Edited by
Karl Esser

THE MYCOTA

A Comprehensive Treatise on Fungi
as Experimental Systems for Basic and Applied Research

Fungal Genomics

XIII

Second Edition

Minou Nowrousian
Volume Editor

Springer

The Mycota

Edited by
K. Esser

The Mycota

I *Growth, Differentiation and Sexuality*
1st edition ed. by J.G.H. Wessels and F. Meinhardt
2nd edition ed. by U. Kües and R. Fischer

II *Genetics and Biotechnology*
1st and 2nd edition ed. by U. Kück

III *Biochemistry and Molecular Biology*
1st and 2nd edition ed. by R. Brambl and G. Marzluf

IV *Environmental and Microbial Relationships*
1st edition ed. by D. Wicklow and B. Söderström
2nd edition ed. by C.P. Kubicek and I.S. Druzhinina

V *Plant Relationships*
1st edition ed. by G. Carroll and P. Tudzynski
2nd edition ed. by H.B. Deising

VI *Human and Animal Relationships*
1st edition ed. by D.H. Howard and J.D. Miller
2nd edition ed. by A.A. Brakhage and P.F. Zipfel

VII *Systematics and Evolution*
1st edition ed. by D.J. McLaughlin, E.G. McLaughlin, and P.A. Lemke†
2nd edition ed. by D. McLaughlin and J.W. Spatafora

VIII *Biology of the Fungal Cell*
1st and 2nd edition ed. by R.J. Howard and N.A.R. Gow

IX *Fungal Associations*
1st and 2nd edition ed. by B. Hock

X *Industrial Applications*
1st edition ed. by H.D. Osiewacz
2nd edition ed. by M. Hofrichter

XI *Agricultural Applications*
1st and 2nd edition ed. by F. Kempken

XII *Human Fungal Pathogens*
1st edition ed. by J.E. Domer and G.S. Kobayashi
2nd edition ed. by O. Kurzai

XIII *Fungal Genomics*
1st edition ed. by A.J.P. Brown
2nd edition ed. by M. Nowrousian

XIV *Evolution of Fungi and Fungal-Like Organisms*
Ed. by S. Pöggeler and J. Wöstemeyer

XV *Physiology and Genetics: Selected Basic and Applied Aspects*
Ed. by T. Anke and D. Weber

The Mycota

A Comprehensive Treatise
on Fungi as Experimental Systems
for Basic and Applied Research

Edited by K. Esser

XIII *Fungal Genomics*
2nd Edition

Volume Editor:
M. Nowrousian

Springer

Series Editor

Professor Dr. Dr. h.c. mult. Karl Esser
Allgemeine Botanik
Ruhr-Universität
44780 Bochum, Germany

Tel.: +49 (234)32-22211
Fax.: +49 (234)32-14211
e-mail: Karl.Esser@rub.de

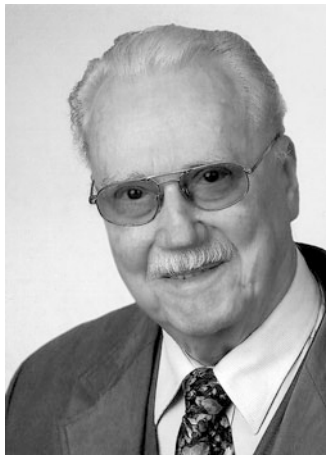
Volume Editor

PD Dr. Minou Nowrousian
Lehrstuhl für Allgemeine und Molekulare Botanik
Ruhr-Universität Bochum
Universitätsstr. 150
Bochum, Germany

Tel.: +49 (234) 32-24588
Fax: +49 (234) 32-14184
e-mail: minou.nowrousian@rub.de

ISBN 978-3-642-45217-8 ISBN 978-3-642-45218-5 (eBook)
DOI 10.1007/978-3-642-45218-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014935370



Karl Esser
(born 1924) is retired Professor of General Botany and Director of the Botanical Garden at the Ruhr-Universität Bochum (Germany). His scientific work focused on basic research in classical and molecular genetics in relation to practical application. His studies were carried out mostly on fungi. Together with his collaborators he was the first to detect plasmids in higher fungi. This has led to the integration of fungal genetics in biotechnology. His scientific work was distinguished by many national and international honors, especially three honorary doctoral degrees.

Minou Nowrousian
(born 1971) is a group leader at the Department for General and Molecular Botany at Ruhr-University Bochum (Germany). A major focus of her work is fungal development and evolution, with a focus on the development of complex multicellular structures like fruiting bodies. Her work combines molecular biology, genetics, genomics, and computational biology.

Series Preface

Mycology, the study of fungi, originated as a sub discipline of botany and was a descriptive discipline, largely neglected as an experimental science until the early years of this century. A seminal paper by Blakeslee in 1904 provided evidence for self incompatibility, termed "heterothallism", and stimulated interest in studies related to the control of sexual reproduction in fungi by mating-type specificities. Soon to follow was the demonstration that sexually reproducing fungi exhibit Mendelian inheritance and that it was possible to conduct formal genetic analysis with fungi. The names Burgeff, Kniep and Lindegren are all associated with this early period of fungal genetics research.

These studies and the discovery of penicillin by Fleming, who shared a Nobel Prize in 1945, provided further impetus for experimental research with fungi. Thus began a period of interest in mutation induction and analysis of mutants for biochemical traits. Such fundamental research, conducted largely with *Neurospora crassa*, led to the one gene: one enzyme hypothesis and to a second Nobel Prize for fungal research awarded to Beadle and Tatum in 1958. Fundamental research in biochemical genetics was extended to other fungi, especially to *Saccharomyces cerevisiae*, and by the mid-1960s fungal systems were much favored for studies in eukaryotic molecular biology and were soon able to compete with bacterial systems in the molecular arena.

The experimental achievements in research on the genetics and molecular biology of fungi have benefited more generally studies in the related fields of fungal biochemistry, plant pathology, medical mycology, and systematics. Today, there is much interest in the genetic manipulation of fungi for applied research. This current interest in biotechnical genetics has been augmented by the development of DNA-mediated transformation systems in fungi and by an understanding of gene expression and regulation at the molecular level. Applied research initiatives involving fungi extend broadly to areas of interest not only to industry but to agricultural and environmental sciences as well.

It is this burgeoning interest in fungi as experimental systems for applied as well as basic research that has prompted publication of this series of books under the title *The Mycota*. This title knowingly relegates fungi into a separate realm, distinct from that of either plants, animals, or protozoa. For consistency throughout this Series of Volumes the names adopted for major groups of fungi (representative genera in parentheses) areas follows:

Pseudomycota

Division:	Oomycota (*Achlya, Phytophthora, Pythium*)
Division:	Hyphochytriomycota

Eumycota

Division:	Chytridiomycota (*Allomyces*)
Division:	Zygomycota (*Mucor, Phycomyces, Blakeslea*)
Division:	Dikaryomycota
Subdivision:	Ascomycotina
Class:	Saccharomycetes (*Saccharomyces, Schizosaccharomyces*)
Class:	Ascomycetes (*Neurospora, Podospora, Aspergillus*)
Subdivision:	Basidiomycotina
Class:	Heterobasidiomycetes (*Ustilago, Tremella*)
Class:	Homobasidiomycetes (*Schizophyllum, Coprinus*)

We have made the decision to exclude from *The Mycota* the slime molds which, although they have traditional and strong ties to mycology, truly represent nonfungal forms insofar as they ingest nutrients by phagocytosis, lack a cell wall during the assimilative phase, and clearly show affinities with certain protozoan taxa.

The Series throughout will address three basic questions: what are the fungi, what do they do, and what is their relevance to human affairs? Such a focused and comprehensive treatment of the fungi is long overdue in the opinion of the editors.

A volume devoted to systematics would ordinarily have been the first to appear in this Series. However, the scope of such a volume, coupled with the need to give serious and sustained consideration to any reclassification of major fungal groups, has delayed early publication. We wish, however, to provide a preamble on the nature of fungi, to acquaint readers who are unfamiliar with fungi with certain characteristics that are representative of these organisms and which make them attractive subjects for experimentation.

The fungi represent a heterogeneous assemblage of eukaryotic microorganisms. Fungal metabolism is characteristically heterotrophic or assimilative for organic carbon and some nonelemental source of nitrogen. Fungal cells characteristically imbibe or absorb, rather than ingest, nutrients and they have rigid cell walls. The vast majority of fungi are haploid organisms reproducing either sexually or asexually through spores. The spore forms and details on their method of production have been used to delineate most fungal taxa. Although there is a multitude of spore forms, fungal spores are basically only of two types: (i) asexual spores are formed following mitosis (mitospores) and culminate vegetative growth, and (ii) sexual spores are formed following meiosis (meiospores) and are borne in or upon specialized generative structures, the latter frequently clustered in a fruit body. The vegetative forms of fungi are either unicellular, yeasts are an example, or hyphal; the latter may be branched to form an extensive mycelium.

Regardless of these details, it is the accessibility of spores, especially the direct recovery of meiospores coupled with extended vegetative haploidy, that have made fungi especially attractive as objects for experimental research.

The ability of fungi, especially the saprobic fungi, to absorb and grow on rather simple and defined substrates and to convert these substances, not only into essential metabolites but into important secondary metabolites, is also noteworthy.

The metabolic capacities of fungi have attracted much interest in natural products chemistry and in the production of antibiotics and other bioactive compounds. Fungi, especially yeasts, are important in fermentation processes. Other fungi are important in the production of enzymes, citric acid and other organic compounds as well as in the fermentation of foods.

Fungi have invaded every conceivable ecological niche. Saprobic forms abound, especially in the decay of organic debris. Pathogenic forms exist with both plant and animal hosts. Fungi even grow on other fungi. They are found in aquatic as well as soil environments, and their spores may pollute the air. Some are edible; others are poisonous. Many are variously associated with plants as copartners in the formation of lichens and mycorrhizae, as symbiotic endophytes or as overt pathogens. Association with animal systems varies; examples include the predaceous fungi that trap nematodes, the microfungi that grow in the anaerobic environment of the rumen, the many insect associated fungi and the medically important pathogens afflicting humans. Yes, fungi are ubiquitous and important. There are many fungi, conservative estimates are in the order of 100,000 species, and there are many ways to study them, from descriptive accounts of organisms found in nature to laboratory experimentation at the cellular and molecular level. All such studies expand our knowledge of fungi and of fungal processes and improve our ability to utilize and to control fungi for the benefit of humankind.

We have invited leading research specialists in the field of mycology to contribute to this Series. We are especially indebted and grateful for the initiative and leadership shown by the Volume Editors in selecting topics and assembling the experts. We have all been a bit ambitious in producing these Volumes on a timely basis and therein lies the possibility of mistakes and oversights in this first edition. We encourage the readership to draw our attention to any error, omission or inconsistency in this Series in order that improvements can be made in any subsequent edition.

Finally, we wish to acknowledge the willingness of Springer-Verlag to host this project, which is envisioned to require more than 5 years of effort and the publication of at least nine Volumes.

Bochum, Germany — Karl Esser
Auburn, AL, USA — Paul A. Lemke
April 1994 — *Series Editors*

Addendum to the Series Preface

During the Fourth International Mycological Congress in Regensburg (1989) while relaxing in a beer garden with Paul Lemke (USA), Dr. Czeschlik (Springer-Verlag) discussed with us the possibility to publish a series about Fungi. We both were at first somewhat reserved, but after a comprehensive discussion this idea looked promising. We decided to name this new series *The Mycota*.

Then Paul Lemke and I created a program involving seven volumes covering a wide area of Mycology. The first volume was presented in 1994 at the Fifth International Mycological Congress in Vancouver (Canada). The other volumes followed step by step. After the early death of Paul Lemke (1995) I proceeded alone as Series Editor. However for Vols. X-XII I received support by Joan Bennett.

Since evidently the series was well accepted by the scientific community and since the broad area of Fungi was not completely covered, it was decided to proceed with eight more volumes. In addition, second editions of twelve volumes were published and one more as well as several third editions are in preparation.

I would like to thank Springer-Verlag, represented by Andrea Schlitzberger for her support and cooperation.

Bochum, Germany
February 2014

Karl Esser

Volume Preface to the Second Edition

Since the first edition of Volume XIII of the *The Mycota* in 2005, a number of technical advances have dramatically changed the field of fungal genomics. Among these are the so-called next generation sequencing techniques that have transformed not only the field of genome sequencing, but can also be applied to address all manners of questions about genome organization and gene expression at a genome-wide scale. Another major development are improvements in gene deletion methods for filamentous fungi, allowing the generation of knockout strains for many, and in the case of *Neurospora crassa* (nearly) all, genes of a fungal species. Thus, while at the time of writing of the first edition most advanced genomics analyses were done with the yeast *Saccharomyces cerevisiae* and other Saccharomycetes, genomics of filamentous fungi has now truly come of age, and is used in both basic and applied research.

This is reflected in this second edition that contains 14 articles covering genomics research with ascomycetes, basidiomycetes, and basal fungal lineages, and topics ranging from basic research to pathogenicity and industrial applications.

Of course, this still does not mean that this volume can comprehensively cover all aspects of this fast-developing field. Similar to the first edition, it is rather meant to illustrate the increasing impact of genomics on fungal research by combining articles on a broad range of species and biological problems. (I apologize to those whose work or favorite topic is not covered).

The volume is divided into four sections, the first of which, *Genome Sequences and Beyond*, contains chapters that illustrate the impact of genome-based information and techniques on research ranging from model organisms like yeast (Chap. 1) to less-studied basal fungal lineages (Chap. 2). Two articles in this section highlight novel types of analyses made possible by multi-genome comparisons (Chap. 3) as well as the impact of genomics on culture collections and vice versa (Chap. 4). The second section, on *Cell and Developmental Biology*, addresses questions that are important for fungal biology, e.g. development of fungal fruiting bodies (Chap. 7), and biology in general, e.g. chromatin organization (Chap. 5) and circadian rhythms (Chap. 6). The third section, *Genomics for Biotechnology*, highlights the search for plant biomass-converting enzymes in fungal genomes (Chap. 8) and work with industrially important fungi (Chaps. 9 and 10). The fourth section focuses on *Pathogenicity* and contains chapters on genomics analyses of plant (Chaps. 11 and 12) and animal/human pathogens (Chaps. 13 and 14). Chapters in this section illustrate how genomics at all levels, from genome to metabolome, is used to study mechanisms of the interactions of fungi with other organisms.

I am extremely grateful to all authors who have made the production of this second edition possible. All are experts in their fields, and I appreciate their willingness to contribute to this project which will be useful for all scientists interested in advances in fungal genomics.

Bochum, Germany
October 2013

Minou Nowrousian
Volume Editor

Volume Preface to the First Edition

The Fungal Genomics era is gathering pace, and this is an enormously exciting time for mycologists. Having trained during an era when whole Ph.D. theses were devoted to the sequencing of a single gene, it is incredibly exciting to run research programmes in the current era, when a fungal genome sequence can be generated in a matter of weeks. Genome sequences for a whole range of fungal species have been made public in the last few years, including those for *Aspergillus fumigatus*, *Candida glabrata*, *Cryptococcus neoformans*, *Debaryomyces hansenii*, *Kluyveromyces lactis*, *Magnaporthe grisea*, *Neurospora crassa* and *Yarrowia lipolytica*. Furthermore, many more genome sequences (for other *Candida*, *Coccidioides*, *Histoplasma*, *Mycosphaerella*, *Pneumocystis* and *Saccharomyces* species, for example) will be released in the near future and, in all probability, the generation of additional fungal genome sequences will continue for some time at a rate of about one every 60 days. These genome sequences are providing a wealth of invaluable data about the evolution, life cycles, cell biology, and virulence of fungi.

Genomic technologies have developed to differing extents, depending upon the fungal species one is interested in. The genome sequences for some fungal species are only just becoming available. For these species, genomic analyses are in the very early stages of development. The initial process of annotating genes within a genome sequence is incomplete for many fungi, and the generation of microarrays for transcript profiling analysis is still some way off. In contrast, a decade has passed since the *Saccharomyces cerevisiae* genome sequence was first made public. This was one of the first genome sequences available for any cellular organism and, thanks to the sophisticated molecular toolbox available for bakers' yeast, this model organism is now leading the way in many areas of functional genomics and systems biology. (Nearly) complete sets of knockout, conditional and epitope tagged mutants are available for the genome-wide analyses of gene function in yeast. Researchers are spoilt for choice with respect to the type and format of yeast microarray which is available – oligonucleotide or gene arrays for transcript profiling, or intergenic arrays for genome-wide chromatin immunoprecipitation studies. Protein microarrays are now available for *S. cerevisiae*, and high-throughput yeast proteomics is now being exploited by numerous research groups. Massively parallel analyses of protein–protein interactions in yeast were published several years ago. Metabolomics is already being applied to system-wide analyses of yeast physiology. Therefore, for many research groups, the problem has shifted from the generation of data to the accurate interpretation and incisive exploitation of massive datasets. Thankfully, increasingly sophisticated software tools are being developed for the integration

of genetic, transcriptomic, proteomic, interactomic and metabolomic datasets. As a result, genomics is now revealing important global perspectives of fungal cell biology which were not obvious using standard reductionist approaches. There were two main challenges in editing this volume on Fungal Genomics. First, it was impossible to provide comprehensive coverage of this immense field. Second, the speed of development of Fungal Genomics precluded the publication of a text which is bang up to date. New developments do not stop for editors! Therefore, this volume does not attempt to cover every topic or to include every last-minute development. (I apologise to those whose favourite topic is not covered.) Rather, my aims have been twofold. In the 13 chapters of this 13th volume of The Mycota, my first aim has been to illustrate the current impact and potential future impact of genomics in different fungal species – in those where genomics is mature, and others where genomics is relatively immature. Second, I wanted to show how genomics is being applied to a diverse range of interesting questions in mycological research. The chapters illustrate fundamental principles of fungal genomics which are universally applicable. Mycota XIII is divided into three sections, the first of which addresses *fungal systems biology and evolution*. The volume starts with a chapter on *Saccharomyces cerevisiae*, as befits its status as the pre-eminent model organism in the genomics era. Also, this chapter addresses the metabolomics and systems biology of bakers' yeast, thereby immediately reminding the reader that there is much more to genomics than genome sequencing and transcript profiling. The other chapters in this section illustrate the huge impact of genome sequencing upon our understanding of fungal evolution, and the exploitation of novel bioinformatic approaches in the identification of genes associated with fungal virulence.

The second section, on *fungal rhythms and responses*, addresses some of the most topical issues in fungal biology – circadian rhythms, programmed cell death, stress responses, secretion and polarised growth, and cellular morphogenesis. These chapters cover a range of fungal species, including filamentous fungi and yeasts. As well as describing the massive impact of fungal genomics in these species, they highlight areas where genomics has the potential to significantly accelerate our research efforts.

The last section focuses on *fungal pathogenesis*, an area of great medical significance. Chapters in this section discuss how both proteomic and transcriptomic approaches have provided, and are helping to provide, important new insights into the pathobiology of some of the major fungal pathogens of humans.

I am very grateful to the authors for their outstanding contributions to Mycota XIII. All are internationally renowned in their chosen fields, and all are extremely busy people. I appreciate their willingness to commit valuable time to this rewarding and interesting project.

Aberdeen, UK
July 2005

Alistair J.P. Brown
Volume Editor

Contents

Pathogenicity

List of Contributors

Kyria Boundy-Mills
Phaff Yeast Culture Collection, University of California, Davis, CA, USA

Axel A. Brakhage
Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knoell Institute (HKI), Jena, Germany; and Department of Molecular and Applied Microbiology, Friedrich Schiller University, Jena, Germany

Alistair J. P. Brown
Aberdeen Fungal Group, School of Medical Sciences, Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK

Juan I. Castrillo
Cambridge Systems Biology Centre and Department of Biochemistry, University of Cambridge, Cambridge, UK

Irina Druzhinina
Research Area Biotechnology and Microbiology, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria

Jay C. Dunlap
Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH, USA

Iuliana V. Ene
Aberdeen Fungal Group, School of Medical Sciences, Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK; and Department of Molecular Microbiology and Immunology, Brown University, Providence, RI, USA

Michael Freitag
Department of Biochemistry and Biophysics, Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR, USA

Kevin Fuller
Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH, USA

Toni Gabaldón
Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Barcelona, Spain; and Universitat Pompeu Fabra (UPF), Barcelona, Spain

Donald M. Gardiner
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Plant Industry; Queensland Bioscience Precinct, St. Lucia, Brisbane, Australia

Thorsten Heinekamp
Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knoell Institute (HKI), Jena, Germany; and Department of Molecular and Applied Microbiology, Friedrich Schiller University, Jena, Germany

Kristiina S. Hildén
Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland

Jennifer Hurley
Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH, USA

Nora Koester-Eiserfunke
Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knoell Institute (HKI), Jena, Germany; and Department of Molecular and Applied Microbiology, Friedrich Schiller University, Jena, Germany

Christian P. Kubicek
Research Area Biotechnology and Microbiology, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria

Jennifer J. Loros
Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH, USA; and Department of Biochemistry, Geisel School of Medicine at Dartmouth, Hanover, NH, USA

Miia R. Mäkelä
Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland

Marina Marcet-Houben
Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Barcelona, Spain; and Universitat Pompeu Fabra (UPF), Barcelona, Spain

Kevin McCluskey
Fungal Genetics Stock Center, University of Missouri–Kansas City, Kansas City, MO, USA

Vera Meyer
Department of Applied and Molecular Microbiology, Institute of Biotechnology, Berlin University of Technology, Berlin, Germany

Benjamin M. Nitsche
Department of Applied and Molecular Microbiology, Institute of Biotechnology, Berlin University of Technology, Berlin, Germany

Minou Nowrousian
Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, Bochum, Germany

Richard P. Oliver
Australian Centre for Necrotrophic Fungal Pathogens, Department of Environment and Agriculture, Curtin University, Bentley, WA, Australia

Stephen G. Oliver
Cambridge Systems Biology Centre and Department of Biochemistry, University of Cambridge, Cambridge, UK

Monika Schmoll
Health and Environment Department, Bioresources Unit, Austrian Institute of Technology GmbH (AIT), Tulln, Austria

Bernhard Seiboth
Research Area Biotechnology and Microbiology, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria

Ekaterina Shelest
Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology– Hans-Knöll-Institute, Jena, Germany

Kar-Chun Tan
Australian Centre for Necrotrophic Fungal Pathogens, Department of Environment and Agriculture, Curtin University, Bentley, WA, Australia

Frances Trail
Department of Plant Biology, Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA

Vito Valiante
Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knoell Institute (HKI), Jena, Germany; and Department of Molecular and Applied Microbiology, Friedrich Schiller University, Jena, Germany

Kerstin Voigt
Jena Microbial Resource Collection, Leibniz Institute for Natural Product Research and Infection Biology–Hans-Knöll-Institute, Jena, Germany; and Department of Microbiology and Molecular Biology, Institute of Microbiology, University of Jena, Jena, Germany

Ronald P. deVries
CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands

Aric Wiest
Fungal Genetics Stock Center, University of Missouri–Kansas City, Kansas City, MO, USA

Genome Sequences and Beyond

1 Yeast as a Model for Systems Biology Studies on Complex Diseases

Juan I. Castrillo[1], Stephen G. Oliver[1]

CONTENTS

I. Introduction

In February 2001, *Nature* and *Science* published two landmark articles presenting the first draft of the human genome sequence (Lander et al. International Human Genome Sequencing Consortium 2001; Venter et al. 2001). On April 14, 2003, the scientific community announced the definitive completion of the **Human Genome Project** (2003), providing the essential reference genome sequence of the three billion letters of genetic code that make up what is considered a reference of every person's DNA (Weiss 2012). Unanimously recognized as one of the greatest milestones in biology, with more refined studies (1000 Genomes Project Consortium 2010; Chen et al. 2013) and the **Genome Reference Consortium** responsible for regular updates and assembly (Genome Reference Consortium), the fact is that we are still far from understanding the exquisite complexity of human biology and the molecular basis and mechanisms underlying complex diseases, even at the basic cellular level (Heard et al. 2010; Alberts 2011; Hayden 2010; Marshall 2011; Vidal et al. 2011; Gonzaga-Jauregui et al. 2012; Castrillo et al. 2013; Human Genome Organisation; Human Genome Project 2013). While it is clear that next-generation sequencing (NGS) and advanced postgenomic technologies (e.g., exome sequencing; whole-genome sequencing, WGS) (Casals et al. 2012; DeFrancesco 2012; Ku et al. 2012) are delivering benchmark examples of applications in diagnostics and treatment of diseases with a basic genomic component (e.g., Mendelian and, as yet, uncharacterized diseases) (Maher 2007, 2013; Choi et al. 2009; Davies 2010; Mayer et al. 2011; Worthey et al. 2011; Auffray et al. 2012; Calvo et al. 2012; Evans and Berg 2012; Gilissen et al. 2012; Li et al. 2012; Puffenberger et al. 2012; Evans et al. 2013; Young 2013), several complex diseases will not be solved solely by

[1]Cambridge Systems Biology Centre and Department of Biochemistry, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK; e-mail: jic28@cam.ac.uk; steve.oliver@bioc.cam.ac.uk

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

the application of more powerful sequencing approaches (Vidal et al. 2011; Castrillo et al. 2013).

The reality is that many **complex diseases** are being revealed as **multifactorial in nature** (Pang et al. 2000; Heard et al. 2010; Monico and Milliner 2011; Vidal et al. 2011; Castrillo and Oliver 2011a; Castrillo et al. 2013), involving a combination of genomic, epigenomic, and environmental factors, with a dynamic interplay at the molecular, cellular, and physiological levels (see below). While advanced sequencing and molecular techniques will continue to be invaluable in the identification of genomic loci and their specific contribution (Pang et al. 2000; Goldstein et al. 2013; Maher 2013; Goh 2013), the challenge of multifactorial diseases in intrinsically complex systems calls for new advanced holistic studies.

Many complex phenotypes may be more directly **related to alterations in global, "systems," "network" properties** than they are to particular DNA sequences (Vidal in Heard et al. 2010). The idea that multiscale complex systems formed by interacting macromolecules arranged in dynamic modular complexes and networks underlie some of the most fundamental aspects of life was proposed by a few visionaries half a century ago (see Vidal 2009 and references therein). The new era of biology is not so much concerned with building blocks or inventories of working parts, but rather with how those parts interact and are arranged into dynamic functional modules, essential complexes, and networks (e.g., DNA-protein; RNA-protein; protein-protein; protein-metabolite networks) to produce units of biological organization whose properties are much greater than the sum of their parts. They result in dynamic, exquisite choreographies that ultimately orchestrate and sustain "life," extant viable "biological systems," with their flexibility and robustness, having been optimized through millions of years of evolution (Darwin and Wallace 1858; Darwin 1859; Dobzhansky 1964; Noble 2006). This is what **systems biology** is all about—its ultimate challenge being what makes complex networks and systems sustainable and viable and how complex diseases can be contemplated as arising from "altered networks states" (Friend 2010) that are the consequence of multifactorial perturbations, whose effects and dynamics can perhaps best be studied initially by experiments in model organisms (Kitano 2002; Nurse 2003; Noble 2006; Vidal 2009; Aitman et al. 2011; Heard et al. 2010; Arkin and Schaffer 2011; Castrillo and Oliver 2011a; Kruger 2011; Walhout et al. 2011; Castrillo et al. 2013) (see below).

In human biology, progress is being made, with metabolic networks carefully reconstructed and refined (Stobbe et al. 2011; Vidal et al. 2011; Bordbar and Palsson 2012; Wang et al. 2012). However, this is limited. As an example, current estimates indicate that actual protein interactome maps may cover less than 10 % of all potential human regulatory protein-protein interactions (Barabsi et al. 2011; Liu et al. 2011a, b). Fundamental principles in systems biology will always need to be studied and validated in several model organisms that represent biological organization from the cellular level to that of tissues, organs, and complete multicellular organisms (Chang and Grieder 2008; Aitman et al. 2011; Murphy 2013; National Institutes of Health). The distinctive properties of each model organism should always be interpreted in the light of evolution (Darwin and Wallace 1858; Darwin 1859; Telford 2013); as Dobzhansky stated: "Nothing makes sense in biology except in the light of evolution" (Dobzhansky 1964).

As a matter of principle, essential discoveries in biology need to start at the cellular level, with advances in the understanding of the exquisite complexity of the basic unit of life, "the eukaryotic cell" (Nurse and Hayles 2011; Castrillo and Oliver 2011a, b). In fact, since the **molecular bases of diseases are likely to be manifested primarily at the cellular level** (e.g., as early dysregulations of cellular networks at the presymptomatic stages of the disease), this makes the eukaryotic cell a fundamental subject of study to elucidate how alterations in specific mechanisms and pathways result in early dysregulation of cellular networks and it is hoped provide candidate strategies with which to overcome or minimize their effects at the earliest stage.

Supracellular characteristics of complex diseases are beyond the reach of investigations with unicellular organisms. Nevertheless, the

optimal characteristics of the yeast *Saccharomyces cerevisiae*, with main molecular mechanisms, biological networks, and subcellular organization essentially conserved in all eukaryotes (Lehninger 1975; Alberts et al. 2008; Castrillo and Oliver 2004, 2011a; Castrillo et al. 2013 and references therein), makes it an optimum model eukaryote for the elucidation of systems biology principles at the cellular level (see below). As a prime example, studies of the dynamics of transcriptional regulatory networks in *S. cerevisiae* have revealed large topological changes depending on environmental conditions, with transcription factors altering their subcellular localization and interactions in response to stimuli, some of them serving as permanent hubs, but most acting transiently in specific conditions only (Luscombe et al. 2004; Balaji et al. 2006; Castrillo and Oliver 2011a).

Comprehensive systems biology experiments with yeast have the potential to uncover the complexity and interplay of eukaryotic networks and short- and long-term effects of perturbations and **dysregulation of networks** that may illuminate **the origin and sequence of events underlying complex diseases**. These approaches may be reproduced and validated in more complex model organisms, including mammalian cells, patient-specific cell lines from induced pluripotent stem cells (iPSCs), animal models, and ultimately, humans (Chang and Grieder 2008; Soucek et al. 2008; Aitman et al. 2011; Ebert et al. 2012; Shi et al. 2012; Dianat et al. 2013; Marx 2013; Murphy 2013; Musunuru 2013; National Institutes of Health). In this chapter, the potential of *S. cerevisiae* as an optimal model for comprehensive studies of altered mechanisms and dynamics of essential eukaryotic networks underlying complex diseases is presented. It focuses on the need for careful experimental designs reproducing in vivo perturbations. These include transitions from initial steady states subjected to short and sustained perturbations in time-course experiments, longitudinal studies monitored at multiple 'omics levels, with effective bioinformatics analyses and data integration. Examples of yeast disease models, latest advanced studies, and comprehensive strategies toward the characterization of early dysregulation and dynamics of modular networks in complex diseases are presented.

II. Yeast for Systems Biology Studies on Essential Mechanisms and Networks Underlying Complex Diseases

A. Yeast as a Model in Eukaryotic Biology and Human Disease

1. Yeast as a Model in Eukaryotic Biology and Systems Biology

Saccharomyces cerevisiae is a species of budding yeast, a group of unicellular fungi belonging to the phylum Ascomycota. This yeast has been used as a model eukaryote because the basic mechanisms of genome replication and repair, cell division, gene expression, RNA processing, translation, signal transduction and regulatory networks, central metabolism, organelles and subcellular organization, protein quality control, vesicular trafficking and autophagic **pathways are essentially conserved between yeast and higher eukaryotes** (Rose and Harrison 1987–1995; Sherman 1998, 2002; Castrillo and Oliver 2004, 2011a; Feldmann 2012a; Castrillo et al. 2013; Westermann and Schleiffer 2013). The basics of eukaryotic biochemistry, as encompassed in a complete map of central metabolic pathways, were first unveiled in *S. cerevisiae* (Lehninger 1975; Alberts et al. 2008), and a wide knowledge of the genetics, biochemistry, and physiology of this yeast is presently available (Rose and Harrison 1987–1995; Oliver 1997, 2002; Sherman 1998, 2002; Castrillo and Oliver 2004, 2011a, b; Stansfield and Stark 2007; Feldmann 2012a). Importantly, the yeast *S. cerevisiae* was the first eukaryotic organism for which the complete genome was sequenced (Goffeau et al. 1996; Kaback 2013) and for which strategies for the proper annotation, curation, and standards initiatives for maintenance of high-quality curated databases and data repositories were implemented (e.g., **Saccharomyces Genome Database**, [SGD], http://www.yeastgenome.org).

Saccharomyces cerevisiae is a particularly suitable organism for biological studies thanks to its status as generally regarded as safe (GRAS), rapid growth, and simple methods of cultivation under defined conditions. It is also a well-defined genetic system with (a) simple techniques of genetic manipulation (Amberg et al. 2005; Stansfield and Stark 2007); (b) continued incorporation of the latest advanced molecular approaches and essential tools (e.g., gateway vectors; genome engineering using CRISPR-Cas systems, see Nagels et al. 2012; Dicarlo et al. 2013; Giuraniuc et al. 2013; RNA interference [RNAi], see Drinnenberg et al. 2009; Moazed 2009; Wang et al. 2013); and (c) the majority of high-throughput postgenomic technologies, including NGS (e.g., ChIP-exo, DNA and RNA sequencing); advanced proteomics; ribonucleoproteins (RNPs); metabolomics; and interactomics being first developed or validated in yeast (Castrillo and Oliver 2006, 2011a, b; Castrillo et al. 2013; Babu et al. 2012; Nookaew et al. 2012; Yen et al. 2012; Ibañez et al. 2013; Klass et al. 2013; Mitchell et al. 2013; Picotti et al. 2013; Song and Singh 2013; Sung et al. 2013). *Saccharomyces cerevisiae* is used as a model organism to study cell growth, the cell cycle, checkpoints and their relation to cancer, evolution, control of pre-mRNA splicing, translation, protein folding, protein homeostasis (proteostasis) and maintenance networks, aging and extension of lifespan, and for basic studies of the molecular basis of diseases (Castrillo and Oliver 2011a; Feldmann 2012b; Castrillo et al. 2013; Lang et al. 2013; see also below).

At this point, a relevant question arises: Can new discoveries in eukaryotic biology, molecular, and "systems" principles relevant to complex diseases still be made using yeast? The answer is yes (Castrillo and Oliver 2011a and references therein). As an example, the groups of Steinmetz, Snyder, and several others are revealing extensive transcriptional heterogeneity and the mechanisms underlying transcriptional directionality in yeast (Tan-Wong et al. 2012; Pelechano et al. 2013; Waern and Snyder 2013), fundamentally conserved from yeast to human (Jacquier 2009; Djebali et al. 2012; Goodman et al. 2013), with essential mechanisms of control of pervasive transcription, directionality, cryptic transcripts, and RNA quality control the subjects of continuing investigations (Neil et al. 2009; Jacquier 2009; Almada et al. 2013; Porrua and Libri 2013a, b). In addition, the latest studies on the dynamics and interplay of interactome networks are revealing not only context-dependent dynamic changes in network topologies (Luscombe et al. 2004; Balaji et al. 2006), but, more importantly, that **essentiality in eukaryotic networks may lie primarily in "modules"** (including protein and RNP complexes); important implications arise from this highly modular landscape (Gstaiger and Aebersold 2013; Ryan et al. 2013; Song and Singh 2013). Thus, Ryan and colleagues claimed: "With more information from organisms and additional species ... it may be possible to define a core set of complexes that are essential in all eukaryotes under all conditions ... and a set of 'conditionally essential' complexes that may be associated with specific environments or morphology" (Ryan et al. 2013).

Such a conclusion was implicit in the results of comprehensive systems biology studies of balanced cell growth in *S. cerevisiae* (e.g., Castrillo et al. 2007), in which a core group of significantly up- and downregulated genes with growth rate, irrespective of the nutrient-limiting condition, appeared highly conserved sharing orthologs in a majority of eukaryotes from yeast to humans. This pointed to the existence of an essentially conserved "core protein machinery" (complexes and interactions) that governs eukaryotic cell growth, with others associated with each specific environmental (nutrient-limiting) condition (Castrillo et al. 2007). This is being confirmed by more recent yeast studies, revealing global "growth" and "stress" modular networks that are remarkably conserved phylogenetically (Thompson et al. 2013; Roy et al. 2013). The concept of **"modular essentiality," with a core machinery common in all eukaryotes, and other context- or organism-specific modules** (e.g., species-specific stress response pathways; Das et al. 2013) is **of central importance for understanding biology at a "systems" level** (Castrillo et al. 2013; Gstaiger and Aebersold

2013; Ryan et al. 2013; Song and Singh 2013) (see also below). For more detailed examples of advanced pre- and postgenomic discoveries in eukaryotic biology made using *S. cerevisiae* as a reference model, refer to the work of Rose and Harrison (1987–1995), Castrillo and Oliver (2004, 2011a, b), Feldmann (2012a), Castrillo et al. (2013) and references therein.

2. Yeast as a Model for Molecular and Systems Biology Studies of Human Diseases

Yeast has already been used extensively as a model organism to investigate the molecular basis of disease (Foury 1997; Johnson et al. 2008; Franssens et al. 2010; De Vos et al. 2011; Munkacsi et al. 2011; Ocampo and Barrientos 2011; Treusch et al. 2011; Mayfield et al. 2012; D'Angelo et al. 2013; Tenreiro et al. 2013; see also below), but can yeast now be used as a model to obtain new insights into multifactorial complex diseases at a systems level? First, it is important to differentiate the following:

1. Complex diseases with genomic components (e.g., genome sequences, mutations, structural variations, possible epigenomic features) significantly contributing to the disease (so-called familial diseases). These include (a) diseases with well-annotated, identified genomic "culprits," and (b) those with as-yet-unidentified components but whose symptoms (e.g., severity or early manifestation at the prenatal or early infancy stages) combined with family history point to a clear genomic/epigenomic contribution, which could potentially be identified by application of the latest NGS technologies.
2. Complex diseases with no clear genomic/epigenomic component contributing significantly to the disease. In these cases, external environmental insults and potential decrease of the efficacy of homeostatic and defense responses will be of most relevance (see below).
3. Multifactorial complex diseases with both genomic/epigenomic and environmental factors contributing to the disease (see also next section).

Saccharomyces cerevisiae, as a reference organism, is well positioned to be a touchstone model for studies of diseases with identified genomic culprits (e.g., diseases 1a and 3 in the list). In fact, the essential core modular architecture and networks, conserved from yeast to human referred to previously (Castrillo et al. 2007, 2013; Castrillo and Oliver 2011a), is reflected in the **existence of a relevant number of human disease genes with orthologs or functional counterparts in yeast** (e.g., see Table 1.1 and references therein for information on relevant congenital, metabolic, and neurological diseases). This remarkable fact, with its huge potential, opens the way to comprehensive studies on the contribution of individual genomic components in this model organism (e.g., using transgenic yeast mutants) in well-designed experiments minimizing bias (Cohen and Medley 2005; Piston 2012; Sarewitz 2012). In most common approaches, phenotypic studies of knockout, knockin mutants (constitutive or conditional) and mutants of overexpression of the endogeneous protein ortholog can be performed and the results analyzed together with those stored in well-curated databases (e.g., SGD). However, since complex biological systems cannot be reduced to a single circuit or network (i.e., the presence or absence of a specific component may result in the activation of others with new dynamic interplay with intertwined networks), the results may not be easy to interpret. Future studies will benefit from the progressive incorporation of more advanced strategies, such as conditional regulatable promoters and strains (Yen et al. 2003; Wishart et al. 2006; Kaufmann and Knop 2011) and the CRISP-mediated RNA-guided platform for tailored regulation of transcription (Gilbert et al. 2013), toward the construction of yeast models with tailored, inducible, finely controlled, tunable expression of the specific "altered" gene that is reported to contribute to the disease. As a general principle, **the main objective of a disease model** (in our case, a yeast model) will be that the alteration of specific mechanism(s) in a simpler organism will **reproduce or recapitulate early-stage** (i.e., cellular) **features and phenotypes observed in humans** and may be followed by more advanced and comprehensive studies (see below).

Table 1.1. Selected congenital, metabolic, and neurological diseases with relevant human disease genes related to *S. cerevisiae* genes and their function

Human disorder/disease: phenotype (gene effect/defect)	Human gene(s)	Yeast gene(s)	Yeast gene annotated function(s)
Aging, premature aging syndrome Hutchinson-Guilford progeria, HPGS	ZNPSTE24	*STE24*	Zinc metalloprotease a-factor maturation; proteolysis
Argininemia: mental retardation, microcephaly, and spasticity	ARG1	*CAR1*	Arginase
Autism features	ADSL	*ADE13*	Adenylosuccinate lyase
Blinding, ocular disorders (see also Niemann-Pick disease C2)			
Galactosemia II: cataract formation	GALK1	*GAL1*	Galactokinase
Retina and choroid atrophy progressive blinding	CAR2	*CAR2*	L-Ornithine transaminase
Retinitis pigmentosa 2	RP2 (cognate G protein-GAP pair conserved in yeast as Cin4-Cin2)	*CIN2*	GTPase-activating protein (GAP) for Cin4p
Citrullinemia	ASS1	*ARG1*	Argininosuccinate synthase
Citrullinemia, type II	SLC25A12 (aralar), SLC25A13 (citrin)	*AGC1*	Mitochondrial amino acid transporter
Cystine transport disorders	CTNS	*ERS1*	Cystine transporter
Diabetes mellitus, non-insulin-dependent	PPP1R3	*GIP2*	Protein phosphatase 1 regulator
Early embryonic lethality	IPK2	*IPK2*	Inositol polyphosphate multikinase
Galactosemia: learning disabilities and speech/motor dysfunction	GALT	*GAL7*	Galactose-1-phosphate uridyltransferase
Galactosemia III: impaired liver function and mental retardation	GALE	*GAL10*	Galactose epimerase
Glia maturation factor	GMF	*AIM7*	Stimulates Arp2/3 complex
Glycosylation disorders Congenital disorders of glycosylation			
CDG-Ia	PMM2	*SEC53*	Phosphomannomutase-2 deficiency; congenital disorder of glycosylation type I-a; Jaeken syndrome
CDG-Ib	MPI (PMI1)	*PMI40*	Mannose-6-phosphate isomerase
CDG-Ic	hALG5	*ALG5*	Transmembrane dolichyl-phosphate beta-glucosyltransferase
CDG-Ic	hALG6	*ALG6*	α-1,3-Glucosyltransferase
CDG-Id	hALG3	*ALG3*	α-1,3-Mannosyltransferase
CDG-Ie	DPM1	*DPM1*	Dolichol phosphate mannose synthase
CDG-Ig	ALG12	*ALG12*	α-1,6-Mannosyltransferase
CDG-Ih	hALG8	*ALG8*	α-1,3-Glucosyltransferase
CDG-Ii	hALG2	*ALG2*	Mannosyltransferase
CDG-Ij	DPAGT1	*ALG7*	Dolichyl-phosphate-dependent N-acetylglucosamine-1-phosphate transferase

(continued)

Table 1.1. (continued)

Human disorder/disease: phenotype (gene effect/defect)	Human gene(s)	Yeast gene(s)	Yeast gene annotated function(s)
CDG-Ik	ALG1	*ALG1*	β-1,4-Mannosyltransferase
CDG-Il	ALG9	*ALG9*	α-1,2-Mannosyltransferase
CDG-Iq	SRD5A3	*DFG10*	Polyprenol reductase
CDG-IIb	GCS1	*GLS1*	Glucosidase
CDG-IIk	GDT1	*TMEM165*	Involved in calcium homeostasis
Glycosylation defects	SIMP, ITM1	*STT3*	Component of OST complex
Degradation of misfolded glycosylated proteins, slowed	GANAB	*ROT2*	Glucosidase II
Glucose-6-phosphate dehydrogenase deficiency: hemolytic anemia	G6PD	*ZWF1*	Glucose-6-phosphate dehydrogenase
Glutamate dehydrogenase defects, hyperammonemia, hyperinsulinism, neurological disorders	GLUD1, GLUD2	*GDH2*	Glutamate dehydrogenase
Hemo biosynthesis. Hereditary coproporphyria (HCP); psychiatric symptoms	CPOX	*HEM12*	Coproporphyrinogen oxidase
Homocystinuria	CBS and human methyltransferases	*CYS4* and yeast methyltransferases	Cystathionine β-synthase; AdoMet-dep. methytransferases
Hyperammonemia/ hyperornithinemia/ homocitrullinuria (HHH) syndrome	SLC25A15	*ORNT1*	Ornithine transporter (MIM)
Hyperglycerolemia	hGUT1	*GUT1*	Glycerol kinase
Hyperhomocysteinemia; psychosomatic symptoms	MHTFR	*MET12*	Methylenetetrahydrofolate reductase
Hyperoxaluria, primary, type 1	AGT	*AGX1*	Peroxisomal alanine: glyoxylate aminotransferase
Hypermethioninanemia: mental retardation	MAT1A/MAT1B	*SAM1/2*	S-Adenosylmethionine synthase
Hyperprolinemia: proline accumulation; mental retardation and psychiatric symptoms	hPUT1 hPUT2	*PUT1* *PUT2*	Proline oxidase Δ1-Pyrroline-5-carboxylate dehydrogenase
Hypertension	SA	*ACS1*	ACS1 acetyl-CoA synthetase
Menke's disease (MNK): occipital horn syndrome; neurodegeneration//and Wilson disease (hepatolenticular degeneration)	ATP7A// ATP7B	*CCC2* *CCC2*	Copper-transporting ATPase Copper-transporting ATPase
Migraine (CACNL1-A4): familial hemiplegic migraine	CACNA1A	*CCH1*	Calcium channel
Neurodegeneration (Alzheimer's disease risk factor)	PICALM	*YAP1802 (YGR241C)*	Protein of the AP180 family, involved in clathrin cage assembly; binds Pan1p and clathrin
Neurodegeneration: atypical amyotrophic lateral sclerosis (ALS) type 8	VAP-A VAP-B VAP-C	*SCS2*	Integral ER membrane protein (VAMP/synaptobrevin-associated protein)

(continued)

Table 1.1. (continued)

Human disorder/disease: phenotype (gene effect/defect)	Human gene(s)	Yeast gene(s)	Yeast gene annotated function(s)
Neurodegeneration: familial amyotrophic lateral sclerosis (FALS); Lou Gehrig's disease	SOD1	*SOD1*	Copper-zinc superoxide dismutase
Neurodegeneration: Friedreich's ataxia; neuropathy; cardiomyopathy; diabetes	FXN, frataxin	*YFH1*	Mitochondrial matrix iron chaperone; frataxin homolog
Neurodegeneration: progressive ataxia-oculomotor, apraxia type 2, and juvenile ALS type 4	SETX	*SEN1*	RNA/DNA helicase; polymerase II termination factor
Neurodegeneration: ataxia; spinocerebellar ataxia 6; episodic ataxia 2	CACNA1A	*CCH1*	Calcium channel
Neurodegeneration: Batten disease	CLN3	*YHC3/BTN1*	Vacuolar membrane protein involved in the transport of arginine into the vacuole
Neurodegeneration: myotonic dystrophy (DM)	DMPK	*CBK1*	Serine/threonine protein kinase
Neurodegenerative peroxisome biogenesis disorders (PBDs): mental retardation; neuronal, hepatic, renal abnormalities; early infancy death	PEX1/PEX6	*PEX1/PEX6*	Peroxisomal AAA proteins
	PEX5	*PEX5*	Peroxisomal receptor (PTS1)
	PEX10	*PEX10*	E3 ubiquitin ligase for Pex5p
	PEX12	*PEX12*	E3 ubiquitin ligase
Neurodegenerative, peroxisomal fatty acid beta oxidation disorder; adrenoleukodystrophy	ABCD1	*PXA1*	Peroxisomal ABC transporter
Neurodegenerative, sphingolipids, lipid storage dysfunction; Niemann-Pick disease C1	NPC1	*NCR1*	Vacuolar membrane protein
Niemann-Pick disease C2; ocular disease; neurodegeneration	NPC2	*NPC2*	Cholesterol-binding protein
Neurodegenerative, spinocerebellar ataxia 17 (trinucleotide repeats)	TBP	*SPT15*	TATA-binding protein
Neurodegenerative, suppression of proteotoxicity in proteinopathies	NMNAT	*NMA1/NMA2*	NAD+biosynthesis pathways: NAD+and nicotinamide riboside salvage pathways
Neurodegenerative tautophaties; tau hyperphosporylation; Alzheimer's disease	GSK-3β	*MDS1*	Protein kinase
	CDK5	*PHO85*	Cyclin-dependent kinase
Neurodegenerative, trinucleotide repeats disorders; Huntington's and fragile X disease	FEN1	*RAD27*	Endonuclease, required for Okazaki fragment processing and maturation
Neurological disorders: epilepsy, familial focal	DEPDC5	*SEA1*	Part of SEA complex involved in membrane trafficking
Neuromuscular: Charcot-Marie-Tooth disease; distal motor neuropathy; hereditary cataracts	Small HSPs	*HSP26/HSP42*	Small heat-shock proteins
Neuron development: migration; brain malformation (lissencephalitis)	LIS1	*PAC1*	Protein involved in dynein/dynactin pathway

(continued)

Table 1.1. (continued)

Human disorder/disease: phenotype (gene effect/defect)	Human gene(s)	Yeast gene(s)	Yeast gene annotated function(s)
Neutropenia; myelofibrosis; progressive bone marrow failure	VPS45	*VPS45*	Protein involved in vacuolar protein sorting, membrane trafficking through the endosomal system
Parkinson disease: misfolded α-synuclein accumulation	DJ-1	*HSP31/32*	Cysteine protease and chaperone
Parkinson disease: familial, late onset	VPS35	*VPS35*	Endosomal subunit of membrane-associated retromer complex for retrograde transport
Succinic semialdehyde dehydrogenase deficiency; developmental and neurological abnormalities	ALDH5A1	*UGA2*	Succinic semialdehyde dehydrogenase for assimilation of gamma-aminobutyrate (GABA)
α-Thalassemia (ATRX); mental/psychomotor retardation	RAD54	*RAD54*	DNA-dependent ATPase, stimulates strand exchange
Triose phosphate isomerase deficiency; hemolytic anemia and progressive neurological disorder	TPI1	*TPI1*	Triose phosphate isomerase
Velo-cardio-facial and DiGeorge syndromes (VCFS, DGS, respectively)	CDC45	*CDC45*	DNA replication initiation factor
X-linked mental retardation-30 disease	PAX3	*STE20*	p21-Protein kinase

For more detailed information, see the following references: Feldmann (2012b) and references therein and Foury (1997), Wilson and Roof (1997), Pearce and Sherman (1998), (1999), Heinicke et al. (2007), Treusch et al. (2011), Mayfield et al. (2012), Morava et al. (2012), Pereira et al. (2012), Demaegd et al. (2013), Ishida et al. (2013), Michaelis and Hrycyna (2013), Ocampo et al. (2013), Porrua and Libri (2013a), Pryor et al. (2013), Stepensky et al. (2013), Tehlivets et al. (2013), Tenreiro et al. (2013); *Saccharomyces* genome database (SGD; www.yeastgenome.org); Online Mendelian Inheritance in Man (OMIM) database, http://www.ncbi.nlm.nih.gov/omim

When the genomic "culprit" has no counterpart in yeast, studies with "humanized" yeasts (i.e., yeast models based on heterologous expression of the human wild-type or mutant coding sequence can be performed) (e.g., Johnson et al. 2008; other examples below). Such models often provide valuable insights and illuminate the molecular basis of the disease. However, the heterologous expression of human genes in yeast, a eukaryotic host with its own intrinsic networks, proteome environment, and so on, may not always be successful and may produce results with many potentially confounding factors. Heterologous expression of the human gene when endogenous orthologs already exist in yeast can be even more challenging. Whatever the case, as a matter of principle, all **yeast studies should be validated in other disease models, higher eukaryotes, and ultimately humans.**

A second challenge is whether *S. cerevisiae* can also be a good model to study complex diseases without a clear genomic culprit. Here, we need to realize that, among the many contributing factors to these complex diseases (which are not well characterized, and are sometimes called "sporadic" or "late onset," if appearing in old age) are environmental external insults combined with intrinsic homeostatic or stress-defense networks, which may decrease in efficacy during prolonged lifetime (i.e., *aging*, defined as "a decrease in efficiency of maintenance"; Holliday 1995, 2006). The most remarkable fact here is that the majority of these essential defense networks (including, e.g., DNA, RNA, and protein quality control; proteostasis; and autophagic mechanisms and pathways) are again essentially conserved from yeast to humans (Holliday 1995, 2006; Castrillo and Oliver 2011a; Breitenbach et al. 2012;

López-Otín et al. 2013). Comprehensive studies in yeast at the cellular level will be invaluable in obtaining proper characterization of such processes. Dynamic studies of these homeostatic networks under specific perturbations in comprehensive experiments (see below), and the identification of factors contributing to their decrease in efficacy, their individual or synergistic effects that need to be counteracted or minimized, will be required.

Finally, can *S. cerevisiae* be a good model to investigate multifactorial complex diseases, with both genomic and environmental contributions? The answer is yes, not only thanks to all the favorable characteristics mentioned, but also because it is in a simple organism like yeast where comprehensive experiments to "dissect" the different contributions of individual genomic components from those arising from specific environmental perturbations can be properly designed and readily implemented. **This requires yeast models that recapitulate the early, cellular, features of the disease, together with dynamic systems biology studies reproducing in vivo perturbations** (e.g., in time-course experiments monitored at different 'omics levels; see below). Toward this objective, we need to go back to basics, rediscovering the architecture and organization of networks in the **eukaryotic cell as a "system"** (Castrillo and Oliver 2011a; Castrillo et al. 2013).

The exquisite complexity of biological systems at the molecular level, with thousands of components (e.g., genes, transcripts, proteins, metabolites) interacting in finely tuned, dynamic networks (Castrillo and Oliver 2006, 2011a) is reflected in Fig. 1.1, which shows the whole spectrum of molecular interactions, essential networks, and interactomes in the eukaryotic cell, with their "**modular" complexes** (e.g., essential RNP and protein complexes). These modules and interactions are responsible for the dynamic interplay between 'omics levels resulting in a particular phenotype.

In a living eukaryotic cell, the basic molecular networks presented in Fig. 1.1 dispose themselves in higher units of organization, supramolecular networks, "modules" and subcellular organelles, with specialized functions and exquisite interplay, with an important remarkable fact: This "function-related" or "function-driven" supramolecular subcellular organization, whose integrity or dysregulation will be of central importance in complex diseases, is once more fundamentally conserved in all eukaryotes. The basic architecture of this "ecosystem of organelles" that constitutes the eukaryotic cell (Klitgord and Segrè 2010) is essentially conserved from yeast to human through millions of years of evolution (Darwin and Wallace 1858; Darwin 1859; Dobzhansky 1964; Lehninger 1975; Alberts et al. 2008). This makes yeast also an excellent model for studies of diseases characterized by subcellular or organellar dysfunction (e.g., mitochondrial disorders; Baile and Claypool 2013; Montanari et al. 2013).

A schematic representation of the main architecture, subcellular organization, and networks of the eukaryotic cell during "balanced" growth (the most oft-studied case) is presented in Fig. 1.2. Basically, environment-sensing networks that monitor external conditions and perturbations and combined with intracellular indexes are integrated in essential control modules, "controllers" (e.g., TOR complexes; Loewith and Hall 2004), which recognize and configure the "cellular state," activating signal transduction pathways with (re)programming of gene expression, metabolic and regulatory networks toward high-energy-yielding metabolism and biosynthetic fluxes sustaining balanced cell growth (Hall et al. 2004; Castrillo et al. 2007). Of particular relevance is the fact that, under these conditions, the essential defense, stress, networks responsible for rapid recovery from perturbations (e.g., DNA repair, RNA and protein quality control homeostasis networks) (Holliday 1995, 2006; Breitenbach et al. 2012; López-Otín et al. 2013) are downregulated, and their essential role in recovery to homeostatic status cannot be properly studied. For complex diseases in which environmental perturbations and counteracting homeostatic defense maintenance networks are relevant, comprehensive time-course experiments in simple model organisms such as yeast can be designed to reproduce in vivo perturbations and reveal the activation of these homeostatic networks and their interplay with the core eukaryotic networks (see below).

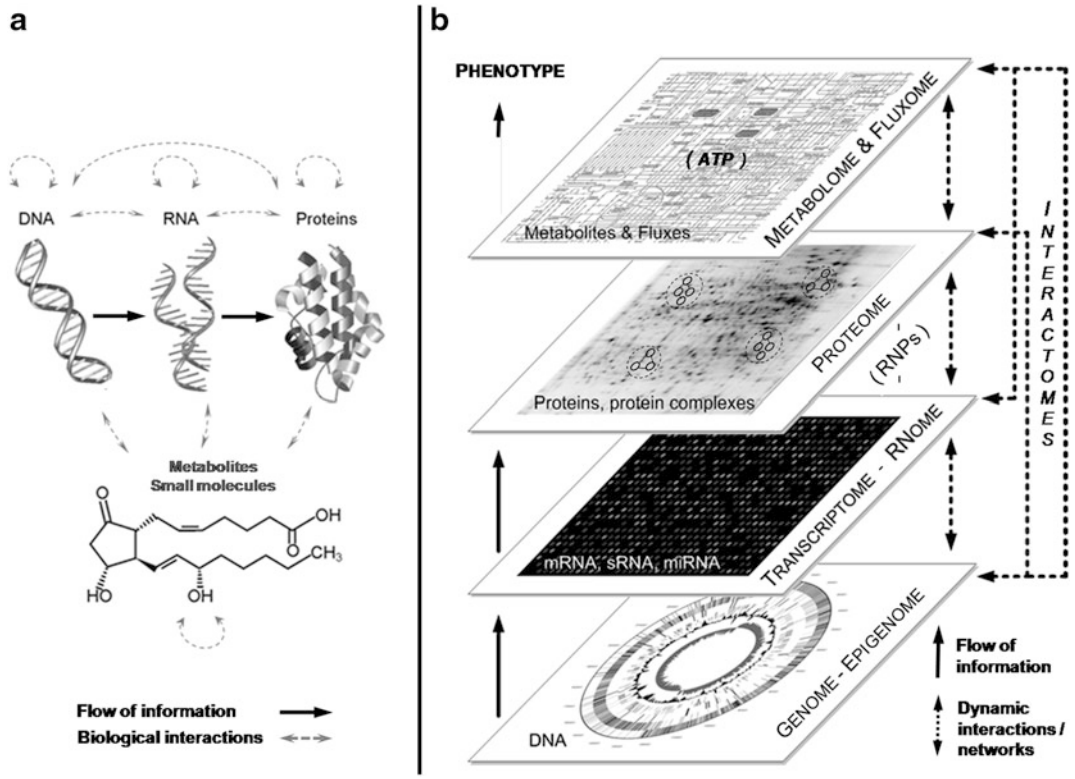


Fig. 1.1. Molecular interactions and networks in the eukaryotic cell, with modular complexes responsible for dynamic interplay between 'omics levels. (**a**) Flow of genetic information, from DNA to RNA to protein (Crick 1958, 1970; Alberts et al. 2008) (*solid arrows*) and spectrum of biological interactions/networks in the eukaryotic cell: DNA-DNA; DNA-RNA; DNA-protein; DNA-metabolites; RNA-RNA; RNA-protein (*RNP*); RNA-metabolites; protein-protein; protein-metabolites and metabolite-metabolite interactions (Adapted from Antunes et al. 2011; F1000 Biol Rep. 3:4 with permission from Faculty of 1000 Ltd [F1000 Ltd; http://f1000.com] and authors; and Castrillo et al. 2013 with permission from Elsevier) (**b**) Schematic representation of 'omics levels, with essential generation of energy (ATP) (Lehninger 1975; Gancedo and Serrano 1989), flow of genetic information (*solid arrows*), and interactome networks ("interactomes") and modular complexes (e.g., protein-protein; RNP complexes) (Castello et al. 2012; Carvunis et al. 2013; Chang et al. 2013; Gstaiger and Aebersold 2013; Mitchell et al. 2013; Ryan et al. 2013; Song and Singh 2013), with the dynamic interplay between levels (*dashed arrows*) responsible for a specific phenotype in close interaction with the environment (Adapted from Kohlstedt et al. 2010 with permission from Springer and authors; and Castrillo et al. 2013 with permission from Elsevier)

B. Yeast Systems Biology for Comprehensive Studies on Essential Mechanisms and Networks Underlying Complex Diseases: Case Examples and Applications—New Models

1. Multifactorial Complex Diseases

Complex diseases are, with the help of new advanced postgenomic technologies, being progressively better characterized and annotated, and this is providing more insight into their primary molecular origin (Ahmed et al. 2006; Vidal et al. 2011). Phenotype and disease ontologies and databases (e.g., OMIM [Online Mendelian Inheritance in Man], PhenOMIM) are beginning to set the standards (Washington et al. 2009; Mungall et al. 2010; Shah et al. 2012; Amberger et al. 2011; van Triest et al. 2011).

Thanks to these advances, it is now widely accepted that many **complex diseases** of

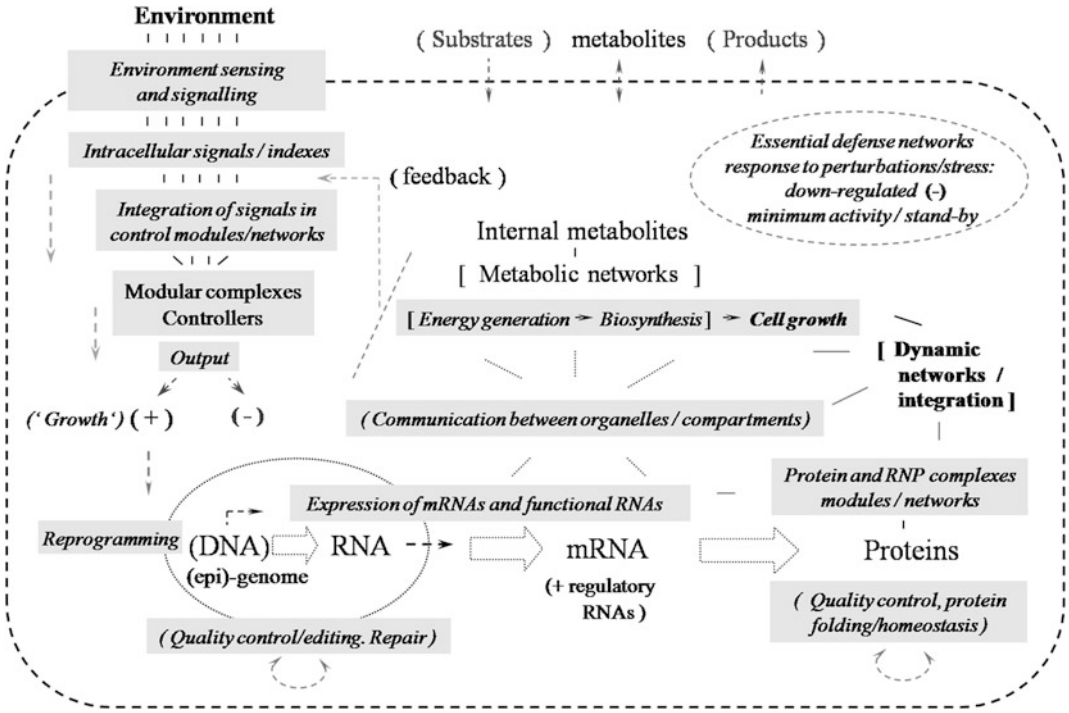


Fig. 1.2. Eukaryotic cell: main architecture, subcellular organization, and networks during balanced growth. Schematic architecture, subcellular organization, and networks in the eukaryotic cell during optimum, balanced growth (e.g., nutrient excess), in direct interaction with the environment (subcellular organelles other than the nucleus omitted for clarity). Environmental sensing of changes in external conditions (e.g., concentrations of external compounds, signals, pH, temperature), combined with intracellular indexes, once integrated in main control modules, "controllers" (e.g., TOR complexes; Loewith and Hall 2004) configure the actual cellular state ("optimum" for growth in this case), with activation (+) of signal transduction pathways and (re)programming of gene expression and metabolic and regulatory networks (e.g., at the (epi) genome, transcriptome, proteome, metabolome levels and their interactions) toward high energy-yielding metabolism and biosynthetic fluxes sustaining balanced cell growth (Hall et al. 2004; Castrillo et al. 2007). Note that under these conditions essential defense, stress networks responsible for rapid recovery from perturbations, maintenance, or survival (e.g., DNA repair, RNA and protein quality control homeostasis networks) are downregulated (—) at their basal, minimum levels of activity or in stand-by (Lehninger 1975; Alberts et al. 2008)

humans (including diabetes, Alzheimer's disease, Parkinson's disease, and most cancers) are far too complex to have a single cause or to rely on genomic variations alone. What really defines these complex diseases is that they (a) are **multifactorial** (i.e., with potential significant contribution of genomic/epigenomic and environmental perturbations); (b) **primarily arise from "altered networks"** affecting "essential modules" and pathways responsible for basic functions (Barrenas et al. 2009; Agarwal et al. 2010; Ulitsky et al. 2010; Kim et al. 2011; Vidal et al. 2011; Gstaiger and Aebersold 2013; Ryan et al. 2013; Song and Singh 2013); (c) are **fundamentally "dynamic,"** with a fine balance between perturbations and defense homeostatic networks underlying the actual phenotype (see below).

We propose that **multifactorial complex diseases should be contemplated essentially as shown in** Fig. 1.3. The genome and epigenome underlie the essential networks, first homeostatic states, and the basic "genomic/epigenomic" susceptibility to dysregulation of an organism, which will be subjected to a particular sequence of environmental perturbations

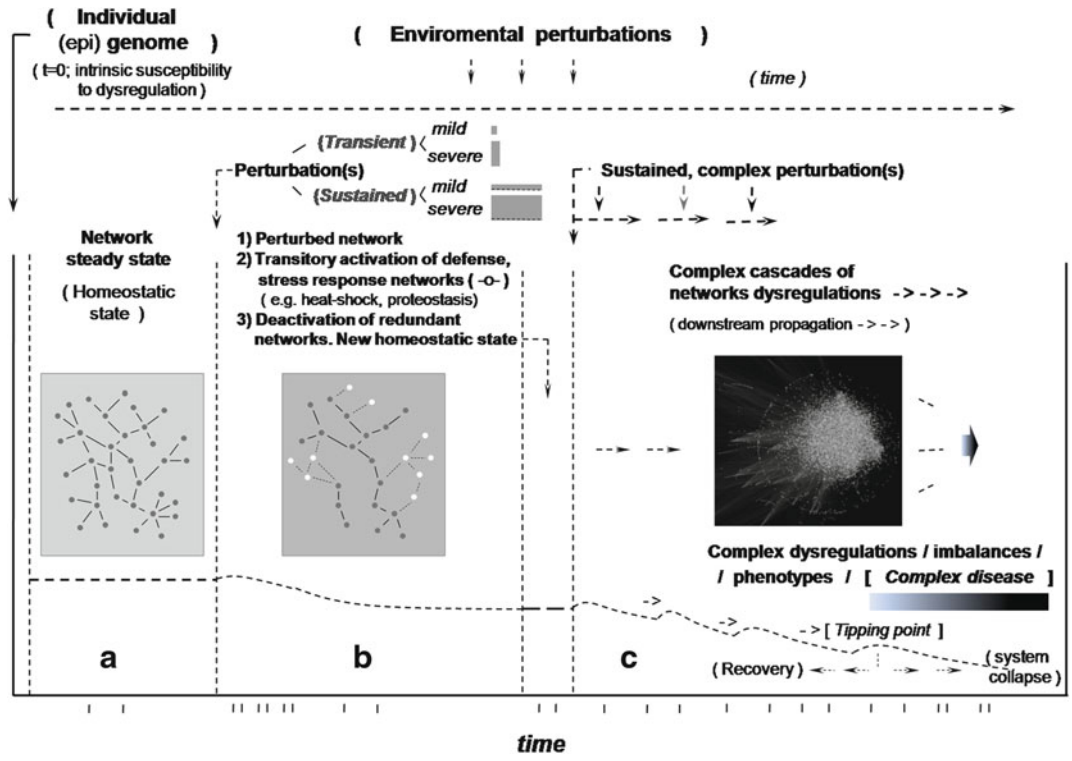


Fig. 1.3. Multifactorial complex diseases and healthy states—a fine balance between periodically perturbed, dysregulated networks and our defense, homeostatic mechanisms during lifetime. The original genome and epigenome underlie the essential networks, first homeostatic states (**a**) and the initial "genomic/epigenomic" susceptibility to dysregulation of an organism, which will be subjected to a specific sequence of environmental perturbations (mild or severe; transient or sustained in time) during its lifetime. Mild perturbations (**b**) result in transient deactivation of redundant networks and subnetworks (*gray nodes and edges*) and activation of intrinsic defense, stress responses and homeostatic networks (e.g., heat shock, protein homeostasis, inflammatory or immunological networks, and others; see new nodes and edges, *–o–*), until a new homeostatic, healthy state is restored. More importantly, severe, complex (e.g., multifactorial) or sustained perturbations (**c**), resulting in accumulated damage (e.g., somatic mutations; severe proteome alterations) can overcome the intrinsic homeostatic defense networks (Holliday 1995, 2006; López-Otín et al. 2013), whose efficiency may also be perturbed or decrease through lifetime, leading to cascades of dysregulations. These can propagate through intertwined networks, resulting in acute imbalances, pleiotropic effects, and complex diseases. Eventually, the system's capacity to maintain homeostasis may be overwhelmed, resulting in irreversible, catastrophic collapse (Holliday 1995, 2006; Dai et al. 2012, 2013; Carpenter 2013). Periodic, longitudinal monitoring at different 'omic levels (e.g., transcriptome, proteome, metabolome) characterizing homeostatic states and early perturbed networks in different genetic backgrounds in molecular and systems biology comprehensive experiments in model organisms (i.e., from yeast to human) (Castrillo et al. 2007, 2013; Dikicioglu et al. 2011; Chen et al. 2012; Li-Pook-Than and Snyder 2013; Pollitzer 2013) have the potential to unveil the origin, early stages, and dynamics of progression of complex imbalances and diseases well before the point of no return, toward early diagnosis (e.g., with combined biomarkers at different 'omic levels) and rational, affordable, counteracting strategies (lifestyle changes or therapeutic) (Adapted from Castrillo et al. 2013 with permission from Elsevier. Human interactome network picture visualized by Cytoscape 2.5. Human microbiome networks (Cho and Blaser 2012) with direct interplay with the human interactome are omitted for clarity. Dataset created by Andrew Garrow at Unilever UK. Author: Keiono, reproduced under GNU Free Documentation License and Creative Commons [CC] licenses, http://en.wikipedia.org/wiki/File:Human_interactome.jpg)

(mild or severe; transient or sustained in time) during its lifetime. Mild perturbations may be counteracted by intrinsic (stress) defense networks (e.g., heat shock, protein homeostasis, inflammatory or immunological networks, and others) restoring the homeostatic, "healthy" state of the individual. However, severe (e.g., multifactorial, genomic, and environmental) or sustained perturbations can overcome these homeostatic defense networks, whose efficiency may decrease through life (Holliday 1995, 2006; López-Otín et al. 2013). This may lead to cascades of dysregulations through intertwined essential networks resulting in acute imbalances, pleiotropic effects, and complex diseases. Eventually, the system's capacity to restore homeostasis may be overwhelmed, resulting in an irreversible catastrophic collapse (Holliday 1995, 2006; Dai et al. 2012, 2013; Carpenter 2013).

Until we understand complex diseases as altered states of human biological networks, in constant relation with the environment (e.g., external insults, perturbations, traumas, infections, our own microbiome; Cho and Blaser 2012), with specific dynamics and interplay, our vision will be incomplete (Friend 2010; Castrillo et al. 2013). Human individuality at the genomic and metabolomic levels (e.g., 1000 Genomes Project Consortium 2010; Chen et al. 2013; Suhre et al. 2011) adds an additional layer of complexity and presents us with an even more challenging picture. Where to start? How can we begin to address most complex human diseases?

Fortunately, for a majority of cases in which the disease manifests itself first primarily at the cellular level (early stage, likely asymptomatic), a good starting point will be to recapitulate the altered mechanisms, system properties, and networks (Figs. 1.1 and 1.2) in simple disease models at the eukaryote cellular level (e.g., yeast). The affected networks, their short- and long-term dynamics, and interplay may then be monitored in well-designed experiments, such as time-course systems biology studies under controlled conditions (Fiechter et al. 1987; Castrillo et al. 2007, 2013; Dikicioglu et al. 2011) (see below).

2. Objectives, Experimental Systems and Designs, and Integrative Data Analysis Strategies

First, by definition, comprehensive studies of complex diseases need to address two **main objectives:** (a) **the construction of reliable models that truly recapitulate altered mechanisms and features of the disease** at the molecular and cellular levels and (b) **analysis of the progression of the altered networks and phenotypes** in experimental systems under controlled conditions, reproducing environmental in vivo perturbations, whose effects are representative of those contributing to the disease. These experimental systems should be easy to implement to analyze the affected networks together with the activation of defense responses, with their dynamics and interplay and enable progress toward direct applications (e.g., early diagnosis and timely intervention).

For the first goal, the large number of molecular tools and huge potential for high-throughput studies available in *S. cerevisiae* (see above) positions this yeast as a first-line approach to generate excellent models for the study of complex diseases. For the second challenge (i.e., the use of proper experimental systems for comprehensive dynamic studies), continuous (e.g., chemostat) cultures in steady state, in which the specific growth rate can be selected and fixed operationally and the cells are growing (and can be long maintained) in a steady state at a constant growth rate in a constant environment (Kubitschek 1970; Fiechter 1975; Fiechter et al. 1987), can be used as representatives of initial homeostatic states (Fig. 1.3a) and may be combined with well-designed perturbations (mild or severe, transient or sustained) followed by time-course experiments and profiling (e.g., monitoring responses at different molecular and 'omics levels; Castrillo et al. 2007) to reproduce the characteristics of multifactorial diseases outlined in the previous section (see also Fig. 1.3). To do this effectively demands the following:

1. The use of **proper experimental design**, minimizing confounding variables and bias and putting in place a **bioinformatic and**

statistical strategy from the outset (Cohen and Medley 2005; Sarewitz 2012; Vaux 2012; Barbash and Soreq 2013). Here, early discussions with bioinformatic experts analyzing goals and expected results, number of experiments, conditions and replicates, total costs, and alternative data analysis strategies to obtain statistically significant results and solid conclusions will be of clear benefit (Veličković 2013).

2. The need for **advanced, comprehensive data analysis and integrative strategies** to be applied, for example, to the integration of data from molecular analyses and a range of 'omics, interactomes, and network datasets. As relevant examples, it is worth mentioning the following: new advanced tools for integrative metabolome and interactome mapping (iMIM) (Dumas and Davidovic 2013); regulatory network reconstructions and modeling (Glass et al. 2013; Manioudaki and Poirazi 2013); integrative strategies to analyze modularity and essential interactomes conserved through evolution (Hodgkinson and Karp 2012, 2013; Thompson et al. 2013; Roy et al. 2013); integrative approaches for measuring and analyzing interactome dynamics (Kristensen et al. 2012; Altelaar et al. 2013); and strategies for integrative personalized 'omics profiling and biomarker discovery (Villoslada and Baranzini 2012; Li-Pook-Than and Snyder 2013). For more exhaustive examples of optimum experimental design, careful implementation of yeast systems biology experiments, the latest techniques, comprehensive data analysis, and integration methods, refer to the work of Castrillo and Oliver (2011a, b) and Castrillo et al. (2013) and references therein.

3. Yeast Systems Biology for Studies on Essential Networks Underlying Complex Diseases: Case Examples and Applications—Enter New Disease Models, from Yeast to Human

a) Yeast Systems Biology for Studies of Complex Diseases

Where will yeast disease models be most useful? Based on the remarkable conservation of the majority of eukaryotic machineries, modules, and networks, even at higher organizational and subcellular levels, the answer is in the study of the core of fundamentally conserved networks (e.g., their dysregulation and dynamics), particularly the essential homeostatic, quality control, defense networks (Holliday 1995; 2006; Breitenbach et al. 2012; López-Otín et al. 2013) and their contribution to complex diseases (see below). As an example of this striking conservation, Table 1.1 shows several congenital, metabolic, and neurological disorders with relevant "human disease genes" (i.e., reportedly contributing to the disease) with yeast orthologs. The fact that such a large number of human disease genes share an ortholog with yeast provides a huge potential from the start. For the majority of the diseases in Table 1.1, yeast models based on their endogenous orthologs can be constructed (e.g., reproducing the altered feature of the human gene or its regulation) for basic characterization and phenotypic studies, with many of them already available. An example is the yeast models of Friedreich ataxia and Batten disease, two autosomal recessive neurodegenerative disorders with reported altered FXN and CLN3 genes encoding frataxin and battenin proteins, respectively (Wilson and Roof 1997; Pearce and Sherman 1998, 1999; other examples in Pereira et al. 2012; Tenreiro et al. 2013; Table 1.1 and references therein). For these yeast models, comprehensive systems biology time-course experiments such as those explained previously (e.g., starting with steady states, initially in the absence of relevant environmental perturbations, to be included later) offer huge potential to monitor and characterize (e.g., at different 'omic levels; Castrillo et al. 2007) the onset and dynamics of diseases at the molecular and cellular levels (i.e., early dysregulation, activation of defense, homeostatic networks, interplay and progression of imbalances, etc.). Moreover, it opens the way to the discovery of true characteristic patterns at the "earliest stages of the disease," at a crucial interval or "window of opportunity," well before the tipping point (Fig. 1.3), for earliest diagnosis and timely intervention (e.g., studies on drug treatments counteracting dysregulation or working synergistically, enhancing the response of our

defense networks) for all diseases in Table 1.1 and for as-yet-uncharacterized diseases with an altered genomic component (sequence, structural variation, other) with a counterpart that can be studied in yeast, to be validated in animal models and humans.

Yeast can also be a good model for the study of diseases in which "human-specific" genomic components and core-conserved networks are altered. Table 1.2 shows relevant examples of disease models based on "humanized" yeasts or yeast orthologs. Provided they truly recapitulate early features of the disease, these models, combined with dynamic time-course experiments, would offer again a unique opportunity to monitor and characterize the onset and progression of these diseases, first in the absence of additional perturbations.

Finally, yeast makes an excellent model to study the specific contributions of environmental perturbations to disease phenotypes (Fig. 1.3). Thus, mild perturbations will allow the study of transient defense responses in a defined period of time, and sustained perturbations (e.g., accumulated, sustained, oxidative stress; Tan et al. 2013) will provide valuable insights into long-term effects (e.g., decrease of efficacy of the homeostatic networks; Holliday 1995, 2006). These will be most relevant to illuminate the onset of complex diseases without a clear genomic culprit, which often appear in old age (late-onset diseases).

b) Case Examples and Applications

Among the most attractive examples of complex diseases and yeast disease models beginning to be exploited to their full potential in dynamic systems biology experiments are the following:

1. Yeast disease models of several congenital disorders and complex neurodegenerative diseases (e.g., Wilson and Roof 1997; Pearce and Sherman 1998, 1999; Pereira et al. 2012; see Table 1.1 and references therein).
2. Yeast disease models of complex diseases with altered essential networks, particularly those responsible for rapid homeostatic stress responses counteracting environmental perturbations (e.g., DNA and RNA repair; proteostasis; endoplasmic reticulum [ER] stress; the unfolded protein response [UPR]; ER-associated degradation [ERAD]; aggregation/disaggregation mechanisms; autophagy; and many others). For examples of relevant studies, complex diseases, and yeast models for which these networks exert significant contribution, see the works of Holliday (1995, 2006), Breitenbach et al. (2012), Pereira et al. (2012), Ocampo et al. (2013), Porzoor and Macreadie (2013), Tenreiro et al. (2013) and Tables 1.1 and 1.2 and references therein.
3. Complex diseases with no yeast counterpart (e.g., no ortholog), but for which "humanized" yeast models can be constructed. Such yeast systems biology studies will be of great interest to advance our insight into the contributing mechanisms and their dynamics. Relevant case examples include yeasts expressing human disease alleles that cause homocystinuria (Mayfield et al. 2012); a yeast model for S-adenosyl-L-homocysteine hydrolase and methylation disorders (Tehlivets et al. 2013); yeast strains with the lipid defect responsible for Niemann-Pick type C (NP-C) disease (Munkacsi et al. 2011); yeast models of human protein aggregation in amyotrophic lateral sclerosis (ALS) (Johnson et al. 2008; Ju et al. 2011); yeast disease models of Aβ toxicity and other relevant neurodegenerative mechanisms (Treusch et al. 2011; Pereira et al. 2012; Duennwald 2013; Ocampo et al. 2013; Porzoor and Macreadie 2013; Tenreiro et al. 2013 and references therein); *S. cerevisiae* yeast models expressing the oncogene-like *RAS2* (Lee et al. 2012); yeast models expressing human mutant isocitrate dehydrogenase (IDH), producing the oncometabolite 2-hydroxyglutarate (2HG), leading to impairment in histone demethylation and heterochromatin modifications in human gliomas and acute myeloid leukemia, for the study of the interdependence of genetic and epigenetic alterations (Chowdhury

Table 1.2. Yeast as a model for human disease: advanced studies and applications

Yeast advanced studies and applications	References
1. Yeast models for comprehensive systems biology studies of dynamics of dysregulated networks in disease, early detection of imbalances and networks dynamics	Gruson and Bodovitz (2010), Chen et al. (2012), Villoslada and Baranzini (2012), Wang et al. (2012)
Yeast cell cycle model/design principles	Ferrell (2011)
Yeast as a model for studies of eukaryotic cell growth (balanced cell growth)	Jorgensen et al. (2004), Castrillo et al. (2007), Gutteridge et al. (2010), Przytycka and Andrews (2010), Pir et al. (2012)
Yeast breast cancer model (yeast with human mutations and increased cell proliferation) (uncontrolled growth)	Li et al. (2009)
Yeast expressing parasite genes and their human orthologs for protein quality control networks studies and identification of antiparasitic agents	Bell et al. (2011), Bilsland et al. (2011, 2013)
Yeast as a model for aging research studies	Breitenbach et al. (2012)
Yeast models to unravel mechanisms of neurodegeneration	Khurana and Lindquist (2010), Tenreiro and Outeiro (2010), Pereira et al. (2012), Pimentel et al. (2012), Ocampo et al. (2013), Tenreiro et al. (2013), Villar-Piqué and Ventura (2013)
Yeast models of protein folding diseases and human neurodegenerative disorders	Wilson and Roof (1997), Pearce and Sherman (1998, 1999), Johnson et al. (2008), Lindquist (2008), Bharadwaj et al. (2010), De Vos et al. (2011), Kryndushkin and Shewmaker (2011), Mason and Giorgini (2011), McGurk and Bonini (2011), Ocampo and Barrientos (2011), Treusch et al. (2011), Pereira et al. (2012) D'Angelo et al. (2013), Tenreiro et al. (2013)
Yeast models of amyotrophic lateral sclerosis (ALS)	Johnson et al. (2008), Ju et al. (2011), Kryndushkin and Shewmaker (2011), Robinson (2011), Sun et al. (2011)
Yeast models of tau, amyloid-β aggregation, and PICALM role in cytotoxicity and Alzheimer's disease	Bharadwaj et al. (2010), De Vos et al. (2011), Treusch et al. (2011), Pereira et al. (2012), D'Angelo et al. (2013), Porzoor and Macreadie (2013), Tenreiro et al. (2013)
Yeast model for S-adenosyl-L-homocysteine hydrolase and methylation disorders	Tehlivets et al. (2013)
2. Yeast molecular, high-throughput systems biology studies and applications in human disease	Castrillo and Oliver (2011a, b), Zhang and Bilsland (2011), Feldmann (2012b), Castrillo et al. (2013)
Yeast genetics and metabolic profiling for characterization of human disease alleles	Mayfield et al. (2012)
High-throughput yeast system (IDESA) for study of protein stability diseases and screening of small molecules	Pittman et al. (2012)
Yeast-based assay for identification of drugs active against human mitochondrial disorders	Couplan et al. (2011)
Yeast as a model system to study the response to anticancer agents	Matuo et al. (2012)
Humanized yeast-based, high-throughput screen to identify small activators for therapeutic intervention in neurodegenerative diseases	Neef et al. (2010)
Yeast functional screen to identify new candidate ALS genes	Couthouis et al. (2011)
Yeast phenotypic screening for compounds that rescue TDP-43, α-synuclein, and polyglutamine proteotoxicity	Tardiff et al. (2012)

(continued)

Table 1.2. (continued)

Yeast advanced studies and applications	References
Phenotypic screening for compounds that target the cellular pathologies underlying Parkinson's disease	Tardiff and Lindquist (2013)
Yeast high-throughput screenings for inhibitors of $A\beta_{42}$ oligomerization and for the study of Aβ peptide and oligomerization-dependent toxicities	Park et al. (2011), López et al. (2012), Porzoor and Macreadie (2013)
Yeast genomics and drug target identification	Bharucha and Kumar (2007), Bilsland et al. (2013)
Chemical genomic (chemogenomic) profiling; screenings of targets genes and biomolecules; yeast as a tool for target-directed, high-throughput screening	Proctor et al. (2011), Andrusiak et al. (2012), Augagneur et al. (2013), Bilsland et al. (2013), Norcliffe et al. (2013), Roemer and Boone (2013), Schenone et al. (2013)

et al. 2011; Xu et al. 2011; McCarthy 2012). For other remarkable yeast disease models, studies, and advanced applications at the forefront of technology (i.e., direct translation into early diagnosis and timely therapeutic strategies, e.g., the most advanced comprehensive chemogenomic screenings starting in yeast, for subsequent testing in animal models and finally in human), see Tables 1.1 and 1.2 and references therein.

c) Enter New Disease Models, from Yeast to Human

Yeast models of human disease, like other models, need to recapitulate the "altered" features and early phenotypes observed in human cells. This is the most important "take-home message," the main requirement of a **good disease model: "to be able to recapitulate molecular and physiological features of the human disease"** (e.g., Johnson et al. 2008; a yeast model of ALS producing the human protein that undergoes pathological conversion from nuclear localization to cytoplasmic aggregation).

In the end, a human disease will be characterized by altered networks in either essentially conserved networks and modules common to all eukaryotes, for which yeast can be an excellent model (e.g., to study them independently isolated, "dissected," or combined with others, with their dynamics) and "human-specific" networks (Das et al. 2013) (with or without altered essential networks as well), for which humanized disease models at the cellular or supracellular levels will be needed. At the cellular level, human-specific altered mechanisms (e.g., in humanized yeast models [see above], combined, e.g., with studies in mammalian cells and patient-specific disease models; Ebert et al. 2012; Dianat et al. 2013; Musunuru 2013; Pollitzer 2013) can be used.

Supracellular characteristics of complex diseases (i.e., at the tissue, organ, and whole-body level) are usually beyond the reach of unicellular model organisms, and **closer-to-human model organisms** are necessary. However, these more complex species need to be chosen with care. As an example, the impossibility of recapitulating all Alzheimer's disease features in mouse models has only recently been explained as likely to be a consequence of the different aggregation properties of the mouse aβ peptide and intracellular mouse proteome compared to that of the human. The construction of a closer-to-human rat model has enabled the recapitulation of all Alzheimer's disease human features, including plaques, tau pathology, behavioral impairment, oligomeric aβ, and neuronal loss (Cohen et al. 2013; Tharp and Sarkar 2013).

Studies in several disease models are exemplifying the most effective strategy for identifying events and pathways with a significant contribution to pathogenesis. Thus, combining data across a range of model species has revealed the role of toxic RNA as a driver of disease in a common form of ALS and dementia (Xu et al. 2013). In the future, progress will come from **confirmation of pathogenetic mechanisms in different disease models and**

integration of data from yeast to humans, including mammalian cell lines, patient-specific cell lines from iPSCs, higher eukaryotes, animal models, and ultimately, human studies and trials (Chang and Grieder 2008; Soucek et al. 2008; Aitman et al. 2011; Ebert et al. 2012; Shi et al. 2012; Dianat et al. 2013; Marx 2013; Mason et al. 2013; Murphy 2013; Musunuru 2013; National Institutes of Health). Finally, **early diagnosis of complex diseases** will come with advances in longitudinal monitoring with guidelines for standardization of clinical trials, application of the most powerful postgenomics analytical techniques, and data integration of multiple biomarkers and patterns (e.g., combined 'omics datasets) at a systems level. Such integrative studies are expensive but are becoming progressively more affordable (Chen et al. 2012; Li-Pook-Than and Snyder 2013; U.S. Food and Drug Administration [FDA] clinical trials guidances, with adherence to the principles of good clinical practices [GCPs], see U.S. FDA 2013, and http://clinicaltrials.gov/; European Medicine Agency [EMA] and European Clinical Trials Database [EudraCT], see https://eudract.ema.europa.eu/; Castrillo et al. 2013; Chiu et al. 2013; Leichtle et al. 2013).

III. Conclusions

The rich complexity of life (Hayden 2010; Heard et al. 2010) is sustained by a continual input of energy (Lehninger 1975; Lever et al. 2013; Aarhus University ScienceDaily News, 2013) and relies on an exquisite interplay between functional networks that has been optimized through millions of years of evolution (Darwin and Wallace 1858; Darwin 1859; Alfred Russel Wallace online and updated, Smith 2000; van Wyhe 2012; Charles Darwin online and updated, van Wyhe 2002). Systems biology takes up the challenges presented by this complexity, and excellent reference model organisms are an integral part of this systems approach. Such model organisms are allowing the following questions to be addressed: (a) What principles make a complex networks system sustainable and robust? (b) How do multifactorial complex diseases arise from altered networks states (Friend 2010)? (c) Can their dynamics be better studied and characterized to enable applications with translational impact (e.g., early diagnosis and affordable therapies that treat complex diseases in an holistic manner)?

Yeast can be a first reference model for systems biology studies on altered essential networks and reveal the basic pathogenetic mechanisms involved in multifactorial complex diseases at the cellular level. The rationale of systems biology studies in well-constructed yeast models of disease that have been presented in this chapter can be reproduced in other model organisms, and effective progress will come from the integration of knowledge from studies in different disease models, from yeast to humans. As an essential prerequisite, the disease models should recapitulate the main human disease features (e.g., Johnson et al. 2008; Cohen et al. 2013).

In order to realize the full potential of yeast as a reference model for complex diseases, the high level of international collaboration between research groups seen in the yeast genome sequencing (Goffeau et al. 1996) and functional genomics projects (Oliver 1997, 2002; Winzeler et al. 1999; Giaever et al. 2002) must be continued and even enhanced, with incorporation of experts from different scientific backgrounds, in comprehensive interdisciplinary efforts. To this end, the yeast (SGD) (http://www.yeastgenome.org; http://wiki.yeastgenome.org) and systems biology (http://systems-biology.org/) communities promote interdisciplinary collaborations, projects, and initiatives between experts in the different fields. A relevant example is the UNICELLSYS project, a systems biology initiative with the overall objective of a quantitative understanding of control of cell growth and proliferation (http://www.unicellsys.eu/).

These joint initiatives and projects can be reproduced in systems biology studies at higher physiological levels (Sperling 2011; Bashan et al. 2012; MacLellan et al. 2012) with more complex model organisms (Stuart et al. 2007; Cox et al. 2009; Mori et al. 2010; Antony et al. 2011; Yu et al. 2012) and, ultimately, humans

(Goto and Tanzi 2002; Joyner and Pedersen 2011). As an example, the Virtual Physiological Human (VPH) initiative represents a network of excellence (Viceconti et al. 2008; Sansom et al. 2011a; VPH 2010), and VPH projects are directed toward modeling a wide range of human organs and systems in both healthy and diseased states (Sansom et al. 2011b and references therein; VPH 2010). In order to advance their objectives, these community efforts will always need to incorporate new knowledge on eukaryotic networks from systems biology studies at the cellular level. On this journey, the humble yeast *S. cerevisiae* can accompany us along the path, with continuous new exciting discoveries ahead. In the words of Honor Fell, who pioneered the study of cells under the microscope: "The more closely we examine a natural object, the more beautiful, exciting and mysterious it becomes. . . . A single living cell is much more beautiful and improbable than the solar system. So far scientists have not explained away the wonder of the cell nor stripped it of its mystery" (Fell 1900–1986; Wellcome Library, The Honor Fell Papers, http://wellcomelibrary.org/using-the-library/subject-guides/genetics/makers-of-modern-genetics/digitised-archives/honor-fell/). As a research community, our responsibility lies in delivering excellent results and outputs to advance both basic knowledge and translational impact for the benefit of society.

Acknowledgments This work was supported by Biotechnology and Biological Sciences Research Council (BBSRC) grants BB/C505140/2 and BB/F00446X/1 as well as by a contract from the European Commission under the FP7 Collaborative Programme, UNICELLSYS. JIC is a beneficiary of a senior postdoctoral aid program 2001 (mode A) of Bizkaia:Xede Foundation in collaboration with GENETADI Biotech Laboratories (Derio, Bizkaia, Spain).

References

1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073, http://www.1000genomes.org/

Aarhus University ScienceDaily News (2013) Life deep within oceanic crust sustained by energy from interior of earth. Science Daily, 14 Mar. http://www.sciencedaily.com/releases/2013/03/130314144340.htm

Agarwal S, Deane CM, Porter MA et al (2010) Revisiting date and party hubs: novel approaches to role assignment in protein interaction networks. PLoS Comput Biol 6:e1000817. doi:10.1371/journal.pcbi.1000817

Ahmed N, Dawson M, Smith C, Wood E (2006) Biology of disease. Garland Science/Taylor and Francis, New York

Aitman TJ, Boone C, Churchill GA et al (2011) The future of model organisms in human disease research. Nat Rev Genet 12:575–582. doi:10.1038/nrg3047

Alberts B (2011) Lessons from genomics. Science 331:511

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2008) Molecular biology of the cell, 5th edn. Garland Science/Taylor and Francis, New York

Almada AE, Wu X, Kriz AJ et al (2013) Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature 499:360–363

Altelaar AF, Munoz J, Heck AJ (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet 14:35–48

Amberg DC, Burke D, Strathern JN (2005) Methods in yeast genetics. A Cold Spring Harbor Laboratory course manual. Cold Spring Harbor Laboratory Press, New York

Amberger J, Bocchini C, Hamosh A (2011) A new face and new challenges for Online Mendelian Inheritance in Man (OMIM). Hum Mutat 32:564–567

Andrusiak K, Piotrowski JS, Boone C (2012) Chemical-genomic profiling: systematic analysis of the cellular targets of bioactive molecules. Bioorg Med Chem 20:1952–1960

Antony PM, Diederich NJ, Balling R (2011) Parkinson's disease mouse models in translational research. Mamm Genome 22:401–419

Antunes LC, Davies JE, Finlay BB (2011) Chemical signaling in the gastrointestinal tract. F1000 Biol Rep 3:4

Arkin AP, Schaffer DV (2011) Network news: innovations in 21st century systems biology. Cell 144:844–849

Auffray C, Caulfield T, Khoury MJ et al (2012) Looking back at genomic medicine in 2011. Genome Med 4:9

Augagneur Y, Jaubert L, Schiovani M et al (2013) Identification and functional analysis of the primary pantothenate transporter, PfPAT, of the human malaria parasite *Plasmodium falciparum*. J Biol Chem 288:20558–20567

Babu M, Vlasblom J, Pu S et al (2012) Interaction landscape of membrane-protein complexes in *Saccharomyces cerevisiae*. Nature 489:585–589

Baile MG, Claypool SM (2013) The power of yeast to model diseases of the powerhouse of the cell. Front Biosci 18:241–278

Balaji S, Babu MM, Iyer LM (2006) Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. J Mol Biol 360:213–227

Barabsi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56–68

Barbash S, Soreq H (2013) Statistically invalid classification of high throughput gene expression data. Sci Rep 3:1102

Barrenas F, Chavali S, Holme P et al (2009) Network properties of complex human disease genes identified through genome-wide association studies. PLoS One 4:e8090

Bashan A, Bartsch RP, Kantelhardt JW et al (2012) Network physiology reveals relations between network topology and physiological function. Nat Commun 3:702. doi:10.1038/ncomms1705

Bell SL, Chiang AN, Brodsky JL (2011) Expression of a malarial Hsp70 improves defects in chaperone-dependent activities in ssa1 mutant yeast. PLoS One 6:e20047

Bharadwaj P, Martins R, Macreadie I (2010) Yeast as a model for studying Alzheimer's disease. FEMS Yeast Res 10:961–969

Bharucha N, Kumar A (2007) Yeast genomics and drug target identification. Comb Chem High Throughput Screen 10:618–634

Bilsland E, Pir P, Gutteridge A et al (2011) Functional expression of parasite drug targets and their human orthologs in yeast. PLoS Negl Trop Dis 5: e1320

Bilsland E, Sparkes A, Williams K et al (2013) Yeast-based automated high-throughput screens to identify anti-parasitic lead compounds. Open Biol 3:120158

Bordbar A, Palsson BO (2012) Using the reconstructed genome-scale human metabolic network to study physiology and pathology. J Intern Med 271:131–141

Breitenbach M, Jazwinski SM, Laun P (2012) Aging research in yeast, vol 57, Subcellular biochemistry. Springer, New York

Calvo SE, Compton AG, Hershman SG et al (2012) Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. Sci Transl Med 4

Carpenter SR (2013) Complex systems: spatial signatures of resilience. Nature 496:308–309

Carvunis A-R, Roth FP, Calderwood MA et al (2013) Interactome networks. In: Walhout M, Vidal M, Dekker J (eds) Handbook of systems biology. Elsevier, New York, pp 45–63

Casals F, Idaghdour Y, Hussin J et al (2012) Next-generation sequencing approaches for genetic mapping of complex diseases. J Neuroimmunol 248:10–22

Castello A, Fischer B, Eichelbaum K et al (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149:1393–1406

Castrillo JI, Oliver SG (2004) Yeast as a touchstone in post-genomic research: strategies for integrative analysis in functional genomics. J Biochem Mol Biol 37:93–106

Castrillo JI, Oliver SG (2006) Metabolomics and systems biology in *Saccharomyces cerevisiae*. In: Esser K (ed) The Mycota. A comprehensive treatise on fungi as experimental systems for basic and applied research. Fungal genomics, vol 13. Springer, Berlin, pp 3–18

Castrillo JI, Oliver SG (2011a) Yeast systems biology: the challenge of eukaryotic complexity. Methods Mol Biol 759:3–28

Castrillo JI, Oliver SG (2011b) Yeast systems biology. Methods and protocols. Methods in molecular biology 759 (MiMB series. Editor-in-chief. JM Walker). Humana Press/Springer, New York

Castrillo JI, Zeef LA, Hoyle DC et al (2007) Growth control of the eukaryote cell: a systems biology study in yeast. J Biol 6:4

Castrillo JI, Pir P, Oliver SG (2013) Yeast systems biology: towards a systems understanding of regulation of eukaryotic networks in complex diseases and biotechnology. In: Walhout M, Vidal M, Dekker J (eds) Handbook of systems biology. Elsevier, New York, pp 343–365

Chang MC, Grieder FB (2008) The baffling multitude of disease models for the study of human disease—how can the scientist navigate the huge amount of data and receive guidance? Dis Model Mech 1:99–102

Chang X, Xu T, Li Y et al (2013) Dynamic modular architecture of protein-protein interaction networks beyond the dichotomy of "date" and "party" hubs. Sci Rep 3:1691

Chen R, Mias GI, Li-Pook-Than J et al (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 148:1293–1307

Chen G, Wang C, Shi L et al (2013) Comprehensively identifying and characterizing the missing gene sequences in human reference genome with integrated analytic approaches. Hum Genet 132:899–911

Chiu IM, Morimoto ET, Goodarzi H et al (2013) A neurodegeneration-specific gene-expression signature of acutely isolated microglia from an amyotrophic lateral sclerosis mouse model. Cell Rep 4:385–401

Cho I, Blaser MJ (2012) The human microbiome: at the interface of health and disease. Nat Rev Genet 13:260–270

Choi M, Scholl UI, Ji W et al (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106:19096–19101

Chowdhury R, Yeoh KK, Tian YM et al (2011) The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. EMBO Rep 12:463–469

Cohen J, Medley G (2005) Stop working and start thinking, 2nd edn. Garland Science/Taylor and Francis, New York

Cohen RM, Rezai-Zadeh K, Weitz TM et al (2013) A transgenic Alzheimer rat with plaques, tau pathology, behavioral impairment, oligomeric aβ, and frank neuronal loss. J Neurosci 33:6245–6256

Couplan E, Aiyar RS, Kucharczyk R et al (2011) A yeast-based assay identifies drugs active against human mitochondrial disorders. Proc Natl Acad Sci U S A 108:11989–11994

Couthouis J, Hart MP, Shorter J et al (2011) A yeast functional screen predicts new candidate ALS disease genes. Proc Natl Acad Sci U S A 108:20881–20890

Cox B, Kotlyar M, Evangelou AI et al (2009) Comparative systems biology of human and mouse as a tool to guide the modeling of human placental pathology. Mol Syst Biol 5:279

Crick FH (1958) On protein synthesis. Symp Soc Exp Biol 12:138–163

Crick F (1970) Central dogma of molecular biology. Nature 227:561–563

D'Angelo F, Vignaud H, Di Martino J et al (2013) A yeast model for amyloid-β aggregation exemplifies the role of membrane trafficking and PICALM in cytotoxicity. Dis Model Mech 6:206–216

Dai L, Vorselen D, Korolev KS et al (2012) Generic indicators for loss of resilience before a tipping point leading to population collapse. Science 336:1175–1177

Dai L, Korolev KS, Gore J (2013) Slower recovery in space before collapse of connected populations. Nature 496:355–358

Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. London. http://darwin-online.org.uk/EditorialIntroductions/Freeman_OntheOriginofSpecies.html

Darwin C, Wallace AR (1858) On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. J Proc Linn Soc Lond Zool 3:45–62. http://darwin-online.org.uk/converted/pdf/1858_species_F350.pdf

Das J, Vo TV, Wei X et al (2013) Cross-species protein interactome mapping reveals species-specific wiring of stress response pathways. Sci Signal 6:ra38

Davies K (2010) Hugh Rienhoff's voyage round his daughter's DNA. Bio-IT World, Sept–Oct 2010. http://www.bio-itworld.com/BioIT_Article.aspx?id=101664

De Vos A, Anandhakumar J, Van den Brande J et al (2011) Yeast as a model system to study tau biology. Int J Alzheimers Dis 2011:428970

DeFrancesco L (2012) Life technologies promises $1,000 genome. Nat Biotechnol 30:126

Demaegd D, Foulquier F, Colinet AS et al (2013) Newly characterized Golgi-localized family of proteins is involved in calcium and pH homeostasis in yeast and human cells. Proc Natl Acad Sci U S A 110:6859–6864

Dianat N, Steichen C, Vallier L et al (2013) Human pluripotent stem cells for modelling human liver diseases and cell therapy. Curr Gene Ther 13:120–132

Dicarlo JE, Norville JE, Mali P et al (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res 41:4336–4343

Dikicioglu D, Karabekmez E, Rash B et al (2011) How yeast re-programmes its transcriptional profile in response to different nutrient impulses. BMC Syst Biol 5:148

Djebali S, Davis CA, Merkel A et al (2012) Landscape of transcription in human cells. Nature 489:101–108

Dobzhansky T (1964) Biology, molecular and organismic. Am Zool 4:443–452

Drinnenberg IA, Weinberg DE, Xie KT et al (2009) RNAi in budding yeast. Science 326:544–550

Duennwald ML (2013) Yeast as a platform to explore polyglutamine toxicity and aggregation. Methods Mol Biol 1017:153–161

Dumas ME, Davidovic L (2013) Metabolic phenotyping and systems biology approaches to understanding neurological disorders. F1000Prime Rep 5:18

Ebert AD, Liang P, Wu JC (2012) Induced pluripotent stem cells as a disease modeling and drug screening platform. J Cardiovasc Pharmacol 60:408–416

Evans JP, Berg JS (2012) The value of your genome. The Scientist, 26 Dec 2012. http://www.the-scientist.com/?articles.view/articleNo/33365/title/The-Value-of-Your-Genome/

Evans JP, Berg JS, Olshan AF et al (2013) We screen newborns, don't we?: realizing the promise of public health genomics. Genet Med 15:332–334

Feldmann H (2012a) Yeast: molecular and cell biology, 2nd edn. Wiley, Weinheim

Feldmann H (2012b) Diseases genes in yeast. In: Feldmann H (ed) Yeast: molecular and cell biology, 2nd edn. Wiley-VCH, Weinheim, pp 335–346

Fell H (1900–1986) Lecture notes. Wellcome Library archive. Archives and manuscripts. Ref. PP/HBF. http://wellcomelibrary.org/using-the-library/subject-guides/genetics/makers-of-modern-genetics/digitised-archives/honor-fell/

Ferrell JE Jr (2011) Simple rules for complex processes: new lessons from the budding yeast cell cycle. Mol Cell 43:497–500

Fiechter A (1975) Continuous cultivation of yeasts. Methods Cell Biol 11:97–130

Fiechter A, Kappeli O, Meussdorfer F (1987) Batch and continuous culture. In: Rose A, Harrison JS (eds) The yeasts, vol 2. Academic, London, pp 99–129

Foury F (1997) Human genetic diseases: a cross-talk between man and yeast. Gene 195:1–10
Franssens V, Boelen E, Anandhakumar J et al (2010) Yeast unfolds the road map toward alpha-synuclein-induced cell death. Cell Death Differ 17:746–753
Friend SH (2010) The need for precompetitive integrative bionetwork disease model building. Clin Pharmacol Ther 87:536–539
Gancedo C, Serrano R (1989) Energy-yielding metabolism. In: Rose AH, Harrison JS (eds) The yeasts, vol 3. Academic, London, pp 205–259
Genome Reference Consortium The Genome Reference Consortium. http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/
Genome Reference Consortium Human genome overview. http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Giaever G, Chu AM, Ni L et al (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418:387–391
Gilbert LA, Larson MH, Morsut L et al (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154:442–451
Gilissen C, Hoischen A, Brunner HG et al (2012) Disease gene identification strategies for exome sequencing. Eur J Hum Genet 20:490–497
Giuraniuc CV, Macpherson M, Saka Y (2013) Gateway vectors for efficient artificial gene assembly in vitro and expression in yeast *Saccharomyces cerevisiae*. PLoS One 8:e64419
Glass K, Huttenhower C, Quackenbush J et al (2013) Passing messages between biological networks to refine predicted interactions. PLoS One 8:e64832
Goffeau A, Barrell BG, Bussey H et al (1996) Life with 6000 genes. Science 274(546):563–567
Goh A (2013) Hunting down culprit genes. BioIT World. http://www.bio-itworld.com/2013/6/7/hunting-down-culprit-genes.html
Goldstein DB, Allen A, Keebler J et al (2013) Sequencing studies in human genetics: design and interpretation. Nat Rev Genet 14:460–470
Gonzaga-Jauregui C, Lupski JR, Gibbs RA (2012) Human genome sequencing in health and disease. Annu Rev Med 63:35–61
Goodman AJ, Daugharthy ER, Kim J (2013) Pervasive antisense transcription is evolutionarily conserved in budding yeast. Mol Biol Evol 30:409–421
Goto JJ, Tanzi RE (2002) The role of the low-density lipoprotein receptor-related protein (LRP1) in Alzheimer's A beta generation: development of a cell-based model system. J Mol Neurosci 19:37–41
Gruson D, Bodovitz S (2010) Rapid emergence of multimarker strategies in laboratory medicine. Biomarkers 15:289–296
Gstaiger M, Aebersold R (2013) Genotype-phenotype relationships in light of a modular protein interaction landscape. Mol Biosyst 9:1064–1067
Gutteridge A, Pir P, Castrillo JI et al (2010) Nutrient control of eukaryote cell growth: a systems biology study in yeast. BMC Biol 8:68
Hall MN, Raff M, Thomas G (2004) Cell growth: control of cell size, vol 42, Monograph. Cold Spring Harbor Laboratory Press, New York
Hayden EC (2010) Human genome at ten: life is complicated. Nature 464:664–667
Heard E, Tishkoff S, Todd JA et al (2010) Ten years of genetics and genomics: what have we achieved and where are we heading? Nat Rev Genet 11:723–733
Heinicke S, Livstone MS, Lu C et al (2007) The Princeton Protein Orthology Database (P-POD): a comparative genomics analysis tool for biologists. PLoS One 2:e766
Hodgkinson L, Karp RM (2012) Algorithms to detect multiprotein modularity conserved during evolution. IEEE/ACM Trans Comput Biol Bioinform 9:1046–1058
Hodgkinson L, Karp RM (2013) Optimization criteria and biological process enrichment in homologous multiprotein modules. Proc Natl Acad Sci U S A 110:10872–10877
Holliday R (1995) Understanding ageing. Cambridge University Press, Cambridge, UK
Holliday R (2006) Aging is no longer an unsolved problem in biology. Ann N Y Acad Sci 1067:1–9
Human Genome Organisation Home page. http://www.hugo-international.org/
Human Genome Project (HGP) (2003) Project completion. http://www.genome.gov/11006929
Human Genome Project (HGP) (2013) The genomics landscape a decade after the Human Genome Project. http://www.genome.gov/HGP10/
Ibañez AJ, Fagerer SR, Schmidt AM et al (2013) Mass spectrometry-based metabolomics of single yeast cells. Proc Natl Acad Sci U S A 110:8790–8794
Ishida S, Picard F, Rudolf G et al (2013) Mutations of DEPDC5 cause autosomal dominant focal epilepsies. Nat Genet 45:552–555
Jacquier A (2009) The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet 10:833–844
Johnson BS, McCaffery JM, Lindquist S et al (2008) A yeast TDP-43 proteinopathy model: exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. Proc Natl Acad Sci U S A 105:6439–6444
Jorgensen P, Tyers M, Warner JR (2004) Forging the factory: ribosome synthesis and growth control in budding yeast. In: Hall MN, Raff M, Thomas G (eds) Cell growth. Control of cell size. Cold Spring Harbor Laboratory Press, New York, pp 329–370
Joyner MJ, Pedersen BK (2011) Ten questions about systems biology. J Physiol 589:1017–1030
Ju S, Tardiff DF, Han H et al (2011) A yeast model of FUS/TLS-dependent cytotoxicity. PLoS Biol 9: e1001052

Kaback DB (2013) The modest beginnings of one genome project. Genetics 194:291–299
Kaufmann A, Knop M (2011) Genomic promoter replacement cassettes to alter gene expression in the yeast *Saccharomyces cerevisiae*. Methods Mol Biol 765:275–294
Khurana V, Lindquist S (2010) Modelling neurodegeneration in *Saccharomyces cerevisiae*: why cook with baker's yeast? Nat Rev Neurosci 11:436–449
Kim YA, Wuchty S, Przytycka TM (2011) Identifying causal genes and dysregulated pathways in complex diseases. PLoS Comput Biol 7:e1001095
Kitano H (2002) Systems biology: a brief overview. Science 295:1662–1664
Klass DM, Scheibe M, Butter F et al (2013) Quantitative proteomic analysis reveals concurrent RNA-protein interactions and identifies new RNA-binding proteins in *Saccharomyces cerevisiae*. Genome Res 23:1028–1038
Klitgord N, Segrè D (2010) The importance of compartmentalization in metabolic flux models: yeast as an ecosystem of organelles. Genome Inform 22:41–55
Kohlstedt M, Becker J, Wittmann C (2010) Metabolic fluxes and beyond-systems biology understanding and engineering of microbial metabolism. Appl Microbiol Biotechnol 88:1065–1075
Kristensen AR, Gsponer J, Foster LJ (2012) A high-throughput approach for measuring temporal changes in the interactome. Nat Methods 9:907–909
Kruger RP (2011) Systems biology. Cell 144(827):829
Kryndushkin D, Shewmaker F (2011) Modeling ALS and FTLD proteinopathies in yeast: an efficient approach for studying protein aggregation and toxicity. Prion 5:250–257
Ku CS, Cooper DN, Polychronakos C et al (2012) Exome sequencing: dual role as a discovery and diagnostic tool. Ann Neurol 71:5–14
Kubitschek HE (1970) Introduction to research with continuous cultures, Prentice Hall biological techniques series. Prentice-Hall, Englewood Cliffs
Lander ES, Linton LM, Birren B, International Human Genome Sequencing Consortium et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
Lang GI, Rice DP, Hickman MJ et al (2013) Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500:571–574
Lee C, Raffaghello L, Brandhorst S et al (2012) Fasting cycles retard growth of tumors and sensitize a range of cancer cell types to chemotherapy. Sci Transl Med 4:124ra27
Lehninger AL (1975) Biochemistry, 2nd edn. Worth, New York
Leichtle A, Dufour JF, Fiedler GM (2013) Potentials and pitfalls of clinical peptidomics and metabolomics. Swiss Med Wkly 143:w13801. doi:10.4414/smw.2013.13801
Lever MA, Rouxel O, Alt JC et al (2013) Evidence for microbial carbon and sulfur cycling in deeply buried ridge flank basalt. Science 339:1305–1308
Li XC, Schimenti JC, Tye BK (2009) Aneuploidy and improved growth are coincident but not causal in a yeast cancer model. PLoS Biol 7:e1000161
Li MX, Gui HS, Kwan JS et al (2012) A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. Nucleic Acids Res 40:e53
Lindquist S (2008) Using yeast to understand protein folding diseases: an interview with Susan Lindquist by Kristin Kain. Dis Model Mech 1:17–19
Li-Pook-Than J, Snyder M (2013) iPOP goes the world: integrated personalized omics profiling and the road toward improved health care. Chem Biol 20:660–666
Liu YY, Slotine JJ, Barabsi AL (2011a) Controllability of complex networks. Nature 473:167–173
Liu YY, Slotine JJ, Barabsi AL (2011b) Liu et al reply to: Müller FJ, Schuppert A (2011) Few inputs can reprogram biological networks. Nature 478:E4; discussion E4–E5
Loewith R, Hall MN (2004) TOR signalling in yeast: temporal and spatial control of cell growth. In: Hall MN, Raff M, Thomas G (eds) Cell growth. Control of cell size. Cold Spring Harbor Laboratory Press, New York, pp 139–165
López LC, Dos-Reis S, Espargaró A et al (2012) Discovery of novel inhibitors of amyloid β-peptide 1-42 aggregation. J Med Chem 55:9521–9530
López-Otín C, Blasco MA, Partridge L et al (2013) The hallmarks of aging. Cell 153:1194–1217
Luscombe NM, Babu MM, Yu H et al (2004) Genomic analysis of regulatory network dynamics reveals large topological changes. Nature 431:308–312
MacLellan WR, Wang Y, Lusis AJ (2012) Systems-based approaches to cardiovascular disease. Nat Rev Cardiol 9:172–184
Maher B (2007) Personal genomics: his daughter's DNA. Nature 449:773–776
Maher B (2013) Father's genetic quest pays off. Nature 498:418–419
Manioudaki ME, Poirazi P (2013) Modeling regulatory cascades using artificial neural networks: the case of transcriptional regulatory networks shaped during the yeast stress response. Front Genet 4:110
Marshall E (2011) Human genome 10th anniversary. Waiting for the revolution. Science 331:526–529
Marx V (2013) Model organisms: beyond the inner circle. Nat Methods 10:471–473
Mason RP, Giorgini F (2011) Modeling Huntington disease in yeast: perspectives and future directions. Prion 5:269–276
Mason RP, Casu M, Butler N et al (2013) Glutathione peroxidase activity is neuroprotective in models of Huntington's disease. Nat Genet. doi:10.1038/ng.2732

Matuo R, Sousa FG, Soares DG et al (2012) *Saccharomyces cerevisiae* as a model system to study the response to anticancer agents. Cancer Chemother Pharmacol 70:491–502

Mayer AN, Dimmock DP, Arca MJ et al (2011) A timely arrival for genomic medicine. Genet Med 13:195–196

Mayfield JA, Davies MW, Dimster-Denk D et al (2012) Surrogate genetics and metabolic profiling for characterization of human disease alleles. Genetics 190:1309–1323

McCarthy N (2012) Metabolism: unmasking an oncometabolite. Nat Rev Cancer 12:229

McGurk L, Bonini NM (2011) Cell biology. Yeast informs Alzheimer's disease. Science 334:1212–1213

Michaelis S, Hrycyna CA (2013) Biochemistry. A protease for the ages. Science 339:1529–1530

Mitchell SF, Jain S, She M et al (2013) Global analysis of yeast mRNPs. Nat Struct Mol Biol 20:127–133

Moazed D (2009) Molecular biology. Rejoice—RNAi for yeast. Science 326:533–534

Monico CG, Milliner DS (2011) Genetic determinants of urolithiasis. Nat Rev Nephrol 8:151–162

Montanari A, Zhou YF, D'Orsi MF et al (2013) Analyzing the suppression of respiratory defects in the yeast model of human mitochondrial tRNA diseases. Gene 527:1–9

Morava E, Vodopiutz J, Lefeber DJ et al (2012) Defining the phenotype in congenital disorder of glycosylation due to *ALG1* mutations. Pediatrics 130: e1034–e1039

Mori MA, Liu M, Bezy O et al (2010) A systems biology approach identifies inflammatory abnormalities between mouse strains prior to development of metabolic disease. Diabetes 59:2960–2971

Mungall CJ, Gkoutos GV, Smith CL et al (2010) Integrating phenotype ontologies across multiple species. Genome Biol 11:R2

Munkacsi AB, Chen FW, Brinkman MA et al (2011) An "exacerbate-reverse" strategy in yeast identifies histone deacetylase inhibition as a correction for cholesterol and sphingolipid transport defects in human Niemann-Pick type C disease. J Biol Chem 286:23842–23851

Murphy MP (2013) Special issue: animal models of disease. Biochim Biophys Acta 1832:1361. doi:10.1016/j.bbadis.2013.05.024

Musunuru K (2013) Genome editing of human pluripotent stem cells to generate human cellular disease models. Dis Model Mech 6:896–904

Nagels Durand A, Moses T, De Clercq R et al (2012) MultiSite Gateway vector set for the functional analysis of genes in the model *Saccharomyces cerevisiae*. BMC Mol Biol 13:30

National Institutes of Health Model organisms for biomedical research. http://www.nih.gov/science/models/; http://www.nigms.nih.gov/Education/modelorg_factsheet.htm

Neef DW, Turski ML, Thiele DJ (2010) Modulation of heat shock transcription factor 1 as a therapeutic target for small molecule intervention in neurodegenerative disease. PLoS Biol 8:e1000291

Neil H, Malabat C, d'Aubenton-Carafa Y et al (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 457:1038–1042

Noble D (2006) The music of life: biology beyond the genome. Oxford University Press, Oxford

Nookaew I, Papini M, Pornputtapong N et al (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. Nucleic Acids Res 40:10084–10097

Norcliffe JL, Alvarez-Ruiz E, Martin-Plaza JJ et al (2013) The utility of yeast as a tool for cell-based, target-directed high-throughput screening. Parasitology 2013:1–9. doi:10.1017/S0031182013000425

Nurse P (2003) The great ideas of biology. Clin Med 3:560–568

Nurse P, Hayles J (2011) The cell in an era of systems biology. Cell 144:850–854

Ocampo A, Barrientos A (2011) Developing yeast models of human neurodegenerative disorders. Methods Mol Biol 793:113–127

Ocampo A, Liu J, Barrientos A (2013) NAD+salvage pathway proteins suppress proteotoxicity in yeast models of neurodegeneration by promoting the clearance of misfolded/oligomerized proteins. Hum Mol Genet 22:1699–1708

Oliver SG (1997) Yeast as a navigational aid in genome analysis. Microbiology 143:1483–1487

Oliver SG (2002) Functional genomics: lessons from yeast. Philos Trans R Soc Lond B Biol Sci 357:17–23

Pang CP, Baum L, Lam DS (2000) Hunting for disease genes in multi-functional diseases. Clin Chem Lab Med 38:819–825

Park SK, Pegan SD, Mesecar AD et al (2011) Development and validation of a yeast high-throughput screen for inhibitors of Aβ oligomerization. Dis Model Mech 4:822–831

Pearce DA, Sherman F (1998) A yeast model for the study of Batten disease. Proc Natl Acad Sci U S A 95:6915–6918

Pearce DA, Sherman F (1999) Investigation of Batten disease with the yeast *Saccharomyces cerevisiae*. Mol Genet Metab 66:314–319

Pelechano V, Wei W, Steinmetz LM (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. Nature 497:127–131

Pereira C, Bessa C, Soares J et al (2012) Contribution of yeast models to neurodegeneration research. J Biomed Biotechnol 2012:941232

Picotti P, Clément-Ziza M, Lam H et al (2013) A complete mass-spectrometric map of the yeast prote-

ome applied to quantitative trait analysis. Nature 494:266–270
Pimentel C, Batista-Nascimento L, Rodrigues-Pousada C et al (2012) Oxidative stress in Alzheimer's and Parkinson's diseases: insights from the yeast *Saccharomyces cerevisiae*. Oxid Med Cell Longev 2012:132146
Pir P, Gutteridge A, Wu J (2012) The genetic control of growth rate: a systems biology study in yeast. BMC Syst Biol 6:4
Piston DW (2012) Research tools: understand how it works. Nature 484:440–441
Pittman AM, Lage MD, Poltoratsky V et al (2012) Rapid profiling of disease alleles using a tunable reporter of protein misfolding. Genetics 192:831–842
Pollitzer E (2013) Biology: cell sex matters. Nature 500:23–24
Porrua O, Libri D (2013a) A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. Nat Struct Mol Biol 20:884–891
Porrua O, Libri D (2013b) RNA quality control in the nucleus: the angels' share of RNA. Biochim Biophys Acta 1829:604–611
Porzoor A, Macreadie IG (2013) Application of yeast to study the tau and amyloid-β abnormalities of Alzheimer's disease. J Alzheimers Dis 35:217–225
Proctor M, Urbanus ML, Fung EL et al (2011) The Automated Cell: Compound and Environment Screening System (ACCESS) for chemogenomic screening. Methods Mol Biol 759:239–269
Pryor EE Jr, Horanyi PS, Clark KM et al (2013) Structure of the integral membrane protein CAAX protease Ste24p. Science 339:1600–1604
Przytycka TM, Andrews J (2010) Systems-biology dissection of eukaryotic cell growth. BMC Biol 8:62
Puffenberger EG, Jinks RN, Sougnez C (2012) Genetic mapping and exome sequencing identify variants associated with five novel diseases. PLoS One 7: e28936
Robinson R (2011) A yeast model for understanding ALS: fast, cheap, and easy to control. PLoS Biol 9: e1001053
Roemer T, Boone C (2013) Systems-level antimicrobial drug and drug synergy discovery. Nat Chem Biol 9:222–231
Rose AH, Harrison JS (1987–1995) The yeasts, vol 1–6. Academic Press, London
Roy S, Wapinski I, Pfiffner J et al (2013) Arboretum: reconstruction and analysis of the evolutionary history of condition-specific transcriptional modules. Genome Res 23:1039–1050
Ryan CJ, Krogan NJ, Cunningham P, Cagney G (2013) All or nothing: protein complexes flip essentiality between distantly related eukaryotes. Genome Biol Evol 5:1049–1059
Sansom C, Mendes M, Coveney P (2011a) Modelling health and disease. Ingenia 47:27–32. http://www.ingenia.org.uk/ingenia/issues/issue47/sansom.pdf
Sansom C, Mendes M, Coveney P (2011b) Modelling the virtual physiological human. BioTechnologia. J Biotech Comput Biol Bionanotechnol 92:225–229. http://www.biotechnologia-journal.org/sites/default/files/journals/fulltext/92_3/03.pdf
Sarewitz D (2012) Beware the creeping cracks of bias. Nature 485:149
Schenone M, Dančík V, Wagner BK et al (2013) Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol 9:232–240
Shah NH, Cole T, Musen MA (2012) Analyses using disease ontologies. PLoS Comput Biol 8:e1002827
Sherman F (1998) An introduction to the genetics and molecular biology of the yeast *Saccharomyces cerevisiae*. http://www.urmc.rochester.edu/labs/Sherman-Lab/publications/books.cfm?redir=dbb.urmc.rochester.edu
Sherman F (2002) Getting started with yeast. Methods Enzymol 350:3–41. Updated http://www.urmc.rochester.edu/labs/Sherman-Lab/publications/pdfs/Getting-Started-With-Yeast.pdf
Shi Y, Kirwan P, Smith J et al (2012) A human stem cell model of early Alzheimer's disease pathology in Down syndrome. Sci Transl Med 4:124ra29
Smith CH (2000) The Alfred Russel Wallace page http://people.wku.edu/charles.smith/index1.htm
Song J, Singh M (2013) From hub proteins to hub modules: the relationship between essentiality and centrality in the yeast interactome at different scales of organization. PLoS Comput Biol 9:e1002910
Soucek L, Whitfield J, Martins CP et al (2008) Modelling Myc inhibition as a cancer therapy. Nature 455:679–683
Sperling SR (2011) Systems biology approaches to heart development and congenital heart disease. Cardiovasc Res 91:269–278
Stansfield I, Stark MJR (2007) Methods in microbiology, vol 36. Yeast gene analysis, 2nd edn. Academic, London
Stepensky P, Saada A, Cowan M et al (2013) The Thr224Asn mutation in the VPS45 gene is associated with congenital neutropenia and primary myelofibrosis of infancy. Blood 121:5078–5087
Stobbe MD, Houten SM, Jansen GA et al (2011) Critical assessment of human metabolic pathway databases: a stepping stone for future integration. BMC Syst Biol 5:165
Stuart LM, Boulais J, Charriere GM et al (2007) A systems biology analysis of the *Drosophila* phagosome. Nature 445:95–101
Suhre K, Shin SY, Petersen AK et al (2011) Human metabolic individuality in biomedical and pharmaceutical research. Nature 477:54–60
Sun Z, Diaz Z, Fang X et al (2011) Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. PLoS Biol 9:e1000614

Sung MK, Lim G, Yi DG et al (2013) Genome-wide bimolecular fluorescence complementation analysis of SUMO interactome in yeast. Genome Res 23:736–746

Tan JL, Li QX, Ciccotosto GD et al (2013) Mild oxidative stress induces redistribution of BACE1 in non-apoptotic conditions and promotes the amyloidogenic processing of Alzheimer's disease amyloid precursor protein. PLoS One 8:e61246

Tan-Wong SM, Zaugg JB, Camblong J et al (2012) Gene loops enhance transcriptional directionality. Science 338:671–675

Tardiff DF, Lindquist S (2013) Phenotypic screens for compounds that target the cellular pathologies underlying Parkinson's disease. Drug Discov Today Technol 10:e121–e128

Tardiff DF, Tucci ML, Caldwell KA et al (2012) Different 8-hydroxyquinolines protect models of TDP-43 protein, α-synuclein, and polyglutamine proteotoxicity through distinct mechanisms. J Biol Chem 287:4107–4120

Tehlivets O, Malanovic N, Visram M et al (2013) S-adenosyl-L-homocysteine hydrolase and methylation disorders: yeast as a model system. Biochim Biophys Acta 1832:204–215

Telford MJ (2013) Evolution. The animal tree of life. Science 339:764–766

Tenreiro S, Outeiro TF (2010) Simple is good: yeast models of neurodegeneration. FEMS Yeast Res 10:970–979

Tenreiro S, Munder MC, Alberti S et al (2013) Harnessing the power of yeast to unravel the molecular basis of neurodegeneration. J Neurochem. doi:10.1111/jnc.12271

Tharp WG, Sarkar IN (2013) Origins of amyloid-β. BMC Genomics 14:290

Thompson DA, Roy S, Chan M et al (2013) Evolutionary principles of modular gene regulation in yeasts. Elife 2:e00603. doi:10.7554/eLife.00603

Treusch S, Hamamichi S, Goodman JL et al (2011) Functional links between Aβ toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. Science 334:1241–1245

U.S. Food and Drug Administration (2013) Regulatory information. http://www.fda.gov/regulatoryinformation/guidances/ucm122046.htm

Ulitsky I, Krishnamurthy A, Karp RM et al (2010) DEGAS: de novo discovery of dysregulated pathways in human diseases. PLoS One 5:e13367

van Triest HJ, Chen D, Ji X et al (2011) PhenOMIM: an OMIM-based secondary database purported for phenotypic comparison. Conf Proc IEEE Eng Med Biol Soc 2011:3589–3592

van Wyhe J (2002) The complete work of Charles Darwin online. http://darwin-online.org.uk

van Wyhe J (2012) Wallace online. http://wallace-online.org/

Vaux DL (2012) Research methods: know when your numbers are significant. Nature 492:180–181

Veličković VM (2013) Statistical misconceptions. Scientist 31 July 2013. http://www.the-scientist.com/?articles.view/articleNo/36781/title/Opinion–Statistical-Misconceptions/

Venter JC, Adams MD, Myers EW et al (2001) The sequence of the human genome. Science 291:1304–1351

Viceconti M, Clapworthy G, Van Sint JS (2008) The virtual physiological human—a European initiative for in silico human modelling. J Physiol Sci 58:441–446

Vidal M (2009) A unifying view of 21st century systems biology. FEBS Lett 583:3891–3894

Vidal M, Cusick ME, Barabsi AL (2011) Interactome networks and human disease. Cell 144:986–998

Villar-Piqué A, Ventura S (2013) Protein aggregation propensity is a crucial determinant of intracellular inclusions formation and quality control degradation. Biochim Biophys Acta 1833:2714–2724

Villoslada P, Baranzini S (2012) Data integration and systems biology approaches for biomarker discovery: challenges and opportunities for multiple sclerosis. J Neuroimmunol 248:58–65

Virtual Physiological Human (VPH) (2010) Network of excellence. Objectives. http://www.vph-noe.eu/objectives

Waern K, Snyder M (2013) Extensive transcript diversity and novel upstream open reading frame regulation in yeast. G3(Bethesda) 3:343–352

Walhout AJM, Aebersold R, Meyer T et al (2011) Systems biology: what's the next challenge? Cell 144:837–838, http://www.sciencedirect.com/science/article/pii/S0092867411002285

Wang X, Wei X, Thijssen B et al (2012) Three-dimensional reconstruction of protein networks provides insight into human genetic disease. Nat Biotechnol 30:159–164

Wang Y, Mercier R, Hobman TC et al (2013) Regulation of RNA interference by Hsp90 is an evolutionarily conserved process. Biochim Biophys Acta 1833:2673–2681

Washington NL, Haendel MA, Mungall CJ et al (2009) Linking human diseases to animal models using ontology-based phenotype annotation. PLoS Biol 7:e1000247

Weiss K (2012) What is the human genome? The Scientist, 17 Aug. http://www.the-scientist.com/?articles.view/articleNo/32446/title/Opinion–What-Is-the-Human-Genome–/

Westermann S, Schleiffer A (2013) Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. Trends Cell Biol 23:260–269

Wilson RB, Roof DM (1997) Respiratory deficiency due to loss of mitochondrial DNA in yeast lacking the frataxin homologue. Nat Genet 16:352–357

Winzeler EA, Shoemaker DD, Astromoff A et al (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285:901–906

Wishart JA, Osborn M, Gent ME et al (2006) The relative merits of the tetO2 and tetO7 promoter systems for the functional analysis of heterologous genes in yeast and a compilation of essential yeast genes with tetO2 promoter substitutions. Yeast 23:325–331

Worthey EA, Mayer AN, Syverson GD et al (2011) Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. Genet Med 13:255–262

Xu W, Yang H, Liu Y et al (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell 19:17–30

Xu Z, Poidevin M, Li X et al (2013) Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. Proc Natl Acad Sci U S A 110:7778–7783

Yen K, Gitsham P, Wishart J et al (2003) An improved tetO promoter replacement system for regulating the expression of yeast genes. Yeast 20:1255–1262

Yen K, Vinayachandran V, Batta K et al (2012) Genome-wide nucleosome specificity and directionality of chromatin remodelers. Cell 149:1461–1473

Young E (2013) We gained hope. The story of Lilly Grossman's genome. Phenomena National Geographic. http://phenomena.nationalgeographic.com/2013/03/11/we-gained-hope-the-story-of-lilly-grossmans-genome/

Yu D, Corbett B, Yan Y et al (2012) Early cerebrovascular inflammation in a transgenic mouse model of Alzheimer's disease. Neurobiol Aging 33:2942–2947

Zhang N, Bilsland E (2011) Contributions of *Saccharomyces cerevisiae* to understanding mammalian gene function and therapy. Methods Mol Biol 759:501–523

2 Genomics to Study Basal Lineage Fungal Biology: Phylogenomics Suggests a Common Origin

Ekaterina Shelest[1], Kerstin Voigt[2,3]

CONTENTS

[1]Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology– Hans-Knöll-Institute, Adolf-Reichwein-Strasse 23, 07745 Jena, Germany
[2]Jena Microbial Resource Collection, Leibniz Institute for Natural Product Research and Infection Biology–Hans-Knöll-Institute, Adolf-Reichwein-Strasse 23, 07745 Jena, Germany; e-mail: Kerstin.Voigt@hki-jena.de
[3]Department of Microbiology and Molecular Biology, Institute of Microbiology, University of Jena, Neugasse 25, 07743 Jena, Germany; e-mail: Kerstin.Voigt@hki-jena.de

I. Introduction

A. Accumulation of Knowledge in the Pregenomic Era: How Many Fungi Do We Know?

The total number of fungal species described so far ranges between 72,000 and 120,000 (Hawksworth and Rossman 1997; Hawksworth 2001). These fungal species are described by a total of 315,796 fungal names (as of September 16, 2013; Hibbett et al. 2009). More than one name is described for ascomycetous developing imperfect (anamorphic) and perfect (teleomorphic) stages or for rust fungi developing morphologically dissimilar thalli on different host plants. Just 5,129 names (ca. 1.6 %) were described for the fungi considered to be members of the basal phylogenetic lineages, the zoosporic and zygosporic fungi, the **Chytridiomycota** and **Zygomycota**, respectively (Figs. 2.1 and 2.2). The **Microsporidia** represent an ancient group of obligately parasitic microorganisms that appears as a sister group to the fungi. Their classification to or apart from the fungi has been extensively studied by a variety of publications performed by the mycologist and zoologist communities, respectively, reporting evidence supporting (Bohne et al. 2000; Fast and Keeling 2001; Keeling and Slamovits 2004; Capella-Gutiérrez et al. 2012) or contradicting (Barron 1987; Patterson and Zölffel 1991; Hirt et al. 1997; Hashimoto et al. 1998; Peyretaillade et al. 1998a, b; Patterson 1999; Delbac et al. 2001; Bacchi et al. 2002; Glockling and Beakes 2002; Tanabe et al. 2002) their fungal affiliation. Since the launch of the International Code of Nomenclature for algae, fungi, and plants (ICNAFP,

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

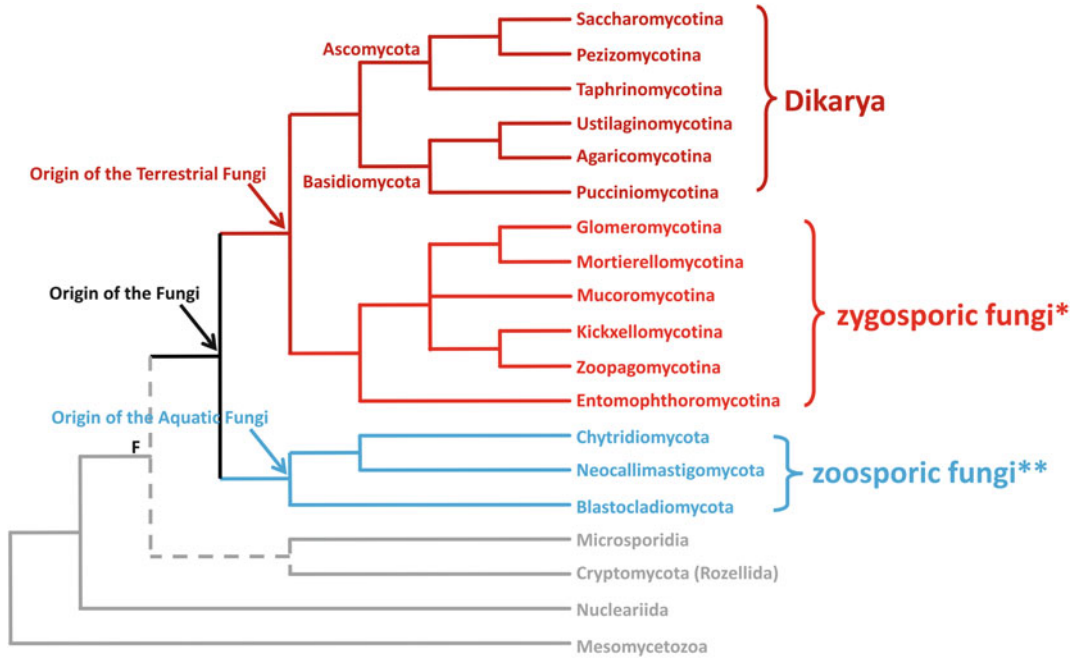


Fig. 2.1. Higher taxon-level phylogeny of the Fungi. Dendrogram of 15 fungal lineages showing consensus relationships in relation to four fungal allied groups. Phylogenetic relationships with high levels of uncertainty are indicated by stippled lines. * indicates the zygosporic fungi, which can be summarized to the phylum 'Zygomycota'; ** indicates the zoosporic fungi, which can be summarized to the phylum 'Chytridiomycota' *sensu lato* as reviewed by Voigt (2012a, b). The traditionally accepted node for delineating Fungi is marked, but there are current attempts to accept the term Fungi in a broader sense as indicated by 'F' (Adl et al. 2005; Lara et al. 2010; Jones et al. 2011a, b; James and Berbee 2011; Capella-Gutiérrez et al. 2012; Schoch et al. 2012). The term *Rozellida* has been superseded by Cryptomycota (Jones et al. 2011b). Both Mesomycetozoa and Ichthyosporea are in common usage

Melbourne Code; McNeill et al. 2012), the microsporidia have been excluded from the ICNAFP and described by the International Code of Zoological Nomenclature (ICZN; Ride et al. 2000). In conjunction with this decision, the Microsporidia and their allies (including the Cryptomycota, formerly Rozellida; Jones et al. 2011a, b) are treated as miscellaneous fungus-like organisms in this review.

