
    31.5.3.3 Use of Flexible Linkers Between the Target Protein and the
      Fluorescent Protein ............................................................................ 521
    31.5.3.4 Gene Targeting Using nkuA Deletion Strains ......................... 521
31.6 Conclusions and Prospects ............................................................. 522
Acknowledgments .................................................................................. 522
References ............................................................................................. 522

31.1 Introduction

The sequencing of the genomes of Aspergillus nidulans, Aspergillus fumigatus, and Aspergillus oryzae has ushered in a new era in research with these organisms. In addition, the development of robust and efficient gene targeting, and fusion PCR protocols has reduced to days or weeks tasks that formerly required weeks or months. Nowhere is this more apparent than in the imaging of intracellular structures and organelles. Tagging of proteins with fluorescent moieties [green fluorescent protein (GFP), monomeric red fluorescent protein (mRFP) and others] is now rapid and efficient. Genes encoding proteins of interest can be identified from the genome sequence data and tagged using procedures that do not even require the gene to be cloned in the conventional sense. This allows proteins to be localized accurately in living cells and the movement of these proteins to be followed over time. Organelles can be observed by tagging proteins specific to those organelles with fluorescent moieties. In addition, immunofluorescence microscopy, which has been used for decades, benefits greatly from the ability to epitope-tag proteins rapidly.

In this review, we will discuss older organelle-labeling procedures briefly, focusing on methods that remain valuable. We will devote the bulk of our efforts to more recently developed methods involving protein tagging based on the genome sequence. In particular, we will discuss the creation of fluorescently
tagged proteins that allow in vivo imaging. We will focus on the model species *A. nidulans*, but it is likely that many of the techniques developed for *A. nidulans* will carry over to other species of *Aspergillus*.

### 31.2 Histological Organelle-Staining Methods and DNA-Binding Fluorochromes

Although they are rarely used today, we would be remiss if we did not mention historically important staining procedures for chromosomes and mitotic spindles such as Feulgen, HCl aceto-orcein, HCl Giemsa, and acid fuchsin staining. These procedures are laborious in comparison with current procedures, but they were instrumental in the identification of mitotic and cell cycle regulatory mutants in *A. nidulans*.² With the advent of fluorescence microscopy, it became possible to visualize chromosomes and nuclei with DNA-binding fluorochromes such as mithramycin and DAPI (4,6-diamidino-2-phenylindole).³,⁶ These were much more rapid and convenient than the older histological stains and remain useful today.

### 31.3 Vital Dyes

Fluorescent vital dyes have been used very effectively to label organelles in fungi including *A. nidulans*. These are particularly useful in instances in which tagging proteins with epitope tags or fluorescent proteins are not effective (e.g., instances in which one wishes to image membranes). The use of such dyes in filamentous fungi in general has been comprehensively reviewed by Hickey et al.⁷ We will restrict our comments to dyes that have already been used in *A. nidulans*, although it is highly likely that other dyes that have proved useful in other fungi will also prove useful in *A. nidulans*.

Among the earliest vital dyes used to stain an organelle in *A. nidulans* was Rhodamine 123, which labels mitochondria and this dye continues to be useful.⁶,⁸ Rhodamine 123 is a cell-permeant dye, the fluorescence of which depends on the electrochemical gradient across the mitochondrial membrane.⁹ Rhodamine 123, thus, only fluoresces at active mitochondria. Another dye that has been used effectively to label mitochondria in living cells is mitotracker.¹⁰,¹¹ The vital membrane dye FM4-64 has also proved to be useful in *A. nidulans*.⁸,¹²,¹³ Although it is a general membrane dye, it can be used in “load and chase” experiments to visualize endocytosis. If it is added to a culture at 0ºC, FM4-64 is incorporated into the plasma membrane, but is not internalized. If it is then washed out and the culture warmed such that the cells become physiologically active (e.g., 25ºC), the internalization of portions of the plasma membrane can be observed. Other variations on this procedure are also useful (e.g., adding FM4-64 to medium for 2 min at room temperature, then washing it out). In addition the lumens of vacuoles can be stained in living cells with 5-(and-6)-carboxy-2,7-dichlorofluorescein diacetate (CDCFDA).⁸ Finally, filipin may be used to stain sterol-rich membrane domains. In *A. nidulans*, filipin highlights a prominent patch at hyphal tips and a ring at septation sites.¹⁴

### 31.4 Immunofluorescence Microscopy

For two decades, the standard approach for localizing proteins in *A. nidulans* (and, thus, the organelles or cellular structures with which they are associated) was immunofluorescence microscopy. This approach employs antibodies (primary antibodies) that bind specifically to the target protein in fixed and permeabilized cells. These antibodies are then imaged by fluorescence microscopy. Rarely, fluorochromes such as fluorescein isothiocyanate (FITC) or tetramethylrhodamine isothiocyanate (TRITC) are coupled directly to the primary antibody. Much more commonly, a fluorescently labeled secondary antibody is used. For example, a FITC-labeled goat antibody that specifically binds to rabbit antibodies might be used as a secondary antibody to detect a primary antibody made in a rabbit.

This approach has required finding or producing a primary antibody that binds specifically to the protein of interest. Some antibodies against proteins from mammals or other phyla cross react with, and
fluorescent labels for intracellular structures and organelles

are specific for, the *A. nidulans* homologs of those proteins. In many cases, however, antibodies made against animal or plant proteins simply do not react with *A. nidulans* proteins. Historically, this meant that the *A. nidulans* protein had to be purified or its gene cloned and expressed in bacteria (usually as a fusion protein). Antibodies then had to be raised in one’s animal of choice and tested for specificity. This approach generally took months of effort and since animals often have fungal infections, sera often contained antibodies against a number of *A. nidulans* proteins in addition to the targeted protein. As a consequence, it was often necessary to affinity purify antibodies to obtain a preparation adequately specific for localization of the protein of interest.

After sufficiently pure antibodies were obtained, the immunofluorescence procedure was, itself, not trivial. It was necessary to digest the cell wall and the cell wall digestive enzymes contained proteases that, if not inhibited, destroyed the protein of interest. A variety of procedures have been developed, however, that give very good results (e.g., Ref. 15).

Many of the limitations of immunofluorescence microscopy can be overcome by using epitope tags. In this approach, a DNA sequence that encodes a short (typically 6–25 amino acids), defined sequence of amino acids is added to the gene that encodes target protein (normally at the C-terminus or N-terminus) by molecular genetic means. A large number of epitope tags have been developed (e.g., C-myc, Flag, HA, V5, etc.) and, in principle, most or all of them can be used in *A. nidulans*. A major advantage of this approach is that epitope tagging a gene is generally much quicker than developing an antibody. Antibodies against epitope tags are available commercially from many sources, although if commercially available polyclonal antibody preparations are used, it is useful to test them to make certain that they do not contain antibodies reactive against *A. nidulans* proteins. The use of epitope tags is greatly facilitated by the sequencing of the genome and the development of rapid and efficient gene-targeting procedures. The sequencing of the genome allows the genes that encode proteins of interest to be identified easily and improvements in gene-targeting procedures allow the target proteins to be epitope tagged rapidly and easily. These procedures are discussed later in our section on creating fluorescent protein fusions.

### 31.5 Fluorescent Protein Tagging

One of the most powerful techniques in cell biology is the use of fluorescent protein fusions to observe proteins in living cells and this is certainly true of *A. nidulans*. The underlying principle is that DNA encoding a fluorescent protein, such as the GFP, is fused in frame to the gene encoding the protein of interest and introduced into the *A. nidulans* genome by transformation. Expression of the gene produces a fusion protein that fluoresces. Since the introduction of this technique by Reinhard Fischer’s group and John Doonan’s group to the *A. nidulans* community, a variety of GFP-fusion proteins have been made (Table 31.1). The fluorescent protein can be fused to the N-terminus or C-terminus of the target protein, or be inserted into the target protein, although this is rarely done.

Fluorescent proteins are much larger than the fluorochromes used for immunofluorescence microscopy. GFP, for example, is 238 amino acids in length and will comprise a substantial fraction of the volume and mass of the fusion protein. It is, thus, somewhat surprising that a large fraction of GFP fusion proteins are functional. Nevertheless, this is clearly the case. Table 31.1 lists fluorescent protein fusions that have been reported for *A. nidulans* along with the organelles and/or cellular structures that have been labeled with the fusion proteins. The list is long and growing. In addition, Nayak et al. have reported that in attempts to tag the C-termini of 28 proteins with GFP, 24 of the fusion proteins were functional, a frequency greater than 85%. If C-terminal fusions are not functional, N-terminal fusions may be, and there are instances in which a fusion of a protein to one fluorescent protein (e.g., GFP) is not functional while a fusion to another fluorescent protein (e.g., mCherry) is functional. It seems safe to say that it will be possible to create functional fluorescent protein fusions with the majority of *A. nidulans* proteins.

Fluorescent protein fusions have many advantages over other protein labeling techniques. First, if expression levels of the protein are normal and the fusion protein is functional, it is relatively safe to assume that one is seeing the real location of the protein in the cell. With immunofluorescence microscopy, there is always a possibility that the distribution of the target protein has changed during fixation.
The Aspergilli

516

The Aspergilli

TABLE 31.1
GFP-Fusions Used to Label Organelles/Cellular Structures in A. nidulans

<table>
<thead>
<tr>
<th>Organelles and/or Cellular Structures</th>
<th>Fusion Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei</td>
<td>GFP-Gal4 (DNA-binding domain),10 GFP-H1,48 CFP-H2A,30 BimG-GFP,49 BimD-GFP,60 An-Rec1-GFP,66</td>
</tr>
<tr>
<td>(interphase)</td>
<td>GFP-StuA(NLS),53 mRFP1-StuA(NLS),15 BFP-StuA(NLS),13 GFP-PacC(5-250),51 An-Trm1-GFP,66</td>
</tr>
<tr>
<td>(mitosis)</td>
<td>An-Nup2-GFP,66</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>Citrate-Synthase (N-terminus)-GFP,10 MdmB-GFP,22 Acu1-GFP,11 IdpA/MTS-RFP,11 EchA-GFP,55</td>
</tr>
<tr>
<td>Peroxisomes</td>
<td>Acu-E-GFP,21 GFP-IdpA/PST1,11 RFP-FoxA,53</td>
</tr>
<tr>
<td>Endoplasmic reticulum</td>
<td>GFP-plant ER retention signal,19 ShrA-sGFP,54</td>
</tr>
<tr>
<td>Plasma membrane/cortex</td>
<td>PmB-GFP,55 UpC-GFP,56 ApsA-GFP,57</td>
</tr>
<tr>
<td>Vacuoles</td>
<td>TpmA-GFP,14</td>
</tr>
<tr>
<td>Actin</td>
<td>GFP-TubA,21,58 CFP-TubA,11 GFP-KipB,24</td>
</tr>
<tr>
<td>Mitochondria plus ends</td>
<td>GFP-NUDA,48 GFP-NUDF,21 GFP-NUDI,60 GFP-NUDM,62 GFP-NUDK,26 GFP-CLIPA,27</td>
</tr>
<tr>
<td>Spindle pole bodies</td>
<td>GFP-MipA,21 YFP-MipA,11 GFP-SNAD,28 BIM-GFP,49 GFP-PLKA,55 GFP-MOBI,21</td>
</tr>
<tr>
<td>Kinetochores</td>
<td>GFP-Nck80,49</td>
</tr>
<tr>
<td>Nuclear pores (mitosis)</td>
<td>NIMA-GFP,25</td>
</tr>
<tr>
<td>Septa</td>
<td>GFP-MYOA,50 SEPA-GFP,48 BIMG-GFP,49 GFP-MOBI,21 GFP-BudA,29 GFP-ApsB,57</td>
</tr>
</tbody>
</table>

and antibody labeling procedure. Second and more importantly, it is possible to view movement of proteins in living cells. For one example (of many), it is possible to observe microtubule dynamics in A. nidulans using GFP fusions to the microtubule protein α-tubulin.21 Microtubules are very dynamic structures, constantly growing and shrinking, and the dynamics of microtubules are important to their function. With GFP-tagged microtubules, it is possible to observe microtubule dynamics and determine how these dynamics are affected by mutations in proteins such as members of the dynein complex. This is simply impossible with immunofluorescence microscopy.

Finally, as we alluded to earlier, the sequencing of the genome greatly facilitates fluorescent protein tagging. Once target genes have been identified, the sequence information greatly facilitates the construction of plasmids or linear molecules for making fluorescent fusion proteins. In many cases the great majority of genes whose products are involved in particular processes can be identified on the basis of homology to proteins in other organisms. For example, over the past several years, A. nidulans researchers have taken advantage of the A. nidulans genome information to identify and study proteins involved in cell biological problems such as septation, polarized growth, nuclear pore dynamics, and microtubule dynamics.22–29,66

31.5.1 Labeling with Multiple Fluorochromes

The development of multiple fluorescent proteins with different emission spectra now allows two or more proteins to be labeled and observed in the same cell (see supplemental Fig. 31.1a on CD). This is often
important; for example, if one can tag a microtubule protein with one fluorescent protein and an organelle with another fluorescent protein (with a different emission wavelength), one can determine if an organelle is moving along a microtubule by straightforward dual wavelength time lapse microscopy.