Until now, 561 microsporidian species were described, which occupy 0.18 % of the total fungal names if implemented among the fungi (Fig. 2.2a). Amongst the basal fungi, 1,804, 449, and 2,876 species were described from the Chytridiomycota, the Glomeromycota, and the Zygomycota, respectively (Fig. 2.2a). About 1,804 species are distributed over nine orders of the Chytridiomycota (Fig. 2.2b). About 2,876 species are distributed over 5 subphyla and 10 orders of the Zygomycota (Fig. 2.2c, d, respectively). The Zygomycota represent the most basal terrestrial fungal phylum forming aplanosporic (non-flagellate) mitospores produced in multispored sporangia or few-spored sporangiola during the reproductive part and zygospores during the conjugative part of their life cycle (Fig. 2.3a, *a–i*). The zoosporic fungi, Chytridiomycota *sensu lato*, including Blastocladiomycota (James et al. 2006b), Chytridiomycota (Barr 2001), Monoblepharomycota (Doweld 2001), and Neocallimastigomycota (Hibbett et al. 2007), compose flagellate (planosporic fungi) producing zoospores aiming at reproduction in the aquatic environment (Fig. 2.3a, *j–n*). Both phyla and the obligate biotrophic, arbuscular endomycorrhizae, the Glomeromycota, are summarized to the basal fungi in this review. During the course of

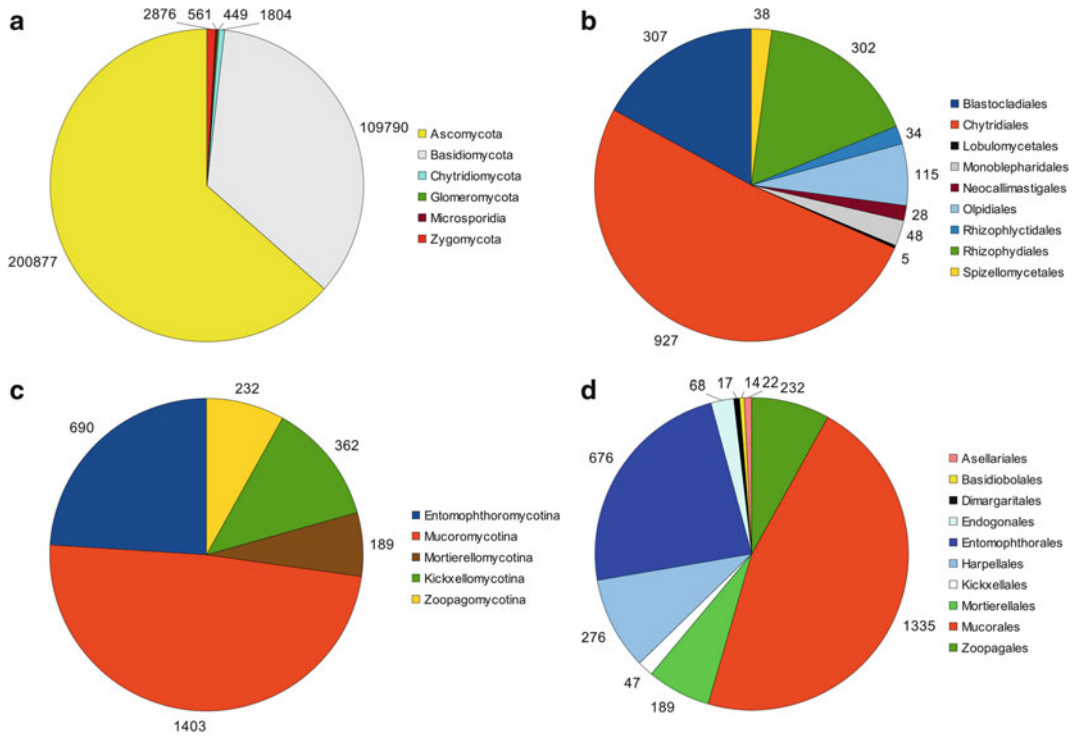


Fig. 2.2. Species distribution among the phyla of fungi (**a**), the orders of the Chytridiomycota (**b**), the subphyla of the Zygomycota (**c**) and the orders of the Zygomycota (**d**)

evolution, basal lineage fungi have adapted to changing environments. Therefore, these lineages are expected to be found in biotops that are characterized by complex interaction schemes that facilitate interplaying patterns with plants, animals, fungi, and humans and are accompanied by a diverse morphology (Fig. 2.3b).

B. Recent Status of the Genomic Era: What Has Been Achieved?

Fungi are nutrition-absorptive eukaryotes that exist in every ecological niche (Alexopoulos et al. 1996). They contribute to human society by either biotechnological importance or serious threats to human health. Because of their ease of handling, rapid life cycle, haploid genetics, and phylogenetic relationship to the animal kingdom (Baldauf and Palmer 1993), fungi are studied as model systems in every aspect of the life sciences (Bennett 2000). Important models for biomedical research originate from fungi, which provide a wide range of evolutionary comparisons at key branch points in the billion years of fungal evolution. Since the release of the first completed fungal genome, which was the genome of the ascomycetous yeast *Saccharomyces cerevisiae* (October 1996; Goffeau et al. 1996), a flood of genome projects was initiated, encompassing a total of 1,899 fungal genomes: 1,014, 752, 51, and 32 genomes from Ascomycota, Basidiomycota, basal fungi, and microsporidia, respectively (as of October 9, 2013; based on Genomes Online Database [GOLD], National Center for Biotechnology Information [NCBI], and European Nucleotide Archive [ENA]; Table 2.1, Fig. 2.4).

At the junction between the past and the new millennium, the fungal and genomics communities have been working together to define a **Fungal Genome Initiative (FGI)**, which is an effort to jumpstart research on the fungal kingdom by prioritizing a set of fungi for genome sequencing (Birren et al. 2002). The FGI,

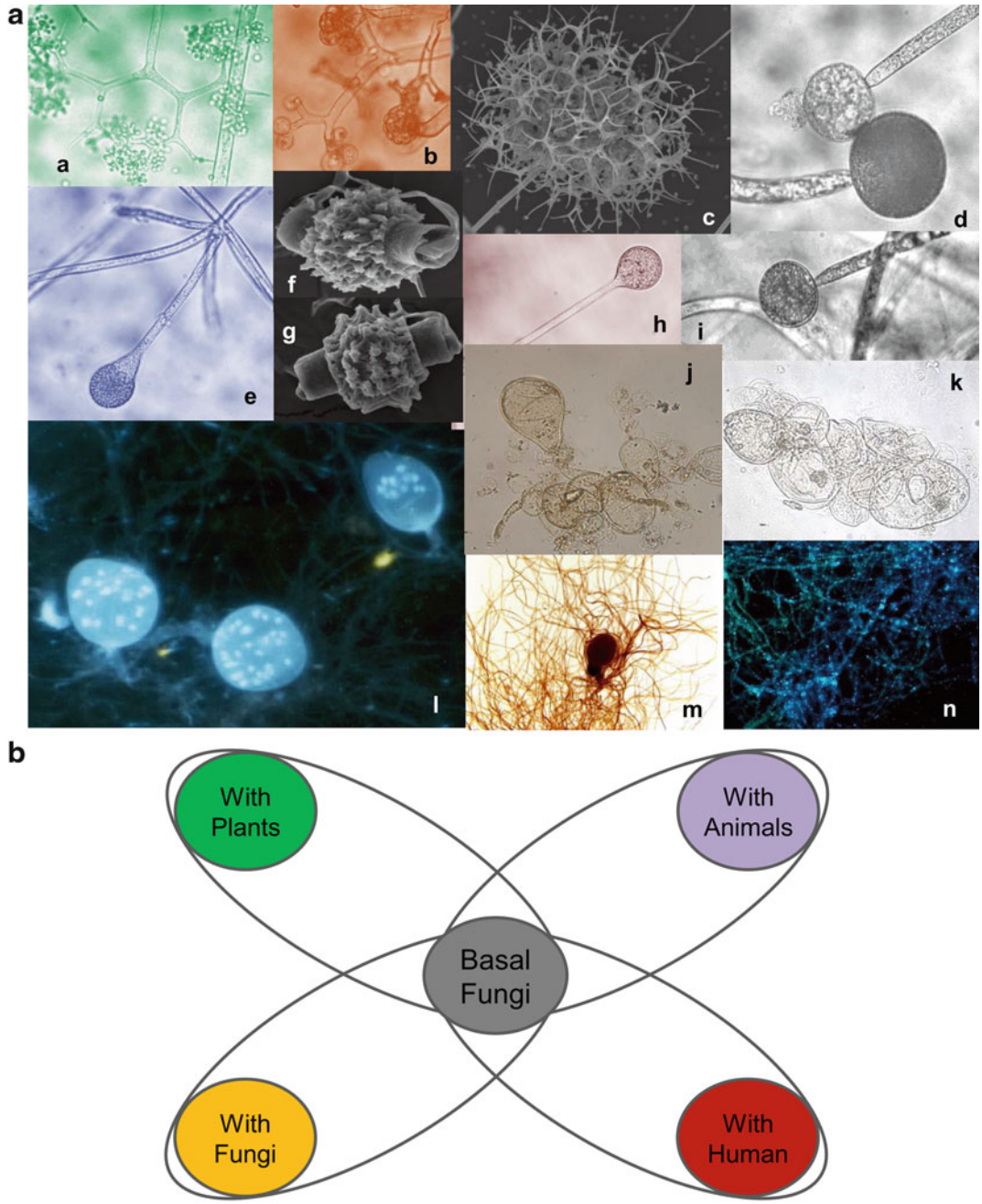


Fig. 2.3. (**a**) The morphological diversity of the basal fungi. *a–i* Zygomycetes (Mucoromycotina; photos: K. Hoffmann; scanning electron microphotographs *c*, *f* and *g*: M. Eckart and K. Hoffmann): *a–c* different types of uni-/few-spored sporangiola; *d, e, h, i* different types of multispored sporangia; *f, g* zygospores; *j–n* reproductive structures (zoosporangia) from anaerobic chytridiomycetes (Neocallimastigomycota; photos: K. Fliegerova) (Reprinted from Gherbawy and Voigt 2010 with permission from Springer-Verlag, Berlin). (**b**) Interaction scheme of basal lineage fungi

Table 2.1. Fungal genome sequencing projects (as of October 9, 2013; source: GOLD [Genomes Online Database], http://www.genomesonline.org)

Phylum	Complete		Incomplete[a]	Targeted[a]
	Finished	Permanent draft		
Ascomycota	36	62	837	79
Basidiomycota	6	21	666	59
Blastocladiomycota	0	1	1	0
Chytridiomycota	1	2	5	3
Cryptomycota	0	0	2	0
Entomophthoromycota	1	0	4	0
Glomeromycota	0	0	1	0
Microsporidia	2	2	28	0
Mortierellomycota	0	1	2	0
Mucoromycota	4	0	17	0
Neocallimastigomycota	0	0	7	2

[a]Only whole-genome sequencing projects

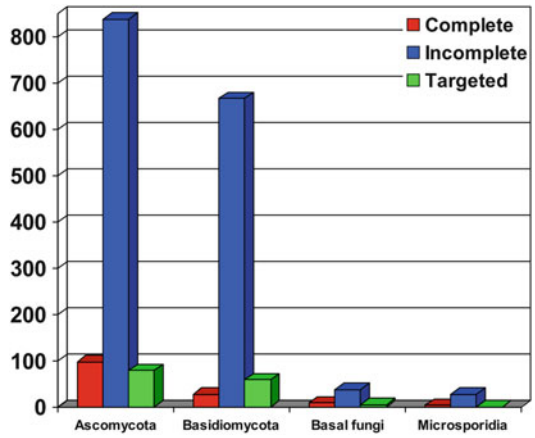


Fig. 2.4. Fungal genome projects provided by the International Nucleotide Sequence Database Collaboration at a glance (as of October 9, 2013)

launched in 2002, is carried out at the Broad Institute of the Massachusetts Institute of Technology (MIT) and Harvard University. In a following initiative, the Broad Institute runs the **Origins of Multicellularity Project (OMP)**, which was launched in 2007 (Ruiz-Trillo et al. 2007). In the year 2010, China's fungal genomics initiative was started with a white paper as a result of the 2009 **China Fungal Genome Initiative** Symposium in Shanghai when the status of China's fungal genome research was examined and, for the first time, focused attention on the extent, range, and quality of fungal genomics in China (An et al. 2010). Following this white paper, a series of recommendations and actions were put in motion to coordinate the various genome projects, ranging from selection of fungi for sequencing, to the development of a bioinformatics infrastructure, and accompanying research ideas, funding and commercialization.

In 2011, the **1000 Fungal Genomes Project (1KFG)**, hosted by the U.S. Department of Energy's Joint Genome Institute (DOE-JGI), was launched as part of the **Fungal Genomics Program (FGP)**. The FGP aims at the future advances made possible by genomic analyses of fungi that have an impact on plant/feedstock health, degradation of lignocellulosic biomass, and fermentation of sugars to ethanol, hydrocarbon biofuels, and renewable chemicals (Grigoriev et al. 2011). The 1KFG project aims at the family-level sampling and sequencing of an unsampled diversity of fungal genomes across the Fungal Tree of Life, an idea that was born as a follow-up of the **Assembling the Fungal Tree of Life (AFTOL)** project (Spatafora 2011). Beyond generating tens of trillions of DNA bases annually, DOE-JGI develops and maintains data management systems and specialized analytical capabilities to manage and interpret complex genomic datasets. Therefore, the JGI Genome Portal (http://genome.jgi.doe.gov) provides **MycoCosm** (http://jgi.doe.gov/fungi), a new integrated fungal genomics resource implemented in the

general organization of the Genome Portal (Grigoriev et al. 2012). MycoCosm is a unified access point to all JGI genomic databases and analytical tools for displaying sequencing projects and their status and searching for and downloading assemblies and annotations of sequenced genomes that are tailored to each particular class of organisms.

Genome projects targeting the Ascomycota, as the species-richest and most prominent phylum, traditionally dominate the past and the present (Fig. 2.4). Of a total of 51 genome projects from the basal fungi, 22 are from zoosporic fungi (Chytridiomycota *sensu lato*), 28 from zygosporic and 1 from arbusclar endomycorrhizal fungi (Table 2.1). The genome projects performed on the zoosporic fungi comprise 2, 11, and 9 from the Blastocladiomycota, Chytridiomycota *sensu stricto,* and Neocallimastigomycota, respectively. None of the genome proejcts refers to the Monoblepharidomycota, which are characterized to sexually reproduce by oogamy, contrary to members of the other phyla of zoosporic fungi, which reproduce by isogamy, anisogamy or schizogamy. The genome projects performed on the zygosporic fungi encompass 5 and 23 for the Entomophthoromycota and Zygomycota, respectively. Among the fungal allies, the Cryptomycota are represented by 2 and the microsporida by 32 genome projects, mostly as incomplete drafts (Table 2.1). The genomes of *Conidiobolus coronatus* JMRC:FSU:4392 (unpublished), *Lichtheimia corymbifera* JMRC:FSU:9682 (PRJEB3978), *Mucor circinelloides* f. *lusitanicus* CBS277.49 (PRJNA46717), *Phycomyces blakesleeanus* NRRL1555 (mating type –; PRJNA61391), *Rhizopus delemar* (*ex: oryzae*) RA99-880 (PRJNA13066) from the Zygomycota and the genomes from *Allomyces macrogynus* ATCC38327 (PRJNA20563), *Batrachochytrium dendrobatidis* JAM81 (PRJNA41157), *Batrachochytrium dendrobatidis* JEL423 (PRJNA13653) and *Spizellomyces punctatus* DAOM BR117 (PRJNA37881) from the Chytridiomycota can be regarded as completed (finalized or permanent draft; Table 2.2).

II. Genomes of the Basal Lineages of the Fungal Kingdom: A Chronicle from the Systematic Point of View

Basal lineages of the fungi traditionally encompass **zoosporic (planosporic)** and **zygosporic (aplanosporic) fungi**. The zoosporic fungi propagate asexually via flagellate spores and compose the Blastocladiomycota, the Chytridiomycota, and the Neocallimastigomycota (Fig. 2.1). The group of the zoosporic fungi is summarized as 'Chytridiomycota' *sensu lato* (Voigt 2012a). The zygosporic fungi are traditionally called the Zygomycota, which are counted as the most basal terrestrial phylum of the kingdom of Fungi (Fig. 2.1). The Zygomycota are not accepted as a valid phylum because of they lack compliance to the International Code of Botanical Nomenclature (now the ICNAFP; Hawksworth 2011) (no Latin diagnosis, no designated type as 'Phylum des Zygomycètes'; Whittaker 1969; Cavalier-Smith 1981; no description), and they lack resolution of the basal fungal clades (James et al. 2006a). Molecular phylogenetic analyses revealed dispersal of Zygomycetes into five subphyla containing one to four orders (Hibbett et al. 2007; Hoffmann et al. 2011). The phylogenetic relationships between these subphyla and their orders are not well resolved yet. However, based on the potential of all five subphyla to produce zygospores during conjugation of two yoke-shaped gametangia, it is treated as a phylogenetically coherent group named the Zygomycota for zygosporic fungi in this review. They share morphological features (Voigt 2012b) but consist of phylogenetically unrelated subphyla (James et al. 2006a; Hibbett et al. 2007; Hoffmann et al. 2011). Therefore, the phylum referred to as 'Zygomycota' is employed in a colloquial sense, for instance, including all basal lineages of terrestrial fungi with the potential to form zygospores or sharing any other of the plesiomorphic morphological characters of the former phylum (Voigt 2012b). Members of the Zygomycota are common and cosmopolitan components of the mycoflora of

Table 2.2. Completed genome sequencing projects from basal fungi

	Finished			Permanent draft		
	GOLD-ID	NCBI accession	NCBI-ID	GOLD-ID	NCBI accession	NCBI-ID
Chytridiomycota	–	–	–	Gi02370	PRJNA37881	37881
				Gi01093	PRJNA41157	41157
				Gi00978	PRJNA13653	13653
				Gi02368	PRJNA20563	20563
Microsporidia	Gc01518	PRJNA42703	42703	Gi03638	PRJNA32971	32971
	Gc00071	PRJNA13833	13833	Gi07638	PRJNA51843	51843
Mucoromycota	Gi01096	PRJNA61391	61391			
	Gi00506	PRJNA13066	13066			
	–	PRJNA46717	46717			
	–	PRJEB3978[a]	–			
Entomophthoromycota	–	*C. coronatus*[b]	–			

As of October 9, 2013; source: GOLD (Genomes Online Database), http://www.genomesonline.org; NCBI (National Center for Biotechnology Information), http://www.ncbi.nlm.nih.gov/genome/; and ENA (European Nucleotide Archive), http://www.ebi.ac.uk/ena/

Numbers for the NCBI-ID refer to NCBI project identifications numbers (IDs)

[a]ENA (European Nucleotide Archive) submission of the genome of the mucoralean fungus *Lichtheimia corymbifera* JMRC:FSU:9682 to be released in 2014 (see Fig. 2.6)

[b]The genome of the entomophthoralean fungus *Conidiobolus coronatus* JMRC:FSU:4392: to be released in 2014 (unpublished)

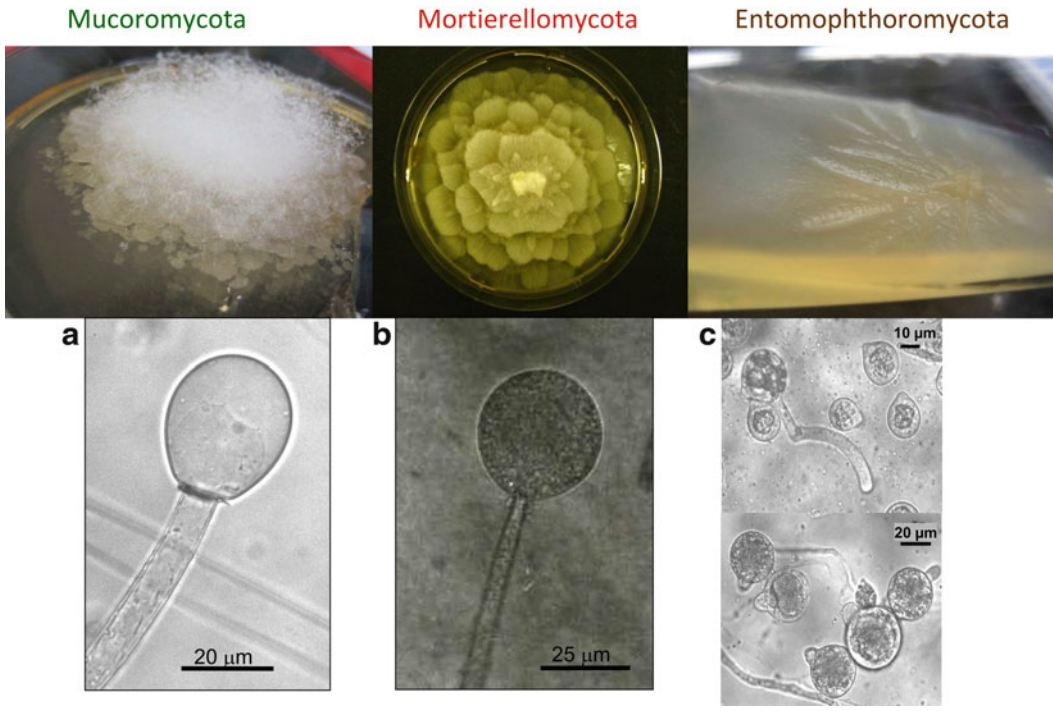


Fig. 2.5. Morphology of the colonies and the sporangia of the (**a**) Mucoromycota (Mucorales): *Mucor mucedo* JMRC:FSU:0620; (**b**) Mortierellomycota (Mortierellales): *Mortierella parvispora* JMRC:FSU: 2736; and (**c**) Entomophthoromycota (Entomophthorales): *Conidiobolus coronatus* JMRC:FSU:4392

dung, soil, and other substrates that support their growth and sporulation. Reproduction occurs asexually via aplanosporic, non-motile mitospores formed in sporangia as shown in the most abundant three orders, which are the Mucorales, the Mortierellales and the Entomophthorales (Fig. 2.5).

A. Aquatic and Planosporic (Zoosporic) Fungi

The zoosporic fungi, like any other fungal group, contain chitin in their cell walls and are primarily osmotrophic (Bartnicki-Garcia 1970, 1987). They **reproduce** either via **planosporic (flagellate) mitospores (zoospores)**, which are **borne in multispored zoosporangia or via planosporic meiospores (gametes) in gametangia arising from diploid mycelium produced after germination of a zygote** (Voigt 2012a).

The zoosporic fungi represent the most basal, and thus most ancient, group of the fungal lineages producing chitin as a structural component of the assimilative cell wall and converting the phagotrophic to an osmotrophic mode of nutrient uptake throughout the main phases of their life cycle (Table 2.1). Whether or not zoospores are able to perform phagotrophic uptake of nutrients has not been clarified yet (Marano et al. 2012). The zoospores are posteriorly uniflagellate or polyflagellate, bearing typically a single flagellum in aerobic chytrids (Blastocladiomycota and Chytridiomycota) and multiple flagella in anaerobic chytrids (Neocallimastigomycota), both always of the whiplash type and posteriorly localized (at their hind end = opisthokont).

Of a total of 22 genome projects targeting the zoosporic true fungi, 2, 11, and 9 are from the Blastocladiomycota, Chytridiomycota *sensu stricto,* and Neocallimastigomycota, respectively (Table 2.1). Just four projects are completed and exist as permanent drafts (Table 2.2). These are the genomes from *Allomyces macrogynus* ATCC38327 (PRJNA20563), *Batrachochytrium dendrobatidis* JAM81 (PRJNA41157), *Batrachochytrium dendrobatidis* JEL423 (PRJNA13653) and *Spizellomyces punctatus* DAOM BR117 (PRJNA37881) composing the Blastocladiomycota (*Allomyces macrogynus*) and the Chytridiomycota *sensu stricto* (the last three). In the following, the fungi and their characteristics qualifying their utilization for genome projects are elucidated in more detail.

1. Blastocladiomycota

The phylum Blastocladiomycota was proposed by Timothy Y. James as a result of phylogenetic analyses based mainly on the nuclear ribosomal DNA combined with ultrastructural data (James et al. 2006b) and multilocus analyses (James et al. 2006a; Liu et al. 2006). Members of the Blastocladiomycota form thalli, which exhibit variable complexity ranging from monocentric (comprising a single cell) to polycentric (typically branched with two or more centres of development; Voigt 2012a). The main characteristic, which is synapopmorhic to the phylum, is **prominent basal cell-bearing rhizoids** and one or more reproductive structures. Reproduction occurs asexually or sexually via zoospores with a single posteriorly directed flagellum, which are released from zoosporangia possessing a single, apical discharge papilla. Isogamous or anisogamous gametes are formed during sexual conjugation. **Generations change between diploid sporothalli and haploid gametothalli**, which are morphologically identical. The sporothalli bear thin-walled mitosporangia (2n) and thick-walled meiosporangia (1n), in which meiosis occurs. The gametothalli bear large and small gametangia, which contain the female and male gametes, respectively. This phenomenon, named **isomorphic alternation of sporothallic and gametothallic generations**, resembles the Eu-type, which is represented by *Allomyces macrogynus*. In special cases, the alternation of sporothallic and gametothallic generation may also be heteromorphic. The gametothallic generation either reduces to an amoeboid structure (cysts) without forming a distinct thallus (Cystogenes type) or is entirely lacking (Mehrotra and Aneja 1990; Webster and Weber 2007). The remarkable nature of the alternations of the generations forming single-cell and multicellular structures and their missing-link position in the Fungal Tree of Life qualifies the Blastocladiomycota as confirmed subjects for genome projects.

Consequently, the genomes of *Allomyces macrogynus* ATCC38327 (PRJNA20563, NCBI-ID 20563, Gi02368) and *Catenaria anguillulae* PL171 (PRJNA69551, NCBI-ID 69551, Gi 13919) were carried out by the Broad Institute and DOE-JGI, respectively.

a) *Allomyces macrogynus* and the Origins of Multicellularity Project

The genome project of *Allomyces macrogynus* ATCC38327 is part of the OMP initiated by the Broad Institute and funded by the National Human Genome Research Institute, which is part of the National Institutes of Health (NIH-NHGRI). The permanent draft of the genome (PRJNA20563), containing 17,500 predicted proteins, was completed in February 2010 (Table 2.2).

The OMP database provides access to multiple genomes, which are sequenced to investigate **commonalities and differences underlying multicellularity in animals and fungi** in order to understand the origins of multicellularity (Ruiz-Trillo et al. 2007). The emergence of multicellular organisms from single-cell ancestors occurred several times and independently during eukaryotic evolution. The emergence of multicellularity is one of the most profound evolutionary transitions in the history of life. It created evolutionary novelties, amongst others the need for cooperation and communication between cells and the division of labour among different cell types. The OMP is a multitaxon genome-sequencing initiative that aims to generate extensive genomic data from ten of the closest extant unicellular relatives of both animals and fungi to gain insights into how multicellularity first evolved. The genetic changes that accompanied the several origins of multicellularity are elucidated at the genomic level.

b) *Catenaria anguillulae* and the 1000 Fungal Genomes Project

Catenaria anguillulae is a facultative parasite belonging to the family Catenariaceae, which forms predominantly polycentric (typically branched with two or more centres of development), tubular, rhizomycelial thalli with catenulate swellings separated by sterile isthmi, thus growing in a hypha-like form, with rhizoids and chains of zoosporangia. Its zoosporangia are intercalary formed from the catenulate swellings (Voigt 2012a). The genome project of *Catenaria anguillulae* PL171 was launched in July 2011 and is embedded in 1KFG initiated by the DOE-JGI. The 1KFG project aims at the **family-level sampling and sequencing of unsampled diversity of fungal genomes across the Fungal Tree of Life**, an idea that was born as a follow-up of the AFTOL project, which was funded by the National Science Foundation (NSF). As an outcome of the AFTOL project, a need became evident for systematic sampling of the Fungal Tree of Life to fully leverage our knowledge of fungal evolution and its application to genome-enabled mycology (Spatafora 2011). To bridge this gap, the DOE-JGI has embarked on a 5-year project to sequence 1000 fungal genomes from across the Fungal Tree of Life—the 1KFG project (Spatafora 2011).

2. Chytridiomycota

The thallus of the Chytridiomycota *sensu stricto* is monocentric or rhizomycelial polycentric up to filamentous. Typical is **the lack of the sterile basal cell**, which is characteristic for the Blastocladiomycota when monocentric thalli are formed (Voigt 2012a). Many species are parasitic (obligately or facultatively) to algae, plants, and animals.

a) *Batrachochytridium dendrobatidis*

The only documented case of a zoosporic fungus parasitizing vertebrates is *Batrachochytrium dendrobatidis* (Bd), one of the most prominent species of the Rhizophydiales. This species is a highly destructive parasite causing **chytridiomycosis in many species of amphibians** (frogs and salamanders) (Longcore et al. 1999; Fisher et al. 2009; Martel et al. 2013). The pathogen infects over 350 species of amphibians and is found on all continents except Antarctica; it causes considerable decline in amphibian populations worldwide (Daszak et al. 1999; Fisher et al. 2009). The populations of parasites can be maintained in long-lived

tadpoles, which are infected but show no symptoms of disease (Briggs et al. 2010). The epidermis of healthy amphibians is known to regulate osmotic balance, which is inhibited during the infection with Bd because of reduction of potassium, sodium and chloride in the blood plasma (Voyles et al. 2009, 2012). Asystolic cardiac arrest follows and results in the death of post-metamorphic frogs (for review, see Voigt et al. 2013). In Central America, there is compelling evidence for the recent spread of pathogenic Bd and for its strong impact on amphibian populations, suggesting an increase in Bd virulence as it spread southwards, a pattern consistent with rapid evolution of increased virulence on Bd's invading front (Philipps and Puschendorf 2013). The impact of Bd on amphibians might therefore be driven by rapid evolution in addition to more proximate environmental drivers. The history and the emergence of the batrachochytridiomycosis causing global decline and extinction of frog and salamander populations at rates exceeding by some 200 times historical extinction levels since more than two decades was reviewed by Collins (2013).

The genomes of 29 isolates of Bd from around the world (with an emphasis on North, Central, and South America as hot spots of infection outbreaks in the New World) were determined to compose a substantial amount of **evolutionary complexity** and **genetic variation** indicating Bd's complex evolutionary history pre-dates recent disease outbreaks (Rosenblum et al. 2013). Dramatic differences among isolates and among genomic regions in chromosomal copy number and patterns of heterozygosity were found, suggesting complex and heterogeneous genome dynamics and evidence for selection acting on the Bd genome, supporting the hypothesis (proposed by Joneson et al. 2011) that protease genes are **important in evolutionary transitions** in this group. To study the mechanisms underlying species heterogeneity and rapid evolution, two more genome projects were initiated on Bd, namely, on *B. dendrobatidis* strain JAM81 (PRJNA41157) and *B. dendrobatidis* strain JEL423 (PRJNA13653) launched by the DOE-JGI and Broad Institute, respectively (Table 2.2). The genome of *B. dendrobatidis* JEL423 comprises 23.722 Mbp (8,818 predicted proteins on 8,806 Open Reading Frames [ORFs]). To understand the molecular mechanisms of Bd's emergence as a pathogen, which is central to mitigate the impacts of novel infectious disease agents, the genome of Bd's closest known relative, the non-pathogenic chytrid *Homolaphlyctis polyrhiza* (Hp), was sequenced (Joneson et al. 2011). This study had four major findings: (a) the genome of Hp is comparable to other chytrid genomes in size and number of predicted proteins; (b) a total of 1,974 Bd-specific genes, a gene set that is enriched for protease, lipase, and microbial effector gene ontology terms was identified; (c) a significant lineage-specific expansion of three Bd protease families (metallo-, serine-type, and aspartyl proteases) occurring after the divergence of Bd and Hp from their common ancestor was described; (d) the timing of the protease gene family expansions pre-dates the emergence of *Bd* as a globally important amphibian pathogen.

b) *Spizellomyces punctatus*

Spizellomyces punctatus is a member of the order Spizellomycetales, which was described by Barr (1980) and is characterized by the formation of monocentric thalli developing endogenously (endobiotic) or exogeneously (epibiotic). The members of this order are cosmopolitan saprobes or parasites, predominantly found in soil on organic substrata, including other fungi, cellulose, keratin, pollen, and plant debris. The genera of the Spizellomycetales differ by cytological variation of their zoospores (Barr 1984). The genome of *Spizellomyces punctatus* strain DAOM BR117 (PRJNA37881) comprises 24.131 Mbp (8,804 predicted proteins on 8,779 ORFs) and was recently completed within the OMP by the Broad Institute (Table 2.2).

3. Neocallimastigomycota

The Neocallimastigomycota are anaerobic gut fungi representing the **only oxygen-intolerant group among the fungi** and a distinct early-branching fungal phylum, with members that reside in the rumen, hindgut, and feces of

ruminant and nonruminant herbivores (Liggenstoffer et al. 2010). The phylum contains a single order and a single family, which are the Neocallimastigales and the Neocallimastigaceae, respectively. All members are strictly anaerobic, **inhabiting the digestive tract and rumen of herbivorous animals**. Their thallus lacks mitochondria but contains hydrogenosomes of mitochondrial origin. Zoospores are posteriorly uniflagellate and often polyflagellate.

The evidence for multiple flagellae in the Neocallimastigomycota contradicts the hypothesis that the flagellum was lost only once in the fungal lineage (Liu et al. 2006). Loss of flagellae took place several times among the neocallimastigomycotan genera and their species. The number of flagellae is species and genus specific (Barr 1978, 2001; Voigt 2012a).

Six genera are known: *Caecomyces*, *Neocallimastix* and *Piromyces*, which form monocentric thalli; and *Anaeromyces*, *Cyllamyces*, and *Orpinomyces*, which form polycentric thalli (Eckart et al. 2010). The anaerobic fungi are remarkable biomass degraders, capable of simultaneous saccharification and fermentation of the cellulosic and hemicellulosic fractions in multiple untreated grasses and crop residues examined, with the process significantly enhanced by mild pretreatments (Noha et al. 2013). Regardless of their important role in biofermentation, none of the completed genome projects refers to the Neocallimastigomycota. There are several genome sequencing programs under way in the USA, but these are hampered by the highly AT-rich genomes (Griffith et al. 2010). In addition to the difficulty in cultivating the anaerobic fungi in axenic cultures apart from their bacterial symbionts specifying the unique ecosystem of the rumen microbiota, particularly this high content of AT-rich stretches in their genomes is creating problems in completion of the genome assembly.

Five whole-genome sequencing projects are currently running, all of them in incomplete status. These are *Neocallimastix frontalis* strain 27 (Gi0047884); *Neocallimastix* sp. strain SC7 (Gi09601); *Neocallimastix* sp. strain 060013 (Gi07700), composing the monocentric-type genus of the phylum; the polycentric *Anaeromyces mucronatus* YE505 (Gi17207); as well as *Orpinomyces* sp. C1A (Gi0048210, Gi0044890, and Gi12080) (based on information from JGI's GOLD at http://www.genomesonline.org). The genome of *Orpinomyces* sp. C1A (Gi12080) was sequenced using a combination of Illumina and PacBio single-molecule real-time (SMRT) technologies (Noha et al. 2013). The genome reveals an extremely large genome size of 100.95 Mb (16,347 genes) and displays extremely low G+C content (17.0 %), large noncoding intergenic regions (73.1 %), proliferation of microsatellite repeats (4.9 %), and multiple gene duplications. Comparative genomic analysis identified multiple genes and pathways that are absent in Dikarya genomes but present in early-branching fungal lineages or nonfungal Opisthokonta.

Among these are genes for posttranslational fucosylation, the production of specific intramembrane proteases and extracellular protease inhibitors, the formation of a complete axoneme and intraflagellar trafficking machinery, and a near-complete focal adhesion machinery.

Analysis of the lignocellulolytic machinery in the C1A genome indicated an extremely rich repertoire, with evidence of horizontal gene acquisition from multiple bacterial lineages. That finding will be of primary importance for the inclusion of anaerobic fungi in biogas fermentation processes, which are often hampered by a large accumulation of excessive lignin as a waste product of crop residue-based biogas production. The capability of simultaneous saccharification and fermentation of the cellulosic and hemicellulosic fractions in multiple untreated grasses and crop residues, acquired during its separate evolutionary trajectory in the rumen, along with its resilience and invasiveness compared to prokaryotic anaerobes, renders this fungus a promising agent for consolidated bioprocessing schemes in biofuels production (Noha et al. 2013). Further genome projects on *Orpinomyces* sp. OUS1 (Gi01756) and *Piromyces* sp.E2 (Gi01097) are targeted. The genome project of *Orpinomyces* sp. OUS1 is conducted within the 1KFG project.

B. Terrestrial and Aplanosporic Fungi

Traditionally, fungi classified into the the **Zygomycota** are counted as the most basal terrestrial phylum of the kingdom of Fungi. Molecular phylogenetic analyses revealed dispersal into **five subphyla** containing one to four orders (Hibbett et al. 2007; Hoffmann et al. 2011). Based on the potential of all five subphyla to produce zygospores during conjugation of two yoke-shaped gametangia, they are referred to as a phylogenetically coherent group named the Zygomycota for zygosporic fungi as a whole, which share morphological features but consist of subphyla whose phylogenetic relationships are not resolved yet. Sexual conjugation results in formation of zygospores.

Thalli are usually filamentous, occasionally dimorphic (yeast and a filamentous stage depending on hypoxia or normoxia, respectively). The colonies develop typically non-septate (coenocytic) syncytium-forming mycelia, which produce aplanosporic (non-flagellate) mitospores borne in multispored sporangia or uni-/few-spored sporangiola arising from different shape sporangiophores (Voigt 2012b). Lifestyle is primarily terrestrial, although also facultatively aquatic species, which are either saprobic or associated with an aquatic host, are observed.

Of a total of 51 genome projects from the basal fungi, more than half (28) target zygosporic fungi (Table 2.1). Zygosporic fungi play an important role as destruents of organic matter in the biosphere. Their natural habitat is the soil. They influence human life and health as opportunistic pathogens causing mucormycosis, an emerging, life-threatening infection characterized by rapid angioinvasive growth with an overall mortality rate that exceeds 50 % in (mostly) immunocompromised patients, such as those with diabetes mellitus, cancer, or AIDS. Several members are of biotechnological importance as biotransformators in the hydroxylation of steroid compounds, carotene producers or microbial additives in food fermentation. Microscopic images from the three most important phyla (Mucoromycota, Mortierellomycota and Entomophthoromycota) are shown in Fig. 2.5a–c, respectively.

1. Mucoromycota

The thalli of the Mucoromycota typically consist of aerial mycelia. The mycelium is usually extremely robust, developing vigorously growing aerial hyphae (Fig. 2.5a). Fertile hyphae are often septate (microporic) at maturity. The sporangiophores arise from trophocysts. Asexual reproduction occurs by spores formed in sporocarps, which are multispored sporangia (Fig. 2.5a), or uni-/few-spored sporangiola or (less commonly) merosporangia produced in soil or on organic material aboveground. Sporangia are occasionally forcibly discharged (e.g. the genera *Pilobolus* and *Utharomyces*). Dimorphism occurs often in the genera *Mucor*, *Benjaminiella* and *Rhizopus*. Sporangia develop a visible, variously shaped columella, which always protrudes into the sporangium. The existence of a visible, well-developed columella in multispored sporangia counts as the trademark and synapomorphic character of the Mucoromycota. Members of the Mucoromycota are ubiquitous and mesotrophic, saprobes or facultative parasites of fungi and plants, occasionally endophytic in plants. The order Mucorales encompasses several **human pathogenic species**, which are opportunistic, systemic infection-causing pathogens of animals (including humans) (Fig. 2.6). **Mucormycoses** are associated with rapid blood vessel invasion and massive necrotic destruction of tissue (Sugar 1992; Ribes et al. 2000). Mortality rates are high (~50 %), and treatment mainly includes a combination of antifungals and extensive surgery (Roden et al. 2005; Chakrabarti et al. 2006; Cornely et al. 2009; Skiada et al. 2011). Of a total of 51 genome projects from the basal fungi, 21 (41 %) target the Mucoromycota. The genomes of *L. corymbifera* JMRC:FSU:9682 (PRJEB3978), *Mucor circinelloides* f. *lusitanicus* CBS277.49 (PRJNA46717), *P. blakesleeanus* NRRL1555 (mating type -; PRJNA61391), *R. delemar* (ex*: oryzae*) RA99–880 (PRJNA13066) are completed (finalized or permanent draft; Table 2.2).

a) *Mucor circinelloides*

Mucor circinelloides develops, as any other *Mucor* species, aerial hyphae, which are some-

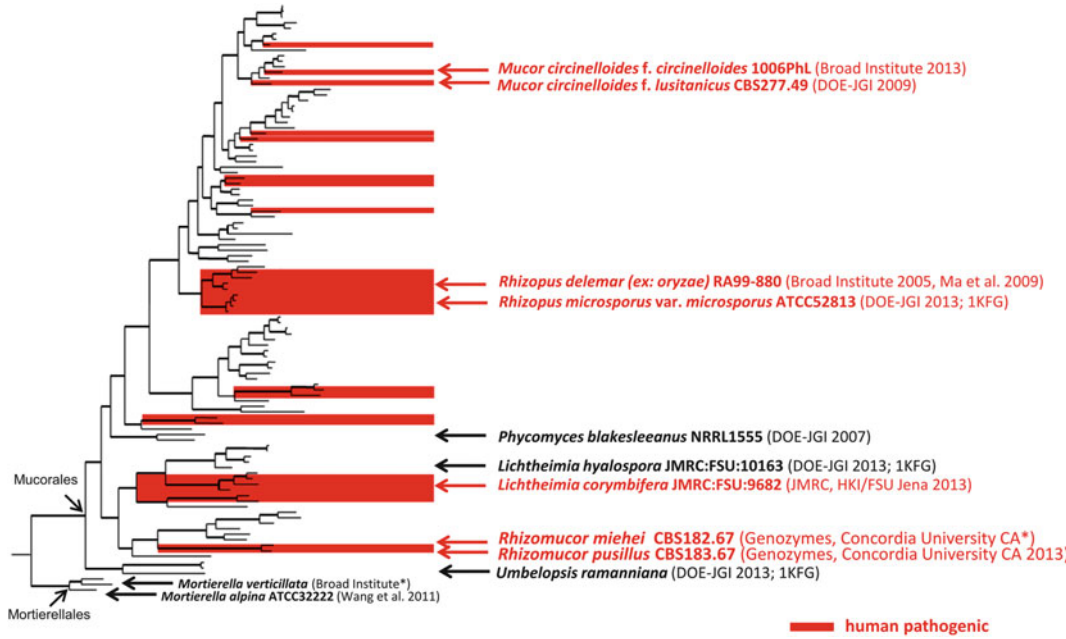


Fig. 2.6. Genome projects among the Mucorales shown on a phylogenetic tree based on a concatenated Bayesian inference analysis using nuclear small-subunit (18S) and large-subunit (28S) ribosomal DNA and exonic nucleotide sequences encoding actin and translation elongation factor 1alpha from 106 taxa. Mortierellales were treated as outgroup. The human pathogenic lineages are indicated in *red*, clinicalla non-relevant lineages are indicated in *black* (Modified after Gherbawy and Voigt 2010 with permission from Springer-Verlag, Berlin)

times positively phototropic. The sporangiophores are branched or unbranched, never bluish-green. Sporangia are lageniform, columellate and multispored. Specialized sporangiola and rhizoids are not formed. Sporangiospores are hyaline and smooth-walled. *Mucor circinelloides* is heterothallic, producing smooth to warty zygospores borne on opposed, naked, non-appendaged suspensors during compatible mating partners. *Mucor circinelloides* is extremely ubiquitously distributed and saprobic in soil. It has airborne, soil-borne and food-borne dispersal. *Mucor circinelloides* is an emerging opportunistic pathogen causing deep and systemic mucormycoses in immunocompromised humans. The spore size is linked to virulence, and the plasticity of the *sex* locus and adaptations in pathogenicity have occurred during speciation of the *M. circinelloides* complex (Li et al. 2011). Phylogenetic analysis supported the existence of three extant subspecies (= formae) populating the *M. circinelloides* complex in nature: *M. circinelloides* f. *circinelloides* (*Mcc*), *M. circinelloides* f. *griseocyanus* (*Mcg*) and *M. circinelloides* f. *lusitanicus* (*Mcl*). *Mcc* was found to be more prevalent among clinical *Mucor* isolates and more virulent than *Mcl* in a diabetic murine model in contrast to the wax moth host. *Mcc* (−) mating type isolates produce larger asexual sporangiospores that are more virulent in the wax moth host compared to (+) isolates that produce smaller, less-virulent sporangiospores. In addition, the larger sporangiospores germinate inside and lyse macrophages, whereas the smaller sporangiospores do not.

Mucor circinelloides exhibits hyphal growth in aerobic conditions but switches to multi-budded yeast growth under anaerobic/high-CO_2 conditions. The **calcineurin pathway** orchestrates the yeast-hyphal and spore size **dimorphic transitions** that contribute to virulence of this common zygomycete fungal pathogen (Lee et al. 2013). Yeast-locked mutants are

attenuated for virulence, illustrating that hyphae or the yeast-hyphal transition are linked to virulence. Since calcineurin governs the yeast-hyphae morphogenic transition, calcineurin inhibitors are attractive candidate compounds for antimucormycosis drugs. The Tor kinase has also received attention as an antifungal target because of its inhibition by the natural product rapamycin in *Mucor circinelloides* (Bastidas et al. 2012). The identification of rapamycin and FK506-binding proteins takes place by searching the *Mucor circinelloides* genome database (http://genome.jgi-psf.org/Mucci2/Mucci2.home.html) for genes encoding proteins containing predicted FK506-binding domains with high homology to the FK506 binding domain of *S. cerevisiae* FKBP12 (*Sc*FKBP12; encoded by *FPR1*) (Bastidas et al. 2012). The genome of *Mucor circinelloides* f. *lusitanicus* CBS277.49 (PRJNA46717) in its first draft and second assembly was finalized in March and December 2009, respectively. There were 10,930 and 11,719 genes structurally and functionally annotated in the first and second version of the draft assembly, repectively, and annotation releases have a genome size of 36.5 Mb (http://genome.jgi.doe.gov/Mucci2/Mucci2.info.html). The first version of the assembly of the genome from *M. circinelloides* f. *circinelloides* strain 1006PhL was released in April 2013 by the Broad Institute (PRJNA172437). The genome size is 36.55 Mb, and the GC content is 39.5 %. Of a total of 12,410 predicted genes, 12,227 were functionally annotated to proteins (based on information from http://www.ncbi.nlm.nih.gov/genome/?term=Mucor+circinelloides+f.+circinelloides).

b) *Rhizopus* spp.

The genus *Rhizopus* is the type genus of the family Rhizopodaceae Schaumann, which develops sporangiophores forming an umbel and each branch forming a secondary umbel that branches dichotomously, one branch terminating in a sporangium and the other in sterile spines, building the rhizoids that are characteristic for the genus and highly abundant (for review, see Voigt 2012b). The sporangiospores are airborne and variously shaped (globose to ovoid). Zygospores possess a pigmented, ornamented, rough wall, which is connected to a pair of more or less equal, opposed, nonappendaged suspensors during heterothallic and homothallic mating interactions.

Species of the genus *Rhizopus* are saprobes in soil and have importance as biotransforming and food-fermenting agents and are also clinically relevant pathogens of plants, animals and humans. They are predominantly **thermotolerant**, rhizoid-forming, multispored-sporangiate, *Mucor*-like fungi. Molecular phylogenetic analyses revealed a split into three clades, a thermophilic *Rhizopus microsporus* group (growth up to 45 °C), a subthermotolerant *Amylomyces-Rhizopus oryzae* species complex group (growth at 37–40 °C) and a mesophilic *R. stolonifer* group (growth below 37 °C; Voigt 2012b). The subthermophilic *R. oryzae* is the primary cause of mucormycosis and thus is the first object of a genome project within the Mucoromycota. The genome sequence of *R. oryzae* strain 99–880 (also termed RA 99–880), isolated from a fatal case of mucormycosis, was determined by the Broad Institute (PRJNA13066) and reported by Ma et al. (2009). This strain was isolated in 1999 from a brain abscess of a patient with diabetes who developed fatal rhinocerebral mucormycosis.

The highly repetitive 45.3-Mb genome assembly contains **abundant transposable elements**, composing approximately 20 % of the genome. Since the original protein prediction comprising 13,895 protein-coding genes, a total of 17,459 proteins were predicted (based on information of the Fungal Genome Collection of the University of Nebraska Lincoln, School of Biological Sciences and Center for Plant Science Innovation, http://em-x8.unl.edu/~canderson/FungalGenomeCollection/). The order and genomic arrangement of the duplicated gene pairs and their common phylogenetic origin provided evidence for an ancestral **whole-genome duplication (WGD)** event (Ma et al. 2009). The WGD resulted in the duplication of nearly all subunits of the protein complexes associated with respiratory electron transport chains, the vacuolar-type H+ -ATPase (V-ATPase) and

the ubiquitin-proteasome systems, as well as in the expansion of multiple-gene families related to cell growth and signal transduction, and secreted aspartic protease and subtilase protein families, which are known fungal virulence factors. The duplication of the ergosterol biosynthetic pathway, such as the lanosterol 14α-demethylase (*ERG11*) as the major azole target, could contribute to the variable responses of *R. oryzae* to different azole drugs, including voriconazole and posaconazole reported by Vitale et al. (2012). Expanded families of cell wall synthesis enzymes, essential for fungal cell integrity but absent in mammalian hosts, reveal potential targets for novel and *R. oryzae*–specific diagnostic and therapeutic treatments (Ma et al. 2009).

The V-ATPase acidifies intracellular organelles by pumping protons across the plasma membranes and also couples energy of ATP hydrolysis to proton transport across intracellular and plasma membranes of eukaryotic cells.

Rhizopus oryzae comprises a morphologically and physiologically heterogeneous species complex encompassing the sibling species *R. oryzae sensu stricto* and *R. delemar*. Phylogenetic analyses of the Internal Transcribed Spacer (ITS) region of the nuclear ribosomal DNA (rDNA), lactate dehydrogenase B, actin, translation elongation factor-1α and genome-wide Amplified Fragment-Length Polymorphisms (AFLPs) resolved the same two exclusive clusters, corresponding to the production of organic acid as effective taxonomic character grouping the species complex into two sibling species, *R. oryzae sensu stricto* (also known as *R. arrhizus*) and *R. delemar,* correlating with the lactic acid and fumaric–malic acid producers, respectively (Abe et al. 2007). Reclassification of strains in the fumaric–malic acid group as *R. delemar*, and therefore reclassification of the genome strain *R. oryzae* 99–880 into *R. delemar,* was proposed (Gryganskyi et al. 2010). More genome projects on *R. oryzae* were initiated on strains 99–892 (PRJNA186020), B7407 (PRJNA184879), HUMC 02 (PRJNA186018), NRRL 13440 (PRJNA186013) and NRRL 21396 (PRJNA186015), the last two both type 1. Besides the genome projects on the *R. delemar–R. oryzae* species complex representing the subthermophilic *Rhizopus* group, further genome projects were initiated on two strains of *R. microsporus* (ATCC52813 and CCTCCM201021) and *Rhizopus stolonifer* strain B9770 (PRJNA184886) representing the thermophilic and the mesophilic group, respectively (classification in accordance to Voigt 2012b). *Rhizopus microsporus* is traditionally used for brewing alcoholic beverages and fermented foods in China (Wang et al. 2013) but is also an important causative agent of mucormycosis (de Hoog et al. 2000). The 45,666,236-bp (ca. 45.7 Mb) draft genome sequence of *R. microsporus* var. *chinensis* CCTCCM201021 (PRJNA179339), isolated from the leaven Daqu, was determined, annotated and released in February 2013 (Wang et al. 2013). The genome of *R. microsporus* var. *microsporus* ATCC52813 (PRJNA205957) was released in August 2013 as part of the 1KFG project. With the exception of the genome from *R. delemar* strain 99–880, all other genomes of *Rhizopus* spp. are incomplete. Genome sizes and GC contents always range around 45 Mb and 36.9 %, respectively.

c) *Lichtheimia* spp.

The genus *Lichtheimia* is the type genus of the family Lichtheimiaceae Kerst. Hoffm., Walther and K. Voigt 2009 (Hoffmann et al. 2009). Colonies of *Lichtheimia* rapidly grow mainly at optimum temperatures between 37 °C and 42 °C and are predominantly **thermotolerant** until a maximum of 55 °C. The hyphae are coenocytic and often form stolons, rhizoids and terminal sporangia.

The sporangiophores are erect or slightly curved. Sporangia are, as typical for any other Mucorales, collumellate, spherical or subpyriform. Giant cells are abundant, pleomorphic with finger-like projections and characteristic for the genus. The genus is heterothallic and requires two compatible mating types to form zygospores, which are globose to oval, thick walled, naked, dark brown, with smooth and less-ornamented epispore. The gametangia are equal in size, non-appendaged, occasionally developing equatorial rings surrounding the zygospore borne on opposed suspensors.

Lichtheimia species not only are ubiquitously distributed, saprobic decomposers in soil and decaying organic matter (e.g. compost, silage, fermented food, etc.) but also are important causative opportunistic agents of **deep-tissue mycoses** in animals and humans. They represent the second and third most common cause of mucormycosis in Europe and worldwide, respectively (Roden et al. 2005; Skiada et al. 2011; Lanternier et al. 2012). The genus *Lichtheimia* encompasses six thermotolerant species, which are *L. corymbifera*, *L. ramosa*, *L. ornata*, *L. hyalospora*, *L. sphaerocystis* (Alastruey-Izquierdo et al. 2010) and *L. brasiliensis* (Santiago et al. 2013). The first three are known to be clinically relevant (Alvarez et al. 2009). Pulmonary *Lichtheimia* infections following solid-organ transplantation seem to be associated with a higher risk to develop disseminated disease (Sun et al. 2009). To gain insight into the genomic differences between these groups of pathogens, the type strain of *L. corymbifera* (JMRC:FSU:9682 > CBS 429.75 > ATCC 46771) was sequenced and compared to published genomes of other fungi composing several phyla. Comparative genomics studies on *Lichtheimia* revealed a large evolutionary distance to all fungal genomes published until now, providing evidence for the hypothesis that *L. corymbifera* has independently evolved its ability to infect humans by developing specific pathogenesis mechanisms (unpublished data). Whether these differences are reflected in the genome may reveal the analysis of the genome of the thermotolerant but not human pathogenic species *L. hyalospora*, whose first-assembly version was released by DOE-JGI and implemented in the 1KFG project in July 2013 (http://genome.jgi.doe.gov/Lichy1/Lichy1.home.html).

d) *Phycomyces blakesleeanus*

The genus *Phycomyces* belongs to the family of the Phycomycetaceae, which produce mycelial thalli with large macroscopically visible, robust sporangiophores developing multispored sporangia (Voigt 2012b). Zygospores are formed during heterothallic mating interactions on coiled, opposed or tong-like suspensors bearing branched appendages. The sporangiophores are large, unbranched, usually bluish-green, positively phototropic, and negatively gravitropic. *Phycomyces* is saprobic in soil or in association with herbivore dung and cosmopolitan in preferably the temperate climatic zones. *Phycomyces blakesleeanus* is sensitive to light, and most research on *Phycomyces* species has been directed towards understanding **phototropic growth** (Popescu et al. 1989) and growth away from stationary objects, a phenomenon known as the *avoidance response* (Johnson and Garnow 1971). The genome (53.9 Mbp in size) of *P. blakesleeanus* NRRL1555 (mating type –, PRJNA61391) was finished and released by DOE-JGI in September 2006 (first version) and March 2010 (second version). The latest version contains a total of 16,528 predicted proteins (Fungal Genome Collection of the University of Nebraska Lincoln, School of Biological Sciences and Center for Plant Science Innovation at http://em-x8.unl.edu/~canderson/FungalGenomeCollection/). Besides NRRL1555 (UBC2[−]), a wide variety of strains, 20 in total (UBC21[+], A893, A905, A909, B2, B16, C6, C21, C47, C68, c107[+], C109, C110, C149, C307, L151, L153, L157, L161, L163), are currently sequenced at the DOE-JGI (based on information from GOLD, http://www.genomesonline.org).

e) Further Mucorales Genome Initiatives

Within the 1KFG, the genomes of *Backusella circina* FSU 941 (genome size 48.65 Mb; http://genome.jgi.doe.gov/Bacci1/Bacci1.info.html) and *Umbelopsis ramanniana* (genome 23.08 Mb; http://genome.jgi-psf.org/Umbra1/Umbra1.home.html) were released in July and September 2013, respectively. Based on information from GOLD (http://www.genomesonline.org), a variety of genome projects on *Rhizomucor* spp. (*R. miehei* CBS182.67 [Gi17195] and *R. pusillus* CBS183.67 [Gi17196]) were initiated by the McGill University and Genome Quebec Innovation Centre (MUGQIC). Both species play important roles in food fermentation as opportunistic human pathogens (de Hoog et al. 2000).

2. Mortierellomycota

The group of the Mortierellomycota (Fig. 2.5b) compose the subphylum Mortierellomycotina (Hoffmann et al. 2011). The mycelium typically develops anastomosing hyphae, which are dichotomously branching and bear stylospores.

The hyphae are sporangiferous with basally inflated sporangiophores elongating towards the sporangiophore apex. The sporangiophores are erect and coenocytic initially but irregularly septated at maturity. Asexual reproduction takes place via sporangia and sporangiola. In contrast to the Mucoromycota, the columella is absent in the sporangia of the Mortierellomycota. Ramifications are gracilous, primarily horizontally expanding, erecting hyphae occasionally terminating with sporangiola. Spores are globose to ellipsoid or irregular, smooth or ornamented. Rhizoids occur only occasionally. Giant cells are absent. Zygospores are rarely reported and, where formed, are naked.

Members of the group are saprobic in soil and occasionally endophytic/'ectomycorrhizal' (Voigt 2012b). One genome of *Mortierella alpina* strain ATCC 32222 (PRJNA41211) was completed and assembled using high-throughput sequencing (Wang et al. 2011). *Mortierella alpina* is an **oleaginous fungus** that produces lipids accounting for up to 50 % of its dry weight in the form of triacylglycerols (Wang et al. 2011). It is used commercially for the **production of arachidonic acid** used as a dietary lipid. The 38.38-Mb *M. alpina* genome shows a high degree of gene duplications. Approximately 50 % of its 12,796 gene models, and 60 % of genes in the predicted lipogenesis pathway, belong to multigene families. Phylogenetic analysis based on genes involved in lipid metabolism suggests that oleaginous fungi may have acquired their lipogenic capacity during evolution after the divergence of Ascomycota, Basidiomycota, Chytridiomycota and Mucoromycota. Also, the first version of the assembly from the genome of *Mortierella elongata* (genome size 49.96 Mb) as implemented within the 1KFG project is available from the DOE-JGI's online facilities (http://genome.jgi-psf.org/Morel1/Morel1.info.html). Furthermore, the genome of *Mortierella verticillata* NRRL 6337 (PRJNA20603) is under way, provided by the OMP of the Broad Institute. The project is still incomplete.

3. Entomophthoromycota

Members of the Entomophthoromycota (Humber 2012) constitute a remarkable group that represents the **most basal lineage among the Zygomycota** (Fig. 2.1). The phylum contains a single order, the Entomophthorales, which splits into six families, among those the Basidiobolaceae and the Entomophthoraceae are the most important (Voigt 2012b). The **Basidiobolaceae** are **pathogenic to amphibians and reptiles**, whereas the **Entomophthoraceae** are entomopathogenic and of great importance as **biocontrol agents against insect pests**. Some of these are so highly adapted to their insect hosts that their lifestyle obligately relies on the close relation to the host insect throughout the entire fungal ontogeny, making a fungal cultivation in axenic cultures impossible. Currently, five genome projects have been initiated targeting entomophthoralean fungi, which are saprobic and thus cultivable in axenic cultures (Fig. 2.5c). All of them have the potential as generalist pathogens to facultatively infect and parasitize animal hosts over a broad host range. Four whole-genome sequencing projects are implemented within the 1KFG project and are still in the incomplete status. These are *Basidiobolus meristosporus* (Gi22570), *Conidiobolus coronatus* strain NRRL 28638 (Gi08779), *Conidiobolus thromboides* (Gi22553) and *Zoophthora radicans* ARSEF 4784 (Gi22538) (based on information from JGI's GOLD, http://www.genomesonline.org). Furthermore, a genome project on C. *coronatus* JMRC:FSU:4392 (Fig. 2.5c, unpublished) has been completed at the Jena Microbial Resource Collection (JMRC) in collaboration with the Leibniz Institute for Age Research–Fritz Lipmann Institute (FLI), the Leibniz Institute for Natural Product Research and Infection Biology–Hans Knöll Institute (HKI) and the Max Planck Institute for Chemical Ecology, all three institutions localized in Jena (Germany). The data are undergoing extensive bioinformatic analyses.

4. Glomeromycota

The **colonization of land by plants** relied on fundamental biological innovations, among them **symbiosis with fungi** to enhance nutrient uptake (Simon et al. 1993; Bidartondo et al. 2011). Since 1975, it has been assumed, largely from the observation that vascular plant fossils of the early Devonian (400 Ma; Remy et al. 1994) and Ordovician (485 to 443 Ma; Redecker et al. 2000) show arbuscule-like structures, that fungi of the Glomeromycota were the earliest to form **mycorrhizas** (Pirozynski and Malloch 1975). Evolutionary trees have, until now, placed Glomeromycota as the oldest known lineage of endomycorrhizal fungi (Heckman et al. 2001; Smith and Read 2008; Wang et al. 2010). ***Endogone*-like fungi**, currently classified among the subphylum Mucoromycotina (Hibbett et al. 2007), are widely associated with the earliest branching land plants and give way to glomeromycotan fungi in later lineages, which provides evidence that members of the Mucoromycota rather than the Glomeromycota enabled the establishment and growth of early land colonists (Bidartondo et al. 2011). Of a total of 51 genome projects, only a single project targets the arbusclar endomycorrhizal fungi, which are represented by ***Rhizophagus irregularis* (*Glomus intraradicis*)**. Also, this project is still incomplete (Table 2.1). The first version of the assembly and annotation of the genome of *Rh. irregularis* strain DAOM181602 was released in July 2013, implemented among the 1KFG (http://genome.jgi.doe.gov/Gloin1/Gloin1.home.html). The production of a completely annotated and assembled *Rh. irregularis* genome has proven to be an especially arduous challenge and is the result of a 9-year effort of hard work (Martin et al. 2008). The problem of accurate genome sequencing of arbuscular mycorrhizal (AM) fungi arises from the endosymbiontic fungi, which localize within the chlamydospores and thus 'contaminate' the genome of the AM fungus (Hijri et al. 2002). Furthermore, the Glomeromycota reproduce by spores containing hundreds of nuclei, which is unique for the AM fungi and represents a departure from the typical eukaryotic developmental pattern, in which a multicellular organism is re-created from a uninucleate propagule (Jany and Pawlowska 2010). The AM fungal-specific developmental pattern of sporogenesis was postulated to have adaptive significance for moderating the accumulation of deleterious mutations and to contribute to the evolutionary longevity of Glomeromycota. This multinuclearity of the AM fungal spores, which are often visible by naked eye, leads to polymorphisms that impede the assembly of the genome.

C. Miscellaneous Microorganisms Sharing a Common Ancestor with the Fungi

Among the fungal allies, the Cryptomycota are represented by 2 and the microsporida by 32 genome projects (Table 2.1).

1. Cryptomycota (Rozellida): *Rozella allomycis*

The Cryptomycota refer to 'hidden fungi' summarizing all the opisthokont flagellate, primarily heterotrophic protists, which were formerly classified to the Rozellida (Lara et al. 2010; Jones et al. 2011a). They differ from the classical fungi *sensu stricto* in that they lack chitinous cell walls in the trophic stage, as reported by Jones et al. (2011a). Because of their phagotrophic lifestyle and lack of chitin in major key phases of their life cycle, the classification of this group to or apart from the fungi is still subject of an extensive scientific debate. Despite their unconventional feeding habits, chitin was observed in the inner layer of resting spores and in immature resting spores for some species of *Rozella*, as indicated with calcofluor-white staining, as well as the presence of a fungal-specific **chitin synthase** (CHS) gene (James and Berbee 2011). However, genes encoding fungal-like CHSs are abundant in members of many non-fungal lineages, such as amoebozoans (*Entamoeba*), trichomonads (*Trichomonas* and *Tritrichomonas*) and diatoms (*Thalassiosira*), stramenopiles and chromalveolates (Campos-Góngora et al. 2004; Kneipp et al. 1998; Durkin et al. 2009; Greco et al. 1990).

The diatom *Thalassiosira pseudonana* possesses six genes encoding three types of CHS (Durkin et al. 2009). Transcript abundance increases when cells resume division after short- and long-term silicic acid starvation and under short-term silicic acid or iron-depletive conditions, which are connected in the environment and known to affect the cell wall. In *Entamoeba*, a pair of genes encoding CHS was identified in each of three single species, *E. histolytica*, *E. dispar* and the reptilian parasite *E. invadens* (Campos-Góngora et al. 2004). Both CHS-1 and CHS-2 are not expressed in *Entamoeba* trophozoites (which are the vegetative cells), but substantial amounts of CHS-1 and CHS-2 RNA were present 4–8 h after induction of cyst formation by glucose deprivation of *E. invadens*, the common model for cyst induction. The expression of chitin as a structural component of *Trichomonas vaginalis* and *Tritrichomonas foetus* was demonstrated by Kneipp et al. (1998). The distribution, synthesis and deposition of chitin in protistan organisms was reviewed by Mulisch (1993).

Cell-wall associated chitin is abundant all across the eukaryotic domain and was reported to be the major structural component of cyst walls from protists not having any fungal affiliation, such as *Giardia* (Diplomonada-Zoo/Sarcomastigophora); *Nephromyces* (classified into Apicomplexa-Chromalveolata by Saffo et al. 2010); *Entamoeba* (Amoebozoa); *Trichomonas* (Trichomonadida-Parabasalia); *Poteriochromonas* (syn. *Ochromonas*, Chrysophyta-Chromalveolata); *Thalassiosira* (Heterokontophyta-Chromalveolata) and *Euplotes* (Ciliata-Chromalveolata) (Ward et al. 1985; Saffo and Fultz 1986; Arroyo-Begovich and Crabez-Trejo 1982; Campos-Góngora et al. 2004; Herth et al. 1977; Brunner et al. 2009; Greco et al. 1990, respectively).

The differences between protists and fungi are that fungi develop cells that are composed of chitin in major key phases during their life cycle, while protists lack cell walls, which are exceptionally formed during encystation. Also, in diatoms (Heterokontophyta), the primary composer of ocean plankton, networks of chitin filaments are integral components of diatom silica shells (Brunner et al. 2009; Durkin et al. 2009).

The potential for chitin production is encoded in the genomes of many eukaryotes (protists, animals and fungi) but seems often not to be expressed during the entire life cycle. Similarly, the presence of genes encoding for enzymes such as cellulases (Zhao et al. 2013) or enzymes participating in secondary metabolism, such as non-ribosomal polyketide synthases (Schroeckh et al. 2009; Brakhage and Schroeckh 2011), were reported to be silent in the genomes of the fungi that are not known to have any enzymatic property encoded by the silent genes (for review, see Brakhage and Schroeckh 2011).

During the past, the Cryptomycota were ambiguously classified to the fungi (Jones et al. 2011a; James and Berbee 2011) or as a sister group to the fungi (Lara et al. 2010). The term was validly described in accordance with the ICNAFP (Jones et al. 2011b). Recently, a cultured aphelid species, *Amoeboaphelidium protococcarum*, was shown to form a monophyletic group with *Rozella* and microsporidia as a sister group to the Fungi and uses, unlike fungi, a non-canonical nuclear genetic code, which provides evidence for *Rozella*'s non-fungal affiliation (Karpov et al. 2013). Aphelids resemble primitive opisthokont protists that produce zoospores and possess, unlike fungi, a phagotrophic mode of nutrition uptake, which qualifies them as descendents of the common ancestor of fungi and multicellular animals right at the divergence point of the fungi from the animal kingdom.

Currently, two genome projects are under way targeting *R. allomycis* strain CSF55, an obligate parasite of zoosporic fungi (here: *Allomyces* spp.). These projects are (based on information from GOLD, www.genomesonline.org): (a) Gi08798 implemented within 1KFG and (b) Gi0045129, a project driven by the University of Michigan, both using a combination of Illumina HiSeq 2000 and 454 next-generation sequencing methods.

2. Microsporidia

The Microsporidia are intracellular parasites that lack mitochondria. Their ribosomes are unusual in (a) being of prokaryotic size (70S) and (b) lacking characteristic eukaryotic nuclear 5.8S rRNA as a separate molecule but which is incorporated into the nuclear 28S rRNA. Comparison of rRNA genes (Vossbrinck et al. 1987) and single-protein phylogenies

(Kamaishi et al. 1996) suggests that microsporidia are ancient eukaryotes. The scientific community is divided into two groups: one arguing for the fungal classification of the microsporidia and one against its fungal alliance (for review, see Voigt and Kirk 2011; pro: Bohne et al. 2000; Fast and Keeling 2001; Keeling and Slamovits 2004; Capella-Gutiérrez et al. 2012; and against: Barron 1987; Patterson and Zölffel 1991; Hirt et al. 1997; Hashimoto et al. 1998; Peyretaillade et al. 1998a, b; Patterson 1999; Delbac et al. 2001; Bacchi et al. 2002; Glockling and Beakes 2002; Tanabe et al. 2002). The new code organizing the description of fungi, the ICNAFP (Melbourne Code; McNeill et al. 2012), excludes the microsporidia. Therefore, new microsporidian species are described by the ICZN (Ride et al. 2000). Currently, 32 genome projects from the Microsporidia are known: 2 are finalized (PRJNA13833/PRJNA155 regarding *Encephalitozoon cuniculi* [Ec] strain GB-M1, and PRJNA42703/PRJNA51607 regarding *Encephalitozoon intestinalis* [Ei] strain ATCC 50506; Katinka et al. 2001; Corradi et al. 2010, respectively); 2 in permanent draft status (PRJNA48321/PRJNA32971 regarding *Nosema ceranae* [Nc] strain BRL01 and PRJNA51843/PRJNA61123 regarding *Nematocida parisii* [Np] strains ERTm1 and ERTm3; Cornman et al. 2009; Cuomo et al. 2012, respectively) and 28 reported as incomplete (Tables 2.1 and 2.2). The genomes of the microsporidia are highly reduced, ranging from 2.22 Mb (Ei), 2.5 Mb (Ec), 4.07 Mb (NpERTm1), 4.15 Mb (NpERTm3) to 7.86 Mb (Nc) in size, containing 1,892 (Ei) 2,029 (Ec), 2,678 (Nc), 2,724 (NpERTm1) and 2,788 (NpERTm3) predicted genes (Katinka et al. 2001; Cornman et al. 2009; Corradi et al. 2010; Cuomo et al. 2012). The draft assembly (7.86 MB) of the Nc genome derived from pyrosequencing data has a strong AT bias (74.7 % A+T, 25.3 % G+C), unlike other microsporidian genomes (41.5 % G+C [Ei], 47.3 % G+C [Ec], 34.5 % G+C [NpERTm3]), and diverse repetitive elements, complicating the assembly (Cornman et al. 2009). Of the 2,614 predicted protein-coding sequences, 1,366 were conservatively estimated to have homologs in the microsporidian Ei (PRJNA13833; Katinka et al. 2001), the most closely related published genome sequence. Microsporidia have a surprisingly wide spectrum of G+C content, genome size and a high variability in genome architecture (e.g. gene syntenies), rendering these microorganisms an exceptional and rapidly evolving group of obligate intracellular parasites that exemplifies the contradictory forces of simplification and complex specialization borne on the adaptation to the parasitic lifestyle as reductive element (Keeling and Slamovits 2004).

3. Nucleariida-Fonticula Group: *Nuclearia simplex* and *Fonticula alba*

The **nucleariids** are a group of amoebae that develop **filose pseudopodia** (from the Greek word ψευδοπόδια, ψευδ 'fake, false'+πόδια 'feet'), which are temporary projections of membranes from eukaryotic cells or unicellular protists (Zettler et al. 2001; Steenkamp et al. 2006; Shalchian-Tabrizi et al. 2008). They are abundant in soils and freshwater. They are distinguished from the morphologically similar vampyrellids (filose-typed Cercozoa: subphylum Filosa) mainly by having discoid cristae-typed mitochondria. ***Fonticula*** is a cellular slime mould that forms a fruiting body in a volcano shape and is a sister genus of the *Nuclearia*, both related to fungi (Brown et al. 2009). *Fonticula*, *Nuclearia*, and Fungi have been united into the Nucletmycea, which is sister to the Holozoa (for review, see Voigt et al. 2013). Sequencing of the genome of *Fonticula alba* strain ATCC 38817 (PRJNA189482, 31.3-Mb genome size and 67.9 % G+C content) is currently in progress, included in the OMP.

4. Mesomycetozoa (Ichthyosporea)

The **DRIP clade**, which is an acronym for the original members: ***Dermocystidium***, **rosette agent**, ***Ichthyophonus***, and ***Psorospermium***, is referred to the class Ichthyosporea, and more recently to the class Mesomycetozoa; it is a heterogeneous group, mostly of parasites of fish and other animals at the boundary of animals and fungi (for review, see Mendoza et al. 2002; Ragan et al. 2003). The Mesomycetozoa seem to be the most masal lineage within the

Holomycota (also: Nucletmycea), which is composed of the fungi, Cryptomycota (Rozellida), Microsporidia, Nucleariida and Mesomycetozoa, summarized as sister group to the **Holozoa** (for review, see Voigt et al. 2013). Two orders were described in the mesomycetozoans: the **Dermocystida** and the **Ichthyophonida**, which form amoeba-like cells and flagellate cells, respectively. So far, all members in the order Dermocystida are pathogens either of fish (*Dermocystidium* spp. and the rosette agent) or of mammals and birds (*Rhinosporidium seeberi*), and most produce posteriorly uniflagellate (opisthokont) zoospores. Fish pathogens also are found in the order Ichthyophonida, but so are saprotrophic microbes. The Ichthyophonida species do not produce flagellate cells, but many produce amoeba-like cells. Descriptions of the genera of the class Mesomycetozoa and their morphological features, pathogenic roles, and phylogenetic relationships were reviewed by Mendoza et al. (2002).

Three genome projects are in progress, all embedded in the OMP as carried out by the Broad Institute: (a) *Amoebidium parasiticum* strain JAP-7–2 (PRJNA189476); (b), *Sphaeroforma arctica* strain JP610 (PRJNA20463) and (c) *Capsaspora owczarzaki* strain ATCC 30864 (PRJNA193613, PRJNA20341).

Amoebidium parasiticum is a unicellular microorganism found on the external cuticule of Crustacea, Amphipods, Isopods, Copepods and Insects. Contrary to its name, this fungus has not been reported to be parasitic. This organism is of high evolutionary interest because, together with other unicellular organisms such as *Sphaeroforma arctica*, it diverged prior to Metazoa and prior to the Choanoflagallata (Lang et al. 2002). *Sphaeroforma arctica* is a unicellular protist originally isolated from an arctic marine amphipod. Both single-cell organisms, *A. parasiticum* and *S arctica*, account for a pool of ancient genes that might be responsible for early development in multicellular animals. Therefore, sequencing of these unicellular organisms originating right at the divergence point of animals and fungi may reveal the origins of genes that are crucial for multicellularity.

Capsaspora owczarzaki is an amoeboid symbiont of a pulmonate snail and was subsequently found not to be a nucleariid or an ichthyosporean as previously described, but is more closely related to the Choanoflagellata as elucidated by molecular phylogenetic analyses of the protein-coding genes of the elongation factor 1-alpha (EF1-alpha) and the heat shock protein 70 (Hsp70) (Ruiz-Trillo et al. 2006).

III. The Evolution of Basal Lineage Fungi: A Phylogenomic Perspective

In the past, the delimitation of the fungi was based on various aspects: classical approaches to identify fungi ranging from comparisons with the fossil record (Hawksworth et al. 1995); the use of physiological (Pitt 1979) and biochemical markers (Paterson and Bridge 1994); the composition of the cell wall (Bartnicki-Garcia 1970, 1987) and isoenzyme patterns (Maxson and Maxson 1990); the existence of pigments (Besl and Bresinsky 1997); secondary metabolite profiles (Frisvad and Filtenborg 1990) to observations on the ultrastructure (James et al. 2006b).

The next, and more difficult, step will be to develop community standards for sequence-based classification, which facilitate naming of species known only from environmental sequences into a new taxonomic category (comparable to the *Candidatus* status for bacteria and archae) or an identifying suffix. This would be the ENAS (environmental nucleic acid sequence) or eMOTU (environmental molecular operational taxonomic unit; for review, see Hibbett and Taylor 2013).

Genomic approaches utilize comparative genomics aimed at the elucidation of biosynthetic gene clusters and other gene syntenies known to be typical for fungi, such as chitin synthesis (James and Berbee 2011), or whole-genome analyses resulting in the reconstruction of phylomes, which represent phylogenetic networks based on single trees (Capella-Gutiérrez et al. 2012, see also Chap. 3 in this volume).

The **resolution of the deep fungal branches** was solely possible by the reconstruction of a robust backbone of the fungal phylogenies using a multitude of genes, which were strongly selected for orthology applying novel bioinformatic tools (Ebersberger et al. 2009b). That bioinformatic approach, called HaMStR (Hidden Markov Model Based Ortholog Search

Tool Using Reciprocity), combines a profile hidden Markov model search with subsequent BLAST (Basic Local Alignment Search Tool) search to extend existing ortholog clusters with sequences from additional taxa to mine EST (Expressed Sequence Tag) and genomic data for the presence of orthologs to a curated set of genes (Ebersberger et al. 2009b). Phylogenomic studies based on more than 100 orthologous protein-coding genes provided evidence for the **reunification of Blastocladiomycota, Chytridiomycota, *s.str.*, Neocallimastigomycota, and the Monoblepharidomycota** into one phylogenetically coherent group named the Chytridiomycota *s.l.* for zoosporic true fungi as a whole. These fungi **share the production of zoospores with posteriorly directed whiplash flagella** and mitochondria with flattened cristae as a common morphological feature with other opisthokonts as common synapomorphies (James et al. 2006b; Ebersberger et al. 2009a, 2012). Despite the success of the novel approaches, the phylogeny of zoosporic true fungi still requires improvement in phylogenetic resolution because of an increase in taxon sampling completion. The increase of genome projects targeting the basal fungal lineages is the first step towards clarification of the phylogenetic relationships right at the divergence nodes, which are still questionable up to date, at a genome-wide scale.

IV. Genomic Features Characteristic for Basal Fungi

A. Whole-Genome Duplication

Because not many basal fungal species have been fully sequenced, not much is known about the peculiarities of their genomic structures. However, even having at hand the limited number of examples, we can discuss some interesting features characteristic for these genomes.

The first characterized genome of the Mucorales was *Rhizopus oryzae*, and the most interesting feature regarding the genome structure was its **whole-genome duplication** (Ma et al. 2009). In this work, the evidence for the WGD was based on the comparison of the duplicated (and non-duplicated) orthologs of *R. oryzae* and *P. blakesleeanus* and on the genomic distribution of the duplicated genes (**segmental duplications**) in the *Rhizopus* genome. This discovery caused increased interest in the possibility of the WGD in the subsequently sequenced genomes, so by the time of the analysis of other representatives of the basal fungi (e.g. *Conidiobolus* and *Lichheimia*), it was tempting to assign the observed duplications to the same mechanism. However, no evidence for a recent WGD could be revealed by the thorough analysis of the corresponding genomes (personal observations, unpublished). In particular, WGD presumes a large (or at least detectable) number of repeated syntenic regions (i.e. regions with the same order of genes) of reasonable size. In *R. oryzae*, the number of segmental duplications covers more than 10 % of the genome (Ma et al. 2009), whereas in other mucoralean genomes this number is comparable with that of the derived fungi (e.g., ascomycetes, with exception of Saccharomycotina). For example, the duplicated regions in *L. corymbifera* cover less than 4 % of its genome (personal observation, unpublished). These observations confirm the conclusions made by Ma et al., who claimed the **WGD as being specific only for *R. oryzae***, not affecting the other mucoralean genomes. The origin of the duplicated genes in mucoralean genomes (besides *Rhizopus*) may be a much more ancient genome duplication, as suggested by Marcet-Houben et al. (2009), which is no longer detectable as regions' duplications. Although some of the duplicated regions can be assumed to be remains of this ancient WGD, we cannot assign all gene duplications to this process. What could be the origin of the gene expansions that we observe in significant amounts in the basal fungal genomes?

B. Gene Duplication, Tandem Duplication

With respect to the basal lineage fungi, large amounts of gene expansions have been reported targeting the proteases. Comparative analyses of the genomes of the amphibian-pathogen Bd and its closest known relative, the non-pathogenic chytrid Hp, provided evi-

dence for a significant lineage-specific expansion of three protease families (metallo-, serine-type, and aspartyl proteases in Bd), which occurred after the divergence of Bd and Hp from their common ancestor and thus are unique for the Bd branch (Joneson et al. 2011). These expansions in protein families count as virulence factors providing the possibility for Bd to become a successful pathogen. The timing of the protease gene family expansions pre-dates the emergence of Bd as a globally important amphibian pathogen supporting the hypothesis (proposed by Joneson et al. 2011) that protease genes are important in the evolutionary transition from non-pathogenic into pathogenic in this group.

Analysing several genomes of basic fungi (e.g. *L. corymbifera*, *C. coronatus*), we noticed a large number of **tandem duplications**, which in a way contradicts the idea of the WGD. Tandem duplications per se cannot be called a specific feature of the basal fungi, as their number is on average not significantly higher as in, for example, asco- or basidiomycete genomes.

The only exception is *C. coronatus*, for which a pilot study revealed the highest so far detected number of protein-coding genes found in tandem duplications (~180 per 1,000 proteins), whereas for other fungi in this study this number does not exceed 69.

But, remarkably, we see a strong overlap of tandems with gene expansions, which suggests that the **majority of expanded gene families could occur through the mechanism of tandem duplications**.

Expansions of particular gene families are lineage specific. For instance, *C. coronatus* is characterized by expansions of the heat shock factors (HSFs), several transcription factor (TF) families and some families of secreted proteases. Remarkably, 24 % of all TFs occur as tandem duplications; the same is true for 60 % of HSFs. All four families of secreted proteases, which are expanded in *C. coronatus* from 42 % (S01) to 65 % (S08), are found in tandem duplications (Table 2.3). Moreover, for some examples the 'excess' of genes can be explained exclusively by tandems: for example, for S08 proteases, if we subtract the number of genes in tandem repeats (24) from the total number of genes in the family (37), we obtain the number (13) that equals the average number of genes in Ascomycetes (11.7±4.6) and in *L. corymbifera* (14), where this family is not expanded.

Table 2.3. Examples of gene families of secreted proteases, expanded in the entomophthoralean fungus *Conidiobolus coronatus* JMRC:FSU:4392 (unpublished data)

Protease family	Number	of them in groups (Number of groups)
M14A	16	10 (3)
M36	9	4 (1)
S01	26	11 (3)
S08	37	24 (5)

Gene expansions can be associated with tandem duplications also in other genomes of basal fungi. In *L. corymbifera*, 42 of 66 gene expansion groups contain tandem duplications, covering 38 % of all genes in the expanded groups. Interestingly, the number of tandem duplications in *R. oryzae* is also quite high (69 per 1,000 proteins, which puts *R. oryzae* in second place after *C. coronatus*). It remains questionable whether these duplications occurred before or after the WGD; in the former case, the tandems would be doubled by the WGD, which might explain this high number.

These findings suggest that, in basal fungi, the mechanism of tandem duplications may be the main force driving the adaptation process, being responsible for gene expansions. As we mentioned in the beginning of this chapter, genomic studies of the basal fungal lineages are in their infancy because of the lack of the fully sequenced genomes until recently. However, the sequencing projects are speeding up, so we can expect more material for comparative studies of these important fungi.

VI. Conclusions

Of a total of about 100,000 fungal species described so far, approximately 5,000 (ca. 6 %) are basal true fungi. Estimates forecast a much higher number of up to 1.5–3.5 million fungal species (Hawksworth 2001), of which basal fungi and miscellaneous fungus-like

organisms may compose the majority. Therefore, the basal lineages of the fungi represent a highly underestimated group of fungi that has been less studied during the past but deserves higher attention in all respects. They significantly contribute to the dynamics of the interaction among all domains of life and provide a pool for the properties of phylogenetically derived fungal lineages. If we take the evolutionary history of the basal fungi into account, we can deduce trends in evolution towards a deep and comprehensive understanding of mechanisms behind adaptation and selection.

Currently, several genome initiatives are under way focusing on different biological questions, such as exploration for possible genetic engineering of fungi to produce higher levels and diverse contents of secondary metabolites, carotenes, dietary lipids, enzymes used for food fermentation or biotransformation of steroids, and so on, as well as facilitate the rational design of antifungal drugs.

Large-scale genome projects provide insight into the evolutionary history of the fungi and their legacy in the evolution of multicellularity. The genomes show a considerable amount of apparent genetic redundancy, and one of the major problems to be tackled during the next stage of completed fungal genome projects is to elucidate the biological functions of the genes encoded in each of the genomes. For example, genome-wide identifications and comparisons of fungal TFs or carbohydrate activity enzymes (CAZymes) can help to gain a better understanding of their regulatory mechanisms and the functionality of the encoding genes and their encoded proteins (Shelest 2008; Zhao et al. 2013). In addition, the large diversity of nutritional modes can be compared with infection mechanisms in terms of adaptation and stress response at the -omics level.

Acknowledgements We wish to express our gratitude to Kerstin Hoffmann and Martin Eckart (University of Jena, Germany) and Katerina Fliegerova (Czech Academy of Sciences, Prague, Czech Republic) for contributing microscopic images shown in Fig. 2.1a and to Paul M. Kirk (Royal Botanic Gardens Kew, UK) for providing the numbers of fungal species based on the Index Fungorum database (http://www.indexfungorum.org) (as of September 16, 2013) shown in Fig. 2.2. We thank Volker U. Schwartze (University of Jena, Germany) for technical assistance with drawing Figs. 2.2 and 2.6. This work was financially supported by the Leibniz Association and the University of Jena.

References

Abe A, Oda Y, Asano K, Sone T (2007) *Rhizopus delemar* is the proper name for *Rhizopus oryzae* fumaric-malic acid producers. Mycologia 99:714–722

Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JA, Bowser SS, Bragerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup Ø, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MFJR (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

Alastruey-Izquierdo A, Hoffmann K, De Hoog GS, Rodriguez-Tudela JL, Voigt K et al (2010) Species recognition and clinical relevance of the zygomycetous genus Lichtheimia (syn. Absidia pro parte, Mycocladus). J Clin Microbiol 48:2154–2170

Alexopoulos CJ, Mims CW, Blackwell M (1996) Introductory mycology, 4th edn. Wiley, New York

Alvarez E, Sutton DA, Cano J, Fothergill AW, Stchigel A, Rinaldi MG, Guarro J (2009) Spectrum of zygomycete species identified in clinically significant specimens in the United States. J Clin Microbiol 47:1650–1656

An Z, Wang C, Liu X, Bennett JW (2010) China's fungal genomics initiative: a whitepaper. Mycology 1:1–8

Arroyo-Begovich A, Crabez-Trejo A (1982) Location of chitin in the cyst wall of *Entamoeba invadens* with colloidal gold racers. J Parasitol 68:253–258

Bacchi CJ, Weiss LM, Lane S, Frydman B, Valasinas A, Reddy V, Sun JS, Marton LJ, Khan IA, Moretto M, Yarlett N, Wittner M (2002) Novel synthetic polyamines are effective in the treatment of experimental microsporidiosis, an opportunistic AIDS-associated infection. Antimicrob Agents Chemother 46:55–61

Baldauf SL, Palmer JD (1993) Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. Proc Natl Acad Sci U S A 90:11558–11562

Barr DJS (1978) Taxonomy and phylogeny of chytrids. Biosystems 10:153–165

Barr DJS (1980) An outline for the reclassification of the Chytridiales, and for a new order, the Spizellomycetales. Can J Bot 58:2380–2394

Barr DJS (1984) Cytological variation in zoospores of *Spizellomyces* (Chytridiomycetes). Can J Bot 62:1202–1208

Barr DJS (2001) Chytridiomycota. In: McLaughlin DJ, McLaughlin EG, Lemke PA (eds) The mycota VII

Part A. Systematics and evolution. Springer, Berlin, pp 93–112
Barron GL (1987) The gun cell of *Haptoglossa mirabilis*. Mycologia 79:877–883
Bartnicki-Garcia S (1970) Cell wall composition and other biochemical markers in fungal phylogeny. In: Harbone JB (ed) Phytochemical phylogeny. Academic, London, pp 81–103
Bartnicki-Garcia S (1987) The cell wall in fungal evolution. In: Rayner ADM, Brasier CM, Moore D (eds) Evolutionary biology of the fungi. Cambridge University Press, New York, pp 389–403
Bastidas RJ, Shertz CA, Lee SC, Heitman J, Cardenas ME (2012) Rapamycin exerts antifungal activity in vitro and in vivo against *Mucor circinelloides* via FKBP12-dependent inhibition of Tor. Eukaryot Cell 11:270–281
Bennett JW (2000) Filamentous fungi. In: Encyclopedia of microbiology. Academic, San Diego, pp 468–477
Besl H, Bresinsky A (1997) Chemosystematics of Suillaceae and Gomphidiaceae (suborder Suillineae). Plant Syst Evol 206:223–242
Bidartondo MI, Read DJ, Trappe JM, Merckx V, Ligrone R, Duckett JG (2011) The dawn of symbiosis between plants and fungi. Biol Lett 7:574–577
Birren B, Fink G, Lander E (2002) The fungal genome initiative. Paper of the Whitehead Institute Center for Genome Research developed by the Fungal Research Community. http://www.broadinstitute.org/annotation/fungi/fgi/FGI_01_whitepaper_2002.pdf. Accessed 19 Oct 2010
Bohne W, Ferguson DJP, Kohler K, Gross U (2000) Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*. Infect Immun 68:2268–2275
Brakhage AA, Schroeckh V (2011) Fungal secondary metabolites—strategies to activate silent gene clusters. Fungal Genet Biol 48:15–22
Briggs CJ, Knapp R, Vrendenburg VT (2010) Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. Proc Natl Acad Sci U S A 107:9695–9700
Brown MW, Spiegel FW, Silberman JD (2009) Phylogeny of the 'forgotten' cellular slime mold, *Fonticula alba*, reveals a key evolutionary branch within Opisthokonta. Mol Biol Evol 26:2699–2709
Brunner E, Richthammer P, Ehrlich H, Paasch S, Simon P, Ueberlein S, van Pée K-H (2009) Chitin-based organic networks: an integral part of cell wall biosilica in the diatom *Thalassiosira pseudonana*. Angew Chem Int Ed Engl 48:9724–9727
Campos-Góngora E, Ebert F, Willhoeft U, Said-Fernndez S, Tannich E (2004) Characterization of chitin synthases from *Entamoeba*. Protist 155:323–330
Capella-Gutiérrez S, Marcet-Houben M, Gabaldón T (2012) Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. BMC Biol 10:47. doi:10.1186/1741-7007-10-47
Cavalier-Smith T (1981) Eukaryote kingdoms: seven or nine? BioSystems 14:461–481
Chakrabarti A, Das A, Mandal J, Shivaprakash MR, George VK, Tarai B, Rao P, Panda N, Verma SC, Sakhuja V (2006) The rising trend of invasive zygomycosis in patients with uncontrolled diabetes mellitus. Med Mycol 44:335–342
Collins JP (2013) History, novelty, and emergence of an infectious amphibian disease. Proc Natl Acad Sci U S A 110:9193–9194
Cornely OA, Vehreschild JJ, Rüping MJGT (2009) Current experience in treating invasive zygomycosis with posaconazole. Treatment options for zygomycosis. Clin Microbiol 15:77–81
Cornman RS, Chen YP, Schatz MC, Street C, Zhao Y, Desany B, Egholm M, Hutchison S, Pettis JS, Lipkin WI, Evans JD (2009) Genomic analyses of the microsporidian *Nosema ceranae*, an emergent pathogen of honey bees. PLoS Pathog 5:e1000466
Corradi N et al (2010) The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. Nat Commun 1:77
Cuomo CA, Desjardins CA, Bakowski MA, Goldberg J, Ma AT, Becnel JJ, Didier ES, Fan L, Heiman DI, Levin JZ, Young S, Zeng Q, Troemel ER (2012) Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Genome Res 22:2478–2488
Daszak P, Berger L, Cunningham AA, Hyatt AD, Green DE, Speare R (1999) Emerging infectious diseases and amphibian population declines. Emerg Infect Dis 5:735–748
De Hoog GS, Guarro J, Gene J, Figueras M (2000) Atlas of clinical fungi, 2nd edn. Centraalbureau voor Schimmelcultures, Utrecht. http://books.google.de/books?id=clv_CWVR1GoC
De Santiago AALCM, Hoffmann K, Lima DX, de Oliviera RJV, Vieira HEE, Malosso E, Maia LC, da Silva GA (2013) A new species of *Lichtheimia* (Mucoromycotina, Mucorales) isolated from Brazilian soil. Mycol Prog. doi:10.1007/s11557-113-0920-8
Delbac F, Peuvel I, Metenier G, Peyretaillade E, Vivares CP (2001) Microsporidian invasion apparatus: identification of a novel polar tube protein and evidence for clustering of *ptp1* and *ptp2* genes in three Encephalitozoon species. Infect Immun 69:1016–1024
Doweld A (2001) Prosyllabus tracheophytorum—Tentamen systematis plantarum vascularium (Tracheophyta). Goes MMI, Moscow, pp 1–111
Durkin CA, Mock T, Armbrust EV (2009) Chitin in diatoms and its association with the cell wall. Eukaryot Cell 8:1038–1050
Ebersberger I, Gube M, Strauss S, Kupczok A, Eckart M, Voigt K, Kothe E, von Haeseler A (2009a) A stable backbone for the fungi. Nat Precedings. hdl:10101/npre.2009.2901.1. Accessed 19 Sept 2012

Ebersberger I, Strauss S, von Haeseler A (2009b) HaMStR: profile hidden Markov model based. BMC Evol Biol 9:157

Ebersberger I, de Matos SR, Kupczok A, Gube M, Kothe E, Voigt K, von Haeseler A (2012) A consistent phylogenetic backbone for the fungi. Mol Biol Evol 29:1319–1334

Eckart M, Eckart M, Fliegerova K, Hoffmann K, Voigt K (2010) Molecular identification of anaerobic rumen fungi. In: Gherbawy Y, Voigt K (eds) Molecular identification of fungi: Part I. Springer, Berlin, pp 297–313

Fast NM, Keeling PJ (2001) Alpha and beta subunits of pyruvate dehydrogenase E1 from the microsporidian *Nosema locustae*: mitochondrion-derived carbon metabolism in microsporidia. Mol Biochem Parasitol 117:201–209

Fisher MC, Garner TWJ, Walker SF (2009) Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. Annu Rev Microbiol 63:291–310

Frisvad JC, Filtenborg O (1990) Secondary metabolites as consistent criteria in *Penicillium* taxonomy and a synoptic key to *Penicillium* subgenus *Penicillium*. In: Samson RA, Pitt JI (eds) Modern concepts in *Penicillium* and *Aspergillus* classification. Plenum Press, New York, pp 373–384

Gherbawy Y, Voigt K (eds) (2010) Molecular identification of fungi. Springer, Berlin, Heidelberg, New York: 1st edition, March 2010, XXI, 501 p. 87 illus., 39 in colour. Hardcover, ISBN 978-3-642-05041-1

Glockling SL, Beakes GW (2002) Ultrastructural morphogenesis of dimorphic arcuate infection (gun) cells of *Haptoglossa erumpens* an obligate parasite of *Bunonema* nematodes. Fungal Genet Biol 37:250–262

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6,000 genes. Science 274(546):563–567

Greco N, Bussers JC, Van Daele Y, Goffinet G (1990) Ultrastructural localization of chitin in the cystic wall of *Euplotes muscicola* Kahl (Ciliata, Hypotrichia). Eur J Protistol 26:75–80

Griffith GW, Baker S, Fliegerova K, Liggenstoffer A, van der Giezen M, Voigt K, Beakes G (2010) Anaerobic fungi: Neocallimastigomycota. IMA Fungus 1:181–185

Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, Kubicek CP, Kuske C, Magnuson JK, Martin F, Spatafora JW, Tsang A, Baker SE (2011) Fueling the future with fungal genomics. Mycology 3:192–209

Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA, Otillar R, Poliakov A, Ratnere I, Riley R, Smirnova T, Rokhsar D, Dubchak I (2012) The genome portal of the Department of Energy Joint Genome Research. Nucleic Acids Res 40(1):D26–D32

Gryganskyi AP, Lee SC, Litvintseva AP, Smith ME, Bonito G, Porter T, Anishchenko IM, Heitman J, Vilgalys R (2010) Structure, function, and phylogeny of the mating locus in the *Rhizopus oryzae* complex. PLoS One 5(12):e15273. doi:10.1371/journal.pone.0015273

Hashimoto T, Snchez LB, Shirakura T, Müller M, Hasegawa M (1998) Secondary absence of mitochondria in *Giardia lamblia* and *Trichomonas vaginalis* revealed by valyl-tRNA synthetase phylogeny. Proc Natl Acad Sci U S A 95:6860–6865

Hawksworth DL (2011) A new dawn for the naming of fungi: impacts of decisions made in Melbourne in July 2011 on the future publication and regulation of fungal names. MycoKeys 1:7–20

Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol Res 105:1422–1432

Hawksworth DL, Rossman AY (1997) Where are all the undescribed fungi? Phytopathology 87:888–891

Hawksworth DL, Kirk PM, Sutton BC, Pegler DN (1995) Ainsworth and Bisby's dictionary of the fungi, 8th edn. International Mycological Institute, Kew, London

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129–1133

Herth W, Kuppel A, Schnepf E (1977) Chitinous fibrils in the lorica of the flagellate chrysophyte *Poteriochromonas stipitata* (syn. *Ochromonas malhamensis*). J Cell Biol 73:311–321

Hibbett DS, Taylor JW (2013) Fungal systematics: is a new age of enlightenment at hand? Nat Rev Microbiol 11:129–133

Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson O, Huhndorf S, James T, Kirk PM, Lücking R, Lumbsch T, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer R, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai Y-C, Gams W, Geiser DM, Griffith GW, Gueidan C, Hawksworth DL, Hestmark G, Hosaka K, Humber RA, Hyde K, Köljalb U, Kurtzman CP, Larsson K-H, Lichtwardt R, Longcore J, Miadlikowska J, Miller A, Moncalvo J-M, Mozley-Standridge S, Oberwinkler F, Parmasto R, Reeb V, Rogers JD, Roux C, Ryvarden L, Sampaio JP, Schuessler A, Sugiyama J, Thorn RG, Tibell L, Untereiner WA, Walker C, Wang A, Weir A, Weiss M, White M, Winka K, Yao Y-J, Zhang N (2007) A higher-level phylogenetic classification of the Fungi. Mycol Res 111:509–547

Hibbett DS, Ohman A, Kirk PM (2009) Fungal ecology catches fire. New Phytol 184:279–282

Hijri M, Redecker D, MacDonald-Comber Petetot JA, Voigt K, Wöstemeyer J, Sanders IR (2002) Identification and isolation of two ascomycete fungi from spores of the arbuscular mycorrhizal fungus *Scutellospora castanea*. Appl Environ Microbiol 68:4567–4573

Hirt RP, Healy B, Vossbrinck CR, Canning EU, Embley TM (1997) A mitochondrial Hsp70 orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria. Curr Biol 7:995–998

Hoffmann K, Telle S, Walther G, Eckart M, Kirchmair M, Prillinger HJ, Prazenica A, Newcombe G, Dölz F, Papp T, Vgvölgyi C, de Hoog S, Olsson L, Voigt K (2009) Diversity, genotypic identification, ultrastructural and phylogenetic characterization of zygomycetes from different ecological habitats and climatic regions: limitations and utility of nuclear ribosomal DNA barcode markers. In: Gherbawy Y, Mach RL, Rai M (eds) Current advances in molecular mycology. Nova Science, New York, pp 263–312

Hoffmann K, Voigt K, Kirk PM (2011) *Mortierellomycotina* subphyl. nov. based on multi-gene genealogies. Mycotaxon 115:353–363

Humber RA (2012) Entomophthoromycota: a new phylum and reclassification for entomophthoroid fungi. Mycotaxon 120:477–492

James TY, Berbee ML (2011) No jacket required—new fungal lineage defies dress code: recently described zoosporic fungi lack a cell wall during trophic phase. Bioessays 34:94–102

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schüler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lücking R, Büdel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006a) Reconstructing the early evolution of the fungi using a six-gene phylogeny. Nature 443:818–822

James TY, Letcher PM, Longcore JE, Mozley-Strandridge SE, Porter D, Powell MJ, Griffith GW, Vilgalys R (2006b) A molecular phylogeny of the flagellated fungi (*Chytridiomycota*) and description of a new phylum (Blastocladiomycota). Mycologia 98:860–871

Jany JL, Pawlowska TE (2010) Multinucleate spores contribute to evolutionary longevity of asexual glomeromycota. Am Nat 175:424–435. doi:10.1086/650725

Johnson DL, Garnow RI (1971) The avoidance response in *Phycomyces*. J Gen Physiol 57:41–49

Jones MDM, Forn I, Gadelha C, Egan MJ, Bass D, Massana R, Richards TA (2011a) Discovery of novel intermediate forms redefines the fungal tree of life. Nature 474:200–203

Jones MDM, Richards TA, Hawksworth DL, Bass D (2011b) Validation and justification of the phylum name Cryptomycota phyl. nov. IMA Fungus 2:173–175

Joneson S, Stajich JE, Shiu S-H, Rosenblum EB (2011) Genomic transition to pathogenicity in chytrid tungi. PLoS Pathog 7(11):e1002338

Kamaishi T, Hashimoto T, Nakamura Y, Nakamura F, Murata S, Okada N, Okamoto KI, Shimizu M, Hasegawa M (1996) Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. J Mol Evol 42:257–263

Karpov SA, Mikhailov KV, Mirzaeva GS, Mirabdullaev IM, Mamkaeva KA, Titova NN, Aleoshin VV (2013) Obligately phagotrophic aphelids turned out to branch with the earliest-diverging Fungi. Protist 164:195–205

Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prenier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivarès CP (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453

Keeling PJ, Slamovits CH (2004) Simplicity and complexity of Microsporidian genomes. Eukaryot Cell 3:1363–1369

Kneipp LF, Andrade AF, de Souza W, Angluster J, Alviano CS, Travassos LR (1998) *Trichomonas vaginalis* and *Tritrichomonas foetus*: expression of chitin at the cell surface. Exp Parasitol 89:195–204

Lang BF, O'Kelly C, Nerad T, Gray MW, Burger G (2002) The closest unicellular relatives of animals. Curr Biol 12:1773–1778

Lanternier F, Dannaoui E, Morizot G, Elie C, Huerre M et al (2012) A global analysis of mucormycosis in France: the RetroZygo Study (2005–2007). Clin Infect Dis 54:35–43. doi:10.1093/cid/cir880

Lara E, Moreira D, López-García P (2010) The environmental clade LKM11 and *Rozella* form the deepest branching clade of fungi. Protist 161:116–121

Lee SC, Li A, Calo S, Heitman J (2013) Calcineurin plays key roles in the dimorphic transition and virulence

of the human pathogenic zygomycete *Mucor circinelloides*. PLoS Pathog 9(9):e1003625. doi:10.1371/journal.ppat.1003625

Li CH, Cervantes M, Springer DJ, Boekhout T, Ruiz-Vazquez RM, Torres-Martinez SR, Lee SC (2011) Sporangiospore size dimorphism is linked to virulence of *Mucor circinelloides*. PLoS Pathog 7 (6):e1002086. doi:10.1371/journal.ppat.1002086

Liggenstoffer AS, Youssef NH, Couger MB, Elshahed MS (2010) Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. ISME J 4:1225–1235

Liu YJ, Hodson MC, Hall BD (2006) Loss of the flagellum happened only once in the fungal lineage: phylogenetic structure of kingdom Fungi inferred from RNA polymerase II subunit genes. BMC Evol Biol 6:74. doi:10.1186/1471-2148-6-74

Longcore JE, Pessier AP, Nichols DK (1999) *Batrachochytrium dendrobatidis* gen. et sp. nov., a chytrid pathogenic to amphibians. Mycologia 91:219–227

Ma LJ, Ibrahim AS, Skory C, Grabherr MG, Burger G et al (2009) Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. PLoS Genet 5:e1000549. doi:10.1371/journal.pgen.1000549

Marano AV, Gleason FH, Bärlocher F, Pires-Zottarelli CL, Lilje O, Schmidt SK, Rasconi S, Kagami M, Barrera MD, Sime-Ngando T, Boussiba S, de Souza JI, Edwards JE (2012) Quantitative methods for the analysis of zoosporic fungi. J Microbiol Methods 89:22–33

Marcet-Houben M, Marceddu G, Gabaldón T (2009) Phylogenomics of the oxidative phosphorylation in fungi reveals extensive gene duplication followed by functional divergence. BMC Evol Biol 9:295

Martel A, Spitzen-van der Sluijs A, Blooi M, Bert W, Ducatelle R, Woeltjes A, Bosman W, Chiers K, Bossuyt F, Pasmans F (2013) *Batrachochytrium salamandrivorans* sp. nov. causes lethal chytridiomycosis in amphibians. Proc Natl Acad Sci U S A 110:15325–15329

Martin F, Gianinazzi-Pearson V, Hijri M, Lammers P, Requena N, Sanders IR, Sachar-Hill Y, Tuskan GA, Young JPW (2008) The long hard road to a completed *Glomus intraradices* genome. New Phytol 180:747–750

Maxson LR, Maxson RD (1990) Immunological techniques. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer, Sunderland, pp 127–155

McNeill J, Barrie FR, Buck WR, Demoulin V, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Marhold K, Prado J, Prud'Homme van Reine WF, Smith GF, Wiersema JH, Turland NJ (2012) International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). Regnum vegetabile 154. Adopted by the Eighteenth International Botnical Congress Melbourne, Australia, July 2011. Koeltz Scientific Books, Koenigstein

Mehrotra RS, Aneja KR (1990) An introduction to Mycology, 1st edn. New Age (P) Ltd. Publishers, New Delhi, pp 141–143

Mendoza L, Taylor JW, Ajello L (2002) The class Mesomycetozoea: a heterogenous group of microorganisms at the animal-fungal boundary. Annu Rev Microbiol 56:315–344

Mulisch M (1993) Chitin in protistan organisms: distribution, synthesis and deposition. Eur J Protistol 29:1–18

Noha H, Youssef NH, Couger MB, Struchtemeyer CG, Liggenstoffer AS, Prade RA, Najar FZ, Atiyeh HK, Wilkins MR, Mostafa S, Elshahed MS (2013) The genome of the anaerobic fungus *Orpinomyces* sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. Appl Environ Microbiol 79:4620–4634. doi:10.1128/AEM.00821-13

Paterson RRM, Bridge PD (1994) Biochemical techniques for filamentous fungi. CAB International, Wallingford

Patterson DJ (1999) The diversity of eukaryotes. Am Nat 154:S96–S124

Patterson DJ, Zölffel M (1991) Heterotrophic flagellates of uncertain taxonomic position. In: Patterson DJ, Larsen J (eds) The biology of free-living heterotrophic flagellates, vol 45, Systematics Association Special Volume. Clarendon, Oxford, pp 427–447

Peyretaillade E, Biderre C, Peyret P, Duffieux F, Méténier G, Gouy M, Michot B, Vivarès CP (1998a) Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. Nucleic Acids Res 26:3513–3520

Peyretaillade E, Broussolle V, Peyret P, Méténier G, Gouy M, Vivarès CP (1998b) Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. Mol Biol Evol 15:683–689

Philipps BL, Puschendorf R (2013) Do pathogens become more virulent as they spread? Evidence from the amphibian declines in Central America. Proc Biol Sci 280:20131290

Pirozynski KA, Malloch DW (1975) The origin of land plants: a matter of mycotrophism. Biosystems 6:153–164

Pitt JI (1979) The genus *Penicillium* and its teleomorphic states *Eupenicillium* and *Talaromyces*. Academic, London

Popescu T, Roessler A, Fukshansky L (1989) A novel effect in *Phycomyces* phototropism. Plant Physiol 91:1586–1593

Ragan MA, Murphy CA, Rand TG (2003) Are Ichthyosporea animals or fungi? Bayesian phylogenetic analysis of elongation factor 1a of Ichthyophonus irregularis. Mol Phylogent Evol 29:550–562

Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. Science 289:1920–1921

Remy W, Taylor TN, Hass H, Kerp H (1994) Four hundred-million year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci U S A 91:11841–11843

Ribes JA, Vanover-Sams CL, Baker DJ (2000) Zygomycetes in human disease. Clin Microbiol Rev 13:236–301

Ride WDL, Cogger HG, Dupuis C, Kraus O, Minelli A, Thompson FC, Tubbs PK (2000) International code of zoological nomenclature. Adopted by the International Union of Biological Sciences. International Trust for Zoological Nomenclature 1999, London

Roden MM, Zaoutis TE, Buchanan WL, Knudsen TA, Sarkisova TA, Schaufele RL, Sein M, Sein T, Chiou CC, Chu JH, Kontoyiannis DP, Walsh TJ (2005) Epidemiology and outcome of zygomycosis: a review of 929 reported cases. Clin Infect Dis 41:634–653

Rosenblum EB, James TY, Zamudio KR, Poorten TJ, Ilut D, Rodriguez D, Eastman JM, Richards-Hrdlicka K, Joneson S, Jenkinson TS, Longcore JE, Olea GP, Toledo LF, Arellano ML, Medina EM, Restrepo S, Flechas SV, Berger L, Briggs CJ, Stajich JE (2013) Complex history of the amphibian-killing chytrid fungus revealed with genome resequencing data. Proc Natl Acad Sci U S A 110:9193–9194

Ruiz-Trillo I, Lane CE, Archibald JM, Roger AJ (2006) Insights into the evolutionary origin and genome architecture of the unicellular opisthokonts *Capsaspora owczarzaki* and *Sphaeroforma arctica*. J Eukaryot Microbiol 53:379–384

Ruiz-Trillo I, Burger G, Holland PWH, King N, Lang BF, Roger AJ, Gray MW (2007) The origins of multicellularity: a multi-taxon genome initiative. Trends Genet 23:113–118

Saffo MB, Fultz S (1986) Chitin in the symbiotic protist *Nephromyces*. Can J Bot 64:1306–1310

Saffo MB, McCoy AM, Rieken C, Slamovits CH (2010) *Nephromyces*, a beneficial apicomplexan symbiont in marine animals. Proc Natl Acad Sci U S A 107:16190–16195

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci U S A 109:6241–6246

Schroeckh V, Scherlach K, Nützmann HW, Ekaterina S, Schmidt-Heck SJ, Martin K, Hertweck C, Brakhage AA (2009) Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. Proc Natl Acad Sci U S A 106:14558–14563

Shalchian-Tabrizi K, Minge MA, Espelund M, Orr R, Ruden T, Jakobsen KS, Cavalier-Smith T (2008) Multigene phylogeny of Choanozoa and the origin of animals. PLoS ONE 3:e2098. doi:10.1371/journal.pone.0002098

Shelest E (2008) Transcription factors in fungi. FEMS Microbiol Lett 286:145–151

Simon L, Bousquet J, Lévesque RC, Lalonde M (1993) Origin and diversification of endomycorrhizal fungi and co-incidence with vascular land plants. Nature 363:67–69

Skiada A, Pagano L, Groll A, Zimmerli S, Dupont B, Lagrou K, Lass-Flörl C, Bouza E, Klimko N, Gaustad P, Richardson M, Hamal P, Akova M, Meis JF, Rodriguez-Tudela JL, Roilides E, Mitrousia-Ziouva A, Petrikkos G, European Confederation of Medical Mycology Working Group on Zygomycosis (2011) Zygomycosis in Europe: analysis of 230 cases accrued by the registry of the European Confederation of Medical Mycology (ECMM) Working Group on Zygomycosis between 2005 and 2007. Clin Microbiol Infect 17:1859–1867

Smith SE, Read DJ (2008) Mycorrhizal symbiosis. Academic, San Diego

Spatafora J (2011) 1000 fungal genomes to be sequenced. Fungi: 1000 genomes. IMA Fungus 2 (2):41

Steenkamp ET, Wright J, Baldauf SL (2006) The protistan origins of animals and fungi. Mol Biol Evol 23:93–106

Sugar AM (1992) Mucormycosis. Clin Infect Dis 14 (Suppl 1):S126–S129. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=90004andtool=pmcentrezandrendertype=abstract

Sun HY, Aguado JM, Bonatti H, Forrest G, Gupta KL, Safdar N, John GT, Pursell KJ, Muñoz P, Patel R, Fortun J, Martin-Davila P, Philippe B, Philit F, Tabah A, Terzi N, Chatelet V, Kusne S, Clark N, Blumberg E, Julia MB, Humar A, Houston S, Lass-Florl C, Johnson L, Dubberke ER, Barron MA, Lortholary O, Singh N, Zygomycosis Transplant Study Group (2009) Pulmonary zygomycosis in solid organ transplant recipients in the current era. Am J Transplant 9:2166–2171

Tanabe Y, Watanabe MM, Sugiyama J (2002) Are *Microsporidia* really related to *Fungi*? A reappraisal based on additional gene sequences from basal fungi. Mycol Res 106:1380–1391

Vitale RG, de Hoog GS, Schwarz P, Dannaoui E, Deng S, Machouart M, Voigt K, van de Sande WWJ, Dolatabadi S, Meis JF, Walther G (2012) Antifungal susceptibility and phylogeny of opportunistic members of the order Mucorales. J Clin Microbiol 50:66–75

Voigt K (2012a) Chytridiomycota. In: Engler A, Frey W (eds) Syllabus of plant families, vol 1/1: Blue-green algae, myxomycetes and myxomycete-like organisms, phytoparqasitic protists, heterotrophic Heterokontobionta and Fungi pp. Borntraeger, Stuttgart, pp 106–129

Voigt K (2012b) Zygomycota. In: Engler A, Frey W (eds) Syllabus of plant families, vol 1/1: Blue-green algae, myxomycetes and myxomycete-like organisms, phytoparqasitic protists, heterotrophic Heterokontobionta and Fungi pp. Borntraeger, Stuttgart, pp 130–162

Voigt K, Kirk PM (2011) Recent developments in the taxonomic affiliation and phylogenetic positioning of fungi: impact in applied microbiology and environmental biotechnology. Appl Microbiol Biotechnol 90:41–57

Voigt K, Marano AV, Gleason F (2013) Ecological and economical importance of parasitic and zoosporic true Fungi. In: Kempken F (ed) The Mycota vol. XI: agricultural applications, 2nd edn. Springer, Berlin, pp 243–270

Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326:411–414

Voyles J, Young S, Berger L, Campbell C, Voyles WF, Dinudom A, Cook D, Webb R, Alford RA, Skerratt LF, Speare R (2009) Pathogenesis of chytridiomycosis, a cause of catastrophic amphibian declines. Science 326:582–585

Voyles J, Vredenburg VT, Tunstall TS, Parker JM, Briggs CJ, Rosenblum EB (2012) Pathophysiology in mountain yellow-legged frogs (*Rana muscosa*) during a chytridiomycosis outbreak. PLoS ONE 7: e35374. doi:10.1371/journal.pone.0035374

Wang B, Yeun LH, Xue JY, Liu Y, Ané JM, Qiu YL (2010) Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol 186:514–525

Wang L, Chen W, Feng Y, Ren Y, Gu Z, Chen H, Wang H, Thomas MJ, Zhang B, Berquin IM, Li Y, Wu J, Zhang H, Song Y, Liu X, Norris JS, Wang S, Du P, Shen J, Wang N, Yang Y, Wang W, Feng L, Ratledge C, Zhang H, Chen YQ (2011) Genome characterization of the oleaginous fungus *Mortierella alpina*. PLoS ONE 6(12):e28319. doi:10.1371/journal.pone.0028319

Wang D, Wu R, Xu Y, Li M (2013) Draft genome sequence of *Rhizopus chinensis* CCTCCM201021, used for brewing traditional Chinese alcoholic beverages. Genome Announc 1(2):e0019512. doi:10.1128/genomeA.00195-12

Ward HHD, Alroy J, Lev BI, Keusch GT, Pereira MEA (1985) Identification of chitin as a structural component of *Giardia* cysts. Infect Immun 49:629–634

Webster J, Weber RWS (2007) Introduction to fungi. Cambridge University Press, Cambridge

Whittaker RH (1969) New concepts of kingdoms of organisms. Science 163:150–160

Zettler LA, Nerad T, O'Kelly C, Sogin M (2001) The nucleariid amoebae: more protists at the animal-fungal boundary. J Eukaryot Microbiol 48:293–297

Zhao Z, Huiquan Liu H, Wang C, Xu JR (2013) Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics 14:274

3 Phylogenomics for the Study of Fungal Biology

Toni Gabaldón[1,2,3], Marina Marcet-Houben[1,2]

CONTENTS

I. Introduction

Phylogenomics has been defined as the intersection between the fields of evolution and genomics (Eisen and Fraser 2003). In other words, this emerging area can be considered as the scientific discipline that studies genome sequences under the dynamic perspective of evolution. Extant organisms are related by common ancestry, and one of the tasks of phylogenomics is to understand what these evolutionary relationships are, with the final aim of reconstructing a **tree of life** (not necessarily strictly bifurcating and perhaps incorporating some reticulated relationships) that would represent the genealogies of species. In addition, all sequences in extant species are the result of billions of years of evolution during which they have been shaped by mutation, selection, recombination, and other evolutionary processes. As a result of the important pressure imposed by natural selection, many of the features present in current genomes are informative of past and recent selective constraints. Phylogenomics also has an interest in how these forces shape genomes and, conversely, on how we can use the comparison of current genomes to understand how they have come about, what selective pressures may have played a role, and which parts of the genome have been more sensitive to them. Such an approach, as we shall see in the last section of this chapter, is sometimes able to shed light on the functional roles of different genomic features.

Since the sequence of the baker's yeast *Saccharomyces cerevisiae* was completed (Goffeau et al. 1996), a growing number of other fungal genome sequences have been determined. With more than 250 genome sequences available in public databases as we write this chapter, fungi is clearly the eukaryotic group that is best sampled in terms of completely sequenced genomes. This – coupled with its medical, ecological, and industrial importance – has

[1]Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Dr. Aiguader, 88, Barcelona 08003, Spain; e-mail: toni.gabaldon@crg.eu
[2]Universitat Pompeu Fabra (UPF), Barcelona 08003, Spain e-mail: toni.gabaldon@crg.eu
[3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, Barcelona 08010, Spain e-mail: toni.gabaldon@crg.eu

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

favored the use of the fungal phylum as a playground to develop and use sophisticated phylogenomic approaches. In this chapter, we overview the main phylogenomic approaches that have been used in fungal organisms and provide some illustrative examples of successful applications. Rather than a comprehensive treaty on fungal phylogenomics, our intention is to spark the interest of researchers considering the use of fungal genomic sequences but may not have considered exploiting a fully fledged phylogenomic approach. After reading this chapter, we hope they will have an accurate picture of what phylogenomics can bring to the understanding of the biology of fungi.

II. Determining Which Gene Is Which Gene Across Genomes: Orthology and Paralogy

A prerequisite for comparing the gene complement of several genomes is to establish gene correspondences, that is, to find which gene "is" which gene in the other genome. The most exact way to delimit such equivalences is to establish orthology relationships because orthologous genes correspond to a single gene in the common ancestor of the species compared (Fitch 1970; Gabaldón 2008b). In addition, given that orthologs, by definition, diverged through a speciation event, they constitute the prime material to reconstruct species relationships. Finally, orthologs are, as compared to paralogs, more likely to perform a similar function, which constitutes the basis of current approaches for functional annotations in the genome era (Gabaldón and Koonin 2013). Hence, prediction of orthology usually constitutes one of the first steps in comparative analyses. In addition, paralogous relationships, occurring among genes that diverged through a duplication event, can also provide useful information when interpreting differences in the gene repertoires of the considered species (see the following text for a discussion of lineage-specific expansions). There are many different approaches to establish orthology (and paralogy) relationships across a set of species (Gabaldón 2008b; Kristensen et al. 2011). Many methods rely solely on pairwise similarity searches, such as best bidirectional hits (Huynen and Bork 1998) or variations thereof, such as inparanoid (Berglund et al. 2008). Others extend these by combining many pairwise genomic comparisons that result in clusters of orthologous groups (Li et al. 2003; Kriventseva et al. 2008; Powell et al. 2011). However, given the fact that orthology and paralogy are defined in purely evolutionary terms, the direct reconstruction of gene family histories is often preferred (Gabaldón 2008b).

There are two main approaches to infer orthology and paralogy relationships from phylogenetic trees, and both resort to annotating the nodes of the tree as either "duplication" or "speciation" events and then simply apply the original definition of orthology to define all possible pairwise relationships. These two phylogeny-based methods differ in how duplications are inferred. In the **"species reconciliation" approach**, the gene tree is "reconciled" with a known species tree in a process in which inconsistencies are solved by introducing duplication nodes followed by eventual gene losses (Arvestad et al. 2003). This method relies on the assumption that both the gene tree and the species tree are known and correct, and that the gene tree has strictly followed the species tree. Because these assumptions are often violated, it is recommended to account for uncertainties in both the gene tree and the species tree. An alternative method is to use a **"species overlap" algorithm**. In this case, nodes in the tree are labeled as speciation or duplication depending on the shared species at either side of the node. Specifically, nodes defining two descending lineages that share species will be considered duplication nodes, while those that do not share species will be considered speciation nodes. This method is simpler, faster, and more robust to phylogenetic noise than a strict reconciliation approach, although it is not totally free of errors. Synteny, despite the fact that it is not implied by the definition of orthology, is often used to clarify orthology relationships, and some databases and methods use this criterion coupled with an approach based on BLAST (Basic Local Alignment Search

Tool) (Byrne and Wolfe 2006; Fitzpatrick et al. 2010) or in combination with tree reconstruction (Wapinski et al. 2007). Finally, some recent initiatives combine predictions from different sources to obtain more accurate relationships (Pryszcz et al. 2011).

III. Phylogenomics Approaches for Species Tree Reconstruction

A. Placing a Genome in an Evolutionary Context

One of the first steps in a phylogenomics study is to situate our species of interest within an evolutionary framework, usually a species tree depicting the relationship of our organism of interest with other related species, especially those for which there is availability of relevant data (e.g., genomic sequences, experimental data, or phenotypic information). This species tree will be an essential tool for subsequent studies, serving a twofold purpose. First, given a species of interest, it allows us to choose additional ones, at appropriate evolutionary distances to be used in a comparative analysis. Second, it allows us to interpret and provide directionality and a temporal scale to numerous evolutionary processes, such as phenotypic transitions, gene duplication, loss, and transfer events that would otherwise remain muddled.

There are numerous approaches to reconstruct a species tree from genome data (Delsuc et al. 2005); the most widely used can be grouped into two main methods: **gene concatenation** and **supertree reconstruction**. The gene concatenation analysis, also known as a supermatrix approach, is based on the concatenation of several alignments into a combined one, which is subsequently analyzed with standard phylogenetic methods. The rationale behind this approach is that, by combining the signal coming from several genes, the signal-to-noise ratio is expected to improve. The genes used in the concatenation approach need to fulfill two basic requirements: being widespread and present in single copy in (most of) the species considered. These requirements can dramatically limit the number of genes suitable for analysis. Fungal genomes are diverse and can vary extensively in gene content, which constitutes a challenge in the search for widespread genes for concatenation analyses. In addition, some species, such as *S. cerevisiae* or *Rhizopus oryzae*, have undergone whole-genome duplications (WGDs) (Wolfe and Shields 1997; Ma et al. 2009), and others such as *Cryptococcus neoformans* (Fraser et al. 2005) contain long segmental duplications. Both events result in a large amount of genes that violate the single-copy requirement. Accurate orthology predictions can in some cases ease the problem by allowing differentiation between orthologs and paralogs in multigene families. This would allow the inclusion of some genes that have undergone gene duplication but for which orthologs can nevertheless be identified in the relevant species.

An alternative strategy for species tree inference is the so-called supertree approach. The basic procedure of this methodology consists of combining a collection of gene tree topologies into a single species tree that is most compatible with them according to some optimization criteria. A plethora of supertree algorithms exist that differ in the optimization criteria and the specific heuristics used to find an optimal solution (Bininda-Emonds 2005). Most popular are those based on maximizing the number of shared tree partitions (or quartets) between the species tree and the gene trees, as in the pioneering program CLANN (Creevey and McInerney 2009), or by minimizing the number of inferred duplications after reconciling the gene trees with the species tree, as in the recently developed DupTree algorithm (Wehe et al. 2008). A clear advantage of the supertree method over the gene concatenation approach is that it does not require that the genes are widespread, as long as the trees share a certain fraction of common species. Some supertree approaches, such as those based on reconciliation, can handle gene phylogenies with duplications. Trees obtained using a supermatrix and a supertree approach tend to have highly similar topologies if the amount of information is sufficient (Snel et al. 2005), although sometimes these slight differences may affect the focus of a study. Overall,

there is a higher tendency to use concatenation methods, with several arguments supporting this preference. First, concatenation methods exploit the sequence information directly, rather than indirectly as in the supertree approach, and are thus expected to be less prone to artifacts. In addition, supertree approaches involve the reconstruction of a large number of phylogenetic trees, which is computationally demanding. In addition, because the combination of the signal is done after phylogenetic reconstruction, artifacts that have a pervasive effect on single-gene phylogenies, such as long-branch attraction (Gribaldo and Philippe 2002), are difficult to correct. Finally, supertrees with few genes tend to give multiple results, which have to be summarized in partially unresolved topologies.

One way of illustrating the importance of establishing the phylogenetic position of newly sequenced genomes is to assess how often a phylogenetic analysis is included in the description of a new genome sequence. Phylogenies depicting the position of newly sequenced fungal genomes have indeed become recurrent in the literature. Since the publication of the yeast genome in 1996, a total of 147 fungal genomes have been published in 94 articles. Of such publications, 57 % contained a tree showing the position of the newly sequenced species in the fungal species tree.

In the last year (i.e., 2012), 13 of the 20 genome papers that involved fungal species contained a phylogenetic tree, and 10 of the papers that contained a species tree used phylogenomic data. The remaining three trees used alternative methodologies, obtained the tree from a previously published analysis, or did not provide information on how the phylogenetic species tree had been reconstructed.

B. Toward a Fungal Tree of Life

The reconstruction of the evolutionary relationships across fungal species in general, the so-called fungal tree of life (fTOL), constitutes a challenge that goes well beyond the individual efforts to situate a recently sequenced species in a phylogeny (Fig. 3.1). Both aims are somewhat related because the availability of a resolved fTOL is extremely helpful in defining the relevant context in which to situate a newly sequenced genome. Conversely, the agglomeration of targeted efforts helps to provide deeper resolution to the fTOL, in a continuous iterative process of expansion and refinement. Of note, our use of the metaphoric term *tree of life*, criticized by some (Merhej and Raoult 2012), should be taken as it is meant here, an ideal representation of species relationships that proceed mainly in a bifurcating manner but that does not preclude the existence of reticulate evolutionary processes such as hybridization or horizontal gene transfer (HGT), discussed further in this chapter. Here, we just provide a brief overview of recent efforts toward the reconstruction of the fTOL, with a particular focus on the use of complete genomes.

One of the most comprehensive fungal species trees published in recent years is the one reconstructed by James and collaborators (2006). They used a gene concatenation method that united six genes that were widespread in 199 fungal species. The aim of the tree was to elucidate the evolutionary history of early-diverging fungi, one of the topics that has attracted more attention in the field in recent times. One of the main points discussed in the paper was the placement of the controversial group of **Microsporidia**. The correct placement of this group of intracellular pathogens within the eukaryotic species tree has puzzled scientists for many years (Keeling and Fast 2002). Through the years, the position of this group has been moving around in the eukaryotic species tree, including the proposition that they belonged to a clade of eukaryotes – Archaeozoa – that diverged prior to the acquisition of mitochondria (Cavalier-Smith 1983). Finally, in the 1990s, the increasing amount of phylogenetic evidence coupled with the discovery of the presence of mitosomes, a relic form of mitochondria, showed that microsporidia were closely related to fungi. The exact relationship has been thoroughly discussed. Phylogenies involving few genes were inconclusive, placing microsporidia in several different positions, such as at the base of dikarya (Gill and Fast

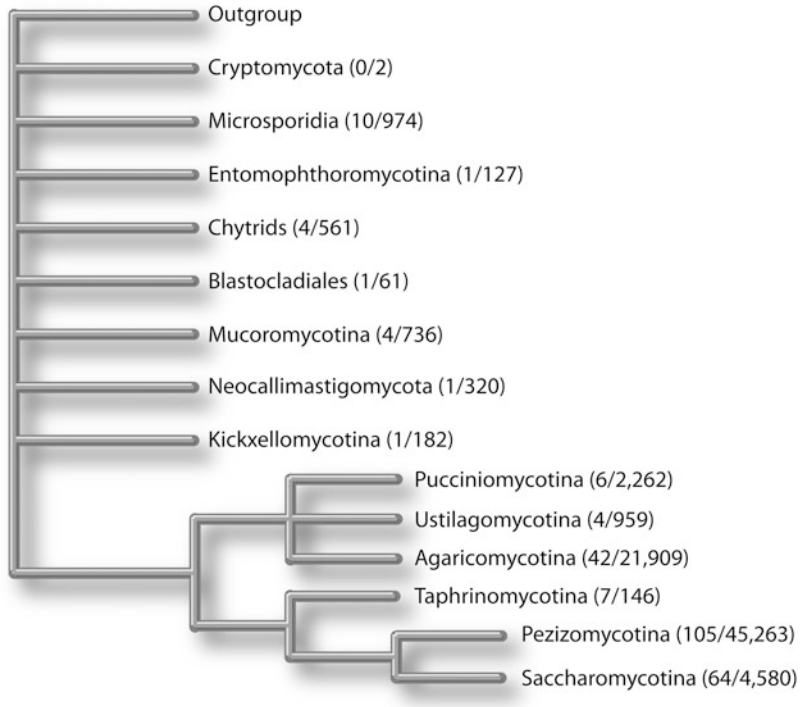


Fig. 3.1. Schematic view of the current understanding of the fungal species tree: *Numbers* represent the amount of currently sequenced genomes and approximate number of identified species according to National Center for Biotechnology Information (NCBI) taxonomy, respectively

2006), within the Zygomycotina (Keeling 2003), as sister to Blastocladiales and Entomophthoromycota (Van de Peer et al. 2000), or as sister group to the chytrid *Rozella allomycis*, as shown by James et al. (2006). This controversy is to be expected in a group of species that has undergone a massive reduction of their genomes coupled with accelerated rates of evolution. The first microsporidian genome sequenced, *Encephalitozoon cuniculi* (Katinka et al. 2001), and those that have followed have enabled the use of phylogenomic analysis on this group of species to resolve their phylogenetic position. In the most extensive phylogenomic analysis of microsporidia so far, involving six complete microsporidian genomes, we showed that this elusive group held an early-branching position among the currently sequenced fungi (Capella-Gutierrez et al. 2012). The species tree, based on the concatenation of 53 single-copy, widespread genes, was highly supported. Statistical tests assessed the possibility that Microsporidia appeared in another part of the tree. All alternatives were discarded because of the lack of statistical support. This result was supported in a second phylogenetic analysis performed by Cuomo et al. (2012). The correct position of microsporidia is not only relevant regarding the group itself but also is important to assess other important aspects of fungal phylogeny, such as which are their closest relatives and what genomic changes may be related to their adaptation to an obligate parasitic lifestyle. The basal position of Microsporidia in the fungal species tree admits the possibility that they are in fact a sister group of fungi and not placed within the fungal kingdom. Nucleariida were also shown to be closely related to fungi in a phylogeny based on 118 orthologous genes (Liu et al. 2009b). Environmental analyses have also shown that there are numerous species left to discover between the Nucleariida and the known fungal species (Jones et al. 2011).

These species would be related to *Rozella* and perhaps form the earliest diverging fungal group. The relationship between Microsporidia and this new group, named the **Cryptomycota**, remains unknown, and only newly sequenced genome data will be able to elucidate their relation at the base of the fungal species tree.

Although recent research on a growing amount of genomic data has helped to resolve many contentious parts of the fungal species tree, others remain controversial. Since the early analysis of ribosomal RNA sequences, three clear domains became evident within Ascomycotina: Saccharomycotina, Pezizomycotina, and Taphrinomycotina. This last group was formed by a clade of previously unrelated fungi. The group contained *Schizosaccharomyces* species, with *Schizosaccharomyces pombe* the most well-known species of the group, *Taphrina*, which comprises several plant pathogens: *Saitoella*, *Neolecta* and *Protomyces*. Yet, the monophyly of Taphrinomycotina is still in question. Some further analyses have in fact created a greater conundrum in grouping *Schizosaccharomyces* with Saccharomycotina (Baldauf et al. 2000) instead of finding it as sister of both Saccharomycotina and Pezizomycotina.

This finally presented two main questions to the scientific community that have still not been fully answered. The first one is whether *Schizosaccharomyces* should be grouped with Saccharomycotina instead of featuring as the basal Ascomycotina group. The second question relates to whether Taphrinomycotina is indeed a monophyletic group. The use of phylogenetic data seems to have answered, at least in part, those questions, with numerous phylogenies supporting the basal position of Taphrinomycotina in the species tree of Ascomycota (Fitzpatrick et al. 2006; Liu et al. 2009a; Marcet-Houben and Gabaldon 2009). Liu et al. (2009a) suggested that the grouping between *Schizosaccharomyces* and Saccharomycotina was due to long-branch attraction and showed, in an analysis of 113 nuclear genes, that Taphrinomycotina opted for the basal position in Ascomycotina with high statistical support. On the other hand, the same analysis performed on 13 mitochondrial proteins resulted in the alternative topology of *Schizosaccharomyces* grouping with Saccharomycotina, a result already reported in an early study (Bullerwell et al. 2003). In this instance, the basal position of *Schizosaccharomyces* could not be discarded as a valid alternative topology. Adding sequences from slow-evolving Taphrinomycotina (*Taphrina* and *Saitoella*) reduced the bootstrap support of the *Schizosaccharomyces*-Saccharomycotina node even further. The complete sequencing of the remaining members of Taphrinomycotina should finally elucidate whether the whole clade is indeed monophyletic.

IV. Focusing on Gene Family Histories: From Genomes to Phylomes

Often, the focus of interest is the evolutionary history of the genes encoded in a given genome or group of genomes, rather than that of the species themselves. Placing the gene complement of each species in the framework of the known or reconstructed species tree (see section IIIb) can lead to interesting insights, such as the loss or enrichment of members of certain families in particular clades (Marcet-Houben et al. 2009; Bushley and Turgeon 2010; Kovalchuk and Driessen 2010). Using some inference methods, one can reconstruct the likely histories of gains and losses across the phylogeny for each gene family (Eirin-Lopez et al. 2012). This, in turn, enables the detection of correlations between the emergence of certain phenotypic traits with events of expansion, compaction, or total loss of certain families or pathways. Such correlations, although clearly not a proof of cause-and-effect relationships, serve to hint at possible genomic changes underlying the emergence or disappearance of a given trait.

A more accurate way of looking into the evolutionary dynamics of genes is to directly reconstruct their phylogenies (Gabaldón 2005). Gene phylogenies can differ from the species tree due to multiple evolutionary processes, such as HGT, differential gene loss, hybridization processes, and incomplete lineage sorting, among others. These processes are difficult to detect and model from the information contained in mapping gene complements on a

species tree but can certainly be considered by careful analysis of gene phylogenies, sometimes in combination with the species tree. Thus, a collection of gene family trees will provide us a more appropriate view of the evolution of a genome. The gene tree approach for the reconstruction of the evolution of entire genomes has been traditionally avoided because of the serious computational challenges involved (Gabaldón et al. 2008). However, recent progress in computer architecture and in computational phylogenetic software and algorithms has enabled the development of highly accurate and fully automated pipelines to reconstruct large collections of gene phylogenies. Two main approaches are used. In the most extended one, a two-step procedure is used in which genes are first grouped into families using some clustering algorithm, and then a phylogeny is reconstructed for each family. A main limitation of this approach is the delimitation of families: Often, slight changes in the clustering parameters result in different groupings of multigene families. This results from the inherent hierarchical nature of gene families, which can often be subdivided into subfamilies or grouped into superfamilies. To circumvent such problems, a second approach is based on the reconstruction of a phylogeny for every gene, ensuring that each gene in the genome of interest is used as a seed in the process of phylogenetic reconstruction. The result of such an approach is a complete collection of evolutionary histories of all genes encoded in a given genome, the so-called **phylome** (Sicheritz-Ponten and Andersson 2001).

A phylome can be reconstructed using automated pipelines that try to balance accuracy and time consumption. The methodology applied follows the typical steps of phylogenetic tree reconstruction: homology search, multiple-sequence alignment, and tree reconstruction. The first step in the construction of the phylome is to define the taxonomic scope of interest, that is, the set of species relevant for the analysis. For this, it is important to have an approximate idea of the position of the genome(s) of major interest (often referred to as the seed genome) within the species tree, even though its exact position may be unknown. This can be ascertained using a phylogeny based on a few genes or even from the analysis of the results of an initial BLAST search. Usually, closely related species are chosen to have a good set of genomes to perform comparative analyses, but the specific set of species may depend on the questions of interest. An important aspect in the selection of species for the analysis is the selection of a few relatively distant species to serve as an out-group. This will allow the trees to be properly rooted and will provide a polarity to the detected events, allowing the establishment of a relative temporal gradient from ancient to recent. Once the set of species has been selected, their gene set will be used to perform homology searches starting from each of the genes encoded in the seed genome. Homology results are filtered based on an e-value threshold and a minimum overlap between the seed sequence and the homologous sequence. Usually, a limit in the number of sequences included in each tree is set. This is done to ensure high tree accuracy and to diminish the time needed to reconstruct each tree. As the phylome will contain a tree for each gene, large gene families will be fully represented in several overlapping trees.

Once the set of homologs has been chosen, the process of **phylogenetic reconstruction** may be common for the gene-based and family-based reconstruction approaches. This generally starts with a multiple-sequence alignment. This alignment should be as accurate as possible because it is the base of the phylogenetic reconstruction. A plethora of alignment algorithms do exist (Notredame 2007), which can even be combined in sophisticated ways to produce alignments as reliable as possible. For instance, the phylomeDB pipeline (Huerta-Cepas et al. 2011) uses several alignment programs in a forward and a reversed version of the sequences to then obtain a consensus alignment. Each alignment algorithm has different biases because of the specific heuristics applied to search for the optimal result. Using a combination of programs and only selecting those columns that are congruent in a given percentage of the alignments will ensure that no biases affect our tree. In addition, having the sequences in forward or

in reverse will also produce different multiple-sequence alignments (Landan and Graur 2007). A complementary strategy to improve the phylogenetic reconstruction is to trim the alignment to remove columns that are unreliably aligned, particularly gappy regions. These regions, while potentially informative, are more likely to introduce noise in the phylogenetic reconstruction process (Talavera and Castresana 2007). Tree construction programs remain mostly unable to correctly model indels; therefore, the information reported by gaps remains unused. Several programs can be used to trim alignments in an automated way, but care should be taken that enough sequence remains to produce a reliable tree. Some trimming programs, such as trimAl, have been developed with large-scale phylogenetic pipelines in mind and include automated heuristics to choose an optimal set of parameters for each particular family (Capella-Gutierrez et al. 2009).

A final step is the reconstruction of the gene tree based on the information contained in the sequence. Again, an enormous variety of methods and programs is available. In practice, most of the currently existing pipelines rely on a few phylogenetic programs, such as phyML (Guindon et al. 2010) and RAxML (Stamatakis et al. 2008), which are able to provide accurate phylogenies with reasonable speed. These programs, using a statistical-based modeling approach for phylogenetic reconstruction called maximum likelihood, are considered more accurate than faster, distance-based methods such as neighbor joining. An important aspect when analyzing gene trees is the potential uncertainty associated to their different partitions. This is generally computed using a bootstrapping analysis of randomly resampled alignments. In practice, this involves repeating the phylogenetic reconstruction for each resampled alignment, which obviously multiplies the necessary amount of time required. Recent implementations of alternative support measures, such as the approximate likelihood ratio test (aLRT), enable the computation of support values in much less time (Guindon et al. 2010).

The phylome reconstruction results in hundreds of trees that cannot possibly be analyzed by direct manual inspection. To extract information from this massive amount of data, trees have to be automatically processed in electronic format by appropriate algorithms and software. One such tool is ETE (Huerta-Cepas et al. 2010), a python Environment for phylogenetic Tree Exploration, which enables running programmatic operations on large collections of phylogenetic trees. Once such tools are in place, one can interrogate a phylome or any other large collection of phylogenies for the existence of relevant signatures. This can include the assessment of the fraction of gene trees that support a given topology (Marcet-Houben and Gabaldon 2009), the detection of orthology and paralogy relationships (Gabaldón 2008b), and the detection of duplication, loss, or transfer events, all of which leave a detectable footprint in the topology of a gene tree (Huerta-Cepas et al. 2007).

V. Study of Evolutionary Processes in Fungi

An immediate application of a comparison of genomic sequences placed on a phylogenetic framework is the study of the underlying evolutionary mechanisms that are shaping them. This is possible because the different evolutionary processes leave specific footprints that become apparent when several sequences are compared. A number of evolutionary processes play a role in the evolution of fungal genomes, and phylogenomics provides innovative approaches to study them. Here, we review those considered most relevant and provide specific examples where they have been successfully used.

A. Gene Family Expansions and Whole-Genome Duplications

The Japanese biologist Susumu Ohno postulated that **gene duplication** constituted one of the main forces in the evolution of species, creating the necessary material for the acquisition of novel functions (Ohno 1970). Subsequent research, often conducted in

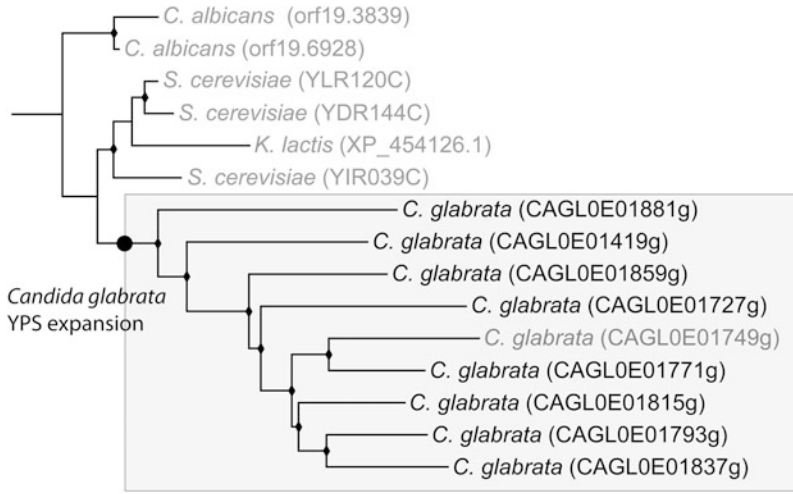


Fig. 3.2. Phylogenetic tree depicting the species-specific expansion that occurred in the YPS proteins family in *Candida glabrata*. *Black-colored leaves* represent those *C. glabrata* paralogous genes that are found in tandem in the genome, *gray-colored leaf* is a sequence not present in tandem. Duplication events are marked with *black diamonds*, and the whole expansion is *framed and shaded*

fungi, has overwhelmingly confirmed the important role of gene duplication in genome evolution. Here, we focus on the use of phylogenomics to study gene duplication events that result in large expansions of families, often indicative of directional selection, and the remarkable process of WGD.

The detection of **lineage-specific gene expansions** is particularly useful because they can be the result of selection toward a higher dosage or diversification of a particular gene family. Detection of lineage-specific expansions in gene trees is straightforward as these will be represented by branches only containing paralogs of a given species or lineage. This information can be optionally coupled with gene order information to assess whether the detected duplications are in tandem. Two such expansions are found, for instance, in the human pathogen *Candida glabrata*, where the MNT (Muller et al. 2009) and YPS (Kaur et al. 2007) families are present in eight tandem copies each in the reference genome. The YPS family encodes extracellular glycosylphosphatidylinositol-linked aspartyl proteases thought to be necessary for cell wall integrity and adherence to mammalian cells, playing a role in the virulence mechanisms of this pathogen in humans. The expansion of this family can be easily detected in the phylogenetic tree (Fig. 3.2), where nine *C. glabrata* proteins are clustered together. Eight of them appear contiguously in the genome (CAGL0E01727g, CAGL0E01749g, CAGL0E01771g, CAGL0E01793g, CAGL0E01815g, CAGL0E01837g, CAGL0E01859g, CAGL0E01881g); the ninth paralog appears in the same chromosome but not in a contiguous location (CAGL0E01419g).

Whole-genome duplication, or polyploidization, is the process by which the genomic complement of an organism is doubled (Ohno 1970). Such events have been described in at least two instances in the fungal species tree, being the best characterized WGD the one affecting the lineage of the yeast *S. cerevisiae*. There, compelling evidence for a paleopolyploidization was found in the form of numerous duplicated stretches of genes that conserved their order and orientation with respect to the centromere (Wolfe and Shields 1997). In support of these, the number of chromosomes in the monophyletic group of species that diverged after such events was doubled with respect to their closest relatives. The event is estimated to have happened roughly 100 million years ago, and only 8 % of the duplicated genes

are conserved in pairs. More recently, evidence for another ancient WGD has been identified in the genome of the Zygomycetes species *R. oryzae* (Ma et al. 2009). In this case, the finding that 50 % of the genes are present in paralogous copies suggests that this event may have been more recent than the one in *S. cerevisiae*.

Phylogenomics can be a great help in identifying and characterizing past events of polyploidization. Besides the finding of duplicated blocks of paralogous genes, the direct analyses of gene phylogenies in a large scale can help identify such events and assess the timing of the event. Of note, analysis of gene family trees can help detect WGD if the conservation of synteny (i.e., gene order) has been lost. In principle, given a procedure to approximately date duplication events detected in a gene family phylogeny (Huerta-Cepas and Gabaldón 2010), an ancient WGD should be detected as an accumulation of duplication events at this point. We explored the detection of the *R. oryzae* WGD event mentioned by reconstructing a phylome centered on this species (Fig. 3.3). The detection of duplication events, after filtering for large family expansions, showed a duplication frequency of 0.48, similar to the predicted rate of conservation of duplicated genes after the WGD event (53 %). An additional WGD event was predicted at the base of Mucoromycotina, where the node contains a duplication rate of 0.99 duplications per branch per gene. This indicates that nearly each gene in the Mucoromycotina genomes experienced a duplication in that lineage.

B. Horizontal Gene Transfer

Horizontal gene transfer, also known as lateral gene transfer (LTG), can be defined as the non-vertical transference of genetic material between two different species. This mechanism has been amply studied in prokaryotes, for which between 2 % and 60 % of the genome is estimated to have been acquired from other species. Until recently, HGT was thought to be almost negligible in eukaryotes, with the remarkable exception of ancient endosymbiotic transfers (Timmis et al. 2004) and those microbial eukaryotes that presented a phagotrophic lifestyle (Doolittle 1998). Fungi, because they are not phagotrophic and present possible barriers to HGT such as the presence of a cell wall, were considered not to be particularly prone to HGT. In recent years, however, increasing amounts of evidence suggest that this mechanism is more prevalent in fungi than previously thought. Perhaps the best-known example of HGT in fungi is the transference of dihydroorotic acid dehydrogenase (URA1) from a group of lactic bacteria to *S. cerevisiae* (Gojkovic et al. 2004). This event is thought to have greatly influenced the lifestyle of the yeast ancestor as it allowed the synthesis of pyrimidines in the absence of oxygen. Other cases of transfer events involve fungi as the donor species. A remarkable example is the transfer of genes involved in the synthesis of carotenoids, pigments involved in processes such as vertebrate night vision, protection from oxidative damage, or coloring of birds or insects. Until recently, it was assumed that all animals obtained such pigments from the diet, as they were incapable of performing this synthesis. However, two independent studies performed on the aphid *Acyrthosiphon pisum* (Moran and Jarvik 2010) and the two-spotted spider mite *Tetranychus urticae* (Altincicek et al. 2011) showed the transference of genes related to carotenoid synthesis from fungal species into these arthropods. Whether both transferences were from fungi or whether there was a transference between the two arthropods is still a matter of debate.

A further analysis (Novakova and Moran 2011) was performed on 38 insect species, 34 belonging to the main aphid groups and 4 belonging to related species: Adelgidae, Psyllidae, and Aleyrodidae. A search for a putatively transferred carotene desaturase in all these species revealed that all aphid species contained between one and seven copies of the gene. The potential donor of carotenoid genes is found in Mucoromycotina, a group of ancestral fungi. The scarcity of sequence data in the early diverging fungal groups renders this result as only provisional because newly sequenced genomes may change it.

Another type of transfer event involves transferences between different fungal lineages. These are known to exist, and many

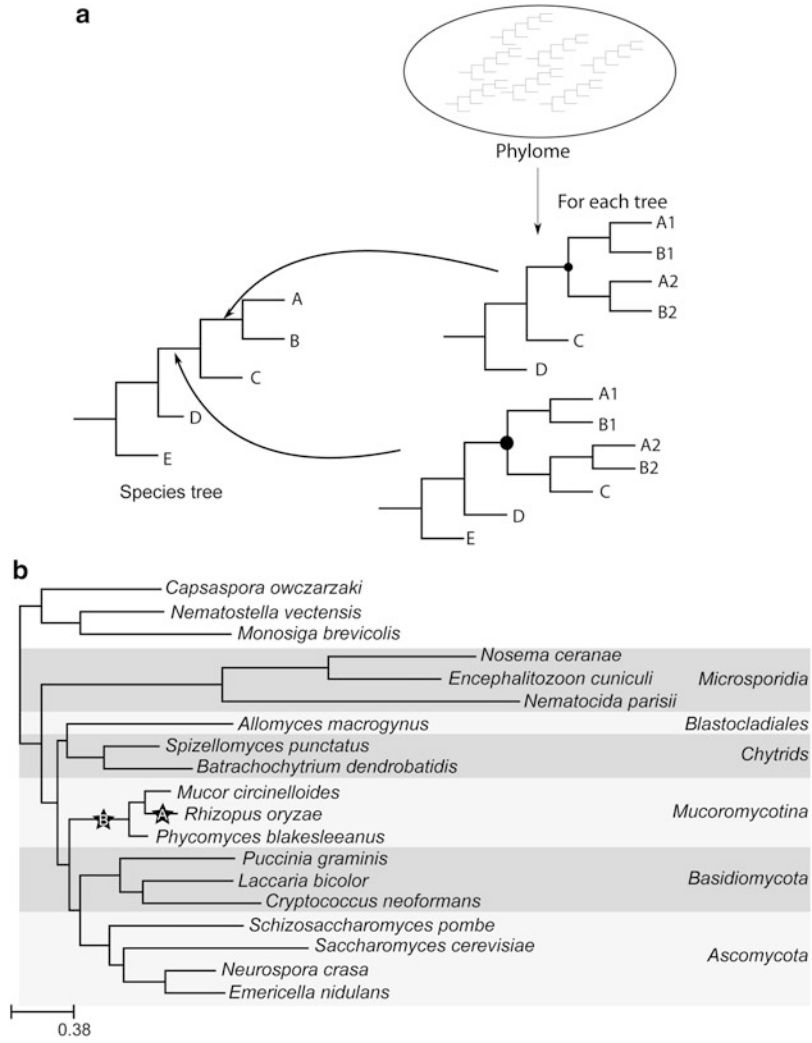


Fig. 3.3. Detecting whole-genome duplications: (**a**) schematic representation of the process to calculate duplication frequency from a phylome; (**b**) duplication frequency of the Mucoromycotina species. *Numbers* represent the duplication frequency at each node in the lineage. Two rounds of whole-genome duplication have been predicted at the nodes marked with a *star*

independent studies have identified a multitude of examples; however, the overall impact of this process in fungal genome evolution remains to be quantified. Known examples include the transfer of multiple genes, whole secondary metabolism clusters (Khaldi and Wolfe 2011; Slot and Rokas 2011; Campbell et al. 2012), or even complete, albeit small in size, chromosomes (Ma et al. 2010). In particular, the acquisition of entire supernumerary chromosomes in *Fusarium* strains has been shown to confer the ability to infect specific plant hosts.

The growing awareness of the importance of HGT events in shaping fungal evolution, coupled with the increasing amount of data, motivated a large-scale phylogenomics study in which **HGT events from prokaryotes to fungi** were detected (Marcet-Houben and Gabaldon 2010). A set of 60 completely sequenced fungal species was scanned, searching for abnormal patterns of gene presence and absence. The selected genes had to be present in less than ten fungal species, could not have homologs in other eukaryotic species, and

needed to have homologs in at least 30 prokaryotic genomes to ensure the directionality of the transfer. There were 713 genes detected that fulfilled those requirements. The genes were grouped into families, and the phylogenetic trees for each family were reconstructed. These trees allowed the splitting of HGT events into 235 monophyletic transference events. They also permitted the mapping of each event to the fungal species tree assuming that the transference had occurred at the last common ancestor of all the fungal species that contained the given transfer. Surprisingly, the events were not evenly distributed across the whole fungal species tree. Roughly two thirds of the events mapped in the Pezizomycotina, one of the main groups of Ascomycotina fungi. In contrast, only 13 events could be mapped to the Saccharomycotina species, representing less than 6 % of the events. While most Pezizomycotina species sequenced at the moment had at least one HGT event, not all of them were equally affected. Groups of species such as *Fusarium* and *Aspergillus* seemed to be particularly prone to receiving transferred genes (*Fusarium oxysporum* was involved in 63 HGT events, while *Aspergillus flavus* was involved in 45). One of the most interesting cases detected during the analysis was that of the arsenic resistance pathway. We have chosen this particular example to illustrate the use of phylogenomics in the study of the role of HGT in the evolution of particular gene families or pathways.

C. Rampant Horizontal Gene Transfer in the Metabolism of Arsenate in Fungi

Arsenic is a toxic compound that is often found as arsenate in soils and water. Because of its similarity to phosphate, the phosphate pumps transport it into the cell, where it accumulates unless the organism has the means to detoxify it. One of the most common ways of arsenate resistance consists of pumping this compound out of the cell after it is reduced from arsenate to arsenite by a specific arsenate reductase enzyme (Rosen 2002). This enzymatic activity has appeared at least three independent times during evolution, twice in prokaryotes and once in eukaryotes. These three protein families are completely unrelated and possess different primary sequences, secondary structures, kinetics, binding sites, and cofactors. In the large-scale phylogenomic analysis mentioned, we identified two HGT events involving the transference of an arsenate reductase protein from prokaryotic origin to two different fungal species: *Yarrowia lipolytica* and *R. oryzae*. We expanded the search of putative transfers in the three arsenate reductase families to other fungal genomes to reveal the overall picture of the evolution of this resistance pathway in fungi. Of the 202 fungi considered, 83 % contained the eukaryotic arsenate reductase (ARR2), which is the expected system to be found in fungi. More surprising was the finding of seven and ten homologs to the prokaryotic arsenate reductases (arsC-gsh and arsC-trx), respectively. Surprisingly, few Saccharomycotina species seemed to contain an ortholog to the ARR2 gene (15.6 %), indicating that the particular arsenic resistance system present in *S. cerevisiae* is nevertheless not widespread in this group of species. Phylogenetic trees reconstructed on the arsC-gsh and arsC-trx protein families, including fungal and prokaryotic homologs, showed that arsC-gsh was involved in at least four independent transference events in Saccharomycotina and early diverging fungi, while the transference of arsC-trx was the result of a single, monophyletic event in a group of Pezizomycotina and also the early diverging fungi (Fig. 3.4). The observed patched distribution is suggestive of multiple transferences between fungi following the prokaryotic gene transfer. Alternatively, a large number of losses would need to be invoked to obtain the same pattern. For instance, a prokaryotic arsC-trx copy can be found in the Pezizomycotina *Metarhizium robertsii*, while none of the other nearly 100 Pezizomycotina species considered contains this gene. Strikingly, several species contain more than one copy of the protein belonging to different families. This is the case for *Aspergillus fumigatus*, which contains both a copy of the eukaryotic ARR2 protein and a copy of arsC-trx, or for *Mortierella verticillioides*, which harbors a member of each of the prokaryotic arsenate reductase genes.

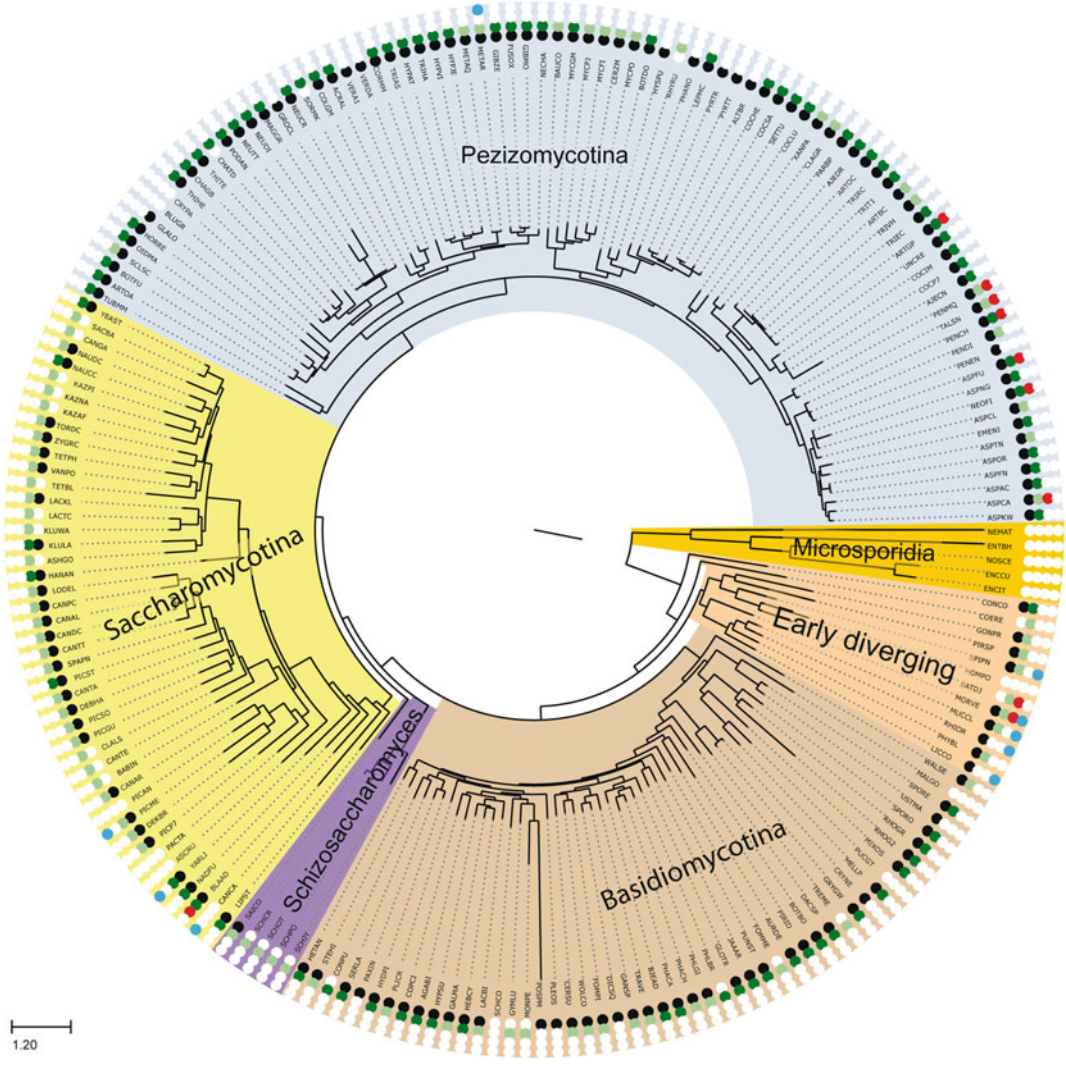


Fig. 3.4. Fungal species tree showing the presence/absence of arsenate reductase proteins. Different fungal groups are shown as different color shades across the tree. Absences are marked with *white dots*. The *four dots* next to each species code represent from the inner to the outer side: presence of the arsenite transporter (*black*), presence of the eukaryotic ARR2 (*green*), presence of the prokaryotic arsC-trx (*red*), and presence of arsC-gsh (*blue*). In the case of the eukaryotic ARR2, orthologs to this gene are shown in *dark green*, while paralogs are shown in *light green*

VI. Comparative Analyses to Understand Fungal Biology

Comparison of genomes from various fungal species can be instrumental to gain insight into the biology of a particular organism. Basically, two complementary approaches can be used. One is based on finding similarities when comparing the genetic complement of two or more species that share a particular phenotypic trait because the common genetic tool kit is expected to constitute the basis of common traits. Conversely, the use of dissimilarities is also informative because phenotypic traits that are distinctive of a particular species

or group are likely to result from genetic differences that are restricted to the corresponding genomes. The use of such approaches is more powerful when used in combination and when based on the analyses of a large number of carefully chosen genomes. Because similarities can be the result of simple phylogenetic proximity, comparisons that are the most informative based on finding similarities in distant species that share a trait or on finding differences in closely related species that differ in a particular phenotype. If the species compared are too distant, however, one may face the risk that the shared trait is the result of a completely distinct molecular basis (i.e., the shared trait is analogous rather than homologous), rendering the comparative approach unsuccessful. Finally, these differences and similarities can be searched for at various levels, from the comparison of residue changes in molecular sequences of proteins or nucleic acids to the study of patterns of presence and absence of genes or complete pathways across species. The choice of the appropriate level of analysis will depend on the biological trait under investigation and on the set of species available.

A recent study focused on the **origin of lignin decomposition** illustrates this approach well (Floudas et al. 2012). The presence of lignin is one of the main reasons why wood is highly resistant to decay. The only organisms capable of substantial lignin decomposition are white-rot fungi in the Agaricomycetes. To shed light on the origins of the important trait of lignin decomposition, the authors of this study compared 31 fungal genomes, 12 of which were generated in the context of that study. The set of compared species contained not only lignin-degrading Agaromycotina but also species from the same group that did not possess this trait, such as lignin-non-degrading brown-rot and ectomycorrhizal species. Their results clearly showed an expansion of lignin-degrading peroxidases in the lineage leading to the ancestor of the Agaricomycetes. This indicates that the common ancestor of the studied species possessed this trait, which was subsequently lost multiple times in parallel in the nondegrading lineages by independent contraction of families of this class of enzymes.

VII. Comparative-Genomics-Based Prediction of Protein Function

Besides using phylogenomics to identify possible genes associated to a given trait, one may, conversely, apply it to the task of deriving specific hypotheses for the function of a given uncharacterized gene. This approach is known as comparative-genomics-based prediction of protein function (Gabaldón 2008a), and the methods are also collectively known as "guilty by association." Basically, all such methods are based on the underlying hypothesis that genes that are functionally related are subjected to similar evolutionary pressures. Hence, the detection of genes following common patterns of evolution, particularly if these are rare, points to the common involvement of these genes in a given biological process. Several approaches can be used, each based in the detection of a different type of evolutionary event, and we describe the main ones in increasing order of subtleness. Perhaps one of the events that is most powerful in predicting specific associations is the detection of **gene fusion or fission events**, that is, the finding of two genes that form a single unit in only a fraction of the genomes where they are present. Such events are usually suggestive of a close physical interaction of the two fused proteins, such as enzymes of the same pathway or subunits of a molecular complex.

Physical proximity of the genes along a chromosome is a somewhat weaker association but can nevertheless be informative of a possible functional relationship. The most extreme case occurs in bacteria in which operons define groups of genes that are coordinately transcribed in polycistronic operons. Although this level of organization is not present as such in eukaryotes, there is accumulating evidence that some higher levels of chromosomal organization do affect the coordinated expression of genes in physical proximity. Such levels range the organization of the genome in chromosomal "territories" in which particular sets of transcription factors preferentially bind to the coordinated regulation of genes in spatial proximity when the three-dimensional struc-

Table 3.1. Useful tools for the phylogenomic analysis of fungal genomes

Resource	Description	URL	Reference
phylomeDB	Phylome repository	phylomedb.org	Huerta-Cepas et al. (2011)
ETE	Tree analysis software	ete.cgenomics.org	Huerta-Cepas et al. (2010)
DupTree	Supertree reconstruction software	genome.cs.iastate.edu/CBL/DupTree/	Wehe et al. (2008)
CLANN	Supertree reconstruction software	bioinf.nuim.ie/clann/	Creevey and McInerney (2009)
phyML	Tree reconstruction software	www.atgc-montpellier.fr/phyml/	Guindon et al. (2010)
RAxML	Tree reconstruction software	sco.h-its.org/exelixis/software.html	Stamatakis et al. (2008)
FastTree	Tree reconstruction software	www.microbesonline.org/fasttree/	Price et al. (2009)
MUSCLE	Alignment reconstruction	www.drive5.com/muscle/	Edgar (2004)
MAFFT	Alignment reconstruction	mafft.cbrc.jp/alignment/software/	Katoh et al. (2009)
T-Coffee	Alignment reconstruction	www.tcoffee.org/	Notredame et al. (2000)
Gblocks	Alignment trimming	molevol.cmima.csic.es/castresana/Gblocks.html	Castresana (2000)
trimAl	Alignment trimming	trimal.cgenomics.org/	Capella-Gutierrez et al. (2009)
metaPhOrs	Orthology database	orthology.phylomedb.org	Pryszcz et al. (2011)

ture of the chromosomes within the nucleus is considered (Simonis et al. 2006; Berger et al. 2008; Homouz and Kudlicki 2013).

The following level of genomic context association corresponds to the consideration of the specific set of genomes in which a particular gene family can be found. This footprint, known as the **phylogenetic profile**, can be informative of function if a limited set of families shares a particular profile. A classical example is subunits of the first complex of the electron transport chain (complex I, or NADH/ubiquinone oxidoreductase), which has been lost, independently, in at least three fungal lineages – Microsporidia, *Schizosaccharomyces*, and *Saccharomyces* lineages (Marcet-Houben et al. 2009) – as well as in other eukaryotic lineages (Gabaldón et al. 2005). This profile is shared by most known complex I subunits, and the discovery of additional gene families sharing this particular phylogenetic profile led to the discovery of novel complex I assembly factors (Gabaldón and Huynen 2005).

Finally, a variety of methods explore the coevolution among gene families by inspecting similarities at the level of gene trees and multiple-sequence alignments (de Juan et al. 2013). The rationale is that proteins participating in the same biological process will be likely subject to similar evolutionary forces. Phylogenomics, through the identification of reliable orthology and paralogy relationships across a set of genomes and through the reconstruction of large-scale phylogenetics datasets, can be of great help in the application of such methods as it provides a context in which to place the appearance/loss of genes. Two distinguishing profiles occurring in two closely related species are likely the results of a common event, while this same profile in two distantly related species will lead to parallel events that led to the same result.

VIII. Concluding Remarks

Phylogenomics is a relatively young field, which nevertheless has already proved useful in transforming increasing amounts of genomic data into biological knowledge. The number of questions that can be addressed through a phylogenomic approach is growing, in parallel with the larger availability and diversity of genomic data. Similarly, new methods and approaches are still being developed with the aim to fulfill specific needs or analyze certain types of data. Here, we provided a brief overview of the main approaches that are currently used in the phylogenomic analyses of fungal genomes. Pointers to some of the tools used in the examples described are provided in Table 3.1. As we

have seen here, the phylogenomic analysis of fungal genomes has served to shed light on the evolution of the main fungal lineages and their traits, and it is emerging as a powerful tool to decode the biological and evolutionary information carried by sequences. As the number and diversity of sequenced fungal species grow, the accuracy, resolution, and utility of phylogenomics methods is expected to increase. We anticipate that phylogenomics analyses will be included more often among the standard genome analysis pipeline, as well as on subsequent analyses of published genomic datasets. In addition, one should not underestimate the challenges that the analysis of increasingly large datasets will impose. We should start working on anticipating the needs for the analysis and visualization of large phylogenomic datasets.

Acknowledgments We wish to thank all members of the Gabaldón group. Group research of T.G. is funded in part by a grant from the Spanish Ministry of Economy and Competitiveness (BIO2012-37161), a grant from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC (Grant Agreement n. ERC-2012-StG-310325), and a Grant from the Qatar National Research Fund grant (NPRP 5-298-3-086).

References

Altincicek B, Kovacs JL, Gerardo NM (2011) Horizontally transferred fungal carotenoid genes in the two-spotted spider mite *Tetranychus urticae*. Biol Lett 8:253–257

Arvestad L, Berglund AC, Lagergren J, Sennblad B (2003) Bayesian gene/species tree reconciliation and orthology analysis using MCMC. Bioinformatics 19(Suppl 1):I7–I15

Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290:972–977

Berger AB, Cabal GG, Fabre E, Duong T, Buc H, Nehrbass U, Olivo-Marin JC, Gadal O, Zimmer C (2008) High-resolution statistical mapping reveals gene territories in live yeast. Nat Methods 5:1031–1037

Berglund AC, Sjolund E, Ostlund G, Sonnhammer EL (2008) InParanoid 6: eukaryotic ortholog clusters with inparalogs. Nucleic Acids Res 36:D263–D266

Bininda-Emonds OR (2005) Supertree construction in the genomic age. Methods Enzymol 395:745–757

Bullerwell CE, Leigh J, Forget L, Lang BF (2003) A comparison of three fission yeast mitochondrial genomes. Nucleic Acids Res 31:759–768

Bushley KE, Turgeon BG (2010) Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evol Biol 10:26

Byrne KP, Wolfe KH (2006) Visualizing syntenic relationships among the hemiascomycetes with the Yeast Gene Order Browser. Nucleic Acids Res 34: D452–D455

Campbell MA, Rokas A, Slot JC (2012) Horizontal transfer and death of a fungal secondary metabolic gene cluster. Genome Biol Evol 4:289–293

Capella-Gutierrez S, Silla-Martinez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973

Capella-Gutierrez S, Marcet-Houben M, Gabaldon T (2012) Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. BMC Biol 10:47

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552

Cavalier-Smith T (1983) A six-kingdom classification and a unified phylogeny. In: Schenk HEA, Schwemmler W (eds) Endocytobiology II. de Gruyter, Berlin, pp 1027–1034

Creevey CJ, McInerney JO (2009) Trees from trees: construction of phylogenetic supertrees using CLANN. Methods Mol Biol 537:139–161

Cuomo CA, Desjardins CA, Bakowski MA, Goldberg J, Ma AT, Becnel JJ, Didier ES, Fan L, Heiman DI, Levin JZ, Young S, Zeng Q, Troemel ER (2012) Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Genome Res 22:2478–2488

de Juan D, Pazos F, Valencia A (2013) Emerging methods in protein co-evolution. Nat Rev Genet 14:249–261

Delsuc F, Brinkmann H, Philippe H (2005) Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6:361–375

Doolittle WF (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113

Eirin-Lopez JM, Rebordinos L, Rooney AP, Rozas J (2012) The birth-and-death evolution of multigene families revisited. Genome Dyn 7:170–196

Eisen JA, Fraser CM (2003) Phylogenomics: intersection of evolution and genomics. Science 300:1706–1707

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst Zool 19:99–113

Fitzpatrick DA, Logue ME, Stajich JE, Butler G (2006) A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. BMC Evol Biol 6:99

Fitzpatrick DA, O'Gaora P, Byrne KP, Butler G (2010) Analysis of gene evolution and metabolic pathways using the Candida Gene Order Browser. BMC Genomics 11:290

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D, Gorecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kues U, Kumar TK, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A, Rokas A, Ruiz-Duenas FJ, Sabat G, Salamov A, Samejima M, Schmutz J, Slot JC, St John F, Stenlid J, Sun H, Sun S, Syed K, Tsang A, Wiebenga A, Young D, Pisabarro A, Eastwood DC, Martin F, Cullen D, Grigoriev IV, Hibbett DS (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719

Fraser JA, Huang JC, Pukkila-Worley R, Alspaugh JA, Mitchell TG, Heitman J (2005) Chromosomal translocation and segmental duplication in *Cryptococcus neoformans*. Eukaryot Cell 4:401–406

Gabaldón T (2005) Evolution of proteins and proteomes, a phylogenetics approach. Evol Bioinform Online 1:51–56

Gabaldón T (2008a) Comparative genomics-based prediction of protein function. Methods Mol Biol 439:387–401

Gabaldón T (2008b) Large-scale assignment of orthology: back to phylogenetics? Genome Biol 9:235

Gabaldón T, Huynen MA (2005) Lineage-specific gene loss following mitochondrial endosymbiosis and its potential for function prediction in eukaryotes. Bioinformatics 21(Suppl 2):ii144–ii150

Gabaldón T, Koonin EV (2013) Functional and evolutionary implications of gene orthology. Nat Rev Genet 14:360–366

Gabaldón T, Rainey D, Huynen MA (2005) Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (complex I). J Mol Biol 348:857–870

Gabaldón T, Marcet-Houben M, Huerta-Cepas J (2008) Reconstruction and analysis of large-scale phylogenetic data: challenges and opportunities. In: Russe A (ed) Computational biology: new research. Nova Science, New York, pp 129–146

Gill EE, Fast NM (2006) Assessing the microsporidia-fungi relationship: combined phylogenetic analysis of eight genes. Gene 375:103–109

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274(546):563–567

Gojkovic Z, Knecht W, Zameitat E, Warneboldt J, Coutelis JB, Pynyaha Y, Neuveglise C, Moller K, Loffler M, Piskur J (2004) Horizontal gene transfer promoted evolution of the ability to propagate under anaerobic conditions in yeasts. Mol Genet Genomics 271:387–393

Gribaldo S, Philippe H (2002) Ancient phylogenetic relationships. Theor Popul Biol 61:391–408

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321

Homouz D, Kudlicki AS (2013) The 3D organization of the yeast genome correlates with co-expression and reflects functional relations between genes. PLoS One 8:e54699

Huerta-Cepas J, Gabaldón T (2010) Assigning duplication events to relative temporal scales in genome-wide studies. Bioinformatics 27:38–45

Huerta-Cepas J, Dopazo H, Dopazo J, Gabaldón T (2007) The human phylome. Genome Biol 8:R109

Huerta-Cepas J, Dopazo J, Gabaldón T (2010) ETE: a python Environment for Tree Exploration. BMC Bioinformatics 11:24

Huerta-Cepas J, Capella-Gutierrez S, Pryszcz LP, Denisov I, Kormes D, Marcet-Houben M, Gabaldon T (2011) PhylomeDB v3.0: an expanding repository of genome-wide collections of trees, alignments and phylogeny-based orthology and paralogy predictions. Nucleic Acids Res 39:D556–D560

Huynen MA, Bork P (1998) Measuring genome evolution. Proc Natl Acad Sci U S A 95:5849–5856

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schussler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lucking R, Budel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006) Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature 443:818–822

Jones MD, Forn I, Gadelha C, Egan MJ, Bass D, Massana R, Richards TA (2011) Discovery of novel intermediate forms redefines the fungal tree of life. Nature 474:200–203

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453

Katoh K, Asimenos G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. Methods Mol Biol 537:39–64

Kaur R, Ma B, Cormack BP (2007) A family of glycosylphosphatidylinositol-linked aspartyl proteases is required for virulence of *Candida glabrata*. Proc Natl Acad Sci U S A 104:7628–7633

Keeling PJ (2003) Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet Biol 38:298–309

Keeling PJ, Fast NM (2002) Microsporidia: biology and evolution of highly reduced intracellular parasites. Annu Rev Microbiol 56:93–116

Khaldi N, Wolfe KH (2011) Evolutionary origins of the fumonisin secondary metabolite gene cluster in *Fusarium verticillioides* and *Aspergillus niger*. Int J Evol Biol 2011:423821

Kovalchuk A, Driessen AJ (2010) Phylogenetic analysis of fungal ABC transporters. BMC Genomics 11:177

Kristensen DM, Wolf YI, Mushegian AR, Koonin EV (2011) Computational methods for gene orthology inference. Brief Bioinform 12:379–391

Kriventseva EV, Rahman N, Espinosa O, Zdobnov EM (2008) OrthoDB: the hierarchical catalog of eukaryotic orthologs. Nucleic Acids Res 36:D271–D275

Landan G, Graur D (2007) Heads or tails: a simple reliability check for multiple sequence alignments. Mol Biol Evol 24:1380–1383

Li L, Stoeckert CJ Jr, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13:2178–2189

Liu Y, Leigh JW, Brinkmann H, Cushion MT, Rodriguez-Ezpeleta N, Philippe H, Lang BF (2009a) Phylogenomic analyses support the monophyly of Taphrinomycotina, including Schizosaccharomyces fission yeasts. Mol Biol Evol 26:27–34

Liu Y, Steenkamp ET, Brinkmann H, Forget L, Philippe H, Lang BF (2009b) Phylogenomic analyses predict sistergroup relationship of nucleariids and fungi and paraphyly of zygomycetes with significant support. BMC Evol Biol 9:272

Ma LJ, Ibrahim AS, Skory C, Grabherr MG, Burger G, Butler M, Elias M, Idnurm A, Lang BF, Sone T, Abe A, Calvo SE, Corrochano LM, Engels R, Fu J, Hansberg W, Kim JM, Kodira CD, Koehrsen MJ, Liu B, Miranda-Saavedra D, O'Leary S, Ortiz-Castellanos L, Poulter R, Rodriguez-Romero J, Ruiz-Herrera J, Shen YQ, Zeng Q, Galagan J, Birren BW, Cuomo CA, Wickes BL (2009) Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. PLoS Genet 5: e1000549

Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie X, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RA, Chapman S, Coulson R, Coutinho PM, Danchin EG, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature 464:367–373

Marcet-Houben M, Gabaldon T (2009) The tree versus the forest: the fungal tree of life and the topological diversity within the yeast phylome. PLoS One 4: e4357

Marcet-Houben M, Gabaldon T (2010) Acquisition of prokaryotic genes by fungal genomes. Trends Genet 26:5–8

Marcet-Houben M, Marceddu G, Gabaldón T (2009) Phylogenomics of the oxidative phosphorylation in fungi reveals extensive gene duplication followed by functional divergence. BMC Evol Biol 9:295

Merhej V, Raoult D (2012) Rhizome of life, catastrophes, sequence exchanges, gene creations, and giant viruses: how microbial genomics challenges Darwin. Front Cell Infect Microbiol 2:113

Moran NA, Jarvik T (2010) Lateral transfer of genes from fungi underlies carotenoid production in aphids. Science 328:624–627

Muller H, Thierry A, Coppee JY, Gouyette C, Hennequin C, Sismeiro O, Talla E, Dujon B, Fairhead C (2009) Genomic polymorphism in the population of *Candida glabrata*: gene copy-number variation and chromosomal translocations. Fungal Genet Biol 46:264–276

Notredame C (2007) Recent evolutions of multiple sequence alignment algorithms. PLoS Comput Biol 3:e123

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205–217

Novakova E, Moran NA (2011) Diversification of genes for carotenoid biosynthesis in aphids following an ancient transfer from a fungus. Mol Biol Evol 29:313–323

Ohno S (1970) Evolution by gene duplication. Allen & Unwin, London

Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, Arnold R, Rattei T, Letunic I, Doerks T, Jensen LJ, von Mering C, Bork P (2011) eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Res 40:D284–D289

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650

Pryszcz L, Huerta-Cepas J, Gabaldón T (2011) MetaPhOrs: orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. Nucleic Acids Res 39:e32

Rosen BP (2002) Biochemistry of arsenic detoxification. FEBS Lett 529:86–92

Sicheritz-Ponten T, Andersson SG (2001) A phylogenomic approach to microbial evolution. Nucleic Acids Res 29:545–552

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 38:1348–1354

Slot JC, Rokas A (2011) Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. Curr Biol 21:134–139

Snel B, Huynen MA, Dutilh BE (2005) Genome trees and the nature of genome evolution. Annu Rev Microbiol 59:191–209

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol 57:758–771

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 56:564–577

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135

Van de Peer Y, Ben Ali A, Meyer A (2000) Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi. Gene 246:1–8

Wapinski I, Pfeffer A, Friedman N, Regev A (2007) Automatic genome-wide reconstruction of phylogenetic gene trees. Bioinformatics 23:i549–i558

Wehe A, Bansal MS, Burleigh JG, Eulenstein O (2008) DupTree: a program for large-scale phylogenetic analyses using gene tree parsimony. Bioinformatics 24:1540–1541

Wolfe KH, Shields DC (1997) Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387:708–713

4 Genome Data Drives Change at Culture Collections

Kevin McCluskey[1], Aric Wiest[1], Kyria Boundy-Mills[2]

CONTENTS

I. Introduction

Research with fungi has been collaborative for many years and has depended on the open sharing of resources either via public collections or via peer-to-peer exchanges. Formal culture collections, such as the Fungal Genetics Stock Center (McCluskey et al. 2010), the Centraalbureau voor Schimmelcultures (Crous et al. 2009), and the U.S. Department of Agriculture (USDA) Agriculture Research Service (ARS) Northern Regional Research Laboratory (NRRL) have facilitated this collaborative community and have ensured that research materials from one era are available as we transition to a modern postgenomics era. Culture collections have been identified as a key infrastructure element to promote the transition to the bioeconomy (Janssens et al. 2010). Collections have leveraged the work of decades of researchers to provide ready-characterized material suitable for whole-genome analysis. Researchers at these collections are in the enviable position of having a veritable treasure trove of materials in their freezers. However, despite this tremendous wealth of biological materials, living collections are facing budget shortfalls that limit their ability to plan for future developments in genome sequence technology, in bioinformatics, and in support of the coming bioeconomy (McCluskey 2013).

II. Culture Collections

Many culture collections have held and distributed taxonomically and genetically characterized fungi for over five decades (Daniel and Prasad 2010; Dugan et al. 2011; Boundy-Mills 2012; Smith 2012). Among these are many publicly supported and university-supported resources (Stern 2004). Culture collections, more formally known as **ex situ microbial germplasm repositories**, have diverse missions, usually related to their funding source. Among the different categories of culture collections are **type collections**, **patent collections**, **genetic stock centers**, and **reference material collections** (McCluskey 2013). Each different category of collection has a different purpose and

[1]Fungal Genetics Stock Center, University of Missouri–Kansas City, Kansas City, MO 64110, USA; e-mail: mccluskeyk@umkc.edu;
[2]Phaff Yeast Culture Collection, University of California, Davis, CA 95616, USA

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

serves a unique scientific community. The most similar categories are the type, patent, and reference collections, and in some instances, these roles are fulfilled by the same institutions. Type collections hold strains used in formal taxonomic descriptions and, for microbes, replace the role of a herbarium or natural history museum (McNeill et al. 2012). While some fungi are deposited into mycological herbaria, others have no significant morphology or are persistently microscopic and hence are not easily included in a mycological herbarium. Reference strain collections specialize in materials used in immunological or other diagnostic tests (ISO Guide 2009). Reference strains can be held in type collections, and this combination is common. Patent collections are like type or reference collections but were established by international treaty to ensure that microbes used in patented processes were available (World Intellectual Property Organization 1980). All of these categories of collections share the characteristic that they may include hugely diverse holdings (Table 4.1). While most collections have grown slowly over many years, growth in holdings exhibits two different modes. Isolates accessioned by the research activities of the collection staff contribute to a slow incremental growth and for some collections can be on the scale of 200–500 isolates per year. This level of growth requires significant support as much of the cost of having isolates in collections is required at the time of accession, and maintaining isolates in collections is straightforward once strains are accessioned, tested, and preserved for long-term maintenance. Much of collection growth, however, can be by large-scale deposits associated with retirement, loss of funding by smaller collections, or collaborative projects with primary research projects outside the collection. For the Fungal Genetics Stock Center (FGSC), over 10,000 *Neurospora* strains were deposited by the functional genomics consortium (Dunlap et al. 2007), 3,150 by David D. Perkins of Stanford University (Turner et al. 2001), and ten additional researchers have deposited more than 100 strains. By way of contrast, 46 individuals have deposited exactly 1 strain, and 211 people have deposited fewer than 100 strains.

Culture collections have been undergoing significant change in the last decade, and much of this is associated with the recognition that the materials they hold are foundational to modern biotechnological advancement. With the publications of best practice guidelines for the operation of culture collections by the Organization for Economic Cooperation and Development (OECD 2007) and for operation of biological repositories by the International Society for Biological and Environmental Repositories (Campbell et al. 2012), much of this change manifests itself as increased professionalism. Similarly, collections now operate in a more complicated regulatory environment, and the increased costs of these regulations fall on collections at a time when collection budgets are falling. Likewise, the impact on genetic engineering and synthetic biological processes have an impact on collections disproportionately. At the FGSC, for example, the vast majority of strains distributed since the advent of systematic gene deletion mutants in filamentous fungi (Ninomiya et al. 2004) have been genetically modified organisms (GMOs). Certainly, not all collections face these challenges, but all collections are faced with increasing regulation on shipping human, animal, or plant pathogens, and these regulations are discontinuous and imposed by diverse regulatory agencies, such as the International Air Transport Authority (IATA), the International Union of Postmasters, and myriad national agriculture and health ministries.

A. Culture Collection Certifications

Many ex situ microbial germplasm collections have obtained external certifications, although no specific certification exists for living microbe collections. The most relevant certifications are the International Organization for Standardization (ISO) 9001:2008 and the ISO Guide 34:2009 (the date refer to the revision under which certification was obtained) for reference material. Many collections seek **ISO certification**, and ISO9001:2008 is the most common. It deals fundamentally with data management, and this certification ensures

Table 4.1. Prominent living fungal collections holding genome and reference strains

Collection name/Acronym	Scope	Support	Website
American Type Culture Collection (ATCC)	Broad/patent	Private nonprofit/ International Depository Authority (IDA)	www.atcc.org
Centraalbureau voor Schimmelcultures (CBS)	Broad	Royal Netherlands Academy of Arts and Sciences	www.cbs.knaw.nl
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ)	Broad/patent	The Leibniz Institute	www.dsmz.de
National Institute of Technology and Evaluation National Biological Resource Center (NITE BRC)	Broad/patent	Government of Japan	www.nbrc.nite.go.jp/e/
Japan Collection of Microorganisms (JCM)	Broad	RIKEN BioResource Center	www.jcm.riken.jp/
Fungal Genetics Stock Center (FGSC)	Genetic	U.S. National Science Foundation	www.fgsc.net
USDA ARS Culture Collection (NRRL)	Broad	U.S. Department of Agriculture/IDA	nrrl.ncaur.usda.gov/
University of Alberta Microfungus Collection and Herbarium	Broad/medical	University of Alberta	http://www.uamh.devonian.ualberta.ca/
Mycology Culture Collection (ACH)	Medical Mycology	University of Adelaide	www.mycology.adelaide.edu.au
Austrian Center of Biological Resources and Applied Mycology (ACBR)	General	University of Biological Resources and Applied Life Sciences	www.acbr-database.at/
Belgian Coordinated Collections of Microorganisms (BCCM)	Broad/network	Belgian Science Policy Office	bccm.belspo.be/
Colecao de Culturas de Fungos Filamentosos (CCFF)	Filamentous fungi	Fundação Oswaldo Cruz	ccff.fiocruz.br/
China General Microbiological Culture Collection Center (CGMCC)	Broad	Institute of Microbiology of the Chinese Academy of Sciences	www.cgmcc.net/
IBT Culture Collection of Fungi (IBT)	Industrial/ taxonomic	Technical University of Denmark	www.biocentrum.dtu.dk/
VTT Culture Collection (VTT)	General	VTT Technical Research Centre of Finland	culturecollection.vtt.fi/
Collection de L'Institut Pasteur (CIP)	Medical	Institut Pasteur	www.crbip.pasteur.fr
National Fungal Culture Collection of India	Agricultural	MACS' Agharkar Research Institute	http://nfcci.dinpl.com/
National Collection of Industrial Microorganisms (NCIM)	Industrial	National Chemical Laboratory (CSIR)	www.ncl-india.org/files/NCIM/
Industrial Yeasts Collection (DBVPG)	Industrial	Italian Ministry of Commerce/IDA	www.agr.unipg.it/dbvpg/
Center for Fungal Genetic Resources (CFGR)	Genetic	Seoul National University	genebank.snu.ac.kr
Korean Collection for Type Cultures (KCTC)	Type cultures	Korea Research Institute of Bioscience and Biotechnology	kctc.kribb.re.kr/
Microbial Strain Collection of Latvia (MSCL)	General	University of Latvia/IDA	mikro.daba.lv
Forest Research Culture Collection (NZFS)	Forest pathogens	New Zealand Crown Research Institute	www.foresthealth.co.nz
Micoteca da Universidade do Minho (MUM)	Fungi	Universidade do Minho	www.micoteca.deb.uminho.pt

(continued)

Table 4.1. (continued)

Collection name/Acronym	Scope	Support	Website
Russian Collection of Agricultural Microorganisms (RCAM)	Agriculture	Russian Federation	www.arriam.spb.ru/
All-Russian Collection of Microorganisms (VKM)	General	Russian Academy of Sciences/IDA	www.vkm.ru/
Coleccion Espanola de Cultivos Tipo (CECT)	General	Universidad de Valencia	www.cect.org/
Fungal Cultures University of Goteborg (FCUG)	General	University of Goteborg	www2.botany.gu.se/database/FCUG.html
Bioresource Collection and Research Center (BCRC)	General	Food Industry Research and Development Institute	www.bcrc.firdi.org.tw
BIOTEC Culture Collection (BCC)	General	National Center for Genetic Engineering and Biotechnology	www.biotec.or.th/bcc/
CABI Genetic Resource Collection (IMI)	General	CAB International/IDA	www.cabi.org/
National Collections of Industrial Food and Marine Bacteria (NCIMB)	Food bacteriology	NCIMB Ltd./IDA	www.ncimb.co.uk
Phaff Yeast Culture Collection (UCD-FST)	Yeast	University of California	phaffcollection.ucdavis.edu

that the collection has implemented best practices for database, client, and quality control management. Similarly, Guide 34 specifies overall requirements that a producer must meet to be certified to produce reference material. In descriptions of the guide, the ISO emphasizes that it is not intended to vouch for the material, but rather that the collection is competent to produce the material. Together, and in the context of ISO/IEC (International Electrotechnical Commission) 17025, which outlines general requirements for testing and calibration laboratories, these standards are valuable for demonstrating that a collection has the appropriate qualifications, management structures, and infrastructure to produce materials for authoritative use in a variety of validation, testing, and reference purposes. While many of the largest international living microbe repositories have obtained independent certification according to ISO norms (Table 4.2), these certifications are difficult to obtain for smaller culture collections. It is important to note, however, that the lack of certification does not imply that the materials in the smaller collections are not fit for use.

Similarly, many collections are coordinating their activities through formal and ad hoc networks, and these have the diverse benefits of pooling expertise, sharing materials for reliability and availability, providing a unified voice for collections, and in the extreme case of formal collaborative networks, of sharing quality management practices and even certification (Stackebrandt 2010).

Table 4.2. Fungal collections with certification

Collection[a]	Independent certification
ATCC	ISO9001:2008, ISO/IEC 17025:2005
CBS	ISO9001:2008
DSMZ	ISO9001:2008
NBRC	ISO9001:2008
CECT	ISO9001:2008
BCRC	ISO9001:2008
NCIMB	ISO9001:2008 current Good Manufacturing Practices (cGMP)
CIP	ISO9001:2005
BCCM	ISO9001:2008
VTT	ISO9001:2008

[a]Acronyms as in Table 4.1

B. The Convention on Biological Diversity

As ex situ microbial germplasm repositories are becoming increasingly professional, outside events are also driving change at collections. Among these are the globalization and professionalization of science and the desirability of having certified reference materials for

taxonomic, diagnostic, and molecular genetic research purposes. Similarly, the drive to promote equitable sharing of the world's biological resources has led to the establishment of the formal international treaty known as the Convention on Biological Diversity (CBD, www.cbd.int). The CBD explicitly mandates in Article 9 that every party to the convention establish a formal ex situ microbial germplasm repository to **conserve the microbial biodiversity** identified through surveys and other bioprospecting endeavors. One significant aspect of the impact of a genome sequence that remains to be evaluated is the extent to which publication of whole-genome sequence data relate to the conditions of the CBD with regard to the distribution of genetic resources, defined in the text of the CBD as "any material of plant, animal, microbial or other origin containing functional units of heredity" (CBD, Article 2). In addition, genetic resources are described as "genetic material of actual or potential value." Whether the publication of a whole-genome sequence and subsequent synthetic preparation of DNA encoded in that information constitutes exchange of genetic resources or access to technology or whether this type of exchange is not related to the CBD has yet to be explored. Similarly, the impact of synthetic biology on the activities of culture collections has yet to mature. While culture collections hold tremendous diversity of organisms, they do not hold every living species, so the ability to use metadata to recapitulate interesting genes by de novo chemical DNA synthesis allows the de facto export of genes via transfer of electronic data. Synthetic biology labs are not culture collections, and their approach of isolating genetic elements for manipulation in simple model organisms means that they can preserve genetic elements, even from taxonomically diverse sources, in simple model organism hosts (Purnick and Weiss 2009). Because the technological framework around the original publication of the CBD in 1992 has shifted, the ultimate relationship between culture collections, genome data storehouses, and synthetic biological practices under the auspices of the CBD remains to be negotiated.

III. Whole-Genome Analysis of Fungi

As the cost of DNA sequencing has made whole-genome analysis of multiple species and even multiple isolates of one species possible, culture collections have become central to the effort to generate whole-genome sequence for strains representing the diversity of life (Grigoriev et al. 2012).

Several such efforts are related to culture collections. The first, known as the **Genomic Encyclopedia of Bacteria and archaea (GEBA)**, led to the elucidation of the whole-genome sequence for nearly 1,000 different bacterial and archael genomes by 2009 (Wu et al. 2009); the pace of sequencing, as well as the informed analysis of the sequence data, has even allowed assembly of unculturable genomes from metagenomic DNA sequence data of a bacterial consortium (Albertsen et al. 2013). Currently, over 3,000 bacterial and archaeal genomes have been sequenced, and this tremendous depth of sequence has enabled novel approaches to understanding bacterial and archaeal phylogenies (Lang et al. 2013). The success of the GEBA project has led to similar efforts by researchers studying phylogeny, population and molecular genetics, and evolutionary biology of fungi, and this project is known as the **1000 Fungal Genomes Project** (http://1000.fungalgenomes.org/home/).

A. Reference Genome Sequence

Among the many ambitious goals of the 1000 Fungal Genomes Project are the sequencing of genomes of organisms that are not amenable to traditional culture and extraction protocols, and these include symbionts, slow-growing organisms, and diploid or dikaryotic fungi (Marmeisse et al. 2013). While these may not enable the kind of complete genomes expected from early studies, the lack of closed contigs where the number of contigs equals the numbers of chromosomes does not impede the use of these incomplete genomes. Indeed, the complement of protein-encoding genes, their specificities, and their association with the biology of the fungus can inform a variety of studies.

Table 4.3. Genome strains in the Fungal Genetics Stock Center (FGSC) and their listing in other collections

Strain number	Characteristics	Alternate numbers[a]
FGSC A1100	*Aspergillus fumigatus*	AF293, ATCC MYA-4609, CBS 101355
FGSC A4	*Aspergillus nidulans*	ATCC 28611, JCM19073
FGSC A1513	*Aspergillus niger*	ATCC MYA-4892, CBS 513.88
FGSC A1121	*Aspergillus niger*	NRRL 3, CBS 120.49
FGSC A1156	*Aspergillus terreus*	JCM19077
FGSC 10389	*Agaricus bisporus*	ATCC MYA-4626, H97
FGSC 10392	*Agaricus bisporus*	ATCC MYA-4627, JB137-S8
FGSC 9003	*Coprinus cinereus*	ATCC 14714, JCM19078
FGSC 9075	*Fusarium graminearum*	ATCC MYA-4620, PH1
FGSC 9935	*Fusarium oxysporum* fsp *lycopersici*	ATCC MYA-4623, NRRL 34936
FGSC 8958	*Magnaporthe grisea*	ATCC MYA-4617, 70-15
FGSC 9013	*Neurospora crassa*	ATCC MYA-4619
FGSC 2489	*Neurospora crassa* 74-OR23-1VA	ATCC MYA-4614
FGSC 8579	*Neurospora discreta*	ATCC MYA-4616, W683
FGSC 2508	*Neurospora tetrasperma*	ATCC MYA-4615, P0657
FGSC 9596	*Nectria haematococca*	ATCC MYA-4622
FGSC 9002	*Phanerochaete chrysosporium*	ATCC MYA-4764, RP-78
FGSC 10004	*Phycomyces blakesleeanus*	NRRL 1555
FGSC 10384	*Podospora anserina*	ATCC MYA-4625, S mat −
FGSC 10383	*Podospora anserina*	ATCC MYA-4624, S mat +
FGSC 9543	*Rhizopus oryzae*	ATCC MYA-4621, RA99880
FGSC 10291	*Rhodotorula graminis*	ATCC MYA-4893, WP1
FGSC 10173	*Stagonospora nodorum*	ATCC 24758, SN15
FGSC 10586	*Trichoderma virens*	ATCC MYA-4894, Tv29-8
FGSC 10136	*Verticillium albo-atrum*	ATCC 201270, VaMs102

This list is not exhaustive and omits classical genetic mutant strains subject to next-generation sequence analysis
[a]Acronyms as in Table 4.1

For example, as genomic data become more accessible, fungi are being evaluated based on the categories of numbers of genes found in their genomes. Martin and colleagues (Martin et al. 2008) demonstrated that fungi that were plant symbionts had fewer genes predicted to encode cell wall-degrading enzymes than comparable saprophytic species.

The data portal for the 1000 Fungal Genomes Project is built on the platform developed by the U.S. Department of Energy Joint Genome Institute (US DOE-JGI), and this includes the integrated fungal genomics resource known as MycoCosm (jgi.doe.gov/fungi) (Grigoriev et al. 2012).

The **Genomes Online Database** lists 65 fungal genome programs, and of these, 49 were of ascomycetes, 9 were basidiomycetes, and 2 were microsporidia (http://www.genomesonline.org/cgi-bin/GOLD/index.cgi, accessed June 12, 2013). This is contrasted with the availability of fungal genome data in the U.S. National Library of Medicine **Short Read Archive (SRA)**, which lists 320 projects for the search term "fungus," 392 projects for "neurospora," and 215 projects for "aspergillus." Admittedly, many of these last projects are transcriptome analyses, but this pattern of large numbers of projects continues as searches are conducted on other common fungal taxa names (such as Fusarium, Schizophyllum, Colletotrichum, or Cochliobolus). Culture collections hold most of the isolates used in these and other genome programs, and a broad goal of collections is to ensure that materials are available, either directly from the host collection or by sharing resources. The FGSC has endeavored both to assemble a collection of reference genome strains and to ensure that these strains are protected from accidental loss by depositing strains with other collections (Table 4.3).

B. Sequence of Multiple Isolates Within One Species

While reference genome data are clearly a valuable resource, much more insight is available

when comparing diverse genomes against a reference. In a preliminary study of the diversity of laboratory isolates of ***Neurospora crassa***, it was found that many strains carried **polymorphisms relative to the reference genome strain** (McCluskey et al. 2011), and this finding was recapitulated in related studies (Pomraning et al. 2011; Nowrousian et al. 2012). The strains that are the subject of these analyses were available because they were deposited in culture collections. The labor required to do the mutant screens that led to the characterization of these, and hundreds of other, anonymous classical genetic mutant loci has been exploited by the ability to rapidly identify the causative mutation in a background of unrelated polymorphisms. The fact that these strains were preserved in culture collections for many decades allows for cross-generational collaboration otherwise impossible without the effort of well-curated collections (Table 4.4). While the large *Neurospora* study identified hundreds of thousands of polymorphisms among 18 strains, many of these polymorphisms were present in 1 strain, which was in a unique genetic background relative to the reference genome strain. In this one comparison, between strain FGSC 2489 and FGSC 821 (the reference and "Emerson" backgrounds, respectively), there were 188,346 single-nucleotide polymorphisms (SNPs) identified in the FGSC 821 sequence. Of these, 112,490 (or nearly 60 %) were noncoding and hence presumably neutral to selection. Of the remaining polymorphisms, 17,193 were nonsynonymous, and 122 were nonsense mutations. The nonsynonymous changes in strain FGSC 821 define unique alleles at 3,512 different loci, and while most Open Reading Frames (ORFs) have no changes or only one change, many ORFs have multiple changes between the reference genome background and the Emerson background. When the 679 ORFs with coding sequence (CDS) indels are also considered, there are a total of 3,575 ORFs that have alleles in FGSC 821 that are unique from the homolog in the reference genome sequence. This amount of diversity in a well-characterized system with active genome defenses (Rountree and Selker 2009) suggests that other systems with less-well-enforced genome defense might generate novel alleles (and hence novel biochemical or regulatory specificities) at a prodigious rate.

Table 4.4. Time between deposit of *Neurospora crassa* strains carrying genetically defined mutations and identification of the mutation by whole-genome analysis

Strain number	Mutant locus	Gap (years)
FGSC 106	*com*	51
FGSC 305	*amyc*	51
FGSC 309	*ti*	51
FGSC 322	*ty-1*	51
FGSC 821	*ts*	50
FGSC 1211	*dot*	46
FGSC 1303	*fi*	45
FGSC 1363	*smco-1*	44
FGSC 2261	*do*	38
FGSC 3114	*Sk-2*	34
FGSC 3246	*fs-n*	34
FGSC 3562	*mb-1*	32
FGSC 3564	*mb-2*	32
FGSC 3566	*mb-3*	32
FGSC 3831	*ff-1*	30
FGSC 3921	*tng*	29
FGSC 7022	*fld*	20
FGSC 7035	*per-1*	20

Because of their importance in industrial production of enzymes, biocontrol (through mycoparastism and other mechanisms), and occasional human facultative pathogenesis, the genomes of several **Trichoderma species** have been sequenced (Table 4.5). A recent publication (Kubicek et al. 2011) described the comparative analysis of two biocontrol species, *T. atroviride* (teleomorph *Hypocrea atroviridis*) and *T. virens* (formerly *Gliocladium virens*, teleomorph *Hypocrea virens*), with *T. reesei* (teleomorph *Hypocrea jecorina*), an industrially important cellulolytic filamentous fungus. This comparison revealed the ancestral nature of many genes involved in mycoparasitism, including those coding for chitinases, chitosanases, and secondary metabolite production, pointing to an early mycoparasitic lifestyle. Other isolates sequenced as biocontrol strains include *T. asperellum* and *T. atroviride*.

Several *T. reesei* genomes have been sequenced, including non- and hypercellulose producers (QM 9136 and

Table 4.5. Trichoderma spp. whole-genome sequencing projects

Project name	Principal investigator	Strain designation(s)[a]
T. reesei RUT C-30	Grigoriev, I (JGI/LBNL)	ATCC 56765, NRRL 11460, NRCC 2906, RUT-C30, VTT-D-86271
T. reesei	Baker, S (PNNL)	ATCC 13631, QM 6a, CBS 383.78, IMI 192654, IMI 45548, T.V. B117
T. reesei RL-P37 reseq.	Gladden, J (JBEI)	NRRL 15709, VTT D-071269
T. reesei MCG77 reseq.	Gladden, J (JBEI)	ATCC 56764, NRRL 11236, CECT 2415, IHEM 5475
T. reesei QMY-1 reseq.	Gladden, J (JBEI)	NRRL 18760, QMY-1, VTT D-071270
T. reesei PC-3–7 reseq.	Gladden, J (JBEI)	ATCC 66589, NRRL 15500
T. reesei PC-1–4 reseq.	Gladden, J (JBEI)	ATCC 66588, NRRL 15499
T. reesei X-31 reseq.	Gladden, J (JBEI)	ATCC 66587, NRRL 15502
T. reesei 25–4 reseq.	Gladden, J (JBEI)	
T. reesei CBS 999.79 reseq.	Baker, S (PNNL)	ATCC 204423, GJS 97–38, CBS 999.97
T. reesei QM 9978 reseq.	Baker, S (PNNL)	NRRL 22832, VTT D-071267
T. reesei QM 9136 reseq.	Baker, S (PNNL)	ATCC 26920, 178, BCRC 31642, CCRC 31642, CDBB 355, DSM 770, IFO 31328, KCTC 6045, NBRC 31328
T. reesei QM 9414 reseq.	Baker, S (PNNL)	ATCC 26921, QM 9414, 3019, CBS 392.92, T.V. B118
T. asperellum	Grigoriev, I (JGI)	ATCC 204424, TS1(c), BBA 70684, CBS 433.97, TR3
T. atroviride	Baker, S (PNNL)	IMI 206040
T. citrinoviride	Grigoriev, I (JGI)	
T. harzianum	Grigoriev, I (JGI)	CBS 226.95, KCTC 16976
T. longibrachiatum	Grigoriev, I (JGI)	ATCC 18648, 13423, BCRC 32895, CBS 816.68, CCRC 32895, IHEM 5747, IMI 352773, KCTC 6379
T. virens Gv29–8	Kenerley, C (Texas A&M University)	ATCC MYA-4894, FGSC 10586 [Tv29–8, Gv29–8]

JBEI US DOE Joint Bioenergy Institute, *JGI/LBNL* US DOE Joint Genome Institute/Lawrence Berkeley National Laboratory, *PNNL* US DOE Pacific Northwest National Laboratory
[a]Acronyms as in Table 4.1

QM 9414, respectively). The genomes of the ethanol- and polyene-resistant RUT C-30 and the L-sorbose-inducible cellulase producers PC-3-7 and X-31 have also been sequenced.

Trichoderma longibrachiatum and *T. citrinoviride* are the only known *Trichoderma* that are opportunistic pathogens of immunocompromised humans. The *T. citrinoviride* fungal genome was sequenced and assembled inadvertantly with the *T. longibrachiatum* genome project. The *T. citrinoviride* genome was a contaminant in the DNA provided to the JGI. The DNA supplier indicated that the two are hardly distinguishable in culture and do not form an incompatible reaction against each other, so it is difficult to differentiate them in cultures. The two species are distinct enough genetically to obtain useful genomic information even from a mixed culture.

Among other species that have been subject to resequencing, perhaps none have the breadth of data than yeast (Table 4.6), and while only-seven ***Saccharomyces cerevisae*** **strains** from public collections have been publicly sequenced, the U.S. National Library of Medicine Sequence Read Archive lists 336 studies and over 9,000 experiments relevant to *Saccharomyces* biology. While many of these later studies are of specifically constructed strains and often include transcriptome or even amplicon sequencing, the vast number of sequences for this one species portends great things as sequence analysis techniques learned from these studies are applied to the diversity of basidiomycete yeast (Table 4.7) and filamentous fungi held in modern culture collections (Table 4.1). Genome sequencing using multiple independent isolates of a yeast species has been used to generate new knowledge in areas such as population genomics and biogeography. *Saccharomyces cerevisiae* has industrial significance and is a model organism to study numerous eukaryotic cellular processes, such

Table 4.6. Ascomycete yeast species and strains sequenced

Species and reference	Note	Strain sequenced[a]
Ascoidea rubsescens	Beech sap; beetle associated	NRRL Y-17699
Ashbya gossypii	Plant pathogen, vitamin B_2 production	FDAG1
Ashbya gossypii	Plant pathogen	ATCC 10895
Aureobasidium pullulans var. *subglaciale*	Cryotolerant	EXF-2481
Aureobasidium pullulans var. *pullulans*	Produces multiple enzymes	EXF-150
Babjeviella inositovora	Phylogenetically isolated	NRRL Y-12698
Candida albicans	Human pathogen	WO-1
Candida albicans	Human pathogen	SC5314
Candida arabinofermentans	Arabinose-fermenting yeast	NRRL YB-2248
Candida caseinolytica	Secretes strong protease	NRRL Y-17796
Candida dublinensis	Human pathogen	CD36
Candida glabrata	Human pathogen	CBS 138
Candida guilliermondii	Human pathogen	ATCC 6260
Candida lusitaniae	Human pathogen	ATCC 42720
Candida orthopsilosis	Human pathogen	90–125
Candida parapsilosis	Human pathogen	CDC 317
Candida tanzawaensis	Unstudied group associated with insects	NRRL Y-17324
Candida tenuis	Studies of D-xylose metabolism enzymes	NRRL Y-1498
Candida tropicalis	Human pathogen	ATCC MYA-3404
Cyberlindnera jadinii	Synonym *C. utilis* used industrially	NRRL Y-1542
Debaryomyces hansenii	Industrial use	CBS 767
Debaryomyces hansenii	Salt tolerant	MTCC 234
Dekkera bruxellensis	Wine yeast	CBS 2499
Hanseniaspora valbyensis	Organic acids and low molecular weight volatiles	NRRL Y-1626
Hansenula polymorpha	Industrial use	NCYC 495 leu1.1
Hyphopichia burtonii	Widespread, competitively uses starch	NRRL Y-1933
Kazachstania africana	Genome evolution	CBS 2517
Kazachstania naganishii	Genome evolution	CBS 8797
Kluyveromyces aestuarii		ATCC 18862
Kluyveromyces lactis	Industrial use	CLIB 210
Kluyveromyces lactis	Industrial use	NRRL Y-1140
Klluyveromyces thermotolerans	Thermotolerant	CBS 6340
Kluyveromyces waltii		NCYC 2644
Kluyveromyces wickerhamii		UCD-FST 54-210
Lachancea (Saccharomyces) kluyveri	Plant pathogen	NRRL Y-12651
Lodderomyces elongisporus	Human pathogen	NRRL YB-4239
Lipomyces starkeyi	High lipid producer (65 % of dry weight)	NRRL Y-11557
Metschnikowia bicuspidata	Aquatic yeast; parasite of *Daphnia*	NRRL YB-4993
Naumovozyma castellii	Genome evolution	AS 2.2404, CBS 4309
Naumovozyma dairenensis	Genome evolution	CBS 421
Pachysolen tannophilus	Pentose fermentation	NRRL Y-2460
Pichia membranifaciens	Organic acids; tolerates low pH	NRRL Y-2026
Pichia pastoris		GS115
Saccharomycse bayanus	Beverage fermentation	
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	M3707
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	M3836
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	M3837
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	M3838
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	S288C
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	Sigma1278b
Saccharomyces cerevisiae	First eukaryotic genome sequenced; model	JAY291
Saccharomyces pastorianus	Industrial, beer	Weihenstephan 34.70
Saitoella complicata	Rare soil yeast, basal to yeast divergence	NRRL Y-17804

(continued)

Table 4.6. (continued)

Species and reference	Note	Strain sequenced[a]
Scheffersomyces (Pichia) stipitis	Pentose- and cellobiose-fermenting yeast	CBS 6054
Schizosaccharomyces pombe	Model eukaryote	972 h
Spathaspora passalidarum	Model xylose fermenter	NRRL Y-27907
Tetrapisispora blattae	Genome evolution	CBS 6284
Tetrapisispora phaffii	Genome evolution	Van der Walt Y 89; CBS 4417
Torulaspora delbrueckii	Genome evolution	CBS 1146
Vanderwaltozyma polyspora		DSM 70294
Wickerhamomyces anomalus	Ethanol, polyols, and glycerol; low water	NRRL Y-366
Yarrowia lipolytica	Industrial use	CLIB 89
Yarrowia lipolytica	Industrial use	CLIB 122
Zygosaccharomyces rouxii	Food spoilage; osmotolerant	CBS 732

Updated list can be obtained from the Department of Energy Joint Genomes Institute (http://genome.jgi.doe.gov/programs/fungi/index.jsf) and Genomes Online Database (http://www.genomesonline.org/). Additional collections that carry many of these strains can be determined in a search of the StrainInfo bioportal (www.straininfo.net). Many strain ID numbers listed in genome-sequencing projects lack a formal culture collection identifier, which makes tracking strains difficult

[a]Acronyms as in Table 4.1

Table 4.7. Basidiomycete yeast species with whole-genome sequence

Species		Strain sequenced[a]
Cryptococcus neoformans var. *grubii*	Human pathogen	H99
Cryptococcus neoformans var. *neoformans*	Human pathogen	JEC21
Cryptococcus vishniacii	Cryotolerant	ANT 03-052
Dacryopinax sp.	Saprophytic	DJM-731 SS1
Dioszegia cryoxerica	Extremophile	ANT03-071
Malassezia globosa	Associated with dandruff	CBS 7966
Malassezia restricta	Associated with dandruff	CBS 7877
Rhodotorula graminis	Plant symbiont	WP1 (FGSC 10291)
Tremella mesenterica	Comparative fungal genomics	DSM 1558
Sporobolomyces roseus	Associated with plants	IAM 13481
Ustilago maydis	Plant pathogen/huitlacoche fungus	UM 521 (FGSC 9021)

[a]Acronyms as in Table 4.1

as the research of Hartwell, Hunt, and Nurse that won the 2001 Nobel Prize (Balter and Vogel 2001).

Initially based on a desire to understand how cells coordinated cell division, the recognition that yeast could be used to make mutants defective in cell cycle control led to the demonstration that genes were required for cell division, first in *S. cerevisae* and subsequently in *Schizosaccharomyces pombe* and the sea urchin *Arbacia punctulata*. Ultimately, more than 100 genes were identified that were required for faithful completion of the cell cycle in a controlled manner. Foundational among these were cell division cycle genes, kinases, and cyclins.

The *S. cerevisiae* genome was the first eukaryotic genome to be sequenced and the second genome of any kind (Goffeau et al. 1996). The **Saccharomyces Genome Resequencing Project**, a collaboration between the Sanger Institute and the Institute of Genetics, University of Nottingham, sequenced the genomes of over 70 diverse wild isolates of *S. cerevisiae* and its close relative *S. paradoxus* from multiple geographic and ecological sources to 1–3X coverage (http://www.sanger.ac.uk/research/projects/genomeinformatics/sgrp.html). They found that phenotypic variation correlated with genome-wide phylogenetic

relationships. While *S. paradoxus* populations were divided along geographic boundaries, the population structure of *S. cerevisiae* indicated multiple domestication events. This study was possible because of participation of a large number of academic researchers and culture collections that provided yeast strains. These yeast strains are available to other researchers from the National Collection of Yeast Cultures (NCYC), either as individual cultures or in sets arrayed in 96-well plate format (http://www.ncyc.co.uk/sgrp.html), to allow further discovery from these isolates. Projects involving sequencing of multiple yeast species, and comparison of genome sequences generated by others, have advanced because of the access to yeast strains from public culture collections. The Génolevures Consortium (http://genolevures.org/), comprises researchers at Laboratoire Bordelais de Recherche en Informatique, Institut National de Recherche en Informatique, Centre de Recherche Bordeaux Sud-Ouest, and Université Louis Pasteur in France, generated nine complete and seven partial genome sequences of Hemiascomycete yeast species (Sherman et al. 2009). The yeast strains used in this project, as well as several hundred thousand genomic clones, can be obtained from the International Center for Microbial Resources (CIRM-Levures).

C. Strain Equivalence in a Postgenomics Era

The material in living microbe collections has a unique status with regard to intellectual property. Strains that are clonally derived from the same progenitor strain are often considered to be equivalent, and with regard to their intended use, this assumption is often valid.

Comparative analysis of the whole-genome sequence from presumably identical strains is possible with next-generation sequencing capabilities. Preliminary analysis of closely related strains shows limited numbers of polymorphisms. For example, among the 18 whole genomes analyzed in the analysis of classical *N. crassa* mutants, the 2 strains with the fewest polymorphisms relative to the reference genome (FGSC 309 and 7035) shared only 3,179 SNPs, suggesting that most of the 31,761 total SNPs in these two strains are not derived from a common ancestor (McCluskey et al. 2011). Ongoing projects to compare even more closely related strains are in progress and promise to help define the relationship between strains that are otherwise identical.

A number of targeted sequencing projects have looked at the **variation in classical mutants** and have largely recapitulated the genetic analysis that defined the various loci. These studies have shown that *N. crassa* mutants at the *al-2* locus (Diaz-Sanchez et al. 2011), the *ad-8* locus (Wiest et al. 2012), and now the *trp-3* locus (Wiest et al. 2013), which have shared lineage and are essentially identical from a physiological or biochemical perspective, can be differentiated by DNA sequence analysis.

Similarly, **mitochondrial genome variability** among closely related strains, a subject that received a great amount of attention as a valuable strain-specific fingerprint (see, e.g., Paquin et al. 1997) is an area that has only been explored at a superficial level in whole-genome analysis of filamentous fungi (McCluskey 2012). Certainly, part of this is because of the early stage of the art, where the whole-genome sequences of one strain per species have yet to be elucidated, and as more studies delve deeply into the diversity within species, mitochondrial diversity will be part of that characterization.

IV. Impact of Genome Technology on Culture Collections

Among culture collections, the impact of whole-genome sequence data has been slowly accumulating. Diversity-based collections have the ability to use **barcode data inherently present in a whole-genome sequence** in revisiting the taxonomic classification of materials deposited under names generated by physiological or morphological criteria. Similarly, genetic collections are expanding to include new diversity as genome technology expands the traditional distinction between observational and genetic studies (McCluskey et al. 2010).

A. A Gene-Ontology-Driven Culture Collection Catalog

Because one of the uses of culture collections is to **identify organisms that possess valuable or otherwise desirable traits**, clients will often obtain multiple isolates from one group of functionally or taxonomically related organisms in the hope that a characteristic found in one such organism will be found in many related isolates. While this is certainly true at a first approximation, once every isolate in a collection has been subject to whole-genome analysis, it should be possible to search culture collection databases not by taxonomic keys, but rather by a search using the Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) or by following a gene ontology classification (Ashburner et al. 2000) to identify strains that share a particular trait, enzymatic activity, or conserved domain. While a diligent investigator can use a functionally similar approach to searching online databases and following the links to the strains desired, this is not closely related to the ability to search a taxonomically diverse culture collection for strains that share a particular characteristic. A BLAST search of the National Center for Biotechnology Information (NCBI) database using the protein sequence from *Neurospora* biotin synthetase (NCU03937, a single-copy gene) identified 100 orthologs in 99 different fungal species among the top 100 scoring-related protein sequences. While this highly conserved protein demonstrates how an ontology-based catalog might function, the real value of the ontological catalog will be in finding orthologs to unusual, rare, or novel protein or genome sequences.

B. Allele Catalogs from Resequencing Efforts

Similarly, an allele database for *N. crassa* listing all of the different yet (presumably) functional alleles of each ORF as well as temperature-sensitive (TS) and loss-of-function alleles would be a valuable resource. This is increasingly important as more and more studies compare their results to the reference genome while there are clearly multiple genetic backgrounds even in the inbred research strain population.

Whole-genome sequencing has also been evaluated for taxonomy in prokaryotes, and although there are some challenges, it is likely that **whole-genome sequencing will ultimately be used for taxonomy**, and for species and gene phylogenies, as well as for unanticipated applications yet to be discovered (Sentausa and Fournier 2013). In an immediate time frame, some of the challenges noted include the fact that, for many species, the strain subject to whole-genome analysis is not the type strain. In addition, many taxa are not sequenced to as great a depth, so coverage is not uniform. This last challenge may go away naturally as multiple isolates of many species are sequenced (Forde and O'Toole 2013), either for cataloging in culture collections, as has been advocated here and elsewhere (McCluskey 2013), or for biogeographic and other studies. Genome resequencing, where multiple isolates of one species are sequenced and compared to a reference genome assembly (Stratton 2008), has been applied at multiple taxonomic levels with great impact.

In fungi, for example, genome resequencing has been applied to basal flagellated chytrid fungi to understand the relationship between geographic distribution and evolutionary history (Rosenblum et al. 2013) and to plant pathogenic fungi demonstrating significant and large-scale genetic differences that are associated with production of pathogenicity-related proteins in *Stagonospora* (Syme et al. 2013). Similar studies with *Fusarium* (Ma et al. 2010) have demonstrated the horizontal transfer of chromosomes bearing pathogenicity-related genes.

Resequencing has also been applied to the **hunt for mutated genes**, which are known genetically and phenotypically, but for which there is no identified ORF in yeasts and filamentous fungi. A number of different approaches have been taken with this same goal. For example, the anonymous *N. crassa* mutant locus *ndc-1* was identified as an allele of the previously characterized biosynthetic locus *spe-1* using an bulk segregant analysis

followed by subtraction of SNPs segregating in the *Neurospora* laboratory mutant strain population (Pomraning et al. 2011). Similarly, developmental mutants of *Sordaria macrospora* were characterized by sequencing DNA pooled from 40 progeny. In this study, ambiguous reads were discarded, and comparative analyses were used to identify candidate mutations responsible for the phenotypes under study; these were subsequently validated by molecular genetic complementation (Nowrousian et al. 2012). By way of contrast, the more well-saturated genetic map of *N. crassa* (Perkins et al. 2000) coupled with the common genetic backgrounds used by this research community (Galagan et al. 2003) meant that bulk segregant analysis was not necessary in a preliminary resequencing project looking for anonymous mutations (McCluskey et al. 2011). In this study, a de facto in silico bulk segregant analysis was used by which any polymorphisms relative to the reference genome present in multiple mutant strains were excluded from the analysis of individual mutant strains. For some mutants, this meant that the number of unique polymorphisms in the genetically defined region carrying the mutant locus was reduced to near unity.

In this last case, a rapid approach to characterizing mutant loci was possible exactly because these strains were preserved in the public culture collection. The laborious approach of mutant identification by protein characterization, library construction, and complementation was not necessary because a public collection was available to preserve the material until a rapid characterization technique was available. The gap between the material being deposited into the public collection and subsequent identification of the mutation was from 20 to 51 years (Table 4.4) and required the genomic technology to accomplish this task. While the ultimate value of this characterization will be demonstrated by requests for strains carrying these mutations, the mere fact of their characterization sets *Neurospora* apart as a model organism. Many other fungi were in use during the early days of genetic analysis, and with the limited exceptions of *Neurospora* and *Aspergillus nidulans*, most genetic fungal systems are no longer studied. For example, dozens of classic mutant strains, including *Allomyces* (Olson 1984), *Ustilago* (Perkins 1949), and *Sordaria fimicola* (Kitani et al. 1961), are available in ad hoc and formal collections, although these materials have not been utilized in as many as 20 years (McCluskey et al. 2010). Clearly, other research systems are in use, as evidenced by the tremendous growth of the Fungal Genetics Conference. In the last 10 years, the number of participants and of abstracts submitted have grown from 550 attendees and 495 poster abstracts in 2003 to nearly 1,000 attendees and 741 poster abstracts in 2013. Similarly, significant research is done with *Ustilago*, *Allomyces*, and *Sordaria*, but it is not in the context of material from a formal culture collection.

V. Conclusions

While it is a tautology to say that culture collections are limited to maintaining and propagating strains that are capable of being grown in axenic culture, the ability to provide a broad foundation of data underlying the known diversity of life will ensure that the materials in culture collections will provide an essential reference for studies of cell biology, taxonomy, phylogeny, and environmental biology. Currently, many metagenomics programs are limited to associating anonymous ORFs with kingdom- or phylum-level taxa, while genomics informed studies could at least allow assignment to family or even genus levels (Grigoriev et al. 2012). Similarly, the ability to characterize neutral, or at least apparently neutral, sequence polymorphisms among laboratory isolates will inform studies of the relationship between phenotype and sequence polymorphisms (McCluskey 2013). Culture collections are the only reliable mechanism of ensuring that research results from one era are available for future researchers. Moreover, researchers who are interested in ensuring the highest impact for their research can expect as many as 150 % more citations to published work when the material is publically available (Furman and Stern 2011). While the global trend is for

formalization of living microbe collections, support for collections in the United States and around the world is fragile. The future availability of living microbial material will require a new rationale for public support or the attention of a visionary philanthropist.

References

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 31:533–538

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29

Balter M, Vogel G (2001) Cycling toward Stockholm. Science 294:502–503

Boundy-Mills K (2012) Yeast culture collections of the world: meeting the needs of industrial researchers. J Ind Microbiol Biotechnol 39:673–680

Campbell L, Betsou F, Garcia D, Giri J, Pitt K, Pugh R, Sexton K, Skubitz A, Somiari S, Astrin J (2012) Best practices for repositories – collection, storage, retrieval and distribution of biological materials for research. Biopreserv Biobank 10:79–161

Crous PW, Verkley GJ, Groenewald JZ, Samson R (2009) Fungal biodiversity. Centraalbureau voor Schimmelcultures, Utrecht

Daniel HM, Prasad GS (2010) The role of culture collections as an interface between providers and users: the example of yeasts. Res Microbiol 161:488–496

Diaz-Sanchez V, Estrada AF, Trautmann D, Limon MC, Al-Babili S, Avalos J (2011) Analysis of al-2- mutations in *Neurospora*. PLoS One 6:e21948

Dugan FM Jr, Wiest A, McCluskey K (2011) Public germplasm collections and revolutions in biotechnology. J Biosci 36:205–209

Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, McCluskey K, Plamann M, Galagan JE, Birren BW, Weiss RL, Townsend JP, Loros JJ, Nelson MA, Lambreghts R, Colot HV, Park G, Collopy P, Ringelberg C, Crew C, Litvinkova L, DeCaprio D, Hood HM, Curilla S, Shi M, Crawford M, Koerhsen M, Montgomery P, Larson L, Pearson M, Kasuga T, Tian C, Basturkmen M, Altamirano L, Xu J (2007) Enabling a community to dissect an organism: overview of the *Neurospora* functional genomics project. Adv Genet 57:49–96

Forde BM, O'Toole PW (2013) Next-generation sequencing technologies and their impact on microbial genomics. Brief Funct Genomics 12:440–453

Furman JL, Stern S (2011) Climbing atop the shoulders of giants: the impact of institutions on cumulative research. Am Econ Rev 101:1933–1963

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274(546):563–567

Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA (2012) The genome portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res 40:D26–D32

ISO Guide 34:2009 (2009) General requirements for the competence of reference material producers. ISO, Geneva

Janssens D, Arahal DR, Bizet C, Garay E (2010) The role of public biological resource centers in providing a basic infrastructure for microbial research. Res Microbiol 161:422–429

Kitani Y, Olive LS, E-Ani AS (1961) Transreplication and crossing over in *Sordaria fimicola*. Science 134:668–669

Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, Zeilinger S, Casas-Flores S, Horwitz BA, Mukherjee PK, Mukherjee M, Kredics L, Alcaraz LD, Aerts A, Antal Z, Atanasova L, Cervantes-Badillo MG, Challacombe J, Chertkov O, McCluskey K, Coulpier F, Deshpande N, von Doehren H, Ebbole DJ, Esquivel-Naranjo EU, Fekete E, Flipphi M, Glaser F, Gomez-Rodriguez EY, Gruber S, Han C, Henrissat B, Hermosa R, Hernandez-Onate M, Karaffa L, Kosti I, Le Crom S, Lindquist E, Lucas S, Lubeck M, Lubeck PS, Margeot A, Metz B, Misra M, Nevalainen H, Omann M, Packer N, Perrone G, Uresti-Rivera EE, Salamov A,

Schmoll M, Seiboth B, Shapiro H, Sukno S, Tamayo-Ramos JA, Tisch D, Wiest A, Wilkinson HH, Zhang M, Coutinho PM, Kenerley CM, Monte E, Baker SE, Grigoriev IV (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of Trichoderma. Genome Biol 12:R40

Lang JM, Darling AE, Eisen JA (2013) Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. PLoS One 8: e62510

Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie X, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RA, Chapman S, Coulson R, Coutinho PM, Danchin EG, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature 464:367–373

Marmeisse R, Nehls U, Öpik M, Selosse MA, Pringle A (2013) Bridging mycorrhizal genomics, metagenomics and forest ecology. New Phytol 198:343–346

Martin F, Aerts A, Ahren D, Brun A, Danchin E, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452:88–92

McCluskey K (2012) Variation in mitochondrial genome primary sequence among whole-genome-sequenced strains of *Neurospora crassa*. IMA Fungus 3:93

McCluskey K (2013) Biological resource centers provide data and characterized living material for industrial biotechnology. Ind Biotechnol 9:117–122

McCluskey K, Wiest A, Plamann M (2010) The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. J Biosci 35:119–126

McCluskey K, Wiest A, Grigoriev IV, Lipzen A, Martin J, Schackwitz W, Baker SE (2011) Rediscovery by whole genome sequencing: classical mutations and genome polymorphisms in *Neurospora crassa*. G3 (Bethesda) 1:303–316

McNeill J, Barrie F, Buck W, Demoulin V, Greuter W, Hawksworth D, Herendeen P, Knapp S, Marhold K, Prado J (2012) International Code of Nomenclature for algae, fungi, and plants (Melbourne Code). Koeltz Scientific Books, Koenigstein

Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101:12248–12253

Nowrousian M, Teichert I, Masloff S, Kuck U (2012) Whole-genome sequencing of *Sordaria macrospora* mutants identifies developmental genes. G3 (Bethesda) 2:261–270

OECD (2007) OECD best practice guidelines for biological resource centres. The Organisation for Economic Development and Co-operation, Paris

Olson LW (1984) Allomyces – a different fungus. Opera Bot 73:1–96

Paquin B, Laforest M-J, Forget L, Roewer I, Wang Z, Longcore J, Lang BF (1997) The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr Genet 31:380–395

Perkins DD (1949) Biochemical mutants in the smut fungus *Ustilago maydis*. Genetics 34:607

Perkins DD, Radford A, Sachs MS (2000) The *Neurospora* compendium: chromosomal loci. Academic, New York

Pomraning KR, Smith KM, Freitag M (2011) Bulk segregant analysis followed by high-throughput sequencing reveals the *Neurospora* cell cycle gene, ndc-1, to be allelic with the gene for ornithine decarboxylase, spe-1. Eukaryot Cell 10:724–733

Purnick PE, Weiss R (2009) The second wave of synthetic biology: from modules to systems. Nat Rev Mol Cell Biol 10:410–422

Rosenblum EB, James TY, Zamudio KR, Poorten TJ, Ilut D, Rodriguez D, Eastman JM, Richards-Hrdlicka K, Joneson S, Jenkinson TS (2013) Complex history of the amphibian-killing chytrid fungus revealed with genome resequencing data. Proc Natl Acad Sci U S A 110:9385–9390

Rountree MR, Selker EU (2009) Genome defense: the *Neurospora* paradigm. In: Ferguson-Smith A, Greally J, Martienssen R (eds) Epigenomics. Springer, New York, pp 321–341

Sentausa E, Fournier PE (2013) Advantages and limitations of genomics in prokaryotic taxonomy. Clin Microbiol Infect 19:790–795

Sherman DJ, Martin T, Nikolski M, Cayla C, Souciet JL, Durrens P (2009) Genolevures: protein families and synteny among complete hemiascomycetous yeast proteomes and genomes. Nucleic Acids Res 37:D550–D554

Smith D (2012) Culture collections. Adv Appl Microbiol 79:73–118

Stackebrandt E (2010) Diversification and focusing: strategies of microbial culture collections. Trends Microbiol 18:283–287

Stern S (2004) Biological resource centers: knowledge hubs for the life sciences. Brookings Institution Press, Washington, DC

Stratton M (2008) Genome resequencing and genetic variation. Nat Biotechnol 26:65–66

Syme RA, Hane JK, Friesen TL, Oliver RP (2013) Resequencing and comparative genomics of *Stagonospora nodorum*; sectional gene absence and effector discovery. G3 (Bethesda) 3:959–969

Turner BC, Perkins DD, Fairfield A (2001) *Neurospora* from natural populations: a global study. Fungal Genet Biol 32:67–92

Wiest A, McCarthy AJ, Schnittker R, McCluskey K (2012) Molecular analysis of mutants of the *Neurospora* adenylosuccinate synthetase locus. J Genet 91:199–204

Wiest A, Barchers D, Eaton M, Henderson R, Schnittker R, McCluskey K (2013) Molecular analysis of intragenic recombination at the tryptophan synthetase locus in *Neurospora crassa*. J Genet 92(3):523–528

World Intellectual Property Organization (1980) Budapest treaty on the international recognition of the deposit of microorganisms for the purposes of patent procedure done at Budapest on April 28, 1977, and amended on September 26, 1980. World Intellectual Property Organization, Geneva. http://www.wipo.int/treaties/en/text.jsp?file_id=283784

Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ (2009) A phylogeny-driven genomic encyclopaedia of bacteria and archaea. Nature 462:1056–1060

Cell and Developmental Biology

5 Fungal Chromatin and Its Role in Regulation of Gene Expression

Michael Freitag[1]

CONTENTS

The histones are commonly regarded as unpleasant proteins for rigorous studies.

—Cook et al. 1956, cited in Van Holde 1988, p. 69

I. Introduction

Gene regulation in fungi must be considered in the context of chromatin. This assembly of DNA, RNA, and proteins undergoes time- and space-dependent compaction or unfolding to restrict or allow access to regulatory regions, the "promoters" and "enhancers" characterized by sequence motifs that are recognized by classical transcription factors. Changes in chromatin structure and composition have received considerable attention in most deeply understood model systems, driven largely by technological advances in studying chromatin as a template for gene regulation. There is a particularly rich literature based on studies with two major models, the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe*. Many recent reviews have addressed chromatin regulation in these organisms (Gasser and Cockell 2001; Rusche et al. 2003; Martienssen et al. 2005; Grewal and Jia 2007; Grewal 2010; Sun et al. 2011; Yankulov 2011; Goto and Nakayama 2012; Rando and Winston 2012). Over the past decade, it has become clear, however, that what is true in either or both yeasts is not necessarily conserved in filamentous ascomycetes or basidiomycetes, as chromatin of these organisms more closely resembles that of "higher" eukaryotes. This chapter contains a brief overview of chromatin components, modification enzymes, and the most recent work with filamentous fungi, without attempting to be comprehensive. Several recent reviews deal with the general subject (Blander and Guarente 2004; Ekwall 2005; Simon and Kingston 2009; Eissenberg and Shilatifard 2010; Beisel and Paro 2011; Margueron and Reinberg 2011; Schuettengruber et al. 2011; Tennen and Chua 2011; Upadhyay and Cheng 2011; Hall et al. 2012; Young and Kirchmaier 2012; Oppikofer et al. 2013; Rudolph et al. 2013), while others cover the topic as it relates to filamentous fungi (Brosch et al. 2008; Shwab and Keller 2008; Palmer and Keller 2010; Rountree and Selker 2010; Brakhage and Schroeckh 2011; Strauss and Reyes-Dominguez 2011; Chang et al. 2012; Gacek and Strauss 2012; Smith et al. 2012a, b). One purpose is to indicate gaps in our current knowledge about chromatin-mediated gene regulation in filamentous fungi and provide examples to illustrate how comparative biology of

[1]Department of Biochemistry and Biophysics, Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR 97331-7305, USA;
e-mail: freitagm@cgrb.oregonstate.edu

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

fungi may uncover unexpected chromatin components and different regulatory networks.

Fungi seem especially well suited for comparative chromatin biology as they are relatively simple organisms with small genomes that usually lack redundancy of genes for chromatin components and DNA or histone modification enzymes. As they are amenable to rapid genetic manipulations, biochemistry, and cytology, many fungi have become important model organisms. With the advent of high-throughput sequencing, however, it is now possible to quickly generate almost-complete genomes and transcriptomes of any fungus (Nowrousian et al. 2010; de Wit et al. 2012; Stukenbrock et al. 2012; Bushley et al. 2013; Umemura et al. 2013; Wiemann et al. 2013; Zhao et al. 2013). To date, there are more than 200 high-quality draft genomes from many different families available, and their number still increases exponentially (see URLs at the end of this chapter). For incisive comparative studies, sequencing depth of specific genera is important; the taxa *Aspergillus* (Galagan et al. 2005; Machida et al. 2005; Nierman et al. 2005; Fedorova et al. 2008; Arnaud et al. 2012; Knuf and Nielsen 2012; Gibbons and Rokas 2013) and *Fusarium* (Cuomo et al. 2007; Ma et al. 2010; Rep and Kistler 2010; Wiemann et al. 2013) are excellent examples. By now, some fungi have been analyzed deeply at the species level, for example, *Neurospora crassa* (Ellison et al. 2011) and *Zymoseptoria tritici* (Goodwin et al. 2011; Stukenbrock et al. 2011; Stukenbrock and Dutheil 2012; Croll et al. 2013). The combination of genomic, transcriptomic, and metabolomic resources allows formulation of new hypotheses with regard to the chromatin landscape and gene regulation in filamentous fungi, as documented recently with *Fusarium fujikuroi* (Wiemann et al. 2013) and *Tolypocladium inflatum* (Bushley et al. 2013).

II. Regulation by Chromatin Structure: Not All Modifications Are "Epigenetic"

It is tempting to equate regulation by chromatin structure with "epigenetic" regulation. ***Epigenetic phenomena are generally defined as heritable but reversible states of regulation that occur in the absence of changes in DNA sequence.*** They include genomic imprinting in mammals and plants, dosage compensation of sex chromosomes in animals, paramutation in plants, transvection in flies, maintenance and inheritance of centromeres, prions, and many more (Gottschling 2007; Williams 2013).

Genomic imprinting refers to systems in which the more typical biallelic expression from genes in diploid cells is changed to parental-specific monoallelic expression (Koerner and Barlow 2010). Imprinted messenger RNA (mRNA) expression is determined by parental origin of the specific locus, and the affected genes have the same parental-specific expression in both male and female progeny. Over 100 maternally or paternally expressed imprinted genes have been identified; many control embryonic development. Both DNA methylation and noncoding RNA are causally involved in this phenomenon (Koerner and Barlow 2010; Lee and Bartolomei 2013). **Dosage compensation** refers to various systems that ensure equal expression from genes on sex chromosomes. In nematodes, sex is determined by the ratio of X chromosomes to autosomes. Genes named *X signal elements* (XSEs) repress a masculinizing sex determination gene, *xol-1* (XO lethal), in a dose-dependent manner. Autosome dosage is determined by *autosomal signal elements* (ASEs) that counter XSEs by stimulating *xol-1* transcription (Meyer 2010; Farboud et al. 2013; Kruesi et al. 2013). In flies, transcriptional upregulation of male X-linked genes is stimulated by the noncoding roX RNA (Simon et al. 2011). A ribonucleoprotein complex called *male-specific lethal* or *dosage compensation complex* (MSL/DCC) forms on the male X chromosome specifically near active genes at high-affinity "entry" sites (HASs), from which it spreads along the chromosome (Straub et al. 2013). Conversely, in mammals one of the two female X chromosomes is randomly chosen and transcriptionally silenced, a process called **X chromosome inactivation**, that is also controlled by a long noncoding RNA, Xist, and several smaller RNAs (Jeon et al. 2012; Lee and Bartolomei 2013). **Paramutation**, mostly studied in maize, refers to heritable expression states "by allelic interactions in trans," mediated by small RNAs and specific patterns of DNA methylation (Stam 2009; Erhard and Hollick 2011; Hollick 2012). **Transvection**, mostly studied in *Drosophila*, refers to interactions of chromosomal loci, where paring or nonpairing can influence expression of genes (Duncan 2002; Kennison and Southworth 2002).

Underlying mechanisms for these phenomena involve DNA methylation, RNA-based

silencing, and protein-mediated inheritance or a combination of all three. Only if these mechanisms result in reversible and heritable states in gene regulation should a process be deemed "epigenetic" in the strict sense. In many cases, however, these criteria are not fulfilled, as chromatin-mediated regulation may not extend across several cell cycles or meioses. In fact, some even argue that the whole concept of epigenetic regulation as applied today should be revised or even abandoned, as accumulating evidence shows that most epigenetic phenomena can be explained by cell- and time-specific transcriptional regulation in the traditional sense (Ptashne 2013a, b). Thus, the idea of epigenetic regulation served as a useful hypothesis that resulted in much deeper understanding of chromatin biology but ultimately it may turn out to be a flawed concept. Epigenetic may, however, still be used to mean all developmental gene regulation, as originally intended (Waddington 1952).

There are a few examples for which the concept of epigenetic inheritance is useful, for example, in genomic imprinting, which seems to be controlled by noncoding RNA and allele-specific DNA methylation (Lee and Bartolomei 2013). Perhaps the best example for epigenetic chromatin regulation by deposition of specific proteins is the **maintenance and inheritance of centromere position** and kinetochore formation. Specific DNA sequences alone are insufficient to generate normal centromeres or aberrant "neocentromeres" in most eukaryotes (Cleveland et al. 2003; Black and Cleveland 2011). Instead, cell-cycle-dependent deposition of a centromere-specific histone H3 variant, CENP-A in mammals (Earnshaw and Migeon 1985; Palmer et al. 1987) and CenH3 in filamentous fungi (Smith et al. 2011, 2012a), marks centromeric DNA for assembly of kinetochores (Verdaasdonk and Bloom 2011; Hori and Fukagawa 2012). For many years, centromeric regions had been viewed as "constitutive heterochromatin," domains of chromosomes not transcribed and condensed even during interphase. Studies in fission yeast, *Drosophila*, and mammals revealed that transcription can occur in centromeric core regions, and that they are associated with active chromatin marks, in this case, dimethylation of canonical histone H3 on lysine residue 4 (H3K4me2) (Sullivan and Karpen 2004; Cam et al. 2005). Indeed, H3K4me2 was required for maintenance of centromeres on a human artificial chromosome and for the deposition of CENP-A (Bergmann et al. 2011). Thus, the view of centromeres as constitutively silenced regions has been challenged by biochemical and cytological analyses of rice, fly, and mammalian chromatin (Sullivan and Karpen 2004; Yan et al. 2005; Ma et al. 2007).

Relatively little is known about this in the filamentous fungi. In *N. crassa*, centromeric nucleosomes carry silencing histone lysine modifications, for example, trimethylation of H3K9me3 (Smith et al. 2011), and the same is true for *Fusarium* species (Wiemann et al. 2013; L.R. Connolly, K.M. Smith, and M. Freitag, unpublished data). Introduction of reporter genes into centromeric DNA in wild-type *Neurospora* strains resulted in gene silencing, presumably by reduced transcription, but the centromeric silencing effect was abrogated in a series of mutants defective in a novel histone deacetylase (HDAC) complex, HCHC (Honda et al. 2012). Some of the comparatively large centromeric regions of *N. crassa* are transcribed, however, and small interfering RNA (siRNA) is generated from the assembly of transposon relics that make up most of the centromeric DNA (Chicas et al. 2004; Lee et al. 2010).

The example of centromeres also illustrates how researchers first encountered regulation by chromatin state, namely, as **position effect variegation** (PEV) in *Drosophila* (Henikoff 1990) or **telomere position effects** (TPEs) in *Saccharomyces* (Gottschling et al. 1990). In these situations, reporter genes were transcribed when placed into chromatin conducive for transcription ("euchromatin") but were silenced when placed into heterochromatin. The field was driven forward by analyses of mating-type switching in *S. cerevisiae*, which uncovered genes for *silent information regulators* (SIRs) (Rine et al. 1979; Rine and Herskowitz 1987; Haber 2012). One of these, Sir2, encodes the founding member of the **sirtuins**, class III

HDACs (Blander and Guarente 2004). Sirtuins are also involved in silencing of rDNA gene clusters (Santoro 2005; Preuss and Pikaard 2007; McStay and Grummt 2008; Salminen and Kaarniranta 2009), another topic on which there is currently little information available from filamentous fungi.

Position effects, specifically TPEs in the widest sense, have received some attention in *Neurospora* and *Aspergillus* (Smith et al. 2008; Palmer and Keller 2010). In *N. crassa*, reporter genes were silenced when placed close to telomere repeats at some chromosome ends, yet silencing was relieved in mutants with faulty heterochromatin assembly (Smith et al. 2008), thus resembling the situation in *S. cerevisiae* but mediated by different protein complexes. In *Aspergillus*, position effects have not been examined this close to the telomeric repeats. Much of the current excitement in the field stems from the realization that silent chromatin controls the expression of large gene clusters that often are required for the **production of secondary metabolites** (SMs) in many fungi. SM gene clusters tend to accumulate near the chromosome ends in most filamentous fungi that have been analyzed (Fedorova et al. 2008; Ma et al. 2010; Andersen et al. 2013; Condon et al. 2013; Gan et al. 2013; Inglis et al. 2013; Wiemann et al. 2013), although they are typically too far away from chromosome ends to be the subject of TPEs in the strict sense. Thus, these phenomena may not be truly epigenetic, but they share mechanistic aspects with heterochromatin silencing. Under most laboratory growth conditions, SM gene clusters remain transcriptionally silent, and one long-standing challenge has been to activate these chromosome domains at will, often by culturing organisms on special sets of growth media (Frisvad 2012; Tormo et al. 2012). Manipulating chromatin structure in fungi has been considered as one strategy that promises to unlock previously untapped biochemical potential by transcriptional activation (Brakhage 2013); thus, many recent studies have been directed at understanding the transcriptional silencing and activation of these large chromosomal regions (Bok et al. 2006; Shwab and Keller 2008; Bok et al. 2009).

III. Building Blocks of Chromatin: The Histone Complement of Filamentous Fungi

Gene regulation by changes in chromatin is dependent on the major structural elements of chromatin, the histones (van Holde 1988). Regulation can be mediated by replacement of the "canonical" core histones (H2A, H2B, H3, H4) with histone variants (Table 5.1) or by numerous posttranslational modifications of core or linker histones (Table 5.2). During cell-specific development in animals, protamines can replace histones; little is currently known about similar potential replacement of histones by histone-like proteins in filamentous fungi.

Overall, fungi carry a **minimal set of genes for histones**, which makes them excellent organisms for the genetic analysis of histone function (Grunstein 1990; Durrin et al. 1991; Ling et al. 1996; Mellone et al. 2003; Dai et al. 2008; Nakanishi et al. 2008; Adhvaryu et al. 2011). In most filamentous fungi, genes for the **core histones** H3, H2A, and H2B, as well as the linker histone H1, are unique (see databases available at the Broad Institute, the Joint Genome Institute (JGI) MycoCosm, and FungiDB; URLs are listed at the end of this chapter). Most fungi, however, contain two copies of H4 (often annotated as H4.1 and H4.2), similar to the hemiascomycetous yeasts, for which two or even three core histone gene clusters exist. Unlike in the *Drosophila* genome (van Holde 1988), histone genes of fungi are not organized into a single cluster but rather dispersed across the genome. H2A and H2B, as well as H3 and H4.1, are always linked and divergently transcribed from one promoter (Ehinger et al. 1990; Hays et al. 2002). H4.2 is differentially regulated, at least in *Aspergillus nidulans*, and has a single amino acid change when compared to H4.1 (Fig. 5.1) (Ehinger et al. 1990). In *N. crassa*, as in most sequenced fungi, H4.1 and H4.2 are identical, even though the genes encode transcripts that have introns of different size and different splice sites (Hays et al. 2002). Nothing is known about the actual function of the second H4 protein. In mammals, H3.1 is incorporated during replication ("replication-

Table 5.1. Fungal core histones and their variants

Group	Name	Homolog	Functions
Histone H2A	H2A	H2A.X	Many; necessary for DNA repair
	H2A.Z	H2A.Z	Many; counteracts gene silencing
Histone H2B	H2B	H2B	Found in all nucleosomes
Histone H3	H3	H3.3	Many; replication-independent type
	CenH3	CENP-A	centromere-specific H3
Histone H4	H4.1	H4	Many; coregulated with H3
	H4.2	H4	Many; independently regulated from H3
	H4v	None	Putative fungal-specific H4

Fungal histones match specific types of histones produced in mammals. "Canonical" mammalian H2A, H3.1, and H3.2 do not exist in fungi, suggesting that H2A.X and H3.3 more closely resemble the precursor of all H2A and H3 proteins. Conversely, fungi seem to produce a larger variety of H4 protein variants

dependent" H3), and H3.3 is used as a "replacement" histone during all other chromatin transactions, for example, when nucleosomes are remodeled during transcription, recombination, or repair (Ahmad and Henikoff 2002; Talbert and Henikoff 2010). Perhaps H4.1 is a replacement H4 while H4.2 is the vestige of a replication-dependent H3-H4 gene pair (Fig. 5.1). It is well known that core histones are the most conserved proteins of eukaryotes; indeed, protein sequence alignments derived from the known or predicted histone genes of all available fungal genomes show close-to-perfect conservation for H3, H4 (Fig. 5.1). There are, however, some minor differences for H2A and H2B (Hays et al. 2002).

There are numerous studies on the **linker histone H1** from many organisms, but its function in chromatin biology is still not completely clear. It is thought to be involved in regulation of heterochromatin structure and DNA methylation (Wolffe et al. 1997; Kasinsky et al. 2001; Vujatovic et al. 2012; Kalashnikova et al. 2013; Lu et al. 2013; Yang et al. 2013). In humans, there are numerous differentially regulated H1 proteins (Kowalski and Palyga 2012). In *Drosophila*, H1 proved essential (Lu et al. 2009), while an ill-conserved homolog is dispensable in *Tetrahymena* (Shen et al. 1995) and *S. cerevisiae* (Patterton et al. 1998; Freidkin and Katcoff 2001). There have been few studies in filamentous fungi, although H1 silencing has been shown to cause DNA hypermethylation in *Ascobolus* (Barra et al. 2000). This was not generally applicable, as no DNA methylation phenotype was detected in *Neurospora* mutants (Folco et al. 2003). Instead, relatively minor transcriptional effects were noticed, exemplified by misregulation of the pyruvate decarboxylase gene. Future research should address the function of histone H1 in fungi.

Aside from the normal set of core or linker histones found in most nucleosomes of a given species, several **histone variants** are known (Table 5.1). These include the centromere-specific H3 (CENP-A or CenH3), already discussed, as well as H2A variants associated with transcription (H2A.Z), repair (H2A.X), or chromosome-wide silencing (macro-H2A) (Ahmad and Henikoff 2002; Talbert and Henikoff 2010). Fungi have genes for H2A.Z, but their regular H2A seems more closely related to H2A.X (Xhemalce et al. 2012). In the wheat pathogen *Pyrenophora tritici-repentis*, a putative DNA transposon has acquired a novel type of H3 gene (Manning et al. 2013), an example of transduplication in fungi; however, when transferred into *N. crassa*, this putative variant was not targeted to the nucleus (S. Solisti, V. Manning, L. Ciuffetti, and M. Freitag, unpublished data). In most eukaryotic genomes, there are no other variants; so far, none have been found for H2B or H4.

In filamentous fungi, however, there is a **potential novel H4 variant**, first identified as *hH4v* in *N. crassa* (Hays et al. 2002). Initially, *hH4v* was considered a pseudogene, as the predicted protein did not seem to share large stretches of identity with canonical H4. Perusal of all fungal genomes suggests otherwise,

Table 5.2. Histone modifications and examples of histone-modifying enzymes

Modification	Residue	Example	Enzyme (fungal homologs)
Acetylation	Lys (K)	H4K5/12ac	KAT1 (ScHat1, NcHAT-1)
		H3K9/14/18/36ac	KAT2 (ScGcn5, NcNGF-1)
		H3ac and H4ac	KAT4 (ScTaf1, NcHAT-3)
		H4K5/8/12/16ac	KAT5 (ScEsa1, NcHAT-4, An ESAA)
		H3K14ac	KAT6 (ScSas3, NcHAT-6)
		H4K16ac	KAT8 (ScSas2, NcHAT-5)
		H3ac	KAT9 (ScElp3, NcELP-3)
		H3K14ac	KAT10 (ScHap2)
		H3K56ac	KAT11 (ScRtt109)
Deacetylation	AcLys (acK)	H3ac, H4ac	Rpd3 (ScRpd3, NcHDA-3, AnRPDA, FgRpd3)
		H3ac?, H4ac?	Hda1 (ScHda1, NcHDA-1, AnHDAA, FgHdF2)
		H3ac, H4ac	Hos2 (ScHos2, NcHDA-2, AnHosA, FgHdF1)
		H3ac, H4ac	Hos3 (ScHos3, NcHDA-4, AnHosB, FgHdF3)
		H3K9/14/56ac, H4K16ac	SIRT1 (ScHst1, NcNST-1,)
		H3K9/14/56ac	SIRT2 (ScHst2, NcNST-2, An HSTA)
		H4ac?	SIRT1 (ScHst4, NcNST-3)
		?	SIRT3 (ScHst3, NcNST-4)
		H4ac?	SIRT4 (NcNST-5)
		?	SIRT5 (NcNST-6)
		Nucleolus	SIRT7 (NcNST-7)
Methylation	Lys (K)	H3K9me2, -3	KMT1 (SpClr4, NcDIM-5, AnCLRD, FgKMT1)
		H3K4me2, -3	KMT2 (SpSet1, NcSET-1, An, FgKMT2)
		H3K36me2, -3	KMT3 (ScSet2, NcSET-2, An, FgKMT3)
		H3K79me	KMT4 (ScDot1, NcDOT-1, FgKMT4)
		H4K20me	KMT5 (NcSET-, FgKMT5)
		H3K27me3	KMT6 (NcSET-7, FgKMT6)
	Arg (R)	H4R3me	PRMT1 (ScHmt1, NcPRM-1)
		Nonhistone proteins	PRMT3 (NcPRM-3)
		H3R8me, H4R3me	PRMT5 (ScHsl7, SpSkb1, NcPP-2)
Demethylation	MeLys (meK)	H3K4me, H3K9me	KDM1 (SpLsd1, AnHDMA)
		H3K36me	KDM2 (ScJhd1)
		H3K9me	KDM3
		H3K9me, H3K36me	KDM4 (ScRph1)
		H3K4me	KDM5 (Jhd2, SpJmj2, SpLid1)
		H3K27me	KDM6
Phosphorylation	Ser (S)	H3S10ph	Aurora B (ScIpl1, AnNIMA)
		H4S1ph	CKII (ScCkII)
	Thr (T)	H3T45ph	CDC7 (ScCdc7)

Emphasis is placed on proteins mentioned in the text. Ubiquitylation, sumoylation, ADP-ribosylation, proline isomerization, and citrullination are not discussed in this chapter. Enzymes are listed by their mammalian names or classes, with fungal homologs where appropriate (Sc *S. cerevisiae,* Sp *S. pombe,* Nc *N. crassa,* An *A. nidulans,* Fg *F. graminearum*). For most enzymes from filamentous fungi, activities still need to be determined, others are from Xhemalce et al. (2012)

however, as H4v genes exist in many taxa (Fig. 5.2). This gene is transcribed in *N. crassa* and several *Fusarium* species, and the resulting *Neurospora* protein, when tagged with green fluorescent protein (GFP) or mCherry, is targeted to the nucleus (J. Galazka and M. Freitag, unpublished results). Compared to canonical H4, this predicted protein has extended and not generally conserved N-terminal and C-terminal tails, but the histone fold domain with the alpha helixes required to fold into the nucleosome (Freeman et al. 1996) are conserved (Fig. 5.2). Splice sites are conserved, and while many silent mutations exist, few missense changes occur between closely related species, suggesting that these genes are under purifying selection. Thus, this may prove to be one example where comparison of the many genomes available from filamentous fungus may yield novel insights into chromatin biology.

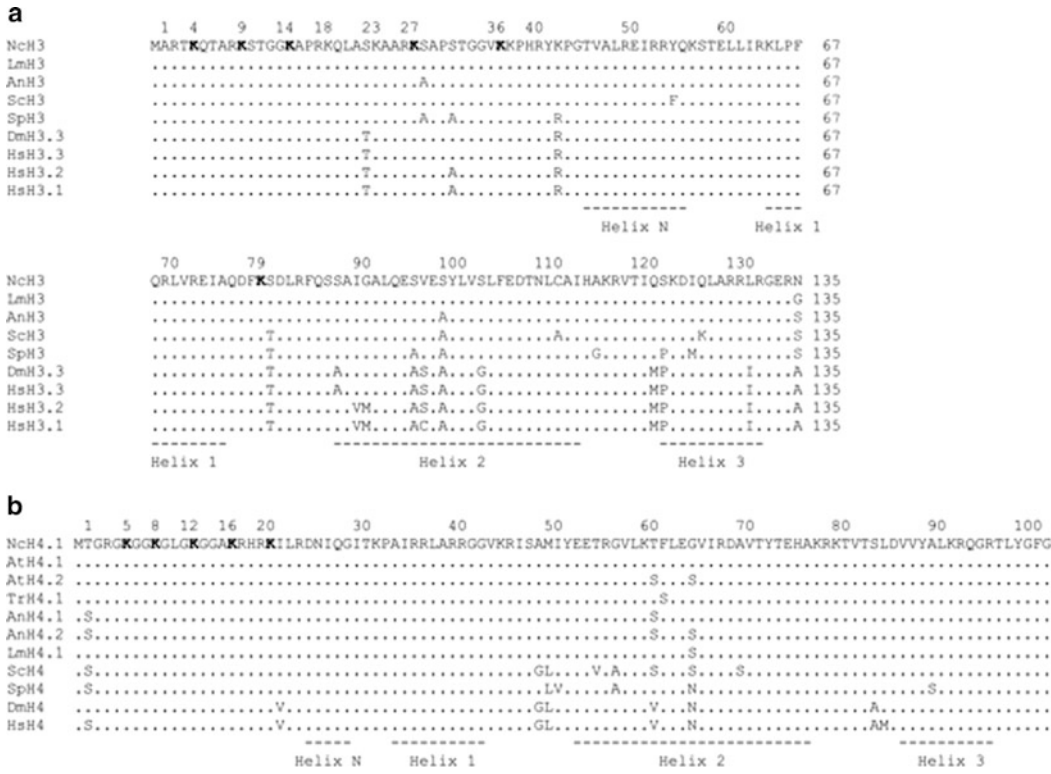


Fig. 5.1. Conservation of histone H3 (**a**) and histone H4 (**b**). A selection of the most commonly found H3 and H4 proteins from filamentous fungi are shown. Modifiable lysine residues discussed in the text are in **bold**. Most changes in both H3 and H4 are conservative substitutions (e.g., S for T or V for I), or they occur in alpha helices for which modeling suggests that substitutions do not change the predicted structure of the H3-H4 tetramers or octameric nucleosome core (J. Galazka and M. Freitag, unpublished data). Abbreviations and accession numbers: Nc, *N. crassa* H3 (XP_956003.1), H4.1 (XP_956597.1); Lm, *Leptospheria maculans* H3 (XP_003845436.1), H4.1 (XP_003845636.1); An, *A. nidulans* H3 (CAA39154.1), H4.1 (AAA20820.1), H4.2 (AAA20821.1); Sc, *S. cerevisiae* H3 (NP_009564.1), H4 (EEU08141.1); Sp, *S. pombe* H3 (NP_595567.1), H4 (NP_595566.1); Dm, *Drosophila melanogaster* H3 (NP_723056.2), H4 (NP_001027352.1); Hs, *Homo sapiens* H3.1 (NP_003522.1), H3.2 (NP_066403.2), H3.3 (NP_005315.1), H4 (EAW55528.1); At, *Aspergillus terreus* H4.1 (XP_001210673.1), H4.2 (XP_001208735.1); Tr, *Trichoderma reesei* H4.1 (EGR44277.1)

Why are histone variants of interest for chromatin regulation? Replacement of canonical histones with variants causes defined changes in gene regulation. In yeasts, H2A.Z (called Htz1 in *S. cerevisiae* and H2A.Z^{Pht1} in *S. pombe*) replaces H2A at promoters of inactive or weakly expressed genes (Wyrick and Parra 2009; Hou et al. 2010), and it has been associated with both silencing and activating activities. It acts as an insulator inhibiting spreading of silencing complexes into euchromatin (Meneghini et al. 2003), and it is required for genome stability and recruitment of RNA Pol II (Adam et al. 2001). Fission yeast H2A.Z^{Pht1} cooperates with the heterochromatin machinery to degrade antisense transcripts from neighboring genes by the exosome (Zofall et al. 2009). The role of H2A.Z in filamentous fungi remains to be explored. Chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) with a GFP-tagged H2A.Z found association with promoters of genes that also had H3K4me2 or -me3, suggesting that *Neurospora* H2A.Z is in promoters of active genes (P.A. Phatale, K.M. Smith, and M. Freitag, unpublished data).

More work has been done in both *N. crassa* and *A. nidulans* on the role of **posttranslational histone modifications** in gene regulation by chromatin structure. In particular, the N-terminal tails of H3 and H4 are required to

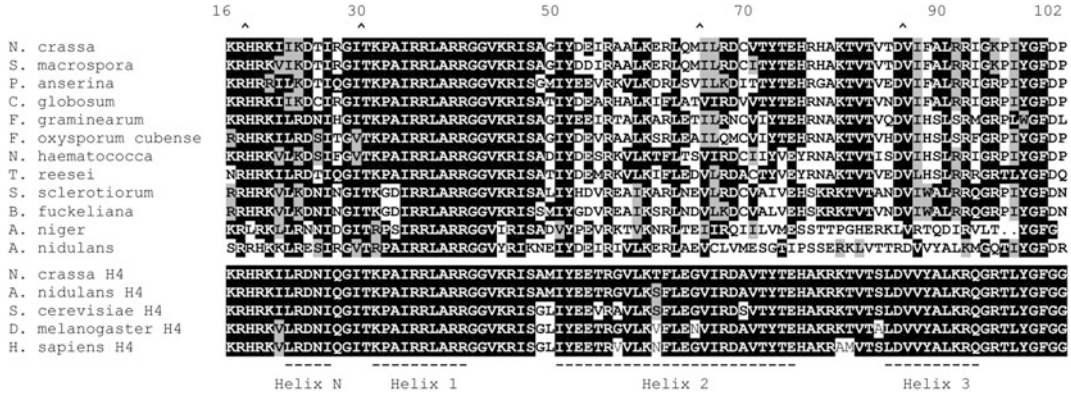


Fig. 5.2. Comparison of fungal H4 variants with canonical histone H4. Only residues within the histone fold domain are shown here as the N-terminal and C-terminal tails of the putative H4 variants are different when comparing different families of fungi and because annotation of these predicted proteins is currently uncertain. Not all H4 variants from filamentous fungi are shown. Numbering is for H4. As for H3 and H4, many changes in H4 variants are conservative substitutions, or they occur in alpha helices for which modeling predicts no changes in structure. Conserved intron splice sites are indicated by ^. Accession numbers for H4 variants: *N. crassa* XP_956319.2; *S. macrospora* XP_003348790.1; *Podospora anserina* XP_001909136.1; *Chaetomium globosum* XP_001224243.1; *F. graminearum* XP_385250.1; *F. oxysporum cubense* Race 1 EGU81449.1; *Nectria haematoccoca* XP_003046553.1; *Trichoderma reesei* Triredraft_55201; *Sclerotinia sclerotiorum* XP_001590757.1; *Botryotinia fuckeliana* XP_001552206.1; *A. niger* XP_001398751.2; *A. nidulans* XP_657721.1). In some cases, the predicted genes were newly annotated as previous annotations were incorrect. Poor alignment of the *Aspergillus* sequences past alpha helix 2 suggests that in these taxa H4 variants are poorly conserved

generate heterochromatin or maintain euchromatin. Lysines in the H3 and H4 tails are hyperacetylated in euchromatin, and H3K4 is trimethylated. In heterochromatin, lysines are hypoacetylated, and H3K9 and H3K27 are trimethylated. Mutating histones of *S. cereviseae* and *S. pombe* revealed a role for specific residues in transcription. For example, changes to H3K9, H3S10, or H3K14 in *S. pombe* caused loss of centromere silencing and defective chromosome segregation, a phenotype similar to deletion of Swi6, a homolog of Heterochromatin Protein 1 (HP1) (Mellone et al. 2003). Similar studies have been carried out in *Neurospora*, where the influence of H3 residues on DNA methylation has been studied (Adhvaryu et al. 2011).

IV. A Primer on Proteins Involved in Histone Modification

Enzymes that modify histone residues are responsible for regulating transcription in addition to other processes involving DNA (recombination, repair, replication, chromatid cohesion, chromosome segregation). Histones can be modified by lysine acetylation, lysine and arginine methylation, phosphorylation, ubiquitylation, sumoylation, and ADP-ribosylation (Voigt and Reinberg 2011; Yun et al. 2011; Zaidi et al. 2011; Zhou et al. 2011). Reversible acetylation is catalyzed by histone acetyltransferases (HATs) and HDACs. Lysine methylation is mediated by SET domain containing histone methyltransferases (HMTs) and a smaller class of proteins homologous to Dot1 (Feng et al. 2002; van Leeuwen et al. 2002; Klose and Zhang 2007; Li et al. 2010), while demethylation is carried out by Jumonji domain proteins or homologs of LSD1 (Whetstine et al. 2006; Li et al. 2010). Arginine methylation is catalyzed by protein arginine methyltransferases (PRMTs) (Gary and Clarke 1998) and opposed by peptidylarginine deiminase enzymes (Klose and Zhang 2007). Phosphorylation of histones is regulated by kinase and phosphatase enzymes that often also act on

other proteins. New nomenclature has been proposed, primarily to counteract the confusing numbering schemes in various organisms but also to reinforce the idea that "HATs" are acting on not only histones but also nonhistone proteins. Thus, HATs are now often referred to as KATs (lysine acetyltransferases) and HMTs as KMTs (lysine methyltransferases), while lysine demethylases are called KDMs (Allis et al. 2007). The confusing schemes from the yeasts and other organisms unfortunately have already been transferred to the few filamentous fungi under study. In general, *N. crassa* naming is based on effects related to DNA methylation, such as DIM-5 (Defective in Methylation-5) as the homolog for the *Drosophila* Su(var3–9) or *S. pombe* Clr4 proteins, or on similarity with *S. cerevisiae* proteins, such as NGF-1 for "Neurospora GCN Five." Conversely, *A. nidulans* nomenclature resembles that of the two yeast systems (Table 5.2).

While histone modifications can change chromatin structure, more often modified histone tails are recognized by regulatory proteins (de la Cruz et al. 2005). Acetylated lysine residues are recognized by bromo domain proteins, often part of complexes involved in transcriptional activation. Methylated lysine residues are bound by chromo domain proteins that can form either transcriptionally silent heterochromatin, such as by use of HP1 or Polycomb (Pc), or may result in transcriptional activation, such as by use of the chromo domains in the chromatin remodeling ATPase CHD1. The remainder of the chapter is concerned with how histone modifications may be involved in generating silent or active chromatin.

V. Control of Heterochromatin Formation

Much of what is known about **constitutive heterochromatin** formation in filamentous fungi was uncovered by incisive studies on the control of DNA methylation in *N. crassa* (Rountree and Selker 2010; Honda et al. 2012). An early link was the finding that the HDAC inhibitor Trichostatin A (TSA) affected DNA methylation, at least in certain regions of the genome (Selker 1998). In subsequent studies, connections between **H3K9 methylation** and **DNA methylation** (Tamaru and Selker 2001; Tamaru et al. 2003; Freitag et al. 2004; Honda and Selker 2008; Lewis et al. 2010a, b) revealed mechanisms that were also found in plants and mammals (Lachner et al. 2001; Jackson et al. 2002; Lehnertz et al. 2003).

Pathways resulting in H3K9 methylation, if not DNA methylation, may be similar in many filamentous fungi as the components of known *Neurospora* complexes are generally conserved in genomes of filamentous fungi. Heterochromatin formation in *Neurospora* relies on **AT-rich DNA**, generated by the action of a premeiotic genome defense system called **repeat-induced point (RIP) mutation** (Selker 1990).

RIP was the first eukaryotic genome defense system discovered (Selker et al. 1987; Cambareri et al. 1989). A genome-wide DNA-based homology search that occurs after fertilization but before karyogamy during the sexual cycle results in mutagenesis of duplicated regions (Selker 1990). The phenomenon was first referred to as "rearrangements induced premeiotically" (Selker et al. 1987), but premeiotic recombination can be separated from mechanisms responsible for mutations (Selker and Garrett 1988). Pairing is required for a still-unknown mutator to introduce C:G-to-T:A transition mutations, an observation that resulted in the new name, repeat-induced point mutation or RIP (Cambareri et al. 1989). Unmethylated DNA segments that have undergone RIP are often methylated (Selker and Stevens 1987; Singer et al. 1995). RIP is not restricted to Neurospora, as *Podospora anserina* (Graia et al. 2001), *Magnaporthe grisea* (Ikeda et al. 2002), *Leptosphaeria maculans* (Idnurm and Howlett 2003), *Gibberella zeae* (*F. graminearum*) (Cuomo et al. 2007; Pomraning et al. 2013), and *Nectria haematococca* (Coleman et al. 2009) produce the hallmarks of RIP in test crosses (i.e., reporter gene inactivation, generation of transition mutations, or cytosine methylation). In the ascomycete *Ascobolus immersus* and the basidiomycete *Coprinus cinereus*, de novo cytosine DNA methylation can occur after crossing strains with repeated DNA (Zolan and Pukkila 1986; Goyon and Faugeron 1989; Faugeron et al. 1990; Rhounim et al. 1992; Freedman and Pukkila 1993). The timing of the methylation events is similar to RIP, and the process is called **methylation-induced premeiotically (MIP)** (Rhounim et al. 1992). Methylation is maintained through many mitoses and has been shown to effectively block recombination between methylated repeats (Goyon and Faugeron 1989; Rhounim et al. 1992; Colot et al. 1996; Maloisel and Rossignol 1998).

DNA mutated by RIP serves as a substrate for DIM-5 (Lewis et al. 2009), which has so far been found in one complex, called DCDC (DIM-5, -7, -9, CUL-4, and DDB-1 complex). DCDC recruits HP1 and DIM-2 (Lewis et al. 2010a, b; Zhao et al. 2010). DIM-7 targets DIM-5 to specific regions and mediates interactions between DIM-5 and DIM-9 (or DDB-1/CUL-4 associated factor, DCAF) (Lewis et al. 2010a; Xu et al. 2010). DDB-1 and CUL-4 are part of a conserved E3 ubiquitin ligase complex (Petroski and Deshaies 2005), for which the substrate is still unknown. DIM-5 is a KMT1 enzyme (Table 5.2) that trimethylates H3K9 (Tamaru and Selker 2001; Tamaru et al. 2003), and H3K9me3 is recognized by the chromo domain of HP1 (Freitag et al. 2004). HP1 recruits DIM-2, the single DNA methyltransferase known in *Neurospora* (Kouzminova and Selker 2001; Honda and Selker 2008). Loss of DIM-2 has little – if indeed any – effect on DIM-5 activity, H3K9me3 levels or HP1 localization.

How histone modification enzymes are attracted to AT-rich DNA remains to be uncovered. An attractive group of proteins that may signal DNA composition are high-mobility group (HMG) proteins, the linker histone H1, and additional proteins with "AT-hook" motifs. Disruption of the *N. crassa* H1 gene had no obvious effect (Folco et al. 2003), although its potential effects on formation or maintenance of heterochromatin should be revisited. AT hooks bind to the minor groove of AT-rich DNA. Treatment of *N. crassa* with distamycin, which also binds to the minor groove, reduced DNA methylation (Tamaru and Selker 2003). This predicts that binding of DIM-5 and distribution of H3K9me3 should be altered in treated cells, as occlusion of the minor groove may have inhibited recognition of AT-rich DNA.

Other histone-modifying enzymes may play a role in heterochromatin formation by acting upstream or separately of DIM-5. The first can be accomplished by modifying histone tails for recognition by DCDC. There is evidence for this pathway, as certain alleles of *ppp-1*, encoding protein phosphatase 1, resulted in increased H3S10 phosphorylation and selective loss of H3K9 methylation from certain regions in the genome (Adhvaryu and Selker 2008), lending evidence to the existence of a K9me-S10ph "binary switch" (Fischle et al. 2003). In addition, mutation of two of the four classical HDAC genes, *hda-1* (yeast Hda1) and *hda-3* (yeast Rpd3), resulted in loss of DNA methylation and H3K9me3 from some regions, associated with increased H3K9 and H3K14 acetylation (Smith et al. 2010). This suggested that HDAC complexes act upstream of or feed back into the DIM-5 pathways. Silencing at telomeres also relies on DIM-5 and HP1, in combination with both NAD-dependent sirtuins and classical HDACs that were shown to cooperate in deacetylating subtelomeric histones (Smith et al. 2008). Combining multiple mutations in different *nst* genes (*Neurospora sir two*; sirtuins) caused enhanced reactivation of silent reporter genes, demonstrating overlapping specificity and cooperative functions of sirtuins and classical HDACs (Smith et al. 2008, 2010). Activation of subtelomeric transgenes was increased in *dim-5* or *hpo* mutants, presumably because there was still residual HDAC activity and because DIM-5 and HP1 act downstream (Smith et al. 2008). DNA methylation does not seem to play a major role in telomeric silencing in *Neurospora*, as *dim-2* mutants resulted in little reactivation (Smith et al. 2008), suggesting that telomeric silencing by H3K9me3 and HP1 is independent of their role in establishing DNA methylation, at least in some regions. This idea suggested the presence of additional complexes that mediate heterochromatin formation or maintenance independently of HP1-DIM-2. One such complex, HCHC, was found to contain HDA-1, CDP-2, HP1, and CHAP (Honda et al. 2012). A third complex that can interact with HP1, DMM (DNA methylation modulator), counteracts the spreading of heterochromatin (Honda et al. 2010). This complex contains DMM-1, a JmjC domain and AT-hook protein, and a close relative of *Arabidopsis* IBM1 (Saze et al. 2008; Miura et al. 2009). DMM acts as an insulating complex as it is recruited to edges of heterochromatin and stops spreading of DNA methylation. The JmjC domain may be involved in demethylation of proteins, but a substrate has not been identified. Growth deficiencies observed by loss of DMM-1 were complemen-

ted by deletion of *dim-2* or treatment with the cytosine methylation inhibitor 5-azacytidine (5AC) (Honda et al. 2010), suggesting that in the absence of DMM-1 DNA methylation silences euchromatic genes, but that the silencing can be reversed by lack of DNA methylation. The interplay between the three *Neurospora* HP1-interacting complexes suggests that competing chromatin complexes generate boundaries between active and silent regions. In this scenario, specific DNA sequences or structures that serve as boundary elements are not required.

VI. Facultative Heterochromatin in Filamentous Fungi

In contrast to more permanently silenced regions (**constitutive heterochromatin**) that are never or hardly ever expressed, some chromatin regions are subject to regulation that makes them switch from active to silent state (or vice versa). Examples of this type of chromatin include the inactive X chromosome of humans and clusters of genes regulated by polycomb proteins in mammals and *Drosophila*. Perhaps the best examples for **facultative heterochromatin** in filamentous fungi are long stretches of chromosomes that encode **SM gene clusters**. To express the usually silent SM gene clusters, investigators attempt to shift the balance from silencing to activating histone modifications, a field now called *chemical epigenetics* (Wang et al. 2010; Brakhage and Schroeckh 2011). To this end, *Aspergillus niger*, *Penicillium citreonigrum*, *Cladosporium cladosporioides*, and *Diatrype disciformis* were treated with 5AC or an HDAC inhibitor, suberoylanilide hydroxamic acid (SAHA), to produce novel SMs (Williams et al. 2008; Fisch et al. 2009; Henrikson et al. 2009; Wang et al. 2010).

Targeted disruption of histone-modifying enzymes has been used; here, research has focused on known components of well-established silencing complexes involved in heterochromatin formation. Transgenes inserted relatively near *A. nidulans* telomeres were silent but reactivated by disruption of the genes for HP1 (*hepA*), KMT1 (*clrD*, the Clr4/DIM-5 H3K9 MTase), Hda1 (*hdaA*), and Ku70 (*nkuA*), but silencing was not affected by lack of at least one sirtuin, encoded by *hstA* (Palmer and Keller 2010). As discussed for *Neurospora*, likely several sirtuins may have to be deleted at the same time to obtain reactivation phenotypes. Mutation of *hepA* also resulted in induction of silent SM gene clusters (Reyes-Dominguez et al. 2010), as did HDAA depletion (Shwab et al. 2007).

Like HP1 and KMT1, **histone deacetylases** are thought to be directly involved in gene silencing. In the absence of yeast Rpd3, acetylation is increased at H4K5, H4K12, and H3K18 (Robyr et al. 2002). The *A. nidulans* Rpd3 homolog (RPDA) is essential, as is *N. crassa* HDA-3 (Smith et al. 2010), but silencing *rpdA* revealed an overall increase in H3 and H4 acetylation (Tribus et al. 2010). A different yeast HDAC, Hda1, deacetylates nucleosomes in subtelomeric regions and promoters of genes that are distinct from those controlled by Rpd3 (Robyr et al. 2002). Deletion of the *A. nidulans hdaA* gene caused reduced growth under conditions of oxidative stress (Tribus et al. 2005) and caused derepression of SM gene clusters (Shwab et al. 2007), which was also found in *Aspergillus fumigatus hdaA* mutants (Lee et al. 2009). Intriguingly, this mutant showed both up- and downregulation of nonribosomal peptide synthase (NRPS) genes, and specificity for telomere-linked regions was not detected. In *F. graminearum*, strains that lack the Hda1 homolog (called HdF2) exhibited only minor effects on regulation of SM gene clusters, while mutants of the yeast Hos2 homolog (called HdF1) showed reduced levels of deoxynivalenol but upregulation of the aurofusarin biosynthetic genes (Li et al. 2011). These results suggest that effects are not simple, and that direct and indirect interactions between different histone modifications are the rule rather than the exception, as discussed further in the following material. Nothing is known about the effects of fungus-specific Hos3-like HDACs (Trojer et al. 2003) on SM gene cluster regulation.

The **putative protein methyltransferase LaeA**, first identified in *A. nidulans* (Bok and Keller 2004), is a regulator of many SM gene

clusters, carbon metabolism, and development in *Aspergillus* and other species (Bok and Keller 2004; Bok et al. 2005; Perrin et al. 2007; Bayram et al. 2008; Kale et al. 2008; Amaike and Keller 2009; Atoui et al. 2010; Wiemann et al. 2010). LaeA is considered a protein methyltransferase because mutagenesis of the *S*-adenosylmethionine binding motif resulted in the same phenotype as *laeA* deletion (Bok et al. 2006). Although it may not be directly modifying histones, LaeA is instrumental in maintaining the active state of SM gene clusters by somehow reversing H3K9 methylation marks (Bayram and Braus 2011; Strauss and Reyes-Dominguez 2011), perhaps by interfering with HP1 binding (Reyes-Dominguez et al. 2010). While there are LaeA homologs in all filamentous fungi, and many of them have at least seven additional LaeA-like predicted proteins (Jiang et al. 2011; Palmer et al. 2013b), pathways may not be conserved, as recent studies on cellulase expression in *Trichoderma reesei* showed (Seiboth et al. 2012; Karimi-Aghcheh et al. 2013). Important challenges include finding the target (or multiple targets) for LaeA and deciphering the role and finding targets of other LaeA-like proteins. In some cases, these proteins may act as "chaperones" rather than enzymes (Palmer et al. 2013b).