Fortunately, a number of useful fluorescent fusion proteins are currently available for use in *A. nidulans*. Many fusions have been made using a version of GFP in which codons have been altered to give brighter and more stable fluorescence.\(^\text{19}\) This GFP variant has, in turn, been subjected to *in vitro* mutagenesis to produce versions of Cyan Fluorescent Protein (CFP) and Yellow Fluorescent Protein (YFP) that work well in *A. nidulans*.\(^\text{30,31}\) mRFP a monomeric derivative of DsRed from *Discosoma* sp. has also been used in *A. nidulans* as has DsRed and the Blue Fluorescent Protein (BFP), a GFP derivative.\(^\text{20,32,33}\) The utility of DsRed is limited, however, because it forms a tetramer *in vivo*. It, thus, tends to link proteins to which it is attached into complexes. Recently, Shaner et al. have created a family of derivatives of mRFP with a variety of absorption and emission spectra and with improved folding, brightness, and photostability characteristics.\(^\text{34}\) Two of these, tdTomato and mCherry, have been tested and found to function well as protein tags in *A. nidulans* (Edgerton et al. 2006 unpublished) and it is likely that others will be equally useful. mCherry is a clearly superior alternative to mRFP. It is brighter, folds faster and fades more slowly.\(^\text{34}\) Among the many fluorescent proteins that are available for use in *A. nidulans*, there is enough variation in absorption and emission spectra, that it is possible to image two, or, in principle, more, proteins in the same cell. CFP- and YFP-labeled proteins have been imaged together,\(^\text{30,31}\) and the combinations of GFP and mRFP or GFP and mCherry also work well (Edgerton unpublished). In principle, combinations such as CFP, YFP, and mCherry should allow three proteins to be imaged simultaneously.

### 31.5.2 Additional Techniques Using Fluorescent Proteins

Fluorescent protein tagging has not only become a tool for observing cellular structures, but new techniques using fluorescently tagged proteins also allow one to study protein–protein interaction and dynamics in live cells. For example, photobleaching of fluorescent proteins in specific cellular areas (FRAP and FLIP) allows protein dynamics to be monitored.\(^\text{35}\) In FRAP, a fluorescent protein is bleached in a small region using a relatively intense light (often a laser) and the recovery of fluorescence in the bleached region is monitored over time. Fluorescence recovery is due to unbleached molecules of the protein moving into the bleached region and thus provides a good, quantifiable index of the dynamics of the protein. In FLIP, a defined region is bleached repeatedly with brief bursts of relatively intense light and the fluorescence intensity of a nonbleached region is monitored. Reduction in fluorescence of the
unbleached region is caused by movement of fluorescently labeled molecules out of the region that is not replenished by fluorescently labeled molecules moving in, and again provides a useful measure of the temporal and spatial dynamics of the protein. Such studies can, in principle, be carried out using a variety of fluorescent protein tags.

Another powerful technique, fluorescence resonance energy transfer (FRET), allows one to determine if two proteins interact in the cell.35,36 This technique relies upon the fact that energy from an excited fluorophore can be transferred to a different fluorophore that is physically very close and the second fluorophore will emit energy at its characteristic emission wavelength. Thus if one wishes to determine if two proteins interact in vivo, one protein can be tagged with CFP and the other with YFP. The specimen is illuminated at the absorption wavelength for CFP (440 nm) and if there is significant output at the emission wavelength for YFP (535 nm) energy must have transferred from CFP to YFP. This energy transfer only occurs over very short distances (typically 3–6 nm). Thus for energy transfer to occur the two proteins must be very close in the cell.

Bimolecular Fluorescence Complementation (BIFC) using a split GFP system can also be used to study protein–protein interaction.37 Recently, Reinhard Fischer’s group has modified the split GFP system so that it could be used to study protein–protein interactions in A. nidulans.38 In this system, each protein is fused to half of the GFP molecule. Neither half GFP is capable of fluorescing on its own, but if the two proteins under study interact physically, the two half GFP moieties will be brought together producing a functional GFP. Thus, one obtains fluorescence only if the proteins interact.

31.5.3 Fluorescent Protein Tagging Approaches

31.5.3.1 Plasmid-Based Approaches

Until recently, most fluorescent protein tagging in A. nidulans involved transformation with plasmids constructed by standard cloning procedures (Fig. 31.1). This approach can be used to tag the N-terminus or the C-terminus. In both cases, all or a portion of the target gene must be cloned and fused to the fluorescent protein sequence. If the plasmid contains the entire coding sequence of the target gene as well as the promoter, integration of the transforming plasmid by homologous recombination will result in two copies of the target gene, one tagged with the fluorescent protein and one untagged. If the plasmid carries only a portion of the gene, integration will result in one functional tagged copy of the gene and a second, partial copy.

N-terminal tagging using plasmids carries an additional complication with respect to promoters. If the normal promoter of the gene is to be used, the fluorescent protein must be inserted between the promoter and the coding sequence of the gene such that it is fused in frame to the coding sequence and is under the control of the promoter. Creating such a precise construct can be difficult and time-consuming, although PCR should, in principle, facilitate the process. A second approach is to create a plasmid in which the fluorescent protein is under the control of a regulable promoter. The highly regulable alcA promoter is often used.21,33,39 The coding sequence, or a portion thereof, can be inserted such that it is in frame with the fluorescent protein-coding sequence.39 This can be greatly facilitated by the use of the GATEWAY system that allows insertion by homologous recombination in vitro.33 When regulable promoters are used the levels of expression of the fusion proteins are a significant concern. The alcA promoter is highly inducible and overexpression of the fusion protein (relative to normal levels) may, in some cases, alter the localization patterns of the protein and/or affect the cellular processes in which the protein is involved. With care, this can be overcome by using a nonrepressing, noninducing carbon source such as glycerol or by carefully balancing the levels of inducer and repressor (e.g., fructose, a weak repressor and threonine, a moderately strong inducer). In any case, since the endogenous levels of different proteins differ greatly, one set of conditions will not give good results for all proteins. In each case, it is important to test the effects of different degrees of induction on growth and localization to guard against misleading results.

31.5.3.2 Fusion PCR-Based Approaches

The advent of fusion PCR cloning procedures in combination with the sequencing of the genome, has greatly facilitated gene tagging and other molecular genetic techniques.40–42 With fusion PCR, one does
not construct plasmids or even use ligases. Rather, one PCR amplifies the fragments to be cloned together using synthetic primers such that the amplified fragments have ends with complementary sequences (Fig. 31.2). One then fuses the fragments together by mixing the fragments and amplifying with primers that anneal to the ends of the fusion PCR product.

This approach readily allows linear molecules to be created that can be used to create C-terminal fluorescent protein fusions. We will use the mipA (γ-tubulin) gene as an example (Fig. 31.2). One first amplifies a portion of the gene and a region immediately downstream of the gene. The amplified fragments are typically about 1000–2000 bp in length. Other sizes may be used, but fusion PCR becomes

![Diagram](image-url)

**FIGURE 31.2** Fusion PCR and creation of a C-terminal GFP fusion by transformation with a fusion PCR product. Note: In panel a, a portion of the *mipA* (γ-tubulin) coding sequence and the 3′ untranslated region (UTR) are amplified from genomic DNA with primers that are based on the genomic sequence. Primers P2 and P3 have “tags” that are not from the genomic sequence but, rather, are identical to end regions of the GFP-AfpyrG cassette that will be fused to this fragment. In panel b, a previously constructed and amplified cassette consisting of the GFP-coding sequence and the *Aspergillus fumigatus* pyrG gene (AfpyrG), which is used as a selectable marker, is mixed with the fragments amplified in panel a as well as primers P1 and P4. PCR then creates a linear molecule consisting of a portion of the *mipA* gene fused in frame to the GFP coding sequence followed by AfpyrG and the 3′ UTR. In practice, replacing P1 and P4 with nested primers (new primers that bind slightly inside P1 and P4) usually reduces the amplification of bands other than the desired product. In panel c, this fragment is used to transform. Homologous recombination results in a C-terminally tagged *mipA* gene.
The Aspergilli

less efficient as size increases and targeting efficiency drops off as size decreases. These fragments can be amplified directly from genomic DNA with primers designed using the genomic sequence. Next, these fragments are mixed with a cassette that carries the fluorescent protein sequence and a selectable marker. The primers used to amplify the fragments from the genome are designed with “tails” that are identical to the ends of the cassette. When the three fragments are mixed and amplified using primers that anneal to the ends of the outside fragments, the resulting amplification product is a linear molecule in which the fluorescent protein is fused in frame to the C-terminus of the target gene. The molecule also contains a selectable marker that can be used to select transformants, and each end of the molecule is homologous to the target region of the chromosome. When such fragments are used to transform, homologous recombination results in a single copy of the target gene, under the control of its normal promoter but fused to a fluorescent protein at its C-terminus (Fig. 31.2).

This approach has a number of advantages. First, conventional cloning is not necessary. The transforming fragment is produced by two sequential PCR amplifications and these can be carried out very quickly. The central cassette does not need to be amplified each time. Enough DNA can be prepared in a single fusion PCR reaction for many subsequent fusion PCR reactions. Second, this process scales up easily because little more work is required to carry out 20 PCR reactions than one. It is easy to prepare 10 or more fusion PCR products at a time and tag 10 or more proteins in a single round of transformations. A third advantage is that the transformant should carry a single copy of the gene and it is the tagged version that cannot be lost by recombination. With plasmid tagging procedures recombination during mitosis or meiosis can cause loss of the fluorescent protein tag. A fourth advantage is that the tagged gene is under the control of its normal promoter and, consequently, expression levels of the gene should be normal in most cases. A fifth advantage is that since only one copy of the gene is present, if the gene is essential and the transformants are robust, one can conclude with confidence that the fusion protein is functional.

There are some disadvantages to this approach as well. First, there is the possibility that PCR will introduce mutations into the transforming fragment. This potential problem can be minimized by using proofreading PCR enzymes that have low error rates. In addition, the problem is self-correcting to a significant extent. PCR reactions produce many molecules, most of which do not have errors and a smaller fraction that do have errors. The fraction having errors will be greater if a PCR enzyme with a higher error rate is used or if a mistake happens in an early PCR round, such that all subsequent descendant molecules carry the mistake. If one transforms with DNA produced by PCR, most transformants should be transformed with a molecule with the correct sequence. A fraction, normally small, will be transformed with molecules that carry mutations caused by PCR errors. If the targeted gene is essential for viability, any mutations that significantly inhibit the function of the targeted protein will result in dead transformants, which will not be recovered or sick transformants, which are easily detected and excluded from further analysis. If the targeted gene is not essential more caution is warranted. It is often useful to observe the fluorescent protein localization pattern in a number of transformants. If they all give the same localization pattern one can have reasonable confidence that the observed pattern is not due to PCR-induced mutations. (The fusion protein could be dysfunctional because of the addition of the fluorescent protein, however, but this is a problem common to all tagging approaches.)

N-terminal tagging using fusion PCR is fairly straightforward if one wishes to simultaneously tag the gene and place the gene under the control of a regulable promoter. In this case, the cassette used for fusion PCR consists of a selectable marker followed by the regulable promoter and the fluorescent protein. The fluorescent protein is fused in frame (by fusion PCR) to the coding region of the targeted gene. The other flanking region is upstream from the gene. N-terminal tagging while retaining the normal promoter is more complex but possible. One separately amplifies a 5' flanking region upstream of the normal promoter, a selectable marker, the normal promoter, the fluorescent protein-coding sequence, and a region of the coding sequence of the target gene beginning with the 5' end of the coding sequence, all with primers that give appropriate overlapping regions. One then fuses them together in a single fusion PCR reaction creating a linear molecule with segments in the aforementioned order, and transforms a target strain. Although conceptually complex, this is operationally straightforward, takes only about two days, and has worked surprisingly well (Oakley, et al., 2006 unpublished).
31.5.3.3 **Use of Flexible Linkers Between the Target Protein and the Fluorescent Protein**

With plasmid or fusion PCR-based targeting procedures, a concern is that the fusion protein may be dys- functional or only partially functional. The fluorescent protein is relatively large, after all, and it is not surprising that it can interfere with the function of the protein to which it is attached. One approach that is helpful in solving this problem is to use a flexible “linker” of several amino acids (e.g., a series of five glycine/alanine repeats) between the target protein and the fluorescent protein.\textsuperscript{41} This allows the two proteins to move freely relative to each other and may improve the likelihood that the fusion will be functional.

31.5.3.4 **Gene Targeting Using nkuA Deletion Strains**

One problem with transformation with plasmids or fusion PCR products has been that in many cases the transforming DNA inserts into the genome by nonhomologous recombination. Thus, instead of inserting at the correct site and producing a correct fusion protein, it integrates elsewhere. In addition, multiple integrations may occur during transformation such that even if one has the correct fusion, one may have additional insertions of the transforming DNA elsewhere. These heterologous insertions are a nuisance at best and, at worst, can cause a misleading phenotype by insertional mutagenesis. In addition, linear molecules may circularize during transformation. These problems have been largely solved by the development of strains that carry a deletion of \textit{nkuA}, the \textit{A. nidulans} homolog of the human KU70 gene.\textsuperscript{20} The KU70 protein and the protein encoded by the KU80 gene form a heterodimer that is involved in DNA repair by nonhomologous end joining and Ninomiya et al. demonstrated that deletion of the \textit{Neurospora crassa} KU70 and KU80 homologs greatly increased the frequency of correct gene replacement.\textsuperscript{43} In \textit{A. nidulans}, deletion of \textit{nkuA} dramatically reduces the frequency of nonhomologous integration as well, although, interestingly, deletion of \textit{nkuA} and \textit{nkuB} (the KU80 homolog), singly or in combination does not increase sensitivity to mutagens including those that cause double strand breaks.\textsuperscript{20} \textit{A. nidulans}, thus, must have a second double strand break repair system that can function to a certain extent when the \textit{nkuA/nkuB} system is inoperative, presumably the homologous repair system. It is worth noting that deletion of the KU70 and/or KU80 homologs in other species of \textit{Aspergillus} also improves gene targeting and, in general, does not increase sensitivity to mutagens or have detrimental effects on growth.\textsuperscript{44–47}

One potential problem with the use of \textit{nkuA} deletion strains is that if the gene used as a selectable marker is from \textit{A. nidulans}, the transforming DNA can integrate by homologous recombination into the chromosomal copy of the selectable marker gene (if the mutant, chromosomal allele of the selectable marker gene is not a deletion). This problem was solved by the cloning of genes from \textit{A. fumigatus} that complement mutant \textit{A. nidulans} alleles but have sufficiently low homology with the \textit{A. nidulans} alleles that they do not direct integration at any particular site. Correct integration will, thus, be directed by the homologous DNA on the transforming linear DNA molecule or plasmid. A glufosinate resistance gene from \textit{Streptomyces hygroscopicus} also works well as a selectable marker in this system, and has the advantage that since it is a dominant drug-resistance marker, it does not require that selectable mutations be crossed into the recipient strain.\textsuperscript{20}

Using \textit{nkuA} deletion strains and nonhomologous selectable markers, approximately 90% of transformants carry a single, correct homologous integration.\textsuperscript{20} The strains give high frequencies of correct integration with linear molecules and with circular plasmids. When linear molecules are used for transformation, 500 base pairs of homologous flanking DNA at each end of the fragment is enough to ensure a high frequency of homologous integration, but the transformation frequency is relatively low. As the sizes of the homologous flanking DNA stretches at the ends of the molecules increase, the transformation frequency increases. This presumably reflects the fact that the \textit{nkuA} deletion minimizes heter- ologous integration and the frequency of homologous integration is a function of the sizes of the regions available for homologous recombination to occur. When using fusion PCR to generate transforming fragments, the size of the flanking regions is a compromise. Longer flanks give higher transformation frequencies and shorter flanks mean that the fusion PCR fragment is shorter and easier to construct. One thousand base pair flanks are a reasonable compromise (Nayak et al., 2006, unpublished) allowing easy
construction of transforming fragments and high transformation frequencies. Recently, the nkuA deletion strain has been successfully used for making multiple deletion strains and GFP or mCherry-tagged strains to study the dynamics of the nuclear pore proteins during mitosis.

Although the nkuA deletion does not appear, in general, to cause detectable alterations in the localization pattern of most proteins, it is important not to forget that it may alter the localization patterns of some proteins or interact synthetically with some fluorescent fusion protein alleles. When this is a concern, one can cross the fluorescent fusion protein allele into an nkuA wild-type (nkuA+) background.

31.6 Conclusions and Prospects

The sequencing of the Aspergillus genomes along with recent advances in the development of fluorescent proteins and gene targeting have ushered in a new era in live imaging in Aspergillus. Projects that formerly required years can now be completed in weeks. Projects that were impossible a few years ago are now carried out routinely. Looking ahead, it is likely that fluorescent fusion proteins that tag all major organelles will be available in the near future. The time is ripe, moreover, to begin to carry out genome-wide gene-targeting projects in A. nidulans. Indeed, a genome wide project to fluorescently tag all A. nidulans proteins is within the realm of possibility.

Acknowledgments

We would like to acknowledge the continuing support from the National Institute of General Medical Sciences. We would also like to thank C. Elizabeth Oakley for her help in preparing the manuscript and Drs. Miguel Peñalva and Reinhard Fischer for helpful comments on the manuscript.

References

Fluorescent Labels for Intracellular Structures and Organelles


Aspergillus at the Fungal Genetics Stock Center

Kevin McCluskey

CONTENTS
32.1 Fungal Genetics Stock Center ................................................................. 527
  32.1.1 Introduction and History .............................................................. 527
  32.1.2 Strains .................................................................................... 528
  32.1.3 Strain Distribution ................................................................... 531
  32.1.4 Molecular Materials ................................................................. 532
  32.1.5 Information Resources ............................................................ 533
32.2 Conclusions ...................................................................................... 534
References .............................................................................................. 535

32.1 Fungal Genetics Stock Center

32.1.1 Introduction and History

The Fungal Genetics Stock Center (FGSC) was established in 1960 at Dartmouth College following a survey carried out by the Committee on the Maintenance of Genetic Stocks of the Genetics Society of America, which recommended that a repository be established to protect strains of Neurospora and other important filamentous fungi. According to the Aspergillus Newsletter, there were about a dozen research laboratories working with Aspergillus at the time. The FGSC has supported these and other labs by holding and distributing strains of Aspergillus and, in recent years, molecular materials to work with Aspergillus.

The FGSC has moved three times in its history. It moved from Dartmouth to California State University at Humboldt in 1975 when Dr. R. Barratt accepted a position as Dean there. In 1984, the FGSC moved to the University of Kansas Medical Center when Dr. Barratt retired as Director of the FGSC. Dr. J. Kinsey was the new director and Mr. C. Wilson took over as curator, a position previously held by Mr. B. Ogata. In 2004, the FGSC moved again, this time to its present home at the University of Missouri-Kansas City where Dr. M. Plamann is the director. The FGSC has continued to expand its mission and currently holds strains of Aspergillus, Neurospora, Fusarium, Magnaporthe, and other fungi.

The materials in the FGSC collection have always been considered to be in the public domain. Other collections are Patent Depositories, according to the Budapest Treaty but the FGSC has never endeavored to become a Patent Depository. The FGSC does, however, meet most of the criteria for acquisition of the status of International Depositary Authority, including key issues such as having continuous existence, necessary staff and facilities, and operating with impartiality and confidentiality.

The FGSC has always sought to provide materials without regard to the ability of a recipient to pay; we have tried to keep our fees as low as possible. As of the writing of this article, the fee for a single strain was US $20 for an academic laboratory and US $50 for a commercial laboratory. There are decreasing fees for academic labs such that the fee for 100 strains would only be US $500. These fee caps have
The Aspergilli

not historically been applied to commercial laboratories. By way of contrast, the fee for the Aspergillus strain FGSC A4 [the American Type Culture Collection (ATCC) 38163] from the ATCC is US $192 and it is accompanied by a 2100-word Material Transfer Agreement. The ATCC is not, however, a genetic repository and while the FGSC has many strains of only a few species of fungi, the ATCC has thousands of species in its collection. Moreover, the ATCC is a Global Bioresource Center serving clients in many diverse fields.

The FGSC is a member collection in the World Federation for Culture Collections and the U.S. Federation for Culture Collections. As such the FGSC is part of a global effort to preserve and make available strains and research materials developed by researchers from all over the world. In 2002, it was estimated that there were 76 fungal culture collections in the world, housing over 385,000 cultures.3 Many such collections have a narrow focus or are mainly for the use of researchers at the particular institution or country. Nevertheless this represents a tremendous resource for research in biological sciences. The FGSC Aspergillus collection is clearly part of a larger body of work with implications for global health and agriculture.

32.1.2 Strains

The first Aspergillus nidulans strain was deposited in October of 1962 by Dr. Etta Kafer with the genotype biA1, choA1.4 This strain, FGSC A1, has been distributed by the FGSC six times. Reflecting the changes in the ability to identify mutations, the genotype of this strain was later updated to reflect both the presence of a suppressor of sB and the translocation T1(I;VII). Ninety-five Aspergillus strains were deposited by Dr. A. Clutterbuck while 461 were deposited by Dr. E. Kafer. Many of these strains duplicate strains in the Glasgow collection: in total, 218 strains are crosslisted with the Glasgow collection while 828 are not.

The one-hundredth A. nidulans stock was entered into the collection in June of 1963. This strain, FGSC A100, had the genotype Acr1, w3; meth1. This strain was retired in March of 1974 and replaced when a new group of strains carrying w3, Acr1, and cha became available. A100 was distributed 26 times and was sent to places such as Harvard University, Ontario, Canada, the University of Liverpool, University of Hong Kong, University of Adelaide, University of Leningrad, USSR, and the Hebrew University of Jerusalem, demonstrating the early global appeal of the resources at the FGSC. Many of the Aspergillus strains in the collection have detailed histories such as this.

The FGSC collection of Aspergillus stocks now includes 1046 active strains of which 878 are A. nidulans and 129 are A. niger. Most strains of Aspergillus at the FGSC are mutant strains and only a few are wild-type. Other collections for wild or industrial strains of Aspergillus exist, such as the U.S. Department of Agriculture collection at the National Center for Agricultural Utilization Research (formerly the Northern Regional Research Laboratory), which holds at least 665 Aspergillus stocks, the American Type Culture Collection, the Centraalbureau voor Schimmelcultures which holds nearly 800 isolates, or the Deutsche Sammlung von Mikroorganismen und Zellkulturen, which holds strains from 39 different Aspergillus species.3 Many of the Aspergillus strains in the FGSC collection are also held in the Glasgow collection managed by A. J. Clutterbuck. This collection houses approximately 650 stocks and those that are not available from the FGSC are available from the BioCentrum-DTU at the Danish Technical University in Lynby. Growth of the Aspergillus collection at the FGSC has been rather slow with many researchers preferring to share strains individually. This is reflected in Table 32.1. The FGSC acquired a set of 1150 temperature-sensitive strains of A. nidulans in 2000.5 These have not been assigned FGSC numbers for a number of reasons including the fact that they are primary mutants and have not been demonstrated to be free of secondary lesions and because they are a set to be used together.

Among the genetically marked Aspergillus strains in the FGSC collection, 37 have only 1 marker while 178 strains have 2 markers (including veA+). Two hundred and sixty strains have 3 markers while 168 strains have 4 markers (Fig. 32.1). Fifty-eight strains have 10 or more markers. The greatest number of markers is 15, present in strains A591 and 593. These are special-purpose strains for generating diploids. The average Aspergillus strain has eight markers and the most common markers are biA1 and ya2 (Table 32.2). Most strains are in the veA background and so this lesion is typically mentioned only when it is absent (veA+).
In the early and mid-1980s the FGSC received a series of *A. niger* mutants. There are 124 total mutants in this collection in addition to the wild type from which they were derived. This strain, FGSC A732 (also known as N400), was first deposited in the FGSC collection in September of 1986, but has been resubmitted twice by various genome-sequencing organizations (Table 32.3). It is also known as FGSC A1121 (deposited by J. Yu) and A1143 (deposited by S. Baker). The more common background is N402 (FGSC A733), which contains the *cspA* marker for ease of use in the laboratory. All 124 of the *A. niger* mutants in the FGSC collection carry this marker. One hundred and fifteen different markers are present in this group of strains with *fwnA1* being present in 40 strains and *pdxA2* and *nicA1* being the next most common lesions. They are both found in 25 of the *A. niger* strains.

Other mutant *Aspergillus* strains in the FGSC collection include *A. flavus* strains carrying markers for the study of aflatoxin production (A1009 and A1010) as well as *A. fumigatus* strains. Among the latter are strains carrying *pyrG1, argB1,* and *lysB1* for transformation, and more recently strains deleted at the *KU80* locus for ease of homologous integration (FGSC A1160).6

The FGSC catalog lists strains in a number of categories to simplify its use. The first category is a listing of lesions that includes 137 different categories based on requirement, resistance, or morphology. Since one category can include multiple loci or alleles, this is only one representation of the breadth of the collection. There are 869 different loci represented among strains in the FGSC collection. This number includes 29 mapped translocation break points. These are organized in the catalog in their own section. The most common translocation is T1(VI;VII), which appears in 15 strains. Overall there are 75 strains with known translocations and 62 strains for which the presence of translocation has not been evaluated.

The limited selection of *Aspergillus* wild-type strains include the *A. nidulans* Glasgow wild-type as well as 31 additional strains. Most are *A. nidulans*, although the FGSC has taken on various strains from sequencing programs. These include *A. fumigatus, A. niger, A. flavus,* and *A. terreus* (Table 32.3).

Historically, groups of strains have been constructed for special purposes. These include the mitotic- and meiotic-mapping strains. The former are comprise 43 strains useful for mapping using parasexual

---

**TABLE 32.1**

Recent Growth of the *Aspergillus* Collection at the FGSC

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of strains added</td>
<td>25</td>
<td>22</td>
<td>73</td>
<td>8</td>
<td>34</td>
<td>40</td>
<td>23</td>
<td>4</td>
<td>1154</td>
<td>33</td>
<td>2</td>
<td>18</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

**FIGURE 32.1** Numbers of markers per *Aspergillus* strain in the FGSC collection.
TABLE 32.2
Most Common Markers

<table>
<thead>
<tr>
<th>Locus</th>
<th># Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>biA1</td>
<td>326</td>
</tr>
<tr>
<td>yA2</td>
<td>285</td>
</tr>
<tr>
<td>pabaA1</td>
<td>203</td>
</tr>
<tr>
<td>aceA1</td>
<td>169</td>
</tr>
<tr>
<td>pyroA4</td>
<td>133</td>
</tr>
<tr>
<td>chaA1</td>
<td>130</td>
</tr>
<tr>
<td>adeE20</td>
<td>128</td>
</tr>
<tr>
<td>cspA1</td>
<td>124</td>
</tr>
<tr>
<td>wA3</td>
<td>106</td>
</tr>
<tr>
<td>choA1</td>
<td>82</td>
</tr>
<tr>
<td>riboA1</td>
<td>80</td>
</tr>
<tr>
<td>sB3</td>
<td>71</td>
</tr>
<tr>
<td>riboB2</td>
<td>65</td>
</tr>
<tr>
<td>suA1 adeE20</td>
<td>62</td>
</tr>
<tr>
<td>ActA1</td>
<td>54</td>
</tr>
<tr>
<td>proA1</td>
<td>52</td>
</tr>
<tr>
<td>nicA2</td>
<td>51</td>
</tr>
<tr>
<td>pyrG89</td>
<td>49</td>
</tr>
</tbody>
</table>

genetics while the latter are comprise 69 strains. Among these are strains useful for mapping on each of the eight linkage groups.