While mutation or inhibition of silencing histone modification enzymes is expected to relieve facultative heterochromatin, overexpression of enzymes involved in generating active chromatin may result in overcoming silencing, similar to overexpression of transcription factors to express a desired region. Transcriptionally active euchromatin is generally thought to be associated with nucleosomes that have trimethylated H3K4, H3K36, and H3K79 residues (Lachner et al. 2003; Sims et al. 2003) and that are overall hyperacetylated. The *A. nidulans* homolog of yeast Bre2 (called CCLA) is part of a KMT2 (or Set1) protein complex, called **COMPASS** (complex proteins associated with Set1). KMT2 homologs are found in all filamentous fungi, as H3K4 di- or trimethylation is thought to be essential for RNA Pol II binding and transcriptional activity in development and differentiation (Eissenberg and Shilatifard 2010). Mutants lacking CCLA have reduced levels of H3K4me2 and -me3 as well as reduced overall H3 acetylation, but surprisingly, also reduced levels of both H3K9 di- and trimethylation (Bok et al. 2009). The function of COMPASS and heterochromatic marks seem to be conserved in regulating fungal SM gene clusters, as deletion of the HP1 homolog *hepA* and *cclA* in *F. graminearum* resulted in altered chemical profiles (Bok et al. 2009; Strauss and Reyes-Dominguez 2011), and deletion of *cclA* in *A. fumigatus* resulted in decrease of at least some SM gene cluster activity, although deletion of a histone demethylase (*hdmA*) had no effect (Palmer et al. 2013a). Coculture of *A. nidulans* with *Streptomyces rapamycinicus* resulted in overproduction of orsellinic acid, sterigmatocystin, terrequinone, and penicillin, apparently caused by triggering genome-wide hyperacetylation of H3K14 (and cluster-wide hyperacetylation of H3K9/K14) by the *A. nidulans* homologue of the yeast Gcn5 KAT (Nutzmann et al. 2011). Overexpression of a different KAT gene, *A. nidulans esaA*, resulted in increased H4K12ac and gene expression in the same four SM gene clusters (Soukup et al. 2012). Taken together, these results do not fit into a simple model where overexpression of activating enzymes results in activation of all SM gene clusters and inactivation of the same enzymes results in silencing. Clearly, more work is required to complete the networks of interactions between activating and silencing modifications at each gene cluster.

Crosstalk between histone modifications is important, yet we know little about this in the filamentous fungi. The role of H3S10 and H3K14 modifications in regulating *Neurospora* DIM-5 activity has been mentioned. Much more is known from *S. cerevisiae*; for example, the H4R3 arginine methyltransferase Hmt1 and the KAT Gcn5 interact functionally (Lacoste et al. 2002). H4K8ac is a preferred substrate for Hmt1p (Kuo et al. 2009), and Gcn5 acetylates H3 and H4, which is required for activation of one target gene, *HIS3*. Mutation of *hmt1* suppresses transcriptional defects of Gcn5 mutation, suggesting potential antagonism between Gcn5 and Hmt1 (Kuo et al. 2009). Thus, putative transcriptional repression by Hmt1 is opposed via transcriptional activation by Gcn5. We should

expect that there are numerous similar systems controlling the expression of SM gene clusters, but so far information is lacking. A global analysis in yeast showed how this may be accomplished in a high-throughput fashion (Weiner et al. 2012). The authors addressed a common observation, namely, that deletion of important chromatin regulators only affects a small number of genes. The effects of 83 histone mutants and 119 chromatin gene deletion mutants on the induction or repression of genes involved in the yeast stress response were analyzed, separating the importance of the genes studied for steady-state levels of transcripts from their function during actual induction or repression. Interestingly, under stress conditions, an "activating" modification (H3K4me3) was actually involved in repression, specifically in promoters of genes involved in ribosome biosynthesis (Weiner et al. 2012).

Can we draw a complete model for how facultative heterochromatin is generated in fungi? Is there even a single pathway that applies to all fungi? It seems that the situation is not simple, as studies on histone methylation in the yeasts, *N. crassa*, *A. nidulans* and *F. graminearum* suggest. While some KMTs are conserved in eukaryotes (e.g., KMT2/Set1, KMT3/Set2, KMT4/Dot1), others are not found in all taxa. KMT1 (Clr4/DIM-5), responsible for H3K9me3, occurs in *S. pombe* and many other fungal genomes analyzed so far, but it is absent from *S. cerevisiae* and its relatives. Similarly, KMT6, responsible for H3K27me3, occurs in *N. crassa* and the fusaria, as well as many other taxa, but it is absent from *A. nidulans*, *Ustilago maydis*, *S. pombe*, and *S. cerevisiae*. In *Drosophila* and mammals, H3K27me3 and the presence of the complex that generates this mark, polycomb repressive complex 2, are responsible for the inheritance of transcriptional states (Simon and Kingston 2009; Margueron and Reinberg 2011). This modification has been identified in *N. crassa* (Smith et al. 2008; Jamieson et al. 2013) and several *Fusarium* species (L.R. Connolly, K.M. Smith, and M. Freitag, in 2013). Lack of H3K27 methylation in *F. graminearum* results in drastic phenotypes (Connolly et al. 2013), but effects in *Neurospora* are subtle (Smith et al. 2008; Jamieson et al. 2013). It seems that the aspergilli have coopted the pathway used by many other fungi to generate constitutive heterochromatin for the formation of facultative heterochromatin, while *Neurospora* and the fusaria maintain two separate pathways. It will be interesting to see how these two pathways interact.

VII. Concluding Remarks

What does the future hold? Improvements in mass spectrometry (MS) will allow the analysis of vanishingly small samples to detect all histone modifications at a certain developmental stage; so far, analyses have been limited to preparations from vegetative cultures (Xiong et al. 2010; Xiong and Wang 2011). Further improvements in "top-down" MS will yield analyses of complete histone molecules, which will allow us to address what modifications occur on one nucleosome at a given time and within a specific chromosomal region. The role of a wide variety of other histone modifications has not been explored. Similarly, chromatin remodeling factors, the large ATP-requiring machines that move nucleosomes around, have not been studied much in filamentous fungi. Currently, the field is driven largely by research on two filamentous fungi, *N. crassa* and *A. nidulans*. While we can extract some general information from what we have learned about these two species, it seems that every group will need to be examined to confirm previous results. Happily, the improvements in genomic, transcriptomic, and metabolomic tools enable us to carry out these studies. The fungi will still hold many surprises for us.

VIII. Useful URLs for Fungal Genomics and Transcriptomics

AspGD: http://www.aspgd.org/

Broad Institute: http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative/fungal-genome-initiative

FungiDB: http://fungidb.org/fungidb/

JGI MycoCosm: http://genome.jgi.doe.gov/programs/fungi/index.jsf

MIPS: http://mips.helmholtz-muenchen.de/genre/proj/fungi/fungal_overview.html

References

Adam M, Robert F, Larochelle M, Gaudreau L (2001) H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Mol Cell Biol 21:6270–6279

Adhvaryu KK, Selker EU (2008) Protein phosphatase PP1 is required for normal DNA methylation in *Neurospora*. Genes Dev 22:3391–3396

Adhvaryu KK, Berge E, Tamaru H, Freitag M, Selker EU (2011) Substitutions in the amino-terminal tail of *Neurospora* histone H3 have varied effects on DNA methylation. PLoS Genet 7:e1002423

Ahmad K, Henikoff S (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell 9:1191–1200

Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, Pillus L, Reinberg D, Shi Y, Shiekhattar R, Shilatifard A, Workman J, Zhang Y (2007) New nomenclature for chromatin-modifying enzymes. Cell 131:633–636

Amaike S, Keller NP (2009) Distinct roles for VeA and LaeA in development and pathogenesis of *Aspergillus flavus*. Eukaryot Cell 8:1051–1060

Andersen MR, Nielsen JB, Klitgaard A, Petersen LM, Zachariasen M, Hansen TJ, Blicher LH, Gotfredsen CH, Larsen TO, Nielsen KF, Mortensen UH (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. Proc Natl Acad Sci U S A 110:E99–E107

Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, Crabtree J, Howarth C, Orvis J, Shah P, Wymore F, Binkley G, Miyasato SR, Simison M, Sherlock G, Wortman JR (2012) The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucleic Acids Res 40:D653–D659

Atoui A, Kastner C, Larey CM, Thokala R, Etxebeste O, Espeso EA, Fischer R, Calvo AM (2010) Cross-talk between light and glucose regulation controls toxin production and morphogenesis in *Aspergillus nidulans*. Fungal Genet Biol 47:962–972

Barra JL, Rhounim L, Rossignol JL, Faugeron G (2000) Histone H1 is dispensable for methylation-associated gene silencing in *Ascobolus immersus* and essential for long life span. Mol Cell Biol 20:61–69

Bayram O, Braus GH (2011) Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. FEMS Microbiol Rev. doi:10.1111/j.1574-6976.2011.00285.x

Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon NJ, Keller NP, Yu JH, Braus GH (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science 320:1504–1506

Beisel C, Paro R (2011) Silencing chromatin: comparing modes and mechanisms. Nat Rev Genet 12:123–135

Bergmann JH, Rodriguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H, Larionov V, Jansen LE, Earnshaw WC (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. EMBO J 30:328–340

Black BE, Cleveland DW (2011) Epigenetic centromere propagation and the nature of CENP-a nucleosomes. Cell 144:471–479

Blander G, Guarente L (2004) The Sir2 family of protein deacetylases. Annu Rev Biochem 73:417–435

Bok JW, Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell 3:527–535

Bok JW, Balajee SA, Marr KA, Andes D, Nielsen KF, Frisvad JC, Keller NP (2005) LaeA, a regulator of morphogenetic fungal virulence factors. Eukaryot Cell 4:1574–1582

Bok JW, Noordermeer D, Kale SP, Keller NP (2006) Secondary metabolic gene cluster silencing in *Aspergillus nidulans*. Mol Microbiol 61:1636–1645

Bok JW, Chiang YM, Szewczyk E, Reyes-Dominguez Y, Davidson AD, Sanchez JF, Lo HC, Watanabe K, Strauss J, Oakley BR, Wang CC, Keller NP (2009) Chromatin-level regulation of biosynthetic gene clusters. Nat Chem Biol 5:462–464

Brakhage AA (2013) Regulation of fungal secondary metabolism. Nat Rev Microbiol 11:21–32

Brakhage AA, Schroeckh V (2011) Fungal secondary metabolites – strategies to activate silent gene clusters. Fungal Genet Biol 48:15–22

Brosch G, Loidl P, Graessle S (2008) Histone modifications and chromatin dynamics: a focus on filamentous fungi. FEMS Microbiol Rev 32:409–439

Bushley KE, Raja R, Jaiswal P, Cumbie JS, Nonogaki M, Boyd AE, Owensby CA, Knaus BJ, Elser J, Miller D, Di Y, McPhail KL, Spatafora JW (2013) The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. PLoS Genet 9:e1003496

Cam HP, Sugiyama T, Chen ES, Chen X, FitzGerald PC, Grewal SI (2005) Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. Nat Genet 37:809–819

Cambareri EB, Jensen BC, Schabtach E, Selker EU (1989) Repeat-induced G-C to A-T mutations in *Neurospora*. Science 244:1571–1575

Chang SS, Zhang Z, Liu Y (2012) RNA interference pathways in fungi: mechanisms and functions. Annu Rev Microbiol 66:305–323

Chicas A, Cogoni C, Macino G (2004) RNAi-dependent and RNAi-independent mechanisms contribute to the silencing of RIPed sequences in *Neurospora crassa*. Nucleic Acids Res 32:4237–4243

Cleveland DW, Mao Y, Sullivan KF (2003) Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell 112:407–421

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S, Schwartz DC, Freitag M, Ma LJ, Danchin EG, Henrissat B, Coutinho PM, Nelson DR, Straney D, Napoli CA, Barker BM, Gribskov M, Rep M, Kroken S, Molnar I, Rensing C, Kennell JC, Zamora J, Farman ML, Selker EU, Salamov A, Shapiro H, Pangilinan J, Lindquist E, Lamers C, Grigoriev IV, Geiser DM, Covert SF, Temporini E, Vanetten HD (2009) The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. PLoS Genet 5: e1000618

Colot V, Maloisel L, Rossignol JL (1996) Interchromosomal transfer of epigenetic states in *Ascobolus*: transfer of DNA methylation is mechanistically related to homologous recombination. Cell 86:855–864

Condon BJ, Leng Y, Wu D, Bushley KE, Ohm RA, Otillar R, Martin J, Schackwitz W, Grimwood J, MohdZainudin N, Xue C, Wang R, Manning VA, Dhillon B, Tu ZJ, Steffenson BJ, Salamov A, Sun H, Lowry S, LaButti K, Han J, Copeland A, Lindquist E, Barry K, Schmutz J, Baker SE, Ciuffetti LM, Grigoriev IV, Zhong S, Turgeon BG (2013) Comparative genome structure, secondary metabolite, and effector coding capacity across *Cochliobolus pathogens*. PLoS Genet 9:e1003233

Connolly LR, Smith KM, Freitag M (2013) The Fusarium graminearum histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLOS Genet 9:e1003916

Cook HA, Eldredge NT, Haley MI, Kupke DW, Luck JM, Rasmussen PS (1956) On the fractionation of thymus histone. Arch Biochem Biophys 65:449–467

Croll D, Zala M, McDonald BA (2013) Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. PLoS Genet 9:e1003567

Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Munsterkotter M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317:1400–1402

Dai J, Hyland EM, Yuan DS, Huang H, Bader JS, Boeke JD (2008) Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. Cell 134:1066–1078

de la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA (2005) Do protein motifs read the histone code? Bioessays 27:164–175

de Wit PJ, van der Burgt A, Okmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, Bahkali AH, Beenen HG, Chettri P, Cox MP, Datema E, de Vries RP, Dhillon B, Ganley AR, Griffiths SA, Guo Y, Hamelin RC, Henrissat B, Kabir MS, Jashni MK, Kema G, Klaubauf S, Lapidus A, Levasseur A, Lindquist E, Mehrabi R, Ohm RA, Owen TJ, Salamov A, Schwelm A, Schijlen E, Sun H, van den Burg HA, van Ham RC, Zhang S, Goodwin SB, Grigoriev IV, Collemare J, Bradshaw RE (2012) The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. PLoS Genet 8:e1003088

Duncan IW (2002) Transvection effects in *Drosophila*. Annu Rev Genet 36:521–556

Durrin LK, Mann RK, Kayne PS, Grunstein M (1991) Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. Cell 65:1023–1031

Earnshaw WC, Migeon BR (1985) Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma 92:290–296

Ehinger A, Denison SH, May GS (1990) Sequence, organization and expression of the core histone genes of *Aspergillus nidulans*. Mol Gen Genet 222:416–424

Eissenberg JC, Shilatifard A (2010) Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Dev Biol 339:240–249

Ekwall K (2005) Genome-wide analysis of HDAC function. Trends Genet 21:608–615

Ellison CE, Hall C, Kowbel D, Welch J, Brem RB, Glass NL, Taylor JW (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. Proc Natl Acad Sci U S A 108:2831–2836

Erhard KF Jr, Hollick JB (2011) Paramutation: a process for acquiring trans-generational regulatory states. Curr Opin Plant Biol 14:210–216

Farboud B, Nix P, Jow MM, Gladden JM, Meyer BJ (2013) Molecular antagonism between X-chromosome and autosome signals determines nematode sex. Genes Dev 27:1159–1178

Faugeron G, Rhounim L, Rossignol J-L (1990) How does the cell count the number of ectopic copies of a gene in the premeiotic inactivation process

acting in *Ascobolus immersus*? Genetics 124:585–591

Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemieux S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, White OR, Whitty BR, Youngman P, Wolfe KH, Goldman GH, Wortman JR, Jiang B, Denning DW, Nierman WC (2008) Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet 4: e1000046

Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12:1052–1058

Fisch KM, Gillaspy AF, Gipson M, Henrikson JC, Hoover AR, Jackson L, Najar FZ, Wagele H, Cichewicz RH (2009) Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. J Ind Microbiol Biotechnol 36:1199–1213

Fischle W, Wang Y, Allis CD (2003) Binary switches and modification cassettes in histone biology and beyond. Nature 425:475–479

Folco HD, Freitag M, Ramon A, Temporini ED, Alvarez ME, Garcia I, Scazzocchio C, Selker EU, Rosa AL (2003) Histone H1 is required for proper regulation of pyruvate decarboxylase gene expression in *Neurospora crassa*. Eukaryot Cell 2:341–350

Freedman T, Pukkila PJ (1993) De novo methylation of repeated sequences in *Coprinus cinereus*. Genetics 135:357–366

Freeman L, Kurumizaka H, Wolffe AP (1996) Functional domains for assembly of histones H3 and H4 into the chromatin of *Xenopus embryos*. Proc Natl Acad Sci U S A 93:12780–12785

Freidkin I, Katcoff DJ (2001) Specific distribution of the *Saccharomyces cerevisiae* linker histone homolog HHO1p in the chromatin. Nucleic Acids Res 29:4043–4051

Freitag M, Hickey PC, Khlafallah TK, Read ND, Selker EU (2004) HP1 is essential for DNA methylation in *Neurospora*. Mol Cell 13:427–434

Frisvad JC (2012) Media and growth conditions for induction of secondary metabolite production. Methods Mol Biol 944:47–58

Gacek A, Strauss J (2012) The chromatin code of fungal secondary metabolite gene clusters. Appl Microbiol Biotechnol 95:1389–1404

Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Penalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115

Gan P, Ikeda K, Irieda H, Narusaka M, O'Connell RJ, Narusaka Y, Takano Y, Kubo Y, Shirasu K (2013) Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum fungi*. New Phytol 197:1236–1249

Gary JD, Clarke S (1998) RNA and protein interactions modulated by protein arginine methylation. Prog Nucleic Acid Res Mol Biol 61:65–131

Gasser SM, Cockell MM (2001) The molecular biology of the SIR proteins. Gene 279:1–16

Gibbons JG, Rokas A (2013) The function and evolution of the Aspergillus genome. Trends Microbiol 21:14–22

Goodwin SB, M'Barek SB, Dhillon B, Wittenberg AH, Crane CF, Hane JK, Foster AJ, Van der Lee TA, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B, Bowler J, Bristow J, van der Burgt A, Canto-Canche B, Churchill AC, Conde-Ferraez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RC, Hammond-Kosack KE, Henrissat B, Kilian A, Kobayashi AK, Koopmann E, Kourmpetis Y, Kuzniar A, Lindquist E, Lombard V, Maliepaard C, Martins N, Mehrabi R, Nap JP, Ponomarenko A, Rudd JJ, Salamov A, Schmutz J, Schouten HJ, Shapiro H, Stergiopoulos I, Torriani SF, Tu H, de Vries RP, Waalwijk C, Ware SB, Wiebenga A, Zwiers LH, Oliver RP, Grigoriev IV, Kema GH (2011) Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genet 7:e1002070

Goto DB, Nakayama J (2012) RNA and epigenetic silencing: insight from fission yeast. Dev Growth Differ 54:129–141

Gottschling DE (2007) Epigenetics: from phenomenon to field. In: Allis CD, Jenuwein T, Reinberg D (eds) Epigenetics. Cold Spring Harbor Laboratory Press, New York, pp 1–14

Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. Cell 63:751–762

Goyon C, Faugeron G (1989) Targeted transformation of *Ascobolus immersus* and de novo methylation of the resulting duplicated DNA sequences. Mol Cell Biol 9:2818–2827

Graia F, Lespinet O, Rimbault B, Dequard-Chablat M, Coppin E, Picard M (2001) Genome quality control: RIP (repeat-induced point mutation) comes to *Podospora*. Mol Microbiol 39:1–11

Grewal SI (2010) RNAi-dependent formation of heterochromatin and its diverse functions. Curr Opin Genet Dev 20:134–141

Grewal SI, Jia S (2007) Heterochromatin revisited. Nat Rev Genet 8:35–46

Grunstein M (1990) Histone function in transcription. Annu Rev Cell Biol 6:643–678

Haber JE (2012) Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. Genetics 191:33–64

Hall LE, Mitchell SE, O'Neill RJ (2012) Pericentric and centromeric transcription: a perfect balance required. Chromosome Res 20:535–546

Hays SM, Swanson J, Selker EU (2002) Identification and characterization of the genes encoding the core histones and histone variants of *Neurospora crassa*. Genetics 160:961–973

Henikoff S (1990) Position-effect variegation after 60 years. Trends Genet 6:422–426

Henrikson JC, Hoover AR, Joyner PM, Cichewicz RH (2009) A chemical epigenetics approach for engineering the in situ biosynthesis of a cryptic natural product from *Aspergillus niger*. Org Biomol Chem 7:435–438

Hollick JB (2012) Paramutation: a trans-homolog interaction affecting heritable gene regulation. Curr Opin Plant Biol 15:536–543

Honda S, Selker EU (2008) Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in *Neurospora crassa*. Mol Cell Biol 28:6044–6055

Honda S, Lewis ZA, Huarte M, Cho LY, David LL, Shi Y, Selker EU (2010) The DMM complex prevents spreading of DNA methylation from transposons to nearby genes in *Neurospora crassa*. Genes Dev 24:443–454

Honda S, Lewis ZA, Shimada K, Fischle W, Sack R, Selker EU (2012) Heterochromatin protein 1 forms distinct complexes to direct histone deacetylation and DNA methylation. Nat Struct Mol Biol 19(471–477):S1

Hori T, Fukagawa T (2012) Establishment of the vertebrate kinetochores. Chromosome Res 20:547–561

Hou H, Wang Y, Kallgren SP, Thompson J, Yates JR 3rd, Jia S (2010) Histone variant H2A.Z regulates centromere silencing and chromosome segregation in fission yeast. J Biol Chem 285:1909–1918

Idnurm A, Howlett BJ (2003) Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the *Dothideomycete Leptosphaeria maculans*. Fungal Genet Biol 39:31–37

Ikeda K, Nakayashiki H, Kataoka T, Tamba H, Hashimoto Y, Tosa Y, Mayama S (2002) Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. Mol Microbiol 45:1355–1364

Inglis DO, Binkley J, Skrzypek MS, Arnaud MB, Cerqueira GC, Shah P, Wymore F, Wortman JR, Sherlock G (2013) Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans, A. fumigatus, A. niger* and *A. oryzae*. BMC Microbiol 13:91

Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416:556–560

Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU (2013) Regional control of histone H3 lysine 27 methylation in *Neurospora*. Proc Natl Acad Sci U S A 110:6027–6032

Jeon Y, Sarma K, Lee JT (2012) New and Xisting regulatory mechanisms of X chromosome inactivation. Curr Opin Genet Dev 22:62–71

Jiang J, Liu X, Yin Y, Ma Z (2011) Involvement of a velvet protein FgVeA in the regulation of asexual development, lipid and secondary metabolisms and virulence in *Fusarium graminearum*. PLoS One 6:e28291

Kalashnikova AA, Winkler DD, McBryant SJ, Henderson RK, Herman JA, DeLuca JG, Luger K, Prenni JE, Hansen JC (2013) Linker histone H1.0 interacts with an extensive network of proteins found in the nucleolus. Nucleic Acids Res 41:4026–4035

Kale SP, Milde L, Trapp MK, Frisvad JC, Keller NP, Bok JW (2008) Requirement of LaeA for secondary metabolism and sclerotial production in *Aspergillus flavus*. Fungal Genet Biol 45:1422–1429

Karimi-Aghcheh R, Bok JW, Phatale PA, Smith KM, Baker SE, Lichius A, Omann M, Zeilinger S, Seiboth B, Rhee C, Keller NP, Freitag M, Kubicek CP (2013) Functional analyses of *Trichoderma reesei* LAE1 reveal conserved and contrasting roles of this regulator. G3 (Bethesda) 3:369–378

Kasinsky HE, Lewis JD, Dacks JB, Ausio J (2001) Origin of H1 linker histones. Faseb J 15:34–42

Kennison JA, Southworth JW (2002) Transvection in *Drosophila*. Adv Genet 46:399–420

Klose RJ, Zhang Y (2007) Regulation of histone methylation by demethylimination and demethylation. Nat Rev Mol Cell Biol 8:307–318

Knuf C, Nielsen J (2012) Aspergilli: systems biology and industrial applications. Biotechnol J 7:1147–1155

Koerner MV, Barlow DP (2010) Genomic imprinting-an epigenetic gene-regulatory model. Curr Opin Genet Dev 20:164–170

Kouzminova EA, Selker EU (2001) *Dim-2* encodes a DNA-methyltransferase responsible for all known cytosine methylation in *Neurospora*. EMBO J 20:4309–4323

Kowalski A, Palyga J (2012) Linker histone subtypes and their allelic variants. Cell Biol Int 36:981–996

Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ (2013) Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. Elife 2:e00808

Kuo MH, Xu XJ, Bolck HA, Guo D (2009) Functional connection between histone acetyltransferase Gcn5p and methyltransferase Hmt1p. Biochim Biophys Acta 1789:395–402

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410:116–120

Lachner M, O'Sullivan RJ, Jenuwein T (2003) An epigenetic road map for histone lysine methylation. J Cell Sci 116:2117–2124

Lacoste N, Utley RT, Hunter JM, Poirier GG, Cote J (2002) Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J Biol Chem 277:30421–30424

Lee JT, Bartolomei MS (2013) X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell 152:1308–1323

Lee I, Oh JH, Shwab EK, Dagenais TR, Andes D, Keller NP (2009) HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. Fungal Genet Biol 46:782–790

Lee HC, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC, Liu Y (2010) Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. Mol Cell 38:803–814

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH (2003) Suv39h-mediated histone h3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13:1192–1200

Lewis ZA, Honda S, Khlafallah TK, Jeffress JK, Freitag M, Mohn F, Schubeler D, Selker EU (2009) Relics of repeat-induced point mutation direct heterochromatin formation in *Neurospora crassa*. Genome Res 19:427–437

Lewis ZA, Adhvaryu KK, Honda S, Shiver AL, Knip M, Sack R, Selker EU (2010a) DNA methylation and normal chromosome behavior in *Neurospora* depend on five components of a histone methyltransferase complex, DCDC. PLoS Genet 6: e1001196

Lewis ZA, Adhvaryu KK, Honda S, Shiver AL, Selker EU (2010b) Identification of DIM-7, a protein required to target the DIM-5 H3 methyltransferase to chromatin. Proc Natl Acad Sci U S A 107:8310–8315

Li KK, Luo C, Wang D, Jiang H, Zheng YG (2010) Chemical and biochemical approaches in the study of histone methylation and demethylation. Med Res Rev. doi:10.1002/med.20228:1-53

Li Y, Wang C, Liu W, Wang G, Kang Z, Kistler HC, Xu JR (2011) The HDF1 histone deacetylase gene is important for conidiation, sexual reproduction, and pathogenesis in *Fusarium graminearum*. Mol Plant Microbe Interact 24:487–496

Ling X, Harkness TA, Schultz MC, Fisher Adams G, Grunstein M (1996) Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: redundant and position-independent functions in assembly but not in gene regulation. Genes Dev 10:686–699

Lu X, Wontakal SN, Emelyanov AV, Morcillo P, Konev AY, Fyodorov DV, Skoultchi AI (2009) Linker histone H1 is essential for *Drosophila* development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. Genes Dev 23:452–465

Lu X, Wontakal SN, Kavi H, Kim BJ, Guzzardo PM, Emelyanov AV, Xu N, Hannon GJ, Zavadil J, Fyodorov DV, Skoultchi AI (2013) Drosophila H1 regulates the genetic activity of heterochromatin by recruitment of Su(var)3–9. Science 340:78–81

Ma J, Wing RA, Bennetzen JL, Jackson SA (2007) Plant centromere organization: a dynamic structure with conserved functions. Trends Genet 23:134–139

Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie X, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RA, Chapman S, Coulson R, Coutinho PM, Danchin EG, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373

Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, Yamada O, Yamagata Y, Anazawa H, Hata Y, Koide Y, Komori T, Koyama Y, Minetoki T, Suharnan S, Tanaka A, Isono K, Kuhara S, Ogasawara N, Kikuchi H (2005) Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438:1157–1161

Maloisel L, Rossignol JL (1998) Suppression of crossing-over by DNA methylation in *Ascobolus*. Genes Dev 12:1381–1389

Manning VA, Pandelova I, Dhillon B, Wilhelm LJ, Goodwin SB, Berlin AM, Figueroa M, Freitag M, Hane JK, Henrissat B, Holman WH, Kodira CD, Martin J, Oliver RP, Robbertse B, Schackwitz W,

Schwartz DC, Spatafora JW, Turgeon BG, Yandava C, Young S, Zhou S, Zeng Q, Grigoriev IV, Ma LJ, Ciuffetti LM (2013) Comparative genomics of a plant-pathogenic fungus, *Pyrenophora tritici-repentis*, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. G3 (Bethesda) 3:41–63

Margueron R, Reinberg D (2011) The Polycomb complex PRC2 and its mark in life. Nature 469:343–349

Martienssen RA, Zaratiegui M, Goto DB (2005) RNA interference and heterochromatin in the fission yeast *Schizosaccharomyces pombe*. Trends Genet 21:450–456

McStay B, Grummt I (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. Annu Rev Cell Dev Biol 24:131–157

Mellone BG, Ball L, Suka N, Grunstein MR, Partridge JF, Allshire RC (2003) Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. Curr Biol 13:1748–1757

Meneghini MD, Wu M, Madhani HD (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112:725–736

Meyer BJ (2010) Targeting X chromosomes for repression. Curr Opin Genet Dev 20:179–189

Miura A, Nakamura M, Inagaki S, Kobayashi A, Saze H, Kakutani T (2009) An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J 28:1078–1086

Nakanishi S, Sanderson BW, Delventhal KM, Bradford WD, Staehling-Hampton K, Shilatifard A (2008) A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. Nat Struct Mol Biol 15:881–888

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafon A, Latge JP, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Pertea M, Price C, Pritchard BL, Quail MA, Rabbinowitsch E, Rawlins N, Rajandream MA, Reichard U, Renauld H, Robson GD, Rodriguez de Cordoba S, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, Vazquez de Aldana CR, Weidman J, White O, Woodward J, Yu JH, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438:1151–1156

Nowrousian M, Stajich JE, Chu M, Engh I, Espagne E, Halliday K, Kamerewerd J, Kempken F, Knab B, Kuo HC, Osiewacz HD, Poggeler S, Read ND, Seiler S, Smith KM, Zickler D, Kuck U, Freitag M (2010) De novo assembly of a 40 Mb eukaryotic genome from short sequence reads: *Sordaria macrospora*, a model organism for fungal morphogenesis. PLoS Genet 6:e1000891

Nutzmann HW, Reyes-Dominguez Y, Scherlach K, Schroeckh V, Horn F, Gacek A, Schumann J, Hertweck C, Strauss J, Brakhage AA (2011) Bacteria-induced natural product formation in the fungus *Aspergillus nidulans* requires Saga/Ada-mediated histone acetylation. Proc Natl Acad Sci U S A 108:14282–14287

Oppikofer M, Kueng S, Gasser SM (2013) SIR-nucleosome interactions: structure-function relationships in yeast silent chromatin. Gene 527:10–25

Palmer JM, Keller NP (2010) Secondary metabolism in fungi: does chromosomal location matter? Curr Opin Microbiol 13:431–436

Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J Cell Biol 104:805–815

Palmer JM, Bok JW, Lee S, Dagenais TR, Andes DR, Kontoyiannis DP, Keller NP (2013a) Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in *Aspergillus fumigatus*. PeerJ 1:e4

Palmer JM, Theisen JM, Duran RM, Grayburn WS, Calvo AM, Keller NP (2013b) Secondary metabolism and development is mediated by LlmF control of VeA subcellular localization in *Aspergillus nidulans*. PLoS Genet 9:e1003193

Patterton HG, Landel CC, Landsman D, Peterson CL, Simpson RT (1998) The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of *Saccharomyces cerevisiae*. J Biol Chem 273:7268–7276

Perrin RM, Fedorova ND, Bok JW, Cramer RA, Wortman JR, Kim HS, Nierman WC, Keller NP (2007) Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. PLoS Pathog 3:e50

Petroski MD, Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol 6:9–20

Pomraning KR, Connolly LR, Whalen JP, Smith KM, Freitag M (2013) Repeat-induced point mutation, DNA methylation and heterochromatin in *Gibberella zeae* (anamorph: *Fusarium graminearum*). In: Brown D, Proctor RH (eds) *Fusarium* genomics and molecular and cellular biology. Horizon Scientific Press, Norwich, p 93

Preuss S, Pikaard CS (2007) rRNA gene silencing and nucleolar dominance: insights into a

chromosome-scale epigenetic on/off switch. Biochim Biophys Acta 1769:383–392
Ptashne M (2013a) Epigenetics: core misconcept. Proc Natl Acad Sci U S A 110:7101–7103
Ptashne M (2013b) Faddish stuff: epigenetics and the inheritance of acquired characteristics. FASEB J 27:1–2
Rando OJ, Winston F (2012) Chromatin and transcription in yeast. Genetics 190:351–387
Rep M, Kistler HC (2010) The genomic organization of plant pathogenicity in *Fusarium* species. Curr Opin Plant Biol 13:420–426
Reyes-Dominguez Y, Bok JW, Berger H, Shwab EK, Basheer A, Gallmetzer A, Scazzocchio C, Keller N, Strauss J (2010) Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. Mol Microbiol 76:1376–1386
Rhounim L, Rossignol J-L, Faugeron G (1992) Epimutation of repeated genes in *Ascobolus immersus*. EMBO J 11:4451–4457
Rine J, Herskowitz I (1987) Four genes responsible for a position effect on expression from *HML* and *HMR* in *Saccharomyces cerevisiae*. Genetics 116:9–22
Rine J, Strathern JN, Hicks JB, Herskowitz I (1979) A suppressor of mating-type locus mutations in *Saccharomyces cerevisiae*: evidence for and identification of cryptic mating-type loci. Genetics 93:877–901
Robyr D, Suka Y, Xenarios I, Kurdistani SK, Wang A, Suka N, Grunstein M (2002) Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. Cell 109:437–446
Rountree MR, Selker EU (2010) DNA methylation and the formation of heterochromatin in *Neurospora crassa*. Heredity 105:38–44
Rudolph T, Beuch S, Reuter G (2013) Lysine-specific histone demethylase LSD1 and the dynamic control of chromatin. Biol Chem 394:1019–1028
Rusche LN, Kirchmaier AL, Rine J (2003) The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. Annu Rev Biochem 72:481–516
Salminen A, Kaarniranta K (2009) SIRT1 regulates the ribosomal DNA locus: epigenetic candles twinkle longevity in the Christmas tree. Biochem Biophys Res Commun 378:6–9
Santoro R (2005) The silence of the ribosomal RNA genes. Cell Mol Life Sci 62:2067–2079
Saze H, Shiraishi A, Miura A, Kakutani T (2008) Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. Science 319:462–465
Schuettengruber B, Martinez AM, Iovino N, Cavalli G (2011) Trithorax group proteins: switching genes on and keeping them active. Nat Rev Mol Cell Biol 12:799–814
Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP (2012) The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol 84:1150–1164
Selker EU (1990) Premeiotic instability of repeated sequences in *Neurospora crassa*. Annu Rev Genet 24:579–613
Selker EU (1998) Trichostatin A causes selective loss of DNA methylation in *Neurospora*. Proc Natl Acad Sci U S A 95:9430–9435
Selker EU, Garrett PW (1988) DNA sequence duplications trigger gene inactivation in *Neurospora crassa*. Proc Natl Acad Sci U S A 85:6870–6874
Selker EU, Stevens JN (1987) Signal for DNA methylation associated with tandem duplication in *Neurospora crassa*. Mol Cell Biol 7:1032–1038
Selker EU, Cambareri EB, Jensen BC, Haack KR (1987) Rearrangement of duplicated DNA in specialized cells of *Neurospora*. Cell 51:741–752
Shen X, Yu L, Weir JW, Gorovsky MA (1995) Linker histones are not essential and affect chromatin condensation in vivo. Cell 82:47–56
Shwab EK, Keller NP (2008) Regulation of secondary metabolite production in filamentous ascomycetes. Mycol Res 112:225–230
Shwab EK, Bok JW, Tribus M, Galehr J, Graessle S, Keller NP (2007) Histone deacetylase activity regulates chemical diversity in *Aspergillus*. Eukaryot Cell 6:1656–1664
Simon JA, Kingston RE (2009) Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10:697–708
Simon MD, Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, Borowsky ML, Kuroda MI, Kingston RE (2011) The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A 108:20497–20502
Sims RJ 3rd, Nishioka K, Reinberg D (2003) Histone lysine methylation: a signature for chromatin function. Trends Genet 19:629–639
Singer MJ, Marcotte BA, Selker EU (1995) DNA methylation associated with repeat-induced point mutation in *Neurospora crassa*. Mol Cell Biol 15:5586–5597
Smith KM, Kothe GO, Matsen CB, Khlafallah TK, Adhvaryu KK, Hemphill M, Freitag M, Motamedi MR, Selker EU (2008) The fungus *Neurospora crassa* displays telomeric silencing mediated by multiple sirtuins and by methylation of histone H3 lysine 9. Epigenetics Chromatin 1:5
Smith KM, Dobosy JR, Reifsnyder JE, Rountree MR, Anderson DC, Green GR, Selker EU (2010) H2B- and H3-specific histone deacetylases are required for DNA methylation in *Neurospora crassa*. Genetics 186:1207–1216
Smith KM, Phatale PA, Sullivan CM, Pomraning KR, Freitag M (2011) Heterochromatin is required for normal distribution of *Neurospora crassa* CenH3. Mol Cell Biol 31:2528–2542

Smith KM, Galazka JM, Phatale PA, Connolly LR, Freitag M (2012a) Centromeres of filamentous fungi. Chromosome Res 20:635–656

Smith KM, Phatale PA, Bredeweg EL, Connolly LR, Pomraning KR, Freitag M (2012b) Epigenetics of filamentous fungi. In: Myers RA (ed) Epigenetic regulation and epigenomics. Wiley-VCH, Weinheim, pp 1063–1105

Soukup AA, Chiang YM, Bok JW, Reyes-Dominguez Y, Oakley BR, Wang CC, Strauss J, Keller NP (2012) Overexpression of the *Aspergillus nidulans* histone 4 acetyltransferase EsaA increases activation of secondary metabolite production. Mol Microbiol 86:314–330

Stam M (2009) Paramutation: a heritable change in gene expression by allelic interactions in trans. Mol Plant 2:578–588

Straub T, Zabel A, Gilfillan GD, Feller C, Becker PB (2013) Different chromatin interfaces of the *Drosophila* dosage compensation complex revealed by high-shear ChIP-seq. Genome Res 23:473–485

Strauss J, Reyes-Dominguez Y (2011) Regulation of secondary metabolism by chromatin structure and epigenetic codes. Fungal Genet Biol 48:62–69

Stukenbrock EH, Dutheil JY (2012) Comparing fungal genomes: insight into functional and evolutionary processes. Methods Mol Biol 835:531–548

Stukenbrock EH, Bataillon T, Dutheil JY, Hansen TT, Li R, Zala M, McDonald BA, Wang J, Schierup MH (2011) The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Res 21:2157–2166

Stukenbrock EH, Christiansen FB, Hansen TT, Dutheil JY, Schierup MH (2012) Fusion of two divergent fungal individuals led to the recent emergence of a unique widespread pathogen species. Proc Natl Acad Sci U S A 109:10954–10959

Sullivan BA, Karpen GH (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol 11:1076–1083

Sun JQ, Hatanaka A, Oki M (2011) Boundaries of transcriptionally silent chromatin in *Saccharomyces cerevisiae*. Genes Genet Syst 86:73–81

Talbert PB, Henikoff S (2010) Histone variants – ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol 11:264–275

Tamaru H, Selker EU (2001) A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature 414:277–283

Tamaru H, Selker EU (2003) Synthesis of signals for de novo DNA methylation in *Neurospora crassa*. Mol Cell Biol 23:2379–2394

Tamaru H, Zhang X, McMillen D, Singh PB, Nakayama J, Grewal SI, Allis CD, Cheng X, Selker EU (2003) Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. Nat Genet 34:75–79

Tennen RI, Chua KF (2011) Chromatin regulation and genome maintenance by mammalian SIRT6. Trends Biochem Sci 36:39–46

Tormo JR, Asensio FJ, Bills GF (2012) Manipulating filamentous fungus chemical phenotypes by growth on nutritional arrays. Methods Mol Biol 944:59–78

Tribus M, Galehr J, Trojer P, Brosch G, Loidl P, Marx F, Haas H, Graessle S (2005) HdaA, a major class 2 histone deacetylase of *Aspergillus nidulans*, affects growth under conditions of oxidative stress. Eukaryot Cell 4:1736–1745

Tribus M, Bauer I, Galehr J, Rieser G, Trojer P, Brosch G, Loidl P, Haas H, Graessle S (2010) A novel motif in fungal class 1 histone deacetylases is essential for growth and development of *Aspergillus*. Mol Biol Cell 21:345–353

Trojer P, Brandtner EM, Brosch G, Loidl P, Galehr J, Linzmaier R, Haas H, Mair K, Tribus M, Graessle S (2003) Histone deacetylases in fungi: novel members, new facts. Nucleic Acids Res 31:3971–3981

Umemura M, Koyama Y, Takeda I, Hagiwara H, Ikegami T, Koike H, Machida M (2013) Fine de novo sequencing of a fungal genome using only SOLiD short read data: verification on *Aspergillus oryzae* RIB40. PLoS One 8:e63673

Upadhyay AK, Cheng X (2011) Dynamics of histone lysine methylation: structures of methyl writers and erasers. Prog Drug Res 67:107–124

van Holde K (1988) Chromatin. Springer, New York

van Leeuwen F, Gafken PR, Gottschling DE (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109:745–756

Verdaasdonk JS, Bloom K (2011) Centromeres: unique chromatin structures that drive chromosome segregation. Nat Rev Mol Cell Biol 12:320–332

Voigt P, Reinberg D (2011) Histone tails: ideal motifs for probing epigenetics through chemical biology approaches. Chembiochem 12:236–252

Vujatovic O, Zaragoza K, Vaquero A, Reina O, Bernues J, Azorin F (2012) Drosophila melanogaster linker histone dH1 is required for transposon silencing and to preserve genome integrity. Nucleic Acids Res 40:5402–5414

Waddington CH (1952) The epigenetics of birds. Cambridge University Press, Cambridge

Wang X, Sena Filho JG, Hoover AR, King JB, Ellis TK, Powell DR, Cichewicz RH (2010) Chemical epigenetics alters the secondary metabolite composition of guttate excreted by an atlantic-forest-soil-derived *Penicillium citreonigrum*. J Nat Prod 73:942–948

Weiner A, Chen HV, Liu CL, Rahat A, Klien A, Soares L, Gudipati M, Pfeffner J, Regev A, Buratowski S, Pleiss JA, Friedman N, Rando OJ (2012) Systematic dissection of roles for chromatin regulators in a yeast stress response. PLoS Biol 10:e1001369

Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, Spooner E, Li E, Zhang G, Colaiacovo M, Shi Y (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell 125:467–481

Wiemann P, Brown DW, Kleigrewe K, Bok JW, Keller NP, Humpf HU, Tudzynski B (2010) FfVel1 and FfLae1, components of a velvet-like complex in *Fusarium fujikuroi*, affect differentiation, secondary metabolism and virulence. Mol Microbiol 77:972–994

Wiemann P, Sieber CM, von Bargen KW, Studt L, Niehaus EM, Espino JJ, Huss K, Michielse CB, Albermann S, Wagner D, Bergner SV, Connolly LR, Fischer A, Reuter G, Kleigrewe K, Bald T, Wingfield BD, Ophir R, Freeman S, Hippler M, Smith KM, Brown DW, Proctor RH, Munsterkotter M, Freitag M, Humpf HU, Guldener U, Tudzynski B (2013) Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. PLoS Pathog 9:e1003475

Williams SC (2013) Epigenetics. Proc Natl Acad Sci U S A 110:3209

Williams RB, Henrikson JC, Hoover AR, Lee AE, Cichewicz RH (2008) Epigenetic remodeling of the fungal secondary metabolome. Org Biomol Chem 6:1895–1897

Wolffe AP, Khochbin S, Dimitrov S (1997) What do linker histones do in chromatin? Bioessays 19:249–255

Wyrick JJ, Parra MA (2009) The role of histone H2A and H2B post-translational modifications in transcription: a genomic perspective. Biochim Biophys Acta 1789:37–44

Xhemalce B, Dawson MA, Bannister AJ (2012) Histone modifications. In: Myers RA (ed) Epigenetic regulation and epigenomics. Wiley-VCH, Weinheim

Xiong L, Wang Y (2011) Mapping post-translational modifications of histones H2A, H2B and H4 in *Schizosaccharomyces pombe*. Int J Mass Spectrom 301:159–165

Xiong L, Adhvaryu KK, Selker EU, Wang Y (2010) Mapping of lysine methylation and acetylation in core histones of *Neurospora crassa*. Biochemistry 49:5236–5243

Xu H, Wang J, Hu Q, Quan Y, Chen H, Cao Y, Li C, Wang Y, He Q (2010) DCAF26, an adaptor protein of Cul4-based E3, is essential for DNA methylation in *Neurospora crassa*. PLoS Genet 6:e1001132

Yan H, Jin W, Nagaki K, Tian S, Ouyang S, Buell CR, Talbert PB, Henikoff S, Jiang J (2005) Transcription and histone modifications in the recombination-free region spanning a rice centromere. Plant Cell 17:3227–3238

Yang SM, Kim BJ, Norwood Toro L, Skoultchi AI (2013) H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. Proc Natl Acad Sci U S A 110:1708–1713

Yankulov K (2011) Dare to challenge the silence? Telomeric gene silencing revisited. Nucleus 2:513–516

Young TJ, Kirchmaier AL (2012) Cell cycle regulation of silent chromatin formation. Biochim Biophys Acta 1819:303–312

Yun M, Wu J, Workman JL, Li B (2011) Readers of histone modifications. Cell Res 21:564–578

Zaidi SK, Young DW, Montecino M, van Wijnen AJ, Stein JL, Lian JB, Stein GS (2011) Bookmarking the genome: maintenance of epigenetic information. J Biol Chem 286:18355–18361

Zhao Y, Shen Y, Yang S, Wang J, Hu Q, Wang Y, He Q (2010) Ubiquitin ligase components Cullin4 and DDB1 are essential for DNA methylation in *Neurospora crassa*. J Biol Chem 285:4355–4365

Zhao C, Waalwijk C, de Wit PJ, Tang D, van der Lee T (2013) RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. BMC Genomics 14:21

Zhou VW, Goren A, Bernstein BE (2011) Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet 12:7–18

Zofall M, Fischer T, Zhang K, Zhou M, Cui B, Veenstra TD, Grewal SI (2009) Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature 461:419–422

Zolan ME, Pukkila PJ (1986) Inheritance of DNA methylation in *Coprinus cinereus*. Mol Cell Biol 6:195–200

6 Photobiology and Circadian Clocks in *Neurospora*

Kevin K. Fuller[1*], Jennifer M. Hurley[1*], Jennifer J. Loros[1,2], Jay C. Dunlap[1]

CONTENTS

*These authors contributed equally

[1]Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, NH 03755, USA; e-mail: Jennifer.Loros@dartmouth.edu; Jay.C.Dunlap@Dartmouth.edu

[2]Department of Biochemistry, Geisel School of Medicine at Dartmouth, Hanover, NH 03755, USA; e-mail: Jennifer.Loros@dartmouth.edu

I. Introduction

Neurospora emerged as a durable model for dissection of the mechanisms underlying circadian regulation in the late 1950s (Pittendrigh et al. 1959). What recommended this system to those researchers, and to subsequent generations, have been the facts that it has a terrific genetic system, it is easy to use for biochemical follow-ups, and it is supported by a relatively large and supportive community. These advantages remain and keep *Neurospora* as a useful and popular model. Although a great deal of work on the photobiology of *Neurospora* emerged independent of circadian research, it is nonetheless true that a large impetus for dissecting the photobiology of *Neurospora* came from the fact that light is a major entraining agent for the circadian system. As a result, much of what we know about the molecular details of photobiology in fungi, nearly all of what we know about circadian rhythms in fungi, and even much of what we know about circadian rhythms in animals (which share a circadian mechanism similar to that of *Neuros-*

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

pora) can be traced to work on this system. For these historical reasons, and because the light-sensing system and circadian system share some components, it makes sense to cover both of these in a single review.

We begin with the photobiology of fungi in general and of *Neurospora* in particular, introducing the White Collar proteins that make up the principle photoreceptor and other real and potential photoreceptive proteins, and then describe how light sensing works. This nicely sets the stage for a discussion of the circadian system, from input via light using the White Collar photoreceptor, to the oscillator mechanism using the same proteins as well as others, to a discussion of how the circadian oscillator acts to control the metabolism of cells.

II. Photobiology of Fungi

A. Why Do Fungi Sense Light?

Despite the fact that fungi can neither visualize adjacent objects nor utilize electromagnetic energy for photosynthesis, it has been estimated that the majority of fungal species respond to light in some form (Idnurm et al. 2010). An intriguing question then arises: How does the perception of light provide a competitive advantage to a fungus? To understand this question, one may first look to processes that are commonly light regulated in diverse fungal species, principle among which is **sporulation**. One obvious role for light is that it likely serves as a positional marker for optimal spore dispersal into the open air. This is strikingly illustrated by the bending of sporangiophores of various Mucormycotina species (e.g., *Phycomyces* and *Mucor*) toward blue light and the phototropic orientation of protoperithecial beaks of *Neurospora crassa* (Linden et al. 1997; Corrochano and Garre 2010). More generally, light frequently serves as a signal to simply induce or repress sporulation in various zygomycete, ascomycete, and basidiomycete species and may even have differential effects on which form of sporulation takes place (Corrochano 2007; Bayram et al. 2010; Idnurm et al. 2010; Rodriguez-Romero et al. 2010). In *Phycomyces blakesleeanus*, for example, light promotes macrophorogenesis while inhibiting microphorogenesis, both of which are forms of asexual development (Corrochano and Cerd-Olmedo 1992).

Ultraviolet (UV) radiation represents a major source of genotoxic and oxidative stress for the cell. As such, the first photosensitive proteins to emerge in the prokaryotes were likely the photolyases, which utilize light in the blue/near-UV range to catalyze the repair of DNA lesions (Hitomi et al. 2000). The involvement of the light response in UV defense is evident in numerous fungal species. For example, various ascomycete, basidiomycete, and zygomycete species demonstrate a light-induced production of carotenoid or melanin **pigments**, which can absorb UV photons and convert the energy to heat (Chen and Loros 2009; Chen et al. 2009; Sano et al. 2009; Avalos and Estrada 2010; Corrochano and Garre 2010; Tisch and Schmoll 2010). In addition, light may also induce the expression of numerous DNA repair enzymes, including photolyases and UV endonucleases. Indeed, the loss of certain photoreceptor genes leads to hypersensitivity to UV stress in numerous species (Idnurm and Heitman 2005a; Bluhm and Dunkle 2008; Ruiz-Roldan et al. 2008). Moreover, light serves as a negative regulator of sexual development in numerous species (e.g., *Aspergillus nidulans*, *P. blakesleeanus*, *C. neoformans*) (Yamazaki et al. 1996; Blumenstein et al. 2005; Idnurm and Heitman 2005a); this may serve to prevent UV-mediated damage of energetically costly meiotic progeny, thereby providing an evolutionary basis for the connection between the regulation of development and genomic protection in fungi.

Essentially all organisms experience predictable changes in their environment associated with the **day/night cycles** of the rotating Earth. As such, the perception of light each morning may serve as an anticipatory signal for daily rise in UV (as discussed), increasing temperature and changes in nutrient availability. In at least a few fungal species, the light signal may entrain the **circadian** clock, which allows an organism to anticipate these daily changes, rather than simply respond to them each day (Greene et al. 2003; Lombardi and

Brody 2005; Bluhm et al. 2010; Baker et al. 2012). The molecular basis for circadian rhythmicity is discussed further in this chapter.

Taken together, these examples underscore the generalized importance of light in promoting fungal fitness. This then leads to the next important question: How does a fungus perceive light? As to be expected, the specific influence of light on an organism's physiology (light output) may vary markedly from species to species; however, the biochemical basis for photoperception (light input) is fundamentally similar and is achieved through specialized and conserved light-sensitive proteins.

B. Fungal Photoreceptors: Biochemistry and Distribution

Fungi are capable of responding to light qualities from the UV to far-red ends of the electromagnetic spectrum. This is accomplished through a variety of wavelength-specific photoreceptor proteins, some of which are highly conserved across the Mycota, whereas others are present in only discrete lineages (Table 6.1). The common feature of all photosensory proteins is that they interact with a **light-absorbing chromophore** that undergoes physicochemical and structural changes on photon absorption (Losi and Gärtner 2011). The protein must then respond to conformational changes in the chromophore by altering its own structural or biochemical properties. Although there are proteins whose function is directly modulated by light, such as photolyases and light-sensitive ion channels, the term *photoreceptor* is reserved for those that absorb light and initiate a signaling cascade to downstream targets and pathways.

1. Flavin-Based Blue Light Receptors

a) White Collar Orthologs

Strictly speaking, the first cloned fungal photoreceptor was White Collar-1 (WC-1) of *N. crassa* (Ballario et al. 1996; Losi and Gärtner 2011), although at the time it was cloned its identity as a photoreceptor was unknown. This was later established through proof that WC-1 bound the flavin-adenine dinucleotide (FAD) chromophore, and that the fluence response for the biochemical and molecular changes in the WC-1 protein in vitro matched the in vivo responses (Froehlich et al. 2002). The name *white collar* stems from the fact that the mutants are defective for light-induced carotenoid pigment production in their hyphae; consequently, in contrast to the completely orange appearance of light-exposed wild-type *Neurospora* test-tube slants, *wc-1* mutants display a nonpigmented mycelial border (a white collar) beneath the constitutively pigmented conidia at the top of the slant. The *Neurospora* WC-1, along with its essential partner WC-2 (so named because its mutants yield the same phenotype as those of *wc-1*) (Linden et al. 1997), is by far the most well-characterized fungal photoreceptor in terms of its biochemistry and molecular biology. The description of the WCC that follows stems from work done in *Neurospora* but is presumed to apply to many White Collar orthologs found in other species based on sequence conservation.

The WC-1 proteins are GATA-like Zn-finger transcription factors that contain three **PAS (Per-Ant-Sim) domains**. The N-terminal-most PAS domain is of a special subclass, called the **LOV domain** (for light, oxygen, and voltage), which noncovalently binds a single blue-light-absorbing chromophore, FAD. Through its C-terminal-most PAS domain, WC-1 interacts with the PAS domain of WC-2 to form a heterodimer known as the **White Collar Complex (WCC)** (Linden et al. 1997; Cheng et al. 2002). WC-2 also contains a Zn-finger domain, but lacks a LOV domain for direct light sensing. A simple model for WCC signaling proposes that when the FAD absorbs blue light it forms a transient cysteinyl adduct with the WC-1 LOV domain, thereby inducing a conformational change that promotes the transcriptional activity of the complex (Froehlich et al. 2002; He et al. 2002; Zoltowski et al. 2007). The WCC binds to specific sequences (called light responsive elements, LREs) in the promoters of its target genes, where it is primarily described as a transcriptional activator; however, the downregulation of genes in response

Table 6.1. Conservation of photoreceptors and FRQ across the Mycota

	WC-1	WC-2	VVD	CRY	OPS	OPS-REL	PHY	FRQ
Ascomycetes								
Saccharomyces cerevisiae						3		
Candida albicans						2		
Fusarium oxysporum	1	1	1	2	3	2	1	1
Magnaporthe oryzae	1	1	1	2		1	1	1
Neurospora crassa	1	1	1	1	1	1	2	1
Trichoderma atroviride	1	1	1	1		1	1	1
Trichoderma reesei	1	1	1	1			1	1
Botrytis cinerea	1	1	1	1	2		2	1
Sclerotinia sclerotiorum	1	1	1	1	2		3	1
Stagonospora nodorum	1	1		2	2		2	1
Histoplasma capsulatum	1	1				1		
Coccidioides immitis						1		
Paracoccidioides brasiliensis	1	1				1		
Aspergillus nidulans	1	1		1[a]		1	1	
Schizosaccharomyces pombe						1		
Basidiomycetes								
Phanerochaete chrysosporium	1	1				5	1	
Coprinus cinerea	1	1					1	
Cryptococcus neoformans	1	1			1		1	
Malassezia globosa								
Ustilago maydis	1	1		2	2	1	1	
Sprobolomyces roseus	2	2		1	2		1	
Puccinia graminis	1	1				5	1	
Zygomycetes								
Phycomyces blakesleeanus	3	4		1				
Rhizopus oryzae	3	5		1				
Mucor circinelloides	3	4		1				
Encephalitozoon cuniculi								
Chytrids								
Batrachochytrium dendrobatidis								
Spizelomyces punctatus	2	2					1	

The distribution of photoreceptor orthologs in selected species from various fungal lineages is shown. The numbers in the chart reflect the number of paralogs present within the organism's genome, and the absence of a number reflects the absence of any detectable ortholog. Orthologs of the *Neurospora* clock protein FREQUENCY (FRQ) are also shown, a discussion of which is found in the section on circadian clocks. *CRY* chryptochrome, *PHY* phytochrome, *OPS* rhodopsin, *OPS-REL* opsin-related protein, *VVD* vivid, *WC-1* White Collar 1, *WC-2* White Collar 2 (Information in this chart is integrated from several sources: Idnurm et al. (2010), Dunlap and Loros (2006), and Salichos and Rokas 2010)

[a]For the cryptochrome category, the CPD photolyase class is not included; the only exception is CryA of *A. nidulans*, which has a demonstrated signaling function

to signaling mediated by the WCC has been reported in both *A. nidulans* and *Trichoderma atroviride* (Rosales-Saavedra et al. 2006; Purschwitz et al. 2008; see Chap. 10 in this volume), although whether this is a direct action of the WCC is unclear.

With orthologs present in the genomes of essentially all fungal lineages, including two species of the anciently diverged chytrids (*Spizellomyces punctatus* and *Allomyces macrogynus*), the White Collars are the most well-conserved fungal photoreceptors in terms of evolutionary depth (Idnurm et al. 2010). There are a few organisms that notably lack *wc-1/2* orthologs, including most of the Saccharomycotina yeasts, which is likely a result of gene loss in those lineages (Dunlap and Loros 2006). This also includes numerous pathogenic species, such as the ascomycete yeast *Candida*, the basidiomycete yeasts *Malassezia* and *Microsporidia*, and numerous dermatophytes (*Microsporum* and *Trichophyton* spp.). As

these organisms represent obligate pathogens with a close association with their hosts, it is possible that genomic reductions have led to the loss of superfluous genes involved in environmental light sensing. Conversely, other fungal pathogens, including *Aspergillus fumigatus*, *Cryptococcus*, *Paraccocidioides*, and *Histoplasma* all have an environmental life cycle as well as a *wc-1* ortholog (Idnurm et al. 2010).

In contrast to the loss of *wc-1* in some lineages, sequence analysis of several Mucormycotina has revealed multiple copies of the *wc-1* orthologs, a likely reflection of gene duplication events early in that lineage (Idnurm et al. 2010). Interestingly, the WC paralogs may display specialized roles in the light response, such as in *Mucor circinelloides*, in which one copy, *mcwc-1a*, regulates phototropism, and a second copy, *mcwc-1c*, regulates the photoinduction of carotenoids (Silva et al. 2006). Interestingly, a third paralog, *mcwc-1b*, is an activator of carotenogenesis but a repressor of sporulation (Silva et al. 2008).

Theoretically, WC-1 should function alone as a photoreceptor, as it contains both light-sensing (LOV) and DNA-binding (Zn-finger) domains. In *Neurospora*, however, WC-2 is essential for the light response because its Zn-finger domain is required for binding of the WCC to DNA (Collett et al. 2002; Froehlich et al. 2002; Liu et al. 2003). This essential interaction seems to have occurred early in fungal evolution, as a *wc-2* ortholog has been found in all species that contain *wc-1*. Indeed, many of the WC-1 proteins of the basidiomycetes, including *C. neoformans*, *Ustilago maydis*, *Phanerochaete chrysosporium*, and *Coprinus cinereus*, have lost their Zn-finger domains altogether (Idnurm and Heitman 2005a).

b) Small LOV Domain Proteins: VIVID and ENVOY

In addition to WC-1 homologs, several species encode a small protein that is essentially a single LOV domain with an N-terminal cap (Zoltowski et al. 2007). The first identified member of this family was the 186-amino-acid protein in *N. crassa* and was named VIVID due to the bright orange pigmentation phenotype of *vvd* mutants grown in light (Heintzen et al. 2001). Like, WC-1, VIVID binds FAD as a chromophore (Schwerdtfeger and Linden 2003).

VIVID homologs seem to be limited to the genomes of a few ascomycetes. An apparent ortholog is the 207-amino-acid protein, ENVOY, in *Hypocrea jecorina*, which is involved in regulation of cellulose in the dark (Schwerdtfeger and Linden 2003). Despite the fact that VIVID and ENVOY contain only a LOV domain, the two genes do not cross complement one another. Moreover, ENVOY seems to play a role in up- and downregulating genes in both the dark and the light (Schuster et al. 2007). Therefore, it is unclear how these small proteins facilitate their downstream function, but as will be discussed for *N. crassa*, likely involves direct protein-protein interactions with WC-1 via the LOV domain.

c) Cryptochromes

Photolyases and cryptochromes are members of a family of blue/UV-A light receptors. Both proteins contain an N-terminal domain (photolyase-related [PHR] region) that binds noncovalently to two chromophores, FAD and 5,10-methenyltetrahydrofolate (MTHF). **Photolyases** are typically characterized by their ability to use light energy to catalyze DNA repair (e.g., pyrimidine dimers), whereas **cryptochromes** have lost DNA repair capability and instead serve as bona fide light-signaling proteins (Losi 2007). Comparative genomics and functional analyses can divide photolyase/cryptochrome proteins into several family members, which are sporadically distributed among species in the ascomycetes, basidiomycetes, and zygomycetes (Mucormycotina). In general, bona fide photolyase orthologs (cyclobutane pyrimidine dimer, CPD, photolyases) are found in essentially all species, where they serve as light-activated DNA repair enzymes (Idnurm et al. 2010). This discussion focuses on those family members with a proposed or demonstrated photosensory function, the cryptochromes; however, as will be seen, some proteins display both repair and signaling capabilities.

The cryptochrome-*Drosophila*, *Arabidopsis*, *Synechocystis*, human (CRY-DASH) family is

believed to represent the earliest divergence of cryptochromes from photolyases as it is the only cryptochrome found in bacteria. Several fungi have CRY-DASH members, including the ascomycetes *N. crassa*, *Sclerotinia sclerotiorum*, and *Stagonospora nodorum* and the basiodiomycete *U. maydis*. CRY-DASH proteins display minimal to no photorepair activity for double-stranded DNA (dsDNA) but do display activity against cyclobutane adducts on single-stranded DNA (ssDNA) (Selby and Sancar 2006). Their involvement in fungal light sensing is poorly understood but has been described in *Sclerotinia sclerotium*, in which it plays a small role in both promoting sclerotium accumulation and repressing hyphal pigmentation in response to UV-A light (Veluchamy and Rollins 2008).

The corn gray leaf spot pathogen, *Cercospora zeae-maydis*, encodes a class 1 CPD photolyase, called CPD1, as well as a member of the so-called 6-4 photolyases, called PHL1. Interestingly, PHL1 not only plays a role in photoreactivation following UV treatment but also regulates the light induction of CPD1 and other DNA repair genes. In addition, *phl1* regulates sporulation and cercosporin biosynthesis. Therefore, PHL1 displays true photosensory functions in the fungus (Bluhm and Dunkle 2008).

Interestingly, several species of *Aspergillus*, including *A. nidulans*, *A. fumigatus*, *A. flavus*, *A. terreus*, and *A. oryzae*, each encode only a single member of the photolyase/cryptochrome family (Bayram et al. 2010). Their sequence most closely resembles a class I CPD photolyase; however, the ortholog of *A. nidulans* demonstrates both photolyase activity and developmental regulation in response to light (Bayram et al. 2008a). It is unclear if the dual function of this protein reflects a photolyase that gained signaling function, or if it represents an ancestral form of photolyase/cryptochrome proteins that displayed both functions before their divergence. Orthologs in other species have not been characterized.

2. Red-Light-Sensing Phytochromes

Although blue light responses dominate the fungal kingdom, many species overtly respond to light above 600 nm (red light). Around 40 years ago, in fact, the ability of red light to reverse the inhibitory effect of blue light on conidiation was described in *Alternaria solani* and *Bortytis cinerea* (Lukens 1965; Tan 1974). Since then, the regulation of growth and development by red light has been described in *A. nidulans* (Bayram et al. 2010), *T. atroviride* (Casas-Flores et al. 2004), *Magnaporthe oryzae* (Lee et al. 2006), *Puccinia graminis* (Schneider and Murray 1979), and *A. fumigatus* (K. Fuller unpublished data), among others.

In both plants and bacteria, red light responses are achieved through phytochromes, which bind the linear tetrapyrrole bilin as a chromophore. The phytochrome protein coverts between two **conformational states**, the red light absorbing Pr confirmation and the far-red light absorbing Pfr confirmation, with the ratio of the two states determining the signaling state within the cell (Idnurm and Heitman 2005b). All phytochromes, regardless of type, have a modular architecture consisting of an N-terminal photosensory module (GAF [cGMP-specific phosphodiesterase, adenylyl cyclase, FhlA] and PHY [phytochrome] domains), which binds the bilin chromophore. The C-terminal output domains can vary considerably between species but will always contain a histidine kinase as part of the regulatory output (Blumenstein et al. 2005).

Phytochromes are also present in the genomes of many fungi, and overall sequence analyses indicate they are most similar to bacterial phytochromes, suggesting perhaps an evolutionary origin (Blumenstein et al. 2005; Karniol et al. 2005). Like the cryptochromes, the phytochromes are found sporadically throughout the Mycota. They are widely conserved among the ascomycete and basidiomycete phyla, and a putative ortholog was even identified in the

chytridiomycete *Spizelomyces punctatus* (Idnurm et al. 2010). However, despite the described red light responses of several species and the presence of phytochromes within their genomes, only FphA of *A. nidulans* has known regulatory functions (Idnurm and Heitman 2005b). Deletion of the phytochrome genes in both *C. neoformans* and *N. crassa* does not lead to obvious photobiological defects (Froehlich et al. 2005; Idnurm and Heitman 2005a); however, neither species is overtly responsive to red light. Molecular studies involving additional red-light-sensing fungi will likely demonstrate a conserved role for phytochrome in the fungal red light response.

3. Rhodopsins and Opsin-Related Proteins

Opsins are a family of seven-transmembrane receptors that bind retinal as a chromophore via a Schiff base linkage to a conserved lysine residue. Based on spectral analyses, opsins typically absorb light in the green to blue-green portions of the spectrum and, on absorption, can use the light signal to facilitate ion transport or for downstream light signaling. Generally, rhodopsins are divided into two categories, distinguished by highly divergent sequence homology between the two. The first is the **type II rhodopsins**, which include the G-protein-coupled receptors used primarily for vision in the higher eukaryotes; these proteins bind 11-*cis*-retinal as the chromophore (Brown 2004). A type II opsin has been reported in the chytrid *Allomyces reticulatus*, and it is clearly implicated in zoospore phototaxis toward green light (Sharma and Foster 1997); however, no orthologs of this type have been identified outside chytrid fungi and may represent an independent gene transfer after chytrid split from the other fungi.

The second type is the **type I rhodopsins**, which are found in the Archea, Bacteria, Fungi, and other lower eukaryotes and have either a transport or sensing role; these proteins bind all-*trans*-retinal as a chromophore. Most fungal rhodopsins are type I, and comparative genome sequencing suggested a horizontal gene transfer event from a haloarchea into an ancestor of the ascomycete and basidiomycete fungi (Sharma et al. 2006). The role for these proteins in light signaling is poorly understood as their disruption in several organisms has not yielded a robust photoresponsive defect, for example, *N. crassa* (Bieszke et al. 1999, 2007), *Gibberella (Fusarium) fujikuroi* (Estrada and Avalos 2009), and *C. neoformans* (Idnurm and Heitman 2005a). Instead, these proteins may have a function in ion transport, such as that described in *Leptosphaeria maculans* (Waschuk et al. 2005).

Understanding the evolutionary basis for fungal opsins is complicated by the fact that most species contain a family of **opsin-related proteins**, which are distinguished based on their absence of the conserved lysine residue required for binding the retinal chromophore (Brown 2004). These proteins display a wide distribution and are even present in multiple copies in the genomes of the Saccharomycotina yeasts, which lack a light response or additional photoreceptors (Idnurm et al. 2010).

4. Conclusions

Although the various photoreceptor families were discussed separately for the sake of exposition, it is important to consider that the light response undoubtedly involves the integration of multiple light inputs via discrete photosensory pathways. For example, the aforementioned positive phototropism of *P. blakesleeanus* sporangiophores is positive for blue light but negative for UV light (Sharma et al. 2006). This illustrates the complexity of the phenotype, considering both wavelengths are likely present at the same time. Moreover, the relative spectral quality and light intensities change over the course of the day, with blue/near-UV peaks near midday and red/far-red peaks at morning and dusk. Regardless, most laboratory experiments are performed under constant conditions. Therefore, photoreceptors that have been reported as having no contribution to an organism's light response (e.g., rhodopsin in *Cryptococcus*) may indeed have subtle, yet important, functions in a natural light environment. Given this, we now discuss what is known about light sensing in two

organisms in which multiple photosensors have been analyzed in some detail.

C. Photobiology of *Neurospora crassa*

1. The Light Response of *Neurospora*

The combination of genetic, physiologic, biochemical, and molecular analyses have made the ascomycete mold *Neurospora crassa* the best-understood fungal system with regard to photosensation. At the phenotypic level, a variety of processes are light regulated in *Neurospora*, including the resetting of the circadian clock, biosynthesis of carotenoid pigments, conidiation, the formation of protoperithecia, and the directionality of ascospore release from perithecia (Lee et al. 2003; Chen and Loros 2009). In recent years, a variety of genome-wide transcriptional analyses have provided novel insight into the influence of light on *Neurospora* beyond those readily observed (Lewis et al. 2002; Chen et al. 2009). Using oligonucleotide arrays, for example, Chen et al. found that approximately 5.6 % of detectable transcripts were influenced by a white light stimulus. Among the light-induced genes were those involved in pigmentation, vitamins and prosthetic groups, secondary metabolism, DNA processing, cell signaling, oxidation of fatty acids, and oxygen detoxification (Chen et al. 2009).

The *Neurospora* genome encodes orthologs to all the photoreceptor types described, including the White Collar genes, a small LOV domain protein, two phytochromes, a cryptochrome, and an archaeal rhodopsin. Interestingly, then, essentially all light responses in *N. crassa*, at both the phenotypic and transcriptional levels, are ablated in either a *wc-1 or wc-2* mutant background. Indeed, screens for mutants that are completely "blind" have yielded over 30 mutants to date, all of which have been alleles of either *wc-1* or *wc-2* (Bieszke et al. 1999; Collett et al. 2002; Froehlich et al. 2005, 2010; Chen and Loros 2009). This suggests the two genes are the only essential regulators of the light response in this fungus, although other photoreceptors such as VVD (vivid) may modify the response. The genetic data agree with the fact that all aforementioned photoreponses in *N. crassa* occur in response to blue light, with the action spectrum peaking at about 465 nm, the same as the absorption spectrum of flavin chromophores. Since the cloning of *wc-1* and *wc-2* in 1996 and 1997, respectively and the proof of their identity as light receptors in 2002, extensive analyses have made the WCC the paradigm for light regulation in fungi.

2. Molecular Basis for Signaling Through the White Collar Complex

As previously described, WC-1 and WC-2 interact to form a heterodimeric complex (WCC) that binds to the LREs in the promoters of light-regulated genes. However, various studies have revealed that the WCC of *Neurospora* actually exists in two forms (Froehlich et al. 2002; Cheng et al. 2003). The first is a simple **WC-1/2 heterodimer** (small complex) that binds to promoter elements most strongly in the dark. It is this small WCC that drives the circadian clock in the dark by promoting the expression of *frq*, the protein of which in turn inhibits WCC expression to form a free-running negative feedback loop that cycles every 22.5 h. The ability of the small WCC to drive *frq* expression is independent of the LOV domain of WC-1 and is thus independent of the photosensing role of the WCC.

On light exposure, the small complex is replaced on the DNA by a larger WCC, which consists of the WC-1/2 heterodimer with the addition of several more WC-1 proteins that interact via their LOV domains. Thus, WC-1 and WC-2 interact to form both a light-independent complex that drives the expression of the circadian clock and a light-dependent complex that drives the photororesponse. Although WC-2 is the limiting protein in the light-induced WCC, its Zn-finger domain is required for DNA binding; therefore, the light-sensing function of WC-1 and the DNA-binding function of WC-2 makes both proteins

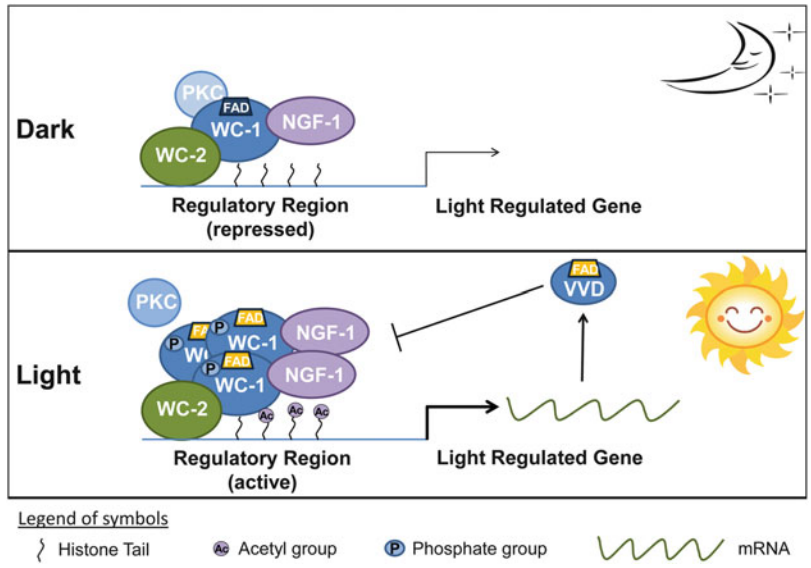


Fig. 6.1. Model for the light-induction of genes by the White Collar Complex (*WCC*). *Top* In the dark, the WC-1/2 heterodimer is preassembled on light-responsive elements (*LREs*) along with NGF-1 and protein kinase C (*PKC*) via their interaction with WC-1. In this state, expression of light-regulated genes is low. *Bottom* In the presence of light, flavin-adenine dinucleotide (*FAD*) absorbs a photon, which leads to conformational changes in the LOV (light, oxygen, and voltage) domain of WC-1. This conformational change promotes LOV domain interactions with additional WC-1 proteins and increases the acetyltrasferase activity of NGF-1. As a result, histone tails are transiently acetylated, leading to chromatin remodeling and an increase in transcriptional activity of the WCC. The activity of the WCC is negatively regulated in two ways: (*1*) WC-1 is rapidly phosphorylated, at first by PKC and then subsequently by others, which leads to its degradation. (*2*) *vvd* expression is induced by the WCC, the protein of which is biochemically activated by light and interrupts the activity of the WCC through a direct LOV domain interaction with WC-1 (This model is modified from a figure in Brenna et al. (2012))

essential in the transmission of the light signal (Linden and Macino 1997; Cheng et al. 2002; Collett et al. 2002).

The induction of light-regulated genes also involves chromatin modifications. For example, histone H3-K14 at light-inducible promoters is transiently acetylated by the **histone acetyltransferase** (HAT) NGF-1 (Grimaldi et al. 2006). Interestingly, acetylation of H3-K14 is transiently decreased at the *frq* promoter (Belden et al. 2007b). The NGF-1-mediated acetylation seems essential for light regulation and is dependent on a direct interaction with WC-1 (Brenna et al. 2012). Interestingly, it was shown that the WCC/NGF-1 complex occurs in the dark. Therefore, the current model for the light induction of target genes is as follows: (a) The small WC-1/2 heterodimer, in complex with NGF-1 via a WC-1 interaction, binds LREs in the dark. (b) Flavin absorbs blue light, which leads to the formation of a cysteine adduct with the WC-1 LOV domain, thereby changing its conformation and allowing for the recruitment of additional WC-1 proteins to the complex (large WCC). (c) The conformational change also enhances the HAT activity of NGF-1 and perhaps changes in other HAT activities, leading to (d) chromatin remodeling and enhanced transcriptional activity of the WC1/2 heterodimer, ultimately leading to the induction of WCC target genes (summarized in Fig. 6.1).

Not all light-induced genes are a direct target of the WCC. Instead, the WCC regulates a subset of genes known as the early-induced genes, which demonstrate transcriptional induction as early as 5 min (average peak at 30 min) post–light transfer (Chen et al. 2009; Smith et al. 2010).

Included in the early-induced genes are additional transcription factors that regulate a second group of late-light-induced genes that show induction as between 45 and 90 min (average peak at 60 min) post–light transfer. For example, Sub-1 is a transcription factor that is directly induced by the WCC, and its deletion results in the loss of most late-light-induced genes (Chen et al. 2009). Thus, the *Neurospora* light response is facilitated by a **hierarchy of transcriptional regulators** but is entirely dependent on the initial perception of light and early transcriptional activity of the WCC.

3. Negative Regulation of the White Collar Complex by Protein Kinases and VVD

Fungi must be able to attenuate the light response to maintain sensitivity to increasing light intensities. To illustrate, carotenoids protect cells from genotoxic and photooxidative stress, but their continued synthesis may represent an unnecessary energy expenditure in low-light conditions; therefore, the amount of carotenoids produced should correlate to the light intensity. This process, termed *photoadaptation*, is obvious in *N. crassa* as the steady-state transcript levels of light-induced genes begin to decrease over time under constant-light conditions (Chen et al. 2009). In *Neurospora*, photoadaptation is achieved primarily through direct inhibition of the WCC, one mechanism of which is through **hyperphosphorylation**, and ultimate turnover, of WC-1. Protein kinase C (PKC), for example, directly interacts with and phosphorylates WC-1 (Franchi et al. 2005). The importance of this in the attenuation of the light response is demonstrated by the fact that PKC inhibition leads to an increase in WC-1 levels and an increase in *albino-3* (involved in carotenogenesis) expression (Arpaia et al. 1999); however, PKC interacts with WC-1 only in the dark, suggesting that PKC can only be responsible for early WC-1 phosphorylation (Fig. 6.1). Therefore, additional kinases are likely involved in the hyperphosphorylation seen after 20 min post–light exposure.

A major advance in understanding photoadaption came with the identification of the small LOV domain protein VIVID (VVD). The name stems from its loss-of-function mutants, which display bright orange conidia when grown in constant light, attributed to the persistence activation of carotenoid pigments (i.e., loss of photoadaptation) (Heintzen et al. 2001). It is a direct target of the WCC, and *vvd* transcript levels are robustly induced following illumination. Various studies have now shown that, on light reception by VVD, its LOV domain interacts with the WC-1 LOV domain, ultimately blocking the transcriptional activity of the WCC (Fig. 6.1) (Chen et al. 2010; Hunt et al. 2010; Malzahn et al. 2010). Ultimately, the amount of WCC activity is dependent on the amount of light-activated VVD present in the system, which itself is a function of the WCC-mediated light response.

4. Additional Photoreceptors in *Neurospora*: All You Need Is LOV?

a) CRY

Neurospora contains a single ortholog to the CRY-DASH family of cryptochromes, called CRY. CRY can bind both FAD and MTHF as chromophores and, interestingly, to both single- and double-stranded DNA and RNA; however, it is not required for photoinduced DNA repair. Similar to *wc-1*, both transcript and protein levels of CRY rapidly accumulate in a WC-1 dependent manner; however, the disruption of the *cry* gene does not lead to any observable photobiology phenotype or seem to interfere with the normal light-regulated transcriptional response. Interestingly, however, CRY may play a role in the circadian clock, as its deletion leads to a small phase delay under light entrainment regimens (Froehlich et al. 2010).

b) PHY-1/PHY-2

Although red light does not regulate the aforementioned growth and developmental programs in *Neurospora*, a half-century ago it was reported that far-red light potentiated X-ray-induced genetic damage in this fungus. Furthermore, this effect was indicative of a

phytochrome response as it could be reversed by exposure to red light following the far-red light treatment (Klein and Klein 1962). *Neurospora* encodes two putative phytochromes for red light sensing, called PHY-1 and PHY-2. The two proteins display similar domain architecture, including an N-terminal domain for chromophore binding and a C-terminal histidine kinase and response regulator motifs, indicating that they may function as two-component sensor kinases. In vitro studies of PHY-2 have indeed shown that the protein can bind bilin or phycocyanobilin, and the holoprotein displays both red and far-red absorption spectra. Unlike *cry-1*, transcripts of neither *phy-1* nor *phy-2* are light induced, and deletion of either does not affect circadian phasing, although the expression of *phy-1* is regulated in a circadian fashion (Froehlich et al. 2005). Similar to CRY, however, the phytochromes have no detectable light-signaling function, as disruption of *phy1/2* does not affect photoinducible phenotypes or gene expression (Froehlich et al. 2005; Chen et al. 2009).

c) Neurospora Opsin (NOP-1)

A single rhodopsin ortholog is present in the *Neurospora* genome, called neurospora opsin (NOP-1). Characteristic of archael rhodopsins, NOP-1 has a seven-transmembrane topology and has been shown to bind all-*trans*-retinal in vitro. The expression of *nop-1* is most prominently induced by conditions that promote conidiation; however, the gene does demonstrate a modest influence by light, with more accumulation of messenger RNA (mRNA) in light-grown cultures (Bieszke et al. 1999). Interestingly, the *nop-1* deletion mutant does display light-dependent morphological differences (with respect to wild type) in the presence of the mitochondrial ATPase inhibitor oligomycin, suggesting it does have a biochemical function as a light-dependent proton pump (Bieszke et al. 1999). Similar to the *cry* and *phy-1/2* mutants, the role of NOP-1 as a light sensor seems to be limited as the mutant displays essentially normal patterns of development (conidiation and catotenogenesis) and gene induction in response to light (Bieszke et al. 1999; Chen et al. 2009).

d) A Role for the Secondary Photoreceptors in Inhibition of the White Collar Complex?

Despite the evidence that all of the *Neurospora* photoreceptors mentioned are expressed and display chromophorebinding capability, their involvement in the photoresponse has remained elusive. However, a study has shown that the gene *con-10* is induced earlier and to a greater extent in a *cry*, *nop-1*, or *phy-2* deletion background, compared to wild-type (WT) (Olmedo et al. 2010). Therefore, while the WCC is essential for driving the expression of light-induced genes, the additional photoreceptors may play a role in inhibiting WCC activity, similar to VVD. A direct interaction of these photoreceptors with the WCC has not been shown, however, so their effect on the WCC presumably requires an intermediary inhibitory protein.

D. Photoperception in *Aspergillus*: Seeing the Rainbow Through Multiple Photoreceptors

1. The *Aspergillus nidulans* Light Response

Aspergillus nidulans has served as a model organism for several decades, but only in recent years has its photobiology been the subject of detailed investigation. *Aspergillus nidulans* distinguishes itself from *Neurospora* by being overtly responsive to both blue and red portions of the visible spectrum and, consequently, having a light response that is mediated by multiple photosensory pathways. Accordingly, *A. nidulans* has quickly emerged as the predominant model for fungal **red-light sensing** and **photosensory cross-talk**. The major recent advances are highlighted.

In contrast to *Neurospora*, in which light induces both the sexual and asexual cycle, the two processes are differentially regulated by light in *A. nidulans*. In the presence of white light, asexual developmental genes are induced, and conidiation represents the dominant form of sporulation (Ruger-Herreros et al. 2011). Either blue or red light alone is capable of inducing conidiation, but only to about one third the level of white light. However, the combination of blue and red light fully restores conidiation,

suggesting the two light qualities are sufficient for the response and interact synergistically. In contrast, cleistothecium formation (sexual development) is repressed by both blue and red light, although red light displays the most dominant role in this response (Purschwitz et al. 2008).

Secondary metabolism, particularly the production of the mycotoxin sterigmatocystin (ST), is tightly linked with sexual development. This may serve to protect sexual structures from predation by competing microorganisms. White light represses ST production, but interestingly, blue and red light play opposing roles; blue light represses ST production, whereas red light induces (Purschwitz et al. 2008).

Microarray analysis has recently demonstrated that light regulates approximately 5 % of the *A. nidulans* genome (Ruger-Herreros et al. 2011), similar to the figure described for *Neurospora*. Beyond development, light was shown to induce genes involved in carbon metabolism and transport, redox reactions, and stress response (Ruger-Herreros et al. 2011).

2. Genetic and Physical Interactions Between the WCC and Phytochrome

With the exception of an archael rhodopsin, the genome of *A. nidulans* contains essentially the same complement of photoreceptors as *Neurospora*. The orthologs to *Neurospora* WC-1/2 in *A. nidulans* are called LreA and LreB (for *L*ight *re*sponse), respectively; both have the predicted domain architecture of their respective *Neurospora* proteins, and they have been shown to interact physically and genetically, suggesting LreA/B also form a codependent heterodimer (WCC) (Purschwitz et al. 2008). For red light sensing, *A. nidulans* encodes for a single phytochrome ortholog, called FphA. The protein shows strong covalent interaction with the tetrapyrrole biliverdin and displays red and far-red absorption peaks typical of bacterial and plant phytochromes. Moreover, the protein displays weak autophosphorylation, thus demonstrating a functional kinase domain (Blumenstein et al. 2005).

Analysis of *lreA*, *lreB*, and *fphA* single-deletion mutants has revealed that the WCC and phytochrome have complimentary roles in regulating development. For example, based on levels of conidiation seen in the mutants with respect to wild type, the WCC and FphA serve as negative and positive regulators of asexual development, respectively (Purschwitz et al. 2008). However, only when *fphA* is deleted along with *lreA* or *lreB* is a complete loss in light induction observed, suggesting that the activation of FphA and derepression of WCC must be coordinated to achieve the proper response. Interestingly, the respective roles for the photoreceptors are reversed with regard to regulating both the sexual cycle and ST production. Notably, cleistothecium formation is severely reduced in the dark in the Δ*lreA* and Δ*lreB* mutants; therefore, the WCC seems to serve as a transcriptional activator and repressor (Purschwitz et al. 2008). Further work will be required to determine if the WCC of *A. nidulans* acts through intermediary transcription factors, as is seen in *Neurospora*.

Strikingly, in addition to the genetic interactions between the WCC and FphA, the two photoreceptor systems also interact physically. Through a combination of bimolecular fluorescence complementation and coimmunoprecipitation assays, it was shown that FphA binds the WCC in the nucleus through a direct interaction with LreB. Within this complex, FphA also interacts directly with a second protein called **VeA (Velvet A)** (Purschwitz et al. 2008, 2009). VeA is a central regulator of development in *A. nidulans* whose loss of function abolishes the requirement of light for conidiation (Mooney and Yager 1990). However, VeA seems not to sense light directly but instead is light regulated through its interaction with FphA. In addition, it is a central regulator of secondary metabolism and is required for ST production (Bayram et al. 2008b). In this way, the complex consisting of VeA/FphA/LreA/LreB likely serves to balance both developmental programming and secondary metabolism through the integration of multiple light inputs.

3. CryA of *Aspergillus nidulans* Functions as Both a CPD Photolyase and a Light-Sensing Cryptochrome

CryA is the only photolyase/cryptochrome family member encoded by *A. nidulans*. This protein is most closely related to the class I CPD photolyases based on sequence comparisons, and it does indeed demonstrate photoactivated DNA repair in vivo and when expressed in *Escherichia coli*. Interestingly, CryA also seems to function as a bona fide cryptochrome in *A. nidulans*, as it represses sexual development in response to blue light and UV-A irradiation. This repressive action of CryA is, in part, achieved by negatively regulating the expression of *veA* and other positive regulators of the sexual cycle (Bayram et al. 2008a). From a physiological standpoint, the combined action of the WCC (blue) and CryA (UV-A) likely imparts a broad action spectrum to which *A. nidulans* can regulate development. Whether the CryA protein interacts with the FphA/WCC/VeA complex has yet to be established.

4. Summary

In summary, the photoresponse of *A. nidulans* displays important parallels and differences with the WCC-dominated system established in *Neurospora*. While *A. nidulans* does indeed respond to blue light through its WCC (LreA and LreB), the fungus is equally responsive to red light via a phytochrome and UV-A light via a cryptochrome. Because these different light qualities may differentially regulate certain processes (e.g., mycotoxin production) and because different light signals are likely detected concurrently, the photosensory pathways must somehow communicate to integrate the light signals in a meaningful way. If recent breakthroughs are any indication, future studies will undoubtedly reveal a highly complex system that is regulated at the transcriptional, posttranscriptional, posttranslational, and epigenetic levels.

E. General Conclusions

Further investigations into how fungi perceive light will undoubtedly have important implications in a variety of different fields. For example, the use of photoreceptors (Christie et al. 2012) and light-inducible promoters (Hurley et al. 2012) as tools to study cell biology in heterologous systems has been an emerging and exciting area of research. Intriguingly, in two divergent fungal pathogens, *C. neoformans* (Idnurm and Heitman 2005a) and *Fusarium oxysporum* (Ruiz-Roldan et al. 2008), deletion of the *wc-1* ortholog leads to an attenuation of virulence in their respective animal infection models. It is, therefore, exciting to speculate that the WCC, and potentially other photosensory pathways, could be conservatively linked to fungal pathogenesis and represent novel targets for antifungal therapy. Recent work has shown that the predominant mold pathogen of immunocompromised patients, *A. fumigatus*, also regulates many virulence-associated pathways in response to light (K. Fuller unpublished data), and research into how this affects its pathogenic potential is under way. Only time will tell to what extent understanding fungal visual systems will help researchers see the world in new and exciting ways.

III. Circadian Rhythms and Fungi

Originally made famous by Beadle and Tatum's one enzyme/one gene work, *Neurospora* was a premier organism in genetics until the use of *E. coli* became common (Davis and Perkins 2002). Circadian rhythms were first documented in *Neurospora* in 1959 by Pittendrigh et al. (1959). These rhythms fulfilled the criteria to be defined as circadian: (a) They maintained a sustained period (approximately 22.5 h under constant conditions); (b) the phase was set by a single transition from light to dark; (c) the rhythms were entrainable; and (d) the period

was not dependent on temperature. Since then, *Neurospora* has become a highly tractable genetic organism, with a full genome sequence, transformation protocols with an extremely high success rate, and near-complete knockout library (Collopy et al. 2010; Colot et al. 2006; Galagan et al. 2003).

A. What Is a Circadian Rhythm?

The origin of circadian studies started with Jean Jacques Ortous de Mairan, a French scientist who described daily leaf movements, noting that the movements persisted when the plant he was studying was in the dark as well as in the light. More important, these movements persisted with a regular period, leading him to the conclusion that they were not a simple response to the sun but in fact a more complex system regulated by an internal mechanism. The study of circadian rhythms in *Neurospora* started with Colin Pittendrigh's work in the 1950s looking at the rhythmic expression of asexual conidia.

To distinguish a reactive response from a clear circadian rhythm, there are four basic tenants. First, the rhythm must persist in the absence of exogenous cues, such as light/dark cycles. This is a clear sign that there is an **internal timekeeping mechanism** that is creating this cycle, and that the cycle is not just a response to external cues.

Second, the duration of a true circadian rhythm must have a cycle length, also called a **period**, of approximately 24 h. This ensures that the organism is at the same state at the same time every day and builds flexibility into the system so that it may be reset by external cues. This allows for better control of the timing of circadian responses.

A third characteristic of a true circadian rhythm is the ability of the rhythm to be reset (or entrained) to external time cues. This **entrainment** allows the organism to adapt to its local environment, as in resetting the clock when crossing time zones. Without this adaptability, a person could not thrive in an environment that did not have conditions that matched their endogenous rhythm.

Finally, as is well established in chemistry, different temperatures affect the kinetics of a reaction. If a clock were to speed up in higher temperatures and slow down in lower ones, it would simply be a thermometer, not a clock. So, a true circadian rhythm must maintain its periodicity over a range of physiological temperatures, a capability referred to as **temperature compensation.**

Beyond these basic rules of circadian rhythms, it is generally accepted that circadian rhythms are nearly ubiquitous in nature, being found in many organisms and biological processes. They are also generated at the cellular level, whether the organism is a unicellular prokaryote or a multicellular, complex eukaryote.

B. Tracking Circadian Rhythms in *Neurospora crassa*

Neurospora propagates both asexually and sexually. Although both are circadianly regulated, it is the asexual cycle that allowed for the first clear insight into the circadian cycle (Bobrowicz et al. 2002; Dunlap and Loros 2004). In the case of asexual spores, a developmental switch leads to the production of conidia in the subjective night. A common method to assay this rhythmic development of conidia is using a race tube, a long glass tube, bent at both ends, holding agar medium onto which mycelia or conidia are inoculated at one end (Fig. 6.2a). The fungus grows toward the other end, and the growth front is marked each day. Cellular clocks are synchronized by germinating the cultures in constant light for a day and then transferring them to constant darkness.

Expression of the overt rhythm in sporulation is not robust in wild-type cultures unless there is gentle airflow over the culture. However, identification of the ***band*** **(*bd*)** strain in the late 1960s alleviated this (Sargent et al. 1966). This strain exhibited rhythmic and robust formation of asexual conidia in a manner that fit the conditions of a true circadian rhythm without the need to flow air through the system (Belden et al. 2007a). The mutation in the strain was identified as a mutation in *ras-1*,

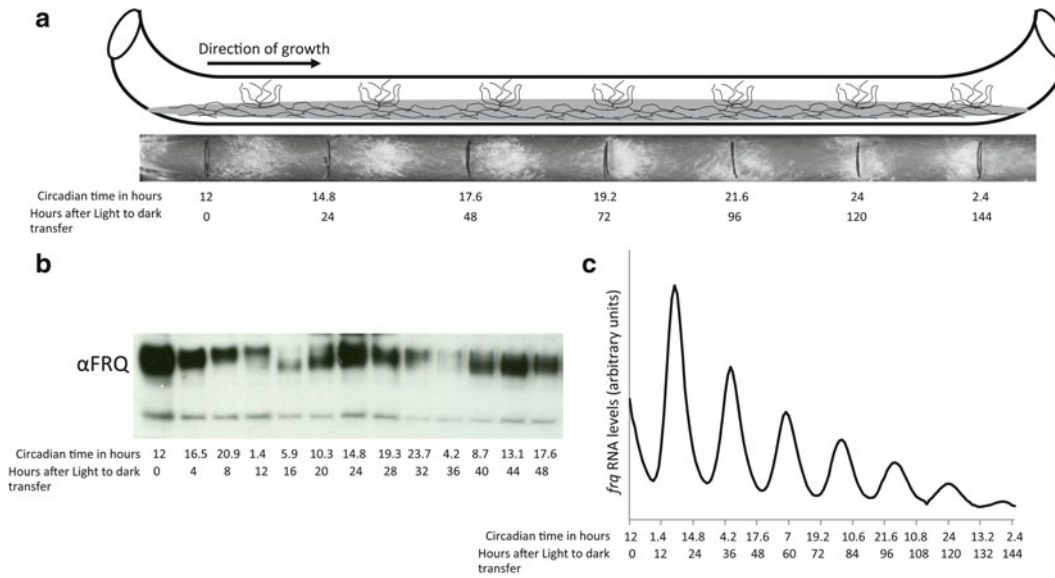


Fig. 6.2. Methods of circadian analysis in *Neurospora*. (**a**) *Neurospora* growing along a race tube with an image of an actual race tube beneath it. Marks represent the growth front (demarcated in sidereal time and in circadian time for the figure) (Unpublished data from J. Hurley) (**b**) Western blot of FRQ (FREQUENCY) protein tracked over 48 sidereal hours (time points taken every 4 sidereal hours) and labeled in circadian hours, highlighting the changes of phosphorylation state of FRQ protein over time (Unpublished data from J. Emerson) (**c**) Luciferase trace of *frq* mRNA expression tracked over 144 sidereal hours and labeled in circadian hours, highlighting the changes of expression levels of *frq* mRNA over time (Unpublished data from J. Emerson)

which increases the levels of reactive oxygen species (ROS) in *Neurospora*. This in turn leads to an increased expression of a subgroup of genes known as the clock-controlled genes (ccgs), in particular the conidial regulation protein, *fluffy*, which accounts for the *bd* phenotype (Belden et al. 2007a).

The *bd* strain has been used in almost all circadian experiments in *Neurospora* since its isolation. However, *bd* turns a subtle circadian regulation into a robust phenotype by increasing transcripts whose gene products control asexual development, suggesting a link between RAS signaling and ROS levels in promotion of sporulation. This means that a subset of the genes and pathways identified as ccgs may be RAS responsive instead of circadianly regulated and be absent in analysis of ccgs from a wild-type strain (Belden et al. 2007a).

In addition to this, rhythms have been tracked when cultures are grown in liquid medium as well, leading to the understanding of the clock on a molecular level as opposed to the overt rhythms followed in the *bd* strain (Aronson et al. 1994; Loros et al. 1989). More recently, in an effort to correct for this potential problem, rhythms have been monitored on a more direct level. An alternative method of monitoring the output of the circadian clock is by following the expression of the luciferase gene from the firefly beetle *Photinus pyralis*, codon optimized for *Neurospora* and driven by the *frq* or *ccg-2* promoter (Gooch et al. 2008; Morgan et al. 2003). Originally designed as a method to follow the endogenous, core rhythms of the *Neurospora* clock on the transcriptional level, this method has been used to understand temperature compensation as well as translational activity in the core *Neurospora* clock (Fig. 6.2b, c) (Gooch et al. 2008; Larrondo et al. 2012).

C. Clock Conservation in Other Fungi

Beyond the well-studied pacemaker in *Neurospora*, there is evidence of conservation of a

circadian circuit in other fungi (Table 6.1). At the phenotypic level, rhythms have been reported in conidiospore formation in the Zygomycete *Pilobolus* (Bruce et al. 1960). Some less-definite growth and developmental rhythms exist in a variety of Ascomycetes (reviewed in Dunlap and Loros 2006), although sclerotia formation of *A. flavus* and enzyme rhythms in *A. nidulans* have been demonstrated to be truly circadianly regulated, with validation of entrainment as well as temperature compensation (Greene et al. 2003).

On the genotypic level, the identification of conserved core clock components has been well investigated. According to a review of 42 sequenced fungal genomes, while light-sensing mechanisms (which typically work as circadian positive-feedback elements; see Molecular Mechanism of a Circadian Oscillator) are widely conserved in the fungi, the clock-exclusive protein FREQUENCY (FRQ) (the negative-feedback element) is less conserved. However, complete circadian feedback loops are seen universally in the Sordariacea, suggesting that many plant and animal pathogens have a functional clock. As described previously, WC-1 is conserved into Basidiomycetes as well as Zygomycetes. Given that rhythms have been reported in *Aspergillus*, which has no FRQ, it may be that rhythms in these strains have a different negative arm protein but still use the positive arm WCC (Dunlap and Loros 2006).

In a more recent study using 64 fungal genomes, it was determined that FRQ was only seen in Sordariomycetes, Leotiomycetes, and Dothideomycetes. Homologs of other clock components such as the WCC, FRH (frequency-interacting helicase), and FWD-1 (F-box and WD40 repeat-containing protein 1) (see FRQ Degradation) have been found in a wider array of organisms, including Zygomycetes, Basidiomycetes, and Ascomycetes. The Saccharomycetes seem to have lost both the WCC and FRQ in a genome size reduction and have never been demonstrated to possess circadian rhythms (Dunlap and Loros 2006; Salichos and Rokas 2010).

D. Molecular Mechanism of a Circadian Oscillator

The mechanism that drives circadian rhythms is a transcriptional/translational feedback loop that is strictly regulated by a series of proteins that interact with a core clock complex. The core complex is made up of two protein complexes, the positive arm (known in *Neurospora* as the WCC), which drives the expression of a second component, the negative arm (known in *Neurospora* as the FRQ/FRH complex or FFC). The negative arm in turn acts to autoregulate its expression on a time delay, which sets that length of the period (Dunlap 1999).

The cycle begins late in the subjective night (Fig. 6.3a–c), as the WCC binds to the promoter of FRQ and induces expression of *frq* mRNA, which reaches its maximum around early subjective morning. About 4 h after the start of *frq* expression, FRQ protein begins to appear; it binds directly to FRH, enters the nucleus, and begins to form homodimers (Dunlap and Loros 2004; Merrow et al. 1997). New FRQ is rapidly phosphorylated in the PEST-1 (a protein domain rich in proline, glutamic acid, serine and threonine) and FFD domains (sites of FRQ/FRQ and FRQ/FRH interactions), with phosphorylation events occurring via interaction with several kinases in the C-terminal region shortly thereafter. These C-terminal phosphorylations stabilize the protein (reviewed in Baker et al. 2012; Heintzen and Liu 2007).

On nuclear entry, FRQ represses *frq* transcription by inhibiting the activity of the WCC while simultaneously increasing the levels of WC-1 (Dunlap and Loros 2004). This inhibition is caused by a direct interaction between FRQ and the WCC, which leads to the phosphorylation of the WCC, inactivating the WCC as well as clearing it from the *frq* promoter (reviewed in Brunner and Kaldi 2008; Liu and Bell-Pedersen 2006). The WCC begins to exit the nucleus at this point in the cycle, perhaps reflecting its phosphorylation status, further decreasing activation of the *frq* promoter

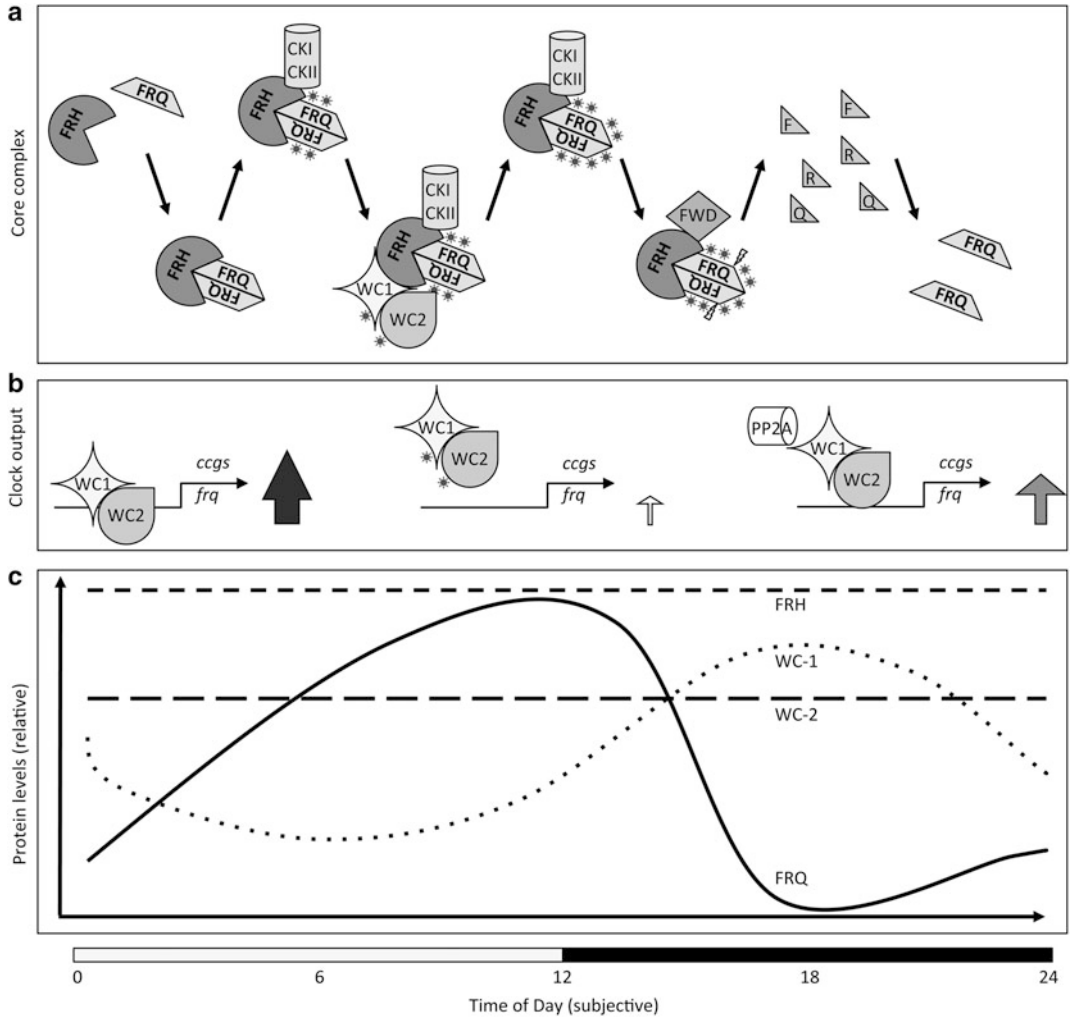


Fig. 6.3. Molecular underpinning of *Neurospora* circadian rhythm. (**a**) In the late subjective night, the WCC induces expression of *frq* mRNA, leading to high levels of FRQ (FREQUENCY) translation. FRQ forms homodimers and binds to FRH. FRQ is phosphorylated via the interaction with several kinases. FRQ is autoregulatory, inhibiting the activity of the WCC by promoting the phosphorylation of the WCC. Lack of WCC leads to a decrease in FRQ synthesis, while old FRQ is increasingly phosphorylated, which leads to ubiquitination facilitated by FWD-1, leading to FRQ degradation. (**b**) The output of the circadian clock. When FRQ levels are low, WCC activity is high, and it subsequently binds to the *frq* promoter as well as other *ccg* promoter and increases expression at those loci. When FRQ is high, it binds to the WCC, promoting its phosphorylation and causing it to become inactive. As FRQ ages and is degraded, phosphatases bind the WCC, dephosphorylating it and allowing it to become active again, leading the WCC to bind to target promoters again. (**c**) Protein levels of the core clock components. While FRH and WC-2 remain constant, FRQ and WC-1 oscillate in opposite phases to one another. Stars represent phosphorylation and lightning bolts represent ubiquitination

(Hong et al. 2008). In the late afternoon, due to the lack of WCC activation, *frq* expression begins to decline, and in accordance, so does FRQ synthesis (Merrow et al. 1997). FRQ is then increasingly phosphorylated at the PEST domain and the N-terminal domain, which leads to FRQ being ubiquitinated by an SCF-ubiquitin ligase complex containing the F-box protein FWD-1 and targeted to the proteasome for degradation (reviewed in Baker et al. 2012;

Heintzen and Liu 2007). The mass of WC-1 that was created and held in the inactive state by FRQ is now released, and the cycle restarts, with the WCC again binding to the *frq* promoter (Dunlap and Loros 2004). The delays between *frq* expression and FRQ synthesis (3–6 h) and FRQ phosphorylation-directed degradation (14–18 h) leads to the nearly 24-h rhythm in *Neurospora* and is how the circuit keeps its specific time (Merrow et al. 1997).

E. Core Clock components

1. The Negative Arm: FRQ and FRH

The use of *Neurospora* as a model for circadian rhythms at the molecular level began with the discovery of several mutants, each of which directly affect the period of banding in *Neurospora*. Eventually, these mutants were all mapped to the *frequency* (*frq*) locus and displayed long, short, or arrhythmic periods; some alleles also altered or disrupted temperature compensation (Gardner and Feldman 1980; Loros and Feldman 1986). *frq* was cloned, and this began the molecular dissection of the core clock of *Neurospora* and a greater understanding of clocks in general (McClung et al. 1989). In *Neurospora*, FRQ expression periodicity matches the conidiation rhythm seen in the *bd* mutant. Altering or inhibiting the FRQ rhythm had a direct and equivalent effect on the clock, demonstrating that FRQ is the driver of *Neurospora* period (Aronson et al. 1994; Belden et al. 2007a; Garceau et al. 1997).

FRQ constitutes the one of two proteins that make up the negative arm of the clock. Full-length FRQ contains 989 amino acids and self-associates via the coil-coil region near the N′ terminus (Aronson et al. 1994; Cheng et al. 2001). *frq* message as well as FRQ protein are rhythmically expressed in a 22.5-h cycle under constant conditions with a phase difference of approximately 4 h (Aronson et al. 1994; Garceau et al. 1997), and it is this rhythmic expression that sets the **period of the clock**. FRQ is a highly transcriptionally, posttranscriptionally, translationally, and posttranslationally controlled protein (see FRQ Transcriptional and Posttranscriptional Regulation) (Baker et al. 2012). FRQ is both nuclear and cytoplasmic; however, the most understood activity of FRQ is nuclear, where it binds to the WCC and blocks the transcriptional activity of the WCC (Liu et al. 2003). FRQ is able to increase WC-1 levels; this probably is the result of inhibiting the activity of WC-1, a protein believed to be unstable when it is active (Shi et al. 2010; reviewed in Baker et al. 2012). FRQ also increases the abundance of *wc-2* through an unknown mechanism (Liu et al. 2003).

a) FRQ Transcriptional and Posttranscriptional Regulation

First, the rhythmic binding of the pertinent transacting factors must occur to maintain a functional circadian rhythm. Transcriptional regulation at the *frq* promoter occurs through binding of the WCC proteins to a pair of *cis*-acting sequences termed the **Clock box** (C box) and the **proximal light-regulated element** (PLRE) (Froehlich et al. 2003). The C box is required for rhythmic expression of *frq* and overall clock function in continual darkness, whereas the PLRE is necessary to establish the proper phase when entrained by light. Both elements are necessary for high levels of light-induced *frq* expression via WCC influence on the *frq* promoter, but each element acts differentially as chromatin is remodeled during the transcriptional activation and deactivation of *frq* (Belden et al. 2007b).

The gene *clockswitch* (*csw-1*) is required for normal *frq* expression and acts directly or indirectly to negatively regulate WCC activity at *frq* by altering chromatin structure, creating a more compact chromatin structure at the C box (Belden et al. 2007b). Chromodomain helicase DNA–binding (CHD-1) contributes to changes in chromatin structure at *frq* and is also needed for normal *frq* expression. DNA methylation at *frq*, which is promoted by loss of CHD-1, is transient and reversible and catalyzed by the DNA methyltransferase DIM-2 (Defective in Methylation-2), which limits the onset of circadian-regulated transcription via regulation of methylation at the *frq* promoter (Belden et al. 2007b, 2011).

The *frq* mRNA regulation extends beyond simple transcriptional regulation. Indeed, there is alternative initiation of translation that results in the production of two distinct FRQ polypeptides, FRQ^{1-989} (L-FRQ) and $FRQ^{100-989}$ (S-FRQ) (Garceau et al. 1997). While both forms of FRQ are able to maintain rhythmicity at 25 °C, the amplitude and robustness of the rhythm are affected across a range of physiological temperatures. Ratios of the two proteins are known to vary with temperature, with more L-FRQ being produced at higher temperatures versus even amounts of long and short FRQ protein at lower temperatures (Liu et al. 1997). Phosphorylation sites on the 100 amino acids on L-FRQ that are not present in S-FRQ have been shown to decrease period length even in the presence of S-FRQ. The **ratio of FRQ polypeptides** is believed to allow the clock to fine-tune the period in response to environmental cues while remaining a robust timekeeping mechanism in their absence (Baker et al. 2009; Diernfellner et al. 2007; Liu et al. 1997).

The selective transcription of either S- or L-FRQ is determined by the temperature-dependent, alternative splicing event of a small intron encompassing the AUG of L-FRQ. The retention of this intron results in the exclusive use of the AUG from L-FRQ, whereas its removal makes the AUG from S-FRQ the first bona fide start codon. Strains unable to splice this intron fail to produce S-FRQ (Colot et al. 2005; Diernfellner et al. 2005).

In addition to this splice site, there are alternative splicing events farther upstream in the 5′ UTR that remove five upstream AUGs with four uORFs from all major *frq* transcripts. Two AUGs remain, and these uORFs may be differentially regulating S- and L-FRQ at the translational level by targeting transcripts for nonsense-mediated decay (NMD), either as a way to remove improperly spliced transcripts or as a mechanism for quantitative control of gene expression (Colot et al. 2005; Diernfellner et al. 2005).

Another regulation of the fine-tuning of the *frq* mechanism is an **antisense transcript**, which comprises the entire length of the FRQ open reading frame (ORF). Elimination of this transcript causes a slight increase in period as well as a loss of rhythmicity at low physiological temperatures and an earlier phase setting on light-to-dark transfer. It is postulated that this antisense RNA may be an additional mechanism to regulate *frq* posttranscriptionally or posttranslationally to further insolate the clock from environmental stresses (Kramer et al. 2003).

b) FRQ Associations and Subcellular Localization

FRQ is only half of the negative arm of the clock. Frequency-interacting RNA helicase (FRH) is a homolog of Mtr4p, a well-studied cofactor of the *Saccharomyces cerevisiae* exosome (Cheng et al. 2005). Mtr4p is a known member of the TRAMP complex and has been suggested to play many roles in a variety of cellular processes (LaCava et al. 2005). FRH is so important to the clock complex that all FRQ is bound to FRH. When FRH is removed from the system via small interfering RNA (siRNA) knockdown (*frh* is an essential gene in *Neurospora*), the clock loses rhythmicity completely, and the FRQ level decreases dramatically while the inverse happens to the mRNA (Cheng et al. 2005).

It has been suggested that the role of FRH in the clock is to regulate the levels of *frq* posttranscriptionally, as when FRH is knocked down, *frq* mRNA is stabilized (Guo et al. 2009). When taken in combination with the demonstrated interaction between FRH and components of the exosome complex in *Neurospora*, the general theory is that FRH directs *frq* mRNA to the TRAMP complex for degradation (Guo et al. 2009).

Beyond this, FRH is known to play a role in the complex interaction between FRQ and the WCC; FRH is integral to the interaction between the FFC and the WCC and is also able to interact with the WCC in the absence of FRQ (Cheng et al. 2005; Guo et al. 2010; Shi et al. 2010). FRH is also implicated in the proper methylation of *frq* (Belden et al. 2011), as well as an essential interactor of VVD in suppression of FRQ expression via interaction with the WCC (Hunt et al. 2010).

FRQ association with FRH is essential for proper phosphorylation (discussed in FRQ Phosphorylations and the Role of Kinases and

Phosphatases in the Clock) as well as FRQ stability, and FRH plays a role in the proper localization of FRQ protein (Cha et al. 2011; Guo et al. 2010). FRQ nuclear localization is essential for its function in the circadian clock, and the nuclear localization signal located downstream of the coiled-coil domain is sufficient to direct the localization of FRQ protein (Luo et al. 1998). However, a large portion of FRQ is cytoplasmic, and the nuclear-cytoplasmic shuttling has been suggested to be dependent on the phosphorylation state of FRQ (Diernfellner et al. 2009). This model stands in contrast to the idea that FRH is the key link to proper subcellular localization of FRQ (Cha et al. 2011).

Independently of the pull-down assay used to identify FRH originally, FRH was identified through a mutagenesis screen for negative-feedback loop mutants. This screen identified a point mutation that was outside the highly conserved helicase region of FRH. This mutation eliminated the interaction of FRH with the WCC but not with FRQ (Shi et al. 2010). This mutation hints that the role of FRH that is specific to the clock may be different from its role in the TRAMP/exosome complex function that FRH plays for overall cell fitness.

c) FRQ Phosphorylations and the Role of Kinases and Phosphatases in the Clock

FRQ undergoes dual molecular rhythms, in both abundance (described previously) and phosphorylation, both of which influence its turnover kinetics (Garceau et al. 1997; Liu et al. 2000). FRQ is phosphorylated immediately after synthesis and further phosphorylated in a highly regulated manner throughout the circadian day (Baker et al. 2009). FRQ is also phosphorylated in constant light, although in a less-specific and regulated manner (Baker et al. 2009; Tang et al. 2009). Mutations that reduce the level of phosphorylated FRQ tend to increase its stability, which in turn leads to increased period lengths (Liu et al. 2000; Ruoff et al. 2005).

No specific phosphorylation event has a role that is essential to protein turnover. Instead, most phosphorylations occur in clusters, added at different locations and times. FRQ is unmodified after synthesis, and phosphorylation events occur in a strict sequence throughout the day. At first, FRQ is rapidly phosphorylated in the middle part of the protein between the PEST-1 and the FFD domain. There has been no function assigned to these events, and mutations at these sites did not alter circadian rhythms. Phosphorylation at the C-terminus is next in the sequence, and mutations at specific sites in the C-terminal region result in FRQ protein that is less stable and in a short-period rhythm. Midday, the PEST-1 domain shows a dramatic increase in phosphorylation. Mutations of sites in this region showed an increase in period and more stable FRQ, suggesting that phosphorylation of these residues is needed to promote turnover of FRQ. Further phosphorylation of FRQ, predominantly in residues specific to the long FRQ isoform, occurs late in the cycle, and mutations in this region result in a longer period, suggesting a role in promoting turnover (Baker et al. 2009).

Beyond stability, **FRQ structure** is predicted to be affected directly by the phosphorylation of FRQ protein (Querfurth et al. 2011). Hyperphosphorylation of FRQ at many sites causes a conformational change caused by charge-charge repulsion. In the hypophosphorylated state, FRQ is in a closed conformation, which opens on increasing phosphorylation, to reveal a degradation signal in the middle portion of FRQ. The key regulation site to this opening is in the N-terminal domain of FRQ; new FRQ adopts preferentially the closed conformation with the positively charged N-terminal domain interacting with the negatively charged remainder portion of the protein. The N-terminal domain of FRQ is progressively phosphorylated, which lowers the pI of the domain, increasing negative surface charge of the N-terminal domain and weakening the interaction with the negatively charged middle and C-terminal domains (Querfurth et al. 2011). While in some ways appealing, this model suffers from the fact that N-terminal phosphorylations were previously shown to be among the last modifications during the circadian cycle rather than among the first (Baker et al. 2009).

Phosphorylation of a protein to this extent, around 100 distinct modifications, needs a complex network of **kinases and phosphatases** for proper modification. Kinases known to play

a role in the clock include casein kinases 1 and 2 (CK1a and CK2), a *Neurospora* homolog of checkpoint kinase-2 (PRD-4), as well as CAMK-1 and basophilic protein kinase A (Klengel et al. 2005); FRQ has been shown to have direct physical interaction with CK1a, CK2, and PRD-4 (Baker et al. 2012; Diernfellner and Schafmeier 2011). CK1a interacts with FRQ via two FRQ/CK1a interacting domains (FCDs), and this interaction not only catalyzes the phosphorylation of FRQ (as many as 41 times), leading to FRQ degradation, but also may catalyze the phosphorylation of the WCC (He et al. 2006; Querfurth et al. 2011), leading to the hypothesis that FRQ acts as a scaffold for major components of the clock. The interaction between CK2 and FRQ is significantly weaker than that of CK1a but plays a unique role in the clock in that the phosphorylations that are attributed to CK2 are involved in maintaining the temperature compensation function of the clock (Mehra et al. 2009).

Several phosphatases are known to be involved in the clock as well, including protein phosphatase-1 (PP1), PP2a, and PP4. Phosphatases play a role in everything from regulating FRQ stability to influencing *frq* transcription, to phosphorylation of the WCC and affecting WCC subcellular localization (Baker et al. 2012).

d) FRQ Degradation

As described previously, proper phosphorylation of FRQ has been shown to regulate the degradation of FRQ as well as affect the function of the clock. But, there are other posttranslational modifications that affect the degradation of FRQ, including **ubiquitination** (He and Liu 2005). One gene that plays a role in ubiquitination is the F-box/WD40 repeat-containing protein FWD-1. FWD-1 has been shown to directly interact with FRQ, particularly the phosphorylated form, and is required for the proper degradation of FRQ. It is predicted that phosphorylated FRQ is a substrate for an FWD-1-containing SCF-type ubiquitin ligase complex, and moreover, that the SCF complex can recognize different phosphorylated motifs within FRQ. Increase in FRQ phosphorylation increases the number of potential FWD-1-binding sites and its overall affinity toward FWD-1 (He et al. 2003). Progressive phosphorylation of FRQ may be a dynamic process that fine-tunes the stability of FRQ through its role in the ubiquitination of FRQ, and this determines the period of the clock. Beyond the FWD-1 role in ubiquitination, it is believed that there may be other FRQ mechanisms of degradation (He et al. 2003).

2. The Positive Arm: The White Collar Complex

Although the WCC was extensively described previously, it is also known to play a major role in addition to its photoreceptor function in *Neurospora*; the WCC is the positive arm of the clock. Knockouts of either of the White Collar genes eliminate clock function (Crosthwaite et al. 1997). The WCC complex binds directly to the *frq* promoter and induces expression of *frq* mRNA (Froehlich et al. 2002). WCC activity is then inhibited when a direct interaction with the FFC causes the complex constituents to be phosphorylated. Individually, WC-1 is necessary for the interaction between the WCC and FRQ. The amount of WC-1 protein also cycles, although this rhythm is not necessary for the clock, and WC-1 is stabilized by WC-2. WC-2 is also necessary for interaction between the WCC and FRQ, shows neither mRNA nor protein rhythms, and is both positively regulated by FRQ and negatively regulated by WC-1 (Liu et al. 2003). Interestingly, WC-1 can be found at the FRQ promoter throughout the day, even when it is inactivated, whereas WC-2 binds cyclically (Belden et al. 2007b).

a) Rhythmic Phosphorylation of the White Collar Complex

Like FRQ, both WC-1 and WC-2 are phosphorylated in a circadian manner, and this phosphorylation regulates the activity of the complex. Direct phosphorylation of WC-2 is rhythmic through the circadian day, and this phosphorylation regulates the binding of the WCC to DNA. Only one site has been identified on WC-2, but it has been suggested that there are many more that could affect both the period

and the stability of the protein (Sancar et al. 2009). WC-1 is also circadianly phosphorylated; phosphorylation of some sites near the Zn-finger DNA binding domain is believed to regulate the ability of WC-1 to activate transcription (He et al. 2005), but otherwise the role of WC-1 phosphorylation is not well understood. What is known is that phosphorylation in the C-terminus of WC-1 is sequential and FRQ dependent and is required for negative feedback (Baker et al. 2012).

F. Clock Inputs and Outputs

1. Inputs

The *Neurospora* clock is designed to be sensitive to many environmental inputs. **Light induction** by the WCC (described previously) drives an increase in *frq* levels as well as entrainment and phase resetting. This effect on *frq* expression by light impacts the clock differently depending on the time in the cycle that the light is seen. At the time of low *frq* expression in the morning, *frq* levels that increase sharply in accordance with exposure to light will advance the clock to the time corresponding to the highest *frq* expression (mid-to late morning), while increasing *frq* levels during the time of declining *frq* (late subjective afternoon and into evening) will lead to phase delays as *frq* levels are forced to return to their maximum midday levels after light exposure (Crosthwaite et al. 1997).

Light input to the clock can be modified by VVD (described previously), which itself is clock regulated and is able to gate the light response on the clock (Heintzen et al. 2001). VVD inhibition of the WCC directs the clock to take its principal cues from the dusk transition as well as contributing to temperature compensation (Elvin et al. 2005; Hunt et al. 2007). Furthermore, VVD levels in the dark are able to inactivate any WCC induced by moonlight and keep the clock on pace in the brightest of nights (Malzahn et al. 2010).

Temperature has a great effect on the clock, both on entrainment and on period, and can be a stronger resetting cue than light. Higher temperatures induce higher levels of L-FRQ as well as higher levels of FRQ overall (Garceau et al. 1997; Liu et al. 1997). When shifting to higher temperatures, the FRQ levels at the shift are lower than the lowest FRQ levels at the higher temperature, so the clock will reset to subjective morning, the time of day corresponding to low FRQ (Liu et al. 1998). Finally, it has been hypothesized that **metabolism** can play a role in clock input via a feedback loop on the positive arm by CSP-1 (Sancar et al. 2012).

2. Outputs

The internal mechanism of the clock has to have a method to provide rhythmic information to regulate the cell. The primary output of the clock is through a subset of rhythmically expressed genes termed the **clock-controlled genes**, or ccgs. Depending on the method of experimentation, anywhere from 1 % to 15 % of the genome is considered to be circadianly regulated in *Neurospora* (Dong et al. 2008; Dunlap and Loros 2004). These genes are not all expressed at the same time but are actually staggered in their expression around the clock, although late night to morning genes are the most common. Not surprisingly, most ccgs are WCC driven, and although most WCC-driven/light-induced genes are also ccgs, there are some distinct subsets of WCC-driven genes in both the light and dark conditions, showing that distinct from its role in light response (described previously), the WCC can have a different set of target genes in the dark (Dunlap and Loros 2004).

Rhythmic gene regulation must occur via either transcription or mRNA turnover, and transcription seems to be the most common mechanism of regulation. The promoter of the well-studied CCG, *ccg-2*, has several regulatory regions that individually confer light, developmental, and circadian regulation. One such element, termed the activating circadian element (ACE), is sufficient to confer clock regulation on these promoters and has been seen in other clock promoters (Bell-Pedersen et al. 1996). The core ACE sequence is distinct from the core LRE sequence (discussed previously), and

both of these elements mediate clock control as they have been identified in some of the promoters of ccgs identified using microarrays. Some genes have neither element, suggesting additional clock control elements and the presence of hierarchical control, in which the oscillator directly regulates oscillator proximal controllers, which in turn regulate more downstream genes, as has been noted for light-responsive genes in *Neurospora* (Chen et al. 2009; Dunlap and Loros 2004).

In addition to transcriptional regulation of ccgs, it has recently been hypothesized that another level of regulation can occur when the FRQ-FRH complex interacts with mRNA and the exosome via an association with the TRAMP complex. This mechanism would allow for control of message stability rather than production (Guo et al. 2009). Beyond this, it has recently been shown that the mitogen-activated protein kinase (MAPK) pathways are regulated by the clock at the transcriptional level by the WCC (Bennett et al. 2013; Lamb et al. 2011, 2012). In addition, MAK-1 has been shown to be phosphorylated in a circadian manner, and its targets are ccgs, demonstrating the circadian clock signal can be propagated outside WCC regulation (Bennett et al. 2013).

G. Molecular Advancements in *Neurospora*

Although historically genetic manipulation of filamentous fungi was difficult, within the past decade methods have been developed to perfect gene replacements and knockouts in *Neurospora*. Wild-type *Neurospora* strains exhibit a low frequency of homologous recombination after transformation. However, *Neurospora* recipient strains in which ectopic insertions are virtually eliminated have been developed. Mutations of either of two *Neurospora* genes required for nonhomologous end-joining DNA repair (*mus-51* and *mus-52*) result in an increase in homologous recombination (Ninomiya et al. 2004). Homologous recombination rates are around 90 % in strains that are mutant at these loci. In addition to this advancement, recombinational cloning in yeast is used to anneal polymerase chain reaction (PCR) fragments, which are synthesized individually with short overlapping ends by PCR and cotransformed into yeast for assembly (Colot et al. 2006). These generated fragments can then be transformed into *Neurospora* strains that are mutant at either of the *mus* loci, creating a streamlined and highly efficient system for gene replacements and knockouts.

In addition to the targeted gene placement system, the ability to insert single-copy DNA sequences into a known locus is useful. In *Neurospora*, integration into the *his-3* gene (replacing a nonfunctional mutant copy of *his-3* with the wild-type allele as well as the DNA sequence of interest) is the most frequently used method. A drawback, however, is that the strain must be backcrossed to obtain a homokaryon, thereby exposing duplicated sequences to RIP. In addition, the *his-3* mutant is a potentially limiting genetic background. To compensate for these limitations, a system has been developed using *csr-1* as a targetable locus. *csr-1* encodes cyclophilin, a gene that, when mutated, confers resistance to cyclosporin A, an antifungal agent (Bardiya and Shiu 2007; Hurley et al. 2012). The advantages of targeting to this locus are that there is a scorable phenotype (resistance to cyclosporin A) and that all transformants are homokaryotic (untransformed *csr-1*$^{+}$ nuclei are selected against based on their abilities to produce cyclosporin A–binding protein).

References

Aronson BD, Johnson KA, Loros JJ, Dunlap JC (1994) Negative feedback defining a circadian clock: autoregulation of the clock gene *frequency*. Science 263:1578–1584

Arpaia G, Cerri F, Baima S, Macino G (1999) Involvement of protein kinase C in the response of *Neurospora crassa* to blue light. Mol Gen Genet 262: 314–322

Avalos J, Estrada AF (2010) Regulation by light in *Fusarium*. Fungal Genet Biol 47:930–938

Baker CL, Kettenbach AN, Loros JJ, Gerber SA, Dunlap JC (2009) Quantitative proteomics reveals a dynamic interactome and phase-specific phosphorylation in the *Neurospora* circadian clock. Mol Cell 34:354–363

Baker CL, Loros JJ, Dunlap JC (2012) The circadian clock of *Neurospora crassa*. FEMS Microbiol Rev 36:95–110

Ballario P, Vittorioso P, Magrelli A, Talora C, Cabibbo A, Macino G (1996) White collar-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein. EMBO J 15:1650–1657

Bardiya N, Shiu PK (2007) Cyclosporin A-resistance based gene placement system for *Neurospora crassa*. Fungal Genet Biol 44:307–314

Bayram Ö, Biesemann C, Krappmann S, Galland P, Braus GH (2008a) More than a repair enzyme: *Aspergillus nidulans* photolyase-like CryA is a regulator of sexual development. Mol Biol Cell 19:3254–3262

Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon NJ, Keller NP, Yu JH, Braus GH (2008b) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science 320:1504–1506

Bayram Ö, Braus GH, Fischer R, Rodriguez-Romero J (2010) Spotlight on *Aspergillus nidulans* photosensory systems. Fungal Genet Biol 47:900–908

Belden WJ, Larrondo LF, Froehlich AC, Shi M, Chen CH, Loros JJ, Dunlap JC (2007a) The band mutation in *Neurospora crassa* is a dominant allele of ras-1 implicating RAS signaling in circadian output. Genes Dev 21:1494–1505

Belden WJ, Loros JJ, Dunlap JC (2007b) Execution of the circadian negative feedback loop in *Neurospora* requires the ATP-dependent chromatin-remodeling enzyme CLOCKSWITCH. Mol Cell 25:587–600

Belden WJ, Lewis ZA, Selker EU, Loros JJ, Dunlap JC (2011) CHD1 remodels chromatin and influences transient DNA methylation at the clock gene frequency. PLoS Genet 7:e1002166

Bell-Pedersen D, Dunlap JC, Loros JJ (1996) Distinct *cis*-acting elements mediate clock, light, and developmental regulation of the *Neurospora crassa* eas (*ccg-2*) gene. Mol Cell Biol 16:513–521

Bennett LD, Beremand P, Thomas TL, Bell-Pedersen D (2013) Circadian activation of the mitogen-activated protein kinase MAK-1 facilitates rhythms in clock-controlled genes in *Neurospora crassa*. Eukaryot Cell 12:59–69

Bieszke JA, Braun EL, Bean LE, Kang S, Natvig DO, Borkovich KA (1999) The *nop-1* gene of *Neurospora crassa* encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. Proc Natl Acad Sci U S A 96:8034–8039

Bieszke JA, Li L, Borkovich KA (2007) The fungal opsin gene *nop-1* is negatively-regulated by a component of the blue light sensing pathway and influences conidiation-specific gene expression in *Neurospora crassa*. Curr Genet 52:149–157

Bluhm BH, Dunkle LD (2008) PHL1 of *Cercospora zeae-maydis* encodes a member of the photolyase/cryptochrome family involved in UV protection and fungal development. Fungal Genet Biol 45:1364–1372

Bluhm BH, Burnham AM, Dunkle LD (2010) A circadian rhythm regulating hyphal melanization in *Cercospora kikuchii*. Mycologia 102:1221–1228

Blumenstein A, Vienken K, Tasler R, Purschwitz J, Veith D, Frankenberg-Dinkel N, Fischer R (2005) The *Aspergillus nidulans* phytochrome FphA represses sexual development in red light. Curr Biol 15:1833–1838

Bobrowicz P, Pawlak R, Correa A, Bell-Pedersen D, Ebbole DJ (2002) The *Neurospora crassa* pheromone precursor genes are regulated by the mating type locus and the circadian clock. Mol Microbiol 45:795–804

Brenna A, Grimaldi B, Filetici P, Ballario P (2012) Physical association of the WC-1 photoreceptor and the histone acetyltransferase NGF-1 is required for blue light signal transduction in *Neurospora crassa*. Mol Biol Cell 23:3863–3872

Brown L (2004) Fungal rhodopsins and opsin-related proteins: eukaryotic homologues of bacteriorhodopsin with unknown functions. Photochem Photobiol Sci 3:555–565

Bruce VG, Weight F, Pittendrigh CS (1960) Resetting the sporulation rhythm in *Pilobolus* with short light flashes of high intensity. Science 131:728–730

Brunner M, Kaldi K (2008) Interlocked feedback loops of the circadian clock of *Neurospora crassa*. Mol Microbiol 68:255–262

Casas-Flores S, Rios-Momberg M, Bibbins M, Ponce-Noyola P, Herrera-Estrella A (2004) BLR-1 and BLR-2, key regulatory elements of photoconidiation and mycelial growth in *Trichoderma atroviride*. Microbiology 150:3561–3569

Cha J, Yuan H, Liu Y (2011) Regulation of the activity and cellular localization of the circadian clock protein FRQ. J Biol Chem 286:11469–11478

Chen CH, Loros J (2009) *Neurospora* sees the light: light signaling components in a model system. Commun Integr Biol 2:448–451

Chen C, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ (2009) Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in *Neurospora*. EMBO J 28:1029–1042

Chen C, DeMay BS, Gladfelter AS, Dunlap JC, Loros JJ (2010) Physical interaction between VIVID and white collar complex regulates photoadaptation in *Neurospora*. Proc Natl Acad Sci U S A 107:16715–16720

Cheng P, Yang Y, Heintzen C, Liu Y (2001) Coiled-coil domain-mediated FRQ-FRQ interaction is essential for its circadian clock function in *Neurospora*. EMBO J 20:101–108

Cheng P, Yang Y, Gardner KH, Liu Y (2002) PAS domain-mediated WC-1/WC-2 interaction is essential for maintaining the steady-state level of WC-1 and the function of both proteins in circa-

dian clock and light responses of *Neurospora*. Mol Cell Biol 22:517–524

Cheng P, Yang Y, Wang L, He Q, Liu Y (2003) WHITE COLLAR-1, a multifunctional *Neurospora* protein involved in the circadian feedback loops, light sensing, and transcription repression of *wc-2*. J Biol Chem 278:3801–3808

Cheng P, He Q, Wang L, Liu Y (2005) Regulation of the *Neurospora* circadian clock by an RNA helicase. Genes Dev 19:234–241

Christie JM, Gawthorne J, Young G, Fraser NJ, Roe AJ (2012) LOV to BLUF: flavoprotein contributions to the optogenetic toolkit. Mol Plant 5:533–544

Collett MA, Garceau N, Dunlap JC, Loros JJ (2002) Light and clock expression of the *Neurospora* clock gene frequency is differentially driven by but dependent on WHITE COLLAR-2. Genetics 160:149–158

Collopy PD, Colot HV, Park G, Ringelberg C, Crew CM, Borkovich KA, Dunlap JC (2010) High-throughput construction of gene deletion cassettes for generation of *Neurospora crassa* knockout strains. Methods Mol Biol 638:33–40

Colot HV, Loros JJ, Dunlap JC (2005) Temperature-modulated alternative splicing and promoter use in the circadian clock gene *frequency*. Mol Biol Cell 16:5563–5571

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103:10352–10357

Corrochano LM (2007) Fungal photoreceptors: sensory molecules for fungal development and behavior. Photochem Photobiol Sci 6:725–736

Corrochano LM, Cerd-Olmedo E (1992) Sex, light and carotenes: the development of *Phycomyces*. Trends Genet 8:268–274

Corrochano LM, Garre V (2010) Photobiology in the zygomycota: multiple photoreceptor genes for complex responses to light. Fungal Genet Biol 47:893–899

Crosthwaite SK, Dunlap JC, Loros JJ (1997) *Neurospora wc-1* and *wc-2*: transcription, photoresponses, and the origins of circadian rhythmicity. Science 276:763–769

Davis RH, Perkins DD (2002) Timeline: *Neurospora*: a model of model microbes. Nat Rev Genet 3:397–403

Diernfellner AC, Schafmeier T (2011) Phosphorylations: making the *Neurospora crassa* circadian clock tick. FEBS Lett 585:1461–1466

Diernfellner AC, Schafmeier T, Merrow MW, Brunner M (2005) Molecular mechanism of temperature sensing by the circadian clock of *Neurospora crassa*. Genes Dev 19:1968–1973

Diernfellner A, Colot HV, Dintsis O, Loros JJ, Dunlap JC, Brunner M (2007) Long and short isoforms of *Neurospora* clock protein FRQ support temperature-compensated circadian rhythms. FEBS Lett 581:5759–5764

Diernfellner AC, Querfurth C, Salazar C, Hofer T, Brunner M (2009) Phosphorylation modulates rapid nucleocytoplasmic shuttling and cytoplasmic accumulation of *Neurospora* clock protein FRQ on a circadian time scale. Genes Dev 23:2192–2200

Dong W, Tang X, Yu Y, Nilsen R, Kim R, Griffith J, Arnold J, Schuttler HB (2008) Systems biology of the clock in *Neurospora crassa*. PLoS One 3:e3105

Dunlap JC (1999) Molecular bases for circadian clocks. Cell 96:271–290

Dunlap JC, Loros JJ (2004) The *Neurospora* circadian system. J Biol Rhythms 19:414–424

Dunlap JC, Loros JJ (2006) How fungi keep time: circadian system in *Neurospora* and other fungi. Curr Opin Microbiol 9:579–587

Elvin M, Loros JJ, Dunlap JC, Heintzen C (2005) The PAS/LOV protein VIVID supports a rapidly dampened daytime oscillator that facilitates entrainment of the *Neurospora* circadian clock. Genes Dev 19:2593–2605

Estrada AF, Avalos J (2009) Regulation and targeted mutation of *opsA*, coding for the NOP-1 opsin orthologue in *Fusarium fujikuroi*. J Mol Biol 387:59–73

Franchi L, Fulci V, Macino G (2005) Protein kinase C modulates light responses in *Neurospora* by regulating the blue light photoreceptor WC-1. Mol Microbiol 56:334–345

Froehlich AC, Liu Y, Loros JJ, Dunlap JC (2002) White collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. Science 297:815–819

Froehlich AC, Loros JJ, Dunlap JC (2003) Rhythmic binding of a WHITE COLLAR-containing complex to the frequency promoter is inhibited by FREQUENCY. Proc Natl Acad Sci U S A 100:5914–5919

Froehlich AC, Noh B, Vierstra RD, Loros J, Dunlap JC (2005) Genetic and molecular analysis of phytochromes from the filamentous fungus *Neurospora crassa*. Eukaryot Cell 4:2140–2152

Froehlich AC, Chen CH, Belden WJ, Madeti C, Roenneberg T, Merrow M, Loros JJ, Dunlap JC (2010) Genetic and molecular characterization of a cryptochrome from the filamentous fungus *Neurospora crassa*. Eukaryot Cell 9:738–750

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G,

Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868

Garceau NY, Liu Y, Loros JJ, Dunlap JC (1997) Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell 89:469–476

Gardner GF, Feldman JF (1980) The *frq* locus in *Neurospora crassa*: a key element in circadian clock organization. Genetics 96:877–886

Gooch VD, Mehra A, Larrondo LF, Fox J, Touroutoutoudis M, Loros JJ, Dunlap JC (2008) Fully codon-optimized luciferase uncovers novel temperature characteristics of the *Neurospora* clock. Eukaryot Cell 7:28–37

Greene AV, Keller N, Haas H, Bell-Pedersen D (2003) A circadian oscillator in *Aspergillus* spp regulates daily development and gene expression. Eukaryot Cell 2:231–237

Grimaldi B, Coiro P, Filetici P, Berge E, Dobosy JR, Freitag M, Selker EU, Ballario P (2006) The *Neurospora crassa* White Collar-1 dependent blue light response requires acetylation of histone H3 lysine 14 by NGF-1. Mol Biol Cell 17:4576–4583

Guo J, Cheng P, Yuan H, Liu Y (2009) The exosome regulates circadian gene expression in a posttranscriptional negative feedback loop. Cell 138:1236–1246

Guo J, Cheng P, Liu Y (2010) Functional significance of FRH in regulating the phosphorylation and stability of *Neurospora* circadian clock protein FRQ. J Biol Chem 285:11508–11515

He Q, Liu Y (2005) Degradation of the *Neurospora* circadian clock protein FREQUENCY through the ubiquitin-proteasome pathway. Biochem Soc Trans 33:953–956

He Q, Cheng P, Yang Y, Wang L, Gardner KH, Liu Y (2002) White collar-1, a DNA binding transcription factor and a light sensor. Science 297:840–843

He Q, Cheng P, Yang Y, Yu H, Liu Y (2003) FWD1-mediated degradation of FREQUENCY in *Neurospora* establishes a conserved mechanism for circadian clock regulation. EMBO J 22:4421–4430

He Q, Shu H, Cheng P, Chen S, Wang L, Liu Y (2005) Light-independent phosphorylation of WHITE COLLAR-1 regulates its function in the Neurospora circadian negative feedback loop. J Biol Chem 280:17526–17532

He Q, Cha J, Lee HC, Yang Y, Liu Y (2006) CKI and CKII mediate the FREQUENCY-dependent phosphorylation of the WHITE COLLAR complex to close the *Neurospora* circadian negative feedback loop. Genes Dev 20:2552–2565

Heintzen C, Liu Y (2007) The *Neurospora crassa* circadian clock. Adv Genet 58:25–66

Heintzen C, Loros JJ, Dunlap JC (2001) The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting. Cell 104:453–464

Hitomi K, Okamoto K, Daiyasu H, Miyashita H, Iwai S, Toh H, Ishiura M, Todo T (2000) Bacterial cryptochrome and photolyase: characterization of two photolyase-like genes of *Synechocystis* sp PCC6803. Nucleic Acids Res 28:2353–2362

Hong CI, Ruoff P, Loros JJ, Dunlap JC (2008) Closing the circadian negative feedback loop: FRQ-dependent clearance of WC-1 from the nucleus. Genes Dev 22:3196–3204

Hunt SM, Elvin M, Crosthwaite SK, Heintzen C (2007) The PAS/LOV protein VIVID controls temperature compensation of circadian clock phase and development in *Neurospora crassa*. Genes Dev 21:1964–1974

Hunt SM, Thompson S, Elvin M, Heintzen C (2010) VIVID interacts with the WHITE COLLAR complex and FREQUENCY-interacting RNA helicase to alter light and clock responses in *Neurospora*. Proc Natl Acad Sci U S A 107:16709–16714

Hurley JM, Chen CH, Loros JJ, Dunlap JC (2012) Light-inducible system for tunable protein expression in *Neurospora crassa*. G3 2:1207–1212

Idnurm A, Heitman J (2005a) Light controls growth and development via a conserved pathway in the fungal kingdom. PLoS Biol 3:e95

Idnurm A, Heitman J (2005b) Photosensing fungi: phytochrome in the spotlight. Curr Biol 15:R829–R832

Idnurm A, Verma S, Corrochano LM (2010) A glimpse into the basis of vision in the kingdom Mycota. Fungal Genet Biol 47:881–892

Karniol B, Wagner JR, Walker JM, Vierstra RD (2005) Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors. Biochem J 392:103–116

Klein RM, Klein DT (1962) Interaction of ionizing and visible radiation in mutation induction in *Neurospora crassa*. Am J Bot 49:870–874

Klengel T, Liang WJ, Chaloupka J, Ruoff C, Schröppel K, Naglik JR, Eckert SE, Mogensen EG, Haynes K, Tuite MF, Levin LR, Buck J, Mühlschlegel FA (2005) Fungal adenylyl cyclase integrates CO_2 sensing with cAMP signaling and virulence. Curr Biol 15:2021–2026

Kramer C, Loros JJ, Dunlap JC, Crosthwaite SK (2003) Role for antisense RNA in regulating circadian clock function in *Neurospora crassa*. Nature 421:948–952

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 121:713–724

Lamb TM, Goldsmith CS, Bennett L, Finch KE, Bell-Pedersen D (2011) Direct transcriptional control

of a p38 MAPK pathway by the circadian clock in *Neurospora crassa*. PLoS One 6:e27149

Lamb TM, Finch KE, Bell-Pedersen D (2012) The *Neurospora crassa* OS MAPK pathway-activated transcription factor ASL-1 contributes to circadian rhythms in pathway responsive clock-controlled genes. Fungal Genet Biol 49:180–188

Larrondo LF, Loros JJ, Dunlap JC (2012) High-resolution spatiotemporal analysis of gene expression in real time: in vivo analysis of circadian rhythms in *Neurospora crassa* using a FREQUENCY-luciferase translational reporter. Fungal Genet Biol 49:681–683

Lee K, Dunlap JC, Loros JJ (2003) Roles for WHITE COLLAR-1 in circadian and general photoperception in *Neurospora crassa*. Genetics 163:103–114

Lee K, Singh P, Chung W, Ash J, Kim TS, Hang L, Park S (2006) Light regulation of asexual development in the rice blast fungus, *Magnaporthe oryzae*. Fungal Genet Biol 43:694–706

Lewis ZA, Correa A, Schwerdtfeger C, Link KL, Xie X, Gomer RH, Thomas T, Ebbole DJ, Bell-Pedersen D (2002) Overexpression of White Collar-1 (WC-1) activates circadian clock-associated genes, but is not sufficient to induce most light-regulated gene expression in *Neurospora crassa*. Mol Microbiol 45:917–931

Linden H, Macino G (1997) White Collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. EMBO J 16:98–109

Linden H, Ballario P, Macino G (1997) Blue light regulation in *Neurospora crassa*. Fungal Genet Biol 22:141–150

Liu Y, Bell-Pedersen D (2006) Circadian rhythms in *Neurospora crassa* and other filamentous fungi. Eukaryot Cell 5:1184–1193

Liu Y, Garceau NY, Loros JJ, Dunlap JC (1997) Thermally regulated translational control of FRQ mediates aspects of temperature responses in the *Neurospora* circadian clock. Cell 89:477–486

Liu Y, Merrow M, Loros JJ, Dunlap JC (1998) How temperature changes reset a circadian oscillator. Science 281:825–829

Liu Y, Loros J, Dunlap JC (2000) Phosphorylation of the *Neurospora* clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. Proc Natl Acad Sci U S A 97:234–239

Liu Y, He Q, Cheng P (2003) Photoreception in *Neurospora*: a tale of two White Collar proteins. Cell Mol Life Sci 60:2131–2138

Lombardi LM, Brody S (2005) Circadian rhythms in *Neurospora crassa*: clock gene homologues in fungi. Fungal Genet Biol 42:887–892

Loros JJ, Feldman JF (1986) Loss of temperature compensation of circadian period length in the *frq-9* mutant of *Neurospora crassa*. J Biol Rhythms 1:187–198

Loros JJ, Denome SA, Dunlap JC (1989) Molecular cloning of genes under control of the circadian clock in *Neurospora*. Science 243:385–388

Losi A (2007) Flavin-based blue-light photosensors: a photobiophysics update. Photochem Photobiol 83:1283–1300

Losi A, Gärtner W (2011) Old chromophores, new photoactivation paradigms, trendy applications: flavins in blue light-sensing photoreceptors. Photochem Photobiol 87:491–510

Lukens RJ (1965) Reversal by red light of blue light inhibition of sporulation in *Alternaria solani*. Phytopathology 55:1032

Luo C, Loros JJ, Dunlap JC (1998) Nuclear localization is required for function of the essential clock protein FRQ. EMBO J 17:1228–1235

Malzahn E, Ciprianidis S, Kldi K, Schafmeier T, Brunner M (2010) Photoadaptation in *Neurospora* by competitive interaction of activating and inhibitory LOV domains. Cell 142:762–772

McClung CR, Fox BA, Dunlap JC (1989) The *Neurospora* clock gene frequency shares a sequence element with the Drosophila clock gene period. Nature 339:558–562

Mehra A, Shi M, Baker CL, Colot HV, Loros JJ, Dunlap JC (2009) A role for casein kinase 2 in the mechanism underlying circadian temperature compensation. Cell 137:749–760

Merrow MW, Garceau NY, Dunlap JC (1997) Dissection of a circadian oscillation into discrete domains. Proc Natl Acad Sci U S A 94:3877–3882

Mooney JL, Yager LN (1990) Light is required for conidiation in *Aspergillus nidulans*. Genes Dev 4:1473–1482

Morgan LW, Greene AV, Bell-Pedersen D (2003) Circadian and light-induced expression of luciferase in *Neurospora crassa*. Fungal Genet Biol 38:327–332

Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101:12248–12253

Olmedo M, Ruger-Herreros C, Luque EM, Corrochano LM (2010) A complex photoreceptor system mediates the regulation by light of the conidiation genes *con-10* and *con-6* in *Neurospora crassa*. Fungal Genet Biol 47:352–363

Pittendrigh CS, Bruce VG, Rosensweig NS, Rubin ML (1959) Growth patterns in *Neurospora*: a biological clock in *Neurospora*. Nature 184:169–170

Purschwitz J, Müller S, Kastner C, Schöser M, Haas H, Espeso EA, Atoui A, Calvo AM, Fischer R (2008) Functional and physical interaction of blue- and red-light sensors in *Aspergillus nidulans*. Curr Biol 18:255–259

Purschwitz J, Müller S, Fischer R (2009) Mapping the interaction sites of *Aspergillus nidulans* phytochrome FphA with the global regulator VeA and the white collar protein LreB. Mol Genet Genomics 281:35–42

Querfurth C, Diernfellner AC, Gin E, Malzahn E, Hofer T, Brunner M (2011) Circadian conformational change of the *Neurospora* clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain. Mol Cell 43:713–722

Rodriguez-Romero J, Hedtke M, Kastner C, Müller S, Fischer R (2010) Fungi, hidden in soil or up in the air: light makes a difference. Annu Rev Microbiol 64:585–610

Rosales-Saavedra T, Esquivel-Naranjo EU, Casas-Flores S, Martínez-Hernndez P, Ibarra-Laclette E, Cortes-Penagos C, Herrera-Estrella A (2006) Novel light-regulated genes in *Trichoderma atroviride*: a dissection by cDNA microarrays. Microbiology 152:3305–3317

Ruger-Herreros C, Rodríguez-Romero J, Fernndez-Barranco R, Olmedo M, Fischer R, Corrochano LM, Canovas D (2011) Regulation of conidiation by light in *Aspergillus nidulans*. Genetics 188: 809–822

Ruiz-Roldan MC, Garre V, Guarro J, Marine M, Roncero MIG (2008) Role of the White Collar 1 photoreceptor in carotenogenesis, UV resistance, hydrophobicity, and virulence of *Fusarium oxysporum*. Eukaryot Cell 7:1227–1230

Ruoff P, Loros JJ, Dunlap JC (2005) The relationship between FRQ-protein stability and temperature compensation in the *Neurospora* circadian clock. Proc Natl Acad Sci U S A 102:17681–17686

Salichos L, Rokas A (2010) The diversity and evolution of circadian clock proteins in fungi. Mycologia 102:269–278

Sancar G, Sancar C, Brunner M, Schafmeier T (2009) Activity of the circadian transcription factor White Collar Complex is modulated by phosphorylation of SP-motifs. FEBS Lett 583:1833–1840

Sancar G, Sancar C, Brunner M (2012) Metabolic compensation of the *Neurospora* clock by a glucose-dependent feedback of the circadian repressor CSP1 on the core oscillator. Genes Dev 26:2435–2442

Sano H, Kaneko S, Sakamoto Y, Sato T, Shishido K (2009) The basidiomycetous mushroom *Lentinula edodes white collar-2* homolog PHRB, a partner of putative blue-light photoreceptor PHRA, binds to a specific site in the promoter region of the *L edodes* tyrosinase gene. Fungal Genet Biol 46: 333–341

Sargent ML, Briggs WR, Woodward DO (1966) Circadian nature of a rhythm expressed by an invertaseless strain of *Neurospora crassa*. Plant Physiol 41:1343–1349

Schneider MJ, Murray BJ (1979) Phytochrome mediation of uredospoore germination in the fungus *Puccinia graminis*. Photochem Photobiol 29:1051–1052

Schuster A, Kubicek CP, Friedl MA, Druzhinina IS, Schmoll M (2007) Impact of light on *Hypocrea jecorina* and the multiple cellular roles of ENVOY in this process. BMC Genomics 8:449

Schwerdtfeger C, Linden H (2003) VIVID is a flavoprotein and serves as a fungal blue light photoreceptor for photoadaptation. EMBO J 22:4846–4855

Selby CP, Sancar A (2006) A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. Proc Natl Acad Sci U S A 103:17696–17700

Sharma AK, Foster KW (1997) Rhodopsin guides fungal phototaxis. Nature 387:465–466

Sharma AK, Spudich JL, Doolittle WF (2006) Microbial rhodopsins: functional versatility and genetic mobility. Trends Microbiol 14:463–469

Shi M, Collett M, Loros JJ, Dunlap JC (2010) FRQ-interacting RNA helicase mediates negative and positive feedback in the *Neurospora* circadian clock. Genetics 184:351–361

Silva F, Torres-Martínez S, Garre V (2006) Distinct *white collar-1* genes control specific light responses in *Mucor circinelloides*. Mol Microbiol 61:1023–1037

Silva F, Navarro E, Peñaranda A, Murcia-Flores L, Torres-Martínez S, Garre V (2008) A RING-finger protein regulates carotenogenesis via proteolysis-independent ubiquitylation of a White collar-1-like activator. Mol Microbiol 70:1026–1036

Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF, Thomas TL, Carrington JC, Stajich JE, Bell-Pedersen D, Brunner M, Freitag M (2010) Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for *Neurospora* white collar complex. Eukaryot Cell 9:1549–1556

Tan KK (1974) Red-far-red reversible photoreaction in the recovery from blue-light inhibition of sporulation in *Botrytis cinerea*. J Gen Microbiol 82:201–202

Tang CT, Li S, Long C, Cha J, Huang G, Li L, Chen S, Liu Y (2009) Setting the pace of the *Neurospora* circadian clock by multiple independent FRQ phosphorylation events. Proc Natl Acad Sci U S A 106:10722–10727

Tisch D, Schmoll M (2010) Light regulation of metabolic pathways in fungi. Appl Microbiol Biotechnol 85:1259–1277

Veluchamy S, Rollins JA (2008) A CRY-DASH-type photolyase/cryptochrome from *Sclerotinia sclerotiorum* mediates minor UV-A-specific effects on development. Fungal Genet Biol 45:1265–1276

Waschuk SA, Bezerra AG, Shi L, Brown LS (2005) *Leptosphaeria* rhodopsin: bacteriorhodopsin-like proton pump from a eukaryote. Proc Natl Acad Sci U S A 102:6879–6883

Yamazaki Y, Kataoka H, Miyazakl A, Watanabe M, Ootakl T (1996) Action spectra for photoinhibition of sexual development in *Phycomyces blakesleeanus*. Photochem Photobiol 64:387–392

Zoltowski BD, Schwerdtfeger C, Widom J, Loros JJ, Bilwes AM, Dunlap JC, Crane BR (2007) Conformational switching in the fungal light sensor Vivid. Science 316:1054–1057

7 Genomics and Transcriptomics to Analyze Fruiting Body Development

Minou Nowrousian[1]

CONTENTS

I. Introduction

Fungal fruiting bodies are multicellular structures that are produced by filamentous ascomycetes and basidiomycetes for the generation, protection, and dispersal of sexual spores (Pöggeler et al. 2006a; Wösten and Wessels 2006). Fruiting body formation involves the spatiotemporal coordination of the differentiation of specialized cell types, many of which are exclusively found in fruiting bodies (Moore 1998; Bistis et al. 2003). These drastic morphogenetic changes, compared to the vegetative mycelium, most likely depend on changes in activity of a large number of genes, similar to developmental processes in animals and plants in which coordinated patterns of gene expression orchestrate the spatiotemporal sequence of events that leads to embryonic or organ development (Guitton and Berger 2005; Scheres 2007). However, as the last common ancestor of fungi and animals was most likely unicellular (Lang et al. 2002; King 2004), multicellular development in fungi evolved independently of other groups; therefore, regulatory genes and networks that govern animal development are not necessarily similar in fungi.

The genetic and molecular basis of fruiting body development has been under investigation for several decades, and much of the progress in this area of research has been made through the analysis of developmental mutants and their corresponding individual genes (Moore 1998; Pöggeler et al. 2006a; Wösten and Wessels 2006). However, in the last approximately 15 years, a number of high-throughput techniques, so-called -omics techniques, have been established in quick succession, enabling the increasingly detailed analysis of complete genomes and their activity, especially at the level of the transcriptome. These technologies include, for example, EST (expressed sequence tag) sequencing, SAGE (serial analysis of gene expression), microarray hybridizations, and

[1]Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, Bochum 44780, Germany; e-mail: minou.nowrousian@rub.de

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

a multitude of applications made possible through next-generation sequencing (NGS) techniques (Velculescu et al. 1995; Churchill 2002; Nowrousian et al. 2004; Ehrenreich 2006; Nowrousian 2007, 2010; Shendure and Li 2008). These technologies have led to a rapid increase in the number of sequenced fungal genomes, as well as genome-wide expression studies to analyze different aspects of fungal biology, including fruiting body development. In addition, the generation of fungal strains with reduced rates of nonhomologous end joining has greatly facilitated targeted gene deletions in many fungal species (Ninomiya et al. 2004; Kück and Hoff 2010), thereby enabling large-scale knockout projects ranging from deletions of all genes from a gene family to genome-wide knockout libraries (see Sect. II.B). Thus, we are now at a point where, in principle, high-throughput expression studies can rapidly identify candidate genes with interesting expression patterns during fruiting body development, and these genes can be characterized in more detail through libraries of deletion mutants. Mutants themselves can also be used for expression studies, thereby establishing molecular phenotypes to characterize mutant strains at the molecular level and to unravel the regulatory networks that govern differentiation processes. While knockout libraries are currently available for only a few model organisms, high-throughput expression studies can be performed also in nonmodel species (Nowrousian 2010). These data can be used for comparative studies to identify evolutionary conserved expression patterns, which in turn are useful for choosing candidate genes for downstream analysis because evolutionary conservation of expression is a powerful criterion for functional significance (Stuart et al. 2003). Thus, high-throughput techniques have made possible the characterization of select model organisms at an unprecedented depth as well as the analysis of a much broader range of nonmodel organisms, both on a genome-wide scale. In this review, I give an overview of how genomics and transcriptomics studies have increased our understanding of fruiting body development and point out some possible future directions in this field.

II. From Genome to Function

Fungi have always been at the forefront of genome-sequencing efforts, with the first sequenced eukaryotic genome that of the yeast *Saccharomyces cerevisiae* (Goffeau et al. 1996), with the first genome sequence of a filamentous fungus, namely that of *Neurospora crassa*, following some years later (Galagan et al. 2003). Because fungal genomes are usually less than 100 Mb in size and therefore smaller and easier to sequence than genomes of many other eukaryotes, fungi (together with animals) are the eukaryotic group with the highest number of sequenced genomes to date (Table 7.1). This also applies to the current rise in the number of genomes sequenced by NGS techniques, where fungi were (again) among the first eukaryotes whose genomes were assembled from NGS sequence reads (DiGuistini et al. 2009; Nowrousian et al. 2010).

The genome sequences of filamentous fungi not only allow downstream analyses like large-scale knockout studies and mutant genome-sequencing projects that yield more insights into fungal development (see Sects. II.B and II.C) but also were the basis for efforts to induce fruiting body formation in supposedly asexual fungi as described in the next section.

A. Sexual Development in "Asexual" Fungi

For an estimated 20 % of described fungal species, no sexual stage is known (Dyer and O'Gorman 2011), and these species were historically grouped together as **deuteromycetes**. First hints that supposedly asexual fungi might have "cryptic" sexual development came from a population genetic study of the pathogenic ascomycete *Coccidioides immitis* and from the isolation of a mating-type gene from the asexual fungus *Bipolaris sacchari* (Burt et al. 1996; Sharon et al. 1996); however, it was not clear at that time whether these findings were the exception or the rule. Therefore, it

Table 7.1. Genome assemblies in the National Center for Biotechnology Information (NCBI) Genome database

Group	Chromosome assemblies[a]	Contig/scaffold assemblies[a]	Total[a]
Fungi	40	165	205
Animals	43	162	205
Plants	19	32	51
Protists	20	53	73
Others	6	0	6
Total	128	412	540

[a]The number of genomes from different eukaryotic groups for which complete chromosomes or assembled contigs or scaffolds are available was calculated from data from http://www.ncbi.nlm.nih.gov/genome/browse/ in October 2012. For each species, only one sequenced genome was counted

came as quite a surprise that virtually all sequenced genomes from supposedly asexual ascomycetes had a full complement of genes known to be involved in sexual differentiation, including **mating-type genes** that were shown to be essential for sexual development in other fungi. This includes, for example, Aspergillus and Penicillium species (Pöggeler 2002; Galagan et al. 2005; Paoletti et al. 2005; Pel et al. 2007; Hoff et al. 2008; van den Berg et al. 2008). These results led to renewed efforts to induce fruiting body formation in supposedly asexual species or strains under laboratory conditions, and the development of sexual structures could be induced, for example, for *Aspergillus fumigatus*, *Aspergillus parasiticus*, *Penicillium chrysogenum*, and the *Trichoderma reesei* strain QM6a (Horn et al. 2009; O'Gorman et al. 2009; Seidl et al. 2009; Böhm et al. 2013).

One key finding from genome sequencing was that the mating-type loci of the sequenced asexual strains indicated that the corresponding species might be heterothallic, that is, not self-fertile. In heterothallic fungi, sexual development can only take place when two strains with different mating types interact (Pöggeler et al. 2006a). Based on this information, researchers screened natural isolates for their mating type and combined strains with different mating types (Horn et al. 2009; O'Gorman et al. 2009; Seidl et al. 2009; Böhm et al. 2013).

The demonstration of sexual development in former deuteromycetes is not only important to increase our knowledge about fungal biology but also can be used as a valuable tool for basic and applied research with these fungi because the ability to perform genetic crosses is a major advantage for any organism under study. This is especially interesting for species that are used in industrial production because strain improvement based on genetic recombination offers a much wider scope than one based solely on conventional mutagenesis or genetic engineering.

B. Large-Scale Knockout Studies

One of the most powerful approaches to analyze gene functions is the inactivation of the gene under investigation and the analysis of the resulting phenotype. Targeted gene deletions in fungi are usually accomplished by **homologous recombination** of a knockout construct containing sequences upstream and downstream of the Open Reading Frame (ORF) flanking a selection marker. The knockout construct is transformed into the wild type or a recipient strain deficient in nonhomologous end joining, and transformants are screened for those where the deletion cassette has replaced the original gene sequence (Ninomiya et al. 2004; Kück and Hoff 2010). The availability of complete genome sequences allows such **forward genetic studies** at a much greater scale because the necessary sequence information to generate gene deletion constructs is readily available for all predicted genes of a sequenced species. *Saccharomyces cerevisiae* was the first eukaryote for which a knockout project for all predicted genes was initiated (Winzeler et al. 1999), and the genome-wide knockout project for the filamentous fungus *N. crassa* followed shortly after its genome was sequenced (Colot et al. 2006; Dunlap et al. 2007). Deletion strains for most *N. crassa* genes are now available

through the Fungal Genetics Stock Center (FGSC) (McCluskey and Wiest 2011), and morphological screens for basic characteristics, including sexual development, are under way. The results are made available through the *Neurospora crassa* Database (http://www.broadinstitute.org/annotation/genome/ neurospora/MultiHome.html). In a pilot study, knockout strains for about 100 transcription factor genes were generated and characterized, and 15 strains displayed phenotypes related to sexual development (Colot et al. 2006). An even higher fraction of serine-threonine protein kinase mutants (32 of 77) had defects in sexual differentiation in a study by Park et al. (2011).

Knockout screens of mutants for certain classes of genes have also been performed with *Gibberella zeae* (anamorph *Fusarium graminearum*). Based on the genome sequence (Cuomo et al. 2007), 15 polyketide synthase genes were identified and disrupted (Gaffoor et al. 2005). All of the mutants were still fertile; however, one of them produced nonpigmented fruiting bodies instead of the usual black ones. A recent study of 657 deletion mutants for predicted transcription factor genes showed that 105 of them were involved in sexual development (Son et al. 2011). Many of these mutants also had other phenotypes, mainly influencing growth, virulence, and toxin production. Generally, transcription factor genes associated with sexual development tended to be conserved throughout the fungal kingdom.

Mutants from such large-scale knockout screens can now be used for detailed functional characterization of the corresponding genes. Furthermore, these strains are also available as marker strains in genetic crosses e.g., to finely map unknown mutations in strains that were generated by classical mutagenesis (Hammond et al. 2011). Such mapping strategies can be used in combination with whole-genome sequencing to identify mutated genes as described in the next section.

C. Genome Sequencing to Identify Mutations

Prior to the possibility of generating knockout strains, many mutagenesis efforts centered around classical radiation or chemical mutagenesis, thereby generating random mutations. For many species, large **mutant collections** are available, and often such mutants have phenotypes different from knockout strains, when the function of the derived protein is not completely lost, but modified. Thus, even in the days of large-scale knockout projects, the identification of the mutated genes in such strains is an important tool for molecular and genetic analysis. The most common ways to identify mutated genes used to be genetic mapping, complementation by transformation, or a combination of both. However, these strategies are often laborious and time consuming. In recent years, the advent of NGS techniques has made the sequencing of mutant genomes and the identification of mutations by subsequent bioinformatics analysis a viable alternative. In filamentous fungi, this approach has been used to identify mutated genes in classical mutants of the ascomycetes *N. crassa* and *Sordaria macrospora* (McCluskey et al. 2011; Pomraning et al. 2011; Nowrousian et al. 2012).

Pomraning and coworkers (2011) crossed a mutant with a wild type with different genetic background, sequenced pools of DNA from individual progeny from the cross, and identified a region where sequence variants from the mutant background accumulated. The mutation causing cell cycle defects could be narrowed down to one of the genes in this genomic region.

In a study by McCluskey et al. (2011), 18 classical mutants for which the mutations had previously been mapped to genomic regions between 80 kb and 4 Mb were sequenced, and candidate genes or regions were identified for 17 of these strains. Six of the mutants for which mutated candidate genes were found are defective in sexual development as they were male sterile (mb-2, mb-3), female sterile (ff-1, fs-n), or producing underpigmented ascospores (ts, perithecial-1). A similar study in *S. macrospora* made use of pooled DNA from progeny of a cross from a mutant with an isogenic wild type, and in this case, causative mutations could be identified from the sequence data without prior genetic mapping (Nowrousian et al. 2012). The three mutants that were studied are all defective in sexual development, with two of

them unable to develop mature fruiting bodies, while the third produces light-brown instead of black ascospores. One of the sterile mutants is allelic to the pro41 mutant carrying a deletion of a gene encoding an endoplasmic reticulum (ER) membrane protein (Nowrousian et al. 2007a), while the second sterile mutant contains a point mutation in a transcription factor gene (*pro44*), and the spore color mutant fus has a defective melanin biosynthesis gene. Thus, mutant sequencing is a quick and efficient method to identify developmental genes from mutants generated by classical mutagenesis.

III. Transcriptomics: The Genome in Action

As mentioned in the Introduction, the profound morphogenetic changes that a mycelium undergoes during fruiting body development depend on corresponding changes in activity of large numbers of genes, and many of these changes are thought to be implemented at the level of transcription. Therefore, transcriptomics approaches (i.e., the analysis of transcript levels for many or all genes in a genome under certain conditions or in different cell types, organs, etc.) are especially suited for the analysis of fruiting body development. A range of transcriptomics methods has been applied in a number of asco- and basidiomycetes to study sexual differentiation (Table 7.2). These approaches have the advantage of allowing a bias-free insight into transcriptional expression because they do not access only a selected subset of genes, but rather yield a global picture. On the other hand, they also allow hypothesis-driven research, for example, the analysis of epistatic relationships or regulatory networks through comparing molecular phenotypes of single and double mutants. In the following sections, results from transcriptomics analyses to study gene expression throughout development or to characterize the molecular phenotype of developmental mutants are summarized.

A. Transcriptomics to Identify Global Spatiotemporal Expression Patterns During Development

Transcriptomics started with EST analyses (Adams et al. 1991) but soon afterward diversified through a number of sequencing- or hybridization-based methods, including SAGE (Velculescu et al. 1995), microarrays (Schena et al. 1995; DeRisi et al. 1997; Breakspear and Momany 2007), and RNA-seq (Nagalakshmi et al. 2008; Wilhelm et al. 2008). All of these methods have by now been applied to analyze fruiting body development (Table 7.2). These studies have not only improved our understanding of the processes that occur during developmental transitions, but also were instrumental in identifying a number of differentially expressed genes whose functions during fruiting body development were determined in subsequent studies.

1. Transcriptomics of Different Developmental Stages

The first transcriptomics approach to learn more about gene expression at different developmental stages was an **EST analysis** of fruiting bodies (perithecia), germinating asexual spores (conidia), and vegetative mycelium from *N. crassa* (Nelson et al. 1997). Nearly 60 % of the ESTs in the mycelial and perithecial libraries showed no match to proteins in public databases, whereas for germinating conidia, this was true for less than 40 %. One reason for this might be that the ESTs from the conidial library contained more genes with putative roles in protein biosynthesis, and that this group of genes was already well characterized, whereas not many genes involved in fungal-specific activities like fruiting body development were known at this time. However, even 15 years later, it is a recurring theme in many expression analyses that among the genes expressed during sexual development are relatively fewer involved in basic cellular functions and metabolism than in nonreproductive mycelia, and that the genes expressed during sexual development have more diverse functions or no characterized function yet. This was found in analyses of the ascomycetes

Table 7.2. Transcriptomics analyses to study sexual development in filamentous ascomycetes and basidiomycetes

Species	Analysis[a]	Subject	References[b]
Ascomycetes			
Pezizomycetes			
Pyronema confluens	E, D, R	Vegetative mycelium and sexually developing mycelium	[1, 2, 48]
Tuber borchii	E, M	Fruiting bodies versus vegetative mycelium	[3, 34]
Tuber melanosporum	E, M, R	Fruiting bodies versus free-living or ectomycorrhizal mycelia	[4, 5]
Eurotiomycetes			
Aspergillus nidulans	M	Wild-type and sterile mutant, vegetative and sexual development	[6]
Sordariomycetes			
Cryphonectria parasitica	M	Virus infected versus noninfected	[7]
	M	wild type versus sterile mutant	[42]
Gibberella fujikuroi	D, M	Wild-type and mating-type gene mutant	[8]
	R	Time course of perithecial development	[9]
Gibberella zeae	E	Vegetative mycelium and perithecia	[39]
	M, R	Time course of perithecial development	[9–11]
	D, M	Wild-type and mating-type gene mutant	[40]
	M	Wild-type versus sterile mutant	[12, 13]
Hypocrea jecorina	M	Sexually potent versus impotent mycelia	[14]
Neurospora crassa	E	Vegetative hyphae, conidia, and perithecia	[15]
	M	Wild-type versus sterile mutants	[41]
	M	Wild-type mat a and mat A mycelia under vegetative conditions	[16]
	M	Time course of perithecial development on two media	[17]
Neurospora intermedia	E, M	Vegetative and sexual mycelium	[46]
Ophiostoma novo-ulmi	D	SSH of perithecia versus synnemata	[18]
Podospora anserina	M	Wild-type mat+ and mat- mycelia, and mating-type gene mutants	[19]
	M	Wild-type and developmental mutants during vegetative growth	[45]
Sordaria macrospora	M	Wild-type versus sterile mutants	[20]
	M	Wild-type versus mating type gene mutants	[21, 22]
	M	Wild-type versus sterile mutants, and vegetative mycelium versus sexually developing mycelium	[23]
	R	Protoperithecia of wild-type versus sterile mutant and vegetative mycelium versus sexually developing mycelium	[24]
Basidiomycetes			
Agaricomycetes			
Agaricus bisporus	E, M	Different stages of fruiting body development	[37, 44]
Coprinopsis cinerea	S, M	Fruiting bodies versus vegetative mycelium	[47]
Ganoderma lucidum	R	Fruiting bodies versus vegetative mycelium	[25]
Laccaria bicolor	M	Fruiting bodies, vegetative mycelium, mycorrhizal root tips	[43]
Lentinula edodes	D	RDA of vegetative mycelium and stages of fruiting body development	[26]
	S	Dikaryotic mycelium and fruiting body primordium	[27]
	D	SSH of fresh versus harvested fruiting bodies	[28]
	R, S	Fruiting bodies before and after sporulation	[29]
Pleurotus ostreatus	E	Different stages of fruiting body development	[30, 35, 36]

(continued)

Table 7.2. (continued)

Species	Analysis[a]	Subject	References[b]
Schizophyllum commune	MP	Time course of fruiting body development	[31]
	M	Monokaryons and compatible and incompatible dikaryotic interactions	[32]
	R	Wild-type versus developmental mutants	[33, 38]

[a]Type of analysis: *E* EST (expressed sequence tag) sequencing, *D* differential hybridization (including *RDA* representational difference analysis and *SSH* suppression subtractive hybridization), *M* microarray analysis including cross-species array hybridizations, *MP* massively parallel signature sequencing, *R* RNA-seq, *S* SAGE (serial analysis of gene expression)
[b]References are as follows: [1] Nowrousian and Kück (2006), [2] Gesing et al. (2013), [3] Gabella et al. (2005), [4] Martin et al. (2010), [5] Tisserant et al. (2011), [6] Nahlik et al. (2010), [7] Allen et al. (2003), [8] Keszthelyi et al. (2006), [9] Sikhakolli et al. (2012), [10] Qi et al. (2006), [11] Hallen et al. (2007), [12] Lee et al. (2010a), [13] Lysoe et al. (2011), [14] Chen et al. (2012a), [15] Nelson et al. (1997), [16] Wang et al. (2012a), [17] Wang et al. (2012b), [18] Jacobi et al. (2010), [19] Bidard et al. (2011), [20] Nowrousian et al. (2005), [21] Pöggeler et al. (2006b), [22] Klix et al. (2010), [23] Nowrousian et al. (2007a), [24] Teichert et al. (2012), [25] Yu et al. (2012), [26] Miyazaki et al. (2005), [27] Chum et al. (2008), [28] Sakamoto et al. (2009), [29] Chum et al. (2011), [30] Joh et al. (2007), [31] Ohm et al. (2010), [32] Erdmann et al. (2012), [33] Ohm et al. (2013), [34] Lacourt et al. (2002), [35] Park et al. (2006), [36] Lee et al. (2002), [37] Ospina-Giraldo et al. (2000), [38] Ohm et al. (2011), [39] Trail et al. (2003), [40] Lee et al. (2006), [41] Li et al. (2005), [42] Deng et al. (2007), [43] Martin et al. (2008), [44] Morin et al. (2012), [45] Bidard et al. (2012), [46] Nygren et al. (2012), [47] Cheng et al. (2013), [48] Traeger et al. (2013)

G. zeae, *Gibberella fujikuroi*, and *Pyronema confluens* as well as in the basidiomycetes *Agaricus bisporus* and *Lentinula edodes* (Ospina-Giraldo et al. 2000; Trail et al. 2003; Miyazaki et al. 2005; Keszthelyi et al. 2006; Nowrousian and Kück 2006; Qi et al. 2006; Gesing et al. 2013; Morin et al. 2012; Sikhakolli et al. 2012). This might indicate that during vegetative growth, a set of well-characterized eukaryotic core genes is active to maintain basic cellular functions, whereas the higher morphological diversity that is achieved during fruiting body development requires the activity of a number of specialized, perhaps less-conserved (or at least less-well-characterized) or even **lineage-specific genes**. This is consistent with a long-standing hypothesis that the vegetative mycelium gathers and stores nutrients until a stage of "competence" is reached, and afterward the energetically costly production of fruiting bodies is supported by transfer of nutrients from the nonreproductive mycelium (Wessels 1993; Pöggeler et al. 2006a). It also fits with a hypothesis that sex-associated genes evolve more rapidly than genes involved in other functions. This has been shown in animals and plants (Swanson and Vacquier 2002; Clark et al. 2006; Brawand et al. 2011), and similar findings have been made for mating-type and pheromone signaling genes in filamentous ascomycetes (Pöggeler 1999; Karlsson et al. 2008). It is also supported by a recent EST analysis of *Neurospora intermedia* and comparison with other *Neurospora* species (Nygren et al. 2012). This study indicated that sex-associated genes (i.e., those genes that are preferentially expressed during sexual development) are also rapidly evolving in fungi. Further support for this hypothesis came from a transcriptomics analysis of *P. confluens* (Traeger et al. 2013). The authors showed at a genome-wide level that the highest percentage of genes upregulated during sexual development was found among the *P. confluens* orphan genes, that is, genes without detectable homolog in the public sequence databases.

Another, related finding is that transcriptional profiles of nonreproductive mycelia in many fungi are different from those of developing fruiting bodies. This was found, for example, in *Tuber borchii*, for which an analysis of two fruiting body stages and vegetative mycelium showed that the differences between the two fruiting body stages were smaller than between either stage of sexual development and vegetative mycelium (Gabella et al. 2005). Even an array of complementary DNAs (cDNAs) derived from vegetative mycelium of *T. borchii* showed large differences when hybridized with targets from vegetative mycelium

versus fruiting body, although strictly fruiting body-specific genes would not even be represented on this array (Lacourt et al. 2002). Interestingly, though, expression studies in *Tuber melanosporum*, a close relative of *T. borchii*, suggested that less than 10 % of all genes are differentially expressed in fruiting bodies of mycorrhizal tissue compared to free-living mycelium (Martin et al. 2010). However, studies in other fungi rather support a hypothesis of **transcriptional reprogramming during sexual development**: Using a genome-wide microarray to analyze a time course of *G. zeae* fruiting body differentiation, Hallen and coworkers found that 12 % of transcripts were specific to sexual development (Hallen et al. 2007). In a recent RNA-seq study of young fruiting bodies of *S. macrospora* that were extracted from surrounding nonreproductive hyphae by laser microdissection, it was found that the transcriptional profile of fruiting body precursors was drastically different from that of nonreproductive hyphae (Teichert et al. 2012), and similar results were obtained in serial analysis of gene expression studies of vegetative mycelium and fruiting body primordia from the basidiomycete *Coprinopsis cinerea* (Cheng et al. 2013).

In an analysis of fruiting body development of *N. crassa* on two different media, Wang and coworkers found that expression of metabolism-related genes was different, as expected, but that known development-specific genes were expressed similarly independent of the growth medium (Wang et al. 2012b). This indicates that expression of developmental genes is robust even in different environments. Overall, the analyses described here already offered first insights into the coordinated patterns of gene expression that underlie developmental transitions in time and space. Several of the genes that are upregulated during sexual development were already characterized in more detail; these are described in Sect. III.C.

2. Comparative Transcriptomics to Identify Evolutionary Conserved Expression Patterns

The rising number of expression studies conducted with different fungal species not only offers insights into the development of single species, but also opens up the possibility of comparing gene expression patterns between species at similar developmental stages. This is important not only to follow **evolutionary trajectories of gene expression** but also as a practical means to narrow down lists of differentially expressed genes for promising candidates for further studies. The necessity for narrowing down such lists has become apparent in the course of a number of expression studies, in which it was found that genes strongly upregulated during a condition of interest are not necessarily involved in the process under investigation (Nowrousian and Cebula 2005; Nowrousian et al. 2007b). Comparative transcriptomics analyses can help focus on genes with a higher likelihood of playing a significant role in the analyzed biological phenomenon because evolutionary conservation of expression is a strong indicator for functional significance (Stuart et al. 2003). This is especially the case if the species under investigation are only distantly related because comparing closely related species might give a high number of false positives because there is insufficient time for the divergence of expression patterns (Romero et al. 2012). Furthermore, comparative analyses can also be used to identify species-specific expression patterns that might help to explain morphological or physiological differences between species (Brawand et al. 2011; Romero et al. 2012).

Comparative expression analyses of fungal development have been performed for several ascomycetes. A small-scale pilot study was based on ESTs from *P. confluens*, whose expression was compared to the corresponding homologs from *S. macrospora* (Nowrousian and Kück 2006). These two ascomycetes are only distantly related, with *P. confluens* positioned at the base of the filamentous ascomycetes, forming apothecia as fruiting bodies, while *S. macrospora* is a derived ascomycete and develops the more complex perithecia.

Genes with conserved expression across the large evolutionary distance between these two species might therefore be among a set of core genes for ascomycete fruiting body formation. Even though this study encompassed only a few genes, it already indicated that there might be a significant amount of overlap in development-specific expression even in distantly related ascomycetes. These analyses were followed up by a comparison of developmentally regulated genes from *P. confluens* identified by suppression subtractive hybridization (SSH) with homologs in *S. macrospora* and *N. crassa* (Gesing et al. 2013). Interestingly, expression tendencies were more conserved between *S. macrospora* and *P. confluens* than between the much more closely related *S. macrospora* and *N. crassa*, both of which are members of the Sordariomycetes and form perithecia. One reason for this might be that *N. crassa* is heterothallic (self-sterile), in contrast to the homothallic (self-fertile) *S. macrospora* and *P. confluens*; therefore, *N. crassa* requires an additional fertilization step for fruiting body formation (Claussen 1912; Davis and deSerres 1970; Pöggeler et al. 2006a). Furthermore, *N. crassa* produces two different types of asexual spores (micro- and macroconidia) in addition to the sexual ascospores that are produced in the fruiting bodies, whereas neither *S. macrospora* nor *P. confluens* produces any asexual spores. Therefore, some differences in gene expression patterns between *N. crassa* and the other two species might be caused by additional differentiation processes in *N. crassa* that occur in parallel to fruiting body formation (Gesing et al. 2013). The recent sequencing of the *P. confluens* genome and several development-dependent transriptomes enabled a genome-wide comparison of transcriptome data between *P. confluens* and *S. macrospora* (Traeger et al. 2013). Interestingly, expression patterns in vegetative or total mycelia in these two fungi were more similar to each other than those of the corresponding mycelia from *S. macrospora* to expression patterns in young fruiting bodies from this species. This suggests that similar tissues or developmental stages in different species are more similar in expression than different tissues or stages from the same species.

Another case for which the differences outweigh the similarities of expression was found in a comparison of genes that are differentially expressed in **mutants of mating-type genes** from *Podospora anserina*, *S. macrospora*, and *G. fujikuroi* (Bidard et al. 2011). The authors compared microarray hybridization results for mutants in the mating-type genes *FMR1* (*MAT1-1-1*) and *FPR1* (*MAT1-2-1*) from *P. anserina* with results from mutants in the corresponding genes *SmtA-1* (*MAT1-1-1*) and *Smta-1* (*MAT1-2-1*) of *S. macrospora* and with results from the *FRP1* homolog *MAT1-2-1* of *G. fujikuroi* from previous studies (Keszthelyi et al. 2006; Pöggeler et al. 2006b; Klix et al. 2010). They found between 16 and 57 orthologous gene pairs among the regulated genes, but the direction of regulation showed little or no correlation between the respective species (Bidard et al. 2011). However, these findings might have several explanations. One caveat for the interpretation is the fact that, in the case of *S. macrospora*, the microarray hybridizations were performed using *N. crassa* microarrays (cross-species microarray hybridizations), which result in fewer genes yielding significant hybridization results. Especially, signals from weakly expressed genes might have been lost in the background noise; therefore, the lists of differentially regulated genes might not be complete (Pöggeler et al. 2006b; Klix et al. 2010). This is also the case for *G. fujikuroi*, for which cDNA macroarrays were used for the analysis that did not cover all potential genes of this fungus (Keszthelyi et al. 2006). Other reasons for the observed differences might be rooted in the biology of the fungi because *S. macrospora* is homothallic, *G. fujikuroi* is heterothallic, and *P. anserina* is pseudohomothallic, but for the analysis by Bidard et al., homokaryotic strains were used that are effectively heterothallic. The mycelia used for RNA extraction were from vegetative, fertilization-competent mycelia in the case of *G. fujikuroi* and *P. anserina* and from already sexually differentiating mycelia in the case of *S. macrospora*. Thus, the analysis for *G. fujikuroi* and *P. anserina* focused on genes that are influenced by the mating-type factors during vegetative growth prior to fertilization

(Keszthelyi et al. 2006; Bidard et al. 2011), while in the case of *S. macrospora*, the role of mating-type genes was studied during sexual development (Pöggeler et al. 2006b; Klix et al. 2010). Furthermore, an additional difference between the three species is that *S. macrospora* does not produce any asexual spores, *P. anserina* produces spores that act only as male spermatia, while *G. fujikuroi* produces asexual conidia in parallel to sexual differentiation.

The previously described studies might indicate that it is important to compare similar stages of fruiting body development, if possible without interference of unrelated differentiation processes, to identify genes with conserved expression patterns. This was done in RNA-seq analyses of six stages of perithecial development in *G. zeae* and *G. fujikuroi* (Sikhakolli et al. 2012). The authors found large overall similarities in expression patterns; for example, a high number of genes encoding proteins involved in transcription are strongly expressed during the time of ascus formation. Expression differences were largely explained by morphological differences (e.g., the different pigmentation of sexually developing cultures in these two species). However, it has to be taken into account that the two species that were compared in this case were both from the genus *Gibberella* and therefore closely related. However, a comparison of gene expression patterns during fruiting body development between *S. macrospora* and *G. zeae* also revealed a number of similarly regulated genes, with functional categories of metabolism and energy overrepresented among the down-regulated genes, and transcription/RNA processing; protein activity regulation; cell fate (differentiation, apoptosis); and cell wall biogenesis overrepresented among the upregulated genes in both species (Gesing et al. 2012). Thus, expression patterns might be to some extent conserved during similar morphological stages, at least among filamentous ascomycetes.

In basidiomycetes, **gene expression during mushroom formation** has been compared between *Schizophyllum commune* and *Laccaria bicolor* (Ohm et al. 2010) and later between these two species and the button mushroom *A. bisporus* (Morin et al. 2012). In the first comparison, it was found that several groups of genes (represented by the same gene ontology terms or other classification methods) were positively correlated with respect to their expression during mushroom differentiation. Among these groups were, for example, metabolic pathways or transcriptional regulation by transcription factors, indicating that regulatory processes might be to some degree conserved during basidiomycete fruiting body development. In the second study, it was found that only 22 and 35 genes were significantly upregulated during development in both *A. bisporus*/*S. commune* and *A. bisporus*/*L. bicolor*, respectively. Further studies including more species, and more distinct stages of development, will be necessary to determine the extent of expression conservation versus species-specific expression during fruiting body development.

B. Transcriptomics to Characterize Developmental Genes and Mutants

While the analysis of overt morphological phenotypes of mutant strains is a valuable tool for the identification of genes that are essential for fruiting body development, this does not necessarily give any information about the molecular actions that the corresponding gene product is performing to ensure the correct progression of differentiation. To understand the molecular mechanisms that are the basis of fruiting body development, a more detailed analysis of gene activities is necessary, and the genome-wide gene expression patterns of a mutant can serve as a molecular phenotype to establish the position of the mutated gene in a regulatory network. Furthermore, genes that are deregulated in a mutant strain are directly or indirectly dependent on the mutated gene for correct expression; therefore, transcriptomics studies of mutants are widely used to identify target gene candidates of transcription factors or regulatory factors.

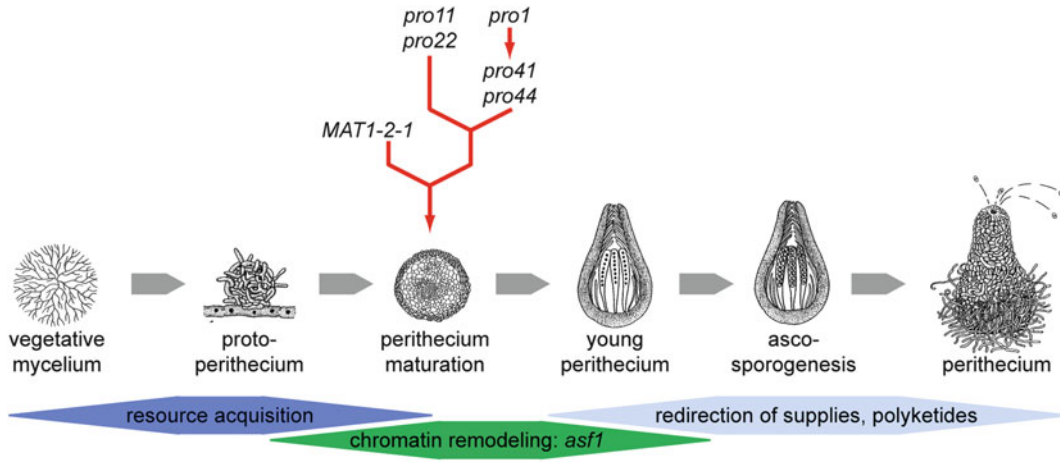


Fig. 7.1. Model for the regulation of fruiting body development in *S. macrospora* as determined through transcriptomics analyses. During vegetative growth, the mycelium gathers nutrients until a stage of competence is reached that allows fruiting body development. Maturation of protoperithecia to perithecia requires the action of several genes that act in a regulatory network that was elucidated by microarray and RNA-seq analysis of the corresponding mutants. Among others, the transcription factors PRO1 and PRO44 are necessary for the transition to perithecia, and the involvement of the histone chaperone ASF1 indicates that, in addition to gene-specific transcriptional changes, extensive chromatin remodeling might play a role in development. In later stages, nutrients are redistributed to the developing fruiting bodies, and polyketides (e.g., melanins) are necessary for timely differentiation of mature perithecia and ascospores. For more information, see text (Pictures of developmental stages are from Kück et al. 2009)

1. Transcriptomics to Analyze Mating-Type-Dependent Gene Expression

Among the first genetic loci that were identified as crucial for sexual development in fungi were the mating-type loci. While the number of loci per fungal genome and the number of genes per locus varies between different fungal groups and species, one common theme is the presence of at least one transcription factor gene among the mating-type genes in each mating system analyzed so far (Casselton 2008; Lee et al. 2010b). The best-studied mating system is that of the yeast *S. cerevisiae*, for which it was shown that the mating-type-encoded transcription factors activate or repress transcription of developmental genes during different stages of sexual differentiation (Galgoczy et al. 2004). In filamentous fungi, several mating-type genes were shown to be involved in fruiting body development; in a number of species, transcriptomics analyses were used to characterize mating-type-dependent gene expression. Most of these studies were performed with Sordariomycetes, but recently a study was carried out with the basidiomycete *S. commune* (Table 7.2).

The earliest large-scale analyses of mating-type-dependent gene expression in filamentous fungi involved comparisons between the wild type and mutants in the **mating-type gene *MAT1-2-1*** in *G. zeae*, *G. fujikuroi*, and *S. macrospora* (Keszthelyi et al. 2006; Lee et al. 2006; Pöggeler et al. 2006b). In all three analyses, up- as well as downregulated genes were found in the mating-type mutants, indicating that the *MAT1-2-1* gene acts both as repressor and as activator, and similar results were found later in the corresponding mutant of *P. anserina* (Bidard et al. 2011). In the yeast *S. cerevisiae*, the mating-type proteins, either individually or as dimers, directly regulate the expression of only about 30 genes, most of which are involved in sexual development themselves (Galgoczy et al. 2004). The number of differentially expressed genes in mating-type mutants of filamentous ascomycetes is higher; however, whether these effects are caused by direct targeting of the respective genes by the

mating-type proteins remains to be elucidated, as the transcriptomics analyses identify both direct and indirect target genes. In the case of *S. macrospora*, expression profiles of the MAT1-2-1 mutant were also compared to those of three other mutants with a block at the same developmental stage, but where the defective genes are not mating type related (Pöggeler et al. 2006b). Interestingly, even though all four mutants had similar phenotypes, the transcription profile of the MAT1-2-1 mutant was significantly different from those of the other mutants, indicating that the mating-type genes might regulate distinct developmental pathways (Fig. 7.1). MAT1-1-1 mutants were also analyzed in *S. macrospora* and *P. anserina* (Klix et al. 2010; Bidard et al. 2011); the comparative analysis was described in previous Sect. III.A.2.

Mating-type genes not only influence sexual development but also can have an effect during the asexual growth phase, as was shown in a microarray hybridization study of *N. crassa* (Wang et al. 2012a). In this fungus, the mat A (MAT1-1) mating type had an overall higher expression of many genes compared to the mat a (MAT1-2) mating type. These differences might be correlated with mating-type-dependent efficiency in fruiting body production (Dettman et al. 2003), possibly because the vegetative growth phase prior to fertilization is important for differentiation of fruiting body precursors (protoperithecia).

While the known mating systems from ascomycetes are bipolar (i.e., there exists only one mating-type locus per genome), many basidiomycetes possess tetrapolar mating systems with two independent mating-type loci in a genome (Casselton and Kües 1994; Kämper et al. 1994). For compatible interactions that lead to fruiting body formation in basidiomycetes with tetrapolar mating systems, alleles at both mating-type loci must be different in the mating partners. Mating-type-dependent gene expression in basidiomycetes was first studied at a genome-wide level in the plant-pathogenic *Ustilago maydis* (Wahl et al. 2010); however, this fungus does not produce fruiting bodies. The first mating-type-dependent transcriptome of a mushroom-forming basidiomycete was analyzed in *S. commune* (Erdmann et al. 2012). Here, it was shown that the **mating-type loci, A and B**, regulate the expression of distinct sets of genes as well as several genes that are under control of both mating-type loci. Interestingly, more genes are differentially regulated by the pheromone/receptor system-encoding B mating-type locus than the transcription factor-encoding A mating-type locus. However, more studies with basidiomycetes will be needed to determine the extent of regulatory influence of different mating-type loci, especially with respect to spatiotemporal developmental effects.

2. Transcriptomics of Photoreceptor Mutants

Light has an effect on fruiting body development in many asco- and basidiomycetes, although the effects of light are species specific, ranging from completely light-dependent development to inhibition of fruiting body differentiation in constant light (Claussen 1912; Perkins 1969; Moore-Landecker 1979; Harding and Melles 1983; Kües 2000; Pöggeler et al. 2006a; Fischer 2008; Chen et al. 2012a). Furthermore, the wavelengths that affect fruiting body development can vary; for example, **blue light** affects differentiation in the ascomycetes *N. crassa* and *Hypocrea jecorina*, as well as the basidiomycete *S. commune*, whereas **red light** represses sexual development in the ascomycete *Aspergillus nidulans*. These effects are mediated by photoreceptors, in the case of blue light by the White Collar Complex (WCC) in *N. crassa* and its homologs in other fungi and the phytochrome FphA for the red light responses of *A. nidulans* (Oda and Hasunuma 1997; Froehlich et al. 2002; Blumenstein et al. 2005; Chen et al. 2012a; Ohm et al. 2013).

The White Collar proteins WC-1 and WC-2 that form the WCC are transcription factors as well as photoreceptors (Ballario et al. 1996; Crosthwaite et al. 1997; Linden and Macino 1997; Talora et al. 1999), and transcriptomics analyses have been performed with several fungi to identify genes that are dependent on the White Collar genes for correct developmental expression.

In *N. crassa*, microarray analyses were performed to identify light-responsive genes by comparing the wild type with several photoreceptor mutants, among them

wc-1 and *wc-2* mutants (Chen et al. 2009). While the conditions under which the mycelia were grown in this case allowed only vegetative development, it is interesting to note that one of the targets of the WCC that was identified is the *sub-1* gene. It encodes a GATA-type transcription factor that is essential for fruiting body development in *N. crassa*, and its orthologs are also essential for sexual development in *A. nidulans* and *S. macrospora* (Han et al. 2001; Colot et al. 2006; Nowrousian et al. 2012). In *N. crassa*, *sub-1* is required for most late light responses (Chen et al. 2009), and one might speculate that the light-dependent orientation of fruiting body necks in this fungus is mediated by *sub-1*.

In *H. jecorina*, it was shown recently that constant light inhibits fruiting body formation, and this effect is mediated by the White Collar homologs *blr1* and *blr2* (Chen et al. 2012a). Light effects are dampened by the *env1* gene, a homolog to the *N. crassa vvd* gene that was shown previously to be involved in modulating light effects in this fungus (Heintzen et al. 2001). Microarray analyses of the wild type and a Δenv1 mutant grown under different light regimes indicated that ENV1 might be involved in balancing expression of genes required for vegetative growth, asexual conidiation, and sexual development (Chen et al. 2012a).

White Collar homologs are also involved in the regulation of sexual development in the basidiomycete *S. commune* (Ohm et al. 2013). In this fungus, fruiting body formation is light dependent, and dikaryotic deletion strains of *wc-1* or *wc-2* are unable to form fruiting bodies, indicating that the White Collar proteins mediate the light responses that are needed for sexual development. RNA-seq analyses of the wild type and a Δwc-2Δwc-2 dikaryon identified a number of genes that are differentially expressed in the mutant. Among the genes downregulated in the mutant were several hydrophobin genes and two transcription factors that were shown previously to be involved in fruiting body formation (Ohm et al. 2011, 2013). So far, transcriptomics data have shown that light effects are integrated with developmental decisions, and that at least to some extent this is mediated at the level of transcription.

3. Transcriptomics of Other Developmental Mutants

Over the last decades, many developmental genes have been identified in several filamentous fungi, among them, for example, genes encoding transcription factors, signaling proteins, and others. A number of the corresponding mutant strains have been subjected to transcriptomics analysis to characterize the molecular phenotype of the mutants on a genome-wide basis. With a complex phenotype like multicellular differentiation, this approach is especially useful to cover the diverse effects of mutations in central developmental genes.

Most of these studies have been performed with Sordariomycetes (Table 7.2). In *N. crassa*, a microarray analysis was conducted with mutants in *mak-2* and *pp-1*, encoding homologs of the yeast mitogen-activated protein (MAP) kinase Fus3p and transcription factor Ste12p, respectively, both of which are part of a signal transduction cascade that is essential for sexual development in yeast (Li et al. 2005). Both genes are essential for fruiting body formation in *N. crassa*, and most genes that are differentially expressed in the mutants compared to the wild type show the same pattern of deregulation in both mutants, confirming that the two genes also might be part of the same signal transduction cascade in *N. crassa*. Interestingly, the pheromone-like gene *poi-2* that was identified in a previous EST analysis (Nelson et al. 1997) is downregulated in both mutants, while the true pheromone gene *ccg-4* is upregulated. Upregulation of pheromone genes was also found in microarray analyses of three different developmental mutants from *S. macrospora* (Nowrousian et al. 2005). It can be hypothesized that in all these cases, the failure to progress through certain developmental stages prevents downregulation of genes like the pheromone genes, something that would occur during normal development in the wild type.

Gene expression patterns constitute **molecular phenotypes** and can therefore be used to determine the position of genes within regulatory networks by the analysis of transcriptomes from single and double mutants. This

was done, for example, for the *Sordaria macrospora* genes *pro1* and *pro41*, which encode a transcription factor and ER membrane protein, respectively, both of which are required for fruiting body formation (Masloff et al. 1999; Nowrousian et al. 2007a). Microarray analysis of single and double mutants and clustering of the mutant strains according to their gene expression patterns revealed that *pro1* acts upstream of *pro41* in a genetic network of development (Fig. 7.1). An even more detailed analysis of gene expression in the mutant *pro1* was performed by combining laser microdissection and RNA-seq to analyze the transcriptome specifically in young fruiting bodies (protoperithecia) because the mutant is able to differentiate these, but cannot progress beyond this step (Teichert et al. 2012). On the one hand, these analyses showed that gene expression in protoperithecia is drastically different from that in nonreproductive hyphae; on the other hand, a number of genes were identified that depend on the transcription factor gene *pro1* for correct expression in young fruiting bodies. One example is the transcription factor gene *pro44* that is essential for fruiting body development (Nowrousian et al. 2012) and strongly expressed in wild-type protoperithecia, but not in pro1 protoperithecia, and therefore probably acts downstream of *pro1* in a regulatory network (Fig. 7.1).

The effect of a transcription factor mutation on the transcriptome was also studied in the sterile mutant of the *FgStuA* gene from *G. zeae* (Lysoe et al. 2011). However, the growth conditions used in these experiments did not support sexual differentiation; therefore, conclusions on the role of FgStuA in sexual development will require additional data.

In *G. zeae*, several microarray analyses were conducted to analyze genes that might be involved in signaling during development (Hallen and Trail 2008; Lee et al. 2010a). The *cch1* gene encodes a calcium channel that is necessary for ascospore discharge from mature perithecia. Transcriptomes were recorded at three developmental stages, and several hundred genes were differentially regulated in the mutant compared to the wild type at each stage, with little overlap between stages (Hallen and Trail 2008). This indicates that the calcium channel CCH1 has a significant impact on the physiological state of the mycelium, and that this effect might be stage specific during development. Microarray hybridizations were also used to compare transcriptomes between the wild type and a mutant of the G protein alpha subunit gene *gpa1* (Lee et al. 2010a). The mutant is sterile, and it was found that several hundred genes are up- or downregulated in the mutant strain 3 days after induction of development. Detailed investigation of 100 downregulated genes showed that the majority of these were also upregulated in the wild type during sexual development compared to vegetative growth. Deletion of 57 of these genes served to identify 11 genes involved in sexual differentiation because the corresponding mutants have phenotypes ranging from sterility to delayed fruiting body maturation. Thus, the signal transduction cascade that contains *gpa1* regulates a number of genes that are necessary for fruiting body formation themselves.

An interesting case is fruiting body development in the chestnut blight fungus *Cryphonectria parasitica*. In this fungus, sexual differentiation (among other traits) is impaired in strains that are infected with certain hypoviruses (Dawe and Nuss 2001). Microarray analysis showed that about 13 % of all genes represented on the arrays were differentially expressed in hypovirus-infected versus noninfected strains (Allen et al. 2003). Thus, hypovirus infection causes a broad reprogramming of the host transcriptome. Interestingly, one of the genes that are downregulated during hypovirus infection is *CpST12*, encoding a homolog of the yeast transcription factor Ste12p. It was shown in subsequent analyses that this gene is also essential for female fertility in *C. parasitica* (Deng et al. 2007).

While most transcriptomics analyses of developmental mutants were conducted with Sordariomycetes, there have also been some reports from other fungal groups (Table 7.2). In the Eurotiomycete *A. nidulans*, a combined transcriptomics/proteomics/metabolomics study was carried out to analyze the genome-wide effects of the deletion of *csnE*, a gene encoding a subunit of the COP9 signalosome (Nahlik et al. 2010). The COP9 signalosome is a multisubunit protein complex that regulates

ubiquitin ligase activity and is conserved in higher eukaryotes. Deletion of individual subunits results in identical phenotypes in *A. nidulans* because the complex can only be assembled when all required subunits are present (Busch et al. 2007). The multilevel gene expression analysis of the csnE mutant showed that during vegetative growth, the COP9 signalosome is involved in protection against oxidative stress and the regulation of hormone levels, whereas later during development, it activates genes for secondary metabolism and cell wall restructuring (Nahlik et al. 2010).

In basidiomycetes, not many developmental mutants have been characterized at the molecular level yet; however, a transcriptome analysis by RNA-seq was performed recently to analyze the effect of transcription factor deletions on gene expression during fruiting body formation in *S. commune* (Ohm et al. 2011). Mutants in the transcription factor genes *fts4* and *hom2* are blocked early during mushroom development, and transcriptomes were analyzed for an early developmental stage in the wild type and the two mutants. At this stage, expression profiles of the mutants were more similar to each other than to the wild type. Interestingly, two other transcription factor genes that are also involved in fruiting body formation, *hom1* and *c2h2*, were downregulated in both mutants. *hom1* regulates number and size of mushrooms, with the mutant forming more, but smaller fruiting bodies, while the c2h2 mutant is arrested during development (Ohm et al. 2011). The transcriptomics data together with information about mutant phenotypes could thus be integrated in a model of the transcription factor network that regulates mushroom differentiation.

C. Detailed Analysis of Genes Differentially Expressed During Fruiting Body Development

Apart from establishing genome-wide transcriptional patterns, gene expression data can also be used to identify target genes for further detailed analyses. By now, a number of genes that are differentially expressed during fruiting body development in one or more fungi have already been analyzed in more detail to determine whether they play a role in this process.

As mentioned in Sect. III.A.1, one of the first transcriptomics analyses in a filamentous fungus was carried out in *N. crassa* (Nelson et al. 1997), and two of the genes, *poi-1* (*mfa-1*) and *poi-2*, that were found to be strongly expressed in starved mycelia and perithecia, but not in conidia, were later shown to be essential for correct fruiting body development (Kim et al. 2002; Kim and Nelson 2005). While *poi-1* encodes a pheromone gene and was therefore later renamed *mfa-1* (*mating factor α-1*), *poi-2* has some structural similarities to fungal pheromones but seems to be involved in a mating response signaling pathway rather than acting as a pheromone itself.

In *S. macrospora*, a number of genes that are regulated differentially in developmental mutants were analyzed in more detail. The first two genes were *tap1* and *app*, encoding a putative lectin and a protein that turned out to be among the most abundant in mature perithecia, respectively (Nowrousian and Cebula 2005; Nowrousian et al. 2007b). Somewhat surprisingly at that time, none of the corresponding deletion mutants showed any developmental phenotype, at least under laboratory conditions. This led to the conclusion that more stringent or specific criteria were needed to identify promising candidate genes from expression studies. Several criteria can be envisioned, among them a comparative approach in which candidate genes are selected from those with evolutionary conserved expression patterns (see also Sect. III.A.2) or a selection of genes upregulated only in morphological structures that are specific to sexual development. The latter criterion requires the analysis of transcriptomes from fruiting body stages only, in contrast to the analyses of total mycelia containing fruiting bodies and vegetative mycelia in varying proportions. Both selection criteria were successfully used in subsequent analyses. Comparative expression analysis of a **cluster of differentially regulated polyketide biosynthesis genes** in *S. macrospora* and *N. crassa* was the basis for the identification of the *fbm1* gene, which was

shown to be required for timely fruiting body maturation in both fungi (Nowrousian 2009) (Fig. 7.1). Comparative analysis of microarray data from *S. macrospora* and *G. zeae* identified the histone chaperone gene *asf1*, and a deletion mutant of the *S. macrospora asf1* turned out to be sterile (Gesing et al. 2012). The second strategy for selection of candidate genes, namely, from expression data for specific structures, is also promising, as indicated by an RNA-seq analysis of young fruiting bodies that were isolated by **laser microdissection** from *S. macrospora* (Teichert et al. 2012). Among the genes that are strongly expressed or upregulated in these structures are the pheromone precursor genes *ppg1* and *ppg2* and the transcription factor genes *mcm1* and *pro44*, all of which were shown previously to be involved in fruiting body formation (Mayrhofer et al. 2006; Nolting and Pöggeler 2006; Nowrousian et al. 2012). Microscopic analysis of *ppg1* expression using promoter-*gfp* fusion constructs confirmed the organ-specific expression of this pheromone precursor gene and revealed a distinct expression pattern within hyphae of the outer layer of developing fruiting bodies (Teichert et al. 2012).

The identification of candidate genes can be significantly improved by the availability of large mutant collections (see Sect. II.B). The collections can be used to quickly screen mutants of differentially expressed genes for interesting phenotypes (e.g., in the case of genes with conserved expression patterns). For example, the corresponding knockout mutants from *N. crassa* were used to screen for sterile phenotypes among homologs to genes that were differentially regulated in *S. macrospora* or *P. confluens*, identifying a polyketide synthase and a putative transporter required for fertility in *N. crassa* (Nowrousian 2009; Gesing et al. 2013). The ability to generate and screen a large number of targeted deletion mutants can also be used to overcome less-than-stringent selection criteria for candidate genes. For example, in an analysis of *G. zeae*, deletion mutants were generated for 57 genes downregulated in a sterile mutant, and of these, 11 showed a development-related phenotype (Lee et al. 2010a).

IV. Other "-Omics" Approaches

Apart from transcriptomics, several other techniques have been established to survey gene expression across a large number of genes. Among these are proteomics and metabolomics, which analyze the protein content and metabolite content, respectively. Because of the chemical diversity of proteins and metabolites, these methods can usually capture only subsets of the complete proteome or metabolome of a cell or organ, in contrast to transcriptomics, by which, at least in principle, it is possible to analyze more or less all transcripts produced from a genome. Nevertheless, proteomics and metabolomics analyses can contribute valuable insights into genome activity because transcript levels often, but not always, correlate with protein levels, and when additional layers of regulation are in place, transcriptomics approaches can miss them. However, with respect to fruiting body formation, not many proteomics or metabolomics approaches have been conducted. Examples are the combined **transcriptomics/proteomics/metabolomics** study of the *csnE* mutant in *A. nidulans* described in Sect. III.B.3 (Nahlik et al. 2010); proteomics analyses of fruiting bodies from *T. borchii*, *Sparassis crispa*, and *Hericium erinaceum* (Pierleoni et al. 2004; Horie et al. 2008); and a study of cell wall proteins from different developmental stages of *Pleurotus tuber-regium* (Chen et al. 2012b). However, further combined "-omics" studies of different stages of fruiting body formation are needed to integrate such data into models of developmental regulation.

In addition to methods that measure gene expression at different levels, a number of techniques have been established that access the chromatin state (e.g., for the analysis of DNA binding sites of transcription factors and other DNA binding proteins, distribution of DNA modifications, or long-range interactions

between different genomic regions) (Lieberman-Aiden et al. 2009; Park 2009; Laird 2010; Nowrousian 2010). In combination with transcriptomics studies, these can be used, for example, to differentiate between direct and indirect targets of transcription factors by analyzing both the genome-wide transcript levels and DNA binding sites. This approach was, for example, applied to unravel the regulatory network of light signaling in *N. crassa* (Chen et al. 2009; Smith et al. 2010); however, similar studies have yet to be performed to analyze sexual development in filamentous fungi.

V. Conclusions and Outlook

Since the first high-throughput studies of fungi by the end of the 1990s, -omics in filamentous fungi has come a long way. Milestones along the way were the sequencing of fungal genomes (Sect. II), EST sequencing and microarrays for filamentous fungi (Sect. III), the development of efficient gene deletion strategies (Sect. II), and the application of NGS to address questions of fungal biology (Sects. II and III). All these advances have already been used to learn more about fruiting body differentiation, for example, to identify genes that are involved in this process and to understand the coordinated patterns of gene expression that underlie developmental events. It has become clear that morphological changes during development are mediated by drastic changes in gene expression patterns, and one of the future challenges will be to refine these analyses by improving the spatiotemporal resolution to analyze regulatory principles that govern the transitions from vegetative mycelium to mature fruiting body. In addition, it is important to study more species from different branches of the phylogenetic tree to establish both core developmental events as well as species-specific aspects of regulation. Currently, a 1,000-genome project for fungi is under way (http://1000.fungalgenomes.org), and among the sequenced genomes will be a number from species that produce fruiting bodies. A corresponding transcriptomics project to survey transcriptomes from different developmental stages for each species would greatly increase our understanding of the evolution of developmental transitions and the diversity of fungal multicellular structures. Thus, looking both deeper (in individual species) and wider (across the phylogenetic spectrum) can lead to a unified theory that explains common traits as well as differences in fruiting body development.

Acknowledgments I would like to thank Prof. Dr. Ulrich Kück for his support and Dr. Sandra Bloemendal for critical reading of the manuscript. Funding for my work comes from the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG, grant no. NO 407/2-1) and the Protein Research Department of Ruhr-University Bochum.

References

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252:1651–1656

Allen TD, Dawe AL, Nuss DL (2003) Use of cDNA microarrays to monitor transcriptional responses of the chestnut blight fungus *Cryphonectria parasitica* to infection by virulence-attenuating hypoviruses. Eukaryot Cell 2:1253–1265

Ballario P, Vittorioso P, Magrelli A, Talora C, Cabibbo A, Macino G (1996) White Collar-1, a central regulator of blue-light responses in Neurospora, is a zinc finger protein. EMBO J 15:1650–1657

Bidard F, Aït Benkhali J, Coppin E, Imbeaud S, Grognet P, Delacroix H, Debuchy R (2011) Genome-wide gene expression profiling of fertilization competent mycelium in opposite mating types in the heterothallic fungus *Podospora anserina*. PLoS One 6:e21476

Bidard F, Coppin E, Silar P (2012) The transcriptional response to the inactivation of the PaMpk1 and PaMpk2 MAP kinase pathways in *Podospora anserina*. Fungal Genet Biol 49:643–652

Bistis GN, Perkins DD, Read ND (2003) Different cell types in *Neurospora crassa*. Fungal Genet Newslett 50:17–19

Blumenstein A, Vienken K, Tasler R, Purschwitz J, Veith D, Frankenberg-Dinkel N, Fischer R (2005) The *Aspergillus nidulans* phytochrome FphA represses sexual development in red light. Curr Biol 15:1833–1838

Böhm J, Hoff B, O'Gorman CM, Wolfers S, Klix V, Binger D, Zadra I, Kürnsteiner H, Pöggeler S,

Dyer PS, Kück U (2013) Sexual reproduction and mating-type mediated strain development in the penicillin-producing fungus *Penicillium chrysogenum*. Proc Natl Acad Sci U S A 110:1476–1481

Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, Albert FW, Zeller U, Khaitovich P, Grützner F, Bergmann S, Nielsen R, Pääbo S, Kaessmann H (2011) The evolution of gene expression levels in mammalian organs. Nature 478:343–347

Breakspear A, Momany M (2007) The first fifty microarray studies in filamentous fungi. Microbiology 153:7–15

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. Proc Natl Acad Sci U S A 93:770–773

Busch S, Schwier EU, Nahlik K, Bayram Ö, Helmstaedt K, Draht OW, Krappmann S, Valerius O, Lipscomb WN, Braus GH (2007) An eight-subunit COP9 signalosome with an intact JAMM motif is required for fungal fruit body formation. Proc Natl Acad Sci U S A 104:8089–8094

Casselton LA (2008) Fungal sex genes-searching for the ancestors. Bioessays 30:711–714

Casselton LA, Kües U (1994) Mating-type genes in homobasidiomycetes. In: Wessels JGH, Meinhardt F (eds) The Mycota I growth, differentiation and sexuality. Springer, Berlin, pp 307–321

Chen CH, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ (2009) Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora. EMBO J 28:1029–1042

Chen CL, Kuo HC, Tung SY, Hsu PWC, Wang CL, Seibel C, Schmoll M, Chen RS, Wang TF (2012a) Blue light acts as a double-edged sword in regulating sexual development in *Hypocrea jecorina* (*Trichoderma reesei*). PLoS One 7:e44969

Chen L, Zhang BB, Cheung PC (2012b) Comparative proteomic analysis of mushroom cell wall proteins among the different developmental stages of *Pleurotus tuber-regium*. J Agric Food Chem 60:6173–6182

Cheng CK, Au CH, Wilke SK, Stajich JE, Zolan ME, Pukkila PJ, Kwan HS (2013) 5′-serial analysis of gene expression studies reveal a transcriptomic switch during fruiting body development in *Coprinopsis cinerea*. BMC Genomics 14:195

Chum WW, Ng KT, Shih RS, Au CH, Kwan HS (2008) Gene expression studies of the dikaryotic mycelium and primordium of *Lentinula edodes* by serial analysis of gene expression. Mycol Res 112:950–964

Chum WW, Kwan HS, Au CH, Kwok IS, Fung YW (2011) Cataloging and profiling genes expressed in *Lentinula edodes* fruiting body by massive cDNA pyrosequencing and LongSAGE. Fungal Genet Biol 48:359–369

Churchill GA (2002) Fundamentals of experimental design for cDNA microarrays. Nat Genet 32 (suppl):490–495

Clark NL, Aagarard JE, Swanson WJ (2006) Evolution of reproductive proteins from animals and plants. Reproduction 131:11–22

Claussen P (1912) Zur Entwicklungsgeschichte der Ascomyceten. Pyronema confluens Zeitschr f Bot 4:1–63

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC (2006) A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103:10352–10357

Crosthwaite SC, Dunlap JC, Loros JJ (1997) *Neurospora wc-1* and *wc-2*: transcription, photoresponses, and the origins of circadian rhythmicity. Science 276:763–769

Cuomo CA, Güldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer GJ, Münsterkötter M, Nelson D, O'donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317:1400–1402

Davis RL, deSerres D (1970) Genetic and microbial research techniques for *Neurospora crassa*. Methods Enzymol 27A:79–143

Dawe AL, Nuss DL (2001) Hypoviruses and chestnut blight: exploiting viruses to understand and modulate fungal pathogenesis. Annu Rev Genet 35:1–29

Deng F, Allen TD, Nuss DL (2007) Ste12 transcription factor homologue CpST12 is down-regulated by hypovirus infection and required for virulence and female fertility of the chestnut blight fungus *Cryphonectria parasitica*. Eukaryot Cell 6:235–244

DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278:680–686

Dettman JR, Jacobson DJ, Taylor JW (2003) Reproductive isolation and phylogenetic divergence in Neurospora: comparing methods of species recognition in a model eukaryote. Evolution 57:2721–2741

DiGuistini S, Liao N, Platt D, Robertson G, Seidel M, Chan S, Docking TR, Birol I, Holt R, Hirst M, Mardis E, Marra M, Hamelin R, Bohlmann J, Breuil C, Jones S (2009) De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. Genome Biol 10:R94

Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, McCluskey K, Plamann M, Galagan JE, Birren BW, Weiss RL, Townsend JP, Loros JJ, Nelson MA, Lambreghts R, Colot HV, Park G, Collopy P, Ringelberg C, Crew C, Litvinkova L, DeCaprio D, Hood HM, Curilla S, Shi M, Crawford M, Koerhsen M, Montgomery P, Larson L, Pearson M, Kasuga T, Tian C, Bastürkmen M, Altamirano L, Xu J (2007) Enabling a community to dissect an organism: overview of the Neurospora functional genomics project. Adv Genet 57:49–96

Dyer PS, O'Gorman CM (2011) A fungal sexual revolution: Aspergillus and Penicillium show the way. Curr Opin Microbiol 14:649–654

Ehrenreich A (2006) DNA microarray technology for the microbiologist: an overview. Appl Microbiol Biotechnol 73:255–273

Erdmann S, Freihorst D, Raudaskoski M, Schmidt-Heck W, Jung EM, Senftleben D, Kothe E (2012) Transcriptome and functional analysis of mating in the basidiomycete *Schizophyllum commune*. Eukaryot Cell 11:571–589

Fischer R (2008) Sex and poison in the dark. Science 320:1430–1431

Froehlich AC, Liu Y, Loros JJ, Dunlap JC (2002) White Collar-1, a circadian blue light photoreceptor, binding to the *frequency* promoter. Science 297:815–819

Gabella S, Abbà S, Duplessis S, Montanini B, Martin F, Bonfante P (2005) Transcript profiling reveals novel marker genes involved in fruiting body formation in *Tuber borchii*. Eukaryot Cell 4:1599–1602

Gaffoor I, Brown DW, Plattner R, Proctor RH, Qi W, Trail F (2005) Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). Eukaryot Cell 4:1926–1933

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap JC, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868

Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, Batzoglou S, Lee S-I, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Penalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115

Galgoczy DJ, Cassidy-Stone A, Llinas M, O'Rourke SM, Herskowitz IL, DeRisi J, Johnson AD (2004) Genomic dissection of the cell-type-specification circuit in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 101:18069–18074

Gesing S, Schindler D, Fränzel B, Wolters D, Nowrousian M (2012) The histone chaperone ASF1 is essential for sexual development in the filamentous fungus *Sordaria macrospora*. Mol Microbiol 84:748–765

Gesing S, Schindler D, Nowrousian M (2013) Suppression subtractive hybridization and comparative expression analysis to identify developmentally regulated genes in filamentous fungi. J Basic Microbiol 53:742–751

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274:546–567

Guitton AE, Berger F (2005) Control of reproduction by Polycomb Group complexes in animals and plants. Int J Dev Biol 49:707–716

Hallen HE, Trail F (2008) The L-type calcium ion channel cch1 affects ascospore discharge and mycelial growth in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). Eukaryot Cell 7:415–424

Hallen HE, Huebner M, Shiu SH, Güldener U, Trail F (2007) Gene expression shifts during perithecium development in *Gibberella zeae* (anamorph *Fusarium graminearum*), with particular emphasis on ion transport proteins. Fungal Genet Biol 44:1146–1156

Hammond TM, Rehard DG, Harris BC, Shiu PK (2011) Fine-scale mapping in *Neurospora crassa* by using genome-wide knockout strains. Mycologia 104:321–323

Han KH, Han KY, Yu JH, Chae KS, Jahng KY, Han DM (2001) The *nsdD* gene encodes a putative GATA-type transcription factor necessary for sexual development of *Aspergillus nidulans*. Mol Microbiol 41:299–309

Harding RW, Melles S (1983) Genetic analysis of phototropism of *Neurospora crassa* perithecial beaks using White Collar and albino mutants. Plant Physiol 72:996–1000

Heintzen C, Loros JJ, Dunlap JC (2001) The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting. Cell 104:453–464

Hoff B, Pöggeler S, Kück U (2008) Eighty years after its discovery, Fleming's Penicillium strain discloses the secret of its sex. Eukaryot Cell 7:465–470

Horie K, Rakwal R, Hirano M, Shibato J, Nam HW, Kim YS, Kouzuma Y, Agrawal GK, Masuo Y, Yonekura M (2008) Proteomics of two cultivated mushrooms *Sparassis crispa* and *Hericium erinaceum* provides insight into their numerous functional protein components and diversity. J Proteome Res 7:1819–1835

Horn BW, Ramirez-Prado JH, Carbone I (2009) Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*. Fungal Genet Biol 46:169–175

Jacobi V, Dufour J, Bouvet GG, Aoun M, Bernier L (2010) Identification of transcripts up-regulated in asexual and sexual fruiting bodies of the Dutch elm disease pathogen *Ophiostoma novo-ulmi*. Can J Bot 56:697–705

Joh JH, Lee JS, Kim KH, Jeong SJ, Youn WH, Kim NK, Son ES, Cho YS, Yoo YB, Lee CS, Kim BG (2007) Isolation of genes expressed during the developmental stages of the oyster mushroom, *Pleurotus ostreatus*, using expressed sequence tags. FEMS Microbiol Lett 276:19–25

Kämper J, Bölker M, Kahmann R (1994) Mating-type genes in heterobasidiomycetes. In: Wessels JGH, Meinhardt F (eds) The Mycota I growth, differentiation and sexuality. Springer, Berlin, pp 323–332

Karlsson M, Nygren K, Johannesson H (2008) The evolution of the pheromonal signal system and its potential role for reproductive isolation in heterothallic Neurospora. Mol Biol Evol 25:168–178

Keszthelyi A, Jeney A, Kerenyi Z, Mendes O, Waalwijk C, Hornok L (2006) Tagging target genes of the MAT-1–2-1 transcription factor in *Fusarium verticilloides* (*Gibberella fujikuroi* MP-A). Antonie Van Leeuwenhoek 91:373–391

Kim H, Nelson MA (2005) Molecular and functional analyses of *poi-2*, a novel gene highly expressed in sexual and perithecial tissues of *Neurospora crassa*. Eukaryot Cell 4:900–910

Kim H, Metzenberg RL, Nelson MA (2002) Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. Eukaryot Cell 1:987–999

King N (2004) The unicellular ancestry of animal development. Dev Cell 7:313–325

Klix V, Nowrousian M, Ringelberg C, Loros JJ, Dunlap JC, Pöggeler S (2010) Functional characterization of *MAT1–1*-specific mating-type genes in the homothallic ascomycete *Sordaria macrospora* provides new insights into essential and non-essential sexual regulators. Eukaryot Cell 9:894–905

Kück U, Hoff B (2010) New tools for the genetic manipulation of filamentous fungi. Appl Microbiol Biotechnol 86:51–62

Kück U, Pöggeler S, Nowrousian M, Nolting N, Engh I (2009) *Sordaria macrospora*, a model system for fungal development. In: Anke T, Weber D (eds) The Mycota XV, physiology and genetics. Springer, Berlin, pp 17–39

Kües U (2000) Life history and developmental processes in the basidiomycete *Coprinus cinereus*. Microbiol Mol Biol Rev 64:316–353

Lacourt I, Duplessis S, Abbà S, Bonfante P, Martin F (2002) Isolation and characterization of differentially expressed genes in the mycelium and fruit body of *Tuber borchii*. Appl Environ Microbiol 68:4574–4582

Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet 11:191–203

Lang BF, O'Kelly C, Nerad T, Gray MW, Burger G (2002) The closest unicellular relatives of animals. Curr Biol 12:1773–1778

Lee SH, Kim BG, Kim KJ, Lee JS, Yun DW, Hahn JH, Kim GH, Lee KH, Suh DS, Kwon ST, Lee CS, Yoo YB (2002) Comparative analysis of sequences expressed during the liquid-cultured mycelia and fruit body stages of *Pleurotus ostreatus*. Fungal Genet Biol 35:115–134

Lee SH, Lee S, Choi D, Lee YW, Yun SH (2006) Identification of the down-regulated genes in a *mat1-2*-deleted strain of *Gibberella zeae*, using cDNA subtraction and microarray analysis. Fungal Genet Biol 43:295–310

Lee J, Park C, Kim JC, Kim JE, Lee YW (2010a) Identification and functional characterization of genes involved in sexual reproduction of the ascomycete fungus *Gibberella zeae*. Biochem Biophys Res Commun 401:48–52

Lee SC, Ni M, Li W, Shertz C, Heitman J (2010b) The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev 74:298–340

Li D, Bobrowicz P, Wilkinson HH, Ebbole DJ (2005) A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in *Neurospora crassa*. Genetics 170:1091–1104

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289–293

Linden H, Macino G (1997) White Collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. EMBO J 16:98–109

Lysoe E, Pasquali M, Breakspear A, Kistler HC (2011) The transcription factor FgStuAp influences spore development, pathogenicity, and secondary metabolism in *Fusarium graminearum*. Mol Plant Microbe Interact 24:54–67

Martin F, Aerts A, Ahren D, Brun A, Danchin EGJ, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buee M, Brokstein P, Canback B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbe J, Lin YC, Legue V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Secq MPO-L, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kues U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouze P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452:88–92

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R, Porcel B, Rubini A, Amicucci A, Amselem J, Anthouard V, Arcioni S, Artiguenave F, Aury J-M, Ballario P, Bolchi A, Brenna A, Brun A, Buee M, Cantarel B, Chevalier G, Couloux A, Da Silva C, Denoeud F, Duplessis S, Ghignone S, Hilselberger B, Iotti M, Marcais B, Mello A, Miranda M, Pacioni G, Quesneville H, Riccioni C, Ruotolo R, Splivallo R, Stocchi V, Tisserant E, Viscomi AR, Zambonelli A, Zampieri E, Henrissat B, Lebrun M-H, Paolocci F, Bonfante P, Ottonello S, Wincker P (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature 464:1033–1039

Masloff S, Pöggeler S, Kück U (1999) The *pro1*+ gene from *Sordaria macrospora* encodes a C_6 zinc finger transcription factor required for fruiting body development. Genetics 152:191–199

Mayrhofer S, Weber JM, Pöggeler S (2006) Pheromones and pheromone receptors are required for proper sexual development in the homothallic ascomycete *Sordaria macrospora*. Genetics 172:1521–1533

McCluskey K, Wiest A (2011) The Fungal Genetics Stock Center in the context of a world wide community of ex situ fungal germplasm repositories. Fungal Biol Rev 25:143–150

McCluskey K, Wiest AE, Grigoriev IV, Lipzen A, Martin J, Schackwitz W, Baker SE (2011) Rediscovery by whole genome sequencing: classical mutations and genome polymorphisms in *Neurospora crassa*. G3 (Bethesda) 1:303–316

Miyazaki Y, Nakamura M, Babasaki K (2005) Molecular cloning of developmentally specific genes by representational difference analysis during the fruiting body formation in the basidiomycete *Lentinula edodes*. Fungal Genet Biol 42:493–505

Moore D (1998) Fungal morphogenesis. Cambridge University Press, Cambridge, UK

Moore-Landecker E (1979) Effect of light regimens and intensities on morphogenesis of the discomycete *Pyronema domesticum*. Mycologia 71:699–712

Morin E, Kohler A, Baker AR, Foulongne-Oriol M, Lombard V, Nagy LG, Ohm RA, Patyshakuliyeva A, Brun A, Aerts AL, Bailey AM, Billette C, Coutinho PM, Deakin G, Doddapaneni H, Floudas D, Grimwood J, Hildén K, Kües U, LaButti KM, Lapidus A, Lindquist EA, Lucas SM, Murat C, Riley RW, Salamov AA, Schmutz J, Subramanian V, Wösten HAB, Xu J, Eastwood DC, Foster GD, Sonnenberg ASM, Cullen D, de Vries RP, Lundell T, Hibbett DS, Henrissat B, Burton KS, Kerrigan RW, Challen MP, Grigoriev IV, Martin F (2012) Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. Proc Natl Acad Sci U S A 109:17501–17506

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320:1344–1349

Nahlik K, Dumkow M, Bayram Ö, Helmstaedt K, Busch S, Valerius O, Gerke J, Hoppert M, Schwier E, Opitz L, Westermann M, Grond S, Feussner K, Goebel C, Kaever A, Meinicke P, Feussner I, Braus GH (2010) The COP9 signalosome mediates transcriptional and metabolic response to hormones, oxidative stress protection and cell wall rearrangement during fungal development. Mol Microbiol 78:964–979

Nelson MA, Kang S, Braun EL, Crawford ME, Dolan PL, Leonard PM, Mitchell J, Armijo AM, Bean L, Blueyes E, Cushing T, Errett A, Fleharty M, Gorman M, Judson K, Miller R, Ortega J, Pavlova I, Perea J, Todisco S, Trujillo R, Valentine J, Wells A, Werner-Washburn M, Yazzie S, Natvig DO (1997) Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*. Fungal Genet Biol 21:348–363

Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101:12248–12253

Nolting N, Pöggeler S (2006) A MADS box protein interacts with a mating-type protein and is required for fruiting body development in the homothallic ascomycete *Sordaria macrospora*. Eukaryot Cell 5:1043–1056

Nowrousian M (2007) Of patterns and pathways: microarray technologies for the analysis of filamentous fungi. Fungal Biol Rev 21:171–178

Nowrousian M (2009) A novel polyketide biosynthesis gene cluster is involved in fruiting body morphogenesis in the filamentous fungi *Sordaria*

macrospora and *Neurospora crassa*. Curr Genet 55:185–198

Nowrousian M (2010) Next-generation sequencing techniques for eukaryotic microorganisms: sequencing-based solutions to biological problems. Eukaryot Cell 9:1300–1310

Nowrousian M, Cebula P (2005) The gene for a lectin-like protein is transcriptionally activated during sexual development, but is not essential for fruiting body formation in the filamentous fungus *Sordaria macrospora*. BMC Microbiol 5:64

Nowrousian M, Kück U (2006) Comparative gene expression analysis of fruiting body development in two filamentous fungi. FEMS Microbiol Lett 257:328–335

Nowrousian M, Dunlap JC, Nelson MA (2004) Functional genomics in fungi. In: Kück U (ed) The Mycota II. Springer, Berlin, pp 115–128

Nowrousian M, Ringelberg C, Dunlap JC, Loros JJ, Kück U (2005) Cross-species microarray hybridization to identify developmentally regulated genes in the filamentous fungus *Sordaria macrospora*. Mol Genet Genomics 273:137–149

Nowrousian M, Frank S, Koers S, Strauch P, Weitner T, Ringelberg C, Dunlap JC, Loros JJ, Kück U (2007a) The novel ER membrane protein PRO41 is essential for sexual development in the filamentous fungus *Sordaria macrospora*. Mol Microbiol 64:923–937

Nowrousian M, Piotrowski M, Kück U (2007b) Multiple layers of temporal and spatial control regulate accumulation of the fruiting body-specific protein APP in *Sordaria macrospora* and *Neurospora crassa*. Fungal Genet Biol 44:602–614

Nowrousian M, Stajich JE, Chu M, Engh I, Espagne E, Halliday K, Kamerewerd J, Kempken F, Knab B, Kuo HC, Osiewacz HD, Pöggeler S, Read ND, Seiler S, Smith KM, Zickler D, Kück U, Freitag M (2010) De novo assembly of a 40 Mb eukaryotic genome from short sequence reads: *Sordaria macrospora*, a model organism for fungal morphogenesis. PLoS Genet 6:e1000891

Nowrousian M, Teichert I, Masloff S, Kück U (2012) Whole-genome sequencing of *Sordaria macrospora* mutants identifies developmental genes. G3 (Bethesda) 2:261–270

Nygren K, Wallberg A, Samils N, Stajich JE, Townsend JP, Karlsson M, Johannesson H (2012) Analyses of expressed sequence tags in Neurospora reveal rapid evolution of genes associated with the early stages of sexual reproduction in fungi. BMC Evol Biol 12:229

O'Gorman CM, Fuller HT, Dyer PS (2009) Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. Nature 457:471–475

Oda K, Hasunuma K (1997) Genetic analysis of signal transduction through light-induced protein phosphorylation in *Neurospora crassa* perithecia. Mol Gen Genet 256:593–601

Ohm RA, de Jong JF, Lugones LG, Aerts A, Kothe E, Stajich JE, de Vries RP, Record E, Levasseur A, Baker SE, Bartholomew KA, Coutinho PM, Erdmann S, Fowler TJ, Gathman AC, Lombard V, Henrissat B, Knabe N, Kües U, Lilly WW, Lindquist E, Lucas S, Magnuson JK, Piumi F, Raudaskoski M, Salamov A, Schmutz J, Schwarze FW, van Kuyk PA, Horton JS, Grigoriev IV, Wösten HA (2010) Genome sequence of the model mushroom *Schizophyllum commune*. Nat Biotechnol 28:957–963

Ohm RA, de Jong JF, de Bekker C, Wösten HAB, Lugones LG (2011) Transcription factor genes of *Schizophyllum commune* involved in regulation of mushroom formation. Mol Microbiol 81:1433–1445

Ohm RA, Aerts D, Wösten HAB, Lugones LG (2013) The blue light receptor complex WC-1/2 of *Schizophyllum commune* is involved in mushroom formation and protection against phototoxicity. Environ Microbiol 15:943–955

Ospina-Giraldo MD, Collopy PD, Romaine CP, Royse DJ (2000) Classification of sequences expressed during the primordial and basidiome stages of cultivated mushroom *Agaricus bisporus*. Fungal Genet Biol 29:81–94

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis JP, Latgé JP, Denning DW, Dyer PS (2005) Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. Curr Biol 15:1242–1248

Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 10:669–680

Park SK, Penas MM, Ramirez L, Pisabarro AG (2006) Genetic linkage map and expression analysis of genes expressed in the lamellae of the edible basidiomycete *Pleurotus ostreatus*. Fungal Genet Biol 43:376–387

Park G, Servin J, Turner GE, Altamirano L, Colot HV, Collopy P, Litvinkova L, Li L, Jones CA, Diala FG, Dunlap JC, Borkovich KA (2011) Global analysis of serine-threonine protein kinase genes in *Neurospora crassa*. Eukaryot Cell 10:1553–1564

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JAE, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EGJ, Debets AJM, Dekker P, van Dijck PWM, van Dijk A, Dijkhuizen L, Driessen AJM, d'Enfert C, Geysens S, Goosen C, Groot GSP, de Groot PWJ, Guillemette T, Henrissat B, Herweijer M, van den Hombergh JPTW, van den Hondel CAMJJ, van der Heijden RTJM, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJEC, Meulenberg R, Menke H, Mortimer MA, Nielsen J, Oliver SG,

Olsthoorn M, Pal K, van Peij NNME, Ram AFJ, Rinas U, Roubos JA, Sagt CMJ, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van de Vondervoort PJJ, Wedler H, Wosten HAB, Zeng AP, van Ooyen AJJ, Visser J, Stam H (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25:221–231

Perkins JH (1969) Morphogenesis in *Schizophyllum commune*. I. Effects of white light. Plant Physiol 44:1706–1711

Pierleoni R, Buffalini M, Vallorani L, Guidi C, Zeppa S, Sacconi C, Pucci P, Amoresano A, Casbarra A, Stocchi V (2004) *Tuber borchii* fruit body: 2-dimensional profile and protein identification. Phytochemistry 65:813–820

Pöggeler S (1999) Phylogenetic relationships between mating-type sequences from homothallic and heterothallic ascomycetes. Curr Genet 36:222–231

Pöggeler S (2002) Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. Curr Genet 42:153–160

Pöggeler S, Nowrousian M, Kück U (2006a) Fruiting-body development in ascomycetes. In: Kües U, Fischer R (eds) The Mycota I. Springer, Berlin, pp 325–355

Pöggeler S, Nowrousian M, Ringelberg C, Loros JJ, Dunlap JC, Kück U (2006b) Microarray and real time PCR analyses reveal mating type-dependent gene expression in a homothallic fungus. Mol Genet Genomics 275:492–503

Pomraning KR, Smith KM, Freitag M (2011) Bulk segregant analysis followed by high-throughput sequencing reveals the Neurospora cell cycle gene, *ndc-1*, to be allelic with the gene for ornithine decarboxylase, *spe-1*. Eukaryot Cell 10:724–733

Qi W, Kwon C, Trail F (2006) Microarray analysis of transcript accumulation during perithecium development in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). Mol Genet Genomics 276:87–100

Romero IG, Ruvinsky I, Gilad Y (2012) Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet 13:505–516

Sakamoto Y, Nakade K, Sato T (2009) Characterization of the post-harvest changes in gene transcription in the gill of the *Lentinula edodes* fruiting body. Curr Genet 55:409–423

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Scheres B (2007) Stem-cell niches: nursery rhymes across kingdoms. Nat Rev Mol Cell Biol 8:345–354

Seidl V, Seibel C, Kubicek CP, Schmoll M (2009) Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci U S A 106:13909–13914

Sharon A, Yamaguchi K, Christiansen S, Horwitz BA, Yoder OC, Turgeon BG (1996) An asexual fungus has the potential for sexual development. Mol Gen Genet 251:60–68

Shendure J, Li H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145

Sikhakolli UR, López-Girldez F, Li N, Common R, Townsend JP, Trail F (2012) Transcriptome analyses during fruiting body formation in *Fusarium graminearum* and *Fusarium verticillioides* reflect species life history and ecology. Fungal Genet Biol 49:663–673

Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF, Thomas TL, Carrington JC, Stajich JE, Bell-Pedersen D, Brunner M, Freitag M (2010) Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for Neurospora White Collar complex. Eukaryot Cell 9:1549–1556

Son H, Seo YS, Min K, Park AR, Lee J, Jin JM, Lin Y, Cao P, Hong SY, Kim EK, Lee SH, Cho A, Lee S, Kim MG, Kim Y, Kim JE, Kim JC, Choi G, Yun SH, Lim JY, Kim M, Lee YH, Choi YD, Lee YW (2011) A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog 7:e1002310

Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discovery of conserved genetic modules. Science 302:249–255

Swanson WJ, Vacquier VD (2002) The rapid evolution of reproductive proteins. Nat Rev Genet 3:137–144

Talora C, Franchi L, Linden H, Ballario P, Macino G (1999) Role of a *white collar-1-white collar-2* complex in blue-light signal transduction. EMBO J 18:4961–4968

Teichert I, Wolff G, Kück U, Nowrousian M (2012) Combining laser microdissection and RNA-seq to chart the transcriptional landscape of fungal development. BMC Genomics 13:511

Tisserant E, Da Silva C, Kohler A, Morin E, Wincker P, Martin F (2011) Deep RNA sequencing improved the structural annotation of the *Tuber melanosporum* transcriptome. New Phytol 189:883–891

Traeger S, Altegoer F, Freitag M, Gabaldon T, Kempken F, Kumar F, Marcet-Houben M, Pöggeler S, Stajich JE, Nowrousian M (2013) The genome and development-dependent transcriptomes of *Pyronema confluens*: a window into fungal evolution. PLoS Genet 9:e1003820

Trail F, Xu JR, Miguel PS, Halgren RG, Corby Kistler H (2003) Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*). Fungal Genet Biol 38:187–197

van den Berg MA, Albang R, Albermann K, Badger JH, Daran JMM, Driessen AJ, Garcia-Estrada C, Fedorova ND, Harris DM, Heijne WHM, Joardar VW, Kiel JAK, Kovalchuk A, Martin JF, Nierman WC, Nijland JG, Pronk JT, Roubos JA, van der Klei IJ,

van Peij NNME, Veenhuis M, von Dohren H, Wagner C, Wortman J, Bovenberg RAL (2008) Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. Nat Biotechnol 26:1161–1168

Velculescu V, Zhang L, Vogelstein B, Kinzler K (1995) Serial analysis of gene expression. Science 270:484–487

Wahl R, Zahiri A, Kämper J (2010) The *Ustilago maydis* b mating type locus controls hyphal proliferation and expression of secreted virulence factors *in planta*. Mol Microbiol 75:208–220

Wang Z, Kin K, López-Girldez F, Johannesson H, Townsend JP (2012a) Sex-specific gene expression during asexual development of *Neurospora crassa*. Fungal Genet Biol 49:533–543

Wang Z, Lehr N, Trail F, Townsend JP (2012b) Differential impact of nutrition on developmental and metabolic gene expression during fruiting body development in *Neurospora crassa*. Fungal Genet Biol 49:405–413

Wessels JGH (1993) Fruiting in the higher fungi. Adv Microb Physiol 34:147–202

Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bähler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453:1239–1243

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285:901–906

Wösten HAB, Wessels JGH (2006) The emergence of fruiting bodies in basidiomycetes. In: Kües U, Fischer R (eds) The Mycota I. Springer, Berlin, pp 393–414

Yu GJ, Wang M, Huang J, Yin YL, Chen YJ, Jian S, Jin YX, Lan XQ, Wong BHC, Liang Y, Sun H (2012) Deep insight into the *Ganoderma lucidum* by comprehensive analysis of its transcriptome. PLoS One 7:e44301

Genomics for Biotechnology

8 Degradation and Modification of Plant Biomass by Fungi

Miia R. Mäkelä[1], Kristiina S. Hildén[1], Ronald P. de Vries[2]

CONTENTS

[1]Department of Food and Environmental Sciences, University of Helsinki, Viikinkaari 9, Helsinki FIN-00014, Finland; e-mail: miia.r.makela@helsinki.fi; kristiina.s.hilden@helsinki.fi
[2]CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, Utrecht 3584 CT, The Netherlands; e-mail: r.devries@cbs.knaw.nl

I. Introduction

Degradation of plant biomass has become one of the focal points of the **bio-based economy**, mainly because plant biomass is the most abundant renewable biomass on Earth that already is extensively used in agriculture and industry. The challenges for the bio-based economy are to use that part of plant biomass that is not used for the production of food or other valuable products and to convert these into biofuels, biochemicals, and bio-based materials. A major step in these processes is the depolymerization of the polymeric components of the plant cell wall, and for that, detailed understanding of the composition of plant biomass is required.

A. Plant Biomass Composition

Plant biomass consists mainly of polymeric components, such as polysaccharides, the aromatic polymer lignin, and proteins. In the polysaccharide fraction, a distinction can be made between cell wall polysaccharides (cellulose, hemicellulose, pectin) and storage polysaccharides (starch, inulin, different gums). Cell wall polysaccharides form a tight network together with lignin that provides strength for the plant, allowing it to grow upright, and is also a major defense against

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

microbial attack. The foundation of this network is formed by the cellulose microfibrils, which are long cables of parallel cellulose polymers. Hemicelluloses and pectin form noncovalent linkages with the microfibrils, creating the structural integrity required by the plant. Hemicelluloses and pectin can be covalently linked to each other by the formation of diferulic-ester bonds, providing additional structure to the network. This network is further strengthened by ester linkages between hemicellulose and lignin, either through ferulic acid or through (4-O-methyl-) glucuronic acid.

B. Cellulose

Cellulose is a linear polymer of β-1,4-linked D-glucopyranose residues, which are noncovalently organized into a cable-like structure called **microfibril** (Kolpak and Blackwell 1976). These microfibrils contain crystalline (highly ordered) and more amorphous regions, the latter of which are more susceptible to enzymatic attack. Degradation of cellulose requires the combined action of at least three types of enzymes: endoglucanases, cellobiohydrolases/exo-glucanases, and β-glucosidases (Lynd et al. 2002). Different families of these enzymes have been identified in fungi and other organisms, which are classified in the Carbohydrate Active enZyme database (CAZy; www.cazy.org) (Cantarel et al. 2009). CAZy families that contain fungal cellulose-active enzymes can be found in Table 8.1. In addition to the three main classes mentioned, other enzymes are involved in efficient cellulose degradation, such as the recently characterized cellulose oxidases, which were originally described as endoglucanases (GH61) (Quinlan et al. 2011).

Cellulose is the most abundant polymer in plant biomass and the most abundant organic compound on Earth. It is already of major importance for several industries, such as paper and pulp, but is currently of particular interest as a substrate for biofuels and biochemicals (Wilson 2009).

C. Hemicelluloses

Three main types of hemicellulose are found in plant cell walls: xylan, xyloglucan, and galactomannan. These are briefly described next.

1. Xylan

Xylan is the most abundant hemicellulose and is found in a wide range of plant species. It consists of a backbone of β-1,4-linked D-xylopyranose units that can be decorated with a range of different residues (de Vries and Visser 2001). The most abundant residues attached to the xylan backbone are L-arabinose and (4-O-methyl-)D-glucuronic acid, but the relative amount of these decorations differs strongly depending on the source. In cereals and grasses, L-arabinose is the most abundant decoration, and up to 50 % of the xylose residues can contain an L-arabinose side group. These xylans are therefore often referred to as arabinoxylans. In contrast, in hardwood (4-O-methyl)-D-glucuronic acid is the most abundant side group, and these xylans are therefore often named glucuronoxylans. Other side groups in xylan are α- and β-linked D-galactose and acetyl groups that are linked to the main-chain xylose residues and feruloyl and *p*-coumaroyl residues that are linked to the L-arabinose side groups (de Vries and Visser 2001). The high variation of the xylan structure makes it a very challenging substrate to degrade for fungi and requires a much larger number of enzymes (Table 8.1) (de Vries and Visser 2001). The main chain of xylan is degraded by endoxylanases and β-xylosidases, while the side groups are removed by arabinoxylan arabinofuranohydrolases, α-arabinofuranosidases, α-glucuronidases, α- and β-galactosidases, feruloyl esterases, and acetyl xylan esterases.

2. Xyloglucan

The backbone of xyloglucan consists of β-1,4-linked D-glucopyranose units, just like the backbone of cellulose. However, these glucose residues can be decorated with α-1,4-linked D-xylose residues in two distinct repeating patterns – XXGG or XXXG – in which two or three decorated glucose

Table 8.1. Fungal enzymatic activities involved in plant polysaccharide degradation with enzymes alphabetically based on the enzyme code

Enzyme	Activity	EC number[a]	Substrate	CAZy families
ABF	α-L-Arabinofuranosidase	3.2.1.55	Xyloglucan, xylan, pectin	GH51,54
ABN	Endoarabinanase	3.2.1.99	Pectin	GH43
ABX	Exoarabinanase	3.2.1.-	Pectin	GH93
AFC	α-L-Fucosidase	3.2.1.51, 3.2.1.63	Xyloglucan	GH29,95
AGD	α-1,4-D-Glucosidase	3.2.1.20	Starch	GH31
AGL	α-1,4-D-Galactosidase	3.2.1.22	Xyloglucan, xylan, galactomannan	GH27,36
AGU	α-Glucuronidase	3.2.1.131, 3.2.1.139, 3.2.2.-	Xylan	GH67,115
AMY	α-Amylase	3.2.1.1	Starch	GH13
AXE	Acetyl xylan esterase	3.1.1.72	Xylan	CE1
AXH	Arabinoxylan arabinofuranohydrolase	3.2.1.55	Xylan	GH62
AXL	α-D-Xylosidase	3.2.1.177	Xyloglucan	GH31
BGL	β-1,4-D-Glucosidase	3.2.1.21	Cellulose, xyloglucan	GH1,3,9
BXL	β-1,4-D-Xylosidase	3.2.1.37	Xylan, pectin	GH3,39,43
CBH	Cellobiohydrolase	3.2.1.91	Cellulose, xyloglucan	GH6,7,9
EGL	β-1,4-D-Endoglucanase	3.2.1.4	Cellulose, xyloglucan	GH5,7,9,12,44,45
FAE	Feruloyl esterase	3.1.1.73	Xylan, pectin	CE1
GAL	β-1,4-Endogalactanase	3.2.1.89	Pectin	GH53
GE	Glucuronoyl esterase	3.1.1.-	Xylan	CE15
GLA	Glucoamylase	3.2.1.3	Starch	GH15
GLN	β-1,6-Endogalactanase	3.2.1.164	Pectin	GH5,30
GMAE	Galactomannan acetyl esterase	3.1.1.-	Galactomannan	
INU	Endo-inulinase	3.2.1.7	Inulin	GH32
INX	Exo-inulinase	3.2.1.80	Inulin	GH32
LAC	β-1,4-D-Galactosidase	3.2.1.23	Xyloglucan, xylan, pectin, galactomannan	GH2,35,42
MAN	β-1,4-D-Endomannanase	3.2.1.78	Galactomannan	GH5,26
MND	β-1,4-D-Mannosidase	3.2.1.25	Galactomannan	GH2
PAE	Pectin acetyl esterase	3.2.1.-	Pectin	
PEL	Pectin lyase	4.2.2.10	Pectin	PL1
RGL	Rhamnogalacturonan lyase	4.2.2.-	Pectin	PL4,11
PGA	Endopolygalacturonase	3.2.1.15	Pectin	GH28
PGX	Exopolygalacturonase	3.2.1.82	Pectin	GH28
PLY	Pectate lyase	4.2.2.2	Pectin	PL1,3,9
PME	Pectin methyl esterase	3.1.1.11	Pectin	CE8
RGAE	Rhamnogalacturonan acetyl esterase	3.1.1.-	Pectin	CE12
RGX	Rhamnogalacturonan galaturonohydrolase/ exorhamnogalacturonase	3.2.1.-	Pectin	GH28
RHA	α-Rhamnosidase/rhamnogalacturonan rhamnohydrolase	3.2.1.40	Pectin	GH78
RHG	Rhamnogalacturonan hydrolase/ endorhamnogalacturonase	3.2.1.171	Pectin	GH28
SUC	Invertase/fructofuranosidase	3.2.1.26	Inulin	GH32
UGH	d-4,5 Unsaturated-glucuronyl hydrolase	3.2.1.-	Pectin	GH88
URH	Unsaturated rhamnogalacturonan hydrolase	3.2.1.172	Pectin	GH105
XEG	Xyloglucan-active β-1,4-D-endoglucanase	3.2.1.151	Xyloglucan	GH12,74
XGAE	Xyloglucan acetyl esterase	3.1.1.-	Xyloglucan	
XGH	xylogalacturonan hydrolase	3.2.1.-	Pectin	GH28
XLN	β-1,4-D-Endoxylanase	3.2.1.8	Xylan	GH10,11
XTG	β-1,3-Exogalactanase	3.2.1.145	Pectin	GH43

Source: Modified from Coutinho et al. (2009)

[a]If the last digit of the EC number is missing, this means that this enzyme does not have an EC number yet, but that based on its activity, we can predict in which group it will be classified

residues are separated by two or one undecorated glucose residues, respectively (Vincken et al. 1997). The xylose residues can be further decorated with L-arabinose, D-galactose, L-fucose, or acetyl residues, creating a further level of variation in the xyloglucan structure (Hisamatsu et al. 1991, 1992; Maruyama et al. 1996; Hantus et al. 1997; Huisman et al. 2000). Degradation of the backbone of xyloglucan requires similar activities as cellulose, and some cellulose-active enzymes are also active on xyloglucan (Table 8.1). However, xyloglucan-specific endoglucanases and exoglucanases have also been described (Pauly et al. 1999; Hasper et al. 2002; Master et al. 2008). The side groups are removed by α-xylosidases (Matsushita et al. 1985, 1987; Yoshikawa et al. 1993a, b), α-arabinofuranosidases (Saha 2000), β-galactosidases, α-fucosidases, and xyloglucan acetyl esterases (Table 8.1).

3. Galactomannan

Galactomannan is mainly found in the cell wall of woody plants and is the most abundant hemicellulose in softwood (Jacobs et al. 2002). Galactomannan consists of a backbone of β-1,4-linked D-mannose residues, which can be decorated with α-1,4-linked D-galactose residues and acetyl groups (Sjöström 1993). In some plants, the backbone is interrupted by D-glucose residues, in which case the polysaccharide is referred to as galactoglucomannan. Low levels of β-1,4-linked D-galactose residues have also been found attached to the mannose residues. The complete degradation of galacto(gluco)mannan involves the action of endo-mannases, β-mannosidases, α- and β-galactosidases, β-glucosidases, and galactomannan acetyl esterases (Gilbert et al. 2008) (Table 8.1). Galactomannans are also found as storage polysaccharides, such as locust bean gum and guar gum, but then do not contain glucose residues in the backbone (Srivastava and Kapoor 2005).

D. Pectin

Pectin is a highly variable polysaccharide that consists of four distinct substructures: homogalacturonan, xylogalacturonan, and rhamnogalacturonan I and II (RG-1 and RG-2; Carpita and Gibeaut 1993; Mohnen 1999; Ridley et al. 2001; Willats et al. 2001). Their characteristics and the enzymes that act on them are briefly discussed next.

1. Homogalacturonan and Xylogalacturonan

Homogalacturonan (also called the pectin smooth region) is a linear polymer of α-1,4-linked D-galacturonic acid residues, which can be methylated or acetylated. Xylogalacturonan is a modified form of homogalacturonan in which some of the galacturonic acid residues are decorated with β-1,3-linked D-xylose residues. The degradation of these polymers involves the action of endo- and exo-polygalacturonases, pectin and pectate lyases, pectin acetyl and pectin methyl esterases, and β-xylosidases (Table 8.1).

2. Rhamnogalacturonan I and II

Rhamnogalacturonan I consists of a backbone of alternating D-galacturonic acid and L-rhamnose residues that are linked through α-1,2- and α-1,4-linkages, respectively. It is also referred to as the hairy region of pectin because the rhamnose residues can be decorated with arabinan or arabinogalactan I side chains. The arabinan side chains consist of α-1,5-linked L-arabinose residues, which can be decorated with α-1,2- and α-1,3-linked L-arabinose residues (Schols and Voragen 1996). Arabinogalactan I consists of a β-1,4-linked D-galactose backbone that is decorated with α-1,3-linked L-arabinose residues (Mohnen 1999; Ridley et al. 2001). Both arabinan and arabinogalactan I can contain terminal feruloyl residues, and the galacturonic acid residues of the RG-I backbone can be acetylated. The degradation of RG-I requires the combined action of a large set of enzymes, such as rhamnogalacturonan hydrolases and lyases, α-rhamnosidases, endo- and exo-arabinanases, α-arabinofuranosidases, endogalactanases, β-galactosidases, feruloyl esterases, and rhamnogalacturonan acetyl esterases (Table 8.1).

Rhamnogalacturonan II (RG-II) is a minor component of pectin and consists of a modified homogalacturonan that contains side chains consisting of 11 uncommon sugars (O'Neill et al. 1996; Vidal et al. 2000). Because of its unusual structure, it is highly resistant to degradation by common pectinolytic enzymes.

E. Lignin

Lignin is an aromatic, amorphous, and heterogenic polymer present in all layers of woody cell walls. It strengthens the plant cell walls by adhesion to the layers of cellulose microfibrils and thereby also enhances water transport and provides resistance to microbial degradation (Higuchi 1997; Boerjan et al. 2003). Lignin is one of the most recalcitrant components of plant biomass. It comprises a three-dimensional network of phenylpropanoid precursors (i.e., coniferyl alcohol, sinapyl alcohol, and *p*-coumaryl alcohol), which are polymerized to guaiacyl-, syringyl-, and hydroxyphenyl-type lignin subunits by the action of laccases and peroxidases (Higuchi 2006). The subunits are joined together with a variety of bond types, mainly carbon-carbon and ether bonds, the β-aryl-ether (β-O-4) bonds being the most abundant (Adler 1977). Heterogeneous ring structures such as the dibenzodioxocin ring also occur (Brunow et al. 1998) (Fig. 8.1).

Lignin structure and composition vary significantly between plant species; therefore, the knowledge of the chemical structure of diverse plant lignins is still incomplete. Lignin precursors are phenolic compounds, but lignin is mostly nonphenolic because of the high occurrence of ether linkages in the lignin polymer (Ralph et al. 2004). In nonwoody plants, acetylated lignin structures are widespread (Martínez et al. 2008). The lignin content of lignocellulose differs depending on whether it is derived from softwood (gymnosperm), hardwood (angiosperm), or grasses (nonwoody or herbaceous plants). Softwood lignin (25–33 % of xylem dry weight) comprises mainly of coniferyl alcohol subunits, with the remainder mainly *p*-coumaryl alcohol units. Hardwood lignin consists of varying ratios of coniferyl and sinapyl alcohol units (Faix et al. 1985; Blanchette 1995; Burlat et al. 1997) and has a lower lignin content (20–25 % of xylem dry weight) (Sjöström and Westermark 1998). Herbaceous plants contain all the three monolignol subunits. The principal monomers are coniferyl and sinapyl alcohols, with modest amounts of *p*-coumaryl alcohol units (Buranov and Mazza 2008; Chundawat et al. 2011), which are attached to hemicelluloses via ester linkages (Eriksson et al. 1990; Hatfield et al. 1999). The lignin content in herbaceous plants varies greatly: Grasses generally contain 5–10 % lignin (van Soest 1982; Schaefer et al. 1985), whereas in crop straw the lignin content can be up to 23 % (Buranov and Mazza 2008).

Because of its complicated structure with nonhydrolyzable intermonomeric bonds, the biological degradation of lignin polymer requires unspecific, extracellular oxidative enzymes or low molecular weight chemical oxidizers, such as activated oxygen species, metal cations, and aromatic radicals (Evans et al. 1994). In addition, fungal metabolites like organic acids were shown to have a crucial role in lignin decay (Mäkelä et al. 2010). Efficient depolymerization of lignin is unique to the basidiomycete **white-rot fungi**, and it has been termed "enzymatic combustion" because the causative agents are oxidative extracellular enzymes (Kirk and Farrell 1987). The oxidoreductases generally recognized as lignin-modifying enzymes are class II heme peroxidases, that is, lignin peroxidases (LiPs), manganese peroxidases (MnPs), and versatile peroxidases (VPs), laccases, and H_2O_2-generating oxidases. In addition, enzymes responsible for quinone redox cycling are implicated to have a role in lignin degradation. Because of this variety of enzymes, degradation involves diverse biochemical reactions to release the basic units from the lignin polymer.

As **brown-rot fungi** lack lignin-modifying enzymes, they rapidly depolymerize cellulose in plant biomass and cause modification, mainly by demethoxylation, of the lignin fraction (Hatakka and Hammel 2010). As a result, the decayed wood remains brown and has lost its strengthening properties. In addition, the remaining modified lignin residues are highly resistant to further decay and thus contribute to the fixed carbon pool in humic soils, playing an important role

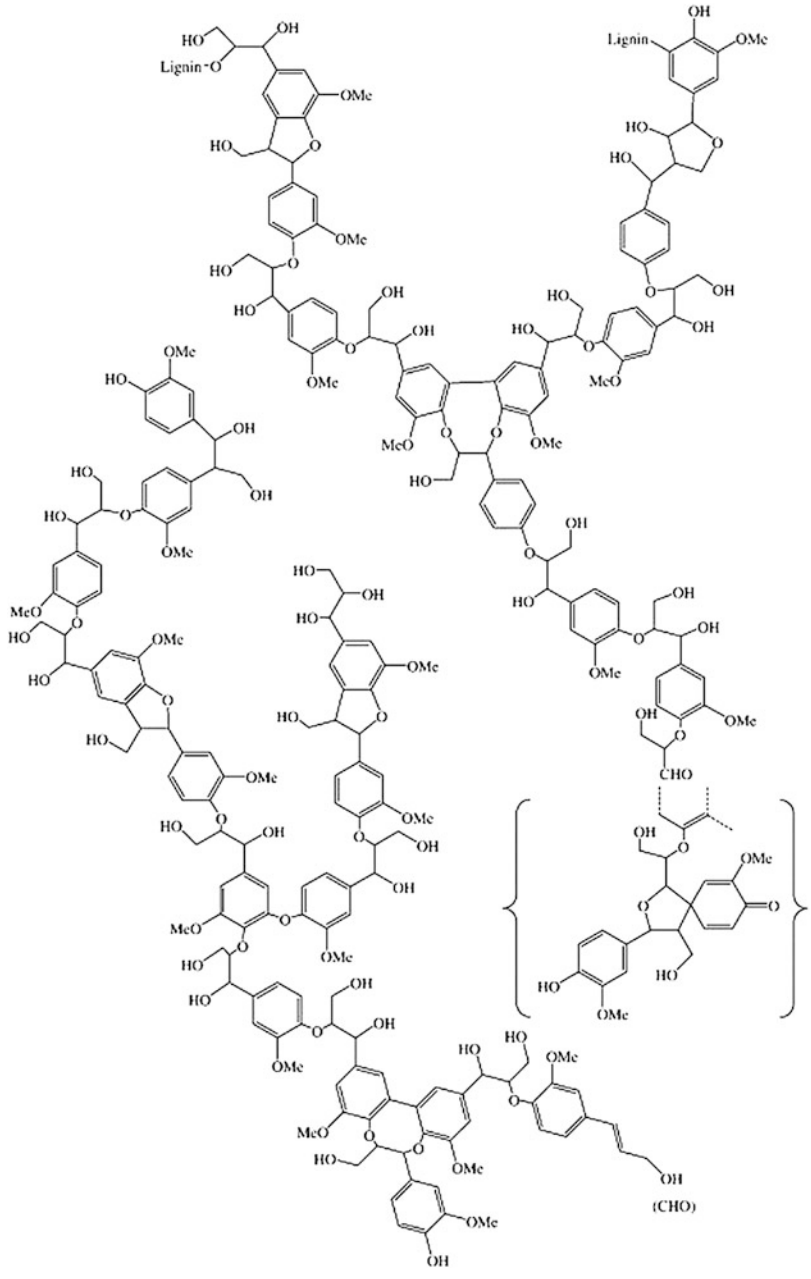


Fig. 8.1. Structural model of lignin (Brunow et al. 1998) (Reproduced with permission of the American Chemical Society)

in terrestrial carbon sequestration. Brown-rot decay of wood involves an initial nonenzymatic attack on the wood cell wall by hydroxyl radicals (·OH) generated extracellularly via a **Fenton reaction** ($Fe^{2+} + H_2O_2 + H^+ \rightarrow Fe^{3+} + \cdot OH + H_2O$) (Hammel et al. 2002). Systems releasing small molecular weight oxidants such as the free ·OH that randomly attack the substrate via Fenton-type chemistry have been thought to act in conjunction with common cellulases in lignocellulose degradation (Kerem et al. 1999; Jensen et al. 2001; Yelle et al. 2008).

II. Plant-Biomass-Modifying Enzymes of Ascomycetes

A. Current Status of Ascomycete Genomes

Within the **Ascomycota**, many of the most commonly used industrial fungi as well as major plant and (opportunistic) human pathogens are found. It is therefore not surprising that most of the initial set of fungal genome sequences were from ascomycete fungi. With the rapid reduction in cost and improvement of sequencing technologies, genome sequences are now available for a large set of fungi covering the entire fungal tree of life, with many more in progress. Genome projects have also moved from detailed analysis of a single genome to species comparisons as well as more habitat-related studies. The major revelation with respect to plant biomass degradation unveiled with the first selection of fungal genomes was that the gene sets related to this topic were much larger than expected. Especially for fungi like *Aspergillus niger*, for which many genes and enzymes involved in plant biomass degradation were already known (de Vries and Visser 2001), the much higher number revealed by the genome was surprising (Pel et al. 2007) and demonstrated that mainly the genes encoding highly produced enzymes under laboratory conditions were identified. Many studies into the **CAZomes** (subset of CAZy genes in a genome) of fungi were performed, revealing strong variations between species. For *A. niger*, such studies were taken to a deeper level by comparing genome content with data available from literature concerning the structures of plant poly- and oligosaccharides, which resulted in a graphical overview of the potential of *A. niger* to degrade these structures (Andersen et al. 2012). Using comparative genomics, this graphical view could be extended to other fungi, highlighting the differences between fungal species for degradation of specific plant biomass components.

In a recent study, the genomes of two species of the plant pathogen *Colletotrichum* were compared, and differences in CAZy content that correlate with the composition of the cell wall of their host plants were observed, such as an expansion of the pectin-related genes in the dicot-associated species (O'Connell et al. 2012). A larger study in which genomes of 18 Dothidiomycetes were compared also revealed significant differences in CAZy numbers in subgroups of these fungi, in particular for CAZy families related to cellulose degradation (Ohm et al. 2012). The expanded number of cellulose-related genes in another Dothidomycete, *Macrophomina phaseolina* (Islam et al. 2012), correlates well with its previously reported high cellulase activity (Kaur et al. 2012). In contrast, *Mycosphaerella graminicola* has a strongly reduced genome content related to plant cell wall degradation, which is supported by its poor growth on cell wall polysaccharides, such as xylan (Goodwin et al. 2011). Correlations between growth and CAZy content were also detected for two other plant pathogens, *Botryotinia fuckeliana* and *Sclerotinia sclerotiorum* (Amselem et al. 2011). While overall these two fungi have a reduced number of genes related to plant biomass degradation compared to other plant pathogens, their number of pectin-related genes is similar to the other species. Comparative **growth analysis** confirmed good growth in particular on pectin for these two species (Amselem et al. 2011).

Such correlations go even deeper, as became obvious when the pectinolytic genome content of several fungi was compared. The two main classes of pectinases are pectin hydrolases and pectin lyases (de Vries and Visser 2001). Pectin hydrolases are mainly active at acidic pH, while pectin lyases prefer neutral to alkaline pH. It was observed that *A. niger* contains higher numbers of pectin hydrolases in its

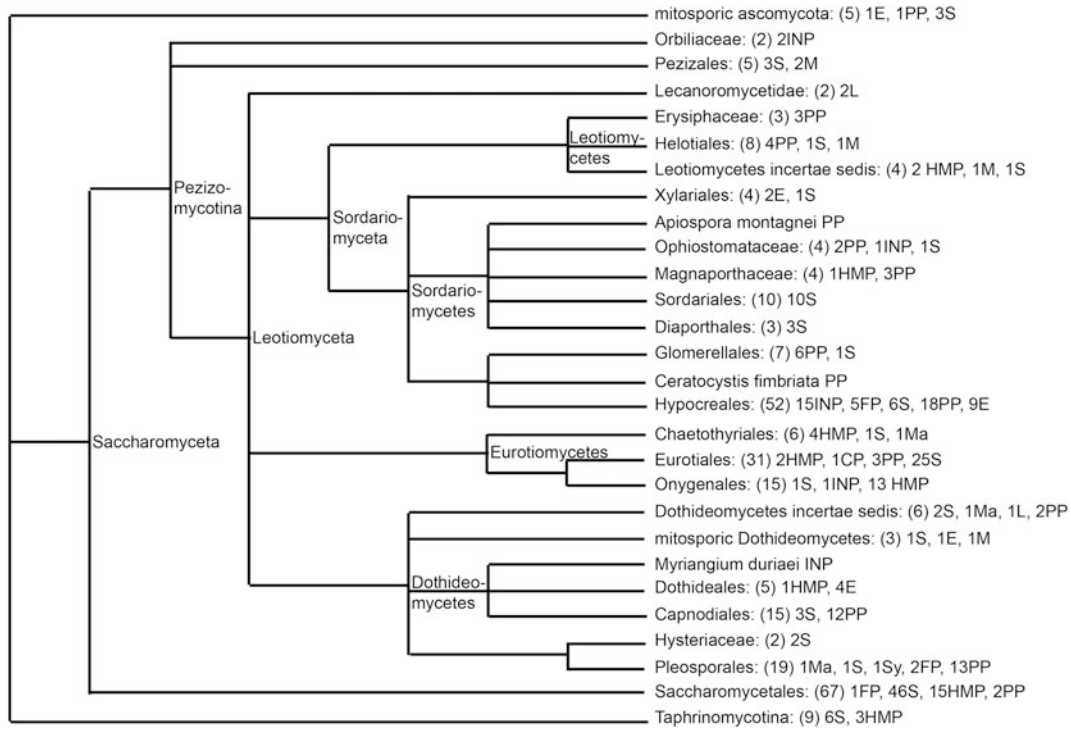


Fig. 8.2. Ascomycete phylogenetic tree indicating the current status of genomics. The number of species with a published or in progress genome sequence is given in parentheses. The number per lifestyle is indicated following this. *E* endophyte, *PP* plant pathogen, *S* saprobe, *INP* insect or nematode pathogen, *M* mycorrhizae, *L* lichen, *HMP* human or mammalian pathogen, *Ma* marine, *FP* fungal pathogen, *Sy* symbiont

genome than *Aspergillus oryzae* and *Aspergillus nidulans*, and it was suggested that this was because of the ability of *A. niger* to acidify its local environment, favoring the action of hydrolases over lyases (Pel et al. 2007).

More recently, a similar phenomenon was demonstrated for two thermophilic fungi, *Myceliophthora thermophila* and *Thielavia terrestris*. The *T. terrestris* genome contains seven pectin hydrolases and two pectin lyases; it is the opposite in the *M. thermophila* genome (Berka et al. 2011). This difference correlates perfectly with their ability to grow on pectin, which is best for *T. terrestris* at acidic pH, while *M. thermophila* grows better at alkaline pH (Berka et al. 2011).

Because of the complexity of the pectin structure, the differences in growth of fungi on this substrate can be more easily compared to fungal genome content. A recent study demonstrated that the presence or absence of specific pectinolytic genes correlated with the ability of the fungi to grow on pectin and to a lesser extent even on substructures of pectin (Benoit et al. 2012).

A survey of ascomycete genomes currently available and in progress resulted in the phylogenetic tree of Fig. 8.2. Although this will not remain a complete list of the available genomes as additional ones are continuously added, it shows the **genome coverage** of the Ascomycota. Behind the species, the lifestyle is indicated, which reflects the relevance of the species for genome sequencing. Many of the species are industrially relevant saprobes, plant pathogens, or human pathogens, although more recently the number of genomes from insect and nematode pathogens, mycorrhizae, and saprobes without industrial relevance are increasing. The tree also demonstrates that some ascomycete subgroups have highly conserved lifestyles, such as saprobes in the Sordariales and plant pathogens in the Glomerellales and Capnodiales.

The relevance of a species for industry or society has been a major driver in the selection of species for genome sequencing. However, the recently granted 1,000 Fungal Genomes Project (http://genome.jgi.doe.gov/pages/fungi-1000-projects-user-guidelines.jsf) aims to generate reference sequences for every clade in the Fungal Tree of Life, which will reduce the heavy bias observed in the fungal genomes sequenced at the time of writing.

Because of the rapidly increasing number of fungal genome sequences, several dedicated web servers for genome analysis have been developed. Some of these are hosted by the sequencing institutes, such as **MycoCosm** of the Joint Genome Institute (JGI) of the U.S. Department of Energy (http://genome.jgi.doe.gov/programs/fungi/index.jsf) and the **Fungal Genome Initiative** of the Broad Institute (http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative/fungal-genome-initiative). In addition, several other resources have been developed, such as **AspGD** (http://www.aspgd.org) for *Aspergillus*, and the yeast genome database (http://www.yeastgenome.org/) and **FungiDB** (http://fungidb.org/fungidb/) for a broader range of fungi. These resources facilitate comparative genomics studies and are therefore of major importance to the community. Tools at these servers include multigenome BLAST (Basic Local Alignment Search Tool), synteny analysis, and others, which are often beyond the reach of individual research groups.

B. Comparison of Saprobic and Human and Plant Pathogenic Fungi and Yeasts

Studies into plant biomass degradation abilities have mainly been performed using saprobic or plant pathogenic species because of the relevance of these enzymes to industry and agriculture. Table 8.2 shows a comparison of the genome content, with respect to plant biomass degradation, of seven filamentous saprobes, four filamentous plant pathogens, three saprobic yeasts, and three human pathogens. From this table, it is clear that the potential for plant biomass degradation is lowest in the saprobic yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pompe* and the human pathogens *Candida albicans* and *Coccidioides immitis*. A slightly higher potential is found in the lignocellulosic yeast *Pichia stipitis*, where the presence of an endoxylanase (GH10/11) (Jeffries et al. 2007) and an α-glucuronidase (GH67/115) (Ryabova et al. 2009) in particular stand out. However, in comparison to the filamentous saprobes and plant pathogens, the numbers of plant-biomass-related genes are almost negligible. In contrast to *C. albicans* and *C. immitis*, *Aspergillus fumigatus* has a broad range of genes related to plant biomass degradation (Nierman et al. 2005). This can be explained by the dual lifestyle of *A. fumigatus* as a saprobe in nature and as an opportunistic pathogen of humans. *A. fumigatus* is therefore more similar in genome content to other aspergilli rather than to the other human pathogens.

A comparison of the filamentous saprobes reveals large differences between these fungi. *A. niger*, *A. nidulans*, and *A. oryzae* contain the broadest set of genes related to plant biomass degradation (Coutinho et al. 2009), while *Penicillium rubens* (recently renamed from *P. chrysogenum*; Houbraken et al. 2011) is somewhat poorer in numbers. While *Podospora anserina* has a similar total number of putative plant-biomass-related genes as ***A. niger***, its distribution is more focused (Table 8.2). This correlates well with the biotope of these two fungi. *A. niger* is found globally in nearly all biotopes and is therefore exposed to a wide range of plant biomass structures. In contrast, *P. anserina* only grows as a late colonizer in herbivore dung, which means that it is mainly exposed to lignocellulose (a mixture of cellulose, xylan, and lignin). The genome of *P. anserina* is enriched in cellulases and xylanases (GH6/7, 10/11, 61, 67/115, CE1, 15) (and reduced in other enzyme classes) compared to the *A. niger* genome, which fits perfectly with its substrate.

Much lower numbers of putative plant-biomass-related genes were detected in ***Trichoderma reesei*** and ***Neurospora crassa*** (Table 8.2). *T. reesei* is a highly efficient cellulose degrader that was originally isolated from army tents in the Pacific (Schuster and Schmoll 2010; see Chap. 10 in this volume). Its efficiency is not caused by a large number of cellulases but by

Table 8.2. Comparison of CAZy numbers for 17 ascomycete fungi and yeasts

	Saprobic filamentous fungi							Plant pathogens				Saprobic yeasts			Human pathogens		
CAZy	AN	An	Ao	Pr	Pa	Tr	Nc	Fg	Bf	Lm	Mo	Ps	Sc	Sp	Ca	Af	Ci
GH1	3	3	3	3	1	2	1	3	3	3	2	0	0	0	0	5	0
GH2	10	6	7	6	7	7	5	10	2	6	6	2	0	0	1	6	1
GH3	20	17	23	17	11	13	9	22	16	13	19	5	1	0	3	18	5
GH5	15	10	13	13	15	11	7	15	15	15	13	4	3	5	5	13	2
GH6/7	5	4	4	3	10	3	8	3	3	6	9	0	0	0	0	5	0
GH10/11	5	5	8	4	14	5	6	8	5	5	10	1	0	0	0	7	0
GH12/74	3	5	4	3	3	3	2	5	4	3	4	0	0	0	0	6	0
GH13	13	18	17	16	9	5	10	8	10	7	10	5	13	9	4	16	5
GH15	2	2	3	3	3	2	2	3	4	3	2	1	2	1	1	5	1
GH26	3	1	1	1	1	0	1	0	2	1	0	0	0	0	0	0	0
GH27/36	7	7	6	3	3	10	1	5	6	6	6	0	1	0	0	8	1
GH28	9	21	21	5	0	4	2	6	18	6	3	0	0	1	0	14	0
GH29/95	3	3	3	1	0	4	0	3	2	2	5	0	0	0	0	2	0
GH31	10	7	10	11	5	3	5	8	4	8	5	3	4	1	3	6	2
GH32	2	6	4	7	0	0	1	5	1	2	5	0	2	1	0	5	0
GH35	4	5	7	2	1	1	2	3	4	4	0	0	0	0	0	5	0
GH43	15	10	20	14	10	2	7	17	4	11	19	0	0	0	0	18	1
GH45	1	0	0	0	2	1	1	1	2	0	1	0	0	0	0	1	0
GH51/54	3	5	4	4	1	2	2	3	4	3	4	0	0	0	0	3	0
GH53	1	2	1	1	1	0	1	1	2	1	1	0	0	0	0	1	0
GH61	9	7	8	4	33	3	14	15	9	20	17	0	0	0	0	7	0
GH62	2	1	2	1	2	1	0	1	1	1	3	0	0	0	0	2	0
GH67/115	2	1	5	1	4	2	2	3	1	3	1	1	0	0	0	2	0
GH78	8	8	9	5	1	1	0	7	8	1	1	0	0	0	0	6	0
GH88	2	1	3	0	0	0	0	1	1	1	1	0	0	0	0	2	0
GH93	2	0	3	2	3	0	2	2	1	2	1	0	0	0	0	3	0
GH105	3	2	4	2	0	1	1	3	1	3	3	0	0	0	0	2	0
PL1	8	6	12	5	4	0	1	9	6	9	2	0	0	0	0	6	0
PL3/9	6	0	4	1	2	0	1	8	2	6	1	0	0	0	0	4	0
PL4/11	5	2	4	3	1	0	1	3	0	4	1	0	0	0	0	4	0
CE1	3	3	5	2	14	3	7	5	3	3	10	1	0	1	1	5	1
CE8	3	3	5	2	1	0	1	6	5	3	1	0	0	0	0	6	0
CE12	2	2	4	2	1	0	1	3	3	3	2	0	0	0	0	3	0
CE15	0	0	0	1	3	1	1	0	0	2	1	0	0	0	0	1	0
Total	192	173	228	149	166	91	106	195	152	166	169	23	26	29	18	199	19

The enzymatic functions in the families are listed in Table 8.1

AN Aspergillus nidulans A4 (Coutinho et al. 2009), *An Aspergillus niger* CBS513.88 (Coutinho et al. 2009), *Ao Aspergillus oryzae* RIB4 (Coutinho et al. 2009), *Pr Penicillium rubens* Wisconsin 54-1255 (van den Berg et al. 2008), *Pa Podospora anserina* S mat+ (Espagne et al. 2008), *Tr Trichoderma reesei/Hypocrea jecorina* QM6A (Martinez et al. 2008), *Nc Neurospora crassa* OR74A (Galagan et al. 2003), *Fg Fusarium graminearum/Gibberella zeae* PH-1 (Cuomo et al. 2007), *Bf Botryotinia fuckeliana* A4 (Amselem et al. 2011), *Lm Leptosphaeria maculans* V23.1.3 (Rouxel et al. 2011), *Mo Magnaporthe oryzae* 70-15 (Dean et al. 2005), *Ps Pichia stipitis* CBS6054 (Jeffries et al. 2007), *Sc Saccharomyces cerevisiae* S288C (Cherry et al. 1997), *Sp Schizosaccharomyces pombe* 972 h (Iben et al. 2011), *Af Aspergillus fumigatus* Af293 (Nierman et al. 2005), *Ci Coccidioides immitis* RS (Sharpton et al. 2009)

high production of a fairly limited set of enzymes (Martinez et al. 2008). Although *T. reesei* is often considered the model fungus for cellulose degradation, the strategy by which it achieves this is in fact highly unusual. A comparative study of more than 80 fungi revealed that, for nearly all fungi, efficient degradation of a plant polysaccharide was related to an increased number of related enzymes (R.P. de Vries et al., unpublished results). Few fungi seem to have a similar strategy as *T. reesei* to produce high levels of a small number of enzymes. The *T. reesei* genome is poor in pectin- and inulin-related genes

(Table 8.2), further supporting the specialization toward cellulose.