Other special purpose strains include color mutants, paired diploids with all homologs marked, teaching strains, translocation mapping strains, and diploids and triploids used for mapping centromeres. The last group of special-purpose strains is the largest, consisting of 1150 temperature-sensitive mutants from Dr. S. Harris.5

The next main catalog division is the numerical listing of stocks. In this section, strains are listed in ascending numerical order with their full genotype and additional information. Because of the historical nature of the FGSC catalog (it has evolved from the strain listings published in the Aspergillus Newsletter), some of the genotypes contain outdated gene names. The numerical listing does not include the large group of A. niger mutants that are listed separately, as are A. awamori, A. oryzae, A. flavus, A. hetero-thallicus, and A. fumigatus strains. An ongoing effort at the FGSC is to move away from the historical document that comprises the FGSC catalog and generate the catalog directly from the FGSC strain database. One goal in doing this is not to lose the information about the use of strains in the categories of the catalog. For example, catalog section V.8. comprises six strains used to select for mutants by starvation for biotin. This information is not presently available on the FGSC website and as we move toward more electronic access, we will build a mechanism to allow for searches that include this information.

TABLE 32.3
Aspergillus Strains from Sequencing Programs

<table>
<thead>
<tr>
<th>Species</th>
<th>FGSC Number</th>
<th>Other Number</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. nidulans</td>
<td>A4</td>
<td>M139</td>
<td>Broad Institute</td>
</tr>
<tr>
<td>A. fumigatus</td>
<td>A1100</td>
<td>AF293</td>
<td>TIGR</td>
</tr>
<tr>
<td>A. flavus</td>
<td>A1120</td>
<td>NRRL3357</td>
<td>Aspergillusflavus.org</td>
</tr>
<tr>
<td>A. niger</td>
<td>A1121</td>
<td>NRRL3</td>
<td>—</td>
</tr>
<tr>
<td>A. niger</td>
<td>A1143</td>
<td>NRRL3, ATCC 9029</td>
<td>DOE JGI</td>
</tr>
<tr>
<td>A. niger</td>
<td>A1144</td>
<td>NRRL 328, ATCC 1015</td>
<td>—</td>
</tr>
<tr>
<td>A. terreus</td>
<td>A1156</td>
<td>NIH2624</td>
<td>Broad Institute</td>
</tr>
</tbody>
</table>
Aspergillus at the Fungal Genetics Stock Center

Strains at the FGSC are maintained in a number of ways. Historically, all strains are grown on appropriate medium upon arrival at the FGSC and are tested for an identifiable phenotype. They are then preserved in two separate formats. Most strains are preserved as freeze-dried spores as described by Wilson and on anhydrous silica gel. While this is robust and reliable, it does not support the preservation of strains that do not sporulate vigorously and so we also preserve aconidial or morphological mutants in 25% glycerol at –80°C and over liquid nitrogen. One hundred and forty-five Aspergillus strains are stored in cryopreservation while 901 are only stored in silica gel and as freeze-dried spores. Among the strains stored in cryopreservation are strains with markers such as snf, fasA, acoA49, or fluG701. While we make every effort to assure that strains are available in perpetuity, some strains do die in storage. When this happens, we try to contact the original depositor and request a replacement. Fortunately, this happens very infrequently and the most common problem that we see with strains is that they are damaged in shipment, either by being subject to dessication or to extreme temperatures.

Unlike other collections, when a strain is requested from the FGSC, it is revived on appropriate medium and when significant growth is evident it is sent by express courier. Because some recipients prefer to receive strains through mail, we have, in certain cases, sent stocks as conidia spotted on filter paper rather than in culture tubes. While some of the strains in the FGSC collection are technically clinical isolates, the fact that they are common environmental fungi makes it possible to send them, albeit in double-walled mailers, without specific clinical pathogen packaging according to International Air Transport Authority regulations. For details please see the 46th edition of the International Air Transport Authority Dangerous Goods Regulations and Addenda II and III (dated March 22, 2005 and July 5, 2005, respectively), which are based on the 13th revised edition of the UN Model Regulations (see www.wfcc.info).

Ongoing work at the FGSC includes periodic strain testing whereby we grow each strain and then verify its genotype. For strains with multiple markers, we attempt to verify each individual marker. This was last carried out for the entire Aspergillus collection in 2005. Only 5 strains out of 1046 were not viable at that time and four of these strains had never been ordered.

32.1.3 Strain Distribution

For a number of reasons, the number of Aspergillus cultures that the FGSC distributes is rather moderate (Table 32.4). These reasons include the size of the research community and the fact that strains are easily preserved in the recipient laboratory. Nevertheless, since being established, the FGSC has sent out nearly 60,000 cultures (including all fungi in the collection). Because of changes in record keeping, the distribution history is broken into three time periods. Prior to 1985, 432 different Aspergillus strains were distributed out of a total of 566 Aspergillus strains in the collection. A total of 2095 cultures were sent out with A139, the most commonly distributed strain being distributed 102 times. This strain (also known as M180) was officially retired in October of 1984 but was maintained in the collection and sent to one researcher between 1985 and 2005. The early distribution of this strain is something of a who’s who of fungal genetics. In the 1960s it was sent to A. Ellingboe, C. Caten, and J. Raper, among others. It was also used in teaching fungal genetics in a number of high school classes in Illinois, California, Wyoming, Texas, North Dakota, and even Shawnee Mission High School in the Kansas City metropolitan area. The second most requested strain, A101, from this early time was also a strain that was subsequently removed. It was also primarily used for teaching.

| TABLE 32.4 | Recent Distribution of Aspergillus Strains from the FGSC |
| Year | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 |
| Strains | 191 | 301 | 183 | 167 | 136 | 133 | 163 | 141 | 230 |
| A. nidulans | NA | NA | 171 | 147 | 124 | 121 | 131 | 108 | 150 |
| A. niger | NA | NA | 11 | 13 | 7 | 8 | 16 | 17 | 38 |
| Recipients | NA | NA | 62 | 68 | 57 | 58 | 62 | 66 | 72 |
| Countries | NA | NA | 17 | 22 | 17 | 16 | 13 | 18 | 16 |
Between 1985 and 1997, 3,539 cultures of 733 different strains were distributed. Of these, 255 strains were sent only once and 519 strains were sent more than once. The strain that was distributed the most is the Glasgow wild type, FGSC A4 (also known as M139). Since 1997, 1,645 Aspergillus cultures have been sent out and among these were 110 A. niger strains. During the last 10 years, strains have been sent to an average of over 63 recipients per year. This represents a tremendous advance compared to the 12 labs in existence in 1960. These strains have been sent to approximately 22 different countries each year, with Aspergillus strains being sent to 38 countries since modern record keeping was established in 1998. In total, nearly 7,300 Aspergillus cultures have been distributed from the FGSC collection since its establishment in 1960. Overall, the one Aspergillus strain that has been distributed the most is FGSC A4, the Glasgow wild type. The FGSC has distributed this strain over 340 times in its history and it is the source of DNA for several genome libraries. It was also the strain used for the A. nidulans genome–sequencing project at the Broad Institute.8

In 2005, 26 A. fumigatus isolates were distributed while that number was 13 in 2004, 5 in 2003, and 3 in 2002. The first A. fumigatus strain, A1100 (also known as A293), was deposited by Michael Anderson in 2002. As of July 2006, we have sent out over 101 Aspergillus cultures including 74 A. nidulans, 6 A. niger, and 17 A. fumigatus strains.

### 32.1.4 Molecular Materials

Unlike most other collections, the FGSC also holds molecular resources for working with Aspergillus. The FGSC began to add resources for molecular genetics after moving to the University of Kansas Medical Center in 1985. Among the earliest resources added were cloning vectors and cloned genes. While some of these cannot be easily associated with one research community, 116 of the 329 plasmids in the FGSC collection are primarily for use in Aspergillus.

In addition to the 329 cloning vectors and cloned genes, 303 RFLP probes for Magnaporthe or Fusarium complete the plasmid collection. Of the Aspergillus plasmids, we have distributed 94 at least once. Since record keeping began in 1997, we have sent out Aspergillus plasmids a total of 441 times (Table 32.5). The most widely requested materials in recent years has been individual clones from the fosmid library that were used at the Broad Institute to order the genome sequence. Since they have been available, we have sent over 650 individual clones from gene libraries. One reason for the relative popularity of these materials is that the fosmid libraries are mapped on the genome and as such one can find a region of the genome of interest and receive the specific clones spanning that region from the FGSC.

<table>
<thead>
<tr>
<th>Table 32.5 Most Commonly Distributed Aspergillus Plasmids (1997–2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid Name</td>
</tr>
<tr>
<td>cosmid An26</td>
</tr>
<tr>
<td>ppyrG</td>
</tr>
<tr>
<td>p3SR2</td>
</tr>
<tr>
<td>pDC1</td>
</tr>
<tr>
<td>pMT-mRFP1</td>
</tr>
<tr>
<td>pAO81</td>
</tr>
<tr>
<td>pRG3-AMA1-NotI</td>
</tr>
<tr>
<td>LH3</td>
</tr>
<tr>
<td>pMT-sGFP</td>
</tr>
<tr>
<td>pJR15</td>
</tr>
<tr>
<td>pHY201</td>
</tr>
<tr>
<td>pXDRFP4</td>
</tr>
<tr>
<td>pFNO3</td>
</tr>
<tr>
<td>pILJ16</td>
</tr>
<tr>
<td>pRF280</td>
</tr>
</tbody>
</table>
In past years, the whole genome ordered libraries (the pWE15 and pLORIST2 libraries) as well as the chromosome-specific sublibraries\(^9\) were very widely distributed. Also widely used were the cDNA libraries\(^10\) that were distributed as pools of a phage stock. Since we began working with gene libraries we have distributed a total of 432 *Aspergillus* gene libraries to recipients in over 25 countries. From 1987 to 1997, there were 109 recipients in 16 countries while from 1998 to 2006 we sent 236 libraries to 122 different recipients in 25 countries. Without regard to the added value of receiving a library with published gene locations or chromosome associations, and compared to the approximately $1500 cost of a gene library from a commercial source such as Stratagene, this represents a value of nearly $650,000. Clearly having the libraries available from a source such as the FGSC has enabled people to do research that would otherwise have been either too expensive or too cumbersome to undertake. This also supports the notion that it was a good thing for the FGSC to branch out and add molecular resources to its holdings.

### 32.1.5 Information Resources

The FGSC holdings are described in the FGSC database, which includes various information for each strain. Originally, strain documentation was only on paper deposit sheets, but when the FGSC moved to the University of Kansas Medical Center, a database was created to allow better curation of the materials in the collection. As the software available for databases has become more capable, the FGSC has sought to add more functionality to its database. The current database operates in a Microsoft Access environment.\(^11\) The original database was in Dbase and there were several issues with migrating to the Access format. Because of the increased functionality and the ease of manipulating relationships in the current system, we are able to view data in ways that were not previously possible. The current database includes information on strain genotypes, strain distribution, and strain requirements. We also maintain records of clients including both their institutional affiliations and their order history.

Because the nomenclature for fungal genetics is complicated, we maintain each strain genotype as a single entry in a large text field. To simplify associating strains with their requirements or other characteristics, they are also associated with a table of markers. This may ultimately allow each strain to be associated with the entries in the genome for each mutated gene in its genome. There are 869 *Aspergillus* markers in the FGSC marker database, but this is both an underestimate of the number of identified markers and, by a factor of 10, of the number of genes in the genome.

The FGSC database, along with other material, is available at the FGSC website that gets more than two million hits per year. In the first seven months of 2006, the FGSC site has received nearly one-and-a-half million hits. Only six pages account individually for more than 1% of all hits. These six pages garner less than 30% of all of the traffic at the FGSC site. The remaining top 94 of the top 100 pages only contribute an additional 20% of all the traffic at the FGSC site demonstrating that the resources are widely utilized. On average, the FGSC website receives over 8000 hits per day (Fig. 32.2).

Over 250 different sites link to the FGSC site and it is regarded as the central site for information in fungal genetics. Moreover, numerous pages link to the *Aspergillus* homepage at the FGSC site. This site, www.fgsc.net/Aspergillus/asperghome.html, includes information regarding the annual *Aspergillus* meetings held in association with the European or Asilomar Fungal Genetics Conferences as well as information on working with *Aspergillus*. A search of the FGSC site identifies 782 mentions of the search term “Aspergillus.” This compares to 145 mentions of Magnaporthe and 199 mentions of Fusarium.