The low number of CAZyme-encoding genes in the genome of *N. crassa* may be related to a similar strategy as observed for *T. reesei*. In addition, the distribution of the genes resembles that of *P. anserina*, suggesting a similar specialization to lignocellulose. *N. crassa* is most commonly found on burned wood, which is rich in lignocellulose. This correlates well with its genome content, in particular with respect to the low number of pectin-related genes and the higher number of lignocellulose-related genes (Table 8.2). The burning of the wood could even be considered a pretreatment that weakens the cell wall structure, allowing easier access for the fungus. Heat treatments are also commonly used in industrial pretreatments of plant biomass before enzymatic hydrolysis (Ruiz et al. 2008). The requirements of this ecological niche may then also explain why *N. crassa* contains a lower diversity of enzymes compared to saprobes such as *Aspergillus*, which typically colonize raw plant material.

The filamentous plant pathogens have a similar number of putative plant-biomass-related genes as the aspergilli, *P. rubens*, and *P. anserina*. The distribution of the genes over the various CAZy families differs significantly for the four plant pathogenic fungi, and correlations with their host plant can be observed. *Magnaporthe oryzae* is a pathogen of rice and other cereals, which contain cell walls consisting mainly of cellulose and xylan. The genome of *M. oryzae* contains an increased number of putative cellulases (GH6/7) and xylanases (GH10/11, CE1) compared to the other plant pathogens (Table 8.2). In contrast, *B. fuckeliana*, often found on fruits that are rich in pectin, contains a significantly higher number of putative pectinases from family GH28 and GH78, while it has a relatively lower number of cellulases. *Leptosphaeria maculans* is a pathogen of rapeseed, which is also rich in pectin and cellulose. In contrast to *B. fuckeliana*, *L. maculans* does not have an elevated number of pectin hydrolases but does contain a higher number of pectin/pectate lyases as well as higher numbers of cellulases (GH6 and GH61). This could fit with the previous observations and suggest that there is a (local) area of high pH where these fungi grow. *Fusarium graminearum* is mainly a cereal pathogen, but its genome appears less specialized to its host plant than observed in the other plant pathogens, as it also contains a significant number of pectinases. Whether this reflects a more recent specialization toward cereal crops is not clear at this point.

C. Transcriptomic and Proteomic Studies into Plant Biomass Degradation by Ascomycete Fungi

The availability of fungal genome sequences opened the door to **transcriptomic** and **proteomic** studies into the genes and enzymes related to plant biomass degradation. While previous studies already demonstrated that many of the genes involved in plant biomass degradation are specifically induced in the presence of these substrates (de Vries 2003), this could now be studied at the genome level, taking into account the much larger set of relevant genes revealed by genome sequencing. Induction of plant-biomass-related genes in the presence of plant biomass seems to be a universal phenomenon in ascomycete fungi, although significant differences in the specific gene/enzyme sets have been observed. Growth of *B. fuckeliana* on pectin resulted in the production of 13 pectin-related enzymes, while only 4 were detected during growth on sucrose (Shah et al. 2009), demonstrating that specific **induction** is the dominant driver of gene expression. Proteomics of extracellular samples from *T. reesei* grown on corn stover revealed cellulases and xylanases as the major enzymes produced (Nagendran et al. 2009), fitting well with this species specialization for these two polysaccharides.

Such a correlation between substrate and secreted enzymes was also reported for *Doratomyces stemonitis*, a coprophilous fungus isolated from koala feces. Growth of this fungus on plant-biomass-related substrates revealed secretion of enzymes related to cellulose and xylan degradation and to a lesser extent to mannan and pectin degradation, which matches the polysaccharides that are not or only partly degraded in the koala's gut (Peterson et al. 2011).

In some cases, the presence of a small number of highly dominant enzymes has been reported. Quantitative proteomics of *N. crassa* on crystalline

cellulose using 13 selected proteins (9 of which were predicted cellulases) demonstrated that 53 % of the total supernatant protein consisted of 2 predicted cellobiohydrolases (Phillips et al. 2011b). This finding is similar to what has been previously shown for *T. reesei* (Kolbe and Kubicek 1990) and supports the suggestion that *N. crassa* may use a similar strategy for plant biomass degradation as *T. reesei* (see Sect. II.B). An even more specialized enzyme set was observed for the beetle-symbiont *Grosmannia clavigera* during growth on pine sawdust. While practically no enzymes involved in lignocellulose degradation were produced, pectin-related enzymes were enriched in the secreted enzyme set (DiGuistini et al. 2011). Pectin degradation of the cell wall or the tracheid-bordered pit membranes could be required for the fungus to access the sapwood, in a similar manner as that described for the basidiomycete *Schizophyllum commune* (Ohm et al. 2010). In contrast, *Fusarium verticillioides* produced enzymes acting on cellulose, xylan, xyloglucan, and pectin during growth on wheat straw (Ravalason et al. 2012), which is a broader set than would be expected based on the composition of wheat straw (mainly cellulose and xylan). This suggests that the **regulation** of the genes encoding these enzymes is more coordinated than in other fungi.

The influence of the nature of the plant biomass on gene expression was particularly well illustrated for two thermophilic fungi. During growth on barley straw, *T. terrestris* and *M. thermophila* both highly expressed genes related to cellulose and xylan degradation, the two main components of barley straw (Berka et al. 2011), while pectinolytic genes were less expressed. However, this correlation and similarity between the species was less obvious for alfalfa straw. Relatively few xylan-related transcripts were observed in both species, but only *M. thermophila* showed high expression of pectin lyases (Berka et al. 2011). This correlates with its lyase-focused pectin-degrading strategy as mentioned previously (Sect. II.A), suggesting a significantly different physiology (in terms of pH) of these two species.

Differences in gene expression or enzyme production on pure plant biomass components and crude plant biomass have been described, demonstrating the complexity of **induction** of the plant-biomass-related genes. Extracellular proteomics of *A. niger* on xylose resulted mainly in xylan- and cellulose-related enzymes, of which the encoding genes were previously shown to be regulated by XlnR (see Section D) (Lu et al. 2010). In contrast, during growth on sugarcane bagasse, 58 % of the (hemi-) cellulolytic genes from *A. niger* were expressed, covering a much broader range of functions (de Souza et al. 2011). A similar observation was reported for *Penicillium purpurogenum*, which produced mainly xylan-related enzymes on acetylated xylan, while on sugar beet pulp, enzymes related to cellulose and pectin were detected (Navarrete et al. 2012). Transcriptome analysis of *Ascocoryne sarcoides* on cellulose and cellobiose revealed significant differences between these conditions with respect to the cellulolytic genes that were expressed, demonstrating that even such similar carbon sources can result in different enzyme sets (Gianoulis et al. 2012).

Differences in enzyme production and gene expression in time have also been reported, demonstrating the plasticity of the plant-biomass-degrading system of fungi. A time course analysis of the extracellular proteome of *A. nidulans* grown on sorghum stover revealed distinct sets of plant-biomass-related enzymes at different stages of cultivation (Saykhedkar et al. 2012), indicating that the fungus modifies its extracellular enzyme set to match the changes in the substrate composition and structure during its degradation. In the plant pathogen *Colletotrichum orbiculare*, expression of many genes related to plant biomass degradation was observed at 7 days past infection (dpi), while a significantly lower number of these genes were expressed at earlier time points (Gan et al. 2012). This indicates that at later stages of infection, a high demand for the nutrients released by plant biomass degradation occurs.

D. Regulation of Plant Biomass-Related Genes

To be able to respond to changes in the substrate, fungi need to be able to change their gene expression rapidly, which is largely mediated by a set of **transcriptional regulators**. Regulation of genes encoding plant-biomass-degrading enzymes has been mostly studied in three industrial species: *A. niger*, *A. oryzae*, and *T. reesei*. In *A. niger* and other aspergilli, **transcriptional activators**

related to degradation of xylan and cellulose (XlnR) (van Peij et al. 1998a, b), starch (AmyR) (Petersen et al. 1999), xylan and arabinan (AraR) (Battaglia et al. 2011a, b), pectin (RhaR) (Gruben 2012), and mannan (Ogawa et al. 2012) have been identified. XlnR is ubiquitous in filamentous ascomycetes, while the other regulators are less broadly distributed (R.B. Todd and R.P. de Vries unpublished results). *T. reesei* also contains an XlnR homolog (Xyr1) as well as a second xylan- and cellulose-related activator (ACEII) (Aro et al. 2001) and a xylanase/cellulase repressor (ACEI) (Aro et al. 2003).

The **carbon repressor protein** CreA (Cre1) is the major regulator repressing the expression of plant-biomass-related genes in the presence of sufficient amounts of monomeric carbon source (Ruijter and Visser 1997; de Vries et al. 1999). Homologs of this protein are found throughout the fungal kingdom (R.B. Todd and R.P. de Vries unpublished results). Recently, a *N. crassa* wild type and *cre-1* mutant were compared on crystalline cellulose using proteomics, demonstrating that the production of several cellulases and hemicellulases had increased in the mutant (Sun and Glass 2011). This study also revealed some novel genes that are under control of CRE-1 and affect cellulase activity. Recently, it was shown that a subset of the genes related to plant biomass degradation is expressed under starvation conditions without the need for these positive regulators (Delmas et al. 2012). The enzymes encoded by these genes are likely responsible for the liberation of the inducing compounds that results in activation of the specific regulators, which then results in higher expression of the related gene sets.

Finally, a role for photoreceptors in cellulase gene expression has been reported for *T. reesei* (Schmoll et al. 2005) and *N. crassa* (Schmoll et al. 2012), suggesting an interplay between environmental conditions and the presence of specific substrates.

III. Plant-Biomass-Modifying Enzymes of Basidiomycetes

A. Current Status of Basidiomycete Genomes

The genomes of over 60 species of **basidiomycete** fungi have been completely sequenced, and several hundred more are in progress. According to the current information, the size of basidiomycetous genomes varies from 9.8 to over 101 Mb. The most compact basidiomycete genome is from the xerotolerant mold-like fungus *Wallemia sebi* (Padamsee et al. 2012), whereas the largest genome is found in the poplar leaf rust fungus *Melampsora laricis-populina* (Duplessis et al. 2011).

Emphasizing the growing interest in plant biomass degradation as an abundant renewable material with vast biotechnological applicability, the first published basidiomycete genome was from ***Phanerochaete chrysosporium*** (Martinez et al. 2004), which is the most intensively studied model fungus for wood white-rot decay and lignin degradation. The genome of *P. chrysosporium* revealed an extensive collection of extracellular hydrolytic and oxidative enzymes, such as class II heme peroxidases and H_2O_2-generating oxidases, which play a role in lignocellulose degradation (Ruiz-Dueñas and Martínez 2009). The genome of *Ceriporiopsis subverispora* represented the first selectively lignin-degrading white-rot fungus with the ability to preferentially degrade lignin and hemicelluloses, leaving cellulose almost intact (Fernandez-Fueyo et al. 2012b). The other white-rot species sequenced are *Ganoderma lucidum* (Liu et al. 2012), which is one of the most famous medicinal fungi, and the forest pathogen and wood decayer *Heterobasidion irregulare* (Olson et al. 2012). Brown-rot fungi represent a different strategy of wood decay, being able to degrade most of the cellulose and hemicelluloses, leaving demethylated lignin. The first brown-rot basidiomycete sequenced was ***Postia placenta*** (Martinez et al. 2009), followed by *Serpula lacrymans* (Eastwood et al. 2011). Brown-rot species lack heme peroxidases but employ several genes encoding methanol oxidases (MOXs) that generate H_2O_2 required for cellulose depolymerization via Fenton reac-

tion (Daniel et al. 2007; Martinez et al. 2009). Other notable landmarks among the plant-biomass-degrading basidiomycetes have been the release of the genomes of *S. commune* (Ohm et al. 2010) and *Coprinopsis cinerea* (Stajich et al. 2010), which represent model systems, such as for mating-type gene function, regulation of multicellular development, and mushroom formation studies. In the order Agaricales, *Agaricus bisporus* was the first species sequenced as a model fungus for the adaptation, persistence, and growth in the humic-rich environment (Morin et al. 2012). Although the repertoire of *A. bisporus* lignin-modifying heme peroxidases was reduced when compared to other sequenced white-rot fungi, a significant expansion of heme-thiolate peroxidase– (HTP-) and β-etherase-encoding genes was found, suggesting adaption for decomposition of nonwoody plant biomass and humic substances in soil.

Hydrolytic enzyme-encoding genes and their expression in *A. bisporus* are significantly different from the taxonomically related ectomycorrhizal (ECM) basidiomycete *Laccaria bicolor*, which was the first symbiotic fungus sequenced (Martin et al. 2008). In contrast to saprobic fungi, genome analyses of ECM fungi have reduced sets of genes encoding both hydrolytic and oxidative plant-biomass-degrading enzymes (Martin et al. 2008; Nagendran et al. 2009). However, the transcriptome of the ECM fungus *Paxillus involutus* revealed upregulated expression of the set of enzymes involved in the oxidative degradation of wood by brown-rot fungi, whereas only two plant cell-wall-degrading CAZymes were upregulated (Rineau et al. 2012). The genetic potential of ECM as plant biomass decomposers is still an intriguing question.

The genome of the corn smut *Ustilago maydis*, which is a model organism in plant-microbe interactions, showed only a few genes encoding plant cell-wall-degrading enzymes, such as polysaccharide hydrolases, polysaccharide lyases, and pectin esterases, which is in contrast to the genomes of the more aggressive pathogenic fungi (Kämper et al. 2006). The minimal set of hydrolytic enzymes found in *U. maydis* correlates with its biotrophic lifestyle, in which damage to the host should be minimized to avoid plant defense responses. However, its genome sequence revealed an unexpected set of small genes with unknown function that might play a role in *U. maydis* virulence.

B. Basidiomycete Enzymes Involved in Lignin Modification

Fungal degradation of lignin is promoted by **class II heme peroxidases**, which are classified into three different groups: **manganese peroxidases** (EC 1.11.1.13), **lignin peroxidases** (EC 1.11.1.14), and **versatile peroxidases** (EC 1.11.1.16). In addition, lignin-degrading fungi secrete various oxidases, which produce the H_2O_2 required by peroxidases.

1. Class II Heme Peroxidases

The so-called **lignin-modifying peroxidases** (LiP, MnP, VP) are exclusively found in white-rot fungi, with the exception of the mushroom forming *S. commune*, regarded as a white rotter, with a limited lignin-degrading capacity and no lignin-modifying peroxidases detected in its genome. *Trametes versicolor* has the highest number of lignin-modifying peroxidase-encoding genes of any genome-sequenced basidiomycete so far. According to the available basidiomycete genome sequences, the duplication rate of class II peroxidase-encoding genes is apparently more rapid in white-rot fungi when compared to other plant-biomass-converting fungal species (Floudas et al. 2012). Supporting their different strategy to degrade plant biomass, lignin-modifying peroxidases are absent in the genomes of brown-rot fungi, the coprophilic fungus *C. cinerea*, and the EMC fungus *L. bicolor* (Table 8.3).

The crystal structures of *P. chrysosporium* MnP1(Sundaramoorthy et al. 1994) and LiP-H8 (Piontek et al. 1993; Poulos et al. 1993), and *Pleurotus eryngii* VPL (Pérez-Boada et al. 2005) have revealed functional differences between the lignin-modifying peroxidases sharing a high degree of primary sequence homology. Both MnP and VP have three conserved acidic amino acid residues (Glu35, Glu39, and Asp175 of *P. chrysosporium* MnP1), which together with one of the heme propionates are involved in Mn^{2+} binding (Sundaramoorthy et al. 1994; Kusters-van Someren et al. 1995). The conserved tryptophan residue (Trp171 of *P. chrysosporium* LiP-H8) is the most important residue in veratryl alcohol oxidation in LiPs and VPs (Piontek et al. 2001; Ruiz-Dueñas et al. 2009).

Table 8.3. Number of putative oxidoreductase gene models in the genomes of plant-biomass-degrading basidiomycete fungi

	White rot											S	Brown rot							ECM
	Pc	Ab	Cs	Ad	Ps	Fm	Ds	Tv	Sh	Hi	Sc	Cc	Pp	Fp	Cp	Gt	Sl	Wc	Dsp	Lb
Class II heme peroxidases																				
LiP	10	0	2	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0
MnP (total)	5	2	13	5	10	16	9	13	5	6	0	0	0	0	0	0	0	0	0	0
Short	0	2	1	0	2	3	5	13	0	na	0	0	0	0	0	0	0	0	0	0
Long	5	0	7	0	5	11	0	0	0	na	0	0	0	0	0	0	0	0	0	0
Extra Long	0	0	5	0	3	0	4	0	0	na	0	0	0	0	0	0	0	0	0	0
Atypical	0	0	0	5	0	2	0	0	5	1	0	0	0	0	0	0	0	0	0	0
VP	0	0	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0
DyP	0	0	0	11	5	3	1	2	2	1	0	4	2	0	0	0	0	0	0	2
HTP	3	24	9	16	8	4	4	3	10	5	3	8	5	4	2	6	3	5	6	5
GMC oxidoreductases										34										
AAO	3	10	5	0	4	2	9	3	14	na	1	18/27	3	1	0	2	0/6	0	0	4
MOX	1	na	4	3	3	2	4	4	7	na	3	0	1	4	2	1	1	4	1	0
POX	1	na	na	2	0	0	0	2	0	na	1	0	0	0	0	0	0	0	0	0
GOX	1	na	2	0	0	0	0	0	0	na	4	1	0	0	0	0	0	0	0	3
Copper-radical oxidases																				
GLX	1	3	1	2	3	0	5	5	3	0	0/2	0	0	0	0	0	0	0	0	0/5
Other copper-radical oxidases	6	6	4	7	6	4	4	4	5	5	2	6	3	4	6	2	3	4	3	11
Multicopper oxidases (total)	5	na	9	10	13	11	13	10	20	17	6	17	5	7	8	4	6	5	5	11
Laccase	0	11	7	7	12	10	11	7	15	5	2	17	3	5	6	3	4	3	0	9
Iron reduction																				
CDH	1	1	1	1	1	1	1	1	1	1	1	1	0	0	2	1	2	0	0	0
QR	4	4	1	4	3	3	1	1	1	2	4	3	1	1	2	3	2	1	1	2
P450	149	109	222	249	144	130	187	190	215	144	115	139	250	190	238	130	164	206	126	101

Based on Martin et al. (2008), Martinez et al. (2009), Ohm et al. (2010), Fernandez-Fueyo et al. (2012a), Floudas et al. (2012), Morin et al. (2012), Olson et al. (2012)

S non-wood-decaying saprotroph, *ECM* ectomycorrhiza, *Pc Phanerochaete chrysosporium*, *Ab Agaricus bisporus*, *Cs Ceriporiopsis subvermispora*, *Ad Auricularia delicata*, *Ps Punctularia strigosozonata*, *Fm Fomitiporia mediterranea*, *Ds Dichomitus squalens*, *Tv Trametes versicolor*, *Sh Stereum hirsutum*, *Hi Heterobasidion irregulare*, *Sc Schizophyllum commune*, *Cc Coprinopsis cinerea*, *Pp Postia placenta*, *Fp Fomitopsis pinicola*, *Cp Coniophora puteana*, *Gt Gloeophyllum trabeum*, *Sl Serpula lacrymans*, *Wc Wolfiporia cocos*, *Dsp Dacryopinax* sp., *Lb Laccaria bicolor*, *LiP* lignin peroxidase, *MnP* manganese peroxidase, *VP* versatile peroxidase, *DyP* dye-decolorizing peroxidase, *HTP* heme-thiolate peroxidase, *AAO* aryl alcohol oxidase, *MOX* methanol oxidase, *POX* pyranose oxidase, *GOX* glucose oxidase, *GLX* glyoxal oxidase, *CDH* cellobiose dehydrogenase, *QR* quinone reductase, *P450* cytochrome P450 monooxygenase, *na* not annotated

MnPs demonstrate diverse substrate affinities and biochemical properties as biocatalysts in the degradation of wood constituents and synthetic lignin by catalyzing the specific oxidation of Mn^{2+} to Mn^{3+} in the presence of H_2O_2 (Hammel and Cullen 2008; Hatakka and Hammel 2010; Lundell et al. 2010). Mn^{3+} ions are stabilized in chelated form by dicarboxylic acids, such as **oxalic acid** to perform oxidative reactions that yield organic radicals from several phenolic substrates, carboxylic acids, and unsaturated lipids. In addition, MnP-mediated **lipid peroxidation** reactions enable the oxidization of the more recalcitrant nonphenolic lignin substructures that compose the majority of the polymer (Jensen et al. 1996; Kapich et al. 1999).

Based on gene structure, MnPs can be divided into four groups by the length of their amino acid sequence and the number of conserved Mn^{2+}-binding amino acid residues: short, long, extra-long, and atypical MnPs (Floudas et al. 2012). All typical MnPs harness experimentally confirmed or putative Mn^{2+} oxidation sites formed by the three conserved acidic amino acid residues but lack the conserved tryptophan residue present in LiPs and VPs. Short MnPs are widely distributed among white-rot fungi (Table 8.3), whereas long MnPs form a monophyletic clade and have primarily been detected in the orders Polyporales, Corticiales, and Hymenochaetales (Table 8.3) (Hildén et al. 2005; Sakamoto et al. 2009). Genes encoding extra-long MnPs, which have a polar C-terminal metal-binding tail providing stability to the enzyme, have so far only been found in *C. subvermispora*, *Dichomitus squalens*, and *Punctularia strigoso zonata* (Table 8.3) (Li et al. 2001). MnPs, which have only one or two acidic amino acids at the Mn^{2+} oxidation site, have been designated as atypical MnPs and the corresponding gene models have been found in the genome of *Auricularia delicata*, *Fomitiporia mediterranea*, and *Stereum hirsutum* (Table 8.3). However, biochemical data are currently lacking to confirm their function as Mn^{2+}-oxidizing enzymes.

LiP catalyzes oxidation of a variety of reducing substrates, such as phenols, aromatic amines, aromatic ethers, and polycyclic aromatics (Kirk and Cullen 1998; Hammel and Cullen 2008). Active LiP isozymes and LiP-encoding gene models have been described from only a few genera of white-rot fungi, including *Phanerochaete* (Glenn et al. 1983; Tien and Kirk 1984; Martinez et al. 2004; Sugiura et al. 2009); *Phlebia* (Niku-Paavola et al. 1988; Vares et al. 1995; Hildén et al. 2006); *Trametes* (Johansson and Nyman 1996; Floudas et al. 2012); and *Bjerkandera* (Kimura et al. 1991; Heinfling et al. 1998a; ten Have et al. 1998). Comparative genomic analysis of two *Phanerochaete* species has shown that *P. carnosa* has only four LiP-encoding genes, while *P. chrysosporium* has ten. However, in both species LiP-encoding genes are tightly clustered, suggesting local gene duplication (Stewart and Cullen 1999; Suzuki et al. 2012). In addition, two *lip* gene models for *Pycnoporus cinnabarinus*, one for *Pleurotus sapidus*, and a putative *lip* gene for genera *Ganoderma* have been reported (Chen et al. 2012). According to the phylogenetic analyses and catalytic properties of heterologously produced enzymes, *C. subvermispora* has been suggested to have two transition types of VP-LiP enzymes (Fernandez-Fueyo et al. 2012b). A unique type of LiP with catalytic tyrosine is present in *Trametes cervina* (Miki et al. 2010).

VP combines the catalytic properties of both MnP and LiP. Accordingly, all the known VPs bear the three conserved acidic amino acid residues involved in Mn^{2+} binding together with the solvent-exposed tryptophan residue (Ruiz-Dueñas et al. 2009). The first VPs were characterized from the genera *Pleurotus* and *Bjerkandera* (Martínez et al. 1996; Heinfling et al. 1998b), while whole-genome sequences have revealed their presence also in the genomes of *D. squalens* and *T. versicolor* (Table 8.3).

In addition to the lignin-modifying peroxidases described, multiple ECM taxa harbor class II peroxidases (Bödeker et al. 2009) that phylogenetically cluster with an *L. bicolor* peroxidase that lacks the conserved Mn^{2+} oxidation site. The only ECM sequence that is placed within one of the lignin-modifying peroxidase groups with strong support is the partial sequence from *Gomphus clavatus* (Morgenstern et al. 2010).

The growing number of basidiomycete genomes has revealed novel peroxidase superfamilies, for instance, **dye-decolorizing peroxidases** (DyPs; EC 1.11.1.19) and **heme-thiolate peroxidases** that were previously characterized from only a few species (Sugano 2009; Hofrichter et al. 2010). It has been suggested that DyPs have hydrolase or

oxygenase catalytic activities as well as typical peroxidase activities. However, there are only a few studies reported on their structure and function. DyPs have been shown to degrade lignin model compounds (Liers et al. 2010), and putative gene models are common in white-rot fungal species but are absent in most brown-rot species (Table 8.3). However, the brown-rot fungus *P. placenta*, the coprophilic fungus *C. cinerea*, and the EMC fungus *L. bicolor* harbor putative *dyp* gene models in their genome.

The HTPs compose another new broadly distributed peroxidase superfamily, including chloroperoxidases (EC 1.11.1.10) and peroxygenases (EC 1.11.2.1), that are able to oxidize aromatic and aliphatic compounds. So far, HTP gene models have been detected in all the genomes of plant-biomass-degrading basidiomycetes with the largest number present in *A. bisporus* (Table 8.3). Whether HTPs have a role in the conversion of plant biomass is still ambiguous, although *Agrocybe aegerita* HTP has been shown to hydroxylate aromatic rings (Hofrichter and Ullrich 2006).

2. Multicopper Oxidases

Laccases (EC 1.10.3.2, *p*-diphenol:oxygen oxidoreductases) are four-copper-containing metalloenzymes that belong to the large and diverse protein superfamily of **multicopper oxidases** (MCOs) (Giardina et al. 2010; Kües and Rühl 2011). They catalyze the oxidation of a variety of phenolic and low-redox potential compounds with the concomitant reduction of molecular oxygen to water. Laccases are widespread in the fungal kingdom, particularly in basidiomycetes, including plant-biomass-degrading species (Table 8.3). Laccase genes often occur in large families coding for different isoenzymes, and paralogous laccase genes can be differentially regulated (Piscitelli et al. 2011). Laccase and MCO proteins display a highly conserved structure with four well-conserved amino acid motifs, which coordinate the copper atoms crucial to activity and also residues essential for the maintenance of protein conformational fold (Hoegger et al. 2006; Kües and Rühl 2011).

Despite their omnipresence, the role of laccases in lignocellulose modification is still ambiguous. While numerous studies demonstrated laccase-catalyzed oxidation of phenolic and nonphenolic lignin model compounds, particularly in the presence of low molecular weight mediators, sensu stricto laccase-encoding genes are not present in the genomes of the white-rot fungi *P. chrysosporium* (Martinez et al. 2004) or *P. carnosa* (MacDonald et al. 2011), implicating that laccase may not be essential for the white-rot decay of wood. However, both species have several MCO-encoding genes, and recent work with *Phanerochaete flavido-alba* has demonstrated that some members of the ferroxidase/laccase group may efficiently oxidize laccase substrates (Rodríguez-Rincón et al. 2010). The MCOs have undergone extensive gene duplication in the genomes of *C. cinerea*, *H. irregulare*, and *L. bicolor*, and thus gene expansions are associated not only with white-rot species (Table 8.3).

3. H_2O_2-Generating Oxidoreductases

For complete decomposition and partial mineralization of plant biomass, additional fungal enzymes are required. H_2O_2 plays a central role in lignocellulose degradation as the precursor of highly reactive hydroxyl radicals (·OH), which are able to depolymerize lignin and polysaccharides. Within the **H_2O_2-generating oxidoreductases**, the **glucose-methanol-choline** (GMC) **superfamily** includes aryl alcohol oxidase (AAO; EC 1.1.3.7); MOX (EC 1.1.3.13); pyranose-2-oxidase (POX; EC 1.1.3.10); and glucose oxidase (GOX; EC 1.1.3.4). The **copper radical oxidase** (CRO) gene group contains glyoxal oxidase (GLX; EC 1.2.3.-) and other CROs.

AAO exhibits a broad substrate specificity catalyzing the oxidation of lignin-derived compounds, such as phenolic aromatic aldehydes and acids (Kirk and Farrell 1987; Shimada and Higuchi 1991), and other aromatic fungal metabolites (Hernndez-Ortega et al. 2012) into their corresponding aldehydes with the concomitant reduction of O_2 to H_2O_2 (Varela et al. 2001). In *Pleurotus* and *Bjerkandera* species AAO has been suggested to be the main peroxide-producing enzyme acting synergistically with VP (Guillén

et al. 1994). AAO has been found in all white-rot genomes, and the number of the gene models identified to date varies from one to ten copies (Table 8.3). The only exceptions are *A. delicata* and the limitedly lignin-degrading *S. commune*, with no AAO gene models present. In contrast, AAOs are rare in brown-rot genomes, with only one to two gene copies in one species (Table 8.3).

MOX is a key methanol-metabolizing enzyme that oxidizes methanol to formaldehyde and H_2O_2 (Ozimek et al. 2005). Brown-rot fungi, such as *P. placenta* and *G. trabeum*, are suggested to employ MOX in H_2O_2 generation by using demethoxylated lignin as a substrate (Daniel et al. 2007; Martinez et al. 2009). In fact, in brown-rot fungi, MOX has been reported to be the main source of H_2O_2 that is used in Fenton chemistry for cellulose depolymerization, but its contribution to white-rot fungi has been less documented. MOX has been found in most white-rot and brown-rot genomes (Floudas et al. 2012).

POX catalyzes the oxidation of a number of common monosaccharides at C2 in the presence of O_2 to the corresponding 2-keto sugars, producing H_2O_2 as a by-product (Giffhorn 2000). As POX can also use various other electron acceptors, it has been suggested to play a role in the reduction of quinones during lignin degradation. *Phanerochaete chrysosporium* POX was found to be strongly upregulated when the fungus grew on lignin, indicating its potential role as a source of H_2O_2 (Manavalan et al. 2011). GOXs are only rarely reported in basidiomycetes. Two GOX-encoding genes have been reported in *P. chrysosporium* (Eriksson et al. 1986; Kelley and Reddy 1986). *Gox1*-corresponding peptides have been identified in media containing crystalline cellulose (Vanden Wymelenberg et al. 2005). However, the function of *gox1* and related genes is not known. In *C. subvermispora*, two GMC oxidases were automatically annotated as GOX-like GMCs (Fernandez-Fueyo et al. 2012b).

GLX is a well-characterized CRO in wood-rotting basidiomycetes. The broad substrate specificity of GLX comprises the oxidation of simple aldehydes to the corresponding carboxylic acids (Whittaker et al. 1996). In *P. chrysosporium*, GLX is associated with lignin decay by generating H_2O_2 for MnP and LiP. It is physiologically coupled with LiP (Kersten 1990), and *glx* transcripts coincide with *lip* and *mnp* mRNAs in defined media (Stewart et al. 1992; Kersten and Cullen 1993), soil (Bogan et al. 1996), and wood (Janse et al. 1998). GLX-encoding genes are present in most white-rot genomes studied but are absent from brown-rot genomes as well as from *C. cinerea* and *L. bicolor*. GLX together with related enzymes (CRO1–6) are the main candidates in H_2O_2 production in *P. chrysosporium*. Sequence homology of other *cro* genes to *glx* is relatively low, but copper-coordinating residues and the constitutive radical redox site are conserved (Whittaker 2002; Martinez et al. 2004; Vanden Wymelenberg et al. 2006b) as well as repeats of the highly conserved WSC domain in *cro3*, *cro4*, and *cro5*, which is suggested to be involved in carbohydrate binding. CRO1–6 encoding genes are widespread in the genomes across Agaricomycotina (Table 8.3).

4. P450 Monooxygenases

Cytochrome P450 monooxygenases are intracellular or membrane-bound enzymes belonging to a superfamily of heme-thiolate proteins that catalyze diverse reactions to transform xenobiotic chemicals and a variety of endogenous compounds (Sono et al. 1996; Bernhardt 2006; Isin and Guengerich 2007). In general, fungal genomes carry vast numbers of P450 genes, with up to 270 copies per species, exceeded only by plant genomes (Suzuki et al. 2012; Syed and Yadav 2012). In basidiomycetes, more expanded gene clans and families are present than in ascomycetes and zygomycetes (Floudas et al. 2012). While P450s were proposed to be involved in further degradation of extracellularly depolymerized lignin fragments, the similar widespread pattern of P450s in white-rot and brown-rot fungi as well as in ECM genomes (Table 8.3) also supports other roles, such as conversion of xenobiotic compounds and modification of fungal secondary metabolites (Lundell et al. 2010). Comparison between two *Phanerochaete* species revealed that *P. carnosa* has nearly twice the amount of P450 genes compared to *P. chrysosporium*, which may indicate a larger substrate range in *P. carnosa* (Suzuki et al. 2012). However, transcriptome analyses in the model white- and brown-rot fungi *P. chrysosporium* and *P. placenta*, respectively, have demonstrated that P450 genes

Table 8.4. Number of putative gene models from selected CAZy families in the genomes of plant-biomass-degrading basidiomycete fungi

	White rot											S	Brown rot							ECM
	Pc	Ab	Cs	Ad	Ps	Fm	Ds	Tv	Sh	Hi	Sc	Cc	Pp	Fp	Cp	Gt	Sl	Wc	Dsp	Lb
GHs																				
GH3	9	8	6	12	12	7	7	11	15	11	11	7	5	12	12	9	9	7	8	2
GH5	5	19	18	8	6	6	5	5	6	7	3	6	5	5	8	5	8	4	5	3
GH6	1	1	1	2	1	2	1	1	1	1	1	5	0	0	2	0	1	0	0	0
GH7	8	1	3	6	5	2	4	4	3	1	2	6	0	0	2	0	0	0	0	0
GH10	6	2	6	4	5	4	5	6	6	2	5	6	4	2	3	3	1	4	3	0
GH11	1	2	1	3	1	0	0	0	1	0	1	6	0	0	0	0	0	0	0	0
GH12	2	2	2	1	2	3	3	5	5	4	1	1	2	2	4	2	1	2	1	3
GH28	4	6	6	10	13	16	7	11	17	8	3	3	7	13	13	10	7	9	6	7
GH61	15	11	9	19	14	13	15	18	16	10	22	35	2	4	10	4	5	2	0	5
GH74	4	1	1	1	2	4	1	1	2	1	1	1	0	0	0	1	1	0	0	0
GH43	4	4	2	26	7	6	7	3	10	4	12	4	1	7	6	5	1	1	5	0
CEs																				
CE1	4	2	3	3	2	0	0	3	1	1	4	3	0	0	0	1	0	0	0	0
CE16	2	10	5	29	8	6	10	7	10	5	10	5	5	11	6	6	3	6	4	3
CE5	0	6	0	3	1	0	0	0	1	0	2	6	0	0	1	0	0	0	0	1
CE8	2	2	2	3	6	3	3	2	4	3	2	0	1	2	2	2	2	1	3	4
CE12	0	3	0	1	0	2	2	0	3	2	1	1	0	0	0	0	0	0	0	0
CE15	2	0	2	6	2	1	2	2	1	1	2	8	1	1	0	1	0	1	1	0

Based on Martin et al. (2008), Martinez et al. (2009), Ohm et al. (2010), Fernandez-Fueyo et al. (2012a), Floudas et al. (2012), Morin et al. (2012), Olson et al. (2012)
For abbreviations, see the Table 8.3 footnote

are differentially regulated in lignocellulose-containing cultures (Martinez et al. 2009; Syed and Yadav 2012).

5. Cellulose-Oxidizing Enzymes

In addition to lignin depolymerization, oxidative enzymes, including **cellobiose dehydrogenase** (CDH; EC 1.1.99.18) and GH61 enzymes, currently described as **copper-dependent polysaccharide monooxygenases** (PMOs) (Horn et al. 2012; Žifčkov and Baldrian 2012), have recently been shown to be important for enzymatic degradation of cellulose and have been identified in multiple transcriptomic and proteomic analyses (Tian et al. 2009; Vanden Wymelenberg et al. 2010; Berka et al. 2011; Eastwood et al. 2011; MacDonald et al. 2011). CDH is an extracellular enzyme that oxidizes cellobiose to cellobiono-1,5-lactone and is the only known example of a secreted flavocytochrome (Zamocky et al. 2006). Typically, only one copy of a CDH-encoding gene is found in the genome of most wood-degrading basidiomycetes (Table 8.3). It is suggested to have a role in degradation and modification of wood polymers via Fenton-type reaction by reducing Fe^{3+} to Fe^{2+} in the presence of H_2O_2 (Kremer and Wood 1992a, b). Supporting this hypothesis, CDH has been shown to be expressed during the growth of *P. chrysosporium* on solid-state wood and cellulose medium (Vanden Wymelenberg et al. 2006a; Sato et al. 2009). Recently, synergistic action between CDH and copper-dependent PMOs (GH61 enzymes) has been observed, indicating a role for CDH as an external electron donor for PMOs (Langston et al. 2011; Phillips et al. 2011a). PMOs are present in most genomes of ascomycetes and basidiomycetes, including white-rot and brown-rot fungi, litter decomposers, mycorrhizal fungi, and plant pathogens (Žifčkov and Baldrian 2012). However, brown-rot species seem to harbor fewer copies of GH61-encoding genes than white-rot species (Table 8.4).

6. Oxalate-Catabolizing Enzymes

Oxalate is a common fungal metabolite that plant-biomass-degrading basidiomycetes typically secrete to their environment in millimolar

quantities. While oxalate apparently has multiple roles in fungal nutritional strategies (Dutton and Evans 1996), accumulating evidence suggests that lignocellulose decay is promoted by the fungal secretion of oxalic acid (Yelle et al. 2008; Watanabe et al. 2010). A number of roles for oxalate in plant biomass degradation have been suggested, but functional studies are scarce. Evidently, oxalate can adjust fungal extracellular pH suitable for the activity of the plant-biomass-modifying enzymes. In addition, it enhances the performance of lignin-modifying peroxidases by chelating Mn^{3+} ions and acts as a substrate for the formation of H_2O_2 that has a central role in both white- and brown-rot decay of wood (Shimada et al. 1997; Urzúa et al. 1998; Varela and Tien 2003; Suzuki et al. 2006).

As oxalate is a toxic compound, regulation of its intra- and extracellular concentration is extremely crucial for fungi and for lignocellulose degradation because high oxalate levels are shown to inhibit the decomposition reactions. Therefore, specific oxalate-converting intracellular enzymes such as **oxalate decarboxylase** (ODC), oxalate oxidase (OXO), and formate dehydrogenase (FDH) are recognized as key enzymes in fungal lignocellulose decay (Mäkelä et al. 2009, 2010). In addition, transcriptome and proteome studies of the white-rot fungus *P. chrysosporium* and the brown-rot fungus *P. placenta* showed the active role of oxalate-degrading enzymes during fungal growth on wood (Vanden Wymelenberg et al. 2006a; Sato et al. 2007; Martinez et al. 2009).

ODC and OXO are evolutionarily closely related members of the cupin protein superfamily, sharing a common fold and structural motifs, but showing different catalytic properties (decarboxylase produces formate, while oxidase produces H_2O_2). ODC seems to be prevalent in fungi, typically with multiple gene copies in one species, whereas OXO is produced mainly in plants, with only two reports on fungal OXO activity (Aguilar et al. 1999; Grąz et al. 2009). The available genome data also show that the primary structure of basidiomycetous ODCs is a bicupin, thus predicting that the gene models encode proteins with decarboxylase activity. Typically, several copies of *fdh* genes, although less than for *odc*, are also present in one species (Mäkelä et al. unpublished results). The reason for this gene redundancy is not known but could point to differences in substrate specificity, kinetic properties, or regulation. In fact, the white-rot fungus *D. squalens* shows differential regulation of *odc* and *fdh* genes when grown on wood (Mäkelä et al. 2014).

C. Comparison of White-Rot and Brown-Rot Fungi

Possibly the most striking difference between white- and brown-rot fungi is the absence of lignin-modifying class II heme peroxidases (LiPs, MnPs, VPs) from the genomes of brown-rot fungi (Table 8.3). The current data show that the common ancestor of the wood-decaying basidiomycetes was a white-rot fungus from which, through a loss of lignin-modifying enzyme-encoding genes, the several lineages of brown-rot fungi evolved (Eastwood et al. 2011; Floudas et al. 2012; Olson et al. 2012). In addition, some of the basidiomycetes that have lost their lignin-modifying peroxidases have evolved other lifestyles, like EMC symbiosis with the roots of vascular plants. Apparently, the lifestyles of plant biomass-modifying basidiomycetes are overlapping and constantly changing during evolution.

Obviously, distinct strategies for lignin decay exist among the white-rot fungi. For instance, *P. chrysosporium* has 15 lignin-modifying peroxidases but no sensu stricto laccases, whereas *S. commune* harbors two laccases but lacks lignin-modifying peroxidase-encoding genes. In addition, the amount and diversity of lignin-modifying peroxidases varies within the different white-rot fungal species, possibly reflecting, for example, the selective and nonselective types of lignin depolymerization.

Compared with the white-rot fungi, brown-rot fungi have fewer hydrolytic CAZyme-encoding genes. Particularly, the genomes of brown-rot fungi almost totally lack the recalcitrant crystalline cellulose-hydrolyzing cellobiohydrolases (GH6/7), with the only exceptions being *Coniophora puteana* and *S. lacrymans* with four and one putative genes, respectively (Table 8.4). This further supports the nonenzymatic strategy of brown-rot fungi for cellulose depolymerization. Similarly, the brown-rot genomes display a reduced number of PMO- (GH61) encoding genes when compared to white-rot

fungi (Table 8.4; Eastwood et al. 2011), possibly emphasizing the fundamental differences in cellulose degradation between these fungi.

D. Transcriptomic and Proteomic Analyses of Plant Biomass-Degrading Basidiomycetes

Because of the limited number of white- and brown-rot fungal genome sequences, only recently transcriptome and proteome studies have given insights into the enzyme patterns needed for plant biomass degradation by basidiomycetes. The first transcriptome and secretome analyses have concentrated on the model wood-decaying basidiomycete species, for instance, *P. chrysosporium* (white rot) and *P. placenta* (brown rot). More recently, a **selective white-rot** fungus, *C. subvermispora*, was compared with *P. chrysosporium* (Fernandez-Fueyo et al. 2012b). Also, studies with another representative of the genus *Phanerochaete*, *P. carnosa*, have been conducted (Mahajan and Master 2010; MacDonald et al. 2011; Suzuki et al. 2012). In addition to the wood-decaying saprobes, the transcriptomes and secretomes of the biotrophic plant pathogenic fungi *M. larici-populina*, *Puccinia graminis*, and *U. maydis* have been recently published (Duplessis et al. 2011; Couturier et al. 2012).

Transcriptomic and comparative genomic studies of the first sequenced basidiomycete species have complemented the existing biochemical data on the fundamental differences in the plant cell-wall-degrading mechanisms. Systematic comparative studies with different lignocellulosic substrates have shown fungal adaptation to the carbon source by altering transcript and secretome profiles (Vanden Wymelenberg et al. 2011; Suzuki et al. 2012).

Comparative transcriptomic study of the white-rot fungus *P. chrysosporium* and the brown-rot fungus *P. placenta* have corroborated the biochemical data that these two decay types employ different approaches for the degradation of plant biomass and have adapted to their specific carbon source (Vanden Wymelenberg et al. 2011). Intriguingly, the importance of H_2O_2 production for both white- and brown-rot wood decay has been emphasized by transcriptome and proteome studies (Martinez et al. 2009; Sato et al. 2009; Vanden Wymelenberg et al. 2009, 2010). The analyses of *P. placenta* have pinpointed oxidases that putatively generate extracellular Fe^{2+} and H_2O_2 to be employed in Fenton chemistry for cellulose depolymerization, whereas the importance of glycoside hydrolases (with and without carbohydrate-binding modules) has been less pronounced compared to the white-rot fungi characterized so far (Martinez et al. 2009; Vanden Wymelenberg et al. 2010). However, upregulation of putative pectinases (polygalacturonase and rhamnogalacturonase from GH28 family) has been reported from *P. placenta* pine cultures, and high expression of an endoglucanase (GH5) was detected with both aspen and pine as carbon sources, thus suggesting its participation in cellulose depolymerization (Vanden Wymelenberg et al. 2011). In addition, the demethylation of lignin by brown-rot fungi has been reinforced by the studies, with *P. placenta* showing the overexpression of MOX that possibly uses the lignin-derived methanol to produce H_2O_2 (Martinez et al. 2009; Vanden Wymelenberg et al. 2010).

The comparison of different lignocelluloses (i.e., aspen and pine) as carbon sources showed significant induction of 13 GH-encoding genes of *P. chrysosporium* during growth only on aspen (Vanden Wymelenberg et al. 2011). Aspen as a carbon source has also resulted in a significant increase in the amount of transcripts and peptides corresponding to both *P. chrysosporium* GHs and **oxidoreductases** (Vanden Wymelenberg et al. 2010). *P. chrysosporium* has altogether nine cellobiohydrolases (GH6/7) (Table 8.4) that are expressed and produced differentially on various carbon sources (Broda et al. 1995; Vallim et al. 1998; Vanden Wymelenberg et al. 2009; Suzuki et al. 2010). Also, the cellulose-induced upregulation of the genes encoding putative endoglucanases (GH5), cellobiohydrolases (GH6/7), xylanases (GH10 and GH11), and GH61 proteins has been detected for *P. chrysosporium* (Vanden Wymelenberg et al. 2009). CDH has been shown to be secreted during the growth of *P. chrysosporium* on solid-state wood (red oak) and cellulose medium (Vanden Wymelenberg et al. 2005; Sato et al. 2009), further supporting its role in wood degradation. A recent study of the *P. chrysosporium* secretome showed that CDH and GH61 proteins are upregulated by xylan (Hori et al.

2011), thus suggesting a role for these oxidative enzymes also in the degradation of plant polysaccharides other than cellulose. Xylan promoted the production of putative endoxylanase (GH10) and glucuronoyl esterase (CE15) as well. In contrast, the addition of starch to the cellulose culture of *P. chrysosporium* was shown to repress the production of cellulolytic and xylanolytic enzymes (Hori et al. 2011).

Comparable patterns of extracellular GHs and oxidative enzymes are produced by *P. chrysosporium* and a related species, *P. carnosa*, when grown on spruce or cellulose (Mahajan and Master 2010). However, at transcript level, CAZymes-encoding genes were the most prevalent in *P. chrysosporium* wood cultures (Sato et al. 2009; Vanden Wymelenberg et al. 2010), while lignin-modifying peroxidase and H_2O_2-producing genes were most abundantly expressed by *P. carnosa* (MacDonald et al. 2011).

When the genome content of the selective white-rot species *C. subvermispora* was compared to *P. chrysosporium*, which simultaneously decays all the plant cell wall polymers, an increased number of MnP- and lipid metabolism-related genes (desaturases) was observed in *C. subvermispora*. That notion combined with the expression data from hardwood (aspen) cultures highlighted the importance of oxidoreductases and lipid peroxidation together with reduced cellulolytic capacity in selective lignin decay (Fernandez-Fueyo et al. 2012b).

When the secretome of a biotrophic maize pathogen *U. maydis* from medium containing maize bran was analyzed, 27 % of the detected proteins were identified as CAZymes targeting plant cell walls (Couturier et al. 2012). Compared to other pathogenic and biotrophic fungi, *M. larici-populina* and *P. graminis*, the repertoire of CAZymes in *U. maydis* was strikingly smaller (Duplessis et al. 2011). One xylanase (GH10) and two arabinofuranosidases (GH51 and GH62) were among the most abundant proteins detected from *U. maydis* cultures. Interestingly, 8 % of the secretome composed of putative oxidoreductases with a potential role in lignocellulose modification and two of the most abundant proteins corresponded to putative GMC oxidoreductases and two to GLX (Couturier et al. 2012). In fact, *U. maydis* GLX has been shown to be essential for successful maize infection (Leuthner et al. 2005).

The genome sequences of wood-decaying basidiomycetes bear a large number of putative extracellular oxidative enzyme-encoding genes, some of which may also be connected to plant biomass degradation. In addition, a set of expressed genes and secreted proteins with unknown function have been detected in basidiomycete wood cultures (Sato et al. 2009; Vanden Wymelenberg et al. 2009, 2011). These data suggest that the whole complexity of plant biomass degradation by basidiomycetes is yet to be unraveled. Furthermore, detailed time course analyses of basidiomycetes are needed to show the possible differences in enzyme sets between the early and later stages of plant biomass degradation.

IV. Conclusions

While ascomycetes and basidiomycete fungi are taxonomically distant, many similarities can be observed in their approach to degrade plant biomass. The type of plant polysaccharide-degrading enzymes and the variation in the CAZyme set encoded in their genomes is similar for ascomycetes and basidiomycetes and often correlates with their biotopes. Expression of the genes encoding plant polysaccharide-degrading enzymes is induced by plant-biomass-related compounds in both ascomycetes and basidiomycetes, and both coexpress genes related to the same plant biomass component (see previous sections). However, significant differences are also present, and the **regulators** that control this process have been observed to be different. In ascomycetes, several regulators involved in plant polysaccharide degradation have been identified (see Sect. II.D), but homologs for these regulators are absent in basidiomycete genomes, with the exception of the general carbon catabolite repressor CreA/CRE1 (R.B. Todd and R.P. de Vries, unpublished results). This suggests **parallel evolution** resulting in a similar regulatory system, but based on other regulators. This hypothesis is further strengthened by the observation that the Zn_2Cys_6 class of regulators is the most dominant class in ascomycetes (and includes all the known plant polysaccharide-degrading activators), while

the C2H2 regulators are the major class in basidiomycetes.

A second difference is the ability to modify **lignin**. White-rot basidiomycetes are particularly efficient at this (see Sect. III), and most studies into lignin degradation have been performed with these fungi. In contrast, few studies addressed the ascomycete potential for lignin degradation, and it is assumed that they are not able to do this because of the absence of the lignin-modifying peroxidases. Growth on lignin-rich carbon sources has been observed for several ascomycete fungi (de Vries et al. unpublished results). In these cultures, no significant lignin-modifying enzyme activities could be measured, suggesting that there may be a different mechanism for lignin modification in these fungi.

With the availability of an increasing number of fungal genome sequences covering the full diversity of fungi and transcriptome and proteome data sets related to plant biomass degradation, our insight in the various strategies for plant biomass degradation by fungi and the key differences between these strategies will deepen significantly over the coming years. This will result in better understanding of the relation between fungi and their biotope and aid in designing new or improved industrial applications of fungi. Further analysis of the genes in fungal genomes without a characterized homolog is also likely to result in new enzymes involved in plant biomass degradation. Recently, several new families have been added to the CAZy system, such as a second family of α-glucuronidases (GH115) (Chong et al. 2011) and a family of glucuronoyl esterases (CE15) (Li et al. 2007; Duranova et al. 2009), but more are likely to be discovered.

References

Adler E (1977) Lignin chemistry – past, present and future. Wood Sci Technol 11:169–218

Aguilar C, Urzúa U, Koenig C, Vicuña R (1999) Oxalate oxidase from *Ceriporiopsis subvermispora*: biochemical and cytochemical studies. Arch Biochem Biophys 366:275–282

Amselem J, Cuomo CA, van Kan JA, Viaud M, Benito EP, Couloux A, Coutinho PM, de Vries RP, Dyer PS, Fillinger S, Fournier E, Gout L, Hahn M, Kohn L, Lapalu N, Plummer KM, Pradier JM, Quevillon E, Sharon A, Simon A, ten Have A, Tudzynski B, Tudzynski P, Wincker P, Andrew M, Anthouard V, Beever RE, Beffa R, Benoit I, Bouzid O, Brault B, Chen Z, Choquer M, Collemare J, Cotton P, Danchin EG, Da Silva C, Gautier A, Giraud C, Giraud T, Gonzalez C, Grossetete S, Guldener U, Henrissat B, Howlett BJ, Kodira C, Kretschmer M, Lappartient A, Leroch M, Levis C, Mauceli E, Neuveglise C, Oeser B, Pearson M, Poulain J, Poussereau N, Quesneville H, Rascle C, Schumacher J, Segurens B, Sexton A, Silva E, Sirven C, Soanes DM, Talbot NJ, Templeton M, Yandava C, Yarden O, Zeng Q, Rollins JA, Lebrun MH, Dickman M (2011) Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet 7:e1002230

Andersen MR, Giese M, Vries PR, Nielsen J (2012) Mapping the polysaccharide degradation potential of *Aspergillus niger*. BMC Genomics 13:313

Aro N, Saloheimo M, Ilmén M, Penttilä M (2001) ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. J Biol Chem 276:24309–24314

Aro N, Saloheimo M, Ilmén M, Penttilä M (2003) ACEI, of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. Appl Environ Microbiol 69:56–65

Battaglia E, Hansen SF, Leendertse A, Madrid S, Mulder H, Nikolaev I, de Vries RP (2011a) Regulation of pentose utilisation by AraR, but not XlnR, differs in *Aspergillus nidulans* and *Aspergillus niger*. Appl Microbiol Biotechnol 91:387–397

Battaglia E, Visser L, Nijssen A, van Veluw J, Wösten HAB, de Vries RP (2011b) Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in the Eurotiales. Stud Mycol 69:31–38

Benoit I, Coutinho PM, Schols HA, Gerlach JP, Henrissat B, de Vries RP (2012) Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. BMC Genomics 13:321

Berka RM, Grigoriev IV, Otillar R, Salamov A, Grimwood J, Reid I, Ishmael N, John T, Darmond C, Moisan M-C, Henrissat B, Coutinho PM, Lombard V, Natvig DO, Lindquist E, Schmutz J, Lucas S, Harris P, Powlowski J, Bellemare A, Taylor D, Butler G, de Vries RP, Allijn IE, van den Brink J, Ushinsky S, Storms R, Powell AJ, Paulsen IT, Elbourne LDH, Baker SE, Magnuson J, LaBoissiere S, Clutterbuck AJ, Martinez D, Wogulis M, de Leon AL, Rey MW, Tsang A (2011) Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. Nat Biotechnol 29:922–927

Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. J Biotechnol 124:128–145

Blanchette RA (1995) Degradation of the lignocellulose complex in wood. Can J Bot 73:S999–S1010

Bödeker IT, Nygren CMA, Taylor F, Olson A, Lindahl BD (2009) Class II peroxidase-encoding genes are

present in a phylogenetically wide range of ectomycorrhizal fungi. ISME J 3:1387–1395

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–546

Bogan B, Schoenike B, Lamar R, Cullen D (1996) Expression of *lip* genes during growth in soil and oxidation of anthracene by *Phanerochaete chrysosporium*. Appl Environ Microbiol 62:3697–3703

Broda P, Birch PRJ, Brooks PR, Sims PFG (1995) PCR-mediated analysis of lignocellulolytic gene transcription by *Phanerochaete chrysosporium*: substrate-dependent differential expression within gene-families. Appl Environ Microbiol 61:2358–2364

Brunow G, Lundquist K, Gellerstedt G (1998) Lignin. In: Sjöström E, Alén R (eds) Analytical methods in wood chemistry, pulping, and papermaking. Springer, Berlin, pp 77–124

Buranov AU, Mazza G (2008) Lignin in straw of herbaceous crops. Ind Crop Prod 28:237–259

Burlat V, Ambert K, Ruel K, Joseleau J-P (1997) Relationship between the nature of lignin and the morphology of degradation performed by white-rot fungi. Plant Physiol Biochem 35:645–654

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active enZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37: D233–D238

Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J 3:1–30

Chen S, Xu J, Liu C, Zhu Y, Nelson DR, Zhou S, Li C, Wang L, Guo X, Sun Y, Luo H, Li Y, Song J, Henrissat B, Levasseur A, Qian J, Li J, Luo X, Shi L, He L, Xiang L, Xu X, Niu Y, Li Q, Han MV, Yan H, Zhang J, Chen H, Lv A, Wang Z, Liu M, Schwartz DC, Sun C (2012) Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. Nat Commun 3:913

Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997) Genetic and physical maps of Saccharomyces cerevisiae. Nature 387:67–73

Chong SL, Battaglia E, Coutinho PM, Henrissat B, Tenkanen M, de Vries RP (2011) The alpha-glucuronidase Agu1 from *Schizophyllum commune* is a member of a novel glycoside hydrolase family (GH115). Appl Microbiol Biotechnol 90:1323–1332

Chundawat SPS, Beckham GT, Himmel ME, Dale BE (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. Annu Rev Chem Biomol Eng 2:121–145

Coutinho PM, Andersen MR, Kolenova K, vanKuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T, Seiboth B, Battaglia E, Aguilar-Osorio G, de Jong JF, Ohm RA, Aguilar M, Henrissat B, Nielsen J, Stalbrand H, de Vries RP (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol 46(Suppl 1):S161–S169

Couturier M, Navarro D, Olivé C, Chevret D, Haon M, Favel A, Lesage-Meessen L, Henrissat B, Coutinho PM, Berrin J-G (2012) Post-genomic analyses of fungal lignocellulosic biomass degradation reveal the unexpected potential of the plant pathogen *Ustilago maydis*. BMC Genomics 13:57

Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Munsterkotter M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317:1400–1402

Daniel G, Volc J, Filonova L, Plihal O, Kubtov E, Halada P (2007) Characteristics of *Gloeophyllum trabeum* alcohol oxidase, an extracellular source of H_2O_2 in brown rot decay of wood. Appl Environ Microbiol 73: 6241–6253

de Souza WR, de Gouvea PF, Savoldi M, Malavazi I, de Souza Bernardes LA, Goldman MH, de Vries RP, de Castro Oliveira JV, Goldman GH (2011) Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. Biotechnol Biofuels 4:40

de Vries RP (2003) Regulation of *Aspergillus* genes encoding plant cell wall polysaccharide degrading enzymes; relevance for industrial production. Appl Microbiol Biotechnol 61:10–20

de Vries RP, Visser J (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microb Mol Biol Rev 65:497–522

de Vries RP, Visser J, de Graaff LH (1999) CreA modulates the XlnR-induced expression on xylose of *Aspergillus niger* genes involved in xylan degradation. Res Microbiol 150:281–285

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu J-R, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeyer C, Liu J, Harding M, Kim S, Lebrun M-H, Bohnert H, Coughlan S, Butler J, Calvo S, Ma L-J, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434:980–986

Delmas S, Pullan ST, Gaddipati S, Kokolski M, Malla S, Blythe MJ, Ibbett R, Campbell M, Liddell S, Aboobaker A, Tucker GA, Archer DB (2012) Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. PLoS Genet 8: e1002875

DiGuistini S, Wang Y, Liao NY, Taylor G, Tanguay P, Feau N, Henrissat B, Chan SK, Hesse-Orce U, Alamouti SM, Tsui CK, Docking RT, Levasseur A, Haridas S, Robertson G, Birol I, Holt RA, Marra MA, Hamelin RC, Hirst M, Jones SJ, Bohlmann J, Breuil C (2011) Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont *Grosmannia clavigera*, a lodgepole pine pathogen. Proc Natl Acad Sci U S A 108:2504–2509

Duplessis S, Cuomo CA, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, Chiu R, Coutinho PM, Feau N, Field M, Frey P, Gelhaye E, Goldberg J, Grabherr MG, Kodira CD, Kohler A, Kües U, Lindquist EA, Lucas SM, Mago R, Mauceli E, Morin E, Murat C, Pangilinan JL, Park R, Pearson M, Quesneville H, Rouhier N, Sakthikumar S, Salamov AA, Schmutz J, Selles B, Shapiro H, Tanguay P, Tuskan GA, Henrissat B, Van de Peer Y, Rouzé P, Ellis JG, Dodds PN, Schein JE, Zhong S, Hamelin RC, Grigoriev IV, Szabo LJ, Martin F (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A 108:9166–9171

Duranova M, Spanikova S, Wosten HA, Biely P, de Vries RP (2009) Two glucuronoyl esterases of *Phanerochaete chrysosporium*. Arch Microbiol 191:133–140

Dutton MV, Evans CS (1996) Oxalate production by fungi: its role in pathogenicity and ecology in the soil environment. Can J Microbiol 42:881–895

Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, Bendiksby M, Blumentritt M, Coutinho PM, Cullen D, de Vries RP, Gathman A, Goodell B, Henrissat B, Ihrmark K, Kauserud H, Kohler A, LaButti K, Lapidus A, Lavin JL, Lee YH, Lindquist E, Lilly W, Lucas S, Morin E, Murat C, Oguiza JA, Park J, Pisabarro AG, Riley R, Rosling A, Salamov A, Schmidt O, Schmutz J, Skrede I, Stenlid J, Wiebenga A, Xie X, Kües U, Hibbett DS, Hoffmeister D, Högberg N, Martin F, Grigoriev IV, Watkinson SC (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science 333:762–765

Eriksson KE, Pettersson B, Volc J, Musilek V (1986) Formation and partial characterization of glucose-2-oxidase, a hydrogen peroxide producing enzyme in *Phanerochaete chrysosporium*. Appl Microbiol Biotechnol 23:257–262

Eriksson K-E, Blanchette RA, Ander P (1990) Microbial and enzymatic degradation of wood and wood components. Springer, Berlin

Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, Porcel BM, Couloux A, Ségurens B, Poulain J, Anthouard V, Grossetete S, Khalili H, Coppin E, Déquard-Chablat M, Picard M, Contamine V, Arnaise S, Bourdais A, Berteaux-Lecellier V, Gautheret D, de Vries RP, Battaglia E, Coutinho PM, Danchin EGJ, Henrissat B, El Khoury R, Sainsard-Chanet A, Boivin A, Pinan-Lucarré B, Sellem CH, Debuchy R, Wincker P, Weissenbach J, Silar P (2008) The genome sequence of the model ascomycete fungus *Podospora anserina*. Genome Biol 9: R77

Evans CS, Dutton MV, Guillén F, Veness RG (1994) Enzymes and small molecular mass agents involved with lignocellulose degradation. FEMS Microbiol Rev 13:235–240

Faix O, Mozuch MD, Kirk TK (1985) Degradation of gymnosperm (guaiacyl) vs. angiosperm (syringyl/guaiacyl) lignins by *Phanerochaete chrysosporium*. Holzforschung 39:203–208

Fernandez-Fueyo E, Ruiz-Duenas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo LF, James TY, Seelenfreund D, Lobos S, Polanco R, Tello M, Honda Y, Watanabe T, Ryu JS, Kubicek CP, Schmoll M, Gaskell J, Hammel KE, St John FJ, Vanden Wymelenberg A, Sabat G, Splinter BonDurant S, Syed K, Yadav JS, Doddapaneni H, Subramanian V, Lavin JL, Oguiza JA, Perez G, Pisabarro AG, Ramirez L, Santoyo F, Master E, Coutinho PM, Henrissat B, Lombard V, Magnuson JK, Kues U, Hori C, Igarashi K, Samejima M, Held BW, Barry KW, LaButti KM, Lapidus A, Lindquist EA, Lucas SM, Riley R, Salamov AA, Hoffmeister D, Schwenk D, Hadar Y, Yarden O, de Vries RP, Wiebenga A, Stenlid J, Eastwood D, Grigoriev IV, Berka RM, Blanchette RA, Kersten P, Martinez AT, Vicuna R, Cullen D (2012a) Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci U S A 109:5458–5463

Fernandez-Fueyo E, Ruiz-Dueñas FJ, Miki Y, Martínez MJ, Hammel KE, Martínez AT (2012b) Lignin-degrading peroxidases from genome of selective ligninolytic fungus *Ceriporiopsis subvermispora*. J Biol Chem 287:16903–16916

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D, Gorecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kues U, Kumar TK, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A, Rokas A, Ruiz-Duenas FJ, Sabat G, Salamov A, Samejima M, Schmutz J, Slot JC, St John F, Stenlid J, Sun H, Sun S, Syed K, Tsang A, Wiebenga A, Young D, Pisabarro A, Eastwood DC, Martin F, Cullen D, Grigoriev IV, Hibbett DS (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-

Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868

Gan P, Ikeda K, Irieda H, Narusaka M, O'Connell RJ, Narusaka Y, Takano Y, Kubo Y, Shirasu K (2012) Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of Colletotrichum fungi. New Phytol. doi:10.1111/nph.12085

Gianoulis TA, Griffin MA, Spakowicz DJ, Dunican BF, Alpha CJ, Sboner A, Sismour AM, Kodira C, Egholm M, Church GM, Gerstein MB, Strobel SA (2012) Genomic analysis of the hydrocarbon-producing, cellulolytic, endophytic fungus *Ascocoryne sarcoides*. PLoS Genet 8:e1002558

Giardina P, Faraco V, Pezzella C, Piscitelli A, Vanhulle S, Sannia G (2010) Laccases: a never-ending story. Cell Mol Life Sci 67:369–385

Giffhorn F (2000) Fungal pyranose oxidases: occurrence, properties and biotechnical applications in carbohydrate chemistry. Appl Microbiol Biotechnol 54: 727–740

Gilbert HJ, Stalbrand H, Brumer H (2008) How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. Curr Opin Plant Biol 11:338–348

Glenn JK, Morgan MA, Mayfield MB, Kuwahara M, Gold MH (1983) An extracellular H_2O_2-requiring enzyme preparation involved in lignin biodegradation by the white rot basidiomycete *Phanerochaete chrysosporium*. Biochem Biophys Res Commun 114: 1077–1083

Goodwin SB, Ben M'barek S, Dhillon B, Wittenberg AH, Crane CF, Hane JK, Foster AJ, Van der Lee TA, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B, Bowler J, Bristow J, van der Burgt A, Canto-Canche B, Churchill AC, Conde-Ferraez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RC, Hammond-Kosack KE, Henrissat B, Kilian A, Kobayashi AK, Koopmann E, Kourmpetis Y, Kuzniar A, Lindquist E, Lombard V, Maliepaard C, Martins N, Mehrabi R, Nap JP, Ponomarenko A, Rudd JJ, Salamov A, Schmutz J, Schouten HJ, Shapiro H, Stergiopoulos I, Torriani SF, Tu H, de Vries RP, Waalwijk C, Ware SB, Wiebenga A, Zwiers LH, Oliver RP, Grigoriev IV, Kema GH (2011) Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genet 7:e1002070

Grąz M, Jarosz-Wilkołazka A, Pawlikowska-Pawlęga B (2009) *Abortiporus biennis* tolerance to insoluble metal oxides: oxalate secretion, oxalate oxidase activity, and mycelia morphology. Biometals 22: 401–410

Gruben BS (2012) Novel transcriptional activators of Aspergillus involved in plant biomass utilization. Microbiology. PhD thesis, Utrecht University

Guillén F, Martínez AT, Martínez MJ, Evans CS (1994) Hydrogen peroxide-producing system of *Pleurotus erygii* involving the extracellular enzyme aryl-alcohol oxidase. Appl Microbiol Biotechnol 41:465–470

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349–355

Hammel KE, Kapich AN, Jensen KA, Ryan ZC (2002) Reactive oxygen species as agents of wood decay by fungi. Enzyme Microb Technol 30:445–453

Hantus S, Pauly M, Darvill AG, Albersheim P, York WS (1997) Structural characterization of novel L-galactose-containing oligosaccharide subunits of jojoba seed xyloglucans. Carbohydr Res 304:11–20

Hasper AA, Dekkers E, van Mil M, van de Vondervoort PJ, de Graaff LH (2002) EglC, a new endoglucanase from *Aspergillus niger* with major activity towards xyloglucan. Appl Environ Microbiol 68:1556–1560

Hatakka A, Hammel KE (2010) Fungal biodegradation of lignocelluloses. In: Esser K, Hofrichter M (eds) The Mycota: a comprehensive treatise on fungi as experimental systems for basic and applied research. Industrial applications. Springer, Berlin, pp 319–340

Hatfield RD, Ralph J, Grabber JH (1999) Cell wall cross-linking by ferulates and diferulates in grasses. J Sci Food Agric 79:403–407

Heinfling A, Martínez MJ, Martínez AT, Bergbauer M, Szewzyk U (1998a) Purification and characterization of peroxidases from the dye-decolorizing fungus *Bjerkandera adusta*. FEMS Microbiol Lett 165:43–50

Heinfling A, Ruiz-Dueñas FJ, Martínez MJ, Bergbauer M, Szewzyk U, Martínez AT (1998b) A study on reducing substrates of manganese-oxidizing peroxidases from *Pleurotus eryngii* and *Bjerkandera adusta*. FEBS Lett 428:141–146

Hernndez-Ortega A, Ferreira P, Martínez AT (2012) Fungal aryl-alcohol oxidase: a peroxide-producing flavoenzyme involved in lignin degradation. Appl Microbiol Biotechnol 93:1395–1410

Higuchi T (1997) Biochemistry and molecular biology of wood. Springer, Berlin

Higuchi T (2006) Look back over the studies of lignin biochemistry. J Wood Sci 52:2–8

Hildén K, Martínez AT, Hatakka A, Lundell T (2005) The two manganese peroxidases Pr-MnP2 and Pr-MnP3 of *Phlebia radiata*, a lignin-degrading basidiomycete,

are phylogenetically and structurally divergent. Fungal Genet Biol 42:403–419

Hildén KS, Mäkelä MR, Hakala TK, Hatakka A, Lundell T (2006) Expression on wood, molecular cloning and characterization of three lignin peroxidase (LiP) encoding genes of the white rot fungus *Phlebia radiata*. Curr Genet 49:97–105

Hisamatsu M, Impallomeni G, York WS, Albersheim P, Darvill AG (1991) The structure of plant cell walls. 31. A new undecasaccharide subunit of xyloglucans with 2 α-L-fucosyl residues. Carbohydr Res 211:117–129

Hisamatsu M, York WS, Darvill AG, Albersheim P (1992) Characterization of seven xyloglucan oligosaccharides containing from seventeen to twenty glycosyl residues. Carbohydr Res 227:45–71

Hoegger PJ, Kilaru S, James TY, Thacker JR, Kües U (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. FEBS J 273:2308–2326

Hofrichter M, Ullrich R (2006) Heme-thiolate haloperoxidases: versatile biocatalysts with biotechnological and environmental significance. Appl Microbiol Biotechnol 71:276–288

Hofrichter M, Ullrich R, Pecyna MJ, Lundell T (2010) New and classic families of secreted fungal heme peroxidases. Appl Microbiol Biotechnol 87: 871–897

Hori C, Igarashi K, Katayama A, Samejima M (2011) Effects of xylan and starch on secretome of the basidiomycete *Phanerochaete chrysosporium* grown on cellulose. FEMS Microbiol Lett 321: 14–23

Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH (2012) Novel enzymes for the degradation of cellulose. Biotechnol Biofuels 5:45

Houbraken J, Frisvad JC, Samson RA (2011) Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*. IMA Fungus 2:87–95

Huisman MMH, Weel KGC, Schols HA, Voragen AGJ (2000) Xyloglucan from soybean (*Glycine max*) meal is composed of XXXG-type building units. Carbohydr Polym 42:185–191

Iben JR, Epstein JA, Bayfield MA, Bruinsma MW, Hasson S, Bacikova D, Ahmad D, Rockwell D, Kittler EL, Zapp ML, Maraia RJ (2011) Comparative whole genome sequencing reveals phenotypic tRNA gene duplication in spontaneous *Schizosaccharomyces pombe* La mutants. Nucleic Acids Res 39:4728–4742

Isin EM, Guengerich FP (2007) Multiple sequential steps involved in the binding of inhibitors to cytochrome P450 3A4. J Biol Chem 282:6863–6874

Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossen QM, Hossain MZ, Ahmed B, Rahim S, Rahman MS, Alam MM, Hou S, Wan X, Saito JA, Alam M (2012) Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. BMC Genomics 13:493

Jacobs A, Lundqvist J, Stalbrand H, Tjerneld F, Dahlman O (2002) Characterization of water-soluble hemicelluloses from spruce and aspen employing SEC/MALDI mass spectroscopy. Carbohydr Res 337:711–717

Janse BJH, Gaskell J, Akhtar M, Cullen D (1998) Expression of *Phanerochaete chrysosporium* genes encoding lignin peroxidases, manganese peroxidases, and glyoxal oxidase in wood. Appl Environ Microbiol 64: 3536–3538

Jeffries TW, Grigoriev IV, Grimwood J, Laplaza JM, Aerts A, Salamov A, Schmutz J, Lindquist E, Dehal P, Shapiro H, Jin YS, Passoth V, Richardson PM (2007) Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. Nat Biotechnol 25:319–326

Jensen KA, Bao W, Kawai S, Srebotnik E, Hammel KE (1996) Manganese-dependent cleavage of nonphenolic lignin structures by *Ceriporiopsis subvermispora* in the absence of lignin peroxidase. Appl Environ Microbiol 62:3679–3686

Jensen KAJ, Houtman CJ, Ryan ZC, Hammel KE (2001) Pathways for extracellular Fenton chemistry in the brown rot basidiomycete *Gloeophyllum trabeum*. Appl Environ Microbiol 67:2705–2711

Johansson T, Nyman PO (1996) A cluster of genes encoding major isozymes of lignin peroxidase and manganese peroxidase from the white-rot fungus *Trametes versicolor*. Gene 170:31–38

Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O, Perlin MH, Wösten HA, de Vries RP, Ruiz-Herrera J, Reynaga-Pena CG, Snetselaar K, McCann M, Pérez-Martín J, Feldbrügge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman ML, Stajich JE, Sentandreu R, Gonzlez-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Greilinger D, Münch K, Rössel N, Scherer M, Vranes M, Ladendorf O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho ECH, Cahill MJ, Boyce KJ, Klose J, Klosterman SJ, Deelstra HJ, Ortiz-Castellanos L, Li W, Sanchez-Alonso P, Schreier PH, Häuser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schlüter T, Margolis J, Platt D, Swimmer C, Gnirke A, Chen F, Vysotskaia V, Mannhaupt G, Güldener U, Münsterkötter M, Haase D, Oesterheld M, Mewes H-W, Mauceli EW, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BW (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444:97–101

Kapich A, Hofrichter M, Vares T, Hatakka A (1999) Coupling of manganese peroxidase-mediated lipid peroxidation with destruction of nonphenolic lignin model compounds and ^{14}C-labeled lignins. Biochem Biophys Res Commun 259:212–219

Kaur S, Dhillon GS, Brar SK, Chauhan VB (2012) Carbohydrate degrading enzyme production by plant pathogenic mycelia and microsclerotia isolates of

Macrophomina phaseolina through koji fermentation. Ind Crop Prod 36:140–148

Kelley RL, Reddy CA (1986) Purification and characterization of glucose oxidase from ligninolytic cultures of *Phanerochaete chrysosporium*. J Bacteriol 166:269–274

Kerem Z, Jensen KA, Hammel KE (1999) Biodegradative mechanism of the brown rot basidiomycete *Gloeophyllum trabeum*: evidence for an extracellular hydroquinone-driven fenton reaction. FEBS Lett 446:49–54

Kersten PJ (1990) Glyoxal oxidase of *Phanerochaete chrysosporium*: its characterization and activation by lignin peroxidase. Proc Natl Acad Sci U S A 87: 2936–2940

Kersten P, Cullen D (1993) Cloning and characterization of a cDNA encoding glyoxal oxidase, a peroxide-producing enzyme from the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Proc Natl Acad Sci U S A 90:7411–7413

Kimura Y, Asada Y, Oka T, Kuwahara M (1991) Molecular analysis of a *Bjerkandera adusta* lignin peroxidase gene. Appl Microbiol Biotechnol 35:510–514

Kirk TK, Cullen D (1998) Enzymology and molecular genetics of wood degradation by white-rot fungi. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York, pp 273–307

Kirk TK, Farrell RL (1987) Enzymatic "combustion": the microbial degradation of lignin. Annu Rev Microbiol 41:465–505

Kolbe J, Kubicek CP (1990) Quantification and identification of the main components of the *Trichoderma* cellulase complex with monoclonal antibodies using an enzyme-linked immunosorbent assay (ELISA). Appl Microbiol Biotechnol 34:26–30

Kolpak FJ, Blackwell J (1976) Determination of the structure of cellulose II. Macromolecules 9:273–278

Kremer SM, Wood PM (1992a) Evidence that cellobiose oxidase from *Phanerochete chrysosporium* is primarily an Fe(III) reductase. Eur J Biochem 205:133–138

Kremer SM, Wood PM (1992b) Production of Fenton's reagent by cellobiose oxidase from cellulolytic cultures of *Phanerochaete chrysosporium*. Eur J Biochem 208:807–814

Kües U, Rühl M (2011) Multiple multi-copper oxidase gene families in Basidiomycetes – what for? Curr Genomics 12:72–94

Kusters-van Someren M, Kishi K, Lundell T, Gold MH (1995) The manganese binding site of manganese peroxidase: characterization of an Asp179Asn site-directed mutant protein. Biochemistry 34: 10620–10627

Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. Appl Environ Microbiol 77:7007–7015

Leuthner B, Aichinger C, Oehmen E, Koopmann E, Müller O, Müller P, Kahmann R, Bölker M, Schreier PH (2005) A H_2O_2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*. Mol Genet Genomics 272:639–650

Li DM, Youngs HL, Gold MH (2001) Heterologous expression of a thermostable manganese peroxidase from *Dichomitus squalens* in *Phanerochaete chrysosporium*. Arch Biochem Biophys 385: 348–356

Li X-L, Špnikov S, de Vries RP, Biely P (2007) Identification of genes encoding microbial glucuronoyl esterase. FEBS Lett 581:4029–4036

Liu D, Gong J, Dai W, Kang X, Huang Z, Zhang HM, Liu W, Liu L, Ma J, Xia Z, Chen Y, Chen Y, Wang D, Ni P, Guo AY, Xiong X (2012) The genome of Ganoderma lucidum provides insights into triterpenes biosynthesis and wood degradation. PLoS One 7:e36146

Liers C, Bobeth C, Pecyna M, Ullrich R, Hofrichter M (2010) DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. Appl Microbiol Biotechnol 85:1869–1879

Lu X, Sun J, Nimtz M, Wissing J, Zeng AP, Rinas U (2010) The intra- and extracellular proteome of *Aspergillus niger* growing on defined medium with xylose or maltose as carbon substrate. Microb Cell Fact 9:23

Lundell T, Mäkelä MR, Hildén K (2010) Lignin-modifying enzymes in filamentous basidiomycetes: ecological, functional and phylogenetic review. J Basic Microbiol 50:5–20

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577

MacDonald J, Doering M, Canam T, Gong YC, Guttman DS, Campbell MM, Master ER (2011) Transcriptomic responses of the softwood-degrading white-rot fungus *Phanerochaete carnosa* during growth on coniferous and deciduous wood. Appl Environ Microbiol 77:3211–3218

Mahajan S, Master ER (2010) Proteomic characterization of lignocellulose-degrading enzymes secreted by *Phanerochaete carnosa* grown on spruce and microcrystalline cellulose. Appl Microbiol Biotechnol 86:1903–1914

Mäkelä MR, Hildén K, Hatakka A, Lundell TK (2009) Oxalate decarboxylase of the white-rot fungus *Dichomitus squalens* demonstrates a novel enzyme primary structure and non-induced expression on wood and in liquid cultures. Microbiology 155: 2726–2738

Mäkelä MR, Hildén K, Lundell T (2010) Oxalate decarboxylase: biotechnological update and prevalence of the enzyme in filamentous fungi. Appl Microbiol Biotechnol 87:801–814

Mäkelä MR, Sietiö O-M, de Vries RP, Timonen S, Hildén KS (2014) Oxalate-metabolising genes of the white-rot fungus Dichomitus squalens are dif-

ferentially induced on wood and at high proton concentration. PLoS One 9:e87959

Manavalan A, Adav SS, Sze SK (2011) ITRAQ-based quantitative secretome analysis of *Phanerochaete chrysosporium*. J Proteomics 75:642–654

Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buee M, Brokstein P, Canback B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbe J, Lin YC, Legue V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kues U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouze P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452: 88–92

Martínez MJ, Ruiz-Dueñas FJ, Guillén F, Martínez AT (1996) Purification and catalytic properties of two manganese-peroxidase isoenzymes from *Pleurotus eryngii*. Eur J Biochem 237:424–432

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, Coutinho PM, Henrissat B, Berka R, Cullen D, Rokhsar D (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695–700

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, del Río JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: a 2D NMR study. Phytochemistry 69:2831–2843

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26:553–560

Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, Hammel KE, Vanden Wymelenberg A, Gaskell J, Lindquist E, Sabat G, Bondurant SS, Larrondo LF, Canessa P, Vicuna R, Yadav J, Doddapaneni H, Subramanian V, Pisabarro AG, Lavin JL, Oguiza JA, Master E, Henrissat B, Coutinho PM, Harris P, Magnuson JK, Baker SE, Bruno K, Kenealy W, Hoegger PJ, Kues U, Ramaiya P, Lucas S, Salamov A, Shapiro H, Tu H, Chee CL, Misra M, Xie G, Teter S, Yaver D, James T, Mokrejs M, Pospisek M, Grigoriev IV, Brettin T, Rokhsar D, Berka R, Cullen D (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci U S A 106:1954–1959

Maruyama K, Goto C, Numata M, Suzuki T, Nakagawa Y, Hoshino T, Uchiyama T (1996) *o*-Acetylated xyloglucan in extracellular polysaccharides from cell-suspension cultures of *Mentha*. Phytochemistry 41: 1309–1314

Master ER, Zheng Y, Storms R, Tsang A, Powlowski J (2008) A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization. Biochem J 411:161–170

Matsushita J, Kato Y, Matsuda K (1985) Purification and properties of an alpha-D-xylosidase from *Aspergillus niger*. J Biochem 98:825–832

Matsushita J, Kato Y, Matsuda K (1987) Characterization of α-D-xylosidase II from *Aspergillus niger*. Agric Biol Chem 51:2

Miki Y, Ichinose H, Wariishi H (2010) Molecular characterization of lignin peroxidase from the white-rot basidiomycete *Trametes cervina*: a novel fungal peroxidase. FEMS Microbiol Lett 304:39–46

Mohnen D (1999) Biosynthesis of pectins and galactomannans. In: Barton D, Nakanishi K, Meth-Cohn O (eds) Comprehensive natural products chemistry. Elsevier, Dordrecht, pp 497–527

Morgenstern I, Robertson DL, Hibbett DS (2010) Characterization of three *mnp* genes of *Fomitiporia mediterranea* and report of additional class II peroxidases in the order Hymenochaetales. Appl Environ Microbiol 76:6431–6440

Morin E, Kohler A, Baker A, Foulongne-Oriol M, Lombard V, Nagy LG, Ohm R, Patyshakuliyeva A, Brun A, Aerts AL, Bailey AM, Billette C, Coutinho PM, Deakin G, Doddapaneni H, Floudas D, Grimwood J, Hildén K, Kües U, LaButti KM, Lapidus A, Lindquist EA, Lucas SM, Murat C, Riley RW, Salamov AA, Schmutz J, Subramanian V, Wösten HAB, Xu JP, Eastwood DC, Foster GD, Sonnenberg ASM, Cullen D, de Vries RP, Lundell T, Hibbett D, Henrissat B, Burton KS, Kerrigan RW, Challen MP, Grigoriev IV, Martin F (2012) The genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. Proc Natl Acad Sci U S A 109:17501–17506

Nagendran S, Hallen-Adams HE, Paper JM, Aslam N, Walton JD (2009) Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus *Amanita bisporigera*, based on the secretome of *Trichoderma reesei*. Fungal Genet Biol 46:427–435

Navarrete M, Callegari E, Eyzaguirre J (2012) The effect of acetylated xylan and sugar beet pulp on the expres-

sion and secretion of enzymes by *Penicillium purpurogenum*. Appl Microbiol Biotechnol 93: 723–741

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafton A, Latge J-P, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Pertea M, Price C, Pritchard BL, Quail MA, Rabbinowitsch E, Rawlins N, Rajandream M-A, Reichard U, Renauld H, Robson GD, de Cordoba SR, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, de Aldana CRV, Weidman J, White O, Woodward J, Yu J-H, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438:1151–1156

Niku-Paavola M-L, Karhunen E, Salola P, Raunio V (1988) Ligninolytic enzymes of the white-rot fungus *Phlebia radiata*. Biochem J 254:877–884

O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmuller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun MH, Lee YH, Miyara I, Moore N, Neumann U, Nordstrom K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stuber K, Sukno SA, Sweigard JA, Takano Y, Takahara H, Trail F, van der Does HC, Voll LM, Will I, Young S, Zeng Q, Zhang J, Zhou S, Dickman MB, Schulze-Lefert P, Loren V, van Themaat E, Ma LJ, Vaillancourt LJ (2012) Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. Nat Genet 44: 1060–1065

O'Neill MA, Warrenfeltz D, Kates K, Pellerin P, Doci T, Darvill AG, Albersheim P (1996) Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently-linked by a borate ester. J Biol Chem 271:22923–22930

Ogawa M, Kobayashi T, Koyama Y (2012) ManR, a novel Zn(II)(2)Cys(6) transcriptional activator, controls the beta-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol 49:987–995

Ohm RA, de Jong JF, Lugones LG, Aerts A, Kothe E, Stajich JE, de Vries RP, Record E, Levasseur A, Baker SE, Bartholomew KA, Coutinho PM, Erdmann S, Fowler TJ, Gathman AC, Lombard V, Henrissat B, Knabe N, Kües U, Lilly WW, Lindquist E, Lucas S, Magnuson JK, Piumi F, Raudaskoski M, Salamov A, Schmutz J, Schwarze FW, vanKuyk PA, Horton JS, Grigoriev IV, Wösten HA (2010) Genome sequence of the model mushroom *Schizophyllum commune*. Nat Biotechnol 28:957–963

Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, Barry KW, Condon BJ, Copeland AC, Dhillon B, Glaser F, Hesse CN, Kosti I, Labutti K, Lindquist EA, Lucas S, Salamov AA, Bradshaw RE, Ciuffetti L, Hamelin RC, Kema GH, Lawrence C, Scott JA, Spatafora JW, Turgeon BG, de Wit PJ, Zhong S, Goodwin SB, Grigoriev IV (2012) Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen dothideomycetes fungi. PLoS Pathog 8:e1003037

Olson Å, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canbäck B, Coutinho PM, Cullen D, Dalman K, Deflorio G, van Diepen LTA, Dunand C, Duplessis S, Durling M, Gonthier P, Grimwood J, Fossdal CG, Hansson D, Henrissat B, Hietala A, Himmelstrand K, Hoffmeister D, Högberg N, James TY, Karlsson M, Kohler A, Kües U, Lee Y-H, Lin Y-C, Lind M, Lindquist E, Lombard V, Lucas S, Lundén K, Morin E, Murat C, Park J, Raffaello T, Rouzé P, Salamov A, Schmutz J, Solheim H, Ståhlberg J, Vélëz H, de Vries RP, Wiebenga A, Woodward S, Yakovlev I, Garbelotto M, Martin F, Grigoriev IV, Stenlid J (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol 194:1001–1013

Ozimek P, Veenhuis M, van der Klei IJ (2005) Alcohol oxidase: a complex peroxisomal, oligomeric flavoprotein. FEMS Yeast Res 5:975–983

Padamsee M, Arun Kumar TK, Riley R, Binder M, Boyd A, Calvo AM, Furukawa K, Hesse C, Hohmann S, James TY, LaButti K, Lapidus A, Lindquist E, Lucas S, Miller K, Shantappa S, Grigoriev IV, Hibbett DS, McLaughlin DJ, Spatafora JW, Aime MC (2012) The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. Fungal Genet Biol 49:217–226

Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, Darvill A (1999) A xyloglucan-specific endo-beta-1,4-glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification and characterization of the recombinant enzyme. Glycobiology 9:93–100

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, van den Berg M, Breestraat S, Caddick

MX, Contreras R, Cornell M, Coutinho PM, Coutinho PM, Danchin EG, Debets AJ, van Dijck PW, van Dijk A, Dijkhuizen L, Driessen AJ, d'Enfert C, Geysens S, Goosen C, Groot GS, de Groot PW, Guillemette T, Henrissat B, Herweijer M, van den Hombergh JP, van den Hondel CA, van der Heijden RT, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJ, Meulenberg R, Menke H, Mortimer MA, Nielsen J, Oliver SG, Olsthoorn M, Pal K, van Peij NN, Ram AF, Rinas U, Roubos JA, Sagt CM, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van de Vondervoort PJ, Wedler H, Wosten HA, Zeng AP, van Ooyen AJ, Visser J, Stam H (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25:221–231

Pérez-Boada M, Ruiz-Dueñas FJ, Pogni R, Basosi R, Choinowski T, Martínez MJ, Piontek K, Martínez AT (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways. J Mol Biol 354:385–402

Petersen KL, Lehmbeck J, Christensen T (1999) A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet 262:668–676

Peterson R, Grinyer J, Nevalainen H (2011) Secretome of the coprophilous fungus *Doratomyces stemonitis* C8, isolated from koala feces. Appl Environ Microbiol 77:3793–3801

Phillips CM, Beeson WT, Cate JH, Marletta MA (2011a) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. ACS Chem Biol 6:1399–1406

Phillips CM, Iavarone AT, Marletta MA (2011b) Quantitative proteomic approach for cellulose degradation by *Neurospora crassa*. J Proteome Res 10:4177–4185

Piontek K, Glumoff T, Winterhalter K (1993) Low pH crystal structure of glyoxylated lignin peroxidase from *Phanerochaete chrysosporium* at 2.5 Å resolution. FEBS Lett 315:119–124

Piontek K, Smith AT, Blodig W (2001) Lignin peroxidase structure and function. Biochem Soc Trans 29:111–116

Piscitelli A, Giardina P, Lettera V, Pezzella C, Sannia G, Faraco V (2011) Induction and transcriptional regulation of laccases in fungi. Curr Genomics 12: 104–112

Poulos T, Edwards S, Wariishi H, Gold M (1993) Crystallographic refinement of lignin peroxidase at 2-Å. J Biol Chem 268:4429–4440

Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JC, Johansen KS, Krogh KB, Jorgensen CI, Tovborg M, Anthonsen A, Tryfona T, Walter CP, Dupree P, Xu F, Davies GJ, Walton PH (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proc Natl Acad Sci U S A 108:15079–15084

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF et al (2004) Lignins: natural polymers from oxidativecoupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29–60

Ravalason H, Grisel S, Chevret D, Favel A, Berrin JG, Sigoillot JC, Herpoel-Gimbert I (2012) *Fusarium verticillioides* secretome as a source of auxiliary enzymes to enhance saccharification of wheat straw. Bioresour Technol 114:589–596

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry 57:929–967

Rineau F, Roth D, Shah F, Smits M, Johansson T, Canbäck B, Olsen PB, Persson P, Grell MN, Lindquist E, Grigoriev IV, Lange L, Tunlid A (2012) The ectomycorrhizal fungus *Paxillus involutus* converts organic matter in plant litter using a trimmed brown-rot mechanism involving Fenton chemistry. Environ Microbiol 14:1477–1487

Rodríguez-Rincón F, Suarez A, Lucas M, Larrondo LF, de la Rubia T, Polaina J, Martínez J (2010) Molecular and structural modeling of the *Phanerochaete flavidoalba* extracellular laccase reveals its ferroxidase structure. Arch Microbiol 192:883–892

Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, Dominguez V, Anthouard V, Bally P, Bourras S, Cozijnsen AJ, Ciuffetti LM, Degrave A, Dilmaghani A, Duret L, Fudal I, Goodwin SB, Gout L, Glaser N, Linglin J, Kema GH, Lapalu N, Lawrence CB, May K, Meyer M, Ollivier B, Poulain J, Schoch CL, Simon A, Spatafora JW, Stachowiak A, Turgeon BG, Tyler BM, Vincent D, Weissenbach J, Amselem J, Quesneville H, Oliver RP, Wincker P, Balesdent MH, Howlett BJ (2011) Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. Nat Commun 2:202

Ruijter GJG, Visser J (1997) Carbon repression in aspergilli. FEMS Microbiol Lett 151:103–114

Ruiz E, Cara C, Manzanares P, Ballesteros M, Castro E (2008) Evaluation of steam explosion pretreatment for enzymatic hydrolysis of sunflower stalks. Enzyme Microb Technol 42:160–166

Ruiz-Dueñas FJ, Martínez AT (2009) Microbial degradation of lignin: how a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. Microb Biotechnol 2: 164–177

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441–452

Ryabova O, Vrsanska M, Kaneko S, van Zyl WH, Biely P (2009) A novel family of hemicellulolytic alpha-glucuronidase. FEBS Lett 583:1457–1462

Saha BC (2000) Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. Biotechnol Adv 18:403–423

Sakamoto Y, Nakade K, Nagai M, Uchimiya H, Sato T (2009) Cloning of *Lentinula edodes lemnp2*, a manganese peroxidase that is secreted abundantly in sawdust medium. Mycoscience 50:116–122

Sato S, Liu F, Koc H, Tien M (2007) Expression analysis of extracellular proteins from *Phanerochaete chrysosporium* grown on different liquid and solid substrates. Microbiology 153:3023–3033

Sato S, Feltus FA, Iyer P, Tien M (2009) The first genome-level transcriptome of the wood-degrading fungus *Phanerochaete chrysosporium* grown on red oak. Curr Genet 55:273–286

Saykhedkar S, Ray A, Ayoubi-Canaan P, Hartson SD, Prade R, Mort AJ (2012) A time course analysis of the extracellular proteome of *Aspergillus nidulans* growing on sorghum stover. Biotechnol Biofuels 5:52

Schaefer D, Steinberger Y, Whitford WG (1985) The failure of nitrogen and lignin control of decomposition in a North American desert. Oecologia 65: 382–386

Schmoll M, Franchi L, Kubicek CP (2005) Envoy, a PAS/LOV domain protein of *Hypocrea jecorina* (Anamorph *Trichoderma reesei*), modulates cellulase gene transcription in response to light. Eukaryot Cell 4:1998–2007

Schmoll M, Tian C, Sun J, Tisch D, Glass NL (2012) Unravelling the molecular basis for light modulated cellulase gene expression – the role of photoreceptors in *Neurospora crassa*. BMC Genomics 13:127

Schols HA, Voragen AGJ (1996) Complex pectins: structure elucidation using enzymes. In: Visser J, Voragen AGJ (eds) Pectin and pectinases. Elsevier Science, Amsterdam, pp 793–798

Schuster A, Schmoll M (2010) Biology and biotechnology of *Trichoderma*. Appl Microbiol Biotechnol 87:787–799

Shah P, Gutierrez-Sanchez G, Orlando R, Bergmann C (2009) A proteomic study of pectin-degrading enzymes secreted by *Botrytis cinerea* grown in liquid culture. Proteomics 9:3126–3135

Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Jordar VS, Maiti R, Kodira CD, Neafsey DE, Zeng Q, Hung CY, McMahan C, Muszewska A, Grynberg M, Mandel MA, Kellner E, Barker BM, Galgiani JN, Orbach MJ, Kirkland TN, Cole GT, Henn MR, Birren BW, Taylor JW (2009) Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives. Genome Res 19:1722–1731

Shimada M, Higuchi T (1991) Microbial, enzymatic and biomimetic degradation of lignin. In: Hon DNS, Shiraishi N (eds) Wood and cellulosic chemistry. Dekker, New York, pp 557–619

Shimada M, Akamatsu Y, Tokimatsu T, Mii K, Hattori T (1997) Possible biochemical roles of oxalic acid as a low molecular weight compound involved in brown-rot and white-rot wood decays. J Biotechnol 53:103–113

Sjöström E (1993) Wood chemistry, fundamentals and applications. Academic, San Diego

Sjöström E, Westermark U (1998) Chemical composition of wood and pulps: basic components and their distribution. In: Sjöström E, Alén R (eds) Analytical methods in wood chemistry, pulping, and papermaking. Springer, Berlin, pp 1–35

Sono M, Roach MP, Coulter ED, Dawson JH (1996) Heme-containing oxygenases. Chem Rev 96: 2841–2888

Srivastava M, Kapoor VP (2005) Seed galactomannans: an overview. Chem Biodivers 2:295–317

Stajich JE, Wilke SK, Ahren D, Au TCH, Birren BW, Borodovsky M, Burns C, Canbäck B, Casselton LA, Cheng CKJ, Deng J, Dietrich FS, Fargo DC, Farman ML, Gathman AC, Goldberg J, Guigó R, Hoegger PJ, Hooker JB, Huggins A, James TY, Kamada T, Kilaru S, Kodira C, Kües U, Kupfer D, Kwan HS, Lomsadze A, Li W, Lilly WW, Ma L-J, Mackey AJ, Manning G, Martin F, Muraguchi H, Natvig DO, Palmerini H, Ramesh MA, Rehmeyer CJ, Roe BA, Shenoy N, Stanke M, Ter-Hovhannisyan V, Tunlid A, Velagapudi R, Vision TJ, Zeng Q, Zolan ME, Pukkila PJ (2010) Genome evolution in mushrooms: insights from the genome and assembled chromosomes of *Coprinopsis cinerea* (*Coprinus cinereus*). Proc Natl Acad Sci U S A 107: 11889–11894

Stewart P, Cullen D (1999) Organization and differential regulation of a cluster of lignin peroxidase genes of *Phanerochaete chrysosporium*. J Bacteriol 181:3427–3432

Stewart P, Kersten P, Vanden Wymelenberg A, Gaskell J, Cullen D (1992) The lignin peroxidase gene family of *Phanerochaete chrysosporium*: complex regulation by carbon and nitrogen limitation, and the identification of a second dimorphic chromosome. J Bacteriol 174:5036–5042

Sugano Y (2009) DyP-type peroxidases comprise a novel heme peroxidase family. Cell Mol Life Sci 66:1387–1403

Sugiura T, Yamagishi K, Kimura T, Nishida T, Kawagishi H, Hirai H (2009) Cloning and homologous expression of novel lignin peroxidase genes in the white-rot fungus *Phanerochaete sordida* YK-624. Biosci Biotechnol Biochem 73:1793–1798

Sun J, Glass NL (2011) Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. PLoS One 6:e25654

Sundaramoorthy M, Kishi K, Gold M, Poulos T (1994) The crystal structure of manganese peroxidase from *Phanerochaete chrysosporium* at 2.06-Å resolution. J Biol Chem 269:32759–32767

Suzuki MR, Hunt CG, Houtman CJ, Dalebroux ZD, Hammel KE (2006) Fungal hydroquinones contribute to brown rot of wood. Environ Microbiol 8:2214–2223

Suzuki H, Igarashi K, Samejima M (2010) Cellotriose and cellotetraose as inducers of the genes encoding cellobiohydrolases in the basidiomycete *Phanerochaete chrysosporium*. Appl Environ Microbiol 76:6164–6170

Suzuki H, MacDonald J, Syed K, Salamov A, Hori C, Aerts A, Henrissat B, Wiebenga A, VanKuyk PA, Barry K, Lindquist E, LaButti K, Lapidus A, Lucas S, Coutinho PM, Gong YC, Samejima M, Mahadevan R, Abou-Zaid M, de Vries RP, Igarashi K, Yadav JS, Grigoriev IV, Master ER (2012) Comparative genomics of the white-rot fungi, *Phanerochaete carnosa* and *P. chrysosporium*, to elucidate the genetic basis of the distinct wood types they colonize. BMC Genomics 13:444

Syed K, Yadav JS (2012) P450 monooxygenases (P450ome) of the model white rot fungus *Phanerochaete chrysosporium*. Crit Rev Microbiol 38:339–363

ten Have R, Hartmans S, Teunissen PJM, Field JA (1998) Purification and characterization of two lignin peroxidase isozymes produced by *Bjerkandera* sp. strain BOS55. FEBS Lett 422:391–394

Tian C, Beeson WT, Iavarone AT, Sun J, Marletta MA, Cate JHD, Glass NL (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. Proc Natl Acad Sci U S A 106:22157–22162

Tien M, Kirk TK (1984) Lignin-degrading enzyme from *Phanerochaete chrysosporium*: purification, characterizationand catalytic properties of a unique H_2O_2-requiring oxygenase. Proc Natl Acad Sci U S A 81:2280–2284

Urzúa U, Kersten PJ, Vicuña R (1998) Manganese peroxidase-dependent oxidation of glyoxylic acid synthetized by *Ceriporiopsis subvermispora* produces extracellular hydrogen peroxide. Appl Environ Microbiol 64:68–73

Vallim MA, Janse BJH, Gaskell J, Pizzirani-Kleiner AA, Cullen D (1998) *Phanerochaete chrysosporium* cellobiohydrolase and cellobiose dehydrogenase transcripts in wood. Appl Environ Microbiol 64:1924–1928

van den Berg MA, Albang R, Albermann K, Badger JH, Driessen AJ, Garcia-Estrada C, Fedorova ND, Harris DM, Heijne WHM, Joardar V, Kiel JAK, Kovalchuk A, Martin JF, Nierman WC, Nijland JG, Pronk JT, Pronk JT, Roubos JA, van der Klei IJ, van Peij NNME, Veenhuis M, von Dohren H, Wagner C, Wortman J, Bovenberg RAL (2008) Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. Nat Biotechnol 26:1161–1168

van Peij N, Gielkens MMC, de Vries RP, Visser J, de Graaff LH (1998a) The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol 64:3615–3619

van Peij NN, Visser J, de Graaff LH (1998b) Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol 27:131–142

van Soest PJ (1982) Nutritional ecology of the ruminant. Cornell University Press, Ithaca

Vanden Wymelenberg AV, Sabat G, Martinez D, Rajangam AS, Teeri TT, Gaskell J, Kersten PJ, Cullen D (2005) The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. J Biotechnol 118:17–34

Vanden Wymelenberg A, Minges P, Sabat G, Martinez D, Aerts A, Salamov A, Grigoriev I, Shapiro H, Putnam N, Belinky P, Dosoretz C, Gaskell J, Kersten P, Cullen D (2006a) Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveals complex mixtures of secreted proteins. Fungal Genet Biol 43:343–356

Vanden Wymelenberg A, Sabat G, Mozuch M, Kersten PJ, Cullen D, Blanchette RA (2006b) Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Appl Environ Microbiol 72:4871–4877

Vanden Wymelenberg A, Gaskell J, Mozuch M, Kersten P, Sabat G, Martinez D, Cullen D (2009) Transcriptome and secretome analyses of *Phanerochaete chrysosporium* reveal complex patterns of gene expression. Appl Environ Microbiol 75:4058–4068

Vanden Wymelenberg A, Gaskell J, Mozuch M, Sabat G, Ralph J, Skyba O, Mansfield SD, Blanchette RA, Martinez D, Grigoriev IV, Kersten P, Cullen D (2010) Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. Appl Environ Microbiol 76:3599–3610

Vanden Wymelenberg A, Gaskell J, Mozuch M, Bondurant SS, Sabat G, Ralph J, Skyba O, Mansfield SD, Blanchette RA, Grigoriev IV, Kersten PJ, Cullen D (2011) Significant alteration of gene expression in wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium* by plant species. Appl Environ Microbiol 77:4499–4507

Varela E, Tien M (2003) Effect of pH and oxalate on hydroquinone-derived hydroxyl radical formation during brown rot wood degradation. Appl Environ Microbiol 69:6025–6031

Varela E, Guillén F, Martínez AT, Martínez MJ (2001) Expression of *Pleurotus eryngii* aryl-alcohol oxidase in *Aspergillus nidulans*: purification and characterization of the recombinant enzyme. Biochim Biophys Acta 1546:107–113

Vares T, Kalsi M, Hatakka A (1995) Lignin peroxidases, manganese peroxidases, and other ligninolytic enzymes produced by *Phlebia radiata* during

solid-state fermentation of wheat straw. Appl Environ Microbiol 61:3515–3520

Vidal S, Doco T, Williams P, Pellerin P, York WS, O'Neill MA, Glushka J, Darvill AG, Albersheim P (2000) Structural characterization of the pectic polysaccharide rhamnogalacturonan II: evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain. Carbohydr Res 326:227–294

Vincken J-P, York WS, Beldman G, Voragen AGJ (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. Plant Physiol 114:9–13

Watanabe T, Shitan N, Suzuki S, Umezawa T, Shimada M, Yazaki K, Hattori T (2010) Oxalate efflux transporter from the brown rot fungus *Fomitopsis palustris*. Appl Environ Microbiol 76: 7683–7690

Whittaker JW (2002) Galactose oxidase. In: Valentine JS, Gralla EB (eds) Advances in protein chemistry. Academic, New York, pp 1–49

Whittaker MM, Kersten PJ, Nakamura N, Sanders-Loehr J, Schweizer ES, Whittaker JW (1996) Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. J Biol Chem 271:681–687

Willats WGT, McCartney L, MacKie W, Knox JP (2001) Pectin: cell biology and prospects for functional analysis. Plant Mol Biol 47:9–27

Wilson DB (2009) Cellulases and biofuels. Curr Opin Biotechnol 20:295–299

Yelle DJ, Ralph J, Lu F, Hammel KE (2008) Evidence for cleavage of lignin by a brown rot basidiomycete. Environ Microbiol 10:1844–1849

Yoshikawa K, Yamamoto K, Okada S (1993a) Isolation of *Aspergillus flavus* MO-5 producing two types of intracellular α-D-xylosidases: purification and characterization of a-D-xylosidase I. Biosci Biotechnol Biochem 57:1275–1280

Yoshikawa K, Yamamoto K, Okada S (1993b) Purification and characterization of an intracellular α-D-xylosidase II from *Aspergillus flavus* MO-5. Biosci Biotechnol Biochem 57:1281–1285

Zamocky M, Ludwig R, Peterbauer C, Hallberg BM, Divne C, Nicholls P, Haltrich D (2006) Cellobiose dehydrogenase – a flavocytochrome from wood-degrading, phytopathogenic and saprotropic fungi. Curr Protein Pept Sci 7:255–280

Žifčkov L, Baldrian P (2012) Fungal polysaccharide monooxygenases: new players in the decomposition of cellulose. Fungal Ecol 5:481–489

9 Transcriptomics of Industrial Filamentous Fungi: A New View on Regulation, Physiology, and Application

Benjamin M. Nitsche[1], Vera Meyer[1]

CONTENTS

I. Introduction

Filamentous fungi are of great economic importance as a source of high-value products. The volume and range of products include (a) chemicals used as ingredients for the food industry and as key building blocks for the chemical industry; (b) secondary metabolites such as antibiotics and statins for the pharmaceutical industry; (c) enzymes for large-scale manufacturing processes, including paper and pulp, food and feed, beverages, wine, detergents, textiles, and biofuel production; and (d) heterologous proteins, including pharmaceuticals for therapeutic use (Lubertozzi and Keasling 2009; Meyer 2008). The main industrial fungi include species of *Aspergillus*, such as *A. niger* for the production of citric acid, proteases, α-amylase and xylanase and *A. oryzae* for the production of proteases, tannases and lipases. Citric acid has been produced by *A. niger* at an industrial scale since 1923 and is now used as an ingredient in beverages, food, cosmetics and pharmaceuticals because of its acidity, flavour and chelating property (Ruijter et al. 2002). *Trichoderma reesei* is a well-established host for the production of hemicellulases and cellulases. The most prominent industrial fungal host for secondary metabolite production is *Penicillium chrysogenum*, which has been used for production of the β-lactam antibiotic penicillin since 1943 (Kardos and Demain 2011).

The key chalenge of the future is to optimally and sustainably utilize industrial fungi by maximizing production at minimal cost, ensuring robustness and accelerating the identification, engineering and development of new products. In contrast to traditional strain improvement programs comprising repetitive cycles of mutagenesis and screening for improved phenotypes, the postgenomic era has now offered new concepts and tools that facilitate knowledge-driven systematic and targeted approaches for strain development. Within the spectrum of -omics technologies, genome-wide transcriptional profiling (transcriptomics) has been used most intensively to investigate industrially important filamentous fungi. The aim of the current chapter is to give an overview of the recently published

[1]Department of Applied and Molecular Microbiology, Institute of Biotechnology, Berlin University of Technology, Gustav-Meyer-Allee 25, Berlin 13355, Germany; e-mail: vera.meyer@tu-berlin.de

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

fungal transcriptomic studies by focusing on the four main industrial cell factories: *A. niger*, *A. oryzae*, *P. chrysogenum* and *T. reesei*. Section II summarises the technological platforms available for fungal transcript profiling and the currently used approaches to cultivate filamentous fungi under controlled or uncontrolled conditions. Transcriptomics insights into the general protein secretion machinery of filamentous fungi are highlighted in Sect. III, whereas the polymer-degrading enzymatic repertoire of filamentous fungi and the underlying regulatory mechanisms are described in Sect. IV. Section V focuses on transcriptomics studies, analysing the metabolic versatility of filamentous fungi by studying selected primary and secondary metabolites. Section VI discusses recent transcriptomic approaches dealing with the identification of gene networks controlling filamentous growth during submerged cultivation in bioreactors. Finally, Sect. VII discusses the current status for fungal integrative data analysis, aiming to combine all -omics levels to reconstruct precise pictures of whole-cellular processes.

II. Experimental Design for Transcriptomics Analyses

A. Platforms for Transcript Profiling

The term ***transcriptomics*** describes the simultaneous analysis and comparison of expression profiles of genome-wide or (very) large and ideally unbiased gene sets. Basically, transcriptomic studies can be considered as a multistep approach divided into several steps: (a) formulation of the major research question; (b) design of the experiment by considering the statistical power, which depends on the variance, the effect size (fold change in gene expression), and the number of replicates; (c) conduction of the experiment followed by rapid sampling and immediate quenching of all cellular activity to obtain a reliable and meaningful snapshot of the transcriptomic status of the biological samples; (d) extraction and analysis of RNA transcripts using an appropriate platform; (e) data processing and statistical analysis for the identification of differentially or coexpressed gene sets; and (f) data interpretation and hypothesis generation followed by experiments to prove or disprove the generated hypothesis.

Although various sequencing approaches have been developed for quantitative transcript profiling, including serial analysis of gene expression (SAGE), cap analysis of gene expression (CAGE), massively parallel signature sequencing (MPSS) and RNA sequencing (RNA-seq) using deep-sequencing technologies (Wang et al. 2009), the vast majority of transcriptomic studies in industrially important filamentous fungi have relied so far on **DNA microarrays**. Here, the complete set of messenger RNA (mRNA) molecules present in a cell is transcribed to complementary DNA (cDNA) molecules, which are subsequently used for hybridization with probes immobilized on the surface of a microarray, which ideally represents all Open Reading Frames predicted for the genome of interest. Basically, one can distinguish different DNA microarray platforms based on two criteria relating to the length of the probes and the labelling protocol (Ahmed 2006). Immobilized probes can either be oligomers ranging from 25 to 60 nucleotides in length (commercially available platforms, e.g. Affymetrix, NimbleGen, Agilent) or spotted cDNA molecules (usually academic platforms). The second criterion, which is related to the labelling and signal readout, has a strong impact on the experimental design of the transcriptomic study. In the case of two-colour (two-channel) DNA microarray platforms, the cDNAs prepared from two samples to be compared are labelled with two different fluorescent colours (e.g. Cy3 and Cy5). Both labelled cDNA populations are mixed and then hybridized to a single microarray. The relative fluorescence intensities are then used to calculate ratio-based intensities and to deduce gene up- or downregulation. In one colour (one-channel) DNA microarray platforms, however, each of the cDNA populations is labelled with the same molecule (e.g. biotin in the case of Affymetrx gene chips), and each microarray is exposed to only one sample. Relative

abundance-based data are then inferred and used to calculate (differential) gene expression levels. In addition to the fact that the one-channel system needs twice as many microarrays when compared to the two-channel system (not considering dye swap and dye switch experimental designs), both microarray platforms differ with respect to price, reproducibility, applicability and subsequent analysis, as discussed by Ahmed (2006).

Table 9.1 summarises **DNA microarray platforms**, which are currently established and in use for transcriptomic studies of *A. niger*, *A. oryzae*, *P. chrysogenum* and *T. reesei*. This compilation implies that the individual research groups have different preferences. Whereas mainly the commercially available one-colour Affymetrix platform has been used for *A. niger* and *P. chrysogenum*, nine different platforms, including commercial and academic ones, have been applied for *A. oryzae*. In general, the decision to specifically exploit only a single one-channel platform strongly facilitates the analysis of different and independent experimental data sets. It allows data reuse for reference conditions (wild-type or standard growth conditions) and considerably improves global normalization and analysis of diverse experimental data sets at a later time point. For example, publically deposited microarray data have been reused to improve the robust multiarray analysis (RMA) preprocessing (Irizarry et al. 2003) for *A. niger* (Nitsche et al. 2012).

At the time of writing, the number of published studies applying RNA-seq for transcriptome profiling of industrial filamentous fungi was limited (Delmas et al. 2012; Novodvorska et al. 2013; Wang et al. 2010), probably because of the high costs and the not-yet-established pipelines for efficient data retrieval, storage and analysis (Baginsky et al. 2010). However, it can be expected that the number of published RNA-seq studies will considerably increase in the near future, as RNA-seq has several advantages over microarray analyses: (a) It is not limited to a priori DNA sequence knowledge; (b) it has almost no background signals and no saturation effects, thus allowing the quantification of very low and very high abundant transcripts; (c) it provides single-nucleotide resolution, allowing the precise definition of transcript borders and splice sites, as well as the identification of single-nucleotide polymorphisms; and (d) it allows the quantification of non-coding transcripts, even in the absence of a reference genome (Wang et al. 2009).

An extended hybridization-based approach for the identification of unknown or non-coding transcripts are **tiling arrays** which are designed to cover the whole-genome sequence, including predicted coding and non-coding regions using overlapping probes. Such tiling arrays can also be applied for genomic hybridization analysis, for example, as shown for (hemi-)cellulose production strains of *T. reesei* (Vitikainen et al. 2010). This study has identified 44 previously uncharacterized genomic alterations introduced by classical strain improvement approaches. An alternative to tiling arrays are so-called sparse arrays, which have been described recently for *T. reesei* (Arvas et al. 2010). Rather than using overlapping probes to cover the complete genome, the authors used a standard DNA microarray targeting known coding transcripts and also designed probes that scattered over predicted intergenic regions with approximately 100 nucleotide gaps. Using this approach, 23 strongly expressed transcripts were identified, which could represent novel genes not predicted by current gene prediction models. This sparse array has also been used in a combinatorial study, which investigated the effect of cell density and the specific growth rate on protein production in *T. reesei*. Here, 16 novel and differentially expressed genes were found (Arvas et al. 2011).

B. Fungal Cultivation Protocols

Costs often limit the number of biological replicates in transcriptome studies. This, however, reduces the statistical power of an experiment ($1 - \beta$; with β the false-negative rate), which is the probability of avoiding false negatives and hence refers to the sensitivity of an experiment. The statistical power is inversely proportional to the variance and directly proportional to the sample size and the effect size, which, in the

Table 9.1. Microarray platforms used for transcriptomic studies

Platform	Specifications	Reference strain	References
A. niger			
Affymetrix GeneChip	~14,000 probes	CBS513.88	Carvalho et al. (2012) De Bekker et al. (2011b) Guillemette et al. (2007) Jacobs et al. (2009) Jørgensen et al. (2010) Kwon et al. (2012) Kwon et al. (2013) MacKenzie et al. (2005) Martens-Uzunova and Schaap (2008) Martens-Uzunova et al. (2006) Meyer et al. (2007) Meyer et al. (2009) Nitsche et al. (2012) Pel et al. (2007) Van de Vondervoort et al. (2007) Van den Berg et al. (2010) Van der Veen et al. (2009) Yuan et al. (2007) Yuan et al. (2008)
Agilent	~14,000 probes	CBS513.88	De Souza et al. (2011)
Agilent	n. i.	n. i.	Poulsen et al. (2012)
A. oryzae			
NimbleGen	~12,000 probes	RIB40	Kimura et al. (2008) Terabayashi et al. (2012)
Affymetrix GeneChip	n. i.	n. i.	Sakamoto et al. (2009)
Affymetrix GeneChip	~15,000	RIB40	Ohno et al. (2011) Wada et al. (2012)
cDNA array	~2,000 probes	n. i.	Maeda et al. (2004)
cDNA array	~3,000 probes	n. i.	Masai et al. (2006) Sakamoto et al. (2008)
cDNA array	~11,000 probes	n. i.	Noguchi et al. (2009) Tamano et al. (2008)
cDNA array	~12,000 probes	n. i.	Imanaka et al. (2010) Terabayashi et al. (2010)
Agilent	n. i.	n. i.	Jin et al. (2010)
Agilent	~13,000 probes	RIB40	Ogawa et al. (2010) Ogawa et al. (2012)
A. niger*, *A. nidulans*, *A. oryzae			
Affymetrix GeneChip	~44,000 probes	ATCC 1015 FGSC A4 RIB40	Andersen et al. (2008b) Andersen et al. (2011) Andersen et al. (2009) Coutinho et al. (2009) Panagiotou et al. (2009) Salazar et al. (2009) Vongsangnak et al. (2009) Vongsangnak et al. (2010) Vongsangnak et al. (2011)
T. reesei			
NimbleGen	Sparse array	n. i.	Arvas et al. (2010) Arvas et al. (2011) Häkkinen et al. (2012)
NimbleGen	~9,000 probes	n. i.	Chen et al. (2012) Karimi-Aghcheh et al. (2013) Metz et al. (2011) Seiboth et al. (2012)

(continued)

Table 9.1. (continued)

Platform	Specifications	Reference strain	References
			Tisch et al. (2011)
cDNA array	~2,000 probes	n. i.	Bonaccorsi et al. (2006)
P. chrysogenum			
Affymetrix GeneChip	~14,000 probes	Wisconsin54-1255	Douma et al. (2011) Gombert et al. (2011) Harris et al. (2009a) Harris et al. (2009b) Hoff et al. (2010a) Hoff et al. (2010b) Koetsier et al. (2010) Sigl et al. (2011) Snoek et al. (2009) Veiga et al. (2012a) Veiga et al. (2012b) Veiga et al. (2012c)
A. terreus			
cDNA array	~5,000 probes	NRRL1960	Li et al. (2011)
Genome fragment microarray		n. i.	Askenazi et al. (2003)

n. i. not indicated

case of comparative transcriptomics, is the fold change. The multiple-hypothesis-testing problem (Sainani 2009) in transcriptome studies and the commonly applied corrections to adjust the false-positive rate α can result in further reduction of the statistical power (Bourgon et al. 2010). Consequently, minimization of variance is the remaining tuneable factor to increase the power if the effect size is kept constant. Variance is introduced during the analytical procedures, including manufacturing of the microarray, RNA extraction, synthesis of labelled cDNA, hybridization and readout, as well as during fungal cultivation and treatment. Variance as part of the analytics (i.e. the technical variance) is largely platform dependent and can be kept at low levels when reproducible and standardized protocols are used by well-trained staff. The biological variance is a stochastic variation, which also can be kept at the lowest level when the cells are cultivated under reproducible and controlled conditions and quick sampling protocols are used to immediately stop any metabolic activity of the cells.

Table 9.2 gives an overview of cultivation protocols used for *A. niger*, *A. oryzae*, *P. chrysogenum* and *T. reesei* to obtain biomass for transcriptomics analyses. Basically, **surface growth cultivations** (solid-state fermentations or agar plate cultures), simulating the natural fungal habitat, and **liquid cultivations** (submerged fermentations), reflecting the majority of industrial production processes, were used. Owing to its use as *kōji-kin* in traditional solid-state fermentations for the production of Japanese food and ingredients thereof (Murooka and Yamshita 2008), the majority of studies applying surface growth conditions were performed for *A. oryzae*. In agreement, cultivation of *A. oryzae* under surface growth conditions has been shown to result in higher hydrolase activities when compared with submerged cultivation (Imanaka et al. 2010). In a similar study, Maeda et al. (2004) compared gene expression profiles of *A. oryzae* cultivated under submerged conditions in glucose-rich and glucose-depleted media with expression levels obtained for *A. oryzae* cultivated under solid-state conditions with industrial substrates such as wheat bran as carbon sources. The data gained led to the hypothesis that high expression of hydrolytic genes under solid-state conditions is a consequence of glucose limitation, which is causatively linked to the low water activity of the medium. Low diffusion rates of substrates and enzymes and therefore slow release of sugar monomers or dimers might lead to

Table 9.2. Cultivation approaches

Species	Research question	Remarks	Reference
Surface growth			
A. oryzae	Germination, stress tolerance	Agar plate	Sakamoto et al. (2009)
A. oryzae	Chromosome minimization	Agar plate	Jin et al. (2010)
A. oryzae	Sexual reproduction	Agar plate	Wada et al. (2012)
A. oryzae	Development	Agar plate	Ogawa et al. (2010)
A. oryzae	Secretion stress	Agar plate	Wang et al. (2010)
A. oryzae	Spatial gene expression	Agar plate (square)	Masai et al. (2006)
A. oryzae	Liquid vs. surface growth	Agar plate and membrane surface	Imanaka et al. (2010)
A. oryzae	Stress tolerance	Solid phase	Sakamoto et al. (2008)
A. oryzae	Non-syntenic block expression	Solid phase	Tamano et al. (2008)
A. oryzae	Hydrolytic gene expression	Solid phase	Maeda et al. (2004)
A. niger	Spatial gene expression	Ring-plate	Levin et al. (2007)
A. niger	Hyphal heterogeneity	Agar plate sandwich	de Bekker et al. (2011a)
T. reesei	Development	Agar plate	Metz et al. (2011)
T. reesei	Development	Agar plate	Chen et al. (2012)
P. chrysogenum	Development	Agar plate	Sigl et al. (2011)
P. chrysogenum	NHEJ pathway	Agar plate	Hoff et al. (2010b)
P. chrysogenum	Velvet complex	Agar plate	Hoff et al. (2010a)
A. niger	Bacterial/fungal confrontation	Agar plate	Mela et al. (2011)
Shake flask			
A. oryzae	Proteases	–	Kimura et al. (2008)
A. oryzae	Liquid vs. surface growth	–	Noguchi et al. (2009)
A. oryzae	Secretion stress	–	Ohno et al. (2011)
A. oryzae	Kojic acid production	–	Terabayashi et al. (2010)
A. oryzae	Hypoxic stress	–	Terabayashi et al. (2012)
A. oryzae	Secretion stress	–	Wang et al. (2010)
A. niger	Cellulase and hemicellulase expression	Mycelial transfer	De Souza et al. (2011)
A. oryzae	Carbon starvation	Mycelial transfer	Maeda et al. (2004)
A. oryzae	XlnR regulon	Mycelial transfer	Noguchi et al. (2009)
A. oryzae	ManR regulon	Mycelial transfer	Ogawa et al. (2012)
A. oryzae	Liquid vs. surface growth	Mycelial transfer	Tamano et al. (2008)
A. niger	Secretion stress	–	Guillemette et al. (2007)
A. niger	Secretion stress	–	MacKenzie et al. (2005)
A. niger	EST annotation	–	Semova et al. (2006)
A. niger	Gene coexpression analysis	–	Van den Berg et al. (2010)
A. niger	Pectin degradation	Mycelial transfer	Martens-Uzunova et al. (2006)
A. niger	InuR regulon	Mycelial transfer	Yuan et al. (2007)
A. niger	Alpha-glucan active enzymes	Mycelial transfer	Yuan et al. (2008)
A. terreus	Lovastatin production	–	Askenazi et al. (2003)
A. niger	Plant biomass degradation	–	Delmas et al. (2012)
T. reesei	Carbohydrate active enzymes	–	Häkkinen et al. (2012)
T. reesei	LAE1 regulation	–	Karimi-Aghcheh et al. (2013)
T. reesei	Light responsiveness	–	Tisch et al. (2011)
Bioreactor			
A. niger	pH response	Batch	Andersen et al. (2009)
A. niger	Strain comparison	Batch	Andersen et al. (2011)
A. niger	Secretion stress	Batch	Carvalho et al. (2012)
A. niger	Morphology	Batch	Kwon et al. (2013)
A. niger	Antifungals, morphology	Batch	Meyer et al. (2007)
A. niger	Morphology	Batch	Meyer et al. (2009)
A. niger	Carbon starvation	Batch	Nitsche et al. (2012)
A. niger	pH response	Batch and chemostat	Poulsen et al. (2012)
A. niger	Gene coexpression analysis	Batch	van den Berg et al. (2010)

(continued)

Table 9.2. (continued)

Species	Research question	Remarks	Reference
A. niger	Analysis of variance	Batch	van der Veen et al. (2009)
A. niger	Secretion stress	Chemostat	Guillemette et al. (2007)
A. niger	Secretion stress	Chemostat	Jørgensen et al. (2009)
A. niger	Secretion stress	Chemostat	Kwon et al. (2012)
A. niger	Protein production	Fed batch	Jacobs et al. (2009)
A. niger	General gene expression	Fed batch	Pel et al. (2007)
A. niger	Carbon starvation	Retentostat	Jørgensen et al. (2010)
A. niger, *A. oryzae*, *A. nidulans*	Xylose utilization	Batch	Andersen et al. (2008b)
A. niger, *A. oryzae*, *A. nidulans*	Glycerol utilization	Batch	Salazar et al. (2009)
A. oryzae, *A. niger*	Plant polysaccharide degradation	Batch	Coutinho et al. (2009)
A. oryzae, *A. niger*	Maltose utilization	Batch	Vongsangnak et al. (2009)
A. oryzae	Secretion stress	Batch	Vongsangnak et al. (2011)
A. tereus	Itaconic acid production	Batch	Li et al. (2011)
P. chrysogenum	Degeneration of penicillin production	Chemostat	Douma et al. (2011)
P. chrysogenum	Oxalacetase and oxalate formation	Chemostat	Gombert et al. (2011)
P. chrysogenum	Phenylacetic acid metabolism	Chemostat	Harris et al. (2009a)
P. chrysogenum	Adipic acid metabolism	Chemostat	Harris et al. (2009b)
P. chrysogenum	Adipic acid activation	Chemostat	Koetsier et al. (2010)
P. chrysogenum	NHEJ pathway	Chemostat	Snoek et al. (2009)
P. chrysogenum	Phenylacetic acid metabolism	Chemostat	Van den Berg et al. (2008)
P. chrysogenum	Engineering β-oxidation	Chemostat	Veiga et al. (2012a)
P. chrysogenum	Velvet complex	Chemostat	Veiga et al. (2012b)
P. chrysogenum	Phenylalanin metabolism	Chemostat	Veiga et al. (2012c)
T. reesei	Secretion stress	Batch and chemostat	Arvas et al. (2006)
T. reesei	Protein production	Chemostat	Arvas et al. (2011)
T. reesei	Secretion stress	Chemostat	Arvas et al. (2010)
T. reesei	Hypoxic stress	Chemostat	Bonaccorsi et al. (2006)
T. reesei	Catabolite repression	Chemostat	Portnoy et al. (2011)
T. reesei	Expression of hydrolases	Fed batch	Foreman et al. (2003)

NHEJ Non-homologous end-joining pathway

derepression of hydrolase genes and therefore high expression levels on solid medium.

Syntenic analysis of *A. oryzae*, *A. nidulans* and *A. fumigatus* has shown the presence of syntenic and non-syntenic blocks distributed across the genome of *A. oryzae* (Machida et al. 2005). Interestingly, Tamano et al. (2008) have demonstrated that *A. oryzae* genes present in non-syntenic blocks (regions specific for *A. oryzae*), including those encoding secreted hydrolases, were specifically upregulated during solid-state cultivations, whereas genes present in syntenic blocks were not. As proposed by the authors, such global transcriptional trends could potentially hint at epigenetic regulatory mechanisms, which might be important for growth of *A. oryzae* on solid substrates.

Low diffusion rates in surface cultures result in gradients, which in turn induce spatial differentiation and **zonal gene expression** as observed for *A. oryzae* (Masai et al. 2006) and *A. niger* (Levin et al. 2007). It is assumed that substrate depletion in the central part of the mycelial colony in combination with exposure of hyphae to air induces asexual development. **Conidiation** has been studied by transcript profiling of *P. chrysogenum* (Sigl et al. 2011), *A. oryzae* (Ogawa et al. 2010), *A. niger* (Jørgensen et al. 2010; Nitsche et al. 2012) and *T. reesei* (Metz et al. 2011). The expression data have provided additional evidence for the conservation of the central regulatory pathway, which includes the transcription factor BrlA in *A. nidulans*, *P. chrysogenum*, *A. oryzae* and

A. niger. Furthermore, the transcriptional regulator StuA has been shown to link conidiation and asexual development in *P. chrysogenum* (Sigl et al. 2011). Besides conidiation, nutrient limitation generally induces transcription of proteases and glycosyl hydrolases in *A. niger* (Jørgensen et al. 2010; Nitsche et al. 2012), and (hemi-)cellulases in *T. reesei* (Metz et al. 2011).

Submerged cultivation of filamentous fungi can be realized in shake flasks or in stirred-tank reactors. The latter are technically sophisticated and allow precise control and monitoring of the physiological state of the cultures, including temperature, pH, dissolved oxygen tension and off-gas composition. Online analysis and control of physiological data help to considerably improve the reliability of replicate cultures, thus improving the reproducibility of transcriptomics data. As opposed to shake flasks, sampling can be synchronized for replicate cultures based on physiological parameters, including titrant addition, dissolved oxygen or off-gas composition (Hrdlicka et al. 2004; Nitsche et al. 2012). Still, a considerable number of fungal cultivations are performed in **shake flasks** (Table 9.2), probably due to the high investment cost for bioreactors, the requirement for trained staff and the experimental complexity. Under certain circumstances, though, the lower reproducibility of shake flask cultivations can be ignored for transcriptomics analyses. For example, if one is only interested in the identification of major transcriptional trends, that is, the identification of genes displaying large fold changes (effect size) in their expression levels. **Mycelial transfer** is an experimental design commonly applied in many shake flask setups. It describes the transfer of mycelial biomass harvested from a preculture into new media containing, for example, an alternative growth-limiting carbon source.

Besides ensuring highly reproducible cultivation conditions, **bioreactor technology** allows continuous cultivations of filamentous fungi, including chemostat and retentostat protocols. With **chemostat cultivations**, the specific growth rate of different strains can be controlled using the dilution rate, thereby allowing the analysis of product formation or stress responses at defined specific growth rates (Table 9.2). Thus, direct physiological responses related to the treatment itself can be decoupled from global physiological responses related to differences in the specific growth rate. Furthermore, the steady-state condition achieved during chemostat cultivation confers time independence and allows repetitive harvesting of identical biological samples. Chemostat cultivations have been used to elucidate the effect of the cell density on the specific protein production rate in *T. reesei*. Here, a combinatorial experimental design was followed that included two specific growth rates (0.03 h^{-1} and 0.06 h^{-1}) and two different substrate concentrations (10 and 40 gL^{-1}) (Arvas et al. 2011). Correlation analysis has shown that the expression of genes involved in major biosynthetic functions correlated negatively with the specific protein production rate. An efficient protein production at low specific growth rates can be explained by the fact that the energy is primarily used for protein biosynthesis. The authors have hypothesised that protein production is regulated by the flux through upper glycolysis or the citric acid cycle, which in turn is controlled by growth rate and cell density.

Retentostat cultivations are carbon-limited chemostat cultivations with cell retention and lead to a biomass-specific carbon and energy limitation over time that eventually adjust extremely low specific growth rates approaching nearly zero (Jørgensen et al. 2010). As shown for *A. niger*, retentostat cultivations represent a well-suited alternative for surface growth conditions as they can be used to stimulate asexual development under submerged conditions. Hence, the underlying genetic network controlling spore formation and its link to secondary metabolism can for the first time be studied under defined and reproducible conditions. It has further been shown that retentostat cultivation might be an interesting tool to specifically produce small cysteine-rich secreted proteins, as the transcriptomic data of *A. niger* cultivated under retentostat conditions indicated an adaptation of *A. niger* and its secretion machinery to such proteins (Jørgensen et al. 2011).

III. Filamentous Fungi as Cell Factories for Enzyme Production

Owing to their saprophytic lifestyle, filamentous fungi like *A. niger*, *A. oryzae*, *A. terreus* and *T. reesei* have high secretion capacities, which is exploited for the production of industrially important hydrolases. The classical **secretory route** for extracellular proteins starts with their translocation into the endoplasmic reticulum (ER) via a transmembrane protein complex, referred to as the translocon (Römisch 1999). Inside the ER, folding of proteins takes place with the help of chaperones and foldases, including the protein disulfide isomerase PdiA, the binding protein BipA and calnexin (ClxA) (Määttänen et al. 2010). During their traffic through the ER, most proteins become glycosylated (van den Brink et al. 2006), whereby oligomeric sugars become attached either to asparagine (N-glycosylation) or to serine/threonine (O-glycosylation) residues. Proteins leave the ER in COPII-coated vesicles and are transported to the Golgi complex, where they undergo additional posttranslational modifications before being packed into vesicles that finally fuse with the cytoplasmic membrane at the hyphal tip to release the proteins into the medium (Taheri-talesh et al. 2008; Wösten et al. 1991).

A. Secretion Stress

The expression of heterologous proteins or overexpression of homologous proteins can lead to an overload of the ER and an accumulation of misfolded proteins. To counteract this and to prevent cell death, two conserved mechanisms have been evolved in eukaryotes, which have been excellently reviewed by Määttänen et al. (2010). In brief, the **unfolded protein response** (**UPR**) induces expression of foldases and chaperones to facilitate proper protein folding, whereas the **ER-associated degradation** (**ERAD**) pathway targets misfolded proteins to the cytosol for proteasome-mediated degradation. An additional cellular strategy has been recently suggested for *A. niger* and *T. reesei* based on the observation that expression levels of certain endogenous secretory proteins decrease under secretion stress conditions. This phenomenon is referred to as **repression under secretion stress** (**RESS**) and has been first described for *T. reesei* (Pakula et al. 2003; Saloheimo and Pakula 2012) and similarly proposed for *A. niger* (Carvalho et al. 2012; Kwon et al. 2012). Based on the transcriptomic signatures obtained for *A. niger* forced to overexpress proteins or stressed with compounds inhibiting the function of ER-resident proteins, it has been speculated that *A. niger* attempts to avoid overload of the secretory machinery by transcriptionally downregulating those genes whose function is less important for growth and survival under a given circumstance (Carvalho et al. 2012; Kwon et al. 2012).

Several approaches have been followed to induce and study secretion stress in filamentous fungi, including treatment with ER-specific stress agents, the introduction of specific mutations or overexpression of homologous and heterologous proteins. Stress agents that have commonly been used to provoke secretion stress in filamentous fungi include dithiothreitol (DTT), a reducing compound that prevents the formation of disulfide bridges, and tunicamycin, which inhibits N-glycosylation of secretory proteins. One such example is the study by Guillemette et al. (2007), who treated *A. niger* with DTT and tunicamycin in shake flask cultures to uncover its genome-wide transcriptional responses to secretion stress. Among the strongly induced genes were genes predicted to encode proteins functioning in the secretory route such as foldases (e.g. PdiA) and chaperones (e.g. BipA). In parallel, expression levels of several secreted proteins were downregulated, hinting at the RESS phenomenon. In a similar study, Wang et al. (2010) used DTT to induce and investigate ER stress in *A. oryzae* when cultivated under submerse and solid-state conditions. The transcriptomic response of *A. oryzae* was captured using RNA-seq, which revealed that more UPR target genes were upregulated by DTT when *A. oryzae* was cultivated under solid-state conditions in comparison to submerged cultivation. It was pro-

posed that the UPR is more efficient under surface growth than during submerged cultivation, suggesting that the UPR of *A. oryzae* constitutes a culture-condition-dependent biological process. Considering the generally increased hydrolase expression and activities under surface growth conditions of *A. oryzae* (Imanaka et al. 2010; Maeda et al. 2004), these observations suggest the UPR can be considered as a homeostatic cellular mechanism that allows the fungus to adapt the capacity of its secretory route to the environmental condition. In agreement, chemostat-based cultivation of *A. niger* at identical growth rates but on two different carbon sources (maltose and xylose) revealed that about 90 secretory pathway genes, including UPR and ERAD genes, were expressed more when maltose was used as a carbon source, resulting in a three times higher biomass-specific protein production rate (Jørgensen et al. 2009). Overall, these observations suggest that filamentous fungi are able to flexibly adapt the protein flux to the actual present capacity of their secretory machinery, thereby preventing protein overload and potential detrimental effects risking cell survival.

An **integrative systems biology approach** has been followed to identify possible limiting steps for α-amylase production by a global transcriptomic comparison of bioreactor batch cultures of *A. oryzae* wild-type and α-amylase overproduction strains. The transcriptomic data have been integrated with putative (orthology-based) protein-protein interactions, metabolic networks and flux calculations, leading to the identification of reporter metabolites and targets in gene expression regulation. Besides identification of the UPR-specific transcriptional factor HacA as an important regulatory protein in response to the α-amylase overexpression, amino acid starvation and Gcn2p/CpcC-mediated cross-pathway control of amino acid biosynthesis have also been suggested to be important for improved α-amylase production. For *A. niger*, Kwon et al. (2012) determined the transcriptomic fingerprint of glycoamylase overexpression in a chemostat setting where the reference and the overproducing strain were cultivated at identical specific growth rates. The physiologically well-defined steady-state conditions allowed the identification of about 1,500 differentially expressed genes by applying an extremely strict cutoff for the statistical significance (false discovery rate, FDR <0.005). FetGOat (Fisher's exact text Gene Ontology annotation tool; http://www.broadinstitute.org/fetgoat/index.html; Nitsche et al. 2011) was used for enrichment analyses of Gene Ontology terms (Ashburner et al. 2000) and identified four higher-order categories that were upregulated in the glucoamylase overexpressing strain: ER translocation, protein glycosylation, vesicle transport and ion homeostasis. In total, 130 genes were predicted to directly function in protein secretion.