The FGSC publishes the *Fungal Genetics Newsletter* and has done so for many years, while the editorship has passed from one individual to another. The FGN includes annual bibliographies for both *Neurospora* and for *Aspergillus*. Dr. J. Clutterbuck is the editor for the *Aspergillus* bibliography. Figure 32.3 shows that from 1985 to 2005 there have been between 82 and 236 articles about *Aspergillus* genetics in the bibliography. The average number of citations per year in the *Aspergillus* bibliographies is 155. The compiled list of articles from the *Aspergillus* bibliography is searchable at the FGSC website. The FGSC also has archived the *Aspergillus Newsletter* and it is available online at the FGSC website. The *Aspergillus Newsletter* was published from spring of 1960 by J. A. Roper through 1981 when J.A. Clutterbuck published volume 15.

It is hoped that the information tools being developed for The Neurospora Functional Genomics program will be portable to other organisms. Because the FGSC holds key materials for work with
Aspergillus, we will endeavor to apply any information tools to Aspergillus as they become available. As information resources become a more significant part of the offerings of the FGSC, we are establishing working partnerships with bioinformaticians in the Computer Science Department at the University of Missouri-Kansas City with the long-term goal of enhancing our information resources. As exemplified by the web-use statistics, this will certainly be a valuable investment.

### 32.2 Conclusions

The FGSC has been and continues to be an important resource to the fungal genetics research community. The number of Aspergillus strains distributed is not as many as the number of Neurospora strains distributed. This is certainly a reflection of the fact that Aspergillus strains are available both directly from other investigators and from the several other existing collections of Aspergillus strains. Most
strains from the Glasgow collection were not made available to the FGSC. In 1963, Professor G. Pontecorvo wrote in response to a request by Dr. Kafer that strains from Glasgow be deposited in the FGSC that their reluctance to providing strains was “simply a matter of not providing stocks to people whom we don’t know.” This, perhaps, is a source of the stronger roots of Aspergillus research in the United Kingdom and western Europe as compared to the strong Neurospora community in the United States. Another demographic apparent in examination of distribution lists is the shifting use of fungi in teaching. In the first 25 years, strain 101, for example, was sent to over 50 high schools. High schools rarely order from the FGSC any longer and the most commonly used teaching protocol is the arginine pathway elucidation in Neurospora. This is, however, taught at the college level. Moreover, the Aspergillus community has a strong commercial component. Of the Aspergillus strains distributed in the 1998–2006 period, over 16% were sent to commercial clients. During the same period fewer than 2% of Neurospora strains were sent to commercial clients demonstrating the greater economic importance of Aspergillus. As the biotechnology revolution continues, the FGSC will continue to serve the needs of its constituency. While other collections exist for Aspergillus, the FGSC has a long-standing reputation as being open and fair in its policies. This is in large part due to the influences of the founders. Nevertheless, there is a growing trend toward establishment of Bioresource Centers that have a critical mass to support both the technical requirements for maintaining a culture collection as well as the logistical and legal aspects of maintaining and distributing strains. While the World Federation of Culture Collections does support the needs of culture collections, having many small collections has increased costs. These costs include the duplication of technical and logistical support. It may be that more and more small repositories are forced, by logistical or legal requirements, to merge and allow centralized oversight. While the arguments for this are straightforward, there are drawbacks. The FGSC is able, because of its small size, to serve the needs of its community. A larger centralized collection would have a larger community, but each individual community would be in competition for the attention and resources of such a merged collection. The need for security may force consolidation upon smaller collections, but for the present the FGSC will maintain its unique position as a collection of genetic diversity for Aspergillus and other filamentous fungi.

References

Index

A
ABPA. See Allergic bronchopulmonary aspergillosis (ABPA) 184
Abplp module, 184
Abplp module, 184
Acetate utilization, 133
Acetobacter, 428
Acetyl-CoA metabolism, 134
ACP. See Acyl carrier protein (ACP) 182
Actin cytoskeleton, 246
Actin dynamics, 183–185
Actin patches, 182
acuK gene, 138
acuM gene, 138
Acyl carrier protein (ACP), 460
Acyltransferase (AT), 460
Adenylate cyclase, 92–93
Advanced Industrial Science and Technology (AIST), 76
Agilent, 479, 489
Ago1, 204
ahhA2 mutation, 244
Air crescent sign, 365
AIST. See Advanced Industrial Science and Technology (AIST) 355
alc gene cluster, 352–355
akR promoter, 333
Allergic asthma vs. ABPA, 391
Allergic bronchopulmonary aspergillosis (ABPA), 360, 363, 378
vs. allergic asthma, 391
Afp alkaline serine protease, 386
Alternating neddlyation status control, 162
AMA1
plasmids, 501–502
replicator, 501
sequence, 501
Amino acids, 143–168
aflatoxin production, 461
Aspergillus, 143–176
Aspergillus nidulans transporters, 307
biosynthesis, 148–155
biosynthesis genes, 149–150
biosynthesis schematics, 154
obtained by fungal cell, 144
genes for proteins, 146
obtained by protein degradation, 156–166
uptake, 143–176
synthesis examples, 153–155
uptake, 145–146
Aminobutyric acid (GABA), 133, 302, 307, 314
Ammonium transporter (AMT), 308
Amphotericin B, 367–368
AMT. See Ammonium transporter (AMT)
Anaphase promoting complex (APC), 238
BimE, 235, 239
hyphal outgrowth, 244
metaphase-anaphase transition, 238
securin, 238–239
septation, 241
Anaphase protein dephosphorylation, 239–240
Animal pathogen, 17
AnNdc1 foci, 266
Annealing product fusion PCR, 518–519
An-Nup2, 268
Antifungal prophylaxis, 367
Antigen GM, 365
Antioxidation affecting aflatoxin formation, 462

537
The Aspergilli

AP-1, 182
AP-2, 182
AP-3, 182
APC. See Anaphase promoting complex (APC)
AReA, 330–332
Argonautes
- RNA silencing genes evolution, 201
- SMS-2, 204
Aromatic amino acid the terrequinone A biosynthesis, 155
ArrayExpress, 479
Ascomycete dicer RNAseIII domains, 202
Ascomycete RNA-dependent RNA polymerases, 203
Ascorbic acid, 463
Aspergilli
- biology, 87–428
- endocytosis, 177–196
- gene regulation, 103–120
- genomics, 3–86
- gluconeogenic carbon metabolism, 129–142
- mitogen-activated protein kinase pathways, 121–128
- RNA silencing, 197–210
- signal transduction, 87–102
Aspergilloma, 360, 363
- first published description, 9
Aspergillosis, 458
- disease forms, 378–379
- mammalian models, 401–412
- mini-host models comparison, 423
- mini-host models emerging role, 413–428
Aspergillum
- morphology, 5
Aspergillus, 3–15
- allergens, 377–400
- amino acid supply, 143–176
- amino acid uptake comparisons, 145–146, 147
- animal pathogen, 9
- bibliography, 534
- biotechnological aspects, 429–474
- chromatin, 321–342
- clinical aspects, 359–376
Fungal Genetics Newsletter, 534
- fungal genetics stock center, 527–536
- genetics, 10
- genome evolution, 43–56
- genomes database, 338
- history, 3–5
- human exploitation, 4
- industry, 6–8
- medically important aspects, 359–474
- microarrays, 475–482
- pathogenicity determinants, 377–400
- protease activity, 441–456
- protein production, 441–456
- publicly available genomes, 44
- secondary metabolites, 8–9
- taxonomy, 3–5
- transporters, 301–320
Aspergillus caespitosus, 5
Aspergillus clavatus
- genome annotation, 37
- mating loci, 36
Aspergillus flavus
- aflinematrix arrays, 20
- aflatoxin contamination, 457, 465
- aflatoxin crop resistance, 464
- comparative genomics, 19
- database, 18
- dClC protein, 201
- ecology, 15–16, 20
- ecotype, 17
- EST sequences, 465
- experimental RNA silencing, 198
FGSC A1009, 529
FGSC A1010, 529
- functional genomics, 20
- genome sequence, 15–24
- genomes, 479
- integrated database, 20
- physical structure, 18, 19
- population analysis, 19
- ppdC protein, 201
- RNA silencing during host infection, 205
- sequence and annotation, 18
- structural genomics, 18
- systemic infections, 9, 10
- unique genes and features, 19
Aspergillus fumigatus, 9
AF293, 27, 32, 33
- allergens cloned by IgE binding, 390
- asexual cycle, 295
- A+T, 348–349
- chromosomal location, 28
- chromosomal rearrangement, 352
- clustering and fragmentation, 349
- comparative genomics, 27
- experimental RNA silencing, 198
- GanB, 91
- genome comparative view, 25–42
- genome sequences, 27–29
- genomes, 479
- global transcription factor CpcA, 151
- GprK, 98
- fong oligonucleotide microarrays, 475
- MAP kinases, 95
- MAPK genes, 122
- mating loci, 36
- mating-type loci, 285
- MpkC, 124
- myoA gene, 184
- pathogenesis, 380
- proteins, 29
- pyrG gene, 505
- repeat-induced point mutation, 343–359
- RNA silencing during host infection, 205
- SakA/HogA signaling pathway, 124–125
- secondary metabolite biosynthetic gene clusters, 32–34
- sequence divergence, 28
- sex and sex genes, 35–36
- sexuality, 352
- synteny, 28
systemic infections, 10
TE, 343–359
TE detection, 344
virulence genes, 29–30
virulence studies, 422
voriconazole, 478
Aspergillus glaucus, 4
Aspergillus nidulans, 229
Abp1p module, 184
actin cytoskeleton, 246–247
actin-dependent motor proteins, 247
affinity tags for protein purifications and proteomics, 505
AfFlbA protein, 97
aflatoxin contamination, 458
AMA1 sequence, 501
A+T, 348–349
auxotrophic markers, 497
branching and cell cycle control, 244–245
carbon metabolism, 129–130
catabolism control, 131
cDNA array, 476, 478
cell cycle, 243
clustering and fragmentation, 349
cellular morphogenesis, 212
ChIP, 111
carbon metabolism, 129–130
cytoskeleton, polarized growth, and cell cycle, 223–260
cytoskeleton, 246
developmental coordinators and modifiers, 293–294
digA gene, 189
dNA targeting cassettes, 494
drug resistance markers, 498
environmental factors, 280–284
fatty acid metabolism, 137
fatty acid utilization, 136
F-box proteins, 162
FGSC A1, 528
FGSC A1100, 528
FGSC distribution, 531
FlbA, 96
fluorescent protein tagging, 505
FM4-64 positive cortical structure, 180
functional annotation, 46
fungicide-resistance genes identifying tubules, 225
fusion PCR, 494–495
GanA, 91
GanB, 91
gap closure in genome sequence, 58
gene calling accuracy, 499
gene deletion, 496, 502, 503
gene manipulations, 493–512
gene regulation, 48, 104
gene targeting confirmation, 500
gene targeting constructs, 494
gene targeting efficiency, 494
gene targeting improvement, 495
genes related in mating process, 285
genes required for polarity, 248
genetic and genomic information, 289–290
genome(s), 43–53, 479
genome annotation, 44
genome assembly, 44
genome characteristics, 44–47, 45
genome comparative view, 43–56
genome protein domains, 47
genome sequence, 57–74
genome sequence gap closure, 58
genome size, 51
genomic evolution, 48–52
geonomics future, 53
hypertonic stress response pathways, 125
hypoxia, 211–222
hypoxic germination and nuclear division cycle, 234
hypoxic morphogenesis, 211–219
hypoxic tip, 214–215
hypothetized genetic model, 286
interphase microtubule cytoskeleton organization, 226
KU70, 521
KU80, 521
labeling organelles in GFP-fusions, 516
long oligonucleotide microarrays, 475
MAPK, 95, 122
mitosis evolution, 270
mitosis regulatory network, 235
mitotic restructuring, 264–265
morphogenetic paradigm, 213
morphogenetic patterns during development, 218
mitosis, 235
morphology, 280
MpkC, 124
MTOC, 228
myoA gene, 184
nonintegrative gene expression, 501
novel system, 313
NPC, 261–278
Nups, 265
organization, 246
Pan1p module, 185
PCl complexes, 165
phylogenetic relationships, 50
physiology, 48
PkaA and PkaB, 93–94
PLC, 95
polarisome, 246
positive and negative selection, 498
promoter rundown, 502, 503–504
putative F-box proteins, 159
Aspergillus nidulans (Continued)
RanGTP gradient, 268–269
recombination frequencies, 62
recycled for multiple gene modification, 498
regulation during mitosis, 262–269
regulators, 288–293
regulatory proteins, 105–106
repeat-induced point mutation, 343–359
repetitive elements, 45
reproductive strategy, 52
RgsB, 97
RgsC, 97
RNA silencing proteins, 200
rrpA, 200, 202
rsdA, 200
SakA signaling pathway, 124–125
secondary metabolite gene clusters, 478
secondary metabolites, 46–47
sensing growth milieu, 308
septum formation, 216–218
sex, 310
sexual development, 279–300
sexual organs, 281
sexuality, 352
SfaD, 91
signal transduction components, 286–288
silencing model, 199
SlcB protein, 147
SlcD proteins, 148
SNAREs, 188, 189
soluble nuclear transporters, 261–273, 272
STRs, 72
structure-function analysis, 311–312
subtelomere domains, 59
systemic infections, 9
taxonomy, 4–5
TCA cycle, 133
TE, 45, 343–359
TE detection, 344
telomeres, 58
telomeric contigs, 71
TLH, 60
trafficking and endocytosis, 310
transcription factors, 104, 288–293
transformation markers, 497–499
transport, 261–278
transporters, 301–307, 309, 310, 313
transposon-related genes, 351
two-step site-specific mutation, 506
ultrastructural morphology, 281
Aspergillus nidulans sexual development
carbon, 281, 282
endogenous inducer, 284
FphA (fungal phytochrome A), 283
GprA, 285
GprD, 286
hypoxia, 282
light, 282
MAP kinase, 287
NADPH oxidase (NoxA), 284
nitrogen, 281, 282
psi factor, 284
ROS, 284
Aspergillus niger
FGSC A732, 529
FGSC A1121, 529
FGSC distribution, 531
food industry, 7–8
genomes, 479
products, 7
zinc binuclear cluster proteins, 106
Aspergillus oryzae
A+T, 348–349
cDNA array, 476
chromosomal rearrangement, 352
clustering and fragmentation, 351
dcIC protein, 201
ecotype of Aspergillus flavus, 17
food products fermented by, 6, 429–440
gene expansion, 78–80
gene prediction, 78
genome, 76
genome sequence, 75–86
genome sequence application, 82
genomes, 479
GprK, 98
H3 gene, 324
MAPK genes, 122
myoA gene, 184
notable characteristics, 81
ppdC protein, 201
repeat-induced point mutation, 343–359
sexuality, 352
SlcD proteins, 148
TE, 343–359
TE detection, 344
Aspergillus parasiticus, 16
aflatoxin contamination, 457
aflatoxin crop resistance, 464
aflatoxin gene cluster, 460
cDNA array, 476
EST sequences, 465
experimental RNA silencing, 198
RNA silencing during host infection, 205
Aspergillus sojae, 6
Aspergillus terreus, 8
H3 gene, 324
Aspergillus versicolor, 458
AT, See Acyltransferase (AT)
A+T content, 348–349
Attacin, 415
Awamori, 436
AzaA, 310
GFP, 309

B

Bacillus subtilis, 302, 310
BAD, See Biotin-Acceptor-Domain (BAD)
BAL, See Bronchoalveolar lavage (BAL)
Barley-Miso, 433
Bean paste, 75, 430–431
Bennett, John Hughes, 9
Beta-D-glucan, 366
Beta-ketoacyl synthase, 460
BFP. See Blue Fluorescent Protein (BFP)
BIFC. See Bimolecular Fluorescence Complementation (BIFC)
BimC, 231–232
BimD, 239
BimE APC, 235
SPB, 239
BimG, 240
BimG11 mutants, 243–244
Bimolecular Fluorescence Complementation (BIFC), 518
Biofilm, 484
Bioprocess engineering protein production, 450
Bioreactors
fungal chemostat fermentations, 485
schematic, 484
Biotin-Acceptor-Domain (BAD), 305
Blue Fluorescent Protein (BFP), 517
Chang, 430
Chemostats and microarrays, 483–492
Chiang, 6
Chimeras, 306
ChIP. See Chromatin immunoprecipitation assays (ChIP)
Chitin assay, 409
CHORD. See Cysteine- and histidine-rich-domain (CHORD)
Chorismate, 415
Caffeic acid, 463
Calcineurin, 384
Calcium, 238
cAMP. See Cyclic adenosine monophosphate (cAMP)
CAND1, 164
Carbon
aflatoxin production, 461
Aspergillus nidulans, 129–130
Aspergillus nidulans sexual development, 281, 282
catabolites, 448
metabolism, 129–139
sources, 131
Carbon metabolite repression (CMR), 307
Carboxypeptidase Y (CPY), 179
Caspofungin, 368
Catabolism control, 131
Caveat lector, 178
Cdk1. See Cyclin-dependent kinase 1 (Cdk1)
cDNA arrays, 475
A. nidulans, 476, 478
A. oryzae, 476
Cecropin, 415
Cell cycle, 244–245
Cell-end markers at the cortex, 233
Cellular expression, 305
Cellular morphology, 212
CenH3, 326
Centromere drive, 326
Centromere H3 histone variant, 324
Cerebral aspergillosis, 362
Cerebral inoculation route, 408
CFP. See Cyan Fluorescent Protein (CFP)
CGD. See Chronic granulomatous disease (CGD)
CHG. See Comparative genome hybridization (CHG)
Chung, 430
Chemostats and microarrays, 483–492
Chiang, 6
Chimeras, 306
ChIP. See Chromatin immunoprecipitation assays (ChIP)
Chitin assay, 409
CHORD. See Cysteine- and histidine-rich-domain (CHORD)
Chorismate, 155
chpA, 424
Chromatin
Aspergillus, 321–342
Neurospora crassa, 322
rearrangements, 336
S. cerevisiae, 321
S. pombe, 321
transcription factors, 330–336
Chromatin immunoprecipitation assays (ChIP)
A. nidulans, 111
techniques, 336
Chromatin rearrangements
alcR promoter, 333
Chronic aspergillosis, 362
Chronic cavitary pulmonary aspergillosis, 363
Chronic fibrosing pulmonary aspergillosis, 363
Chronic granulomatous disease (CGD), 361
Chronic invasive pulmonary aspergillosis, 362
Chronic necrotizing aspergillosis, 360
Chronic necrotizing pulmonary aspergillosis, 363
chu, 429
Church, Mable, 5
Citrlic acid, 7
Classical inbred mice, 404
Clathrin-dependent and -independent endocytosis, 180
Cleistothecium, 280, 290–291
CMR. See Carbon metabolite repression (CMR)
CMV. See Cytomegalovirus (CMV)
COGEME, 479
Compactin, 8
Comparative genome hybridization (CGH), 465
Confocal microscopy, 304
Congenic mice, 404
Conidia, 5
Conidiophore, 5
Conserved H4, 322
Consomic mice, 404
COP9 signalosome (CSN), 166
Aspergillus nidulans sexual development, 293
components, 163
subunit, 166
Corn
aflatoxin contamination, 464
oligo microarray, 464
Cotton
aflatoxin contamination, 464
oligo microarray, 464
CPC. See Cross-pathway control (CPC) system
The Aspergilli

cpa, 385
Aspergillus nidulans sexual development, 292
master regulator degradation, 156
mRNA control, 151
recognition elements, 151
translational control, 152
cpb, 292
cpc, 385
CPRE. See cpa
CPY. See Carboxypeptidase Y (CPY)
CreA, 332, 335
CreloxP-based method, 498
CRL. See Cullin-RING ligases (CRL)
Crmi, 273
Cross-pathway control (CPC) system, 151–152
CSN. See COP9 signalosome (CSN)
Cullin-RING ligases (CRL), 145, 161
Currie, James N., 5, 7
Cutaneous aspergillosis, 362
Cyan Fluorescent Protein (CFP), 517
Cyclic adenosine monophosphate (cAMP), 92–93
Cyclin-dependent kinase 1 (Cdk1), 506
Cyclosome, 238
Cys-scanning mutations, 305
Cysteine- and histidine-rich-domain (CHORD), 242
Cystic fibrosis transporters, 303
Cytomegalovirus (CMV), 360
D

dclB, 200
dclC
Aspergillus flavus, 201
Aspergillus oryzae, 201
Drcl, 204
Defatted soybeans, 431
Demethylhydroxysterigmatocystin (DMDHST), 460
Demethylsterolignocystin (DMST), 460
DHOMST. See Dihydro-O-methylsterolignocystin (DHOMST)
DHST. See Dihydrosterolignocystin (DHST)
Diastrophic dysplasia transporters, 303
Dicers
DCL-1, 204
RNA silencing genes evolution, 201
diga, 189
Dihydro-O-methylsterolignocystin (DHOMST), 460
Dihydrosterolignocystin (DHST), 458–459
Diptericin, 415
Disorders
cultures, 364
diagnosis, 363–366
epidemiology and risk factors, 360–363
immunomodulatory therapy, 368
prognosis, 369
radiology, 364–366
treatment, 366–369
Dithiothreitol (DTT), 478
DMDHST. See Demethylhydroxysterigmatocystin (DMDHST)
DMST. See Demethylsterolignocystin (DMST)
DNA, 476
A. nidulans, 494
binding fluorochromes, 514
damage-checkpoint pathway, 241
integration, 494
Neurospora crassa, 494
NimX CDC, 241
S. cerevisiae sequences, 494
targeting cassettes, 494
Doburoku, 434
DopA, 294
Double hairy-type bristle, 417, 418
Downstream signaling branches, 92–94
Drosomycin, 415
Drosophila
candidate compound screening, 420
infection models, 417–422
virulence testing, 421
Drosophila IA model, 421
DTT. See Dithiothreitol (DTT)
Dye flip labeling design, 488
Dynamin independent endocytosis, 181
Dynecin, 223

E

Early sexual developmental stage (ESD), 295
Economic impact, 458
EIF3, 166
Emericella nidulans, 4
EMTOC. See Equatorial MTOC (EMTOC)
Endo, Akira, 8
Endocytic adaptors, 182
Endocytic internalization proteins, 184
Endocytosis, 177–190
Abplp nodule, 184
accessory proteins and scaffolds, 183–185
actin dynamics, 183–185
adaptors, 181
caveat lector, 178
dynamin independent, 181
endocytic vesicle formation, 180
functional characterization, 190
internalization step, 180
Las17p module, 183
membrane identity across endocytic pathway, 183–187
Panlp module, 185
pathway, 178–182
plasma membrane lipid composition, 178
Saccharomyces cerevisiae, 179
signaling, 190
Endogenous inducer, 284
Endpoints, 408–409
Environment, 462
Epifluorescence, 304
Index

Epitope tagging
Aspergillus nidulans transporters, 304
immunofluorescence microscopy, 515
Equatorial MTOC (EMTOC), 227
ER-associated degradation (ERAD), 310
Erythocyte sedimentation rate (ESR), 361
Escherichia coli transporters, 302
ESCRT complexes, 187
ESD. See Early sexual developmental stage (ESD)
EST. See Expressed Sequence Tag (EST)
Eukaryotic translation initiation factor eIF3, 152
Eurotium herbariorum, 4
Experimental RNA silencing, 198–199
genetic analysis, 199–200
during infection, 199
Expressed Sequence Tag (EST), 465
sequence project, 464
F
F-actin. See Filamentous actin (F-actin)
FadA, 90
Farmer's lung, 9
Fatty acids
beta oxidation, 136
metabolism, 137
F-box proteins, 162
FC&C. See Federal Food Drug and Cosmetic Act (FC&C)
Federal Food Drug and Cosmetic Act (FC&C), 7
Fermentations
Acetobacter, 428
batch vs. chemostat, 483–486
fermentation, 485
main mash, 434
shochu, 436
vinegars, 428
FGSC. See Fungal genetics stock center (FGSC)
FibA, 96
Filamentous actin (F-actin), 223
Filamentous fungi, 487
Filtrated sake (seishu), 434
FITC. See Fluorescein isothiocyanate (FITC)
FLIP, 517–518
Fluorescein isothiocyanate (FITC), 514
Fluorescence proteins, 517–518
Fluorescence resonance energy transfer (FRET), 518
Fluorescent labeled organelles, 513–526
Fluorescent microscopic imaging, 304
Fluorescent protein tagging, 515–521
A. nidulans, 505
flexible linkers, 521
fusion PCR-based approaches, 518–520
gene targeting using nkuA deletion strains, 521
immunofluorescence microscopy, 515
multiple fluorochromes labeling, 516
plasmid-based approaches, 518
FM4-64, 514
positive cortical structure, 180
Food industry, 7–8
Food products fermented
Aspergillus oryzae, 6
Aspergillus sojae, 6
Food safety
aflatoxicosis, 459
aflatoxin contamination, 458
FphA, 283
FRAP, 517
FRET. See Fluorescence resonance energy transfer (FRET)
Fruit flies
A. fumigatus mutants with attenuated virulence, 421
drug delivery, 421
high throughput screens, 421
Fungal chemostat fermentations, 485
Fungal Genetics Newsletter, 534
Fungal genetics stock center (FGSC), 527–534
A1, 528
A4, 532
A732, 529
A1009, 529
A1010, 529
A1100, 528
A1121, 529
A. flavus, 529
A. nidulans, 528, 531
A. niger, 529, 531
common markers, 530
database, 533
distribution, 531
growth, 529
history, 527
information resources, 533
markers per strain, 529
molecular materials, 532
repeat nups during mitosis, 267
strains, 528–530
strains distribution, 531
strains from sequencing programs, 530
website, 533–534
Fungal inoculum size, 406
Fungal phytochrome A, 283
Fungicide-resistance genes identifying tubules, 225
Fusion PCR
annealing product, 518–519
fluorescent protein tagging, 518–520
gene targeting constructs, 494–495
generating linear replacement cassettes, 504
generating null alleles, 503
GFP fusion C-terminal, 519
targeting pyrG-alcA cassette, 505
G
G2/M transition, 238
G protein
alpha subunits, 89–90
beta subunit, 91
coupled receptors, 88, 285
gamma subunit, 92
GABA. See Aminobutyric acid (GABA)
Galactomannan (GM), 365, 389
antigen, 365
Galleria mellonella, 422
virulence studies, 423
Gallic acid, 463
Gamma-tubulin, 225, 227, 229
GanA, 91
GanB, 91
Aspergillus fumigatus, 91
Aspergillus nidulans, 91
GBrowse, 19
G-CSF. See Granulocyte colony-stimulating factor (G-CSF)
Gene Expression Omnibus, 479
Gene Ontology data, 465
Generally Recognized as Safe (GRAS), 7, 75
Genes
encoding proteins, 187
pathogenicity, 389
regulation levels, 104
GeneSpring (Agilent), 479, 489
Genisphere, 476
Genome surveillance, 242
Genomics, 11
Aspergillus nidulans sexual development, 294
Germ tube emergence, 244
GFP fusions
C-terminal, 519
labeling organelles/cellular structures in A. nidulans, 516
Ginjo-shu, 82
Glel
during mitosis, 267
RNA transport, 267
Gliotoxin, 386, 387
Global transcription factor CpcA, 151
Gluconeogenic carbon metabolism, 129–139
acetate utilization, 133
fatty acid catabolism regulation, 137
fatty acid utilization, 136
gluconeogenesis control, 139
induction, 132
organelle localization, 134
peroxins, 135
TCA cycle, 133
Gluconeogenic carbon sources
Aspergillus nidulans, 139
Saccharomyces cerevisiae, 139
Gluconobacter, 428
Glufosinate resistance gene, 521
GM. See Galactomannan (GM)
Golgi syntaxin (t-SNARE) tlg2p, 179
GpaA, 90
GpaB, 91
GPCR. See G protein, coupled receptors
GpaA, 92
GpRA, 285
GpRB, 285
GpRD, 286
GpRK, 97
Graft-versus-host disease (GVHD), 360
Granulocyte colony-stimulating factor (G-CSF), 368–369
GrAS. See Generally Recognized as Safe (GRAS)
Growth tests, 303
Gs/M transition regulation, 236
GVHD. See Graft-versus-host disease (GVHD)
H
H2A.X histones, 327
H2A.Z histones, 327
H3 histones, 323
Hair-type bristle, 417, 418
Halo sign, 365
Health risk, 458
Hebrew Bible, 4
HECT-type E3s, 160
Hemoptysis, 361
Heterotrimeric G protein signaling, 87–91
elements, 89
hhA genes, 324
hhtA genes, 324
High-mobility group (HMG), 36
Histidinol biosynthesis, 154
Histologic analysis, 409
Histological organelle-staining methods, 514
Histone H2A, 326
Histone H2B, 326
HMG. See High-mobility group (HMG)
HogA signaling pathway, 124
Homologs, 181
Host
contacting, 382
feeding, 384
fighting, 385
sensing, 383
sensitizing, 388–390
Hulle cells, 291
Human estrogen receptor, 505
Human recombinant granulocyte colony-stimulating factor, 368–369
Hypertonic stress response pathways, 125
Hyphal germination and nuclear division cycle, 234
Hyphal morphogenesis
Aspergillus nidulans, 211–219
cell cycle, 243
Hyphal outgrowth, 244
Hypoxia, 282
I
IA. See Invasive aspergillosis (IA)
Immune suppression, 404–405
Immunocompromised sinusitis, 361
Immunodetection, 304
Immunofluorescence microscopy, 514
epitope tags, 515
Immunomodulatory therapy, 368
Index