A common strategy to overcome **bottlenecks in heterologous protein secretion** and to improve the production of heterologous proteins is their fusion to a homologous secreted carrier protein (Gouka et al. 1997; Tsuchiya and Nagashima 1994). In *A. oryzae*, a comparative transcriptomic study has been conducted comparing the gene expression profiles of strains expressing carrier-fused and non-carrier-fused bovine chymosin to a non-expressing control strain. Concomitantly with a twofold increase of extracellular chymosin compared to the non-carrier-fused protein, the carrier-fused protein accumulated intracellularly and significantly induced transcription of UPR genes. Expression of the non-carrier-fused chymosin, however, only slightly induced UPR genes (Ohno et al. 2011). A search for lead genes most important for strengthened protein production has been followed for *A. niger* (Jacobs et al. 2009). The authors assumed that genes being coregulated at both the transcriptome and the proteome levels constitute key candidates for strain improvements. Furthermore, the authors made a distinction between overexpression of homologous and heterologous proteins. Although both challenge the folding capacity of the secretion machinery, it is more likely a quantitative challenge in the case of homologous proteins but a quantitative and qualitative one in the case of heterologous proteins. Coregulation on the transcriptome and proteome levels has been shown for genes/proteins involved in protein folding and responses to reactive oxygen species. A β-glucuronidase reporter strain has

been applied to validate two leads, demonstrating that *doaA* deletion in combination with *sstC* overexpression can indeed increase β-glucuronidase production (Jacobs et al. 2009).

The regulatory mechanisms controlling UPR induction are well conserved in eukaryotes, and the central role of the **transcription factor HacA** is well established in yeast (Sidrauski and Walter 1997), filamentous fungi (Guillemette et al. 2007; Mulder et al. 2006; Ohno et al. 2011; Valkonen 2004) and mammalian cells (Calfon et al. 2002). Activation of HacA is mediated by splicing of a small unconventional intron from the *hacA* mRNA, which normally prevents its own translation. Carvalho et al. (2012) applied this knowledge to generate a defined *A. niger* mutant that carried a *hacA* gene at its endogenous locus that lacked the intron sequence thus being constitutively expressed. Bioreactor-based comparative transcriptomics clearly demonstrated a concerted induction of HacA target genes and of genes predicted to function in protein glycosylation, phospholipid biosynthesis, protein transport and exocytosis in the constitutively active HacA strain. Most interestingly, this activation was paralleled by downregulation of genes belonging to the AmyR regulon, providing an additional hint to the RESS phenomenon. For a condition-independent view of how *A. niger* ensures high-level secretion on the transcriptional level, Kwon et al. (2012) compared the transcriptomic data set of several independent studies performed with *A. niger*: (a) the GlaA overexpression transcriptome (Kwon et al. 2012), (b) the maltose versus xylose chemostat transcriptome (Jørgensen et al. 2009), (c) the constitutively active HacA transcriptome reflecting permanent activation of UPR (Carvalho et al. 2012), and (d) three UPR stress transcriptomes obtained from *A. niger* when stressed with DTT or tunicamycin or forced to express a heterologous protein (Guillemette et al. 2007). This comparative transcriptomic analysis uncovered a core set of 40 genes whose expression is crucial for *A. niger* to cope with secretion stress. This gene set includes ER chaperones and foldases (*prpA*, *clxA*, *lhs1*, *pdiA*, *bipA*, *tigA*); genes important for translocation of secretory proteins into the ER (*sec63*, *sec11*, *sss1*, *spc3*, *sec71*); and genes important for protein glycosylation and COPII-based vesicle trafficking (Kwon et al. 2012).

B. Proteolysis

The efficient proteolytic machinery of filamentous fungi may be considered as advantageous or disadvantageous, depending on the context. **Secreted proteases** themselves are commercially attractive products that are industrially produced by filamentous fungi for food applications or as biocatalysts; however, their activities seriously limit the yield of homologous or heterologous proteins because of product degradation (Archer and Peberdy 1997). Extracellular proteases account for approximately 1 % of the predicted genes of *A. oryzae* (Machida et al. 2005), and their expression often coincides with nutrient limitations as shown for *A. niger* (Nitsche et al. 2012). Knowledge of the transcriptomic or proteomic profiles of nutrient-limited growth phases is extremely valuable for the identification of proteases whose deletions might result in improved product yields. For example, transcriptomic profiling of 132 predicted proteases of *A. oryzae* helped to identify the *nptB* gene encoding a neutral protease II as a potential target. Its deletion indeed improved human lysozyme production in *A. oryzae* (Kimura et al. 2008). This transcriptomic data set was further exploited for the prediction of meaningful deletions of protease genes. Eventually, a quintuple protease disruption strain of *A. oryzae* was generated that displayed 30 % improved lysozyme production (Yoon et al. 2009). In addition to the extracellular proteases, cytoplasmic and vacuolar protein degradation has been suggested to be a bottleneck for secreted enzyme production (Jacobs et al. 2009); however, this has not been proven experimentally so far.

C. Hyphal Heterogeneity

Heterogeneity of mycelial cultures is a well-known phenomenon. Fluorescently labelled

reporter strains have demonstrated the existence of two hyphal subpopulations in the context of protein secretion. One subpopulation is actively secreting proteins, while the other is not (Vinck et al. 2005, 2011). It is thus conceivable that the productivity of a certain process could be improved by increasing the fraction of actively secreting hyphae. Interestingly, heterogeneity within isogenic microbial populations has been discussed as an evolvable trait contributing to the survival of environmental stress conditions (Veening et al. 2008). However, because of technical challenges (i.e. the availability of single-cell analytics), respective single-cell transcriptome studies of microbes have not been published so far. Still, the recently published method for single-cell transcriptomics by de Bekker et al. is an important milestone towards understanding heterogeneity of hyphal gene expression (de Bekker et al. 2011a).

IV. Filamentous Fungi and Plant Biopolymer Degradation

In their natural habitats, filamentous fungi are confronted with homopolymeric (e.g. starch and cellulose) or heteropolymeric (e.g. pectin, hemicellulose and lignin) carbon sources of plant origin. The versatile spectrum of secreted hydrolases required for breakdown of these plant biopolymers constitutes a valuable collection of industrially exploitable enzymes for the food industry (e.g. amylases and pectinases) and biofuel production (e.g. lignocellulolytic enzymes). Unlike starch-based first-generation feedstocks, second-generation feedstocks (i.e. waste plant biomass) do not compete with food or feed supply, are cheaper, and contribute to environmentally friendlier bioprocesses (Rumbold et al. 2009). For example, bioethanol production in Brazil relies so far exclusively on first-generation sugarcane feedstocks (sucrose extracted from sugarcane), and it has been calculated that the use of sugarcane waste for bioethanol production would reduce the required agricultural land by up to 38 % (de Souza et al. 2011).

Transcriptional regulation of fungal plant cell-wall-degrading enzymes has been reviewed comprehensively by Aro et al. (2005). In brief, expression of genes encoding secreted hydrolases is either subjected to CreA-mediated **carbon catabolite repression** or is induced by carbon-specific transcription factors, including AmyR (activator of amylolytic genes) and XlnR (activator of [hemi-]cellulolytic genes) (Tsukagoshi et al. 2001). Carbon catabolite repression in *T. reesei* has been studied by genome-wide transcriptional profiling applying a chemostat-based combinatorial design (Portnoy et al. 2011). CreA wild-type and deletion strains have been propagated at high (0.07 h^{-1}; repressing) and low (0.0025 h^{-1}; non-repressing) specific growth rates, leading to the identification of 250 differentially expressed genes. Clustering defined gene sets repressed (47 %) and induced (29 %) by CreA as well as those regulated independently of CreA (17 %). This study showed that carbon catabolite repression predominantly affected the uptake of (nitrogenous) substrates into the cell. Only a minority of the carbohydrate active enzymes were induced in the Δ*creA* strain, indicating that CreA affects their induction rather than their basal expression levels (Portnoy et al. 2011).

XlnR-mediated regulation is partly conserved in filamentous fungi, as demonstrated by a trispecies microarray analysis for *A. nidulans*, *A. niger* and *A. oryzae* (Andersen et al. 2008b). A set of 23 xylose-responsive genes is conserved across these aspergilli and in situ promoter analysis uncovered a conserved XlnR motif for 22 of the induced genes. These data provided evidence that xylose is sensed as an inducer molecule for the presence of (hemi-) cellulose in the environment (Andersen et al. 2008b). The regulatory role of XlnR was also investigated in *A. oryzae* by comparative genome-wide expression profiling of strains overexpressing or lacking XlnR (Noguchi et al. 2009). XlnR overexpression led to an increase of extracellular xylanolytic and cellulolytic activities and enhanced transcription of 75 genes when compared with the *xlnR* deletion strain.

The three-species microarray approach has also been used to study conserved transcriptional responses of *A. niger* and *A. oryzae* when subjected to maltose, the inducer molecule of AmyR and amylolytic gene expression

(Vongsangnak et al. 2009). This study showed differences in the response to maltose between the two species. For *A. oryzae*, 16 differentially expressed genes were identified, whereas for *A. niger*, no gene showed statistically significant differential expression.

Maltose utilization in *A. oryzae* has been suggested to involve the *MAL* gene cluster, which encodes a maltose permease, a maltase and a cluster-specific transcription factor. For *A. niger*, however, no such *MAL* gene cluster is present in its genome, suggesting that maltose utilization in *A. niger* depends on extracellular cleavage of maltose and subsequent uptake of its glucose monomers, instead of intracellular maltose degradation as proposed for *A. oryzae*. Further insights into the physiology of *A. niger* during maltose and xylose utilization has been provided by a chemostat-based transcriptomic study by Jørgensen et al. (2009). Here, the transcriptomic fingerprint clearly demonstrated induced expression of AmyR-dependent amylolytic genes and XlnR-dependent genes, respectively. Microarray-based transcriptional profiling in combination with functional genomics studies also helped to identify InuR as a transcriptional activator of inulinolytic genes in *A. niger* and ManR as a transcriptional activator of genes important for β-mannan utilization in *A. oryzae* (Ogawa et al. 2012; Yuan et al. 2007).

The genomes of *T. reesei* and *A. niger* have been screened for the presence of predicted putative **carbohydrate active enzymes**. Overall, 228 and 281 enzymes belonging to more than 60 different families as defined by the Carbohydrate Active Enzyme database have been annotated in both fungi (Cantarel et al. 2009; Delmas et al. 2012; Häkkinen et al. 2012). Which of those predicted hydrolases function in response to complex lignocellulose-related substrates is largely unknown. To obtain a glimpse of the polymer-degrading potential of *A. niger*, De Souza et al. (2011) performed mycelial transfer experiments to uncover genes important for the degradation of sugarcane bagasse. Using microarrays in combination with a secretome analysis, it was demonstrated that more than half of the predicted cellulases and hemicellulases were highly expressed during growth on sugarcane bagasse. The lignocellulose-degrading potential of *A. niger* has been studied using RNA-seq by Delmas et al. (2012). Mycelia of *A. niger* were transferred from glucose medium to wheat straw, to which additional glucose was added after some time of cultivation. Comparative RNA-seq profiles determined prior to the transfer and before and after glucose addition allowed the identification of genes under general starvation control, substrate-specific regulation and CreA-mediated carbon catabolite repression, respectively. This study led to two interesting findings. Firstly, about 20 % of the total mRNA of *A. niger* was encoding known or predicted plant cell-wall-degrading enzymes when the fungus was cultivated on wheat straw, demonstrating not only the power but also the investment cost of *A. niger* to utilize complex plant polymers. Secondly, *A. niger* secretes a broad spectrum of different hydrolytic enzymes into the environment when encountering carbon starvation. These enzymes supposedly act as scouts testing for the presence of polysaccharides in the environment. By liberating inducing sugars, the subsequent induction of the majority of hydrolases becomes triggered (Delmas et al. 2012).

The lignocellulolytic degradation potential of *T. reesei* has also been analysed using microarray-based transcriptomics. In doing so, *T. reesei* was cultivated on bagasse, wheat straw, spruce, xylan and cellulose (Häkkinen et al. 2012). Of the 228 predicted carbohydrate active enzymes, 179 were induced by at least one of the substrates tested. Hemicellulosic substrates induced the largest fraction of hydrolases, in contrast to cellulosic substrates.

An example demonstrating how transcriptomics can be used to predict the functions of genes has been given by Yuan et al. (2008). The authors extended the collection of predicted α-glucan active enzymes (which belong to the same glycosyl hydrolase families as those required for starch breakdown) using a hidden Markov model and predicted their functions by phylogenetic clustering and microarray analysis of wild-type and Δ*amyR* strains grown on xylose and maltose. Interestingly, most of the 17 newly presumptive α-glucan active enzymes were not predicted to function

in starch degradation but rather in fungal cell wall remodelling or protein glycosylation.

Pectin is another heterogeneous and abundant plant polymer (~30 % of total cell wall) that can be used as a carbon source during fungal fermentation. Its degradation requires a diverse set of enzymes, many of which are industrially produced by *A. niger*. Genome sequencing uncovered the inventory of its pectinolytic system, and mycelial transfer experiments combined with microarray-based transcriptomics were used to study the pectin degradation network of *A. niger*. Based on its transcriptomic fingerprint when cultivated on sugar-beet pectin, on its main monomeric (galacturonic acid, rhamnose, xylose) and polymeric (polygalacturonic acid) constituents or on non-repressing (sorbitol) and repressing (fructose) carbon sources (Martens-Uzunova and Schaap 2009; Martens-Uzunova et al. 2006; Pel et al. 2007), a conserved galacturonic acid catabolic pathway and a phylogenetically conserved promoter motif were identified, suggesting the existence of a specific pectinolytic regulator in *A. niger* (Martens-Uzunova and Schaap 2008).

V. Filamentous Fungi as Hosts for Metabolite Production

Filamentous fungi are not only extraordinary producers of secreted proteins but also are exploited in biotechnology as efficient cell factories for the production of primary and secondary metabolites such as citric acid (using *A. niger*) and penicillin (using *P. chrysogenum*). Genome-wide metabolic models have been shown to be an excellent tool to improve metabolite production, and a few examples are discussed in this section. These models can be used to calculate maximal theoretical yields, identify putative bottlenecks and generally facilitate the interpretation of transcriptomic data to identify targets for strain engineering (Andersen et al. 2008a). Besides, genome-wide metabolic models and flux calculations can also be used to study protein expression, as exemplarily shown for glucoamylase in *A. oryzae*, where it was predicted that its overexpression results in cellular shortage of four amino acids (Vongsangnak et al. 2011).

A. Primary Metabolites

Industrial **citric acid** production with *A. niger* is an extremely efficient process, yielding up to 95 g citric acid per 100 g sugar (Karaffa and Kubicek 2003). Although such yields can only be achieved with highly optimized production strains under optimal conditions, they are based on the naturally high capacity of *A. niger* to produce organic acids such as oxalic acid, gluconic acid and citric acid. *A. niger* can grow over a wide range of ambient pH (<2 and >8), whereby different pH values affect its organic acid profile of *A. niger* (Andersen et al. 2009; Hesse et al. 2002). To gain insights into the molecular basis of this regulatory mechanism, Andersen et al. (2009) conducted an integrative transcriptomic study by comparing expression profiles of *A. niger* when cultivated at three different pH values (2.5, 4.5 and 6). Clustering of genes with similar expression profiles led to the prediction of several putative transcription factor binding sites, among which was the *A. nidulans* PacC consensus sequence. In addition, genome-scale metabolic modelling provided evidence that citric acid and oxalic acid production have evolved as efficient acidification traits, while gluconic acid production has not. In a follow-up study by Poulsen et al. (2012), a set of knockout strains of putative pH-responsive transcription factors was generated, which helped to identify the oxalic acid repression factor OafA. Deletion of *oafA* led to an 87 % increase of oxalic acid, and a chemostat-based comparative transcriptomic analysis of the deletion strain and its isogenic wild-type identified 241 differentially expressed genes at pH 5. Altogether, the integration of transcriptomic and physiological data indicated that reuse of gluconic acid resulted in an increased flux through the glycolytic but not through the pentose phosphate pathway, thus explaining enhanced oxalic acid production by *A. niger* when depleted for OafA (Poulsen et al. 2012).

Itaconic acid is a multifunctional organic acid that is exploited by the chemical industry as a building block for the production of plastics and synthetic fibres. It is currently produced by *A. terreus* at suboptimal yields, reaching only about 40 % of its theoretical maximum (Li et al. 2011). Interestingly, citric acid is a precursor for itaconic acid, which is not naturally produced by *A. niger*. Li and coworkers (2011) thus aimed to engineer *A. niger* as producer of itaconic acid by modification of its citric acid biosynthetic pathway. The authors used cDNA microarrays to identify *A. terreus* genes important for low and high itaconic acid levels and used the deduced gene set to predict candidate genes potentially involved in itaconic acid metabolism of *A. niger*. Constitutive expression of a candidate gene of *A. terreus* (*cis*-aconitate decarboxylase) in *A. niger* resulted indeed in itaconic acid production – although only at low levels, but demonstrating a proof of concept of this approach.

B. Secondary Metabolites

Only three enzymes are required for the synthesis of a secondary metabolite that has revolutionized medicine: **penicillin G** (Fleming 1929). Classical strain improvement and metabolic engineering of *P. chrysogenum* considerably increased the yields of penicillin and its derivative **cephalosporin** (Thykaer and Nielsen 2003). Shortly after publication of the *P. chrysogenum* genome sequence (Berg et al. 2008), a number of chemostat-based transcriptomic studies were published aiming to understand the molecular basis of penicillin production. Two independent chemostat-based studies were performed to identify genes important for high-level β-lactam biosynthesis and to understand the cellular fate of the fed side chains phenylacetic acid (PAA) and adipic acid (ADA) (Harris et al. 2009a, b). Basically, fed side chains can either become incorporated into the β-lactam intermediate or become catabolized by *P. chrysogenum* for energy generation. Penicillin-producing strains and non-producing control strains were cultivated in chemostats, in both the presence or the absence of the respective side chain. PAA feeding resulted in transcriptional induction of the homogentisate pathway, which is thus a likely candidate route for PAA catabolism. The expression data also suggested that the amino acid pools drain during penicillin G production (Harris et al. 2009a). Similarly, the transcriptomic data suggested that most of the fed ADA becomes catabolised via β-oxidation instead of being used for cephalosporin synthesis (Harris et al. 2009b). Veiga et al. (2012a) have therefore deleted two genes of the β-oxidation pathway (a mitochondrial acyl-CoA [coenzyme A] oxidase and a peroxisomal acyl-CoA dehydrogenase) and could thereby improve cephalosporin yields by a factor of 1.6 and 3.7, respectively. In another study, the importance of an acyl-CoA ligase for incorporation of ADA into the β-lactam ring was demonstrated. The respective gene was identified through comparative transcriptomics of *P. chrysogenum* chemostat cultures, which were fed with different side chains (ADA, PAA and without). Deletion of the predicted candidate gene caused a significant drop in ADA consumption and cephalosporin production but did not affect penicillin production (Koetsier et al. 2010).

VeA and LaeA are proteins of the heteromeric **velvet complex** and function as global regulators of secondary metabolism in *Aspergillus* and *Penicillium* (Bayram et al. 2008; Brakhage and Schroeckh 2011; Hoff et al. 2010a). Interestingly, conflicting results have been reported for their regulatory roles in penicillin production. During surface growth and in submerged batch cultures, penicillin titres were strongly reduced in *veA* or *laeA* deletion strains in agreement with a transcriptional downregulation of two genes belonging to the penicillin cluster (Hoff et al. 2010a). However, in glucose-limited chemostat cultures, neither a transcriptional downregulation of penicillin cluster genes nor a significant decrease of penicillin titres has been observed for both strains when compared with the wild-type strain (Veiga et al. 2012b). This suggests that the regulatory role of both LeaA and VeA differs depending on the growth conditions, a conclusion supported by the observation that the number of transcripts affected by deletion of

veA is about ten times larger during surface growth when compared with glucose-limited chemostat cultures. Interestingly, further evidence for such context-dependent regulatory roles of LaeA comes from microarray studies investigating LaeA deletion and overexpression strains of *T. reesei*. The studies showed that besides positively regulating nonribosomal peptide synthethases and polyketide synthases (Karimi-Aghcheh et al. 2013), LaeA is essential for cellulase expression in *T. reesei* (Karimi-Aghcheh et al. 2013).

Another commercially exploited fungal secondary metabolite is **kojic acid**. This compound is used in Asia as a skin lightener and as an additive to prevent browning of food. Kojic acid was discovered in the beginning of the twentieth century and is industrially produced by *A. oryzae*, although the knowledge about its biosynthesis was limited for a long time (Terabayashi et al. 2010). Comparative transcriptomics of kojic acid producing and non-producing conditions suggested nine candidate genes with unknown function. Disruption of two candidate genes resulted in a loss of kojic acid production. Interestingly, both genes are located close to each other in the genome, hinting at the existence of a kojic acid gene cluster. An early publication using genome-wide transcript profiling of *A. terreus* described an association study of transcript and metabolite profiles providing mechanistic insight into lovastatin biosynthesis and identifying promoters for use in reporter strains to improve selection of mutants with enhanced production levels (Askenazi et al. 2003).

In general, many more secondary metabolites can be potentially produced by filamentous fungi and commercially exploited. As secondary metabolite genes are often associated in gene clusters, their expression can easily be induced (Brakhage 2013). In addition, the number of putative secondary metabolite clusters exceeds by far the number of known secondary metabolites as the vast majority of these gene clusters are silent under laboratory conditions (Brakhage and Schroeckh 2011). One of the strategies to activate such **cryptic secondary metabolite clusters** is the one-strain-many-compounds (OSMAC) approach, which tries to activate silent genes by varying cultivation conditions, including pH, nutrient source and nutrient concentration (Bode et al. 2002; Brakhage and Schroeckh 2011). For example, transient induction of the putative *A. niger* fumonisin gene cluster was uncovered by transcriptomics when the strains encountered carbon limitation (Jørgensen et al. 2010; Nitsche et al. 2012). Similarly, the expression of six putative secondary metabolite clusters has been shown to be pH dependent in *A. niger* (Andersen et al. 2009).

VI. Fungal Morphology and Productivity

Polarised growth is a defining attribute of filamentous fungi because phases of apolar growth are restricted to sporulation and spore swelling. Immediately after spore swelling, a polarity axis becomes established and maintained, leading to the outgrowth and elongation of a germ tube. Apical and subapical branches are formed through splitting of the existing polarity axis at hyphal tips and the initiation of new polarity axes at lateral regions, respectively. The branching frequency directly correlates with the number of hyphal tips and is an important factor defining the micromorphology of filamentous fungi. It in turn influences the macromorphology of filamentous fungi during submerged cultivation, which can be pelleted or dispersed as the two extremes. Together, micro- and macromorphologies have substantial impact on submerged fungal fermentations by affecting the rheology and productivity (Meyer et al. 2009; Papagianni 2004; Papagianni and Mattey 2006).

To improve fungal morphology for industrial bioprocesses, the underlying genetic network was investigated in *A. niger* using comparative transcriptomics approaches. Different strategies were followed to induce **hyperbranching phenotypes**. For example, antifungal compounds (caspofungin and fenpropimorph) were applied at sublethal concentrations, which increased the branching frequency of germlings (Meyer et al. 2007). The transcriptomic finger-

print deduced suggested that, besides antifungal-specific responses, the TOR signalling pathway and the cell wall integrity (CWI) pathway were involved in the formation of new branches.

In another approach, the transcriptome of the dichotomous branching mutant *ramosa-1* was compared with its respective wild-type strain (Meyer et al. 2009). Bioreactor-based gene expression profiling revealed 136 differentially expressed genes, and a model for polarity control and apical branching was reconstructed, predicting that at least four signalling cascades are part of the underlying genetic network in *A. niger*: TOR signalling, CWI signalling, calcium signalling and phospholipid signalling.

Additional support for the involvement of these pathways was gained by a recently published comparative transcriptomic study in *A. niger* (Kwon et al. 2013). First, Kwon et al. rationally induced a hyperbranching and an apolar growing phenotype by targeting RacA, the major Rho GTPase controlling actin dynamics in *A. niger*. Loss of RacA function provoked hyperbranching, whereas dominant activation of RacA caused loss of polarised growth in *A. niger* (Kwon et al. 2011). The transcriptomic signatures of these mutants were compared with the transcriptomic fingerprint of the dichotomous branching mutant *ramosa-1* (Kwon et al. 2013). This analysis confirmed the hypothesis that polar tip growth in *A. niger* is most likely orchestrated by the activities of several signalling pathways, including phospholipid signalling, TOR signalling, calcium signalling and CWI signalling.

VII. What More Is to Come?

The -omics technologies have revolutionized biological sciences. The vast amount and diversity of accumulating data constitutes a challenge but offers even greater possibilities. Multidisciplinary collaborations linking biologists, mathematicians, chemists and engineers have brought forward interesting studies that have deepened our understanding of fungal biology and physiology. Functional annotations as applied in many of the cited publications can be considered a prerequisite for effective and unbiased analysis of -omics data. Systems-level investigations of microbial cell factories are facilitated by the integration of different -omics data types, for example, as demonstrated by the use of genome-wide metabolic models for transcriptome data analysis (Andersen et al. 2008a; Vongsangnak et al. 2011). Other promising approaches are association studies that integrate transcript profiles with metabolome profiles (Askenazi et al. 2003), transcript profiles with proteomics data (Jacobs et al. 2009) or coregulation studies and physical association studies of genes for the identification of putative gene and metabolite clusters (Andersen et al. 2009).

Furthermore, the prediction of biologically relevant modules within gene coexpression networks and the search for enriched putative transcription factor binding sites constitutes a reasonable approach for understanding higher-order regulatory structures (van den Berg et al. 2010; Meyer et al. 2009). In addition, mapping of putative transcription factors to such modules might be considered a promising approach to predict. Their functions, as exemplarily shown in Fig. 9.1. Construction of gene coexpression networks is especially interesting because of accumulating transcriptomic data that can easily be combined and screened for underlying networks, without the need to conduct wet experiments. Even inferring putative genome-wide interactome networks has become feasible and has been proven to be of value for identifying putative regulatory key players (Vongsangnak et al. 2011). Hence, an exciting period of fungal systems biology research is awaiting us that will considerably improve our understanding of fungal physiology and lead to new concepts for rational improvement of fungal cell factories.

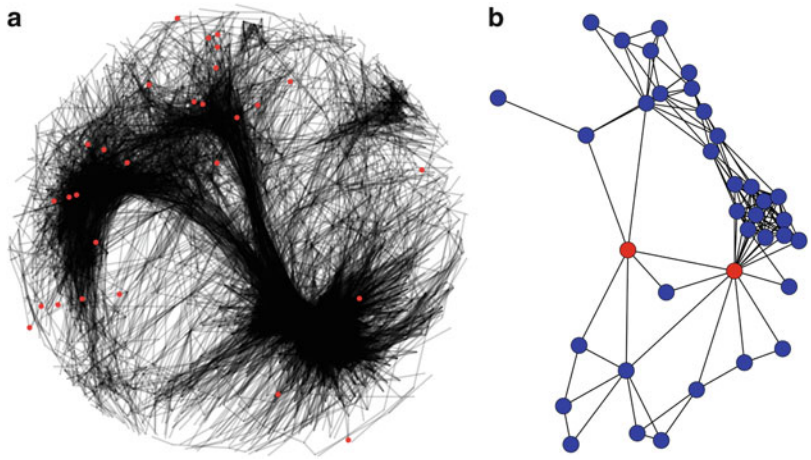


Fig. 9.1. Mapping of putative transcription factors to gene coexpression networks. (**a**) Complete network constructed using all microarray data publically available for *A. niger* (platform: GPL6758) in December 2009. Mapped transcription factors are highlighted in *red*. (**b**) Subnetwork extracted with the Cytoscape plug-in MCODE (Shannon et al. 2003). Carbon catabolism is associated with the subnetwork, as indicated by enrichment analysis of Gene Ontology annotations. One of the two (putative) transcription factors is AmyR; the function of the second transcription factor is unknown, but its expression pattern indicates a possible role in carbon catabolism (guilty by association) (Nitsche, Ram, Meyer, unpublished data)

References

Ahmed FE (2006) Microarray RNA transcriptional profiling: part i. Platforms, experimental design and standardization. Expert Rev Mol Diagn 6:535–550. doi:10.1586/14737159.6.4.535

Andersen MR, Nielsen ML, Nielsen J (2008a) Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol 4:178. doi:10.1038/msb.2008.12

Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J (2008b) A trispecies Aspergillus microarray: comparative transcriptomics of three *Aspergillus* species. Proc Natl Acad Sci U S A 105:4387–4392. doi:10.1073/pnas.0709964105

Andersen MR, Lehmann L, Nielsen J (2009) Systemic analysis of the response of *Aspergillus niger* to ambient pH. Genome Biol 10:R47. doi:10.1186/gb-2009-10-5-r47

Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJI, Culley D, Thykaer J, Frisvad JC, Nielsen KF, Albang R, Albermann K, Berka RM, Braus GH, Braus-Stromeyer SA, Corrochano LM, Dai Z, van Dijck PWM, Hofmann G, Lasure LL, Magnuson JK, Menke H, Meijer M, Meijer SL, Nielsen JB, Nielsen ML, van Ooyen AJJ, Pel HJ, Poulsen L, Samson RA, Stam H, Tsang A, van den Brink JM, Atkins A, Aerts A, Shapiro H, Pangilinan J, Salamov A, Lou Y, Lindquist E, Lucas S, Grimwood J, Grigoriev IV, Kubicek CP, Martinez D, van Peij NNME, Roubos JA, Nielsen J, Baker SE (2011) Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21:885–897. doi:10.1101/gr.112169.110

Archer DB, Peberdy JE (1997) The molecular biology of secreted enzyme production by fungi. Crit Rev Biotechnol 17:273–306

Aro N, Pakula T, Penttilä M (2005) Transcriptional regulation of plant cell wall degradation by filamentous fungi. FEMS Microbiol Rev 29:719–739. doi:10.1016/j.femsre.2004.11.006

Arvas M, Pakula T, Lanthaler K, Saloheimo M, Valkonen M, Suortti T, Robson G, Penttilä M (2006) Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. BMC Genomics 7:32. doi:10.1186/1471-2164-7-32

Arvas M, Haiminen N, Smit B, Rautio J, Vitikainen M, Wiebe M, Martinez D, Chee C, Kunkel J, Sanchez C, Nelson MA, Pakula T, Saloheimo M, Penttilä M, Kivioja T (2010) Detecting novel genes with sparse arrays. Gene 467:41–51. doi:10.1016/j.gene.2010.07.009

Arvas M, Pakula T, Smit B, Rautio J, Koivistoinen H, Jouhten P, Lindfors E, Wiebe M, Penttilä M, Saloheimo M (2011) Correlation of gene expression and protein production rate – a system wide study. BMC Genomics 12:616. doi:10.1186/1471-2164-12-616

Ashburner M, Ball C, Blake J, Botstein D (2000) Gene Ontology: tool for the unification of biology. Nature 25:25–30

Askenazi M, Driggers EM, Holtzman DA, Norman TC, Iverson S, Zimmer DP, Boers M-E, Blomquist PR, Martinez EJ, Monreal AW, Feibelman TP, Mayorga ME, Maxon ME, Sykes K, Tobin JV, Cordero E, Salama SR, Trueheart J, Royer JC, Madden KT (2003) Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. Nat Biotechnol 21:150–156. doi:10.1038/nbt781

Baginsky S, Hennig L, Zimmermann P, Gruissem W (2010) Gene expression analysis, proteomics, and network discovery. Plant Physiol 152:402–410. doi:10.1104/pp. 109.150433

Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon N-J, Keller NP, Yu J-H, Braus GH (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science (New York, NY) 320:1504–1506. doi:10.1126/science.1155888

Bode HB, Bethe B, Höfs R, Zeeck A (2002) Big effects from small changes: possible ways to explore nature's chemical diversity. Chembiochem 3:619–627. doi:10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9

Bonaccorsi E, Ferreira A (2006) Transcriptional response of the obligatory aerobe *Trichoderma reesei* to hypoxia and transient anoxia: implications for energy production and survival in the absence. Biochemistry 45:3912–3924

Bourgon R, Gentleman RC, Huber W (2010) Independent filtering increases detection power for high-throughput experiments. Proc Natl Acad Sci U S A 107:9546–9551. doi:10.1073/pnas.0914005107, 0914005107 [pii]

Brakhage A (2013) Regulation of fungal secondary metabolism. Nat Rev Microbiol 11:21–32. doi:10.1038/nrmicro2916

Brakhage A, Schroeckh V (2011) Fungal secondary metabolites—strategies to activate silent gene clusters. Fungal Genet Biol 48:15–22. doi:10.1016/j.fgb.2010.04.004

Calfon M, Zeng H, Urano F, Till J (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. Nature 415:1–6

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238. doi:10.1093/nar/gkn663

Carvalho N, Jørgensen TR, Arentshorst M, Nitsche BM, van den Hondel CAMJJ, Archer DB, Ram AF (2012) Genome-wide expression analysis upon constitutive activation of the HacA bZIP transcription factor in *Aspergillus niger* reveals a coordinated cellular response to counteract ER stress. BMC Genomics 13:350. doi:10.1186/1471-2164-13-350

Chen C, Kuo H, Tung S, Hsu P (2012) Blue light acts as a double-edged sword in regulating sexual development of *Hypocrea jecorina* (*Trichoderma reesei*). PLoS One 7:9. doi:10.1371/journal.pone.0044969

Coutinho PM, Andersen MR, Kolenova K, VanKuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, Van Solingen P, Pakula T (2009) Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol 46:S161–S169. doi:10.1016/j.fgb.2008.07.020

De Bekker C, Bruning O, Jonker MJ, Breit TM, Wösten HB (2011a) Single cell transcriptomics of neighboring hyphae of *Aspergillus niger*. Genome Biol 12:R71. doi:10.1186/gb-2011-12-8-r71

De Bekker C, van Veluw GJ, Vinck A, Wiebenga LA, Wösten HAB (2011b) Heterogeneity of *Aspergillus niger* microcolonies in liquid shaken cultures. Appl Environ Microbiol 77:1263–1267. doi:10.1128/AEM.02134-10

De Souza WR, de Gouvea PF, Savoldi M, Malavazi I, de Souza Bernardes L a, Goldman MHS, de Vries RP, de Castro Oliveira JV, Goldman GH (2011) Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. Biotechnol Biofuels 4:40. doi:10.1186/1754-6834-4-40

Delmas S, Pullan ST, Gaddipati S, Kokolski M, Malla S, Blythe MJ, Ibbett R, Campbell M, Liddell S, Aboobaker A, Tucker GA, Archer DB (2012) Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. PLoS Genet 8:e1002875. doi:10.1371/journal.pgen.1002875

Douma RD, Batista JM, Touw KM, Kiel JAKW, Krikken AM, Zhao Z, Veiga T, Klaassen P, Bovenberg RAL, Daran J, Heijnen JJ, Van Gulik WM (2011) Degeneration of penicillin production in ethanol-limited chemostat cultivations of *Penicillium chrysogenum*: a systems biology approach. BMC Syst Biol 5:132. doi:10.1186/1752-0509-5-132

Fleming A (1929) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. Br J Exp Pathol 10:226–236

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem 278:31988–31997. doi:10.1074/jbc.M304750200

Gombert A, Veiga T, Puig-Martinez M, Lamboo F, Nijland JG, Driessen a JM, Pronk JT, Daran JM (2011) Functional characterization of the oxaloacetase encoding gene and elimination of oxalate formation in the β-lactam producer *Penicillium*

chrysogenum. Fungal Genet Biol 48:831–839. doi:10.1016/j.fgb.2011.04.007

Gouka RJ, Punt PJ, van den Hondel CAMJJ (1997) Glucoamylase gene fusions alleviate limitations for protein production in *Aspergillus awamori* at the transcriptional and (post) translational levels. Appl Environ Microbiol 63:488–497

Guillemette T, van Peij NNME, Goosen T, Lanthaler K, Robson GD, van den Hondel CAMJJ, Stam H, Archer DB (2007) Genomic analysis of the secretion stress response in the enzyme-producing cell factory *Aspergillus niger*. BMC Genomics 8:158. doi:10.1186/1471-2164-8-158

Häkkinen M, Arvas M, Oja M, Aro N, Penttilä M, Saloheimo M, Pakula TM (2012) Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. Microb Cell Fact 11:134. doi:10.1186/1475-2859-11-134

Harris DM, van der Krogt ZA, Klaassen P, Raamsdonk LM, Hage S, van den Berg MA, Bovenberg RAL, Pronk JT, Daran J-M (2009a) Exploring and dissecting genome-wide gene expression responses of *Penicillium chrysogenum* to phenylacetic acid consumption and penicillinG production. BMC Genomics 10:75. doi:10.1186/1471-2164-10-75

Harris DM, Westerlaken I, Schipper D, van der Krogt Z a, Gombert AK, Sutherland J, Raamsdonk LM, van den Berg M a, Bovenberg R a L, Pronk JT, Daran J-M (2009b) Engineering of *Penicillium chrysogenum* for fermentative production of a novel carbamoylated cephem antibiotic precursor. Metab Eng 11:125–137. doi:10.1016/j.ymben.2008.12.003

Hesse SJ, Ruijter GJ, Dijkema C, Visser J (2002) Intracellular pH homeostasis in the filamentous fungus *Aspergillus niger*. Eur J Biochem 269:3485–3494

Hoff B, Kamerewerd J, Sigl C, Mitterbauer R, Zadra I, Kürnsteiner H, Kück U (2010a) Two components of a velvet-like complex control hyphal morphogenesis, conidiophore development, and penicillin biosynthesis in *Penicillium chrysogenum*. Eukaryot Cell 9:1236–1250. doi:10.1128/EC.00077-10

Hoff B, Kamerewerd J, Sigl C, Zadra I, Kück U (2010b) Homologous recombination in the antibiotic producer *Penicillium chrysogenum*: strain DeltaPcku70 shows up-regulation of genes from the HOG pathway. Appl Microbiol Biotechnol 85:1081–1094. doi:10.1007/s00253-009-2168-4

Hrdlicka PJ, Sørensen AB, Poulsen BR, Ruijter GJG, Visser J, Iversen JJL (2004) Characterization of nerolidol biotransformation based on indirect on-line estimation of biomass concentration and physiological state in batch cultures of *Aspergillus niger*. Biotechnol Prog 20:368–376. doi:10.1021/bp034137f

Imanaka H, Tanaka S, Feng B, Imamura K, Nakanishi K (2010) Cultivation characteristics and gene expression profiles of *Aspergillus oryzae* by membrane-surface liquid culture, shaking-flask culture, and agar-plate culture. J Biosci Bioeng 109:267–273. doi:10.1016/j.jbiosc.2009.09.004

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15

Jacobs DI, Olsthoorn MMA, Maillet I, Akeroyd M, Breestraat S, Donkers S, Van Der Hoeven RAM, van den Hondel CAMJJ, Kooistra R, Lapointe T, Menke H, Meulenberg R, Misset M, Müller WH, van Peij NNME, Ram AFJ, Rodriguez S, Roelofs MS, Roubos JA, Van Tilborg MWEM, Verkleij AJ, Pel HJ, Stam H, Sagt CMJ (2009) Effective lead selection for improved protein production in *Aspergillus niger* based on integrated genomics. Fungal Genet Biol 46:S141–S152. doi:10.1016/j.fgb.2008.08.012

Jin FJ, Takahashi T, Utsushikawa M, Furukido T, Nishida M, Ogawa M, Tokuoka M, Koyama Y (2010) A trial of minimization of chromosome 7 in *Aspergillus oryzae* by multiple chromosomal deletions. Mol Genet Genomics 283:1–12. doi:10.1007/s00438-009-0494-y

Jørgensen TR, Goosen T, van den Hondel CAMJJ, Ram AFJ, Iversen JJL (2009) Transcriptomic comparison of *Aspergillus niger* growing on two different sugars reveals coordinated regulation of the secretory pathway. BMC Genomics 10:44. doi:10.1186/1471-2164-10-44

Jørgensen TR, Nitsche BM, Lamers GEM, Arentshorst M, van den Hondel CAMJJ, Ram AFJ (2010) Transcriptomic insights into the physiology of Aspergillus niger approaching a specific growth rate of zero. Appl Environ Microbiol 76:5344–5355. doi:10.1128/AEM.00450-10, AEM.00450-10 [pii]

Jørgensen TR, Nielsen KF, Arentshorst M, Park J, van den Hondel CAMJJ, Frisvad JC, Ram AFJ (2011) Submerged conidiation and product formation by *Aspergillus niger* at low specific growth rates are affected in aerial developmental mutants. Appl Environ Microbiol 77:5270–5277. doi:10.1128/AEM.00118-11

Karaffa L, Kubicek CP (2003) *Aspergillus niger* citric acid accumulation: do we understand this well working black box? Appl Environ Microbiol 61:189–196. doi:10.1007/s00253-002-1201-7

Kardos N, Demain AL (2011) Penicillin: the medicine with the greatest impact on therapeutic outcomes. Appl Environ Microbiol 92:677–687. doi:10.1007/s00253-011-3587-6

Karimi-Aghcheh R, Bok JW, Phatale P a, Smith KM, Baker SE, Lichius A, Omann M, Zeilinger S, Seiboth B, Rhee C, Keller NP, Freitag M, Kubicek CP (2013) Functional analyses of Trichoderma reesei LAE1 reveal conserved and contrasting roles of this regulator. G3 (Bethesda) 3:369–378. doi:10.1534/g3.112.005140

Kimura S, Maruyama J, Takeuchi M, Kitamoto K (2008) Monitoring global gene expression of proteases and improvement of human lysozyme production

in the nptB Gene disruptant of *Aspergillus oryzae*. Biosci Biotechnol Biochem 72:499–505. doi:10.1271/bbb.70582

Koetsier MJ, Gombert AK, Fekken S, Bovenberg RAL, van den Berg MA, Kiel JAKW, Jekel PA, Janssen DB, Pronk JT, van der Klei IJ, Daran J-M (2010) The *Penicillium chrysogenum* aclA gene encodes a broad-substrate-specificity acyl-coenzyme A ligase involved in activation of adipic acid, a side-chain precursor for cephem antibiotics. Fungal Genet Biol 47:33–42. doi:10.1016/j.fgb.2009.10.003

Kwon MJ, Arentshorst M, Roos ED, van den Hondel CAMJJ, Meyer V, Ram AFJ (2011) Functional characterization of Rho GTPases in *Aspergillus niger* uncovers conserved and diverged roles of Rho proteins within filamentous fungi. Mol Microbiol 79:1151–1167. doi:10.1111/j.1365-2958.2010.07524.x

Kwon MJ, Jørgensen TR, Nitsche BM, Arentshorst M, Park J, Ram AFJ, Meyer V (2012) The transcriptomic fingerprint of glucoamylase over-expression in *Aspergillus niger*. BMC Genomics. doi:10.1186/1471-2164-13-701

Kwon MJ, Nitsche BM, Arentshorst M, Jørgensen TR, Ram AFJ, Meyer V (2013) The transcriptomic signature of RacA activation and inactivation provides new insights into the morphogenetic network of *Aspergillus niger*. PLoS One 8:68946

Levin AM, de Vries RP, Conesa A, de Bekker C, Talon M, Menke HH, van Peij NNME, Wösten HAB (2007) Spatial differentiation in the vegetative mycelium of *Aspergillus niger*. Eukaryot Cell 6:2311–2322. doi:10.1128/EC.00244-07

Li A, van Luijk N, ter Beek M, Caspers M, Punt PJ, van der Werf M (2011) A clone-based transcriptomics approach for the identification of genes relevant for itaconic acid production in *Aspergillus*. Fungal Genet Biol 48:602–611. doi:10.1016/j.fgb.2011.01.013

Lubertozzi D, Keasling JD (2009) Developing *Aspergillus* as a host for heterologous expression. Biotechnol Adv 27:53–75. doi:10.1016/j.biotechadv.2008.09.001

Määttänen P, Gehring K, Bergeron JJM, Thomas DY (2010) Protein quality control in the ER: the recognition of misfolded proteins. Semin Cell Dev Biol 21:500–511. doi:10.1016/j.semcdb.2010.03.006

Machida M, Asai K, Sano M, Tanaka T, Kumagai T et al (2005) Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438:1157–1161. doi:10.1038/nature04300

MacKenzie DA, Guillemette T, Al-Sheikh H, Watson AJ, Jeenes DJ, Wongwathanarat P, Dunn-Coleman NS, van Peij NNME, Archer DB (2005) UPR-independent dithiothreitol stress-induced genes in *Aspergillus niger*. Mol Genet Genomics 274:410–418. doi:10.1007/s00438-005-0034-3

Maeda H, Sano M, Maruyama Y, Tanno T, Akao T, Totsuka Y, Endo M, Sakurada R, Yamagata Y, Machida M, Akita O, Hasegawa F, Abe K, Gomi K, Nakajima T, Iguchi Y (2004) Transcriptional analysis of genes for energy catabolism and hydrolytic enzymes in the filamentous fungus *Aspergillus oryzae* using cDNA microarrays and expressed sequence tags. Appl Microbiol Biotechnol 65:74–83. doi:10.1007/s00253-004-1608-4

Martens-Uzunova ES, Schaap PJ (2008) An evolutionary conserved d-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45:1449–1457. doi:10.1016/j.fgb.2008.08.002

Martens-Uzunova E, Schaap P (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol 46 (Suppl 1):S170–S179

Martens-Uzunova ES, Zandleven JS, Benen JA, Awad H, Kools HJ, Beldman G, Voragen AG, van den Berg JA, Schaap PJ (2006) A new group of exo-acting family 28 glycoside hydrolases of Aspergillus niger that are involved in pectin degradation. Biochem J 400:43–52. doi:10.1042/BJ20060703, BJ20060703 [pii]

Masai K, Maruyama J, Sakamoto K, Nakajima H, Akita O, Kitamoto K (2006) Square-plate culture method allows detection of differential gene expression and screening of novel, region-specific genes in *Aspergillus oryzae*. Appl Microbiol Biotechnol 71:881–891. doi:10.1007/s00253-006-0429-z

Mela F, Fritsche K, de Boer W, van Veen J a, de Graaff LH, van den Berg M, Leveau JHJ (2011) Dual transcriptional profiling of a bacterial/fungal confrontation: *Collimonas fungivorans* versus *Aspergillus niger*. ISME J 5:1494–1504. doi:10.1038/ismej.2011.29

Metz B, Seidl-Seiboth V, Haarmann T, Kopchinskiy A, Lorenz P, Seiboth B, Kubicek CP (2011) Expression of biomass-degrading enzymes is a major event during conidium development in *Trichoderma reesei*. Eukaryot Cell 10:1527–1535. doi:10.1128/EC.05014-11

Meyer V (2008) Genetic engineering of filamentous fungi—progress, obstacles and future trends. Biotechnol Adv 26:177–185. doi:10.1016/j.biotechadv.2007.12.001

Meyer V, Damveld RA, Arentshorst M, Stahl U, van den Hondel CAMJJ, Ram AFJ (2007) Survival in the presence of antifungals: genome-wide expression profiling of Aspergillus niger in response to sublethal concentrations of caspofungin and fenpropimorph. J Biol Chem 282:32935–32948. doi:10.1074/jbc.M705856200, M705856200 [pii]

Meyer V, Arentshorst M, Flitter SJ, Nitsche BM, Kwon MJ, Reynaga-Pena CG, Bartnicki-Garcia S, van den Hondel CAMJJ, Ram AFJ, Reynaga-Peña CG (2009) Reconstruction of signaling networks regulating fungal morphogenesis by transcriptomics. Eukaryot Cell 8:1677–1691. doi:10.1128/EC.00050-09, EC.00050-09 [pii]

Mulder HJ, Nikolaev I, Madrid SM (2006) HACA, the transcriptional activator of the unfolded protein response (UPR) in *Aspergillus niger*, binds to partly palindromic UPR elements of the consensus sequence 5′-CAN(G/A)NTGT/GCCT-3′. Fungal Genet Biol 43:560–572. doi:10.1016/j.fgb.2006.02.005

Murooka Y, Yamshita M (2008) Traditional healthful fermented products of Japan. J Ind Microbiol Biotechnol 35:791–798. doi:10.1007/s10295-008-0362-5

Nitsche BM, Crabtree J, Cerqueira GC, Meyer V, Ram AFJ, Wortman JR, Fj A (2011) New resources for functional analysis of omics data for the genus *Aspergillus*. BMC Genomics 12:486. doi:10.1186/1471-2164-12-486

Nitsche BM, Jørgensen TR, Akeroyd M, Meyer V, Ram AFJ (2012) The carbon starvation response of *Aspergillus niger* during submerged cultivation: insights from the transcriptome and secretome. BMC Genomics 13:380. doi:10.1186/1471-2164-13-380

Noguchi Y, Sano M, Kanamaru K, Ko T, Takeuchi M, Kato M, Kobayashi T (2009) Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol 85:141–154. doi:10.1007/s00253-009-2236-9

Novodvorska M, Hayer K, Pullan S (2013) Transcriptional landscape of *Aspergillus niger* at breaking of conidial dormancy revealed by RNA-sequencing. BMC Genomics 14:246

Ogawa M, Tokuoka M, Jin FJ, Takahashi T, Koyama Y (2010) Genetic analysis of conidiation regulatory pathways in koji-mold *Aspergillus oryzae*. Fungal Genet Biol 47:10–18. doi:10.1016/j.fgb.2009.10.004

Ogawa M, Kobayashi T, Koyama Y (2012) ManR, a novel Zn(II)(2)Cys(6) transcriptional activator, controls the β-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol 49:987–995. doi:10.1016/j.fgb.2012.09.006

Ohno A, Maruyama J, Nemoto T, Arioka M, Kitamoto K (2011) A carrier fusion significantly induces unfolded protein response in heterologous protein production by *Aspergillus oryzae*. Appl Microbiol Biotechnol 92:1197–1206. doi:10.1007/s00253-011-3487-9

Pakula TM, Laxell M, Huuskonen A, Uusitalo J, Saloheimo M, Penttilä M (2003) The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei*. Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. J Biol Chem 278:45011–45020. doi:10.1074/jbc.M302372200

Panagiotou G, Andersen MR, Grotkjaer T, Regueira TB, Nielsen J, Olsson L (2009) Studies of the production of fungal polyketides in *Aspergillus nidulans* by using systems biology tools. Appl Environ Microbiol 75:2212–2220. doi:10.1128/AEM.01461-08

Papagianni M (2004) Fungal morphology and metabolite production in submerged mycelial processes. Biotechnol Adv 22:189–259. doi:10.1016/j.biotechadv.2003.09.005

Papagianni M, Mattey M (2006) Morphological development of *Aspergillus niger* in submerged citric acid fermentation as a function of the spore inoculum level. Application of neural network and cluster analysis for characterization of mycelial morphology. Microb Cell Fact 5:3

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G et al (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25:221–231

Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sndor E, Hartl L, Karaffa L, Druzhinina IS, Seiboth B, Le Crom S, Kubicek CP (2011) The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. BMC Genomics 12:269. doi:10.1186/1471-2164-12-269

Poulsen L, Andersen MR, Lantz AE, Thykaer J (2012) Identification of a transcription factor controlling pH-dependent organic acid response in *Aspergillus niger*. PLoS One 7:e50596. doi:10.1371/journal.pone.0050596

Römisch K (1999) Surfing the Sec61 channel: bidirectional protein translocation across the ER membrane. J Cell Sci 112:4185–4191

Ruijter GJG, Kubicek CP, Visser J (2002) Production of organic acids by fungi. In: Osiewacz HD (ed) The Mycota X industrial applications. Springer, Berlin, pp 213–230

Rumbold K, van Buijsen HJJ, Overkamp KM, van Groenestijn JW, Punt PJ, van der Werf MJ (2009) Microbial production host selection for converting second-generation feedstocks into bioproducts. Microb Cell Fact 8:64. doi:10.1186/1475-2859-8-64

Sainani KL (2009) The problem of multiple testing. PMR 1:1098–1103. doi:10.1016/j.pmrj.2009.10.004, S1934-1482(09)01460-9 [pii]

Sakamoto K, Arima T, Iwashita K, Yamada O, Gomi K, Akita O (2008) *Aspergillus oryzae* atfB encodes a transcription factor required for stress tolerance in conidia. Fungal Genet Biol 45:922–932. doi:10.1016/j.fgb.2008.03.009

Sakamoto K, Iwashita K, Yamada O, Kobayashi K, Mizuno A, Akita O, Mikami S, Shimoi H, Gomi K (2009) *Aspergillus oryzae* atfA controls conidial germination and stress tolerance. Fungal Genet Biol 46:887–897. doi:10.1016/j.fgb.2009.09.004

Salazar M, Vongsangnak W, Panagiotou G, Andersen MR, Nielsen J (2009) Uncovering transcriptional regulation of glycerol metabolism in *Aspergilli* through genome-wide gene expression data analysis. Mol Genet Genomics 282:571–586. doi:10.1007/s00438-009-0486-y

Saloheimo M, Pakula TM (2012) The cargo and the transport system: secreted proteins and protein

secretion in *Trichoderma reesei* (*Hypocrea jecorina*). Microbiology 158:46–57. doi:10.1099/mic.0.053132-0

Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP (2012) The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol 84:1150–1164. doi:10.1111/j.1365-2958.2012.08083.x

Semova N, Storms R, John T, Gaudet P, Ulycznyj P, Min XJ, Sun J, Butler G, Tsang A (2006) Generation, annotation, and analysis of an extensive *Aspergillus niger* EST collection. BMC Microbiol 6:7. doi:10.1186/1471-2180-6-7

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. doi:10.1101/gr.1239303, 13/11/2498 [pii]

Sidrauski C, Walter P (1997) The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. Cell 90:1009–1031

Sigl C, Haas H, Specht T, Pfaller K, Kürnsteiner H, Zadra I (2011) Among developmental regulators, StuA but not BrlA is essential for penicillin V production in *Penicillium chrysogenum*. Appl Environ Microbiol 77:972–982. doi:10.1128/AEM.01557-10

Snoek ISI, van der Krogt ZA, Touw H, Kerkman R, Pronk JT, Bovenberg RAL, van den Berg MA, Daran JM (2009) Construction of an hdfA *Penicillium chrysogenum* strain impaired in non-homologous end-joining and analysis of its potential for functional analysis studies. Fungal Genet Biol 46:418–426. doi:10.1016/j.fgb.2009.02.008

Taheri-talesh N, Horio T, Araujo-baza L, Dou X, Espeso EA, Pen MA, Osmani SA, Oakley BR (2008) The tip growth apparatus of *Aspergillus nidulans*. Mol Biol Cell 19:1439–1449. doi:10.1091/mbc.E07

Tamano K, Sano M, Yamane N, Terabayashi Y, Toda T, Sunagawa M, Koike H, Hatamoto O, Umitsuki G, Takahashi T, Koyama Y, Asai R, Abe K, Machida M (2008) Transcriptional regulation of genes on the non-syntenic blocks of *Aspergillus oryzae* and its functional relationship to solid-state cultivation. Fungal Genet Biol 45:139–151. doi:10.1016/j.fgb.2007.09.005

Terabayashi Y, Sano M, Yamane N, Marui J, Tamano K, Sagara J, Dohmoto M, Oda K, Ohshima E, Tachibana K, Higa Y, Ohashi S, Koike H, Machida M (2010) Identification and characterization of genes responsible for biosynthesis of kojic acid, an industrially important compound from *Aspergillus oryzae*. Fungal Genet Biol 47:953–961. doi:10.1016/j.fgb.2010.08.014

Terabayashi Y, Shimizu M, Kitazume T, Masuo S, Fujii T, Takaya N (2012) Conserved and specific responses to hypoxia in *Aspergillus oryzae* and *Aspergillus nidulans* determined by comparative transcriptomics. Appl Microbiol Biotechnol 93:305–317. doi:10.1007/s00253-011-3767-4

Thykaer J, Nielsen J (2003) Metabolic engineering of β-lactam production. Metab Eng 5:56–69. doi:10.1016/S1096-7176(03)00003-X

Tisch D, Kubicek CP, Schmoll M (2011) The phosducin-like protein PhLP1 impacts regulation of glycoside hydrolases and light response in *Trichoderma reesei*. BMC Genomics 12:613. doi:10.1186/1471-2164-12-613

Tsuchiya K, Nagashima T (1994) High level secretion of calf chymosin using a glucoamylase-prochymosin fusion gene in *Aspergillus oryzae*. Biosci Biotechnol Biochem 58:895–899

Tsukagoshi N, Kobayashi T, Kato M (2001) Regulation of the amylolytic and (hemi-)cellulolytic genes in aspergilli. J Gen Appl Microbiol 47:1–19

Valkonen SM (2004) The ire1 and ptc2 genes involved in the unfolded protein response pathway in the filamentous fungus *Trichoderma reesei*. Mol Genet Genomics 272:443–451. doi:10.1007/s00438-004-1070-0

Van de Vondervoort PJI, Langeveld SMJ, Visser J, van Peij NNME, Pel HJ, van den Hondel CAMJJ, Ram AFJ (2007) Identification of a mitotic recombination hotspot on chromosome III of the asexual fungus *Aspergillus niger* and its possible correlation with [corrected] elevated basal transcription. Curr Genet 52:107–114. doi:10.1007/s00294-007-0143-0

Van den Berg MA, Albang R, Albermann K, Badger JH, Daran J-M, Driessen AJM, Garcia-Estrada C, Fedorova ND, Harris DM, Heijne WHM, Joardar V, Kiel J a KW, Kovalchuk A, Martín JF, Nierman WC, Nijland JG, Pronk JT, Roubos JA, van der Klei IJ, van Peij NNME, Veenhuis M, von Döhren H, Wagner C, Wortman J, Bovenberg R a L (2008) Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. Nat Biotechnol 26:1161–1168. doi:10.1038/nbt.1498

Van den Berg RA, Braaksma M, van der Veen D, van der Werf MJ, Punt PJ, van der Oost J, de Graaff LH (2010) Identification of modules in Aspergillus niger by gene co-expression network analysis. Fungal Genet Biol 47:539–550. doi:10.1016/j.fgb.2010.03.005, S1087-1845(10)00046–0 [pii]

Van den Brink HJM, Petersen SG, Rahbek-Nielsen H, Hellmuth K, Harboe M (2006) Increased production of chymosin by glycosylation. J Biotechnol 125:304–310. doi:10.1016/j.jbiotec.2006.02.024

Van der Veen D, Oliveira JM, van den Berg W a M, de Graaff LH (2009) Analysis of variance components reveals the contribution of sample processing to transcript variation. Appl Environ Microbiol 75:2414–2422. doi:10.1128/AEM.02270-08

Veening J-W, Smits WK, Kuipers OP (2008) Bistability, epigenetics, and bet-hedging in bacteria. Annu

Rev Microbiol 62:193–210. doi:10.1146/annurev.micro.62.081307.163002

Veiga T, Gombert AK, Landes N, Verhoeven MD, Kiel J a KW, Krikken AM, Nijland JG, Touw H, Luttik M a H, van der Toorn JC, Driessen AJM, Bovenberg R a L, van den Berg M a, van der Klei IJ, Pronk JT, Daran J-M (2012a) Metabolic engineering of β-oxidation in *Penicillium chrysogenum* for improved semi-synthetic cephalosporin biosynthesis. Metab Eng 14:437–448. doi:10.1016/j.ymben.2012.02.004

Veiga T, Nijland JG, Driessen AJM, Bovenberg R a L, Touw H, van den Berg M a, Pronk JT, Daran J-M (2012b) Impact of velvet complex on transcriptome and penicillin G production in glucose-limited chemostat cultures of a β-lactam high-producing *Penicillium chrysogenum* strain. OMICS 16:320–333. doi:10.1089/omi.2011.0153

Veiga T, Solis-Escalante D, Romagnoli G, ten Pierick A, Hanemaaijer M, Deshmuhk A, Wahl A, Pronk JT, Daran J-M (2012c) Resolving phenylalanine metabolism sheds light on natural synthesis of penicillin G in *Penicillium chrysogenum*. Eukaryotic Cell 11:238–249. doi:10.1128/EC.05285-11

Vinck A, Terlou M, Pestman WR, Martens EP, Ram AFJ, van den Hondel CAMJJ, Wösten HAB (2005) Hyphal differentiation in the exploring mycelium of *Aspergillus niger*. Mol Microbiol 58:693–699. doi:10.1111/j.1365-2958.2005.04869.x

Vinck A, de Bekker C, Ossin A, Ohm RA, de Vries RP, Wösten HAB (2011) Heterogenic expression of genes encoding secreted proteins at the periphery of *Aspergillus niger* colonies. Environ Microbiol 13:216–225. doi:10.1111/j.1462-2920.2010.02322.x

Vitikainen M, Arvas M, Pakula T, Oja M, Penttilä M, Saloheimo M (2010) Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. BMC Genomics 11:441. doi:10.1186/1471-2164-11-441

Vongsangnak W, Salazar M, Hansen K, Nielsen J (2009) Genome-wide analysis of maltose utilization and regulation in aspergilli. Microbiology 155:3893–3902. doi:10.1099/mic.0.031104-0

Vongsangnak W, Nookaew I, Salazar M, Nielsen J (2010) Analysis of genome-wide coexpression and coevolution of *Aspergillus oryzae* and *Aspergillus niger*. OMICS 14:165–175. doi:10.1089/omi.2009.0118

Vongsangnak W, Hansen K, Nielsen J (2011) Integrated analysis of the global transcriptional response to α-amylase over-production by *Aspergillus oryzae*. Biotechnol Bioeng 108:1130–1139. doi:10.1002/bit.23033

Wada R, Maruyama J-I, Yamaguchi H, Yamamoto N, Wagu Y, Paoletti M, Archer DB, Dyer PS, Kitamoto K (2012) Presence and functionality of mating type genes in the supposedly asexual filamentous fungus *Aspergillus oryzae*. Appl Environ Microbiol 78:2819–2829. doi:10.1128/AEM.07034-11

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63. doi:10.1038/nrg2484

Wang B, Guo G, Wang C, Lin Y, Wang X, Zhao M, Guo Y, He M, Zhang Y, Pan L (2010) Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. Nucleic Acids Res 38:5075–5087. doi:10.1093/nar/gkq256

Wösten HAB, Moukha SM, Sietsma JH, Wessels JG (1991) Localization of growth and secretion of proteins in *Aspergillus niger*. J Gen Microbiol 137:2017–2023

Yoon J, Kimura S, Maruyama J, Kitamoto K (2009) Construction of quintuple protease gene disruptant for heterologous protein production in *Aspergillus oryzae*. Appl Microbiol Biotechnol 82:691–701. doi:10.1007/s00253-008-1815-5

Yuan XL, Roubos JA, van den Hondel CAMJJ, Ram AFJ (2007) Identification of InuR, a new Zn(II)2Cys6 transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol Genet Genomics 279:11–26. doi:10.1007/s00438-007-0290-5

Yuan X, Van Der Kaaij R, van den Hondel CAMJJ, Punt PJ, van der Maarel MJEC, Dijkhuizen L, Ram AFJ (2008) *Aspergillus niger* genome-wide analysis reveals a large number of novel alpha-glucan acting enzymes with unexpected expression profiles. Mol Genet Genomics 279:545–561. doi:10.1007/s00438-008-0332-7

10 Genomics Analysis of Biocontrol Species and Industrial Enzyme Producers from the Genus *Trichoderma*

Monika Schmoll[1], Bernhard Seiboth[2], Irina Druzhinina[2], Christian P. Kubicek[2]

CONTENTS

[1]Health and Environment Department, Bioresources Unit, Austrian Institute of Technology GmbH (AIT), Konrad Lorenz Strasse 24, 3430 Tulln an der Donau, Austria; e-mail: monika.schmoll@ait.ac.at

[2]Research Area Biotechnology and Microbiology, Institute of Chemical Engineering, Vienna University of Technology, Gumpendorferstrasse 1a, 1060 Vienna, Austria

I. Introduction

Trichoderma is a genus with diverse habitats in diverse climate zones (Jaklitsch 2009, 2011). Key features include rapid growth, bright green or white conidial pigments, and a repetitively branched conidiophore structure (Gams and Bissett 1998). *Trichoderma* spp. are especially successful in their habitats because of their diverse metabolic capabilities and their aggressively competitive nature (Gams and Bissett 1998).

The genus *Trichoderma* has a long history of investigation and application since its first occurrence in literature as early as 1794 (Persoon 1794). *Trichoderma* spp. are frequently isolated from soil, dead wood, or sporocarps of other fungi, which reflects their biological relevance as important contributors to the carbon cycle as well as mycoparasites. However, also such unexpected habitats as indoor environments or marine invertebrates are colonized by these fungi. Their role in various ecosystems also includes interactions with plants such as rhizosphere colonization or endophytism. *Trichoderma* spp. are even able to tolerate a number of toxic organic chemicals (Argumedo-Delira et al. 2012) and degrade fungicides (Ahlawat et al. 2010). Together with their capability to accumulate heavy metals (Joshi et al. 2011), *Trichoderma* spp. are promising candidates for bioremediation of contaminated soils and wastewater. Although the majority of species of this genus are versatile resources for application from agriculture to industry, some are opportunistic pathogens on animals and humans (Hatvani et al. 2013; Kredics et al. 2003).

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

Because of their various applications, different *Trichoderma* species have become model organisms for a broad spectrum of physiological phenomena, such as **plant cell wall degradation** and **enzyme production** (Kubicek et al. 2009), **mycoparasitism** and **biocontrol** (Hermosa et al. 2012), as well as **light response** (Schmoll et al. 2010a).

Plant cell wall degradation by fungi is currently a major focus of research, as sustainable technologies and protection of the environment become increasingly important in our society. The use of cellulases produced by *Trichoderma reesei* for degradation of cellulosic waste material to simple sugars and subsequent biosynthesis of ethanol by yeast has boosted research with *T. reesei*. Although the production of second-generation biofuels represents the most immediate application of fungal enzymes, additional uses in production of chemicals or modification of industrial processes currently done by organic chemistry provide promising perspectives as well.

Trichoderma reesei is a paradigm for the industrial production of enzymes, and these applied aspects were also the driving forces for basic research on the genes and their corresponding enzymes. Over the past decades, different aspects of the regulation of their formation have been described, including many different regulatory components and some of their interactions. However, these data cannot offer a convincing explanation why this fungus is today one of the most successful hosts for plant-cell-wall-degrading enzyme production. With the sequencing and annotation of the *T. reesei* genome (Martinez et al. 2008), a fundament was established for a more holistic or systems biology approach to understand this question.

At the same time, *Trichoderma* spp. are also of importance for biological protection of plants in agriculture; hence, this application complements industrial enzyme production. Consequently, *Trichoderma* spp. provide a toolbox from plant protection to biofuel production, and the increasing number of species sequenced will help reveal step by step the molecular basis for the specific characteristics of diverse *Trichoderma* strains. Sequencing of several genomes of *Trichoderma* spp. has already significantly increased our understanding of the evolution of this genus, and first steps toward connections between phenotype and genotype have already been made (Kubicek et al. 2011).

The industrial application of fungi and their enzymes, however, also influenced the way research is done on this topic. In nature, fungi mostly grow on solid substrates, and for those specialized in plant cell wall degradation, this substrate mainly consists of cellulosic material resulting from plant growth. Conditions applied to grow fungi in industry are considerably different from these natural environments, as are the artificial conditions used for research with fungi in the lab, where liquid cultures in fermentors or shake flasks are common.

Nevertheless, the vigorous growth and development of filamentous fungi under artificial conditions indicate that their evolutionary heritage provides them with the capabilities to thrive given use of appropriate substrate conditions. Signal transduction mechanisms as well as their output pathways are still at work and allow the fungus to cope with its man-made environment. The necessity to use defined, constant parameters for investigation of fungal physiology stands in contrast to the desire to explain their behavior in nature. Hence, for interpretation of the situation of a fungus in its natural habitat, the differences compared to an artificial environment are to be considered and can be exploited for strain improvement.

In this review, we summarize current knowledge on the genomic inventory of different *Trichoderma* spp. and provide an overview of genome-wide analyses. Discussion of results from a diverse array of species with different characteristics of the genus *Trichoderma* is intended to stimulate investigation of novel hypotheses inspired by various perspectives on this genus.

II. *Trichoderma* Genomes and Their Characteristics

A. General Characteristics

Trichoderma reesei QM6a, then ancestor of all currently used cellulase-producing mutants, was the first *Trichoderma* species whose genome was sequenced (Martinez et al. 2008). Three years later, the genomes of two other *Trichoderma* spp. that are vigorous mycoparasites – *Trichoderma virens* and *Trichoderma atroviride* – were also published (Kubicek et al. 2011). The genomes of two further strong mycoparasites, *Trichoderma harzianum* sensu stricto and *Trichoderma asperellum*, and the two facultative human pathogens *Trichoderma longibrachiatum* and *Trichoderma citrinoviride* have also been sequenced, and analysis is currently in progress.

The genome of *T. reesei* approximates the fungal average in size (34.1 Mbp) and is comprised of 9,129 genes. At least a further 23 putative genes are encoded in the genome as detected using sparse arrays. All of these genes are orphans, unique to *T. reesei* or at least only found within the genomes of close relatives. Interestingly, many of them are located next to genes encoding regulators such as the transmembrane serine/threonine kinase inositol requiring protein 1 (IRE1) involved in the unfolded protein response (UPR) or the carbon catabolite repressor (CRE1) (Arvas et al. 2010).

The genomes of the other *Trichoderma* spp. sequenced so far are in the same range (36.1 and 38.8 Mbp, for *T. atroviride* and *T. virens*), giving rise to 11,000–12,000 genes in asexual and about 9,000 in sexually reproducing taxa. A similar small genome size was also detected for *T. citrinoviride* (unpublished data), which in temperate climates is also frequently found in the form of its teleomorph *Hypocrea schweinitzii* (Jaklitsch 2009).

The analysis of the *T. reesei* genome sequence further revealed that only a limited number of paralogues are present, and that mobile elements contain multiple in-frame stop codons and are apparently inactive. This suggests that, similar to *Neurospora crassa*, a mechanism such as **repeat induced point mutation (RIP)** is active, which detects duplicated DNA sequences in the genome. RIP mutation is a homology-based defense mechanism that actively eliminates duplicated sequences, leading to an accumulation of G:C-to-A:T transitions often accompanied by epigenetic silencing of the mutated sequences through DNA methylation (Galagan and Selker 2004). All genes known to be necessary for operation of RIP are available in *T. reesei* (Kubicek et al. 2011). RIP occurs during the sexual phase of the life cycle in haploid nuclei, and although *T. reesei* QM6a has no full sexual cycle (Seidl et al. 2009a), newly isolated *T. reesei* strains are all able to propagate sexually (Druzhinina et al. 2010a).

Trichoderma reesei has **seven chromosomes** (Carter et al. 1992; Mäntylä et al. 1992). Herrera-Estrella et al. (1993) reported the presence of six chromosomes in *T. atroviride* (which was at that time mistaken as *T. harzianum*), but this smaller number is likely the result of overlooking a doublet chromosomal band because it has been similarly overlooked in *T. reesei*. No chromosome numbers are yet available for *T. virens*. Genome mapping has not been performed yet because *Trichoderma* spp. do not generally mate in vitro except for *T. reesei*, as recently shown by Seidl et al. (2009a).

The vast majority of the genes occur in all three *Trichoderma* species. Yet, *T. virens* and *T. atroviride* contain about 2,215 and 1,856 genes, respectively, that have no *T. reesei* ortholog, whereas *T. reesei* has only three unique genes. *Trichoderma atroviride* and *T. virens* share 1,873 orthologs that are not present in *T. reesei*, which could thus be the factors that make *T. atroviride* and *T. virens* strong environmental opportunists and successful mycoparasites (Kubicek et al. 2011).

A comparison of the genomic organization of *T. atroviride*, *T. virens*, and *T. reesei* showed that most of the genes are in **synteny**: Only 1 % of the genes of *T. reesei* and 5.5 % and 8.5 % of the genes of *T. virens* and *T. atroviride*, respectively, are located in areas containing breaks in synteny by a series of three or more genes. They also display extensive rearrangements, mostly represented by small inversions, as also observed between genomes of other fungi, and their size distribution correlates with the

operation of a random breakage model (Kubicek et al. 2011; Nadeau and Taylor 1984). The sequence similarity between orthologous and syntenic proteins, representative of the evolutionary distance between the three *Trichoderma* spp., is around 70 % (Kubicek et al. 2011), which is in the same range, for example, as between *Aspergillus fumigatus* versus *Aspergillus niger* (69 %) and *Aspergillus nidulans* (68 %) and resembles the evolutionary relationship between fish and humans (Fedorova et al. 2008).

The genomes of the three *Trichoderma* spp. also contain sequences with significant similarity to known transposable elements from other eukaryotes, which however were fragmented and are highly divergent from one another, indicating rather ancient events (Kubicek et al. 2011). Apparently, *Trichoderma* spp. are only occasionally subject to infection or invasion by transposable elements, which are rapidly rendered unable to replicate and rapidly accumulate mutations. More recently, a mutator-like element was detected in an L-xylulose reductase paralogue gene of two natural isolates of *T. reesei* (M. Flipphi and L. Atanasova, personal communication).

B. Gene Families

The gene families that are present in *Trichoderma* in the highest number of members comprise the Pfam groups for fungal specific **Zn (2)Cys(6) transcription factors** (PF00172, PF04082); solute transporters of the **major facilitator superfamily** (PF07690, PF00083); and enzymes potentially involved in the synthesis and modification of **secondary metabolites** (oxidoreductases, monooxygenases, AMP activation of acids, phosphopantetheine attachment, synthesis of isoquinoline alkaloids; Kubicek et al. 2011). However, these Pfam groups are also among the largest gene families in several other Pezizomycotina. A Markov cluster analysis (MCL) that identifies ortholog groups in genomes made for the three *Trichoderma* species, together with 44 Pezizomycotina genomes, showed that *T. virens* and *T. atroviride* have a unique genome inventory among all *Peziziomycotina* by harboring the highest number of genes that encode proteins with ankyrin and heterokaryon incompatibility (HET)/ankyrin/NACHT domains among all other fungi (Druzhinina and Kubicek, unpublished).

The **ankyrin repeat** has been shown to play important roles in microbial pathogenesis in bacteria, including endosymbiosis (Walker et al. 2007). Genes encoding proteins with HET/NACHT domains are part of the genetic systems that lead to recognition of and response to nonself during cell fusion between different individuals belonging to the same species but may also represent a niche adaptation for reaction to stimuli associated with defense against pathogens, self/nonself recognition, or differentiation (Fedorova et al. 2008; Paoletti and Saupe 2009).

The abundance of these two gene families in *Trichoderma* and their even higher amplification in strongly mycoparasitic species (*T. atroviride*, *T. virens*, *T. harzianum*, *T. asperellum*; Druzhinina and Kubicek, unpublished) suggest that they may contribute to the mycotrophic and opportunistic lifestyle.

III. Tools for Genome-wide Investigation and Beyond

Research with *Trichoderma* spp. has a long tradition; consequently, the development of genome-wide analysis methods also is reflected in the literature published. Early approaches involved investigation of different **cDNA (complementary DNA) libraries** from various growth conditions or strains and their sequencing (Chambergo et al. 2002; Foreman et al. 2003; Vizcaino et al. 2006), which was cost intensive, and expression levels of individual genes could hardly be assessed directly. However, availability of such sets of cDNA libraries enabled construction of **microarrays**, which then allowed for investigation of at least a subset of several thousand genes to be studied simultaneously (Rosales-Saavedra et al. 2006). Alternatively, prior to the availability of fully sequenced genomes, methods for genome-wide **cDNA subtraction** and analysis were applied. Although these analyses only yielded a small number of regulated genes compared to microarrays, they nevertheless revealed involvement of novel

genes (Mukherjee et al. 2007; Schmoll and Kubicek 2005; Schmoll et al. 2004; Schuster et al. 2011). A medium-scale approach for transcript analysis in multiple samples was provided by the TRAC (transcript analysis with the aid of affinity capture) system, which allows for simultaneous analysis of 96 samples for around 30 genes (Kataja et al. 2006; Rautio et al. 2006, 2007). After publication of the fully sequenced genome of *T. reesei* (Martinez et al. 2008) and later also of other *Trichoderma* spp., construction of full-genome microarrays became possible (Arvas et al. 2011; Häkkinen et al. 2012; Tisch et al. 2011b), followed by **large-scale sequencing** and transcriptome analysis, which is a standard technique today. Besides genomic and transcriptomic studies, a few proteomic (Adav et al. 2012; Herpoel-Gimbert et al. 2008; Jun et al. 2011) and metabolomic investigations were reported, although these tools are far less frequently applied.

The plethora of novel regulators and targets as revealed by genome-wide analysis of *Trichoderma* spp. also necessitated more efficient tools for genome manipulation to evaluate the numerous hypotheses raised by these studies. Although some of these genes encode proteins with well-characterized orthologs in other fungi, which at least allows a functional prediction, their physiological relevance still needs to be evaluated. However, many of the newly discovered genes are annotated as "hypothetical proteins" or even orphans, which makes it impossible to predict any function. A key technique to assess gene functions and to alter the characteristics of fungal strains is the inactivation of genes by targeted deletion or knockout. Functional genetic studies on a large-scale basis depend on the construction of well-defined gene deletion strains with an efficient gene-targeting system. Therefore, it was necessary to adapt the existing methods for gene manipulation of *Trichoderma* in a way that they are compatible to high-throughput methods.

A major obstacle for functional genetic studies is the general low frequency of **homologous recombination** (HR) of filamentous fungi in comparison, for example, to the yeast *Saccharomyces cerevisiae*. On average, only 1–10 % of the transformants show an integration of the knockout cassette at the homologous locus. Integration of DNA fragments in the genome requires the action of double-strand repair mechanisms and is mediated by two main pathways: the HR pathway and the nonhomologous end-joining pathway (NHEJ). Consequently, different approaches were developed to either improve HR or prevent the NHEJ pathway (Krappmann 2007). Construction of a *T. reesei* strain defective in the NHEJ pathway represents a milestone toward high-throughput gene deletion (Guangtao et al. 2009), and later such a strain became available also for *T. virens* (Catalano et al. 2011). The strongly increased efficiency of homologous integration paved the way to construction of a knockout library as already available for *N. crassa* (Park et al. 2011). This approach can be further refined by adding recombination-mediated marker reuse systems (Hartl and Seiboth 2005; Steiger et al. 2011).

For *T. reesei*, discovery of **sexual development** was another milestone in research with *Trichoderma* spp. (Seidl et al. 2009a). Crossing of strains bearing different characteristics or mutations is now possible. With this new tool in hand, a tool kit for streamlined high-throughput gene deletion was developed that also includes a primer database for construction of gene deletion vectors using three different markers (Schuster et al. 2012b). Also, *T. reesei* can serve as a blueprint for attempts to achieve sexual development under laboratory conditions in other species of the genus.

IV. Environmental Regulation of Gene Expression

The habitats where *Trichoderma* spp. are found are diverse. Some of these ecological niches will represent preferred habitats, but others will require considerable adaptation for survival. This reaction is accomplished after sensing the environment by transmission and interpretation of signals, which result in an adjustment of resource management and hence regulation of growth and primary and secondary metabolism as well as secretion of metabolites and hydrolytic enzymes (Fig. 10.1). Thereby,

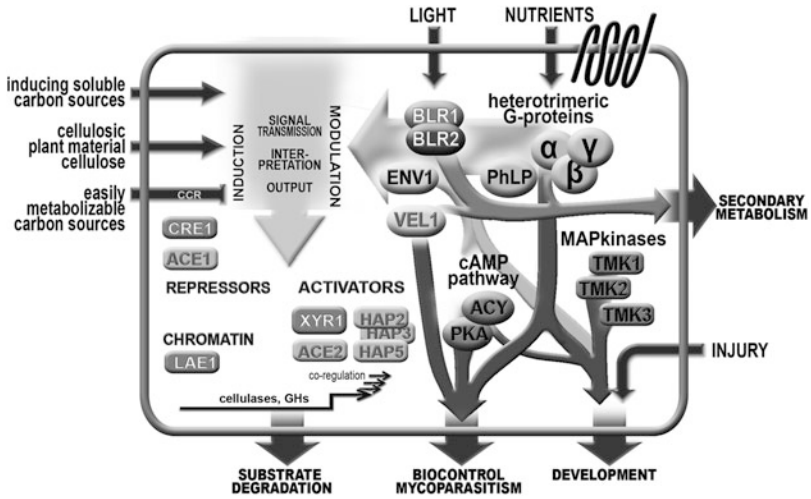


Fig. 10.1. Gene regulation mechanisms in *Trichoderma*. Names of regulators represent the orthologs of all relevant species if data are available

considerable cross talk among transmission systems of environmental signals occurs, which determines the response of a fungus to its environment (Bahn et al. 2007). The signal transduction machinery of *Trichoderma* spp. is well studied with respect to both its influence on enzyme production and interaction with other fungi (Carreras-Villasenor et al. 2012; Schmoll 2008).

A. Light Response

Light is an important environmental cue for fungi in general (Schmoll 2011), and virtually every metabolic pathway is influenced by light (Tisch and Schmoll 2010). *Trichoderma* spp. were among the first fungi to be studied in detail for their response to light and were used as model organisms to study photoconidiation and morphological alterations on illumination (Schmoll et al. 2010a). In *Trichoderma*, light initiates asexual reproduction, causes decreased mycelial growth, alters carbon metabolism, and influences gene expression (Herrera-Estrella and Horwitz 2007). Considerable cross talk was observed between nutrient and light signaling (Schmoll et al. 2010a).

As in other fungi, blue light is sensed in *Trichoderma* by **photoreceptors** that comprise the PAS/LOV (Per-Arndt-Sim domains of the light-oxygen-voltage type) domains as well as GATA-type zinc fingers, which enable them to act as transcription factors (Casas-Flores et al. 2004; Castellanos et al. 2010; Schmoll et al. 2005; see Chap. 6 in this volume).

The photoreceptor blue light regulator (BLR1) is conserved in all three *Trichoderma* spp. characterized so far at the genomic level and contains three PAS domains, with the first ones a PAS/LOV domain and a GATA-type zinc finger DNA-binding domain. However, these genes lack the transcriptional activation domain described for *N. crassa* WC-1. The BLR2 proteins of *Trichoderma* are characterized by one PAS domain and a GATA-type zinc finger DNA-binding domain (Schmoll et al. 2010a) and are assumed to be a limiting factor in blue light perception (Esquivel-Naranjo and Herrera-Estrella 2007). The putative third photoreceptor, the light regulatory protein ENVOY1 (ENV1), contains one PAS/LOV domain and is involved in photoadaptation (Castellanos et al. 2010; Schmoll et al. 2005; Schuster et al. 2007). BLR1 and BLR2 are involved in regulation of cellulase gene expression (Castellanos et al. 2010), development (Casas-Flores et al. 2004; Chen et al. 2012; Seibel et al. 2012b), and peptaibol production (Komon-Zelazowska et al. 2007; Sanchez-Arreguin et al. 2012) in *Trichoderma* spp.

BLR1 and BLR2 were found to be active in darkness as well and exert their functions at both the transcript and protein levels. Thereby, they act individually and as a complex (Sanchez-Arreguin et al. 2012).

Candidates for additional transmitters of light signals are the microbial opsins, which are assumed to sense green light (Spudich and Jung 2005). However, only one putative microbial opsin has been found in *T. atroviride*, but not in *T. reesei* or *T. virens* (Schmoll et al. 2010a). The *Trichoderma* genomes also comprise cyclobutane pyrimidine dimer (CPD) photolyases, cry-dash cryptochromes, and cryptochrome/6-2 photolyases. Interestingly, the cry-dash cryptochromes of *Trichoderma* show unusually long COOH termini, suggesting that novel or additional functions might have developed for these proteins. Moreover, phytochromes have been detected (Schmoll 2008), which indicates a sensitivity to red and far-red light as observed previously (Rosales-Saavedra et al. 2006).

In addition, genomes of *T. reesei*, *T. atroviride*, and *T. virens* comprise of homologs of the major components of the network governing light response and circadian rhythms (Brunner and Kaldi 2008), such as frequency (FRQ1), F-box/WD-40 repeat containing protein (FWD1), blue light induced 1 (BLI1), SHAGGY, or TIMELESS (Schmoll et al. 2010a).

In *T. atroviride* and *T. reesei*, around 3 % of the genome are regulated by light (Rosales-Saavedra et al. 2006; Tisch et al. 2011b). In contrast to the situation in *N. crassa* (Chen et al. 2009), also genes repressed by light have been observed in *Trichoderma*. Accordingly, the photoreceptors BLR1 and BLR2 as well as ENV1 positively and negatively regulate gene expression (Rosales-Saavedra et al. 2006; Schuster et al. 2007). However, for ENV1 and BLR2, it was shown that despite overexpression in darkness, light was required for its regulatory effect (Esquivel-Naranjo and Herrera-Estrella 2007; Schuster et al. 2007).

Genes involved in **carbohydrate metabolic processes** were found to be significantly enriched in the transcriptome sampled on growth in light in *T. reesei*, as were those responsible for regulation of oxidoreductase activity and sulfate transport (Tisch et al. 2011b). Recently, components of the oxidative stress pathway were reported to be responsive to light in *T. atroviride* (Carreras-Villasenor et al. 2012), hence supporting the hypothesis that oxidative stress response correlates with photostimulation, which is dependent on BLR1 (Friedl et al. 2008).

The recent genome-wide studies on *T. reesei* revealed an unexpected phenomenon: Lack of photoreceptors, components of the heterotrimeric G-protein pathway, or the cAMP pathway increased the number of genes differentially expressed between light and darkness as well as the extent of differential expression (Schuster et al. 2012a; Tisch et al. 2011b; M. Schmoll, unpublished results). Hence, light responsiveness of *T. reesei* physiology seems to be enhanced if modulating factors are missing. The physiological relevance of this phenomenon warrants further investigation.

Trichoderma spp. also have an ortholog of VELVET, a light-dependent regulator of morphogenesis, development, and secondary metabolism (Carreras-Villasenor et al. 2012; Schmoll 2008), which was mainly studied in *Aspergillus* spp. (Bayram and Braus 2012). Deletion of *vel1* results in a total loss of conidiation and loss of gliotoxin production in *T. virens* (Mukherjee and Kenerley 2010).

B. Heterotrimeric G-Protein Signaling

As one of the crucial signal reception and transmission pathways, the heterotrimeric G-protein pathway has also been studied extensively in *Trichoderma* spp. Thereby, the signal perceived by the seven-transmembrane domain G-protein-coupled receptors (GPCRs) is transmitted by G-protein alpha, beta, and gamma subunits, which form a heterotrimer in their inactive state. On activation by a ligand binding to the GPCR, they dissociate, and the individual subunits exert their regulatory functions on downstream output pathways (Cabrera-Vera et al. 2003). *Trichoderma* spp. have 3 G-protein alpha subunits and 1 G-protein beta and gamma subunit each, as well as more than 50 GPCRs (Brunner et al. 2008; Schmoll 2008).

Trichoderma atroviride **G-alpha subunit** TGA1 is essential for mycoparasitism and efficient production of secondary metabolites (Reithner et al. 2005). TGA1 was further reported to influence cyclic AMP (cAMP) levels, although the results in the literature do not allow for consistent interpretation, and an involvement of the G-protein beta and gamma subunits could not be excluded (Reithner et al. 2005; Rocha-Ramirez et al. 2002).

In contrast, however, the deletion of *tgaA* (*tga1*) in *T. virens* resulted only in a somewhat reduced mycoparasitic activity on *Sclerotium rolfsii* (Mukherjee et al. 2004). These differences are in accordance with the different modes of mycotrophy between *T. atroviride* and *T. virens*, in which the induction (and thus signaling) of a mycoparasitic response is much less important for *T. virens* because of constitutive transcription of the necessary genes (Atanasova et al. 2013)

In *T. reesei*, the G-alpha subunit GNA1 is involved in light-dependent regulation of cellulase and hydrophobin gene expression (Seibel et al. 2009). In contrast to many fungi studied, including *T. atroviride*, deletion of *T. virens* TGAA (Mukherjee et al. 2004) or *T. reesei* GNA1 (Seibel et al. 2009) only has a minor or no effect on growth and sporulation.

Also, the *T. atroviride* G-alpha subunit TGA3 has an impact on conidiation, growth, and intracellular cAMP levels as well as mycoparasitism (Mukherjee et al. 2004; Zeilinger et al. 2005) and is involved in regulation of peptaibol production (Komon-Zelazowska et al. 2007). Similarly, the defects observed on deletion of *T. atroviride* TGA3 (Zeilinger et al. 2005) do not occur in *T. reesei* (Schmoll et al. 2009; Schuster et al. 2012b). Consequently, it seems that, despite comparable genomic equipment with respect to heterotrimeric G-protein signaling, the downstream targets of this pathway differ even in closely related *Trichoderma* spp.

Trichoderma reesei GNA3, like GNA1, regulates cellulase gene expression but exerts its function specifically in light (Schmoll et al. 2009). Light-dependent regulation of heterotrimeric G-protein signaling is at least in part mediated by ENV1, which downregulates transcription of *gna3* and is essential for positive-feedback regulation of *gna1* (Tisch et al. 2011a). For the genes related to *N. crassa gna2*, no function could be determined yet because the phenotypes of deletion mutants are comparable to the wild type (Mukherjee et al. 2004; Schuster et al. 2012b).

Functions of the G-protein beta and gamma subunits as well as their putative cochaperone PhLP1, a class I phosducin-like protein, have been studied in *T. reesei*. *Gnb1*, *gng1*, and *phlp1* were found to be light induced and to regulate more than 600 genes, including 21 glycoside hydrolases (GHs) in light (Tisch et al. 2011b). GPCRs were mostly found to be involved in mycoparasitism in *Trichoderma* spp. (Brunner et al. 2008; Omann and Zeilinger 2010).

C. The cAMP Pathway

Major factors of the cAMP pathway are adenylate cyclase, which synthesizes cAMP, phosphodiesterases, which degrade cAMP as well as catalytic and regulatory subunits of protein kinase A (D'Souza and Heitman 2001). In addition to these components, the genome of *T. reesei* also contains an adenylate-cyclase-associated protein (Schmoll 2008). In mammalian systems, adenylate cyclases are assumed to represent molecular coincidence detectors integrating various signals to a defined output (Anholt 1994). Also, the data available on *Trichoderma* spp. support such a role because of the involvement of the cAMP pathways in physiological functions from growth, enzyme production, conidiation, to secondary metabolite production as well as the regulatory effect of upstream signaling pathways.

In all *Trichoderma* spp. tested, lack of **adenylate cyclase** caused strongly reduced growth on solid media (but not liquid media) and enhanced conidiation. Despite the crucial function of adenylate cyclase in biosynthesis of cAMP, amendment of the medium with cAMP could not restore the severe phenotype of these mutants (Mukherjee et al. 2007; Schuster et al. 2012a). Hence, the function of adenylate cyclase in regulation of growth in *Trichoderma* spp. likely goes beyond production of this second-

ary messenger. Involvement of cAMP in regulation of cellulases in *T. reesei* has been known for quite some time (Sestak and Farkas 1993), and its function in light response was shown even earlier (Farkas et al. 1990). However, only recently molecular data supporting these findings were provided. Protein kinase A plays an important role in the regulation of light responses (Casas-Flores et al. 2006), and adenylate cyclase regulates cellulase gene expression in a light-dependent manner in *T. reesei* (Schuster et al. 2012a).

Subtractive hybridization revealed that adenylate cyclase positively regulates at least nine genes involved in secondary metabolism, including several cytochrome P450 genes (Mukherjee et al. 2007).

D. Signaling by Mitogen-Activated Protein Kinases

Mitogen-activated protein (MAP) kinase cascades are composed of three serine/threonine protein kinase cascades that act in series (Chen and Thorner 2007). *Trichoderma* spp. have all three of the classes common in filamentous fungi: the pheromone/pathogenicity pathway, the cell integrity pathway, and the stress response pathway (Carreras-Villasenor et al. 2012; Schmoll 2008).

Deletion of the pheromone/pathogenicity pathway MAP kinase *tmkA* (*tvk1*) in a "P" strain of *T. virens* resulted in a loss of mycoparasitism and biocontrol against *S. rolfsii*, but not *Rhizochtonia solani* (Mukherjee et al. 2003; Viterbo et al. 2005). On the contrary, deletion of *tmkA* in a "Q" strain of *T. virens* resulted in improved biocontrol against both *R. solani* and *Pythium ultimum* (Mendoza-Mendoza et al. 2003). Deletion of the this gene in *T. virens* further caused a growth phenotype with severely reduced hydrophobicity of conidia in liquid and solid media. Subtractive hybridization revealed differential regulation of not only several hydrophobin genes but also one clock-controlled gene. Hence, TVK1 regulates morphology and cell wall composition during development (Mendoza-Mendoza et al. 2007).

Loss of TMKA in *T. virens* caused decreased colonization and lysis of host hyphae and reduced antagonistic properties. Moreover, TMKA is responsible for repression of conidiation in the dark (Mukherjee et al. 2003). In addition, induction of systemic resistance in plants by *T. virens* and *T. asperellum* is dependent on the presence of TMKA/TIPK (Shoresh et al. 2006; Viterbo et al. 2005). Similar to the *T. virens* "Q strains," the deletion of TMK1 in *T. atroviride* resulted in improved biocontrol against *R. solani*, paralleled by mycoparasitism-related morphogenetic events, and increased production of chitinase activity and antifungal compounds (Reithner et al. 2007).

The role of the two other MAP kinases is less well understood because mutants in them are also characterized by poor growth that precludes successful antagonism; consequently, *T. virens* mutants in the cell-integrity kinase TmkB were also defective in mycoparasitism on *S. rolfsii* (Kumar et al. 2010), and the HOG1 MAPK in *T. atroviride* that is involved in osmotic and oxidative stress tolerance was also necessary for its mycoparasitic ability (Delgado-Jarana et al. 2006).

Interestingly, the *T. reesei* homolog of this MAP kinase, TMK3, was found to be located in close vicinity to the genes encoding the G-protein alpha subunit GNA3 and the glycogen phosphorylase GPH1, called the MGG cluster (Schmoll et al. 2009). The expression of glycogen phosphorylase is regulated by the HOG-MAP kinase pathway (Sunnarborg et al. 2001), and its activity is triggered by cAMP-mediated phosphorylation in yeast (Lin et al. 1995). As GNA3 positively influences cAMP levels and this cluster is present in several ascomycetes, functional relevance of this gene order is predicted (Schmoll et al. 2009).

V. Enzyme Production

A. The Genomic Inventory for Substrate Degradation

The most interesting results for the *T. reesei* community are naturally related to plant cell wall degradation (see also Chap. 8 in this volume). Enzymes that degrade cellulose are

formed on diverse carbon sources, albeit in slightly different proportions and total amounts (Dashtban et al. 2011; Margolles-Clark et al. 1997). Analysis of the genome produced at least two surprises (Martinez et al. 2008): First, the annotated sequence contains a comparatively small number (200) of **glycoside hydrolases**, which is comparable to *N. crassa*, but more surprising was that also the group of genes related to plant cell wall degradation, including cellulases, hemicellulases, and pectinases, was reduced. The second surprise was that a considerable number of these GH genes are organized in distinct clusters.

Martinez et al. (2008) identified only seven cellulases, including cellobiohydrolases and endoglucanases and three GH61 family members. The latter family contains lytic polysaccharide monooxygenases (PMOs), which promote cellulose breakdown by acting on the surfaces of the insoluble substrate, introducing chain breaks in the cellulose chains. Beside cellulases, the group of hemicellulases (16 genes) and pectinases (5 genes) are strongly reduced. Although *T. reesei* has four pectin hydrolases that attack the main chain, it misses, for example, lyases and arabinases, which act on the pectin side chains (Akel et al. 2009). Many of these GHs involved in plant polysaccharide depolymerization carry an additional carbohydrate-binding module (CBM), which promotes the binding of the catalytic domain GHs domain to the target polysaccharide. Similar to the trend found for plant-cell-wall-degrading enzymes, the number of CBM-encoding proteins is reduced in *T. reesei*. It seems that *T. reesei* has lost several enzymes that degrade plant cell walls during its evolution to a saprotrophic organism because *T. virens* and *T. atroviride*, which evolved earlier than *T. reesei* (Kubicek 2013; Kubicek et al. 2011), have a considerably higher number of GHs (260 and 257, respectively), and their number of these enzymes is higher.

Since *T. reesei* has specialized toward saprotrophism on predecayed wood (Druzhinina et al. 2011), it can be concluded that these additional genes are not needed for efficient breakdown of the predigested lignocelluloses found in its habitat. Support for the hypothesis that already a small set of enzymes is sufficient for the breakdown of certain substrates came from Banerjee et al. (2010). They developed an optimized mixture of six "core" enzymes, defined as cellobiohydrolase 1, cellobiohydrolase 2, endo-β-1,4-glucanase 1, β-glucosidase 1, endo-β-1,4-xylanase 3, and β-xylosidase for the hydrolysis of pretreated biomass. The hydrolytic efficiency of this core set could be significantly improved by the addition of other "accessory" enzymes to match the efficiency of contemporary commercial enzyme mixtures (Accellerase 1000 or Spezyme CP).

The annotation of the *T. reesei* CAZyme genes was recently updated (Häkkinen et al. 2012). In total, 201 GH genes, 22 CE (carbohydrate esterase) genes, and 5 PL (polysaccharide lyase) genes were found. With regard to plant cell wall degradation, the members discussed next can be identified (Häkkinen et al. 2012; Martinez et al. 2008).

Cellulases are found within several GH families: The main cellobiohydrolase CBH1/CEL7A is found within the GH7 family together with the endoglucanase EGL1/CEL7B. The other cellobiohydrolase CBHII/CEL6A is assigned to the GH6 family. Other endoglucanases are found in the families GH5, GH12, and GH45. Besides EGLII/CEL5A, GH5 includes the putative membrane-bound endoglucanase CEL5B and an additional endoglucanase candidate. The GH12 family includes the characterized EGLIII/CEL12A together with another endoglucanase. EGV and xyloglucanase CEL74A are found in families GH45 and GH74, respectively. PMOs were originally classified to family GH61 because of the erroneous determination of some hydrolytic activity. Häkkinen et al. (2012) assigned five members to this group, two more than previously reported (Martinez et al. 2008). Eleven β-glycosidases have been identified, including the major extracellular BGLI and an intracellular BGLII and nine further candidate β-glycosidases. They are found in family GH1 (two genes) and GH3 (nine genes).

Because of the high degree of substrate diversification, hemicellulases are naturally more diverse than cellulases. Xylanases are found in several families, including GH10 (XYN3), GH11 (XYN1, XYN2, XYN5), and GH30 (XYN4). A putative fifth member is

found in GH30. Within the GH3 family, two β-xylosidases (including BXL1) are found. Further, GH3 β-xylosidases are found in family GH39, and two β-xylosidase/α-L-arabinofuranosidases are found in GH4. Arabinofuranosidases ABFI and ABFIII are found in family GH54 and ABFII in family GH62.

With two members of the GH30 family of glucuronyl-xylanases, *Trichoderma* spp. are among those fungi with the highest number of these enzymes, which reflects specialization toward efficient hydrolysis of certain hemicellulose side chains (Kubicek 2013). Interestingly, both genes are targets of the heterotrimeric G-protein pathway, and in the respective mutants, their transcript abundance decreases in light (Tisch et al. 2011b).

Carbohydrate esterase families CE3 and CE5 contain acetyl xylan esterases. CE5 includes, beside the acetyl xylan esterases, cutinases. Further hemicellulases are present in families GH2 and GH27. The family GH2 includes five candidate β-mannosidases, candidate exo-β-D-glucosaminidase, and β-galactosidase/β-glucuronidase. Six putative α-galactosidases are found in family GH27. Family CE16 contains an additional member besides the characterized acetyl esterase AES1 (Häkkinen et al. 2012). Also, a class of PLs, PL20, has a second enzyme related to the glucuronan hydrolase TRGL.

In addition to the known GH67 α-glucuronidase (GLRI), a GH115 type of α-glucuronidase [α-1,2- or α-(4-O-methyl)-glucuronidase] was found. Four GH79 β-glucuronidases probably involved in proteoglycan hydrolysis are present. Only four proteins are found that encode putative petinases (GH28), including an endopolygalacturonase (PGA1), exopolygalacturonase (PGX1), exorhamnogalacturonase (RGX1), and endoxylogalacturonase (XGA1). Two members of the GH105 family are putative rhamnogalacturonylhydrolases. The GH5 family includes, besides cellulases, a β-mannanase (MANI) and a candidate β-1,3-mannanase/endo-β-1,4-mannosidase, glucan β-1,3-glucosidase, endo-β-1,6-glucanase, and a β-glycosidase.

The comparison of the enzyme content of *T. reesei* to other fungal species revealed several interesting details (Häkkinen et al. 2012; Kubicek 2013; Martinez et al. 2008). Despite the rather small number of GH in *T. reesei*, some families are enriched, such as GH 18 chitinases, GH27 α-galactosidases, including AGLIII and four further candidate α-galactosidases unique to *T. reesei*, GH64 β-1,3-glucanase, GH30 endoxylanase (glucuronoxylan xylanohydrolases) and GH54 α-L-arabinofuranosidases/β-xylosidases, and GH79 β-glucuronidases, GH89 N-aceytlglucosaminidase, and GH95 α-fucosidases. One of the two clusters that contain genes encoding members of family GH43 is significantly reduced in *T. reesei* compared to other Pezizomycotina species; reduction is also visible in members from the family GH61.

These expansions and reductions are often found in other *Trichoderma* spp. as well and indicate that *Trichoderma* has apparently specialized toward efficient hydrolysis of some hemicelluloses and other carbohydrate side chains. Some of these enzymes are encoded by genes that have no homologs in other fungi but in bacteria and may have been obtained by horizontal gene transfer (Akel et al. 2009; Häkkinen et al. 2012).

Among these enriched GH families, GH family 18, which contains all fungal chitinases, is the largest family of carbohydrate-active enzymes in mycoparasitic *Trichoderma* species. Thereby, functions in mycoparasitic attack of other fungi and remodeling of the fungus' own cell wall appear to overlap. The recognition between self and nonself functions is suggested to be achieved by protection in healthy hyphae versus deprotection during mycoparasitic attack, hyphal aging, or autolysis (Gruber and Seidl-Seiboth 2012).

As for the hydrolysis of the prey's cell wall upon a mycoparasitic attack, the respective enzymes must be able to preferentially hydrolyze chitin, β-1,3-, and α-1,3/1,4-glucans that make up most of it (Latge 2007). The excellent ability of *Trichoderma* spp. to degrade the carbohydrate armor of the prey hyphae is reflected in the findings that *T. atroviride* and *T. virens* have the highest number of chitinases of all ascomycetes (29 and 36, respectively). Chitinases belong to GH family 18 and can be categorized into subgroups A, B, and C based on the

amino acid sequences of their GH 18 modules (Seidl et al. 2005). The proteins of the respective subgroups differ in their substrate binding geometry: Subgroup A and C proteins have tunnel-shaped substrate-binding clefts (typical for exoactivity), whereas subgroup B proteins have more shallow and open binding clefts, typical for endoactivity (Horn et al. 2006). The chitinases from the three subgroups also have different modular structures (Seidl et al. 2005): Many subgroup B chitinases have CBMs at their C-terminal ends. The CBMs of subgroup B chitinases from *Trichoderma* spp. belong exclusively to the CBM family 1 (CBM1) and possess cellulose- and chitin-binding properties. Subgroup C chitinases contain both CBM 18 (chitin-binding) as well CBM 50 ("LysM") modules. CBMs are completely lacking from subgroup A. CBMs enable the chitinases to bind more tightly to insoluble substrates, which may be a limiting factor in chitin hydrolysis, and cleave it in a processive way (Eijsink et al. 2008). In fact, addition of a CBM to CHI18-15 resulted in both enhanced chitinase activity and biocontrol ability of the respective recombinant strain of *T. harzianum* (Limon et al. 1999). The expansion of chitinase genes in *Trichoderma* is observed only with subgroups B and C, while the numbers of subgroup A chitinases have remained rather constant (Karlsson and Stenlid 2008), indicating that these fungi have an improved ability to attack insoluble chitinous substrates. Interestingly, gene expression patterns of subgroup C chitinase indicate individual and complex regulatory responses (Gruber et al. 2011), suggesting that they are involved in many different physiological events. Ihrmark et al. (2010) showed that CHI18-13 (subgroup B) and CHI18-15 (which does not belong to groups A–C because it has been acquired by horizontal gene transfer from soil bacteria (Karlsson and Stenlid 2009), evolved under positive selection typical of a coevolutionary arms race between plant host and fungal pathogen, which also underlines the role of chitinases in antagonism and mycotrophy.

Another GH family that is expanded in *Trichoderma* and that is likely involved in the degradation of chitinous carbohydrates is the GH75 family chitosanases. Chitosan is the partially deacetlyated form of chitin and can also be found in fungal cell walls, where it is formed by the action of chitin deacetylases (CE family 4).

The second most abundant polymer in fungal cell walls is β-1,3-glucan (Latge 2007) with β-1,6-branches; β-1,3-glucanases can be found in GH families 55, 64, and 81. For all of these families, the mycoparasitic *Trichoderma* spp. again have the highest number of members. The β-1,6-glucanase Tvbgn3 (Djonovic et al. 2006a) was shown to be important for the mycoparasitism of *T. virens* on *P. ultimum*, which does not contain chitin but solely glucans as the main structural cell wall components. A role in biocontrol has been suggested for lytic enzymes (Seidl et al. 2006b; Viterbo et al. 2002).

B. CAZyme Gene Clusters in *T. reesei*

Although the overall number of genes involved in plant cell wall degradation is lower in *T. reesei* compared to other fungi, this might be compensated by a higher order of organization within the genome. Genes involved in a common pathway or function are frequently found near each other on bacterial chromosomes, but usually such clusters are less frequently found in eukaryotes (Keller and Hohn 1997). However, natural selection may favor such gene clusters through a variety of mechanisms, such as increased efficiency of coregulation. Of the 316 total CAZymes genes, 130 (41 %) are found in 25 clusters ranging from 14 to 275 kb (roughly 2.4 Mb, or 7 % of the genome). On average, the CAZymes are fivefold enriched in these clusters, and 73 % of them occur in gaps between syntenic blocks (Martinez et al. 2008). This means that in these gaps the preserved co-localization of genes on chromosomes is different for closely related species. Of the 95 CAZyme genes in these clusters, 68 have orthologs in *Fusarium graminearum*, but only 16 orthologs are in synteny with *F. grami-*

nearum. This indicates that these clusters emerge from gene movement rather than by gene duplications as only a few CAZyme paralogues are located in the clusters. The dominant CAZyme genes in these clusters encode GHs and make up 46 % of all GH genes present in the genome. Their majority is related to plant biomass degradation, including, for example, the two major cellobiohydrolases, CBH1 and CBH2. Also, all genes that encode a CBM1, implicated in cellulose binding, are exclusively found in these syntenic gaps. Glycosyltransferases found in CAZyme clusters are believed to be involved in synthesizing fungal cell walls.

Genome-wide transcript analysis confirmed that these clusters are coregulated. Of 49 CAZyme genes that were confirmed by genomic cluster analysis to be clustered in the *T. reesei* genome, 13 were more than twofold upregulated during growth on cellulose, and 6 were upregulated during growth on lactose (Kubicek 2013). Most of these clusters occur near the end of the respective *T. reesei* scaffolds, which indicates that these regions are next to either the telomeric regions of the chromosomes or long repetitive sequences found in dynamic regions of the genome.

Several of the regions of high CAZyme gene density also contain genes encoding proteins involved in secondary metabolism, including polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS) gene (Martinez et al. 2008). Coregulation of the formation of such antagonistic metabolites with plant-cell-wall-degrading enzymes could be a further strategy of *T. reesei* to compete with other organisms in its habitat. A number of genes involved in secondary metabolite synthesis, particularly in *Aspergillus* spp., are regulated by the **putative protein-methyltransferase LaeA** (Bok and Keller 2004). Although the exact mechanism of LaeA function is still unknown, it is assumed that LaeA is involved in the reversion of gene repression at the heterochromatin level (Reyes-Dominguez et al. 2010). In *T. reesei*, its ortholog LAE1 is reported to be essential for cellulase and hemicellulase formation in *T. reesei* and regulates only a minor number of genes involved in secondary metabolism (Karimi-Aghcheh et al. 2013). However, an examination of the histone methylation pattern around the cellulase genes showed that the promoters of these genes are still accessible for regulatory proteins (Seiboth et al. 2012). The mechanism of how LAE1 controls expression of these plant-cell-wall-degrading enzymes therefore seems to be more complex than it would be anticipated from the data obtained with *Aspergillus* spp.

C. Regulators of Enzyme Production

The regulation of plant-cell-wall-degrading enzymes in *T. reesei* (Fig. 10.1) was studied for decades and detailed reviews are available (Aro et al. 2005; Kubicek et al. 2009). Production of extracellular enzymes is a process that consumes a high amount of energy; therefore, different sensing and signaling networks communicate the presence of the substrates in the fungal environment to ensure an economic production of the enzymes. Cellulase and related genes are therefore not produced constitutively but depend on specific induction and are usually regulated in a coordinated way (Foreman et al. 2003; Häkkinen et al. 2012). The major antagonistic control circuit to induction is general carbon catabolite repression (CCR). Usually, repression is connected to glucose, although other carbon sources are able to provoke CCR. In *T. reesei*, most CAZymes related to plant cell wall degradation are repressed by the presence of fast metabolizable monomeric carbon sources. With respect to their regulation, at least three transcriptional activators, including XYR1, ACE2, and the HAP2/3/5 complex, as well as the two repressors CRE1 and ACE1 are involved in cellulase and hemicellulase regulation. XYR1 is clearly the major activator of cellulolytic and xylanolytic enzymes, and its absence eliminates their expression by all known inducers (Herold et al. 2013; Stricker et al. 2006). XYR1 is also necessary to activate the multifunctional aldose reductase XYL1 necessary for the catabolism of D-xylose, L-arabinose, and lactose (Akel et al. 2009; Seiboth et al. 2007). Although necessary for both cellulase and xylanase, induction it is not clear how XYR1 is able to selectively induce either cellulases or xylanases in

response to different inducers. Intriguingly, while regulation of both cellulases and xylanases by an XYR1 homolog is conserved in *Aspergillus* spp., this is not the case in *Fusarium* or *Neurospora*, for which only xylanases are XYR1 targets (Znameroski and Glass 2013). Further important regulators of GHs are the photoreceptors BLR1, BLR2, and ENV1, of which BLR1 and BLR2 are supposed to form a transcription factor complex (Castellanos et al. 2010). Light-dependent regulation of cellulases and GHs is at least in part achieved by cross talk of the light-signaling pathway (BLR1, BLR1, and ENV1) with other signaling pathways (Schmoll et al. 2010a).

A large-scale screening for further genes of impact on cellulase gene expression (Pakula et al. 2011a, b, c) recently revealed 15 genes whose overexpression or deletion increased cellulase and hemicellulase production in *T. reesei* QM6a. Most of them make up Zn2Cys6 transcription factors, but two other GCN5-related acetyltransferases are included. The function of these genes within the regulatory network of plant cell wall degradation remains to be studied.

D. Tracing Hypercellulase Production

The extensive use of *T. reesei* in biotechnology is accompanied by extensive efforts for strain improvement using diverse strategies (Seidl and Seiboth 2010). One approach to better understand the biology underlying cellulase expression and its hyperproduction is the analysis of improved producer strains derived from classical mutagenesis programs. It is reported that improved **industrial strains** secrete today more than 100 g of cellulases per liter in industrial fermentations (Cherry and Fidantsef 2003). With the annotated genome sequence of the ancestor strain QM6a and the development of high-throughput methods such as massively parallel DNA sequencing and comparative genomic hybridization (CGH), it is now possible to trace the genomic changes that occurred during the different mutation steps.

These mutation programs yielded strains with considerable genome rearrangements, single-nucleotide variants (SNVs), insertions, and losses of genomic regions of more than 85 kb in case of RutC30, an early high-producing mutant (Le Crom et al. 2009; Seidl et al. 2008; Vitikainen et al. 2010).

Interestingly, with the exception of CRE1, none of the known transcriptional regulators of cellulase expression were affected in these strain lines by the different rounds of mutagenesis (Vitikainen et al. 2010). These results not only highlight additional mechanisms that can be of importance for cellulase hyperproduction but also underscore the complexity of the problem to find out which mutations are detrimental, neutral, or beneficial for cellulase production. Interestingly, three genes mutated in the early cellulase-enhanced mutant QM 9414 (81136, 124295, and 108540) encode putative cell wall or plasma membrane proteins. It is possible that an altered cell wall/membrane structure may have been beneficial for spore survival and germination during mutagenesis, but it cannot be ruled out that these genes are also relevant for improved cellulase production. Such an assumption would be supported by two other mutated genes involved in cytoskeleton function (2439 encoding a ARP2/3 complex protein and 35386 encoding actin-interacting protein 3) and suggest that the improvement of cellulase production in QM 9414 is caused by mutations in cellulase trafficking and secretion.

The genome of one cellulase-negative *T. reesei* mutant (QM 9136) revealed a frameshift mutation in the gene encoding the transcriptional activator of cellulase and hemicellulase gene expression, *xyr1* in one strain (B. Seiboth et al., unpublished data).

E. Genome-wide Analysis of Enzyme Regulation

Transcriptomic investigations with *T. reesei* focused on the response of plant-cell-wall-degrading enzymes to different inducing carbon sources (cellulose, xylan, and derivatives)

and recombinant protein production. Hence, the most detailed data on regulation of GHs and enzyme expression are from this fungus.

Comparative transcriptomic studies of induced, noninduced, and repressed *T. reesei* cultures not only provide valuable insights in the regulation of already characterized genes but are well suited to identify novel genes whose expression correlates with cellulase and hemicellulase gene expression and are therefore likely to be involved in plant cell wall degradation. Before the availability of the *T. reesei* QM6a genome sequence, Foreman et al. (2003) sequenced different *T. reesei* cDNA clones from cultures growing on the cellulase inducer lactose and identified 12 previously unknown enzymes related to biomass degradation. Their analysis revealed numerous coregulated genes under different cellulase-inducing conditions (growth on lactose, induction by sophorose) in the wild-type QM6a and the RL-P37 cellulase hyperproducer. Genes involved in protein processing and secretion were not significantly enriched among the genes regulated under cellulase-inducing conditions (Foreman et al. 2003).

Interestingly, recent studies showed that coregulation of cellulases is not reflected in secretome analysis of *T. reesei* hyperproducers (Herpoel-Gimbert et al. 2008; Jun et al. 2011). The most obvious reason for this finding could be that the regulatory machinery in the mutants used might have been altered because of repeated mutagenesis; alternatively, proteolysis may influence the result and interfere with synergy in the enzymatic mixture.

However, there is also support for the hypothesis that posttranslational regulation might occur. A study on the function of a novel dehydrogenase revealed discrepancies between regulation of cellulase transcription, abundance of CBH1 in the culture medium, and activity (Schuster et al. 2011). Studies investigating the effect of photoreceptors on cellulase gene expression in *T. reesei* (Gyalai-Korpos et al. 2010) and *N. crassa* (Schmoll et al. 2012) indicated that at least with respect to light response, posttranslational regulation of cellulase gene expression is a conserved phenomenon.

Today, a number of whole-genome microarray analyses are available. Häkkinen et al. (2012) used **oligonucleotide microarrays** to study the transcriptomic **response to different substrates** in the hypercellulolytic *T. reesei* strain RutC30. The largest set of induced genes was detected in the cultures with the hemicellulosic material (68–124 genes in 39–47 CAZy families), whereas cultivation in the presence of pure cellulosic or cellulose-derived materials, including the inducer sophorose, resulted in a clearly smaller number of genes induced (43–58 genes in 28–36 families). A common set of genes was defined that is usually expressed in the presence of the lignocellulose substrates. It included GH6 cellobiohydrolase CEL6A; GH5 endoglucanase; xylanases of families GH10, GH11, and GH30; GH5 β-mannanase; GH3 family β-glucosidases and β-xylosidases; GH27 α-galactosidases; GH2 β-mannosidases; acetyl xylan esterases of families CE3 and CE5; glucuronoyl and acetyl esterases of families CE15 and CE16; GH31 α-glucosidases/α-xylosidases; GH54 and GH43 α-L-arabinofuranosidases (or β-xylosidase/α-L-arabinofuranosidases); GH61 lytic PMOs; GH55 β-1,3-glucanases; GH67 α-glucuronidase; GH79 β-glucuronidase; GH105 rhamnogalacturonylhydrolase; GH95 α-L-fucosidase; GH89 α-N-acetylglucosaminidase; and chitinases of family GH18 and β-N-actetylhexosaminidases of GH20. Cluster analysis also provided a more refined pattern of coexpressed genes on various substrates. Interestingly, with *T. harzianum*, the major hydrolytic enzymes secreted on growth on cellulosic medium were chitinases, although also cellulases were detected (Do Vale et al. 2012).

Oxidative pathways were found to participate in plant cell wall degradation in fungi (Beeson et al. 2012; Harris et al. 2010). However, this process is not studied in detail in *Trichoderma*. Expression of a heterologous laccase in *T. reesei* caused downregulation of the native cellulases in continuous culture (Rautio et al. 2006). However, because lactose was used as a carbon source in this study, it remains to be elucidated whether there is indeed a mechanism that governs hydrolytic versus oxidative degradation of cellulosic material.

Among the GHs, the group of chitinases was investigated in detail as well (Hartl et al. 2012). Chitinases of subgroups A and B and the new subgroup C were found to be upregulated before or on contact with *Fusarium* spp. by an efficient biocontrol isolate of *Trichoderma gamsii* (Matarese et al. 2012). However, considerable differences in regulation of these genes were observed on contact with a host or in the presence of different nutrients in *T. virens* and *T. atroviride* even if they are orthologs in the two fungi (Gruber et al. 2011).

In addition to regulation in response to different carbon sources and inducers, GHs, including cellulases, are regulated by light and the photoreceptors BLR1, BLR2, and ENV1 (Castellanos et al. 2010; Schmoll et al. 2005). This effect is also conserved in *N. crassa* (Schmoll et al. 2012), hence indicating a general regulation of extracellular polysaccharide degradation by light in fungi, which is also supported by earlier findings in fungi in general (Tisch and Schmoll 2010). The connection of light response to nutrient signaling is at least in part established by cross talk with the heterotrimeric G-protein pathway. Light-dependent regulation of cellulase gene expression was observed for the two G-protein alpha subunits GNA1 and GNA3 (Schmoll et al. 2009; Seibel et al. 2009). The G-protein beta and gamma subunits GNB1 and GNG1 as well as their putative cochaperone, the phosducin-like protein PhLP1, have an impact on light-dependent regulation of up to 51 GHs. In summary, 99 GHs were found to be potentially regulated by light in the wild type or on deletion of a signaling factor (up to 34-fold up or down), which indicates condition-dependent light regulation of substrate degradation in *T. reesei* (Tisch et al. 2011b). Also, for the cAMP pathway, which acts downstream of the heterotrimeric G-protein pathway, a light-dependent function in regulation of cellulase gene expression was detected. While adenylate cyclase has a consistently positive effect on cellulase gene expression, protein kinase A acts positively in light but negatively in darkness (Schuster et al. 2012a). This effect can in part be explained with the putative function of protein kinase A as a priming kinase for regulation of the photoreceptor complex as shown in *N. crassa* (Huang et al. 2007), which in turn regulates cellulases. For the components of the heterotrimeric G-protein pathway as well as the cAMP pathway, a dampening effect on light responsiveness of cellulase gene expression was detected (Schuster et al. 2012a; Tisch et al. 2011b).

VI. Physiology

Detailed knowledge on the physiology of a fungus is crucial for its application in biotechnology. For *Trichoderma* spp., the main focus of most studies was growth and enzyme production on diverse substrates. However, response to stress, carbon concentration, injury, or low levels of oxygen were also studied at a genome-wide level.

Trichoderma spp. are adapted to a natural habitat with predominantly plant cell wall material as substrate. Glucose-dependent regulation of gene transcription in *T. reesei* differs considerably from *S. cerevisiae*. Hence, adaptation to an environment dominated by recalcitrant substrates such as cellulose and hemicellulose is suggested to correlate with a shift of ATP-producing pathways from fermentation to respiration (Chambergo et al. 2002).

Interestingly, conditions of low growth rates in carbon-limited **chemostat cultivations** correlate with high specific enzyme production, which reflects response of *T. reesei* to carbon limitation on a cellulase-inducing substrate (lactose). At the same time, an enrichment of GCN5-related genes among those upregulated under these conditions occurs (Arvas et al. 2011). Consequently, chromatin rearrangement is likely to be involved in regulation of enzyme production for carbon acquisition, which is in accordance with studies on correlation of cellulase regulation and epigenetic events (Seiboth et al. 2012; Zeilinger et al. 2003). In addition, two GCN5-related N-acetyltransferases were found to positively influence cellulase gene expression (summarized in Kubicek 2013).

A significant nonrandom distribution of significantly expressed genes in the genome was observed, which agrees with the clustering of CAZyme expression (summarized in Kubicek 2013). Based on the presence of GCN5-related acetyltransferase genes, these specific chromosomal regions might be regulated by histone acetylation. Histone modifications could therefore be a major factor that determines the accessibility of the genome to more specific transcription factors. This would also explain why XYR1 dependent on the inducer can selectively activate cellulases or xylanases (Kubicek et al. 2009).

The expression of many genes encoding secreted proteins and such involved in secondary metabolism, as well as various lineage-specific, mostly orphans and unknown, genes, is positively correlated with the specific extracellular protein production rate, which in turn was highest at the lowest growth rate (0.03 h^{-1}). The major biosynthetic activities, in contrast, were negatively correlated with extracellular protein production. Transcriptomic and proteomic analysis suggested that the growth rate and the cell density strongly influenced the flux through the upper part of glycolysis or the TCA (tricarboxylic acid) cycle. At low specific growth rates, this results in a low flux to biomass, which could be the fundamental determining factor of protein production, by induction of lineage-specific genes and regulatory factors required for this condition (Arvas et al. 2011).

Portnoy et al. (2011) studied the role of the specific growth rate on CCR in *T. reesei*. Using chemostat cultures with glucose as a carbon source and a recombinant strain in which the carbon catabolite repressor gene *cre1* was removed, a complex interplay between CCR and the growth rate occurred. A CRE1-independent repression of gene expression by high specific growth rates was observed, which most strongly affected genes involved in C1 and carbohydrate metabolism. In this regard, it is important to stress that the results obtained by Arvas et al. (2011) were obtained with strain RUTC30, which harbors a truncated *cre1* gene; therefore, the effects of the specific growth rates on the transcriptome must be viewed with caution. Interestingly, none of the cellulase genes was significantly (>log 1.5-fold) expressed on glucose in the *cre1* deletion strain and at low growth rates, indicating that induction is essential for their expression (Portnoy et al. 2011).

Protein production at high levels, production of heterologous proteins, or addition of different oxidizing chemicals stimulates the **unfolded protein response**, a mechanism that enables the cell to activate the production of molecular chaperones involved in protein folding. Investigation of secretion stress in fermentations with *T. reesei* revealed 457 genes that are assumed to be involved in this process. Of these genes, 20 were found to be UPR-specifically regulated. Among the genes upregulated under secretion stress are different histone genes (encoding H2A and H4) as well as the cross-pathway controller *cpc1* (the *S. cerevisiae GCN4* ortholog), which is activated on amino acid deprivation. Accordingly, also the homolog of a transcriptional coactivator of *S. cerevisiae* Gcn4p, MBF1, was enhanced under UPR conditions. Hence, CPC1 seems to induce genes involved in glutathione biosynthesis to alleviate the lack of reducing power. The cellulases *cbh1* and *egl1* are repressed under DTT treatment and in tissue plasminogen activator (tPA) producing culture (Arvas et al. 2006).

Investigation of physiological processes in *T. reesei* revealed a delicate regulation of transcription during chemostat cultivations. Even small perturbations that did not cause measurable changes in biomass or product concentrations were detectable as changes in transcript levels (Rautio et al. 2006). In addition, the ratio between polyA RNA and biomass as well as that between polyA RNA and total RNA was found to be growth phase dependent in batch cultures (Rautio et al. 2007).

Although the response to **mechanical injury**, which triggers conidiation in *T. atroviride* (Casas-Flores et al. 2004), was not studied in a biotechnological setting, it should be kept in mind that in a fermentor, mycelial injury is likely to occur. In *T. atroviride*, injury stimulates reactive oxygen species (ROS) formation by the NADPH oxidase NOX1. Early after injury, genes involved in calcium signaling

and transport, redox balance, stress responses, cell cycle, and cell death as well as transcription factors are upregulated, while metabolic genes and those involved in ROS scavenging are downregulated. Later, induced genes comprise those involved in DNA damage response and cell cycle as well as several genes encoding proteins with oxidoreductase activity. Again, at later stages after injury, metabolic genes are downregulated (Hernandez-Onate et al. 2012). Hence, on injury, expression of genes for control of damage is enhanced at the expense of metabolism.

VII. Development

Development in fungi occurs in response to environmental conditions (Fig. 10.1) and is balanced between vegetative growth, production of chlamydospores or conidia (asexual development), and induction of sexual development. All these stages of development have been found in nature for *Trichoderma* spp. but are often not easily reproduced under laboratory conditions. As a primary means of dissemination, sexual and asexual spores serve as inoculum causing fungal diseases as well as plant protection by fungi.

A. Asexual Development

The induction of conidiation (i.e. asexual development) occurs as a response to a multitude of environmental conditions, such as light, injury, availability and type of carbon and nitrogen source, pH, calcium concentration, and circadian rhythms (Carreras-Villasenor et al. 2012).

Blue light represents one of the major environmental cues to induce sporulation. The blue light photoreceptors BLR1 and BLR2 are essential for light-induced conidiation in *T. atroviride* (Casas-Flores et al. 2004) and in *T. reesei* (Castellanos et al. 2010). Lack of the light regulatory protein ENV1, however, causes not only strongly reduced conidiation in light but also a severe growth defect in *T. reesei* wild-type and cellulase-enhanced mutant strains (Schmoll et al. 2005; Seibel et al. 2012b).

In *T. atroviride*, more than 200 genes showed a specific response to blue light, with enrichment in functions in metabolism, stress, cellular transport, cell cycle, and DNA processing (Carreras-Villasenor et al. 2012). Among these genes, also more than ten transcription factors were detected, which supports a flat hierarchy of gene regulation in response to light as observed in *N. crassa* (Smith et al. 2010) also for *Trichoderma* spp.

Conidiation is also dependent on the carbon source (Chovanec et al. 2001), and light enhances the extent of conidiation in many, but not all, cases (Friedl et al. 2008). In accordance with the response of metabolic genes to blue light, an increased abundance of genes encoding glycosyl hydrolases, especially of cellulases, was observed during conidiation. In addition, genes involved in secondary metabolism were found to be enriched among the genes upregulated during conidiation (Metz et al. 2011). The **spore surface** of *T. reesei* was found to be covered with cellulolytic enzymes (Kubicek et al. 1988), which explains the enhanced production of cellulases at this stage. The strong presence of different CAZymes might, however, be a peculiarity of *T. reesei* because, during conidiation of other fungi (i.e., *N. crassa* or *A. fumigatus*), no upregulation of cellulases or hemicellulase genes was detected (Greenwald et al. 2010; Lamarre et al. 2008).

Although light represents an important cue for initiation of conidiation and BLR1 and BLR2 are necessary for conidiation when induced by glucose deprivation, hence indicating cross talk between light response and carbon utilization, the photoreceptors were found not to be required for **conidiation after injury** (Casas-Flores et al. 2004). Investigation of the mechanism underlying the induction of conidiation by injury revealed a dependence of this response on NADPH oxidases (NOXs), especially NOX1, while production of ROS was found to be transient. However, NOX mutants are not defective in light-induced conidiation; hence, the effect of NOX1 on injury-induced

conidiation is not generally targeted to this process (Hernandez-Onate et al. 2012). Deletion of the regulatory gene noxR targeted oxylipin biosynthesis (Hernandez-Onate et al. 2012), which regulates conidiation in *A. nidulans* and *Penicillium chrysogenum* (Hegedus et al. 2011). Hence, oxylipins are likely to play an important role in injury response in *T. atroviride* through the formation of lipoperoxides on injury.

B. Sexual Development

Already in 1865, asexual *Trichoderma* spp. were linked to the sexual state of *Hypocrea* spp. (Tulasne and Tulasne 1865). Nevertheless, it took more than 100 years until sexual development, as it obviously occurs frequently in nature, could be reproduced in the lab (Seidl et al. 2009a). Investigation of sexual development in *Trichoderma* is still in its beginnings (for a detailed review, see Schmoll 2013). *Trichoderma reesei* was found to be **heterothallic** because of the presence of two different idiomorphs at the mating-type locus, which is also true for all other *Trichoderma* spp. for which the genome sequence has been published (M. Schmoll, unpublished). Besides an alpha type peptide pheromone precursor gene (*ppg1*), the pheromone system of *T. reesei* also comprises a **novel type of peptide pheromone precursors** belonging to the h-type, *hpp1*, which assumes a-type function (Schmoll et al. 2010b). Two pheromone receptors, HPR1 and HPR2, are responsible for reception of the pheromone signal and crucial for fruiting body formation in their cognate mating type (Seibel et al. 2012a). Along with several other genes assumed to be involved in sexual development, also *hpp1* is upregulated in response to light and regulated by the phosducin-like protein PhLP1 as well as the G-protein alpha and beta subunits GNB1 and GNG1 (Tisch et al. 2011b), which is in agreement with the preferential occurrence of sexual development in light in *T. reesei* (Seidl et al. 2009a). Accordingly, also the photoreceptors BLR1 and BLR2 as well as ENV1 are involved in regulation of sexual development (Seibel et al. 2012b). Investigating different light conditions and photoreceptor mutants, Chen et al. (2012) found a correlation between upregulation of plant-cell-wall-degrading enzymes and conditions enabling sexual development.

A major drawback with sexual development in *T. reesei* is that it can only undergo sexual development with fertile wild-type strains because QM6a and all its derivatives were found to be female sterile (Seidl et al. 2009a). Although lack of pheromone precursors as well as lack of *env1* in light cause female sterility in *T. reesei*, these mutations are not responsible for female sterility of QM6a (Seibel et al. 2012a, b). While numerous mutations are known to cause this defect in fungi, the molecular basis of female sterility of QM6a has not been elucidated so far and is still under investigation.

VIII. Interactions of *Trichoderma* spp. with Fungi and Plants

The best-known biotrophic interaction of *Trichoderma* is its ability to antagonize or even parasitize and kill other fungi (a form of predation or necrotrophic hyperparasitism conventionally called mycoparasitism). This ability is today widely used in agriculture to combat phytopathogenic fungi in biological control of plant pests ("biocontrol") (Benitez et al. 2004; Lorito et al. 2010; Mukherjee et al. 2013). However, more recently it became clear that some soil *Trichoderma* species also have direct positive effects on the plants either by stimulating their self-defense against the pathogenic fungi or by stimulating their growth and tolerance to environmental stresses (Harman et al. 2004).

A. Mycoparasitism

Mycoparasitism is the innate property of the genus *Trichoderma* (Druzhinina et al. 2011). It consists of several phases: recognition of the prey; defending against its action; attachment to it; and finally killing and feeding on the prey. Recent transcriptome analyses revealed common and species-specific changes during the precontact phase of antagonism, when recognition is assumed to take place. Moreover, it has

been shown that individual species can use different strategies for mycotrophy: *Trichoderma virens* and *T. atroviride* display completely different transcriptomic responses when confronted with *Rhizoctonia solani*. *Trichoderma virens* only overexpressed 78 genes (1 % of the total expressed genes), of which those involved in formation of gliotoxin biosynthesis and its precursor metabolites accounted for the largest group, whereas *T. atroviride*, in contrast, overexpressed 400 genes (about 5 % of all expressed genes), and almost no orthologs were shared between both species. As *T. atroviride* resembles a more ancestral state of *Trichoderma* than *T. virens* (Kubicek et al. 2011), this dissimilarity in mycotrophic strategies suggests that *T. atroviride* follows the "traditional" way of parasitism, characterized by "weak" predation and by delaying the death of the host organism. *Trichoderma virens*, in contrast, is more specialized and aims at true predation and killing of its prey.

However, there were also general responses shared between *T. virens* and *T. atroviride* when confronted with *R. solani*, suggesting that they may represent general features of mycotrophy and antagonism: one was the overexpression of a high number of genes for **proteolytic enzymes** and **oligopeptide transporters**. Most of these proteases belonged to the subtilisin-like serine protease group, which is expanded in *Trichoderma* (Kubicek et al. 2011). These data are also consistent with an analysis of ESTs accumulating during the onset of contact between *T. atroviride* and *R. solani* and *Sclerotinia sclerotiorum* (Seidl et al. 2009b) and genetic proof for the involvement of one of these proteases from *T. atroviride* (*prb1*) in mycoparasitism (Flores et al. 1997). Seidl et al. (2009b) hypothesized that the arising oligopeptides are bound by receptors, which sense the nitrogen status of the medium, a mechanism reminiscent of nematophagous fungi, where trapping of the prey is induced by oligopeptides from the nematodes (Dijksterhuis et al. 1994). This so-far-unknown nitrogen sensor could be a member of the class IV nitrogen sensors of the nine groups of GPCRs present in *T. atroviride* (Seidl et al. 2009b). Candidates for this function are available in the genomes of *Trichoderma* spp. In addition, *T. atroviride* GPR1 (Triat2:160995), which belongs to the class of cAMP receptor-like proteins, is required for antagonism (Brunner et al. 2008).

Another event, common to both *T. virens* and *T. atroviride*, was the induction of genes of the **heat shock response** such as HSP23, HSP70, HSP90, and HSP104; genes of oxidative stress response (cytochrome C peroxidase, proline oxidase, and ER-bound glutathione-S-transferases); and genes for **detoxification** processes (ABC efflux transporters, the pleiotropic drug resistance [PDR] transporters, and the multidrug resistance MDR-type transporters). Accordingly, an ABC transporter from *T. atroviride* (TAABC2) is involved in the biocontrol of *R. solani* (Ruocco et al. 2009).

Rhizochtonia solani has been shown to use radical oxygen species as signaling molecules during sclerotia formation (Papapostolou and Georgiou 2010) and to excrete antifungal components (Aliferis and Jabaji 2010), both of which may have elicited this response. This illustrates that the prey pose considerable defense toward the mycoparasite. Production of ROS represents a common defense response in fungi (Silar 2005). Besides positively affecting genes encoding hydrolytic enzymes such as proteases, cellulases, or chitinases, the NADPH oxidase NOX1 contributes to efficient biocontrol of *P. ultimum* (Montero-Barrientos et al. 2011). The relevance of genes involved in metabolic functions was also confirmed by investigation of *T. atroviride* in the presence of different fungal hosts, with proteases, acetyl xylan esterase 1 (*axe1*), and swollenin (*swo1*) (Reithner et al. 2011).

The actual final killing of the prey is brought about by the synergistic action of **antimycotic secondary metabolites** and cell wall hydrolytic enzymes. Their importance to the lifestyle of *Trichoderma* is also reflected in the amplification of the respective genes in *Trichoderma* (Kubicek et al. 2011). Peptaibols have been shown to inhibit the growth of prey fungi in synergism with cell wall lytic enzymes (Schirmböck et al. 1994), which has been explained by an inhibition of resynthesis of the cell wall polymers by the respective

membrane-bound synthases (Lorito et al. 1996). However, with the exception of gliotoxin, no other nonribosomal peptides – including the peptaibols – or polyketide has been shown by reverse genetics to be functionally involved in killing of the prey.

B. *Trichoderma*-Plant Interactions

Trichoderma spp. have been known for decades to be "rhizosphere competent," that is, to grow and develop within the plant rhizosphere without causing diseases but eventually antagonizing other pathogenic microorganisms (Lewis and Papavizas 1984). However, metagenomic analyses showed that *Trichoderma* spp. are rather infrequent in soils and the rhizosphere (Buee et al. 2009; Lim et al. 2010). Friedl and Druzhinina (2012) demonstrated low in situ diversity of *Trichoderma* in temperate soil and proposed that soil is not an original habitat of these fungi. It is likely that *Trichoderma* spp. in soil are those that developed high environmental opportunistic abilities and may also develop in several other habitats or those species that followed their prey (other fungi) into the soil (Druzhinina et al. 2011). Nevertheless, several *Trichoderma* spp. can trigger an **induced systemic resistance** (ISR) in the plants whose roots they meet in the rhizosphere. This induction starts with the recognition of microbe-associated molecular patterns by a pattern recognition receptor of the plant, which subsequently activates a primary defense response in the plant. *Trichoderma* molecules that have been shown to trigger ISR include secreted xylanases, cellulases, and the cellulose-binding protein swollenin (for review, see Shoresh et al. 2010); small cysteine-rich secreted proteins (Djonovic et al. 2006b; Seidl et al. 2006a); peptaibols (Leitgeb et al. 2007; Viterbo et al. 2007); and an unknown PKS-NRPS product (Mukherjee et al. 2012b). In all these cases, knockout in the respective genes did not impair the ability of *Trichoderma* to colonize the roots, although the induction of ISR was abolished in most cases. Thus, *Trichoderma* does not seem to benefit from the plant's response on first glance.

Harman et al. (2004) advocated the interpretation that *Trichoderma* spp. are plant symbionts, but based on available data, this hypothesis can neither be proved nor rejected (Druzhinina et al. 2011). However, some *Trichoderma* taxa (including several novel species) have been reported to live inside the plants as **endophytes** and thereby contribute to the health of the plants by promoting plant growth and delaying onset of drought stress and inhibition of pathogens (Bae et al. 2009; Hanada et al. 2008, 2010; Moy et al. 2002; Samuels 1996; Tejesvi et al. 2006). Druzhinina et al. (2011) speculated that mycotrophs may have become endophytes by entering the plant roots by parasitism of the mycorrhizae of the plant as described by de Jaeger et al. (2010).