Imo-shochu, 436
Importin beta1, 273
Importin beta-like super family, 272
Importin-alpha, 273
Inbred stocks, 403
Indigestion, 6
Infection assay, 417–418
Inhalational inoculation route, 407
Innate immune system, 379
Invasive aspergillosis prevention, 26
Inoculation route, 406–408
Insects
  cellular immunity against, 416
  humoral response against, 415
Insects innate immunity pathways, 415
Interphase microtubule cytoskeleton organization, 226
Intracellular amino acid pool sensing, 148
Intranasal inoculation route, 406
Intratracheal inoculation route, 406
Intravenous inoculation route, 407
Invasive aspergillosis, 25
  host-pathogen system, 390
Invasive aspergillosis (IA), 401
  fly model, 417
  in vitro model pathogenesis, 414
  invertebrate models, 416
  murine models, 402–407
Invasive aspergillosis prevention, 26
Invasive pulmonary aspergillosis (IPA), 378
Invasive sinusitis, 361
Invertebrates
  antifungal compound mass screening, 423–424
  host immune response, 424
  inverted repeat transgenes (IRT), 198–199
  IPA. See Invasive pulmonary aspergillosis (IPA)
  IRT. See Inverted repeat transgenes (IRT)
Itraconazole, 367
K
Kanjang, 6
KapA (importin-alpha), 273
KapK, 273
Kecap, 6
Ketamine, 406
Ketoacyl synthase (KS), 460
Ketoacyl-thiolase enzymes, 137
KfsA (kinase for septation) gene, 245
KinA, 230
Kinase for septation gene, 245
Kinesin, 223
King Tut’s curse, 10
Kip2, 229, 230
KipA, 232–233
KipB, 232
Koikuch-shoyu, 432
Koji, 4, 6, 429–430
  molds, 432–433
  preparation, 435–436
  production, 431
Kome-shochu, 436
KS. See Ketoacyl synthase (KS)
Ku70
  A. nidulans, 521
  N. crassa, 495–496
Ku80
  A. nidulans, 521
  N. crassa, 495–496
Kuki, 430
L
Laboratory mice categories, 404
LaeA, 156, 461
Las17p module
  endocytosis, 183
Saccharomyces cerevisiae, 182–183
Late sexual developmental stage (LSD), 295
Leaf Permease 1 (LPE1), 313
Leishmania major transporters, 302
Leucine Zipper protein, 294
Leviticus, 4
Light in sexual development, 282
Linker histone HI, 328–239
Lipids, 461
Long oligonucleotide microarrays, 475
Long terminal repeats (LTR), 344
Loop design, 488
Lovastatin, 8
LPE1. See Leaf Permease 1 (LPE1)
t-phenylalanine, 155
LSD. See Late sexual developmental stage (LSD)
LTR. See Long terminal repeats (LTR)
t-tryptophan, 155
Lysine and penicillin biosynthesis, 154

Kasuga, 434
Koizumi, 436
Koizumi-shochu, 436
Kosaka, 436
Kumagai, 436
Kuri, 430
Kusakabe, 436
Kwong, 10
L
Laboratory mice categories, 404
LaeA, 156, 461
Las17p module
  endocytosis, 183
Saccharomyces cerevisiae, 182–183
Late sexual developmental stage (LSD), 295
Leaf Permease 1 (LPE1), 313
Leishmania major transporters, 302
Leucine Zipper protein, 294
Leviticus, 4
Light in sexual development, 282
Linker histone HI, 328–239
Lipids, 461
Long oligonucleotide microarrays, 475
Long terminal repeats (LTR), 344
Loop design, 488
Lovastatin, 8
LPE1. See Leaf Permease 1 (LPE1)
t-phenylalanine, 155
LSD. See Late sexual developmental stage (LSD)
LTR. See Long terminal repeats (LTR)
t-tryptophan, 155
Lysine and penicillin biosynthesis, 154

M
MAB. See Monoclonal antibody (MAb)
Macrophage-based immune suppression, 405
Madison Chamber, 407
Main mash (moromi), 436
Mammalian models, 402–409
Mammalian Mx proteins, 181
MAPK. See Mitogen activated protein kinases (MAPK)
Master regulator degradation, 156
mat A, 285
mat B, 285
Mating-type loci, 285
MC. See Mitochondrial transporters (MC)
MedA, 293
Medusa protein (MedA), 293
Meiotic silencing, 203
Medch, 285
Mep metalloprotease, 386, 445
MET, 308
Metabolites, 307
Metabolomics, 451
Metaphase–anaphase transition, 238
Metchnikowin, 415
Mevastatin, 8
Micafungin, 368
Micheli, Pier A., 4, 25
Microarray, 465
analysis software, 479–480
data interpretation, 476
image analysis, 476
labeling and hybridization, 476
normalization, 476
platforms, 475
resources, 479–480
RNA isolation, 475
studies, 476–478, 477
Microarray experiments
batch vs chemostat fermentation, 483–486
design, 487–488
dye flip labeling design, 488
filamentous fungi, 487
growth characteristics, 486
loop design, 488
universal reference design, 488
Microbial Genome Sequencing Project, 18
Microtubules (MT), 223, 227–232
dependent motor proteins, 231
lattice, 231
plus end, 229
plus-end localizing proteins, 227–229
S. cerevisiae, 229
S. pombe, 229
Saccharomyces cerevisiae, 226
Microtubules organizing centers (MTOC), 223, 225, 227, 228
MIDAS, 489
Miso (bean paste), 6, 75, 430–431
fermentation, 433
manufacturing process, 434
Mitochondrial transporters (MC), 303
Mitogen activated protein kinases (MAPK), 94, 124
aspergillus pathways, 121–128
Aspergillus fumigatus, 95, 122
Aspergillus nidulans, 95, 122
Aspergillus nidulans sexual development, 287
Aspergillus oryzae, 122
signal transduction pathways, 122
Mitotic exit, 241
Mitotic restructuring, 264–265
ML236, 8
Mod5, 233
Monoclonal antibody (MAb), 405
Monoubiquitin, 181
Moromi, 436
Moromi-mash, 430
Mortality, 378
Moto preparations, 436
mpkA gene, 123
mpkB pathway and mating, 123
mpkC
Aspergillus fumigatus, 124
pathway and carbon utilization, 124
MPN domain proteins, 164–165
MT. See Microtubules (MT)
MTOC. See Microtubules organizing centers (MTOC)
Mugi-shochu, 436
Multiple fluorochromes labeling, 516
Multivesicular body pathway (MVB), 179–180, 188
Multivesicular endosomes (MVE), 180
Mummy’s curse, 10
Murine models, 402
MVB. See Multivesicular body pathway (MVB)
MVE. See Multivesicular endosomes (MVE)
MxA, 181–182
Mycoses, 458
Mycotoxins, 386, 457–474
contamination, 458
myoA, 190
Aspergillus fumigatus, 184
Aspergillus nidulans, 184
Aspergillus oryzae, 184
Myosin, 223, 247

N
N. crassa
Ku70, 495–496
Ku80, 495–496
NADPH oxidase (NoxA)
Aspergillus nidulans sexual development, 284
S. cerevisiae, 284
S. pombe, 284
Nama-shoyu, 433
National Food Research Institute (NFRI), 76
National Institute of Technology and Evaluation (NITE), 77, 82
National Research Institute of Brewing (NRIB), 76
Natto (soybeans), 430, 431
Natural killer (NK) cells, 402, 404
NCR. See Nitrogen catabolite repression (NCR)
NE. See Nuclear envelope (NE)
NEDD 8 homolog, 160
Negative regulators of G protein signaling, 96
Neosartorya fischeri  
mating loci, 36  
NRRL181, 28  
ppBR, 201  
proteins, 29  
NES. See Nuclear export sequences (NES)  
*Neurospora crassa*  
chromatin, 322  
DNA integration, 494  
*Neurospora* protein, 158  
Neutropenia, 360  
Neutrophil-based immune suppression, 405  
NFRI. See National Food Research Institute (NFRI)  
NHEJ. See Nonhomologous DNA end-joining (NHEJ)  
Nicander of Colophon, 3  
niiA-niaD bidirectional promoter, 331  
nimA gene, 262  
inactivation, 240  
kine, 235, 236–238  
Nimblegen, 479  
nino, 238–239  
nino, 238–239  
nino, 239  
ninX  
CDC, 241  
CDC2 kinase, 236–237, 244  
NirA, 330–332  
NITE. See National Institute of Technology and Evaluation (NITE)  
Nitrate assimilation gene cluster, 330–331  
Nitrogen  
aflatoxin production, 461  
*Aspergillus nidulans* sexual development, 281, 282  
metabolites, 448  
Nitrogen catabolite repression (NCR), 307  
NK. See Natural killer (NK) cells  
NLS. See Nuclear localization sequences (NLS)  
NMD. See Nonsense-mediated decay (NMD) pathway  
Nonhomologous DNA end-joining (NHEJ), 495–496  
Nonmurine models, 403  
Nonoxidized purines, 307  
Nonribosomal peptide synthetases (NRPS), 465  
Nonsense-mediated decay (NMD) pathway, 114  
Nontoxicigenic *Aspergillus flavus*, 463  
NOR. See Norsorlorinic acid (NOR)  
Norsorlorinic acid (NOR), 200  
NosA, 291  
NosA. See NADPH oxidase (NoxA)  
NPC. See Nuclear pore complex (NPC)  
NRI Microbial Genome Sequencing Project, 18  
NRIB. See National Research Institute of Brewing (NRIB)  
NRPS. See Nonribosomal peptide synthetases (NRPS)  
NsdD, 288–290  
N-terminal GFP tagging, 517  
Ntr2. See Nuclear transport factor 2 (Ntr2)  
Nuclear division cycle, 235  
Nuclear envelope (NE), 266  
Nuclear export pathway, 273  
Nuclear export sequences (NES), 271  
Crm1, 273  
Nuclear localization sequences (NLS), 271  
Nuclear pore complex (NPC)  
*Aspergillus nidulans*, 262–269  
regulation during mitosis, 262–269  
structural core during mitosis, 266  
Nuclear pore complex proteins (Nups)  
identification from filamentous fungi, 262  
localization, 263  
Nuclear transport factor 2 (Ntr2), 274  
Nucleosome  
core histones, 322–327  
structure, 322–327  
Nups. See Nuclear pore complex proteins (Nups)  
Nutrition  
affecting aflatoxin formation, 461  
*Aspergillus nidulans* sexual development, 282  
Nxl transporters, 274  
One-dimensional (1D) SDS-PAGE, 450–451  
Organelles, 513–526  
Outbred stocks, 403  
Outcome measurement, 408  
Oxidative stress, 462  
PAF, 190  
Panlp module  
*Aspergillus nidulans*, 185  
edocytosis, 185  
Pathogen fungal genomics resource center (PFGRC), 295, 475–476, 479  
Pathogenicity, 377–391  
determinants, 381  
PCI, 164–166  
*Aspergillus nidulans*, 165  
PCR. See Polymerase chain reaction (PCR)  
Peanuts  
aflatoxin contamination, 464  
Penicillin biosynthesis, 154  
*Penicillium citrinum*, 8  
Pep aspartic protease, 386  
Phosphoinositide(s), 135  
gluconeogenic carbon metabolism, 135  
PGE. See Post-Golgi endosome (PGE)  
Phenylalanine, 155  
PhLP. See Phosducin-like proteins (PhLP)  
Phosducin-like proteins (PhLP), 92  
Phosphoinositide(s), 185  
binding proteins, 186  
metabolism, 186
Aspergillus nidulans sexual development, 283

Phosphoinositide phosphatases, 186
Phosphorylated substrates ubiquitylation, 157–159
pH protein production, 447
Phytochrome homolog
Aspergillus nidulans sexual development, 283
Pickles, 430
PinA, 238
PKA. See Protein kinases (PKA)
PkaA and PkaB, 93–94
PKC. See Protein kinase C (PKC)
Plant pathogen, 16
Plasma membrane lipid composition, 178
transporters, 301
Plasmid integration, 517
Plasmid-based approaches, 518
Plasmids, 532
Plasmodium falciparum transporters, 302
PLC
Aspergillus nidulans, 95
Pmd-pmb bidirectional promoter, 335
PMN. See Polymorphonuclear neutrophils (PMN)
Polarisome, 246
Polymerase chain reaction (PCR), 366–367
Polymorphonuclear neutrophils (PMN), 388
Pontecorvo, Guido, 10, 57
Positive sexual regulator, 288–291
Post-Golgi endosome (PGE), 179
ppdB, 201
ppDC
Aspergillus flavus, 201
Aspergillus oryzae, 201
Prevacuolar-endosome (PVE), 179
prn, 333–334
prnB
GFP, 309
intergenic region, 336
prnD, 335
intergenic region, 336
Proline, 307
Promoter replaced strains, 504
Propionate, 133
Protease, 449
Protease carboxypeptidase Y, 179
Protease genes, 443
Protease mutants
classical screening, 442
molecular genetic construction methods, 443–444
Protease regulators, 443
Protease-deficient fungi isolation, 446
Pteumose Lid, 164–165
Protein degradation, 157, 158
Protein dephosphorylation, 239–240
Protein induction, 449
Protein kinase C (PKC), 95
Protein kinases (PKA), 92–93
Protein production
Aspergillus, 441–456
bioprocess engineering, 450
fermentation conditions, 446–450
nitrogen metabolites, 448
strain development, 442–445
sulfur and phosphorus metabolite, 449
Proteomic analysis, 450–451
Proteomics, 113
Pseudallescheria boydii, 364
psi factor, 284
Pulmonary aspergillosis, 361
Putative amino acid transporter functions, 146
Putative DNA-binding domains, 104, 107–109
Putative GPCR, 286
Putative pheromone receptors, 285
PVE. See Prevacuolar-endosome (PVE)
Index

RNA
- dependent RNA polymerases, 201
- directed RNA polymerase complex, 204
- induced silencing complex, 197

RNA silencing, 197–210
- dicers and argonautes, 201
- future, 205
- genes evolution, 200–201
- during host infection, 205
- meiotic silencing, 203
- nature roles, 203–205
- proteins, 200
- quelling, 204

RNA-dependent RNA polymerases, 201

RNAi-mediated heterochromatic silencing, 204

Rolling and ingestion assays, 418–419

ROS. See Reactive oxygen species (ROS)

Rpa
- Aspergillus nidulans, 200, 202
- degeneration, 202

rdaA, 200

Rub1p/NEDD 8 homolog, 160

Rvs167p, 184

Saccharomyces cerevisiae
- actin patches, 182
- AP-1, 182
- AP-3, 182
- centromere He histone variant, 324–325
- chromatin, 321
- endocytic adaptors, 182
- endocytic internalization proteins, 184
- endocytosis, 179
- gluconeogenesis, 130
- gluconeogenic carbon sources, 139
- H1 histone, 328
- homologs, 181
- Las17p module, 182–183
- MT, 226
- MT plus ends, 229
- NADPH oxidase (NoxA), 284
- short DNA sequences, 494

SAD-1, 204

SAD-2, 204

SakA/HogA
- Aspergillus nidulans sexual development, 287
- signaling pathway, 124–125

Sake (Japanese alcohol beverage), 75, 430, 434–436
- koji preparation, 435–436
- main mash (moromi), 436
- manufacturing process, 435
- pasteurization and aging, 436

pressing and filtration, 436
- rice milling, 435
- rice steaming, 435
- seed mash (moto) preparations, 436

Salt, 431

Saprophytic pathogen, 377

SAT. See System A transporters (SAT)

SCF
- alternating neddylation status control, 162
- cycle regulation, 157

Schizosaccharomyces pombe
- Ago1, 204
- chromatin, 321
- Der1, 204
- interphase cells, 226
- MT plus ends, 229
- NADPH oxidase (NoxA), 284
- polarized growth model, 231
- Rdpl, 204

Sclerotia, 462

SDA. See Standardized diagnostic antigen (SDA)

SDS-PAGE, 450–451

Securin, 238–239

Seed mash (moto) preparations, 436

SekH gene, 240

Septation initiation network, 241

Septation initiation network (SIN), 245

Septum formation, 216–218

Serine palmitoyltransferase (SPT), 244

SfaD, 91

Shi, 430

Shikimate pathway, 155

Sho, 430–431

Sho-chu (Japanese spirits), 430

Shoyu, 430

Shoyu-kasu, 433

Signal transduction pathways, 88

Silencing model, 199

SIN. See Septation initiation network (SIN)

SlcA, 147

SlcB protein, 147

SlcC protein, 149

SnAREs, 185, 188

Sod, 241

SNAREs, 185, 188

Aspergillus nidulans, 188, 189

Soba-shochu, 436

SOD. See Superoxide dismutases (SOD)

Soluble nuclear transport machinery, 269

SogA gene, 262–263

SogB gene, 262–263

Southern Regional Research Center (SRRC), 465
Soybean paste (MISO) fermentation, 433, manufacturing process, 434
Soybeans, 430, 431
Soy sauce, 430 fermention, 430–431
history, 430 koji molds, 432–433
manufacturing process, 431, 432
mash pressing, 433
mash production and aging, 433
procedure, 432 raw materials, 431
refining, 433
Spatio-temporal coordinators, 293
SPB. See Spindle-pole bodies (SPB)
Spindle-pole bodies (SPB), 227, 228
AnNdc1 foci, 266
BimE APC, 239
SnaD, 241
Split marker, 507
sporulation, 462
SPT. See Serine palmitoyltransferase (SPT)
SRRC. See Southern Regional Research Center (SRRC)
Ssylp-Ptr3p-Ssy5p (SPS) system, 145
ST. See Seriogastromocystost (ST)
S-Tag, 506
Standardized diagnostic antigen (SDA), 389
SteA, 291
SteC, 287
Sterigmatocystost (ST), 199, 458
STR. See Short tandem repeats (STR)
STRE. See Stress response elements (STRE)
Streptomyces hygroscopicus, 521
Stress, 462
Stress response elements (STRE), 96
StuA. See Stunted protein (StuA)
Stunted protein (StuA), 293
Su (vinegar), 75
Subacute invasive pulmonary aspergillosis, 363
Sugars, 461
Sulfur and phosphorus metabolite, 449
Supercontig 1, linkage group VIII-R, 64
Supercontig 2, linkage group VII-R, 65
Supercontig 3, linkage group VI-R, 66
Supercontig 4, linkage group II-R, 66
Supercontig 5, linkage group III-L, 67
Supercontig 6a, linkage group IV, 67
Supercontig 6b, linkage group C-R, 67
Supercontig 7, linkage group I-L, 68
Supercontig 8, linkage group I-R, 68
Supercontig 9, linkage group IV-L, 68
Supercontig 10, linkage group IV-R, 69
Supercontig 11, linkage group II-L, 69
Supercontig 12, linkage group V-L, 69
Supercontig 13, linkage group III-R, 70
Supercontig 14, linkage group II-L, 70
Supercontig 15, linkage group VI-L, 70
Supercontig 16, linkage group III-L, 71
Superoxide dismutases (SOD), 387
SwoC1 mutant, 190
Symptomatic pulmonary aspergillosis, 363
Synaptotagmin, 186
System A transporters (SAT), 148
System biology approach, 450

T
Takadiastase, 6–7, 438
Takamine, Jokichi, 6, 438
Tannins, 463
Tao1, 347
Tao-tjo, 6
TAP/Nxtl transporters, 274
TAP-Tag, 506
TCA. See Tricarboxylic acid (TCA) cycle
TE. See Thioesterase (TE); Transposable elements (TE)
Telcontigs identified by TERMINUS, 71
Telomere-linked helicases (TLH), 60
Telomeric contigs, 71
Temperature, 462
Tetramethylrhodamine isothiocyanate (TRITC), 514
TN. See Trans Golgi network (TGN) derived vesicles
The Institute for Genomic Research (TIGR), 18, 465, 479
Thermophyl, 381
Thioesterase (TE), 460
Thioredoxins, 391
Thom, Charles, 5
TIGR. See The Institute for Genomic Research (TIGR)
TinA, 237
TinC, 240
TLH. See Telomere-linked helicases (TLH)
TLR. See Toll-like receptor (TLR)
Th4732/Th-RX A Drosophila mutants, 417, 418
TMD. See Transmembrane domain (TMD)-containing permeases
Toll pathway, 416
Toll-like receptor (TLR), 380
TOR pathway, 148 amino acid activation, 151
Trace elements, 461
Trans Golgi network (TGN) derived vesicles, 188
Transcription factors, 104–109
Aspergillus nidulans sexual development, 291 containing An(II)2Cys6, 291
transport into nucleus, 153
Transcriptional regulatory code, 110
Transcriptomics, 113
Translation initiation factor 3, 166
Transmembrane domain (TMD)-containing permeases, 177
Transposable elements (TE), 35
Transpositional rearrangement, 352
detection, 344
genome expansion, 352
graphical lists, 346
sexuality, 352
transcription, 352
Index

Transposons. See Transposable elements (TE)
Tricarboxylic acid (TCA) cycle, 130, 138–139, 154
Aspergillus nidulans, 133
carbon sources, 131
gluconeogenic carbon metabolism, 133
TRITC. See Tetramethylrhodamine isothiocyanate (TRITC)
TRizol, 475
Trypanosoma brucei transporters, 302
Tryptophan, 155
aflatoxin production, 461
precursor of terrequinone A, 155–156
T-SNARE, 179
Tumor necrosis factor-alpha, 369
Turbid sake (doburoku), 434
Turkey X disease, 9
Two-step gene replacement, 506
Tyrosine, 155
aflatoxin production, 461

U
UapA, 310
GFP, 309
UapC transporter, 311
xanthine-uric acid transporter, 311
UapC-GFP, 309
Ubiquitin ligases, 160–161
Unfolded protein response (UPR), 478
Uptake measurements, 303
Uric acid, 307
USDA/NRI Microbial Genome Sequencing Project, 18

V
Variable H3, 323
Velvet (VeA), 273
Aspergillus nidulans sexual development, 282
Versicolorin A (VERA), 460
Versicolorin B (VERB), 460
Vinegar, 75
Virulence, 380
Virulence factor, 388
Virulence studies
A. fumigatus, 422
Galleria mellonella, 423
Vital dyes, 514
Voriconazole, 367, 370
A. fumigatus, 478

W
Water
aflatoxin production, 462
soy sauce, 431
WD repeat protein, 294
WGS. See Whole genome shotgun (WGS)
Wheat, 431
Whole genome sequencing, 465
Whole genome shotgun (WGS), 77
Wild-derived inbred mice, 404

X
X-linked adrenoleukodystrophy transporters, 303
Xylazine, 406

Y
Yeasts for fermenting alcoholic beverages, 428
Yellow Fluorescent Protein (YFP), 517

Z
Zinc binuclear cluster proteins, 106
With high-quality genome sequences for the important and ubiquitous Aspergilli now available, increased opportunities arise for the further understanding of their gene function, interaction, expression, and evolution. The *Aspergilli: Genomics, Medical Aspects, Biotechnology, and Research Methods* provides a comprehensive analysis of the research that reveals the main biological attributes of these species. The co-editors are a particularly proficient and prolific pair with long track records of scientific productivity.

The book sets the stage with a discussion of basic biology, examining the data on the structure of genomes and comparing the genetic map and annotation methodologies. It includes a comparison of metabolic abilities among different *Aspergillus* spp. and other species, then covers areas such as comparative biology, pathogenic properties, and metabolic capabilities of the Aspergilli. The book reviews established techniques and new methodologies for the post-genomic era in *Aspergillus* spp. It comes with a CD containing color illustrations to supplement the text.

Filling the need for centralized information on a genus that has important economic impacts on agriculture, human health, industry, and pharmacology, the book presents a wide range of data, collected and arranged into one convenient resource. Written by a team of international experts, this is the first in-depth and exhaustive analysis of the genomics of the Aspergilli.