Trichoderma association with roots can also result in a promotion of plant growth. For example, *T. virens* has been shown to increase *Arabidopsis* biomass production and lateral root growth, and plant mutants in the auxin-mediated response pathways are reduced in these effects (Contreras-Cornejo et al. 2009). As for the molecules mediating this effect, best evidence is available for the plant hormone ethylene. It is formed in the plant from S-adenosyl-L-methionine and then converted to 1-aminocyclopropane-1-carboxylic-acid (ACC) by the enzyme ACC synthase before being converted to ethylene by ACC-oxidase (Wang et al. 2002). Viterbo et al. (2010) showed that T203 (reidentified as *T. asperelloides*; Samuels et al. 2010) expresses an α-1-aminocyclopropane-1-carboxylate (*acc1*) deaminase gene during interaction with roots of *Brassica napus* (canola). A knockout in this gene, whose enzyme product cleaves ACC and thereby lowers the level of ethylene, reduced the ability of the fungus to promote root elongation. A transient high concentration of ethylene is required to break seed dormancy, but following germination, a sustained high level of ethylene inhibits root elongation. Thus, *T. asperellum*, like plant growth-promoting bacteria that contain the enzyme ACC deaminase (Glick 2005), possesses a mechanism for ensuring that the ethylene level in the plant does not become elevated to the point at which root growth is impaired and thus facilitates the formation of longer roots, thus positively influencing its own habitat.

IX. Secondary Metabolism

Numerous secondary metabolites from various *Trichoderma* spp. have been reported (for review, see Mukherjee et al. 2012a; Vinale et al. 2012), but because of a high level of uncertainty about the species identity of the producing organisms (most *Trichoderma* strains were until recently only identified by insufficient morphological characters), their production by *T. reesei*, *T. virens*, and *T. atroviride* cannot be deduced. Also, the genes for synthesis of most metabolites have not yet been identified. However, genes encoding nonribosomal peptide synthetases (NRPS) and the PKSs can be predicted in silico because of their unique protein structure. *Trichoderma reesei*, *T. atroviride*, and *T. virens* contain 11, 18, and 18 PKS and 10, 16, and 28 NRPS genes, respectively. While the numbers for PKS are average within fungi, the NRPS genes of *Trichoderma*, particularly *of T. virens*, are among the highest of all fungi. Kubicek et al. (2011) suggested that this is apparently because of recent duplications of cyclodipeptide synthases, cyclosporin/enniatin synthase-like proteins, and NRPS-hybrid proteins. Half of the NRPS genes present in *T. atroviride* or *T. virens* are unique for the respective species and occur within nonsyntenic islands of the genome, indicating their origin by recent genome rearrangements, which is also reflected in a higher nucleotide dissimilarity (about 30 %) than the average of genes between the three *Trichoderma* spp.

A unique feature of *Trichoderma* and close relatives from such families as Hypocreaceae, Clavicipitaceae, and Bionectriaceae is the presence of two NRPSs that synthesize modified peptides, termed **peptaibols**; these are small (500–2,000 Da) linear peptides containing a high number of nonproteinogenic, α,α′-dialkylated α-amino acids like isovaline and α-aminoisobutyric acid (Aib). In addition, their N-terminal amino acid is acetylated, and the C-terminus is reduced to an amino alcohol, mostly phenylalaninol. These properties have given rise to the name peptaibol (***pept***ide, ***Aib***, and amino alcoh***ol***). Peptaibols form a helical structure with the hydrophobic side chains exposed to the surface, which allows them to interact with natural and artificial bilayers and form pores or voltage-dependent ion channels, increasing membrane permeability (Duclohier 2007). *Trichoderma* spp. secrete a broad spectrum of peptaibols, which can be summarized as heterogeneous peaks falling into three size categories (11, 14, and 18–20 kDa). One of the two synthases thereby synthesizes both small peptaibols (Degenkolb et al. 2012; Mukherjee et al. 2011), probably by performing module skipping (Degenkolb et al. 2012). The larger enzyme synthesizes a single peptaibol only. Further kinetic regulation by, for example, precursor supply then gives rise to the multitude of peptides produced by these two enzymes. Stoppacher et al. (2008), using LC/MS, detected 20 trichorzianes and 15 trichoatrokontins (representing the large and the two smaller peptaibols, respectively) in culture filtrates of *T. atroviride*.

Two other NRPS-encoding genes are present in *T. virens* (and one of them in *T. reesei*) but not in *T. atroviride*. They synthesize the epipolythiodioxopiperazine-type peptides **gliotoxin** and **gliovirin** (Patron et al. 2007). In *T. virens*, the former is exclusively produced by so-called P strains, whereas the latter is produced only by the Q strains (Mukherjee et al. 2012b). The gliotoxin synthase and its auxiliary biosynthetic enzymes are located in a cluster containing eight genes that are closely similar to those of *A. fumigatus* (Patron et al. 2007). Gliotoxin has a fungistatic action but is also known from other fungi, including the opportunistic human pathogen *A. fumigatus*, where it acts as a virulence factor (Dagenais and Keller 2009). Its toxicity is due to the presence of a disulfide bridge in the molecule, which can inactivate proteins via reaction with thiol groups and generate ROS by redox cycling (Gardiner and Howlett 2005).

Two further NRPSs of *Trichoderma* have been identified to encode genes for **siderophore biosynthesis**. Despite the presence of only 2 genes, Lehner et al. (2013) reported that *Trichoderma* spp. produced 12–14 siderophores, with 6 common to all species tested. The production of diverse products by a single NRPS synthase therefore seems to be a general

property of these enzymes. None of the other *Trichoderma* NRPSs have as yet been characterized; consequently, their products are unknown.

As for PKS-encoding genes, 29 of a total of 47 in the three sequenced *Trichoderma* spp. fall into a single orthologous group (Baker et al. 2012). The loci of these PKSs are frequently flanked by cytochrome P450 monooxygenases, FAD-dependent monooxygenases, short-chain dehydrogenases/reductases, or epimerases next to PKS-encoding genes, stressing the previous suggestion that these are auxiliary components in the biosynthesis of the respective metabolites. The *T. reesei* PKS gene Trire2:82208 has been shown by gene deletion to be responsible for the green color of the conidia but, interestingly, also is involved in protection against other fungi but less in the direct attack on them (Atanasova et al. 2013). The genomes of the three *Trichoderma* spp. also harbor genes encoding PKS-NRPS hybrids (two, four, and four in *T. reesei*, *T. atroviride*, and *T. virens*, respectively).

Apart from NRPS and PKS, fungi also synthesize isoprenoid secondary metabolites. The genomes of *T. reesei*, and *T. atroviride* do not contain genes that can obviously be related to these activities. However, Mukherjee et al. (2006) identified a cluster in the genome of *T. virens* that includes three cytochrome P450 genes and one terpene cylase. Because these genes are underexpressed in the mutant that is unable to produce viridin and viridiol, they proposed that this cluster could be responsible for the production of viridin.

Trichothecenes belong to the more prominent isoprenoid-derived secondary metabolites claimed to be synthesized by *Trichoderma* . These are compounds that have a family of over 200 secondary metabolites with a common cyclic 12,13-epoxytrichothec-9-ene (EPT) structure. They bear a characteristic internal disulfide bridge, which is responsible for all known toxic effects of these molecules, such as induction of apoptosis, inhibition of the proteasome, or inhibition angiogenesis. Exposure to trichothecenes can cause a number of symptoms, from feed refusal, immunological problems, vomiting, and skin dermatitis to immunosuppressive effects and neurotoxicity (Ueno 1985). The reputation of *T. reesei* as a trichothecene producer is because of the problems with *Trichoderma* species identification (Druzhinina and Kubicek, unpublished) and eventually strain mixup (Watts et al. 1988). In any case, the genome sequence of *T. reesei* lacks the necessary *tri5* ortholog encoding the trichothecene synthase and is thus principally unable to initiate trichothecene biosynthesis. Nielsen et al. (2005) showed that species from *Trichoderma brevicompactum* (now defined as a clade, including *T. brevicompactum*, *T. arundinaceum*, *T. turrialbense*, and *T. protrudens*) are in fact the only trichothecene producers of the genus *Trichoderma*.

X. Conclusions and Outlook

Environmental problems as well as the anticipated shortage of fossil fuels have initiated efforts to increase sustainability in virtually all aspects of life and enhance green economy. *Trichoderma* spp. can support this movement with environmentally friendly applications from production of second-generation biofuels to bioremediation and biological control of plant diseases.

Because of their broad applicability, investigation of *Trichoderma* species will benefit multiple research fields and represents a key research area for a sustainable future. With the huge amount of genome-wide data generated in recent years, we are now facing the challenge of interpreting them in a broader context. In many cases, the respective studies showed genes relevant under biotechnological conditions also were crucial for plant protection or mycoparasitism and vice versa. Considering the natural environment of *Trichoderma* spp., these findings reflect the need of *Trichoderma* to combat its competitors and kill basidiomycetes when feeding on the wood predegraded by them and should now stimulate enhanced networking between researchers focused on biocontrol and biotechnology.

One crucial aim in research with *Trichoderma* is the construction of a whole-genome

knockout library, as already available for *N. crassa*. It will take a community effort and collaboration between several labs working on different species and a contribution of industry to coordinate this important milestone.

In this respect, it will also be of importance to investigate the conditions for sexual development in other species of *Trichoderma*, which have been shown to display a sexually recombining history (e.g. see Druzhinina et al. 2010b, 2012) as crossing of strains with different characteristics has proven to be a versatile tool for both research and industrial strain improvement.

However, with increasing application of plant-protecting strains in the environment, the consequences of large-scale application of a single fungal species in nature should be considered, and potential production of secondary metabolites by these fungi should not be neglected. Accordingly, common and potential biocontrol agents should be rigorously tested for potential ability as opportunistic human pathogens to make sure that these environmentally friendly applications are also safe for the people applying them.

References

Adav SS, Chao LT, Sze SK (2012) Quantitative secretomic analysis of *Trichoderma reesei* strains reveals enzymatic composition for lignocellulosic biomass degradation. Mol Cell Proteomics 11: M111 012419

Ahlawat OP, Gupta P, Kumar S, Sharma DK, Ahlawat K (2010) Bioremediation of fungicides by spent mushroom substrate and its associated microflora. Indian J Microbiol 50:390–395

Akel E, Metz B, Seiboth B, Kubicek CP (2009) Molecular regulation of arabinan and L-arabinose metabolism in *Hypocrea jecorina* (*Trichoderma reesei*). Eukaryot Cell 8:1837–1844

Aliferis KA, Jabaji S (2010) Metabolite composition and bioactivity of *Rhizoctonia solani* sclerotial exudates. J Agric Food Chem 58:7604–7615

Anholt RR (1994) Signal integration in the nervous system: adenylate cyclases as molecular coincidence detectors. Trends Neurosci 17:37–41

Argumedo-Delira R, Alarcon A, Ferrera-Cerrato R, Almaraz JJ, Pena-Cabriales JJ (2012) Tolerance and growth of 11 *Trichoderma* strains to crude oil, naphthalene, phenanthrene and benzo[a]pyrene. J Environ Manage 95(Suppl):S291–S299

Aro N, Pakula T, Penttila M (2005) Transcriptional regulation of plant cell wall degradation by filamentous fungi. FEMS Microbiol Rev 29:719–739

Arvas M, Pakula T, Lanthaler K, Saloheimo M, Valkonen M, Suortti T, Robson G, Penttila M (2006) Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. BMC Genomics 7:32

Arvas M, Haiminen N, Smit B, Rautio J, Vitikainen M, Wiebe M, Martinez D, Chee C, Kunkel J, Sanchez C, Nelson MA, Pakula T, Saloheimo M, Penttila M, Kivioja T (2010) Detecting novel genes with sparse arrays. Gene 467:41–51

Arvas M, Pakula T, Smit B, Rautio J, Koivistoinen H, Jouhten P, Lindfors E, Wiebe M, Penttila M, Saloheimo M (2011) Correlation of gene expression and protein production rate – a system wide study. BMC Genomics 12:616

Atanasova L, Le Crom S, Gruber S, Coulpier F, Seidl-Seiboth V, Kubicek CP, Druzhinina I (2013) Comparative transcriptomics reveals different strategies of *Trichoderma* mycoparasitism. BMC Genomics 14:121

Bae H, Sicher RC, Kim MS, Kim SH, Strem MD, Melnick RL, Bailey BA (2009) The beneficial endophyte *Trichoderma hamatum* isolate DIS 219b promotes growth and delays the onset of the drought response in *Theobroma cacao*. J Exp Bot 60:3279–3295

Bahn YS, Xue C, Idnurm A, Rutherford JC, Heitman J, Cardenas ME (2007) Sensing the environment: lessons from fungi. Nat Rev Microbiol 5:57–69

Baker SE, Perrone G, Richardson NM, Gallo A, Kubicek CP (2012) Phylogenomic analysis of polyketide synthase-encoding genes in *Trichoderma*. Microbiology 158:147–154

Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Walton JD (2010) Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. Biotechnol Biofuels 3:22

Bayram O, Braus GH (2012) Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. FEMS Microbiol Rev 36:1–24

Beeson WT, Phillips CM, Cate JH, Marletta MA (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J Am Chem Soc 134:890–892

Benitez T, Rincon AM, Limon MC, Codon AC (2004) Biocontrol mechanisms of *Trichoderma* strains. Int Microbiol 7:249–260

Bok JW, Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell 3:527–535

Brunner M, Kaldi K (2008) Interlocked feedback loops of the circadian clock of *Neurospora crassa*. Mol Microbiol 68:255–262

Brunner K, Omann M, Pucher ME, Delic M, Lehner SM, Domnanich P, Kratochwill K, Druzhinina I,

Denk D, Zeilinger S (2008) *Trichoderma* G protein-coupled receptors: functional characterisation of a cAMP receptor-like protein from *Trichoderma atroviride*. Curr Genet 54:283–299

Buee M, Reich M, Murat C, Morin E, Nilsson RH, Uroz S, Martin F (2009) 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytol 184:449–456

Cabrera-Vera TM, Vanhauwe J, Thomas TO, Medkova M, Preininger A, Mazzoni MR, Hamm HE (2003) Insights into G protein structure, function, and regulation. Endocr Rev 24:765–781

Carreras-Villasenor N, Sanchez-Arreguin JA, Herrera-Estrella A (2012) *Trichoderma* – sensing the environment for survival and dispersal. Microbiology 158(Pt 1):3–16

Carter GL, Allison D, Rey MW, Dunn-Coleman NS (1992) Chromosomal and genetic analysis of the electrophoretic karyotype of *Trichoderma reesei*: mapping of the cellulase and xylanase genes. Mol Microbiol 6:2167–2174

Casas-Flores S, Rios-Momberg M, Bibbins M, Ponce-Noyola P, Herrera-Estrella A (2004) BLR-1 and BLR-2, key regulatory elements of photoconidiation and mycelial growth in *Trichoderma atroviride*. Microbiology 150:3561–3569

Casas-Flores S, Rios-Momberg M, Rosales-Saavedra T, Martinez-Hernandez P, Olmedo-Monfil V, Herrera-Estrella A (2006) Cross talk between a fungal blue-light perception system and the cyclic AMP signaling pathway. Eukaryot Cell 5:499–506

Castellanos F, Schmoll M, Martinez P, Tisch D, Kubicek CP, Herrera-Estrella A, Esquivel-Naranjo EU (2010) Crucial factors of the light perception machinery and their impact on growth and cellulase gene transcription in *Trichoderma reesei*. Fungal Genet Biol 47:468–476

Catalano V, Vergara M, Hauzenberger JR, Seiboth B, Sarrocco S, Vannacci G, Kubicek CP, Seidl-Seiboth V (2011) Use of a non-homologous end-joining-deficient strain (delta-ku70) of the biocontrol fungus *Trichoderma virens* to investigate the function of the laccase gene *lcc1* in sclerotia degradation. Curr Genet 57:13–23

Chambergo FS, Bonaccorsi ED, Ferreira AJ, Ramos AS, Ferreira JRJJR, Abrahao-Neto J, Farah JP, El-Dorry H (2002) Elucidation of the metabolic fate of glucose in the filamentous fungus *Trichoderma reesei* using expressed sequence tag (EST) analysis and cDNA microarrays. J Biol Chem 277:13983–13988

Chen RE, Thorner J (2007) Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta 1773:1311–1340

Chen CH, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ (2009) Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in *Neurospora*. EMBO J 28:1029–1042

Chen CL, Kuo HC, Tung SY, Hsu PW, Wang CL, Seibel C, Schmoll M, Chen RS, Wang TF (2012) Blue light acts as a double-edged sword in regulating sexual development of *Hypocrea jecorina* (*Trichoderma reesei*). PLoS One 7:e44969

Cherry JR, Fidantsef AL (2003) Directed evolution of industrial enzymes: an update. Curr Opin Biotechnol 14:438–443

Chovanec P, Hudecova D, Varecka L (2001) Vegetative growth, aging- and light-induced conidiation of *Trichoderma viride* cultivated on different carbon sources. Folia Microbiol (Praha) 46:417–422

Contreras-Cornejo HA, Macias-Rodriguez L, Cortes-Penagos C, Lopez-Bucio J (2009) *Trichoderma virens*, a plant beneficial fungus, enhances biomass production and promotes lateral root growth through an auxin-dependent mechanism in *Arabidopsis*. Plant Physiol 149:1579–1592

D'Souza CA, Heitman J (2001) Conserved cAMP signaling cascades regulate fungal development and virulence. FEMS Microbiol Rev 25:349–364

Dagenais TR, Keller NP (2009) Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis. Clin Microbiol Rev 22:447–465

Dashtban M, Buchkowski R, Qin W (2011) Effect of different carbon sources on cellulase production by *Hypocrea jecorina* (*Trichoderma reesei*) strains. Int J Biochem Mol Biol 2:274–286

De Jaeger N, Declerck S, de la Providencia IE (2010) Mycoparasitism of arbuscular mycorrhizal fungi: a pathway for the entry of saprotrophic fungi into roots. FEMS Microbiol Ecol 73:312–322

Degenkolb T, Karimi Aghcheh R, Dieckmann R, Neuhof T, Baker SE, Druzhinina IS, Kubicek CP, Bruckner H, von Dohren H (2012) The production of multiple small peptaibol families by single 14-module peptide synthetases in *Trichoderma/Hypocrea*. Chem Biodivers 9:499–535

Delgado-Jarana J, Sousa S, Gonzalez F, Rey M, Llobell A (2006) ThHog1 controls the hyperosmotic stress response in *Trichoderma harzianum*. Microbiology 152:1687–1700

Dijksterhuis J, Veenhuis M, Harder W, Nordbring-Hertz B (1994) Nematophagous fungi: physiological aspects and structure-function relationships. Adv Microb Physiol 36:111–143

Djonovic S, Pozo MJ, Kenerley CM (2006a) Tvbgn3, a beta-1,6-glucanase from the biocontrol fungus *Trichoderma virens*, is involved in mycoparasitism and control of *Pythium ultimum*. Appl Environ Microbiol 72:7661–7670

Djonovic S, Pozo MJ, Dangott LJ, Howell CR, Kenerley CM (2006b) Sm1, a proteinaceous elicitor secreted by the biocontrol fungus *Trichoderma virens* induces plant defense responses and systemic resistance. Mol Plant Microbe Interact 19:838–853

Do Vale LH, Gomez-Mendoza DP, Kim MS, Pandey A, Ricart CA, Ximenes FFE, Sousa MV (2012) Secretome analysis of the fungus *Trichoderma harzianum* grown on cellulose. Proteomics 12:2716–2728

Druzhinina IS, Komon-Zelazowska M, Atanasova L, Seidl V, Kubicek CP (2010a) Evolution and ecophysiology of the industrial producer *Hypocrea*

jecorina (anamorph *Trichoderma reesei*) and a new sympatric agamospecies related to it. PLoS One 5:e9191

Druzhinina IS, Kubicek CP, Komon-Zelazowska M, Mulaw TB, Bissett J (2010b) The *Trichoderma harzianum* demon: complex speciation history resulting in coexistence of hypothetical biological species, recent agamospecies and numerous relict lineages. BMC Evol Biol 10:94

Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV, Kubicek CP (2011) *Trichoderma*: the genomics of opportunistic success. Nat Rev Microbiol 9:749–759

Druzhinina IS, Komon-Zelazowska M, Ismaiel A, Jaklitsch W, Mullaw T, Samuels GJ, Kubicek CP (2012) Molecular phylogeny and species delimitation in the section Longibrachiatum of *Trichoderma*. Fungal Genet Biol 49:358–368

Duclohier H (2007) Peptaibiotics and peptaibols: an alternative to classical antibiotics? Chem Biodivers 4:1023–1026

Eijsink VG, Vaaje-Kolstad G, Varum KM, Horn SJ (2008) Towards new enzymes for biofuels: lessons from chitinase research. Trends Biotechnol 26:228–235

Esquivel-Naranjo EU, Herrera-Estrella A (2007) Enhanced responsiveness and sensitivity to blue light by *blr-2* overexpression in *Trichoderma atroviride*. Microbiology 153:3909–3922

Farkas V, Gresik M, Kolarova N, Sulova Z, Sestak S (1990) Biochemical and physiological changes during photo-induced conidiation and derepression of cellulase synthesis in *Trichoderma*. In: Kubicek CP, Eveleigh DE, Esterbauer W, Steiner W, Kubicek-Pranz EM (eds) *Trichoderma reesei* cellulase: biochemistry, genetics, physiology, and application. Graham House, Cambridge, pp 139–155

Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemieux S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, White OR, Whitty BR, Youngman P, Wolfe KH, Goldman GH, Wortman JR, Jiang B, Denning DW, Nierman WC (2008) Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet 4: e1000046

Flores A, Chet I, Herrera-Estrella A (1997) Improved biocontrol activity of *Trichoderma harzianum* by over-expression of the proteinase-encoding gene *prb1*. Curr Genet 31:30–37

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJ, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem 278:31988–31997

Friedl MA, Druzhinina IS (2012) Taxon-specific metagenomics of *Trichoderma* reveals a narrow community of opportunistic species that regulate each other's development. Microbiology 158:69–83

Friedl MA, Schmoll M, Kubicek CP, Druzhinina IS (2008) Photostimulation of *Hypocrea atroviridis* growth occurs due to a cross-talk of carbon metabolism, blue light receptors and response to oxidative stress. Microbiology 154:1229–1241

Galagan JE, Selker EU (2004) RIP: the evolutionary cost of genome defense. Trends Genet 20:417–423

Gams W, Bissett J (1998) Morphology and identification of *Trichoderma*. In: Harman GE, Kubicek CP (eds) *Trichoderma* and Gliocladium, vol 1. Taylor and Francis/CRC Press, London, pp 3–34

Gardiner DM, Howlett BJ (2005) Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of *Aspergillus fumigatus*. FEMS Microbiol Lett 248:241–248

Glick BR (2005) Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. FEMS Microbiol Lett 251:1–7

Greenwald CJ, Kasuga T, Glass NL, Shaw BD, Ebbole DJ, Wilkinson HH (2010) Temporal and spatial regulation of gene expression during asexual development of *Neurospora crassa*. Genetics 186:1217–1230

Gruber S, Seidl-Seiboth V (2012) Self versus non-self: fungal cell wall degradation in *Trichoderma*. Microbiology 158:26–34

Gruber S, Kubicek CP, Seidl-Seiboth V (2011) Differential regulation of orthologous chitinase genes in mycoparasitic *Trichoderma* species. Appl Environ Microbiol 77:7217–7226

Guangtao Z, Hartl L, Schuster A, Polak S, Schmoll M, Wang T, Seidl V, Seiboth B (2009) Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. J Biotechnol 139:146–151

Gyalai-Korpos M, Nagy G, Mareczky Z, Schuster A, Reczey K, Schmoll M (2010) Relevance of the light signaling machinery for cellulase expression in *Trichoderma reesei* (*Hypocrea jecorina*). BMC Res Notes 3:330

Häkkinen M, Arvas M, Oja M, Aro N, Penttila M, Saloheimo M, Pakula TM (2012) Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. Microb Cell Fact 11:134

Hanada RE, de Jorge ST, Pomella AW, Hebbar KP, Pereira JO, Ismaiel A, Samuels GJ (2008) *Trichoderma martiale* sp. nov., a new endophyte from sapwood of *Theobroma cacao* with a potential for biological control. Mycol Res 112:1335–1343

Hanada RE, Pomella AW, Costa HS, Bezerra JL, Loguercio LL, Pereira JO (2010) Endophytic fungal diversity in *Theobroma cacao* (cacao) and

T. grandiflorum (cupuacu) trees and their potential for growth promotion and biocontrol of black-pod disease. Fungal Biol 114:901–910

Harman GE, Howell CR, Viterbo A, Chet I, Lorito M (2004) *Trichoderma* species – opportunistic, avirulent plant symbionts. Nat Rev Microbiol 2:43–56

Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry 49:3305–3316

Hartl L, Seiboth B (2005) Sequential gene deletions in *Hypocrea jecorina* using a single blaster cassette. Curr Genet 48:204–211

Hartl L, Zach S, Seidl-Seiboth V (2012) Fungal chitinases: diversity, mechanistic properties and biotechnological potential. Appl Microbiol Biotechnol 93:533–543

Hatvani L, Manczinger L, Vagvölgyi C, Kredics L (2013) *Trichoderma* as a human pathogen. In: Mukherjee PK, Horwitz BA, Singh US, Mukherjee M, Schmoll M (eds) *Trichoderma*: biology and applications. CAB International, Wallingford, pp 292–313

Hegedus N, Sigl C, Zadra I, Pocsi I, Marx F (2011) The *paf* gene product modulates asexual development in *Penicillium chrysogenum*. J Basic Microbiol 51:253–262

Hermosa R, Viterbo A, Chet I, Monte E (2012) Plant-beneficial effects of *Trichoderma* and of its genes. Microbiology 158:17–25

Hernandez-Onate MA, Esquivel-Naranjo EU, Mendoza-Mendoza A, Stewart A, Herrera-Estrella AH (2012) An injury-response mechanism conserved across kingdoms determines entry of the fungus *Trichoderma atroviride* into development. Proc Natl Acad Sci U S A 109:14918–14923

Herold S, Bischof R, Metz B, Seiboth B, Kubicek CP (2013) Prepared for the possible: xylanase gene transcription in *Trichoderma reesei* is triggered by different inducers representing different hemicellulosic pentose polymers. Eukaryot Cell 12:390–398

Herpoel-Gimbert I, Margeot A, Dolla A, Jan G, Molle D, Lignon S, Mathis H, Sigoillot JC, Monot F, Asther M (2008) Comparative secretome analyses of two *Trichoderma reesei* RUT-C30 and CL847 hypersecretory strains. Biotechnol Biofuels 1:18

Herrera-Estrella A, Horwitz BA (2007) Looking through the eyes of fungi: molecular genetics of photoreception. Mol Microbiol 64:5–15

Herrera-Estrella A, Goldman GH, van Montagu M, Geremia RA (1993) Electrophoretic karyotype and gene assignment to resolved chromosomes of *Trichoderma* spp. Mol Microbiol 7:515–521

Horn SJ, Sorbotten A, Synstad B, Sikorski P, Sorlie M, Varum KM, Eijsink VG (2006) Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. FEBS J 273:491–503

Huang G, Chen S, Li S, Cha J, Long C, Li L, He Q, Liu Y (2007) Protein kinase A and casein kinases mediate sequential phosphorylation events in the circadian negative feedback loop. Genes Dev 21:3283–3295

Ihrmark K, Asmail N, Ubhayasekera W, Melin P, Stenlid J, Karlsson M (2010) Comparative molecular evolution of *Trichoderma* chitinases in response to mycoparasitic interactions. Evol Bioinform Online 6:1–26

Jaklitsch WM (2009) European species of *Hypocrea* part I. The green-spored species. Stud Mycol 63:1–91

Jaklitsch WM (2011) European species of *Hypocrea* part II: species with hyaline ascospores. Fungal Divers 48:1–250

Joshi PK, Swarup A, Maheshwari S, Kumar R, Singh N (2011) Bioremediation of heavy metals in liquid media through fungi isolated from contaminated sources. Indian J Microbiol 51:482–487

Jun H, Kieselbach T, Jonsson LJ (2011) Enzyme production by filamentous fungi: analysis of the secretome of *Trichoderma reesei* grown on unconventional carbon source. Microb Cell Fact 10:68

Karimi-Aghcheh R, Bok JW, Phatale PA, Smith KM, Baker SE, Lichius A, Omann M, Zeilinger S, Seiboth B, Rhee C, Keller NP, Freitag M, Kubicek CP (2013) Functional analyses of *Trichoderma reesei* LAE1 reveal conserved and contrasting roles of this regulator. G3 (Bethesda) 3:369–378

Karlsson M, Stenlid J (2008) Comparative evolutionary histories of the fungal chitinase gene family reveal non-random size expansions and contractions due to adaptive natural selection. Evol Bioinform Online 4:47–60

Karlsson M, Stenlid J (2009) Evolution of family 18 glycoside hydrolases: diversity, domain structures and phylogenetic relationships. J Mol Microbiol Biotechnol 16:208–223

Kataja K, Satokari RM, Arvas M, Takkinen K, Soderlund H (2006) A highly sensitive and multiplexed method for focused transcript analysis. J Microbiol Methods 67:102–113

Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. Fungal Genet Biol 21:17–29

Komon-Zelazowska M, Neuhof T, Dieckmann R, von Dohren H, Herrera-Estrella A, Kubicek CP, Druzhinina IS (2007) Formation of atroviridin by *Hypocrea atroviridis* is conidiation associated and positively regulated by blue light and the G protein GNA3. Eukaryot Cell 6:2332–2342

Krappmann S (2007) Gene targeting in filamentous fungi: the benefits of impaired repair. Fungal Biol Rev 21:25–29

Kredics L, Antal Z, Doczi I, Manczinger L, Kevei F, Nagy E (2003) Clinical importance of the genus

Trichoderma. A review. Acta Microbiol Immunol Hung 50:105–117

Kubicek CP (2013) Systems biological approaches towards understanding cellulase production by *Trichoderma reesei*. J Biotechnol 163:133–142

Kubicek CP, Mühlbauer G, Grotz M, John E, Kubicek-Pranz EM (1988) Properties of the conidial-bound cellulase system of *Trichoderma reesei*. J Gen Microbiol 134:1215–1222

Kubicek CP, Mikus M, Schuster A, Schmoll M, Seiboth B (2009) Metabolic engineering strategies for improvement of cellulase production by *Hypocrea jecorina*. Biotechnol Biofuels 2:19

Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, Zeilinger S, Casas-Flores S, Horwitz BA, Mukherjee PK, Mukherjee M, Kredics L, Alcaraz LD, Aerts A, Antal Z, Atanasova L, Cervantes-Badillo MG, Challacombe J, Chertkov O, McCluskey K, Coulpier F, Deshpande N, von Dohren H, Ebbole DJ, Esquivel-Naranjo EU, Fekete E, Flipphi M, Glaser F, Gomez-Rodriguez EY, Gruber S, Han C, Henrissat B, Hermosa R, Hernandez-Onate M, Karaffa L, Kosti I, Le Crom S, Lindquist E, Lucas S, Lubeck M, Lubeck PS, Margeot A, Metz B, Misra M, Nevalainen H, Omann M, Packer N, Perrone G, Uresti-Rivera EE, Salamov A, Schmoll M, Seiboth B, Shapiro H, Sukno S, Tamayo-Ramos JA, Tisch D, Wiest A, Wilkinson HH, Zhang M, Coutinho PM, Kenerley CM, Monte E, Baker SE, Grigoriev IV (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. Genome Biol 12:R40

Kumar A, Scher K, Mukherjee M, Pardovitz-Kedmi E, Sible GV, Singh US, Kale SP, Mukherjee PK, Horwitz BA (2010) Overlapping and distinct functions of two *Trichoderma virens* MAP kinases in cell-wall integrity, antagonistic properties and repression of conidiation. Biochem Biophys Res Commun 398:765–770

Lamarre C, Sokol S, Debeaupuis JP, Henry C, Lacroix C, Glaser P, Coppee JY, Francois JM, Latge JP (2008) Transcriptomic analysis of the exit from dormancy of *Aspergillus fumigatus* conidia. BMC Genomics 9:417

Latge JP (2007) The cell wall: a carbohydrate armour for the fungal cell. Mol Microbiol 66:279–290

Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106:16151–16156

Lehner SM, Atanasova L, Neumann NK, Krska R, Lemmens M, Druzhinina IS, Schuhmacher R (2013) Isotope-assisted screening for iron-containing metabolites reveals a high degree of diversity among known and unknown siderophores produced by *Trichoderma* spp. Appl Environ Microbiol 79:18–31

Leitgeb B, Szekeres A, Manczinger L, Vagvolgyi C, Kredics L (2007) The history of alamethicin: a review of the most extensively studied peptaibol. Chem Biodivers 4:1027–1051

Lewis JA, Papavizas GC (1984) A new approach to stimulate population proliferation of *Trichoderma* species and other potential biocontrol fungi introduced into natural soils. Phytopathology 74:1240–1244

Lim YW, Kim BK, Kim C, Jung HS, Kim BS, Lee JH, Chun J (2010) Assessment of soil fungal communities using pyrosequencing. J Microbiol 48:284–289

Limon MC, Pintor-Toro JA, Benitez T (1999) Increased antifungal activity of *Trichoderma harzianum* transformants that overexpress a 33-kDa chitinase. Phytopathology 89:254–261

Lin K, Hwang PK, Fletterick RJ (1995) Mechanism of regulation in yeast glycogen phosphorylase. J Biol Chem 270:26833–26839

Lorito M, Farkas V, Rebuffat S, Bodo B, Kubicek CP (1996) Cell wall synthesis is a major target of mycoparasitic antagonism by *Trichoderma harzianum*. J Bacteriol 178:6382–6385

Lorito M, Woo SL, Harman GE, Monte E (2010) Translational research on *Trichoderma*: from 'omics to the field. Annu Rev Phytopathol 48:395–417

Mäntylä AL, Rossi KH, Vanhanen SA, Penttilä ME, Suominen PL, Nevalainen KM (1992) Electrophoretic karyotyping of wild-type and mutant *Trichoderma longibrachiatum* (*reesei*) strains. Curr Genet 21:471–477

Margolles-Clark E, Ilmen M, Penttilä M (1997) Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. J Biotechnol 57:167–179

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barbote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26:553–560

Matarese F, Sarrocco S, Gruber S, Seidl-Seiboth V, Vannacci G (2012) Biocontrol of *Fusarium* head blight: interactions between *Trichoderma* and mycotoxigenic *Fusarium*. Microbiology 158:98–106

Mendoza-Mendoza A, Pozo MJ, Grzegorski D, Martinez P, Garcia JM, Olmedo-Monfil V, Cortes C, Kenerley C, Herrera-Estrella A (2003) Enhanced biocontrol activity of *Trichoderma* through inactivation of a mitogen-activated protein kinase. Proc Natl Acad Sci U S A 100:15965–15970

Mendoza-Mendoza A, Rosales-Saavedra T, Cortes C, Castellanos-Juarez V, Martinez P, Herrera-Estrella A (2007) The MAP kinase TVK1 regulates conidiation, hydrophobicity and the expression of genes encoding cell wall proteins in the fungus *Trichoderma virens*. Microbiology 153:2137–2147

Metz B, Seidl-Seiboth V, Haarmann T, Kopchinskiy A, Lorenz P, Seiboth B, Kubicek CP (2011) Expression of biomass-degrading enzymes is a major event during conidium development in *Trichoderma reesei*. Eukaryot Cell 10:1527–1535

Montero-Barrientos M, Hermosa R, Cardoza RE, Gutierrez S, Monte E (2011) Functional analysis of the *Trichoderma harzianum nox1* gene, encoding an NADPH oxidase, relates production of reactive oxygen species to specific biocontrol activity against *Pythium ultimum*. Appl Environ Microbiol 77:3009–3016

Moy M, Li HM, Sullivan R, White JF Jr, Belanger FC (2002) Endophytic fungal beta-1,6-glucanase expression in the infected host grass. Plant Physiol 130:1298–1308

Mukherjee PK, Kenerley CM (2010) Regulation of morphogenesis and biocontrol properties in *Trichoderma virens* by a VELVET protein, Vel1. Appl Environ Microbiol 76:2345–2352

Mukherjee PK, Latha J, Hadar R, Horwitz BA (2003) TmkA, a mitogen-activated protein kinase of *Trichoderma virens*, is involved in biocontrol properties and repression of conidiation in the dark. Eukaryot Cell 2:446–455

Mukherjee PK, Latha J, Hadar R, Horwitz BA (2004) Role of two G-protein alpha subunits, TgaA and TgaB, in the antagonism of plant pathogens by *Trichoderma virens*. Appl Environ Microbiol 70:542–549

Mukherjee M, Horwitz BA, Sherkhane PD, Hadar R, Mukherjee PK (2006) A secondary metabolite biosynthesis cluster in *Trichoderma virens*: evidence from analysis of genes underexpressed in a mutant defective in morphogenesis and antibiotic production. Curr Genet 50:193–202

Mukherjee M, Mukherjee PK, Kale SP (2007) cAMP signalling is involved in growth, germination, mycoparasitism and secondary metabolism in *Trichoderma virens*. Microbiology 153:1734–1742

Mukherjee PK, Wiest A, Ruiz N, Keightley A, Moran-Diez ME, McCluskey K, Pouchus YF, Kenerley CM (2011) Two classes of new peptaibols are synthesized by a single non-ribosomal peptide synthetase of *Trichoderma virens*. J Biol Chem 286:4544–4554

Mukherjee PK, Horwitz BA, Kenerley CM (2012a) Secondary metabolism in *Trichoderma*—a genomic perspective. Microbiology 158:35–45

Mukherjee PK, Buensanteai N, Moran-Diez ME, Druzhinina IS, Kenerley CM (2012b) Functional analysis of non-ribosomal peptide synthetases (NRPSs) in *Trichoderma virens* reveals a polyketide synthase (PKS)/NRPS hybrid enzyme involved in the induced systemic resistance response in maize. Microbiology 158:155–165

Mukherjee PK, Horwitz BA, Herrera-Estrella A, Schmoll M, Kenerley C (2013) *Trichoderma* research in the genome era. Annu Rev Phytopathol 51:105–129

Nadeau JH, Taylor BA (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. Proc Natl Acad Sci U S A 81:814–818

Nielsen KF, Grafenhan T, Zafari D, Thrane U (2005) Trichothecene production by *Trichoderma brevicompactum*. J Agric Food Chem 53:8190–8196

Omann M, Zeilinger S (2010) How a mycoparasite employs g-protein signaling: using the example of *Trichoderma*. J Signal Transduct 2010:123126

Pakula T, Saloheimo M, Häkkinen M, Westernholm-Parvinen A, Penttilä ME (2011a) Improved production of proteins in filamentous fungi. International Patent WO2011/151515 A1

Pakula T, Saloheimo M, Häkkinen M, Westernholm-Parvinen A, Penttilä ME, Vitikainen M (2011b) Method for protein production in filamentous fungi. International Patent WO2011/151513A1

Pakula T, Saloheimo M, Häkkinen M, Westernholm-Parvinen A, Penttilä ME, Vitikainen M (2011c) Method for protein production in filamentous fungi. International Patent WO2011/151512 A1

Paoletti M, Saupe SJ (2009) Fungal incompatibility: evolutionary origin in pathogen defense? Bioessays 31:1201–1210

Papapostolou I, Georgiou CD (2010) Hydrogen peroxide is involved in the sclerotial differentiation of filamentous phytopathogenic fungi. J Appl Microbiol 109:1929–1936

Park G, Colot HV, Collopy PD, Krystofova S, Crew C, Ringelberg C, Litvinkova L, Altamirano L, Li L, Curilla S, Wang W, Gorrochotegui-Escalante N, Dunlap JC, Borkovich KA (2011) High-throughput production of gene replacement mutants in *Neurospora crassa*. Methods Mol Biol 722:179–189

Patron NJ, Waller RF, Cozijnsen AJ, Straney DC, Gardiner DM, Nierman WC, Howlett BJ (2007) Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol Biol 7:174

Persoon CH (1794) Disposita methodica fungorum. Römer's Neues Mag Bot 1:81–128

Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, Hartl L, Karaffa L, Druzhinina IS, Sei-

both B, Le Crom S, Kubicek CP (2011) The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. BMC Genomics 12:269

Rautio JJ, Kataja K, Satokari R, Penttila M, Soderlund H, Saloheimo M (2006) Rapid and multiplexed transcript analysis of microbial cultures using capillary electophoresis-detectable oligonucleotide probe pools. J Microbiol Methods 65:404–416

Rautio JJ, Bailey M, Kivioja T, Soderlund H, Penttila M, Saloheimo M (2007) Physiological evaluation of the filamentous fungus *Trichoderma reesei* in production processes by marker gene expression analysis. BMC Biotechnol 7:28

Reithner B, Brunner K, Schuhmacher R, Peissl I, Seidl V, Krska R, Zeilinger S (2005) The G protein alpha subunit Tga1 of *Trichoderma atroviride* is involved in chitinase formation and differential production of antifungal metabolites. Fungal Genet Biol 42:749–760

Reithner B, Schuhmacher R, Stoppacher N, Pucher M, Brunner K, Zeilinger S (2007) Signaling via the *Trichoderma atroviride* mitogen-activated protein kinase Tmk1 differentially affects mycoparasitism and plant protection. Fungal Genet Biol 44:1123–1133

Reithner B, Ibarra-Laclette E, Mach RL, Herrera-Estrella A (2011) Identification of mycoparasitism-related genes in *Trichoderma atroviride*. Appl Environ Microbiol 77:4361–4370

Reyes-Dominguez Y, Bok JW, Berger H, Shwab EK, Basheer A, Gallmetzer A, Scazzocchio C, Keller N, Strauss J (2010) Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. Mol Microbiol 76:1376–1386

Rocha-Ramirez V, Omero C, Chet I, Horwitz BA, Herrera-Estrella A (2002) *Trichoderma atroviride* G-protein alpha-subunit gene *tga1* is involved in mycoparasitic coiling and conidiation. Eukaryot Cell 1:594–605

Rosales-Saavedra T, Esquivel-Naranjo EU, Casas-Flores S, Martinez-Hernandez P, Ibarra-Laclette E, Cortes-Penagos C, Herrera-Estrella A (2006) Novel light-regulated genes in *Trichoderma atroviride*: a dissection by cDNA microarrays. Microbiology 152:3305–3317

Ruocco M, Lanzuise S, Vinale F, Marra R, Turra D, Woo SL, Lorito M (2009) Identification of a new biocontrol gene in *Trichoderma atroviride*: the role of an ABC transporter membrane pump in the interaction with different plant-pathogenic fungi. Mol Plant Microbe Interact 22:291–301

Samuels GJ (1996) *Trichoderma*: a review of biology and systematics of the genus. Mycol Res 100:923–935

Samuels GJ, Ismaiel A, Bon MC, De Respinis S, Petrini O (2010) *Trichoderma asperellum* sensu lato consists of two cryptic species. Mycologia 102:944–966

Sanchez-Arreguin A, Perez-Martinez AS, Herrera-Estrella A (2012) Proteomic analysis of *Trichoderma atroviride* reveals independent roles for transcription factors BLR-1 and BLR-2 in light and darkness. Eukaryot Cell 11:30–41

Schirmböck M, Lorito M, Wang YL, Hayes CK, Arisan-Atac I, Scala F, Harman GE, Kubicek CP (1994) Parallel formation and synergism of hydrolytic enzymes and peptaibol antibiotics, molecular mechanisms involved in the antagonistic action of *Trichoderma harzianum* against phytopathogenic fungi. Appl Environ Microbiol 60:4364–4370

Schmoll M (2008) The information highways of a biotechnological workhorse – signal transduction in *Hypocrea jecorina*. BMC Genomics 9:430

Schmoll M (2011) Assessing the relevance of light for fungi implications and insights into the network of signal transmission. Adv Appl Microbiol 76:27–78

Schmoll M (2013) Sexual development in *Trichoderma* – scrutinizing the aspired phenomenon. In: Mukherjee PK, Horwitz BA, Singh US, Mukherjee M, Schmoll M (eds) *Trichoderma*: biology and applications. CAB International, Wallingford, pp 67–86

Schmoll M, Kubicek CP (2005) *ooc1*, a unique gene expressed only during growth of *Hypocrea jecorina* (anamorph: *Trichoderma reesei*) on cellulose. Curr Genet 48:126–133

Schmoll M, Zeilinger S, Mach RL, Kubicek CP (2004) Cloning of genes expressed early during cellulase induction in *Hypocrea jecorina* by a rapid subtraction hybridization approach. Fungal Genet Biol 41:877–887

Schmoll M, Franchi L, Kubicek CP (2005) Envoy, a PAS/LOV domain protein of *Hypocrea jecorina* (Anamorph *Trichoderma reesei*), modulates cellulase gene transcription in response to light. Eukaryot Cell 4:1998–2007

Schmoll M, Schuster A, do Nascimento Silva R, Kubicek CP (2009) The G-alpha protein GNA3 of *Hypocrea jecorina* (anamorph *Trichoderma reesei*) regulates cellulase gene expression in the presence of light. Eukaryot Cell 8:410–420

Schmoll M, Esquivel-Naranjo EU, Herrera-Estrella A (2010a) *Trichoderma* in the light of day – physiology and development. Fungal Genet Biol 47:909–916

Schmoll M, Seibel C, Tisch D, Dorrer M, Kubicek CP (2010b) A novel class of peptide pheromone precursors in ascomycetous fungi. Mol Microbiol 77:1483–1501

Schmoll M, Tian C, Sun J, Tisch D, Glass NL (2012) Unravelling the molecular basis for light modulated cellulase gene expression – the role of photoreceptors in *Neurospora crassa*. BMC Genomics 13:127

Schuster A, Kubicek CP, Friedl MA, Druzhinina IS, Schmoll M (2007) Impact of light on *Hypocrea jecorina* and the multiple cellular roles of ENVOY in this process. BMC Genomics 8:449

Schuster A, Kubicek CP, Schmoll M (2011) Dehydrogenase GRD1 represents a novel component of the cellulase regulon in *Trichoderma reesei* (*Hypocrea jecorina*). Appl Environ Microbiol 77:4553–4563

Schuster A, Tisch D, Seidl-Seiboth V, Kubicek CP, Schmoll M (2012a) Roles of protein kinase A and adenylate cyclase in light-modulated cellulase regulation in *Trichoderma reesei*. Appl Environ Microbiol 78:2168–2178

Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP, Schmoll M (2012b) A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. Biotechnol Biofuels 5:1

Seibel C, Gremel G, Silva RD, Schuster A, Kubicek CP, Schmoll M (2009) Light-dependent roles of the G-protein subunit GNA1 of *Hypocrea jecorina* (anamorph *Trichoderma reesei*). BMC Biol 7:58

Seibel C, Tisch D, Kubicek CP, Schmoll M (2012a) The role of pheromone receptors for communication and mating in *Hypocrea jecorina* (*Trichoderma reesei*). Fungal Genet Biol 49:814–824

Seibel C, Tisch D, Kubicek CP, Schmoll M (2012b) ENVOY is a major determinant in regulation of sexual development in *Hypocrea jecorina* (*Trichoderma reesei*). Eukaryot Cell 11:885–895

Seiboth B, Pakdaman SB, Hartl L, Kubicek CP (2007) Lactose metabolism in filamentous fungi: how to deal with an unknown substrate. Fungal Biol Rev 21:42–48

Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP (2012) The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol 84:1150–1164

Seidl V, Seiboth B (2010) *Trichoderma reesei*: genetic approaches to improving strain efficiency. Biofuels 1:343–354

Seidl V, Huemer B, Seiboth B, Kubicek CP (2005) A complete survey of *Trichoderma chitinases* reveals three distinct subgroups of family 18 chitinases. FEBS J 272:5923–5939

Seidl V, Marchetti M, Schandl R, Allmaier G, Kubicek CP (2006a) Epl1, the major secreted protein of *Hypocrea atroviridis* on glucose, is a member of a strongly conserved protein family comprising plant defense response elicitors. FEBS J 273:4346–4359

Seidl V, Schmoll M, Scherm B, Balmas V, Seiboth B, Migheli Q, Kubicek CP (2006b) Antagonism of *Pythium* blight of zucchini by *Hypocrea jecorina* does not require cellulase gene expression but is improved by carbon catabolite derepression. FEMS Microbiol Lett 257:145–151

Seidl V, Gamauf C, Druzhinina IS, Seiboth B, Hartl L, Kubicek CP (2008) The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85kb (29 gene-encoding) region of the wild-type genome. BMC Genomics 9:327

Seidl V, Seibel C, Kubicek CP, Schmoll M (2009a) Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci U S A 106:13909–13914

Seidl V, Song L, Lindquist E, Gruber S, Koptchinskiy A, Zeilinger S, Schmoll M, Martinez P, Sun J, Grigoriev I, Herrera-Estrella A, Baker SE, Kubicek CP (2009b) Transcriptomic response of the mycoparasitic fungus *Trichoderma atroviride* to the presence of a fungal prey. BMC Genomics 10:567

Sestak S, Farkas V (1993) Metabolic regulation of endoglucanase synthesis in *Trichoderma reesei:* participation of cyclic AMP and glucose-6-phosphate. Can J Microbiol 39:342–347

Shoresh M, Gal-On A, Leibman D, Chet I (2006) Characterization of a mitogen-activated protein kinase gene from cucumber required for *Trichoderma*-conferred plant resistance. Plant Physiol 142:1169–1179

Shoresh M, Harman GE, Mastouri F (2010) Induced systemic resistance and plant responses to fungal biocontrol agents. Annu Rev Phytopathol 48:21–43

Silar P (2005) Peroxide accumulation and cell death in filamentous fungi induced by contact with a contestant. Mycol Res 109:137–149

Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF, Thomas TL, Carrington JC, Stajich JE, Bell-Pedersen D, Brunner M, Freitag M (2010) Transcription factors in light and circadian clock signaling networks revealed by genome wide mapping of direct targets for *Neurospora* white collar complex. Eukaryot Cell 9:1549–1556

Spudich JL, Jung KH (2005) Microbial rhodopsins: phylogenetic and functional diversity. In: Briggs WR, Spudich JL (eds) Handbook of photosensory receptors, vol 1. Wiley, Weinheim, pp 1–23

Steiger MG, Vitikainen M, Uskonen P, Brunner K, Adam G, Pakula T, Penttila M, Saloheimo M, Mach RL, Mach-Aigner AR (2011) Transformation system for *Hypocrea jecorina* (*Trichoderma reesei*) that favors homologous integration and employs reusable bidirectionally selectable markers. Appl Environ Microbiol 77:114–121

Stoppacher N, Zeilinger S, Omann M, Lassahn PG, Roitinger A, Krska R, Schuhmacher R (2008) Characterisation of the peptaibiome of the biocontrol fungus *Trichoderma atroviride* by liquid chromatography/tandem mass spectrometry. Rapid Commun Mass Spectrom 22:1889–1898

Stricker AR, Grosstessner-Hain K, Wurleitner E, Mach RL (2006) Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. Eukaryot Cell 5:2128–2137

Sunnarborg SW, Miller SP, Unnikrishnan I, LaPorte DC (2001) Expression of the yeast glycogen phosphorylase gene is regulated by stress-response elements

and by the HOG MAP kinase pathway. Yeast 18:1505–1514

Tejesvi MV, Mahesh B, Nalini MS, Prakash HS, Kini KR, Subbiah V, Shetty HS (2006) Fungal endophyte assemblages from ethnopharmaceutically important medicinal trees. Can J Microbiol 52:427–435

Tisch D, Schmoll M (2010) Light regulation of metabolic pathways in fungi. Appl Microbiol Biotechnol 85:1259–1277

Tisch D, Kubicek CP, Schmoll M (2011a) New insights into the mechanism of light modulated signaling by heterotrimeric G-proteins: ENVOY acts on *gna1* and *gna3* and adjusts cAMP levels in *Trichoderma reesei* (*Hypocrea jecorina*). Fungal Genet Biol 48:631–640

Tisch D, Kubicek CP, Schmoll M (2011b) The phosducin-like protein PhLP1 impacts regulation of glycoside hydrolases and light response in *Trichoderma reesei*. BMC Genomics 12:613

Tulasne LR, Tulasne C (1865) Selecta fungorum carpologia. Jussu, Paris

Ueno Y (1985) The toxicology of mycotoxins. Crit Rev Toxicol 14:99–132

Vinale F, Sivasithamparam K, Ghisalberti EL, Ruocco M, Wood S, Lorito M (2012) *Trichoderma* secondary metabolites that affect plant metabolism. Nat Prod Commun 7:1545–1550

Viterbo A, Ramot O, Chemin L, Chet I (2002) Significance of lytic enzymes from *Trichoderma* spp. in the biocontrol of fungal plant pathogens. Antonie Van Leeuwenhoek 81:549–556

Viterbo A, Harel M, Horwitz BA, Chet I, Mukherjee PK (2005) *Trichoderma* mitogen-activated protein kinase signaling is involved in induction of plant systemic resistance. Appl Environ Microbiol 71:6241–6246

Viterbo A, Wiest A, Brotman Y, Chet I, Kenerley C (2007) The 18mer peptaibols from *Trichoderma virens* elicit plant defence responses. Mol Plant Pathol 8:737–746

Viterbo A, Landau U, Kim S, Chernin L, Chet I (2010) Characterization of ACC deaminase from the biocontrol and plant growth-promoting agent *Trichoderma asperellum* T203. FEMS Microbiol Lett 305:42–48

Vitikainen M, Arvas M, Pakula T, Oja M, Penttila M, Saloheimo M (2010) Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. BMC Genomics 11:441

Vizcaino JA, Gonzalez FJ, Suarez MB, Redondo J, Heinrich J, Delgado-Jarana J, Hermosa R, Gutierrez S, Monte E, Llobell A, Rey M (2006) Generation, annotation and analysis of ESTs from *Trichoderma harzianum* CECT 2413. BMC Genomics 7:193

Walker T, Klasson L, Sebaihia M, Sanders MJ, Thomson NR, Parkhill J, Sinkins SP (2007) Ankyrin repeat domain-encoding genes in the wPip strain of Wolbachia from the Culex pipiens group. BMC Biol 5:39

Wang KL, Li H, Ecker JR (2002) Ethylene biosynthesis and signaling networks. Plant Cell 14(Suppl): S131–S151

Watts R, Dahiya J, Chaudhary K, Tauro P (1988) Isolation and characterization of a new antifungal metabolite of *Trichoderma reesei*. Plant and Soil 107:81–84

Zeilinger S, Schmoll M, Pail M, Mach RL, Kubicek CP (2003) Nucleosome transactions on the *Hypocrea jecorina* (*Trichoderma reesei*) cellulase promoter *cbh2* associated with cellulase induction. Mol Genet Genomics 270:46–55

Zeilinger S, Reithner B, Scala V, Peissl I, Lorito M, Mach RL (2005) Signal transduction by Tga3, a novel G protein alpha subunit of *Trichoderma atroviride*. Appl Environ Microbiol 71:1591–1597

Znameroski EA, Glass NL (2013) Using a model filamentous fungus to unravel mechanisms of lignocellulose deconstruction. Biotechnol Biofuels 6:6

Pathogenicity

11 Application of Genomics to the Study of Pathogenicity and Development in *Fusarium*

Frances Trail[1], Donald M. Gardiner[2]

CONTENTS

I. Introduction

The year 2013 marked the tenth anniversary of the public release of the genome sequence of *Fusarium graminearum* on the Broad Institute website (http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiHome.html; Cuomo et al. 2007). The *Fusarium graminearum* sequence was the first *Fusarium* genome sequence to be released and the third publically available genome from a filamentous fungus, following the release of *Neurospora crassa* and *Magnaporthe oryzae* (formerly *M. grisea*) genomes (Dean et al. 2005; Galagan et al. 2003). The *F. graminearum* genome sequence was followed in 2006 with the release of the *F. verticillioides* genome sequence (Ma et al. 2010) and in 2007 with that of *F. oxysporum* f. sp. *lycopersici* (Ma et al. 2010) and *F. solani* f. sp. *pisi* (formerly known as *Nectria haematococca* MP VI) in 2009 (Coleman et al. 2009). Each of these genome sequences was generated with traditional Sanger sequencing and used optical or genetic maps to scaffold the assemblies into nearly complete chromosomes. Next-generation technologies are now generating genomes at a fraction of the cost and in much shorter periods of time than these first releases. Genome sequences for *F. pseudograminearum* and *F. oxysporum* strain 5176 (isolated from *Brassica oleracea*) based entirely on next-generation technologies have been published (Gardiner et al. 2012; Thatcher et al. 2012). These technologies have also been used to fill gaps in the Sanger-derived sequences (Cavinder and Trail 2012). However, the value of these four founding genome sequences cannot be overstated, and these genomes will continue to drive our understanding of the biology of this genus.

The *F. graminearum* genome sequence, as one of the first fungal genomes, stimulated researchers all over the world to concentrate research efforts on this species. The *F. graminearum* release had the most dramatic effect of any filamentous fungal genome in the literature. By 2005, the number of publications listing *F. graminearum* in the title had more than

[1]Department of Plant Biology, Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA; e-mail: trail@msu.edu
[2]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Plant Industry; Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, Brisbane 4067, Australia; e-mail: Donald.Gardiner@csiro.au

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

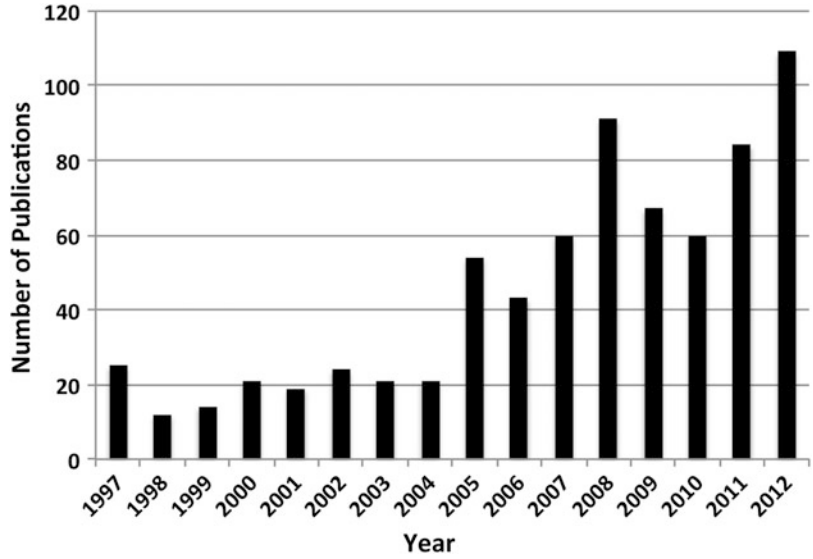


Fig. 11.1. Annual number of publications with '*Fusarium graminearum*' in the title. In 2003, the genome sequence was publicly released (Data from the Web of Science; accessed January 13, 2013 http://www.wokinfo.com)

doubled over publications in previous years, and the publication rate continues to increase (Fig. 11.1). Similarly, the number of publications with *F. verticillioides* in the title had increased by 2008, 2 years after the release of the genome, and continues to increase, but the effect has not been as great. The *F. graminearum* genome release served as a beacon for those interested in *Fusarium* diseases, particularly those working on mycotoxin contamination in agricultural crops. The genome sequence, together with the U.S. Department of Agriculture (USDA) Wheat and Barley Scab Initiative, which has been funding research on *F. graminearum* since 1997 (http://www.scabusa.org; McMullen et al. 2012), encouraged U.S. researchers in particular to focus on this species, and our current knowledge and understanding of this species compared to the other fusaria reflects this.

Availability of the genome sequence accelerates research progress and greatly increases the kinds of questions that can be investigated. The progress before and after the *F. graminearum* genome release is an example of the power of this technology. In 2002, researchers had identified a few pathogenicity genes, including the trichothecene biosynthesis gene *TRI5* (Proctor et al. 1995) and mitogen-activated kinase gene *MGV1* (Hou et al. 2002; Jenczmionka et al. 2003; Urban et al. 2003), mainly by knowing their functions in other pathogens, particularly in *M. oryzae*. Without mutants or knowledge of homologs, gene identification was challenging. For example, the trichodiene synthetase gene had been identified by screening an expression library with antibodies to the purified protein from *F. sporotrichioides* (Hohn and Beremand 1989). The trichothecene gene cluster was then identified through the complementation of several trichothecene mutants across the pathway by overlapping cosmids (Hohn et al. 1993). Knockout technology was available, but knockouts were laborious, as genetic sequences were not known, so cloning and sequencing of individual genes was essential to design knockout vectors. Thus, our knowledge of the genome and genomic potential was fragmentary.

The release of the *Fusarium* spp. genome sequences provides a catalogue of genes and a map showing how they are physically organised. In addition, because the genome sequence provides genetic sequences flanking each gene, gene knockouts are fairly easily obtained in several species, permitting large-scale knockout projects. Instead of the handful

of gene knockouts we obtained in 2002, we now have about 2,000 among the fusaria with sequenced genomes. Large-scale knockout projects can focus on gene families, resulting in information on how related genes are used for diverse purposes. In addition, with the great reduction in price of genome sequencing (by 1,000-fold since 2000), multiple isolates of some species have been sequenced, providing important information on population genetics and diversity. This has also been improved by the availability of high-throughput transcriptomics. Transcriptomics has allowed us to identify the genes expressed at a particular stage or condition. It allows us to identify genes and processes involved in a life-cycle stage. *Fusarium* researchers are also beginning to apply other -omics techniques to understand these important pathogens.

The *Fusarium* species have a wide range of life history characteristics, and those that now have genome sequences available are representative of that diversity. *Fusarium graminearum*, *F. pseudograminarum* and *F. verticillioides* are pathogens of grains. These three species produce mycotoxins that contaminate harvested grain and cause economic and health problems worldwide (CAST Report 2003). *Fusarium graminearum* causes head blight of wheat, barley and other small-grain cereals as well as stalk and ear rot of maize, and the species complex was recently divided into several species endemic to regions of the world (O'Donnell et al. 2000). *Fusarium verticillioides* causes grain molds primarily in corn and sorghum and is seed transmitted as an endophyte (Marin et al. 2004). Whereas *F. graminearum* is common in more temperate regions of the world, *F. verticillioides* affects plants stressed by heat and drought. Recent studies suggest warming trends in Europe will increase the occurrence of this species there (Doohan et al. 2003). In Australia and other arid areas of the world, *F. pseudograminearum* is the primary causal agent of wheat and barley crown and root rots (Burgess et al. 2001). Both *F. solani* and *F. oxysporum* are species complexes that have historically been broadly defined. *Fusarium oxysporum* is a wilt pathogen with a wide host range but is subdivided into *formae speciales* that have prominent host specialization (Armstrong and Armstrong 1981; Michielse and Rep 2009). Similarly, *F. solani* causes stem base infections of a wide variety of hosts. Both *F. oxysporum* and *F. solani* are known to cause opportunistic diseases in humans (Alastruey-Izquierdo et al. 2008). The diversity of life cycles of these species on a hypothetical plant is represented in Fig. 11.2.

Mycotoxins are fungal secondary metabolites that are toxic to humans and animals, and filamentous ascomycetes produce a large and diverse array of these metabolites. The mycotoxins of major concern to human health are the aflatoxins, trichothecenes, fumonisins, zearalenone, ochratoxins and ergot alkaloids (CAST Report 2003). *Fusarium* species produce the trichothecenes, fumonisins and zearalenone. In many first-world countries, strict regulations and risk-based monitoring systems for mycotoxin contamination of grain are in place, but in poorer regions of the world, such as parts of Africa, mycotoxin contamination is a major problem, particularly fumonisins in contaminated maize, given the importance of maize as a major dietary component (Siame and Nawa 2008; Dutton 2009; Kpodo and Bankole 2008).

The genes important for biosynthesis of secondary metabolites are often clustered in the genome and can be recognized by the presence of three classes of proteins that are typically involved in their biosynthesis: terpene synthetases (TSs); polyketide synthetases (PKSs); and non-ribosomal peptide synthetases (NPSs). Analysis of genes for these enzymes in the *F. graminearum* genome has revealed the potential to produce an estimated 50 different secondary metabolites (Cuomo et al. 2007), and *F. graminearum* has one of the best-studied mycotoxin arsenals of any of the filamentous fungi. A recent publication has summarized what is known about the PKS and NPS genes and their products from *Fusarium* species with sequenced genomes (Hansen et al. 2012). The Fusaria sequenced to date each contain between 13 and 15 PKS genes and between 13 and 19 NPS genes (Hansen et al. 2012). The numbers of unique PKS, NPS and TS genes compared to the known metabolites suggests we understand only a small proportion of the compounds that these species produce. The

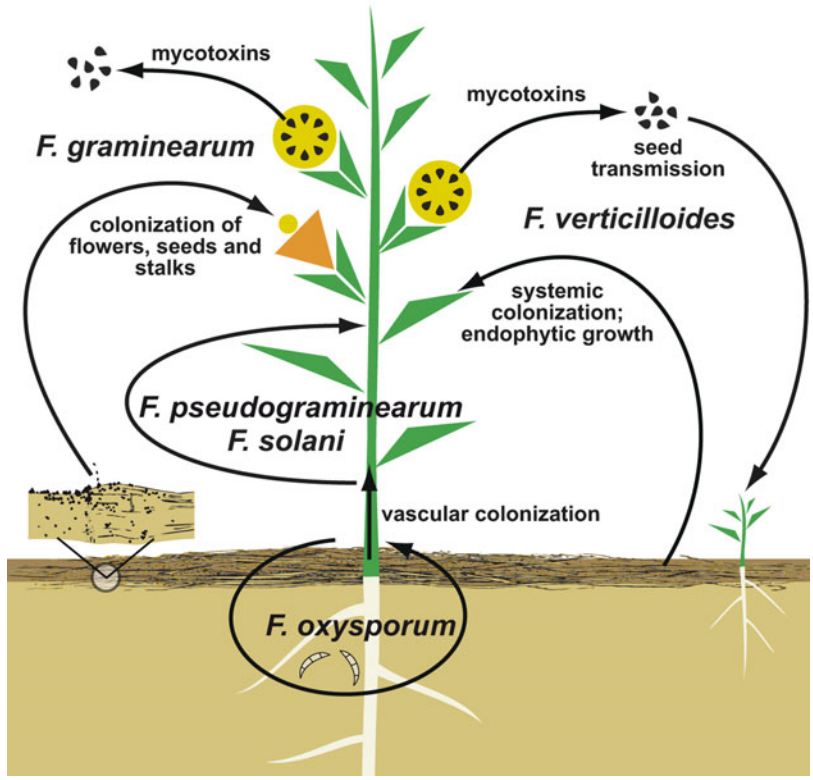


Fig. 11.2. Life cycle of the five core *Fusarium* species with sequenced genomes on a hypothetical plant

role of mycotoxins in virulence and the developmental process are discussed in the relevant sections that follow.

II. Comparative Genomics

For the four founding *Fusarium* genomes with higher-order assemblies, the chromosome number and genome size vary considerably among species. The *F. graminearum* genome has 4 chromosomes and is 36 Mb; that of *F. verticillioides* is 42 Mb arranged in 11 chromosomes. *Fusarium solani* f. sp. *pisi* has 17 chromosomes covering 54 Mb, and the *F. oxysporum* f. sp. *lycopersici* genome is 60 Mb distributed among 15 chromosomes (Cuomo et al. 2007; Coleman et al. 2009; Ma et al. 2010). The variable size of the genomes can be attributed predominantly to the rich repeat content in *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi*. In addition, the *F. solani* f. sp. *pisi* and *F. oxysporum* f. sp. *lycopersici* chromosome complement includes dispensable chromosomes that are variably present in different isolates and encode many host-specific pathogenicity factors.

One striking feature of the *Fusarium* genomes is the **macrosynteny** observed between genomes, despite the genus being more than 90 million years old (O'Donnell et al. 2013). Pairwise alignment between the four first-generation *Fusarium* genomes is represented in Fig. 11.3, and with the exception of the **dispensable chromosomes** in *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi*, large regions of colinearity can be observed. Although there is variation in chromosome number, these alignments have been used to infer that *F. graminearum* chromosomes were formed by fusion between the smaller chromosomes in an ancient lineage (Cuomo et al. 2007). For example, in Fig. 11.3, *F. graminearum* chromosome 1 can be seen to align with *F. verticillioides* chromosomes 1 and 5. In addition, much of the repeat content of each of the genomes

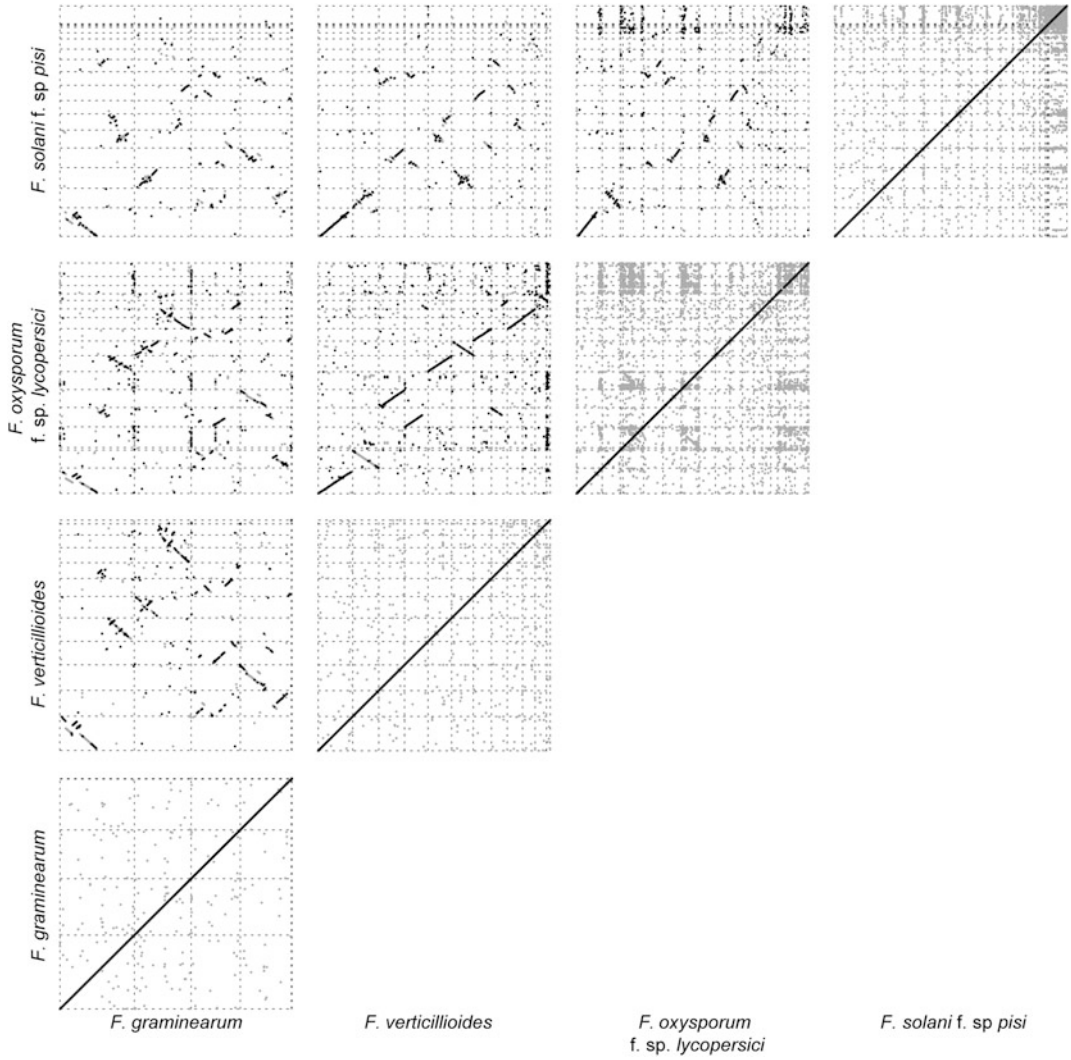


Fig. 11.3. Syntenic conservation between four *Fusarium* spp. *Black* and *grey* colours represent alignment between the sequences in forward and reverse directions, respectively. Chromosome boundaries for each species are marked by *horizontal* and *vertical lines* on each of the plots, with unscaffolded sequences appearing at the extreme *right* and *top* of each plot. The syntenic conservation observed between species has been used to infer the fusion of chromosome in the *F. graminearum* lineage (Cuomo et al. 2007)

seems to be species specific, represented by significant alignments detected away from the diagonals in the self-comparisons, suggesting repeat expansions in these genomes occurred postseparation of the four lineages (particularly in *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi*; Fig. 11.3). Furthermore, this repeat content is higher in the dispensable chromosomes (3, 6, 14, and 15 for *F. oxysporum* f. sp. *lycopersici* and 14, 15, and 17 for *F. solani* f. sp. *pisi*). The synteny that is observed between distantly related fusaria can be applied to the assembly of other *Fusarium* genomes (Fig. 11.3). For example, *Fusarium* genomes generated using next-generation technologies with no higher-order assembly can be aligned to their nearest *Fusarium* cousin to facilitate assembly of pseudochromosomal sequences. The synteny between *Fusarium*

species has also been used to define the core and dispensable chromosomes in *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi* (Coleman et al. 2009; Ma et al. 2010).

Analysis of the additional *F. oxysporum* genome sequences generated using Illumina technology available on the Broad website and that of an *Arabidopsis* pathogen *F. oxysporum* isolate 5176 generated with 454 data (Thatcher et al. 2012) will draw heavily on this founding *F. oxysporum* f. sp. *lycopersici* genome to define the dispensable chromosomes from each of these additional isolates.

Despite the differing genome sizes and chromosomal organization, a common feature of all *Fusarium* genomes to date is the subdivision of the genome into two (conceptual) compartments: one encodes many basic fungal functions (basal genome) and another adaptive genome that consists of genes that encode attributes such as virulence towards particular specific hosts or mechanisms of niche-specific survival. This organization provides a suggestion of how the *Fusarium* genomes have evolved. Indeed, there are specific characteristics of the two regions that suggest differing evolutionary constraints on the genes that reside in each. In these fungi, the **basal genome**, in contrast to the **adaptive genome**, is characterized by reduced numbers of polymorphisms among isolates of the same species, fewer repetitive sequences, and a relatively low density of genes for secreted proteins (Cuomo et al. 2007). In the sexually recombining *F. graminearum*, these regions are also low in genetic recombination rate per unit length of the genome (Gale et al. 2005), and it is likely that this will hold for the other sexually recombining species, such as *F. solani* and *F. verticillioides*, but the concept has yet to be demonstrated. Between-species comparisons have demonstrated high levels of synteny in these regions, which also supports the hypothesis that the small number of chromosomes in the *F. graminearum* lineage is due to chromosomal fusions (discussed previously; Cuomo et al. 2007).

The adaptive genome of the *Fusarium* species contrasts greatly with the basal genome. The adaptive regions feature increased levels of polymorphisms among isolates, higher representation of secreted protein-encoding genes, lower representation of between-species gene orthologs, higher genetic recombination rates in the sexually recombining species, and in *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi*, higher transposon density. In *F. oxysporum* f. sp. *lycopersici* and *F. solani* f. sp. *pisi*, these regions are predominantly contained on the dispensable chromosomes that are transferred horizontally between isolates and impart pathogenicity on specific hosts (Han et al. 2001; Ma et al. 2010; Rodriguez-Carres et al. 2008; Temporini and VanEtten 2004). In *F. graminearum*, the adaptive genome is located in subtelomeric regions and the ancient subtelomeric regions, now located internally in the chromosomes because of chromosomal fusion that is likely to have occurred in this lineage (Cuomo et al. 2007). Alignment between the relatively closely related species *F. graminearum* and *F. pseudograminearum* highlighted the highly variable nature of these regions. In this comparison, approximately 6 % of the *F. graminearum* genome was not found in *F. pseudograminearum*, and the gaps in the alignment were almost exclusively found in the terminal areas of each chromosome or in regions of ancient chromosome fusion (Gardiner et al. 2012). The subtelomeric arrangement of genome variable regions is a feature shared with other fungi, including the aspergilli (Fedorova et al. 2008).

Whole-genome comparisons between species have identified the adaptive genome component in individual species, and genes important for niche or host specialization/preference have been demonstrated to be contained within these regions. For example, all 19 nonribosomal peptide synthetase genes in *F. graminearum* except NPS2 and NPS13 are located in regions of high polymorphism or high genetic recombination rates. NPS2 encodes a siderophore for intracellular sequestration of iron, which can be considered a fundamental function in fungi (Tobiasen et al. 2007).

Another example of specialization functions present in the adaptive genome component is the pea pathogenicity gene cluster and genes for pea rhizosphere colonization, which are on a dispensable chromosome of *F. solani* f. sp. *pisi* (Han et al. 2001; Rodriguez-Carres et al. 2008).

However, for nearly all genes in the adaptive genome, biological roles are yet to be defined, and this represents a major challenge in understanding the unique biology of individual *Fusarium* species. An emerging trend in the literature is that these hypervariable or adaptive genome components contain many genes with unexpected evolutionary histories, suggesting that these regions may also be hosting places for genes acquired by horizontal transfer. Examples of genes important for virulence that have been shown to be horizontally transferred into or out of *Fusarium* spp., many of which are found in the adaptive component of the genomes, are presented in Table 11.1 with their chromosomal locations, if known.

III. Genomic Contribution to Understanding Fusaria as Plant Pathogens

Fusarium species are important agricultural crop pathogens worldwide. Genomics in the *Fusarium* genus is beginning to shed light on mechanisms of pathogenicity and of the variety of plant-*Fusarium* interactions (Fig. 11.2). From a whole-system point of view, the global gene expression analysis in the pathogen has driven our understanding of the processes occurring during infection. This is best characterized in *F. graminearum*, for which well-annotated Affymetrix **microarrays** have been developed (Güldener et al. 2006), including a multispecies microarray with probe sets for genes from each of the four founding *Fusarium* genomes (Jonkers et al. 2012a, b). In *F. graminearum*, experiments that have profiled the global expression during infection on wheat and barley heads and wheat stem bases have been undertaken (Güldener et al. 2006; Lysøe et al. 2011a; Stephens et al. 2008; Zhang et al. 2012). Through these reports, it is clear that processes of carbohydrate metabolism, degradation of host tissues for nutrient acquisition and utilization, and secondary metabolism are important to the infection process. A common feature of these analyses is that for genes specific to *in planta* activities, those with unknown function are overrepresented compared to the genome as a whole, suggesting that our understanding of infection processes is far from complete. For example, *F. graminearum* genes expressed specifically in wheat (but not barley and axenic culture) are preferentially encoded by the adaptive genome (Lysøe et al. 2011a).

In an elegant set of experiments, a highly detailed analysis of the infection processes in wheat coleoptiles was undertaken (Zhang et al. 2012). These authors described stage-specific infection of initial intercellular growth with minimal hyphal branching, followed by more invasive and intracellular growth with increased branching, which further increased over time to ultimately result in full intracellular ramification. Using laser microdissection microscopy, gene expression profiles at each of these infection stages were analyzed, revealing highly coordinated gene expression programs corresponding to differing cell-wall-degrading enzymes, secreted proteins and reactive oxygen species scavenging and production. The cell-wall-degrading enzyme expression profiles were related to the tissue invasion patterns at each infection stage. Likewise, extracellular reactive oxygen species scavenging was suggested to be occurring during initial infection followed by the active extracellular reactive oxygen species production later in the infection to kill the plant cells. The genome sequence availability and annotation have been instrumental in providing the gene expression analysis tools for studies such as these that transform our understanding of the interaction with the host. Extending these analyses to the other *Fusarium* spp. will undoubtedly inform these systems in a similar manner.

Detailed individual gene functional analysis via the generation of **gene knockouts** has also been fundamental to enhancing our understanding of pathogenicity processes, and the genome availability has greatly enhanced our ability to undertake these experiments. Recent species-specific reviews provide an overview of genes known to be involved in pathogenicity in *Fusarium* spp. (Michielse and Rep 2009; Walter et al. 2010; Kazan et al. 2012). A core group of common pathogenicity factors has been identified in several *Fusarium* species. In general,

Table 11.1. Examples of horizontal genetic transfer events with probable roles in virulence in *Fusarium* spp.

Event type	Description	Effect on pathogenicity	Mode of action	Species involved	Genomic location	Reference
Whole chromosome						
	Transfer of entire pathogenicity chromosome (14) under experimentally controlled circumstances	Acquisition renders recipient strain virulent on tomato	Chromosome encodes many effector genes	*F. oxysporum* f. sp. *lycopersici* (donor) to other *F. oxysporum* isolates	Chromosome 14	Ma et al. (2010)
Individual genes						
	Amidohydrolase gene (FpAH1)	Mutant shows reduced virulence towards wheat and barley	Possibly involved in detoxification of host metabolite(s)	Unknown bacterium (donor) to *F. pseudograminearum*	Possibly subtelomeric region	Gardiner et al. (2012)
	FoAve1	Not tested directly in *F. oxysporum* f. sp. *lycopersici*, but plants lacking the cognate resistance gene are more susceptible	The *Verticillium* homolog contributes to virulence on multiple hosts	Likely plant donor. Present in a small number of phytopathogenic fungi and one bacterium	Chromosome 14	de Jonge et al. (2012)
	Glucan glucosyltransferase	Unknown in *F.* oxysporum, but in *Verticillium dahliae* that shares this gene, it is important for wilt on tobacco		Bacterial donor (Rhizobiales). Present only in four fungi, *F. oxysporum* f. sp. *lycopersici*, two *Verticillium* spp. and *Metarhizium anisopliae*	Chromosome 8 in *F. oxysporum* f. sp. *lycopersici*	Klosterman et al. (2011)
Gene clusters						
	Dienlactone hydrolase (FpDLH1) and amidase	FpDLH1 mutant shows reduced virulence towards wheat and barley	Hypothesised detoxification of host defence compounds	Present in three cereal pathogens: *F. pseudograminearum*, *F. verticillioides* and *Colletotrichum graminicola*	Subtelomeric region of chromosome 10 in *F. verticillioides*	Gardiner et al. (2012)

Auxin biosynthesis	Not described		*F. proliferatum and F. verticillioides* (recipient) and unknown bacterium	Subtelomeric region of chromosome 3 in *F. verticillioides*	Tsavkelova et al. (2012)
Pea pathogenicity gene cluster (PEP)	The cluster is required for pathogenicity on pea. Four of the genes can independently increase a strain's ability to cause disease of pea	One of the genes detoxifies the pea phytoalexin pisatin	*F. solani f., sp. pisi*, AND *F. oxysporum f. sp. pisi*	Chromosome 14 in *F. solani*	Han et al. (2001), Milani et al. (2012), Temporini and VanEtten (2004)

these are central cell-signalling components that often have pleiotropic effects on the fitness of the organism. A number of **signalling pathways** have been demonstrated as important for virulence. The importance of these genes for pathogenicity was identified before most genomes were available using homologs identified in other species. Specific mitogen-activated protein kinase (MAPK) signalling pathways (Di Pietro et al. 2001; Hou et al. 2002; Jenczmionka et al. 2003; Urban et al. 2003) and other components that may interact with MAPK pathways such as the Ras GTPase RAS2 (Bluhm et al. 2007) have been shown to be important for virulence in multiple fusaria by multiple research teams. G-protein signalling components and downstream pathways are important for virulence, although the importance of specific components may differ from species to species (Jain et al. 2002, 2003, 2005; Sagaram and Shim 2007; Yu et al. 2008; Park et al. 2012). For example, the β subunit, which is typically a single gene in fungi, is necessary for virulence in *F. oxysporum* but not *F. verticillioides* (Jain et al. 2003; Sagaram and Shim 2007). Recently, the velvet complex, known to regulate key developmental steps and mycotoxin production in aspergilli through chromatin structure modification (Bayram et al. 2008), has been shown to be important for virulence in pathogenic fusaria, including the rice pathogen *F. fujikuroi*, *F. oxysporum* and *F. graminearum* (Merhej et al. 2012; Wiemann et al. 2010; Lee et al. 2012; López-Berges et al. 2013). The genes for many of these common pathogenicity factors are found in the basal genome (Beacham et al. 2009). Interactions between these signalling pathways are beginning to be teased out through large knockout projects of gene families (Wang et al. 2011) and will be aided by studies of transcriptional profiling of individual mutants. In addition, transcriptional networks built from profiling transcription factor mutants (Son et al. 2011) will link genes together.

The pleiotropic effects observed in many signalling pathway mutants make it difficult to dissect out the functional mechanisms that are controlled by these pathways. However, these mutants have been useful in characterizing the key control points in plant invasion and the downstream genes that are regulated by them. For example in *F. graminearum*, mutation of the BDM1 phosphotransducin-encoding gene (downstream component of G-protein signalling pathways) results in impaired colonization of maize silks but normal kernel infection, demonstrating different infection programs in these two tissue types, which are presumably mediated by distinct downstream functional pathway outputs (Horevaj and Bluhm 2012). Likewise, deletion of the *F. verticillioides VeA* gene (a velvet complex component) has developmental consequences, such as growth habit and cell wall defects in culture (Li et al. 2006), but in maize, asymptomatic colonization stages are unaffected, whereas pathogenicity and fumonisin production are abolished (Myung et al. 2012). Mutants such as these, which alter the interaction of the fungus with the host plant, will be crucial for understanding key differences between asymptomatic and pathogenic interactions. Such interactions are particularly important in *F. verticillioides*, which acts as both an endophyte and a pathogen of maize.

Signalling components important in plant diseases caused by fusaria have also been demonstrated to be important in mammalian mycoses, such as the Fmk1 MAP kinase and G-protein β-subunit, which act in concert in mice infection by *F. oxysporum* (Prados-Rosales et al. 2006). Understanding the commonalities and differences in pathogenicity of plant versus animal hosts will also allow us to distinguish genes in common with parasitizing both hosts and those unique to the environment in each host. Dissecting the differences and similarities of infection in different hosts has been particularly enlightening in bacterial pathogens (Staskawicz et al. 2001).

Fusaria possess a variety of **cell-wall-degrading** and other hydrolytic enzymes undoubtedly used during infection and colonization of other substrates to facilitate extraction of nutrients. Zhang et al. (2012) recently showed that a cellulose β-cellobiosidase and an endo-β-glucanase contribute to virulence on wheat in a coleoptile infection assay, but not on inoculated heads. The secreted lipase *FGL1* of *F. graminearum* is another hydrolytic

enzyme required for pathogenicity. Mutants of *FGL1* are severely reduced in pathogenicity towards wheat, barley and maize (Voigt et al. 2005; Ilgen et al. 2008). When expressed constitutively, *FGL1* partially restores virulence to the apathogenic mitogen-activated kinase 1 mutant (Δgpmk1) (Salomon et al. 2012), which is known to regulate the expression of a variety of cell-wall-degrading enzymes (Jenczmionka and Schäfer 2005). In addition, restoration of conidiation in the Δgpmk1 mutant by constitutive *FGL1* expression indicates a role for lipase activity in developmental processes as well as pathogenicity (Salomon et al. 2012).

A. Specialized Virulence Strategies: Genome Innovation to Access a Specific Niche

Genes underlying species-specific virulence strategies often reside in the adaptive genome. The expression of these virulence components that act in a host- or pathogen-specific manner are often regulated by central signalling pathways. However, in many *Fusarium*-host interactions, these outputs quantitatively contribute to virulence, making them somewhat difficult to functionally characterize. Examples include mycotoxin production, mechanisms to detoxify host defences, and the use of host molecule mimicry to induce susceptibility. Progress in these areas is discussed next.

1. Mycotoxins in Virulence

The **trichothecenes** are produced by *F. graminearum* and related species and not only are toxic to humans and animals but also remain the mycotoxin most clearly shown to be a virulence factor in plant-fungal interactions (Desjardins et al. 1992; Proctor et al. 1995; Bai et al. 2002; Ilgen et al. 2008; Jansen et al. 2005). Trichothecenes are inhibitors of the protein translational apparatus (Pestka 2007) and have elicitor-like activity in stimulating plant defences and cell death (Desmond et al. 2008; Nishiuchi et al. 2006). Delineation of the trichothecene biosynthetic gene cluster was accomplished before fungal genome availability through mutant analysis in *F. sporotrichioides* (Hohn et al. 1993), first with the identification of *TRI5*, the trichodiene synthetase gene (Hohn and Beremand 1989). All the genes involved in trichothecene biosynthesis and regulation are given a '*TRI*' designator followed by a numeral.

Trichothecene-producing *Fusarium* species, and different strains within a species, vary in the patterns of oxygenation and acetylation (chemotypes) produced, and the differential patterns of decoration are reflected in the genes present in the organism. Of the trichothecenes produced by *Fusarium* species, DON (deoxynivalenol) and NIV (nivalenol) and the DON acetylated derivatives (3-acetylaldeoxynivalenol, 3-ADON; 15-acetyldeoxynivalenol, 15-ADON) are the most important. *TRI13* and *TRI7* determine the production of ADON- or NIV-based trichothecenes (Lee et al. 2002). Recently, variants in *TRI8* were found to determine the production of 3-ADON or 15-ADON (Alexander et al. 2011). These two chemotypes, and their comparison to NIV-producing strains have been the basis for tracking population shifts worldwide (Miedaner et al. 2008; Starkey et al. 2007; Xu and Nicholson 2009; Ward et al. 2008; Suga et al. 2008; Yang et al. 2008; Zhang et al. 2010; Waalwijk et al. 2003).

Since the availability of genome sequences, comparative analyses within the fusaria and to other trichothecene-producing species has led to a thorough understanding of how these clusters have evolved and how genetics affects variation in trichothecene mycotoxins produced (Proctor et al. 2009; Cardoza et al. 2011). From a genome evolution perspective, trichothecene biosynthetic genes provide a good example of how other parts of the *Fusarium* genomes may evolve. The genes for biosynthesis in *F. sporotrichioides* are arranged in three genomic locations: one described as the core cluster containing most of the necessary genes (*TRI3* through *TRI14*) and two other loci containing either one gene (*TRI101*) or two genes (*TRI16* and *TRI1*). This arrangement is retained in *F. graminearum*, but the *TRI16* is non-functional and is not required for DON and NIV chemotypes produced by *F. graminearum* (Peplow et al. 2003; McCormick et al. 2004). In the *F.*

graminearum genome, the core cluster is located centrally on chromosome 2 near what appears to be an ancient chromosome fusion site. *TRI7* and *TRI13* are also pseudogenes in strains that produce 3-ADON and 15-ADON but are functional in strains that produce NIV (Lee et al. 2002). Polymorphisms within the cluster genes (e.g. *TRI7* and *TRI13*) are also observed in other trichothecene-producing species, indicating that the alleles of these genes evolved prior to speciation (Ward et al. 2002). In *F. equiseti*, *TRI1* and *TRI101* are located within the main cluster, and evolutionary analysis suggests that the genes were transferred into the cluster in this lineage (Proctor et al. 2009). Presumably similar genetic changes and genomic rearrangement mechanisms act on other genes in *Fusarium* spp. with important implications for the evolution of pathogenesis.

Fumonisins are polyketide-derived mycotoxins produced by *F. verticillioides.* They are sphingolipid analogs that may interfere with regular biosynthesis of these molecules in exposed cells by inhibiting ceramide synthetase (Abbas et al. 1994). The role of fumonisins in disease has been addressed with non-producing mutant strains of *F. verticillioides* with mixed results that may be dependent on the host maturity and tissue type. Neither deletion of the *FUM1* gene, encoding the PKS involved in fumonisin biosynthesis, nor mutations in key fumonisin biosynthesis genes affect maize ear rot (Desjardins and Plattner 2000; Desjardins et al. 2002). However, for disease in seedlings, the ability to produce fumonisins seems to be important (Desjardins et al. 1995). When the entire fumonisin gene cluster was introduced into a fumonisin non-producing *F. verticillioides* strain, which was apathogenic on maize, both disease and stunting were observed when the transformed strain was inoculated onto maize plants, strongly suggesting a role for fumonisin in host colonization (Glenn et al. 2008). The equivocal evidence for the role of fumonisins in the interaction with maize may be because of their activity as both a pathogenicity factor and an elicitor of host responses. It has been suggested that the balance between these two properties determines the interaction outcome (Arias et al. 2012).

The fumonisin biosynthetic gene cluster, consisting of 15 genes, has been identified and resides 800 Kb from the chromosome 1 terminus in *F. verticillioides* (Proctor et al. 1999, 2003). In two other fumonisin-producing species, *F. oxysporum* isolate O-1890 and *F. proliferatum*, the gene orders in the clusters are conserved with the order in *F. verticillioides*, but the flanking genes in all three species are different (Waalwijk et al. 2004; Proctor et al. 2008), suggesting the cluster may be derived by horizontal transfer. Whether this cluster resides in an adaptive genome region remains to be determined, and determining the evolutionary origin of the fumonisin cluster will greatly enhance our understanding of the evolution and role of these clusters in *Fusarium* biology.

The depsipeptide beauvericin is produced by some strains of *F. oxysporum*, has cytotoxic activity towards mammalian cells and insecticidal properties, and may act via interference ion homeostasis (Bryden et al. 2001; Wang and Xu 2012). Beauvericin has recently been identified as a virulence factor for *F. oxysporum* towards both tomato and mammalian hosts (López-Berges et al. 2013). In *F. oxysporum*, the beauvericin biosynthetic apparatus is located about 330 Kb from the end of chromosome 10.

Beauvericin is biochemically related to the enniatins from *F. avenaceum*, which are occasionally associated with diseases of small-grain cereals (Bottalico and Perrone 2002) and are known to contribute to disease towards potato (Herrmann et al. 1996). Genomic comparisons between *F. oxysporum*, *F. avenaceum* and other depsipeptide producers will undoubtedly reveal how the biosynthesis of these molecules evolved.

Other mycotoxins produced by the fusaria have also been identified and the clusters for their production characterized at the genomic level, including the gibberellins, fusaric acid, fusarin C and zearalenone (Tudzynski and Hölter 1998; Gaffoor et al. 2005; Brown et al. 2012a; Gaffoor and Trail 2006). Roles for fusarin C and zearalenone in pathogenicity were not found (Gaffoor et al. 2005). A role for fusaric acid in plant disease is yet to be explored; however, the phytotoxicity of fusaric acid would strongly

suggest an important role (Löffler and Mouris 1992). Gibberellins are discussed in the section on host susceptibility. An unknown metabolite is also involved in virulence towards wheat coleoptiles based on the mutation of the *NPS9* gene in *F. graminearum*, but the metabolite is unknown (Zhang et al. 2012). Given the diverse potential of any one *Fusarium* sp. to produce mycotoxins, dissecting the roles for every compound in disease processes and pathogen fitness will be challenging.

2. Detoxifying the Host Environment

A major component of the host response to pathogen invasion is the production of metabolites that have evolved to inhibit pathogen growth. Some of the almost-universal components of this response include **reactive oxygen species** and simple **phenolic compounds** (e.g. ferulic acid), and mechanisms to overcome or avoid these responses are likely to be universal to pathogens. For example, *F. oxysporum* f. sp. *lycopersici* uses the β-ketoadipate pathway to degrade plant-derived phenolics to allow invasion of the host (Michielse et al. 2012); this observation, based on the conservation of this pathway, is expected to be important in other interactions. The production of **phytoalexins** is common in nearly all plants in response to pathogen attack, but the chemical nature of phytoalexins is usually specific to the host species. Therefore, to be a pathogen on a particular host one must be able to overcome these specific antimicrobial compounds. One of the best-studied examples of this is the ability of *F. solani* f. sp. *pisi* to detoxify the pea phytoalexin pisatin. *F. solani* f. sp. *pisi* does this using a combination of a cytochrome P450 monooxygenase (pisatin demethylase, PDA) and an ATP-binding cassette (NhABC1) transporter protein, which both quantitatively contribute to pisatin tolerance and virulence (Wassman and VanEtten 1996; Coleman et al. 2011). *PDA* is located on chromosome 14, one of the dispensable chromosomes in *F. solani* f. sp. *pisi*, and is located within a cluster of at least three genes shown to be important for virulence on pea (Han et al. 2001). Other isolates of *F. solani* detoxify the related phytoalexin maackiain, the genes for which are also likely to be located on a dispensable chromosome (Miao and VanEtten 1992). An ABC transporter in *F. sambucinum*, homologous to NhABC1, is also involved in tolerance to an unrelated phytoalexin produced by potato, rishitin, where mutation of the transporter results in reduced pathogenicity towards potato (Fleißner et al. 2002). Homologs of these transporters are also present in other *Fusarium* spp., suggesting that a conserved mechanism for dealing with phytoalexins may be present, and functional characterization of these ABC transporter genes in *Fusarium* pathogens of other hosts may reveal the importance of phytoalexins in these interactions, which, particularly in the monocot hosts, are largely unknown.

Another example of phytoalexin detoxification contributing to virulence in the fusaria is the degradation of the tomato saponin α-tomatine by the tomatinase produced by *F. oxysporum* f. sp. *lycopersici* (Pareja-Jaime et al. 2008). The chromosomal location of the tomatinase gene (*tom1*) is unknown as this sequence is in the unpositioned scaffolds of the *F. oxysporum* f. sp. *lycopersici* genome.

Genomic studies have also begun to dissect the detoxification pathway in *F. verticillioides* for the benzoxazolinones (BOA and MBOA), preformed antimicrobials from maize. Using subtractive hybridization, a cluster of genes highly expressed in response to BOA exposure was identified (Glenn and Bacon 2009). A gene in this cluster encodes an acetyltransferase involved in the biochemical transformation of BOA. An adjacent gene, encoding a transcription factor, was proposed to regulate the expression at this locus; indeed, transformants harbouring a mutation in this gene were partially reduced in the ability to detoxify BOA. This two-gene cluster for one step of BOA detoxification is located in the subtelomeric region of chromosome 10, within a few genes of the *FvDLH1-FvAMD1* gene cluster, which is likely to be horizontally transferred between species (Table 11.1) (Gardiner et al. 2012). However, the ability to detoxify BOA is not necessary for virulence towards maize in seedling assays or in endophytic colonization assays

(Glenn et al. 2002). Benzoxazolinones are also produced by wheat and rye (Niemeyer 1988), and fusaria that cause disease on these hosts, such as *F. culmorum*, are known to effectively metabolize them (Friebe et al. 1998). However, the mechanism of degradation of these toxins by *F. culmorum* and the role in virulence are yet to be elucidated. Comparative genomics between genomes of strains of *F. culmorum*, which are in progress, and *F. verticillioides* will be interesting (Scherm et al. 2013).

3. Inducing Host Susceptibility

A common feature of many *Fusarium*-host interactions is the manipulation of the host via the endogenous hormone-signalling pathways. Although the pathways and mechanisms differ between *Fusarium* spp., there are a number of key examples that highlight the exquisite hijacking of endogenous host processes to facilitate disease, including the auxin, jasmonate and gibberellin.

The most striking manipulation of the host via phytohormone production is that of the **gibberellins** by *F. fujikuroi* mating population C (MP-C), the causal agent of bakanae disease of rice (also known as foolish seedling disease). The fungal-produced gibberellins act on the endogenous host gibberellin receptors and pathway to induce a hyperextension phenotype. Indeed, the gibberellin hormones take their name from the (historical) teleomorph epithet of *F. fujikuroi*, *Gibberella fujikuroi*, and this fungus has been instrumental in the discovery and understanding of the role of gibberellins in normal plant development (Bömke and Tudzynski 2009 and references therein). The biosynthetic pathway and underlying genes for gibberellin biosynthesis in *F. fujikuroi* have been elucidated and shown to be different from the pathway in plants (Bömke and Tudzynski 2009). The genes are clustered in the genome, resembling a secondary metabolite gene cluster (Tudzynski and Hölter 1998). Comparative analysis of the other members of the species complex in which *F. fujikuroi* MP-C resides identified the presence of the complete gene cluster within nearly all species, but only *F. fujikuroi* MP-C actually synthesizes gibberellins (Malonek et al. 2005).

Specific isolates of *F. oxysporum* are known to produce **jasmonates** (Miersch et al. 1999), which presumably trigger responses of this pathway in the host. How production of these molecules is encoded in the host genome is unknown, but a secondary metabolite biosynthetic gene cluster that includes a PKS gene would be one possibility. Indeed, in the *F. oxysporum-Arabidopsis* interaction, mutations in key regulators of the jasmonate response in the host alter the outcome of the interaction. Mutations in the jasmonate receptor gene *COI1* abolish jasmonate defence responses and render the host resistant, although colonization by the pathogen still occurs (Thatcher et al. 2009). The transcription factor MYC2, which differentially regulates outputs of the jasmonate response pathway in *Arabidopsis* (Dombrecht et al. 2007), is also an important susceptibility factor for interactions with *F. oxysporum* (Anderson et al. 2004). In contrast, host biosynthesis of jasmonate is not required for susceptibility (Thatcher et al. 2009). Taking these findings together, Thatcher et al. (2009) postulated that the pathogen produces a jasmonate mimic to trigger jasmonate-mediated senescence responses in the host, which leads to susceptibility.

The salicylic acid, cytokinin, and abscisic acid responses have not been described to date as targets of *Fusarium* pathogens, but pathogens from other genera are known to manipulate these pathways, so it is possible that specific fusaria may also do so (e.g., see Djamei et al. 2011; Bruce et al. 2011). Likewise, ethylene perception is required for susceptibility to *F. oxysporum* f. sp. *lycopersici* in tomato (Lund et al. 1998), and auxin signalling and transport perturbations impart increased resistance to *F. oxysporum* infection in *Arabidopsis* (Kidd et al. 2011), although whether there is a specific manipulation of these pathways by the pathogen remains to be elucidated.

Some fusaria synthesize **auxins** with the genes encoded in a biosynthetic cluster acquired from bacteria (Tsavkelova et al. 2012). In *F. verticillioides*, which produces small amounts of auxin, the biosynthetic genes for auxin production are located 90 Kb

from the terminus of chromosome 3, which is presumably a component of the adaptive genome. With the examples described previously, it seems modulation of phytohormone signalling by fusaria is a common theme of pathogenesis.

4. Secreted Proteins in the Interaction Between Fusaria and Host Plants

Among the fusaria, only *F. oxysporum* isolates are known to interact with their host plants in a gene-for-gene manner. That is, specific host resistance genes are known to interact genetically with specific pathogen avirulence genes, resulting in resistance to the pathogen. Recently, genes encoding SIX (Secreted in Xylem) proteins in *F. oxysporum* have been shown to be avirulence genes. Three proteins – Avr3 (SIX1), Avr2 (SIX3) and Avr1 (SIX4) – identified by their presence in the xylem of infected tomato plants via proteomic studies, have been shown to be the products of avirulence genes in *F. oxysporum* f. sp. *lycopersici*, where they interact with cognate resistance genes in the host; however, in tomato this does not seem to result in a hypersensitive response (Rep et al. 2004; Houterman et al. 2008, 2009). Despite being identified by their production in xylem, the avirulence proteins are likely to have their effect in the xylem parenchyma (Houterman et al. 2009). Each of these genes also has virulence functions based on mutant studies. Avr1 inhibits resistance initiated by recognition of the Avr2 and Avr3 determinants (Houterman et al. 2008), but the mechanisms of virulence for Avr2 and Avr3 are yet to be identified (Rep et al. 2005). Through whole-genome sequencing, an almost-identical copy of Avr1 was identified in the *Arabidopsis*-infecting *F. oxysporum* strain Fo5176, where it contributes to virulence. However, resistance in *Arabidopsis* seems to be quantitative and not mediated by classical resistance genes (Thatcher et al. 2012). Thus, the mechanism of virulence in tomato and *Arabidopsis* may be different for Avr1. In addition to the three avirulence proteins described, *F. oxysporum* f. sp. *lycopersici* secretes a number of other small proteins into the xylem vessels, which are presumed also to have roles in infection. Most of these seem to be encoded on the dispensable chromosomes of *F. oxysporum* f. sp. *lycopersici* (Houterman et al. 2007).

Secreted protein effectors that act inside the host cell have yet to be identified in other *Fusarium* spp. However, it is difficult to conceive that secreted proteins would not be involved, as in many fungal-plant interactions secreted proteins are known to be important as toxic proteins or proteins that manipulate host responses. Paper et al. (2007) isolated 120 proteins secreted during *F. graminearum* infection by isolating apoplastic fluid of infected wheat heads. Functional analysis of 37 genes identified in that study showed no change in virulence for 35 of the genes, with the remaining 2 showing slight reduction in virulence (Paper 2011). Recently, the predicted **secretome** of *F. graminearum* was re-analysed (Brown et al. 2012b), and previous observations of the concentration of secreted proteins encoding genes in genomic regions that are undergoing higher rates of genetic change were confirmed (Cuomo et al. 2007). Brown et al. (2012b) hypothesized, based on the infection patterns of and host responses to various *F. graminearum* mutants, that factors other than the mycotoxin DON are clearly involved in the interaction and host recognition of this fungus. The expression of secreted proteins during infection, some of them exclusively during infection, strongly suggests a role for them in virulence (Lysøe et al. 2011a; Zhang et al. 2012). Two different studies examining different tissue showed that more than 40 % of the proteins described as exclusively expressed *in planta* are predicted to be secreted, compared to the genome as a whole, for which only 13 % of proteins have signal peptides (Lysøe et al. 2011a; Zhang et al. 2012). One of the potential problems in understanding how secreted proteinaceous effectors are involved in an interaction is the possibility that each individual effector may only have a small quantitative contribution to virulence, which is difficult to detect using loss-of-function mutants. However, quantitative reduction in an isolate's ability to invade a plant will undoubtedly be selected against over generations in the field.

Recent publications from other pathosystems are now redefining what an effector of virulence might look like. Historically, small cysteine-rich secreted proteins have been thought of as prime candidates, and indeed they have been. However, a number of lines of evidence suggest an expanded view of effectors is necessary. First, alternative mechanisms for secretion of proteins, independent of a signal peptide, are likely to exist (Yang et al. 2012); Paper et al. (2007) identified *F. graminearum* proteins in the apoplast of infected wheat heads that were not predicted to contain standard secretion signals. Second, effectors from both bacterial and fungal pathogens are known to include relatively large proteins with enzymatic activities (e.g. see Djamei et al. 2011; Feng et al. 2012).

IV. From Pathogen to Saprotroph: Overwinter Survival

All known *Fusarium* species that cause disease on plants can be described as facultative pathogens. Most species survive on crop residue remaining in the field after harvest or present in the soil from previous years. The two alternating stages of the life cycle, a pathogenic stage on the host and a dormant or saprobic stage, are linked by dispersal to the new host, and selection likely occurs during both stages. Relatively little is known about the molecular mechanisms of *Fusarium* spp. survival and competition against other microbes in these dormant or saprobic stages, but this may be a key component of pathogen fitness. However, we are beginning to understand the processes involved in the development and differentiation of the spores and other structures *Fusarium* spp. uses both as inoculum and to survive adverse conditions.

Fusarium species can produce potentially four types of spores: macroconidia, microconidia, chlamydospores and the sexual ascospores. Spores are the primary inoculum source for *Fusarium* diseases, being formed on the colonized residues of the previous crop. The macroconidia and microconidia are relatively short-lived and probably serve to link the pathogen from crop to crop. Thick-walled chlamydospores are likely to survive well in the soil. Further investigations are needed to determine their relative contribution to disease initiation in species that produce them. Unlike the fruiting bodies of many ascomycetous fungi, those of *Fusarium* species are ephemeral and do not serve as the overwintering stage. In *F. graminearum*, it has been shown that lipid-filled perithecium initials overwinter in substomatal cavities in wheat straw and develop perithecia in warmer temperatures (Guenther and Trail 2005). Ascospores seem to provide the main inoculum source for head blight in many areas of the world. These airborne spores are dispersed more broadly than the limited range of macroconidia. However, in Australia, where head blight is less common than crown rot, macroconidia are the main airborne inoculum source, and isolates of *F. graminearum* do not produce the copious numbers of perithecia that are common in isolates from other parts of the world (Mitter et al. 2006). The limited dispersal of macroconidia may be partially responsible for the less-prominent presence of *F. graminearum* and head blight in Australia (Chakraborty et al. 2006). In the next sections, the contribution of genomics to understanding sporulation and sexual development in the fusaria and how this relates to the overall fitness of the pathogens are discussed.

V. Genomics of Sporulation

A. Asexual Sporulation

As with most fungi, sporulation is the most important means of dispersal in *Fusarium* spp. Asexual sporulation (conidial formation) is always present in these species, but sexual sporulation is known only in some species. Asexual cycles can consist of **microconidia** and **macroconidia** or just macroconidia. For fungi without prominent sexual cycles, conidia are the primary agents of dispersal and the major infective agents (Dahlberg and Van Etten 1982). In *Fusarium* species, macroconidia are formed in **phialides**, flask-shaped conidiogenous cells. The macroconidia

are canoe shaped and are held together in a slimy matrix within the **sporodochia**, a cluster of phialides and the conidia they produce. The shape of these spores, as well as the slimy matrix, helps these spores disperse by rain (Deacon 2006). Microconidia are also produced by phialides and generally accumulate in small spheres of spores called **false heads**, although in *F. verticillioides*, they also appear in chains emerging from the phialides (Leslie and Summerell 2006). Microconidia have roles as spermatia, for fertilizing female structures, but can also serve as germinating spores. Microconidia are produced by *F. oxysporum* and *F. verticillioides*. Germination can be slow or unreliable if they serve primarily as spermatia, but in the *Fusarium* spp., microconidia seem to be reliable germinators. Chlamydospores are prominent in *F. oxysporum* and are produced distinct from the conidia, intercalated in the hyphae. In *F. graminearum*, chlamydospores are produced from enlarged individual cells within the macroconidia. Identification of genes that affect production of one or both types of conidia or production of chlamydospores could be used to study the importance of spore types in the disease cycle and pathogen survival. In addition, the availability of genomes of multiple *Fusarium* species offers a way to study the evolution of conidiogenesis, including mechanisms that result in the absence of a spore type or morphological variations among spores of different species.

The role of microconidia has traditionally been regarded as one of fertilization (Dodge 1930; Maheshwari 1999); however, the species of *Fusarium* that produce microconidia, including *F. verticillioides*, *F. oxysporum* and *F. solani*, produce microconidia that germinate and can thus function in dispersal, similar to the macroconidia. Furthermore, there is no known sexual stage for *F. oxysporum*, so the role of microconidia in this species is likely not limited to fertilization. We are not aware of any studies that focused on different roles of macro- and microconidia in the field. However, all *Fusarium* strains pathogenic to humans produce microconidia, and many are associated with the formation of biofilms in plumbing drains (Short et al. 2011). The majority of clinical isolates (78 %) fall into the *F. solani* and *F. oxysporum* species complexes (Short et al. 2011).

The study of the roles of micro- and macroconidia in *Neurospora intermedia*, which is found in burnt sugarcane fields, found that they were produced at different stages of the life cycle (Pandit and Maheshwari 1996). In the early stages of colonization of crop residues, when the substrate was rich in nitrogen and sugars, macroconidia dominated. In the depleted, desiccated tissue, several weeks after macroconidia had ceased to be produced and the substrate was moistened by rain, the microconidia predominated. At this time, perithecia began to form as well, suggesting a role in fertilization.

Aspergillus and *Neurospora* have long served as model systems for the genetics of conidiation (see reviews by Park and Yu 2012; Etxebeste et al. 2010). The *con* genes are a group of genes originally identified by their expression during macroconidiation (Springer and Yanofsky 1992). They are also associated with microconidia formation. Studies of conidiation in the *Fusarium* species have been initiated to understand how the pathways differ from these models. Because of the current availability of genome sequences, however, these studies have begun to move beyond identification of orthologs to studies of transcriptional regulation in the *Fusarium* species. Transcription factors MedA and StuA modulate gene regulation across development. In *F. oxysporum*, REN1 has high similarity to MedA of *Aspergillus*, which is a transcriptional regulator of conidiogenesis (Ohara et al. 2004). The *REN1* mutants lack conidiophores and phialides and form conidia-like cells that are rod shaped instead of canoe shaped and are formed directly on the hyphae. Production of the microconidia is also affected in the *REN1* mutants. The nitrite reductase gene in *F. oxysporum* (*FoNIIA*) is regulated by REN1 (Iida et al. 2008). Mutants of *FoNIIA* produce morphologically normal micro- and macroconidia but produce significantly fewer microconidia than the wild type. The role of FoNIIA in conidiation is not clear, but direct nutritional roles were eliminated using reduced nitrogen sources in the experiments. In *F. graminearum*, the transcription factor FgStuA was shown to affect conidiation, with mutants lacking spores and phialides. This mutant also did not produce perithecia. As was shown in studies of the ortholog in *Aspergillus nidulans* (Clutterbuck 1969), StuA is a global

regulator, affecting expression of thousands of genes (Lysøe et al. 2011b). In *F. oxysporum*, deletion of the *FoSTUA* ortholog increased chlamydospore production but reduced macroconidium production. In these mutants, phialide and conidiophore formation were affected for macroconidia, but microconidial formation was not altered. Originally identified in *Magnaporthe grisea*, the transcription factor Htf1, a homeodomain protein, has been found to be essential for phialide development and conidiogenesis in *F. graminearum*, *F. verticillioides* and *F. oxysporum* (Zheng et al. 2012). The mutants produced clusters of aberrant phialides that did not give rise to conidia but produced aberrant conidia directly on the tips of vegetative hyphae. It would be interesting to pursue the role of this gene in other genera, perhaps to reveal how phialide function has evolved.

Light is an integral part of sporulation for many fungi. Red-to-yellow Light wavelengths increased conidiation at high intensity, but the greatest conidiation was observed under short-wave blue light in *F. verticillioides* (Fanelli et al. 2012). The White Collar protein (Wc-1), first studied in *N. crassa*, is known to be involved in light-sensing mechanisms in filamentous fungi. Functional studies of the Wc-1 homolog in *F. fujikuroi* (wco-1) has demonstrated its importance in conidiation rates, although this is likely not caused by an interaction with light, as light does not seem to affect conidiation (Estrada and Avalos 2007). The velvet gene *VeA* has been well studied in *A. nidulans*, where it integrates developmental responses to light, along with *VelB* and *LaeA*, inhibiting sexual development (Bayram et al. 2008). In *F. verticillioides*, deletion of *Ve1* affected colony development, including formation of aerial mycelia, increased macroconidium formation, resulted in fewer microconidia, and eliminated the formation of microconidia in chains (Li et al. 2006). The addition of osmotic stabilizers to the medium restored the wild-type phenotype. *VeA* has also been studied in *F. graminearum* (Jiang et al. 2011a), for which mutants exhibited an increase in conidial production but delayed germination. Interestingly, conidia of the mutants accumulated a higher number of lipid droplets, potentially resulting from a higher basal level of glycerol than the wild-type parent. A large number of fatty acid biosynthesis genes were upregulated in the mutant. Mutants of the type 2C protein phosphatase gene were also reported to result in large lipid droplets in conidia of *F. graminearum* (Jiang et al. 2011b).

Analysis of gene expression during conidiation has been done in several species. Seong et al. (2008) measured gene expression during *F. graminearum* conidial germination at conidial swelling, germ tube emergence, hyphal elongation and branching. This comprehensive analysis of gene expression, done using Affymetrix GeneChips, revealed the processes involved in each stage of conidial development. In *F. oxysporum*, a study of comparisons between gene expression in vegetative hyphae versus conidiating cultures using expressed sequence tags, revealed genes that were specific to conidiation (Iida et al. 2006). Changes in gene expression patterns in *REN1* and *FoSTUA* mutants enabled the authors to identify genes regulated by these transcription factors (Iida et al. 2007). These two studies resulted in the identification of the nitrite reductase gene as involved in macroconidium formation (see Section VA). *FoNIIA* is positively regulated by REN1, and transcripts accumulate to much higher levels in conidiating cultures (Iida et al. 2008). With the increasing power of high-throughput sequencing, comparative expression studies during conidiation are needed for all of these species and for all spore types to understand transcriptional changes that have resulted in evolution of these distinct morphologies and life cycles. In addition, the recent identification of differentiated infection structures in *F. graminearum* on hosts wheat and *Brachypodium* (Peraldi et al. 2011; Boenisch and Schaefer 2011) indicates the potential for comparisons of prepenetration differentiation among species growing on different hosts.

B. Sexual Development

The availability of the *Fusarium* genome sequences from species across a spectrum of life histories provides an opportunity to understand the genetics and evolution of reproductive

processes. Of the species with fully sequenced genomes, sexual cycles are present in *F. graminearum*, *F. verticillioides*, *F. pseudograminearum* and *F. solani*. *Fusarium graminearum* is **homothallic**, meaning that it is self-compatible; thus, a single strain can produce fruiting bodies. *Fusarium solani*, *F. pseudograminearum* and *F. verticillioides* are **heterothallic**, meaning that two strains, harbouring different mating-type genes, are required to produce fertile fruiting bodies. Of the four fertile species, the sexual stage seems to play a prominent role in the disease cycle of *F. graminearum*, for which ascospores are major dispersal agents (Shaner 2003; Desjardins et al. 2006), and some formae speciales of *F. solani*. For example, in *F. solani* on pepper, perithecia have been shown to form on vines and be a major source of inoculum (Ikeda 2010). In *F. verticillioides*, perithecia are seen occasionally on crop residues or in stored contaminated material (Marin et al. 2004). There is evidence of recombination in *F. pseudograminearum*, as indicated by the presence of both mating types in similar frequencies (Bentley et al. 2008). *Fusarium graminearum* has become a model for understanding development of perithecia in the ascomycetes, and there are now many studies of the roles of individual genes involved in sexual development.

Many individual genes that have been functionally characterized in *F. graminearum* have been determined to affect sexual development. A comprehensive review of these genes and their roles in development is published elsewhere (Trail 2013). Genomics has been most useful in identifying gene families, on which functional studies can be performed on the members. Prior to the availability of genome sequences, genes identified as involved in perithecium development were mainly those that arrested sexual development entirely. Since the availability of methods to perform large-scale functional studies, more genes have been identified that are involved in the intermediate stages of perithecium development (Trail 2013). Studies of gene families in multiple species offer insight into the evolution of functional changes during speciation. In addition, gene expression studies can be performed to identify potentially important processes, which can then be used for further investigation.

Genes that control mating compatibility (MAT genes) have been intensively studied in fungi (reviewed in Heitman et al. 2007). The *MAT1* locus may harbour idiomorphs MAT1-1 or MAT1-2, determinants of mating type (Glass et al. 1988). The term *idiomorph* is used because *MAT1-1* and *MAT1-2* are not alleles. The *MAT1-1* locus consists of three genes (*MAT1-1-1*, *MAT1-1-2*, *MAT1-1-3*), and *MAT1-2* consists of one gene, *MAT1-2-1* (Butler 2007). These genes regulate expression of mating-type-specific functions. Strains with different idiomorphs are compatible. In homothallic species, the presence of both idiomorphs allows sexual development to progress in a single strain. The genome of *F. graminearum* harbours both the *MAT1-1* and the *MAT1-2* loci, rendering it homothallic. Replacement of the *MAT1* locus in *F. graminearum* with either idiomorph results in strains that are functionally self-incompatible and will outcross (Lee et al. 2003). This strategy has been used to generate strains for genetic analyses (Metzenberg and Glass 1990).

Strains of *F. verticilliodes*, *F. pseudograminearum* and *F. solani* are heterothallic and harbour either the *MAT1-1* genes or the *MAT1-2* gene, and strains with different MAT genes are sexually compatible (Arie et al. 2000; Bentley et al. 2008). *Fusarium oxysporum* strains also harbour either *MAT1-1* or *MAT1-2* (Yun et al. 2000; Arie et al. 2000), indicating the presence of a sexual cycle, but the sexual stage of *F. oxysporum* has not been reported. Expression studies of genes under regulation of the MAT genes have been performed in *F. graminearum* (Lee et al. 2006). Transcriptomic analyses of *F. graminearum* and *F. verticillioides* during mating and fruiting body development revealed surprising patterns of expression of the MAT1 genes (Sikhakolli et al. 2012).

In *F. graminearum*, all four *MAT1* genes showed high expression levels through perithecium maturation. In contrast, *MAT1-2* showed detectable expression only during the mating/early development period in *F. verticillioides*, whereas the *MAT1-1* genes showed similar expression to those of *F. graminearum*. These

results indicate a broader role for MAT1 genes during fruiting body development, and may indicate a role for *MAT1-2* in maintaining nuclear identity in homothallic species during sexual development. Furthermore, MAT genes may have more extensive roles during sexual development. More work in this area, among both the *Fusarium* species and a broader range of fungi, will enhance our understanding of the evolution and role of MAT genes in regulating mating and sexual development.

The functional studies of gene families have focused on regulatory genes, including transcription factors, G-proteins, G-protein regulators and kinases. G-proteins are important signal-transducing proteins that often relay signals from membrane-bound receptors to cytoplasmic proteins. G-proteins comprise three subunits: α, β, and γ. A functional study of G-proteins in *F. graminearum* revealed three Gα proteins (GPA1, 2, and 3), one Gβ protein, and one GpGγ gene called GpG1 (Yu et al. 2008). Interestingly, the study revealed that the Gα- and Gβ-proteins regulated either sexual development or pathogenicity, but there was no crossover regulation across both processes by a single protein. *GPA1* regulates female fertility, and none of the other G-proteins shows a functional relationship to sexual development when deleted. A study of *GBB1*, a G-protein *B*-subunit gene in *F. verticillioides*, revealed that it also was not involved in sexual development (Mukherjee et al. 2011). Regulators of G-protein signalling (RGSs) proteins control heterotrimeric G-protein signalling by specifically interacting with the Gα subunit, usually by turning off activated G-proteins through enhancing the GTPase activity of the subunit. In *F. graminearum*, seven RGS genes were identified and functionally characterized (Park et al. 2012). Deletion mutants of *FgFlbA*, *FgRgsB* and *FgRgsC* were affected in sexual development. Single strains of the *FgFlbA* mutant did not develop perithecia and were not successful as female parents. Mutants of *FgRgsB* developed perithecia that matured at a slower rate, and mutants of *FgRgsC* produced abnormal ascospores, although perithecia seemed normal. *FgFlbA* also seemed to regulate the ratio between sexual and asexual development. Thus, the Gα subunit has a global effect on sexual reproduction, and the regulators have narrower effects.

Kinases are the middlemen in the signal transduction system, relaying signals coming in from G-proteins through protein phosphorylation. Functional analysis of the kinome, the complete set of kinases, in *F. graminearum* has been completed (Wang et al. 2011). Some groups of kinases are closely linked on the chromosomes and retain synteny with *F. oxysporum* and *F. verticillioides*. Of 117 kinases identified in the genome, 20 protein kinases yielded mutants that no longer underwent sexual development, mutants of 6 asci no longer produced, mutants of 13 resulted in defective ascospore development, and mutants of another 6 no longer forcibly discharged spores. These results indicate that a large proportion of the 117 are in some way involved in sexual development, reflecting both the complexity of the developmental process and the importance of this process in *F. graminearum*. It would be interesting to compare these functions to other species within the genus and outside.

A large-scale functional analysis was performed of all predicted transcription factors in the genome of *F. graminearum* (Son et al. 2011), with 657 predicted transcription factor genes characterized for 17 phenotypes. Effects on perithecium development and function were shown by 170 transcription factors. The authors categorized the mutants of 107 genes as having effects on sexual development: Mutants of 44 did not initiate the sexual cycle, 4 produced larger numbers of perithecia, 12 were arrested early in development, and 23 produced fewer numbers of normal perithecia. Several additional mutants produced fewer perithecia and had other abnormalities also: Mutants of 9 had abnormal ascospore shape, and 19 were arrested in ascospore formation. For those that produced normal perithecium abundance, mutants of five genes formed perithecia containing abnormal spores, and mutants of an additional gene did not produce spores. Of the 657 genes examined, only 4 of the transcription factors affected conidiation specifically. In contrast, 13 affected sexual development specifically. This comprehensive study

reflected both the complexity of regulation of the sexual cycle and the importance to the biology of *F. graminearum*.

Network analyses done on the regulatory pathways in *F. graminearum* (Son et al. 2011; Wang et al. 2011) were based on extending the protein-protein interactions in yeast to the orthologs in *F. graminearum*. More work on these interactions in filamentous fungi directly is essential, as regulatory pathway gene families have greatly expanded in these organisms.

Several global expression studies have been done over six stages of perithecium development in *F. graminearum* and *F. verticillioides*. A study using the Affymetrix *Fusarium* GeneChip across five stages of development revealed about 2,000 genes that were expressed uniquely in sexual development (Hallen et al. 2007). A second study examined the same stages of sexual development of *F. graminearum* on infected wheat stalks (Guenther et al. 2009). Together, these studies support the model that lipids accumulate early in substrate colonization, and that these lipid stores are depleted during development of perithecia. The ephemeral nature of perithecia in this species suggests that the lipid-filled hyphae function as the overwintering stage for *F. graminearum*. The transcriptional shifts across the same five stages of perithecium development were examined in both *F. graminearum* and *F. verticillioides* using Illumina-based RNA sequencing, revealing that the majority of genes expressed during sexual development fell into the category 'unclassified proteins', reflecting our poor understanding of these important structures and the need for functional genomics studies of these genes to understand their roles (Sikhakolli et al. 2012). More cross-species and cross-genera comparisons of gene expression in sexual development not only will yield important information on the genes involved but also will reveal the evolutionary processes that are reflected in their diverse morphologies.

1. Secondary Metabolism and Sporulation

Several secondary metabolites have been associated with sporulation in *Fusarium* species. The most striking are the **polyketides** that produce **pigments**, associated with different stages of *Fusarium* life cycles. The availability of several *Fusarium* genomes has enabled identification and characterization of these pigments, as well as the study of their evolutionary relatedness. The characterization of these pigments was one of the first outcomes of the release of the *F. graminearum* genome. The black perithecium wall pigment in *F. graminearum*, *F. verticillioides* and *F. fujikuoi* is synthesized by a PKS encoded by *PGL1* (Gaffoor et al. 2005; Brown et al. 2012a; Proctor et al. 2007; Studt et al. 2012). The structure of that pigment was recently determined to be a naphthoquinone called fusarubin (Studt et al. 2012). In contrast, *F. solani* produces reddish perithecia, and the pigment, produced by the polyketide synthetase PKSN, is closely related to the reddish mycelial pigments of many *Fusarium* species (Graziani et al. 2004; Studt et al. 2012). Although *F. solani* has a *PGL1* homolog, the lack of black pigmentation together with the results from the study of *PKSN* indicate that this gene either is nonfunctional or produces a different pigment not yet recognized (Proctor et al. 2007; Studt et al. 2012). The function of the black perithecial pigment is not known, but its colour suggests protection of the developing meiospores from light, particularly UV (ultraviolet) rays, and reactive oxygen species (Graziani et al. 2004).

The most recognizable feature of *Fusarium*-infected corn and wheat straw is the appearance of pink regions associated with mycelia. *Fusarium verticillioides*, *F. fukijuroi* and *F. oxysporum* produce a related reddish pigment, the polyketide bikaverin (Hansen et al. 2012; Wiemann et al. 2009; Tatum et al. 1987). In *F. graminearum*, the reddish pigment is produced by *PKS12 (Aur1)*, which is unique to *F. graminearum*, and is associated with the polyketide aurofusarin and its dimer rubrofusarin (Gaffoor et al. 2005; Kim et al. 2005a; Malz et al. 2005). *Aur1* and *PKSN* of *F. solani* are closely related and more distant from the genes for bikaverin (Studt et al. 2012). Knockout strains for *PGL1* in *F. graminearum* produce perithecia that are pale but have a reddish hue, whereas knockouts of the aurofusarin PKS gene (*PKS12*) have a flat, black appearance (Gaffoor et al. 2005). Therefore, the purplish-blue colour

characteristic of the *F. graminearum* perithecia is caused by the combination of both of these pigments (Gaffoor et al. 2005). The *PGL1* knockouts of *F. fujikuroi* are colourless (Studt et al. 2012), indicating that the fusarubin contributes the entire perithecium pigmentation in that species.

Although pathogenic hyphae in host plants must protect themselves from host responses, the fruiting bodies and overwintering hyphae and spores in the crop residues and soil are susceptible to predation by other microbes, insects, animals, and the environment. The arsenal of secondary metabolites can function to protect the vulnerable resting stages, and there is some indirect evidence that this occurs. The antimicrobial activity of aurofusarin (Medentsev et al. 1993), combined with its production on colonized straw, suggests it may be used as a protective mechanism for overwintering perithecium initials and hyphae. Previously, production of zearalenone Mycotoxins had been hypothesized to stimulate sexual development in *F. graminearum* (Wolf and Mirocha 1973) and in other ascomycetes (Nelson 1971). Gene deletion experiments have resulted in identification of two PKS genes (*PKS4* and *PKS13*) that are essential for production of zearalenone (Kim et al. 2005b; Gaffoor and Trail 2006; Lysøe et al. 2006). The absence of zearalenone in these knockout strains did not affect perithecium production. The ability of zearalenone to initiate sexual development in other species could be potentially beneficial if it resulted in dormancy for the target species, thus allowing *Fusarium* to continue to colonize the shared substrate. An understanding of how microbes interact in association with hosts will come from elucidation of the host microbiomes. As has been demonstrated by recent studies of the human microbiome (Lepage et al. 2013; Morgan et al. 2013), our understanding of microbes associated with host plants is essential to our understanding of crop health.

In relation to the polyketides, the characterization, genetics and biosynthesis of **non-ribosomal peptides** and **terpenes** have been relatively unexplored in fungi, although there is also a great diversity of biosynthetic genes for these compounds. NPSs have been shown to have a role in sporulation. *Fusarium graminearum* has the most unique NPS genes of the sequenced *Fusarium* species and has a total of 19 NPS genes (Cuomo et al. 2007; Hansen et al. 2012). Several NPS genes have been shown to function as siderophores (Oide et al. 2007; Tobiasen et al. 2007), and one of them is essential for conidium germination (Bushley et al. 2008). Of the 39 different NPS genes identified in *Fusarium* genome sequences, only 4 have known products (Hansen et al. 2012). This family, and its possible role in other aspects of *Fusarium* biology, is ripe for exploration.

VI. Future Prospects

A number of genomic tools, such as association and comparative genomics, population-scale whole-genome sequencing, network analysis and large-scale gene knockout collections are now becoming available that will provide new routes for investigating the biology of these important pathogens. The potential applications of these are discussed next.

Association and comparative genomics are promising routes for identifying virulence gene candidates. **Association genomics** (within species), in which genetic variants (usually single-nucleotide polymorphisms, SNPs) within many sequenced individuals are used to identify genes associated with a particular variant, was recently used to identify the avirulence gene product Ave1 and its homolog in *F. oxysporum* f. sp. *lycopersici* (de Jonge et al. 2012). This strategy is likely to become an important tool as genome sequencing costs decrease and tools for genome analysis improve. **Comparative genomics** studies (between species) rely on the central hypothesis that if a gene (or virulence strategy) is shared among organisms with a common biological feature (e.g. lifestyle or host) and is absent from species where the feature is absent, that gene or virulence strategy is likely to be involved in that biological process. Two recent studies that were either *Fusarium* focused or included fusaria successfully used comparative genomics to identify completely novel virulence factors that were proven via gene knockouts (Gardiner et al. 2012; Klosterman et al. 2011).

In the analysis of the *F. pseudograminearum* genome, an amidohydrolase (FpAH1) and a two-gene cluster encoding an amidase (FpAMD1) and a dienelactone hydrolase (FpDLH1) were identified as only present in other cereal pathogens and absent from all non-cereal pathogens in a collection of 27 fungal genomes (Gardiner et al. 2012). Likewise, in the analysis of wilt-causing pathogens, presence of a glucan glucosyltransferase (FOXG_02706) in the wild pathogens and absence of this gene from non-wilt-causing fungi was used to identify this gene as important for wilt diseases, at least on some hosts (Klosterman et al. 2011).

Both of these studies used BLASTp comparisons to establish homology between species. Other approaches that use clustering of homologous proteins or query tools that are more adept at detecting remote homologies will allow more robust analysis of shared virulence genes or sporulation processes between species. Extending comparative analyses to species that are closely related to pathogenic fusaria, but may act as saprotrophs or endophytes, will also increase the power of such approaches.

The comparative genomic analyses conducted across diverse species (Gardiner et al. 2012; Klosterman et al. 2011) both identified bacteria as likely sources of virulence genes in the fusaria. Indeed, 16 additional genes in *F. pseudograminearum* were identified that had some support for being of bacterial origin. For example, a gene encoding an aminotransferase present in only *F. graminearum*, *F. pseudograminearum* and *F. verticillioides*, but no other fungi, was 75 % identical to a protein from a plant endophytic bacterium. These analyses suggested bacteria are a rich source for 'new' genes in the fungi, and these may have shaped the evolution of virulence in this genus. To understand the biological species that may have acted as donors for these genes, metagenome analyses of host-associated microbes will be important.

Network analyses and coregulation also offer an opportunity to better predict genes with roles in virulence and developmental programs. Particularly in *F. graminearum*, ever-expanding gene expression data in the public sphere (including the Plant Expression Database, http://plexdb.org) can be mined to identify sets of genes that are coregulated or anticoregulated to predict genes that might share roles in similar processes. Recently, a network of transcription factor expression in *F. graminearum* was built as part of a larger study (Son et al. 2011), and expanding this network to all genes in the genome may associate more genes to known biological processes. Higher-resolution time courses of many biological processes, including development, growth on various substrates and infection, will increase the power of these network analyses.

Coregulation of genes induced when mycotoxin biosynthesis was high was used to identify two genes in *F. graminearum* that, when mutated, result in strains with higher virulence that could be attributed to increased trichothecene biosynthesis (Gardiner et al. 2009). The mechanism at play occurred most likely via alterations in substrate availability, as one of the genes was later shown to be involved in biosynthesis of culmorin, another terpenoid secondary metabolite (McCormick et al. 2010). Coregulation of these genes by the *TRI6* transcriptional regulator was confirmed using mutants of *TRI6* in independent studies (Gardiner et al. 2009; Seong et al. 2009).

Analysis of the gene expression alterations in the transcription factor mutants will also be a fruitful method to understand the downstream regulons of every one of these transcription factors. Mutants in nearly every *F. graminearum* transcription factor are now available (Son et al. 2011). Indeed, such an analysis of the two regulators of trichothecene biosynthesis, TRI6 and TRI10, revealed an extensive network of genes outside the trichothecene biosynthesis pathway that were also under the control of these transcription factors (Seong et al. 2009). This study also revealed an interesting example of genome adaptation; the promoters of genes for primary metabolite pathways that feed into trichothecene biosynthesis were regulated by (and contained binding sites for) the TRI6 transcription factor, and these binding sites were absent in non-trichothecene-producing fusaria. Similar analyses have been undertaken to understand how the Wor1-like transcription factor regulates trichothecene production, virulence, morphological changes and both asexual and sexual

reproduction (Jonkers et al. 2012a). Similarly analyses of other transcription factor mutants will undoubtedly reveal their downstream regulons, and this approach is being applied to other genes and many aspects of pathogen biology (Lysøe et al. 2011b; Iida et al. 2007).

Next-generation genome sequencing of whole populations of mutated isolates is now feasible. These types of approaches will facilitate an increased ability to ascribe roles for candidate genes in mechanisms of virulence/pathogenicity, developmental processes or any other biological phenomenon for that matter. However, mutant population-based screens will need to utilize multiple alleles of the same gene and be coupled with follow-up experiments, such as complementation or genetic purification of mutations. For these studies, there will clearly be a continuing need for improved methods for large-scale functional analysis, such as streamlined knockout procedures (Paz et al. 2011) and improved phenotyping methods (Son et al. 2011).

Most of our understanding of *Fusarium* (molecular) biology to date has been via reductionist approaches using forward or reverse genetics. However, in many simple systems, the understanding of organism biology has moved to much larger functional genomics approaches, which can capture higher-resolution quantitative information about the role of every gene in a particular aspect of biology. For example, in yeast, genome-wide deletion sets of strains individually barcoded can be used to understand both simple traits like growth rate (Winzeler et al. 1999) or more complex traits like the genes that are important for tolerance to particular xenobiotics (Hillenmeyer et al. 2008). Deploying a similar approach in any *Fusarium* spp. will be extremely challenging but has the potential to massively expand our understanding of a variety of biological processes.

With the recent publication of the wheat, barley and tomato genomes (Brenchley et al. 2012; Barley Genome Consortium 2012; Tomato Genome Consortium 2012) along with the maize genome (Schnable et al. 2009), we now have corresponding **host gene catalogues** for every published *Fusarium* genome except *F. solani* f. sp. *pisi* (Smýkal et al. 2012). In addition, *Arabidopsis thaliana* has been used in numerous studies as a model host, particularly in the study of host resistance mechanisms (e.g. Kidd et al. 2011; Thatcher et al. 2009; Arabidopsis Genome Initiative 2000). *Brachypodium* has recently been shown to be a model host for study of pathogen ingress by *F. graminearum* (Peraldi et al. 2011). Availability of this grass model system is particularly important because the barley and maize genomes are large, and the wheat genome is so complex because of its hexaploid nature. Studies of host pathogen systems from both sides of the interaction will greatly enhance our ability to decode diverse interaction outcomes. This knowledge will lead to the most promising applications of fungal genomics: the use of transgenics and manipulation of the microbiome to control these diseases.

Although much work has been done to understand the biology of these core *Fusarium* species, they remain a challenge for agriculture worldwide. Shifts in weather patterns caused by global warming will cause changes in species and mycotoxin accumulation because these species are quite adaptable. The use of triazoles for control of head blight has greatly expanded in recent years, resulting in the challenge of managing fungicide resistance. The potential for increased aggressiveness and shifts in mycotoxin chemotype frequencies is real and threatening, particularly with the common use of moderately resistant and susceptible cultivars. A recent study by Foroud et al. (2012) compared traditional and emergent populations of *F. graminearum* on different wheat lines of varying resistance and susceptibility. Although highly resistant lines were resistant to all strains, moderately resistant lines were susceptible to emergent 15-ADON and NIV populations, which were more aggressive and accumulated higher levels of trichothecenes. The presence of actively changing populations in a changing environment highlights the need for continued and enhanced efforts to develop more effective and diverse control strategies, and understanding the basic biology of these organisms is critical to delivering informed control strategies.

In a recent review, McMullen et al. (2012) stated: 'When the U.S. Wheat and Barley Scab Initiative was created in 1997, breeders assumed that it was just a matter of time until resistant cultivars would solve the problem of FHB. This has proved not to be the case.' Indeed, neither traditional breeding nor the study of fungal genomics has yet provided a durable solution to these diseases. Throughout the last decade, fungal genomics has provided the first systematic approaches to discovery of genes and processes essential for all stages of the fungal life cycle: spores, toxins, pathogenicity, host-pathogen interactions. Without the genome sequence, we would not have nearly the insight into these pathogens that we now have. We must now apply our knowledge in creative ways to resolve these challenges over the next decade.

References

Abbas HK, Tanaka T, Duke SO, Porter JK, Wray EM, Hodges L, Sessions AE, Wang E, Merrill AH Jr, Riley RT (1994) Fumonisin- and AAL-toxin-induced disruption of sphingolipid metabolism with accumulation of free sphingoid bases. Plant Physiol 106:1085–1093

Alastruey-Izquierdo A, Cuenca-Estrella M, Monzón A, Mellado E, Rodríguez-Tudela JL (2008) Antifungal susceptibility profile of clinical *Fusarium* spp. isolates identified by molecular methods. J Antimicrob Chemother 61:805–809

Alexander NJ, McCormick SP, Waalwijk C, van der Lee T, Proctor RH (2011) The genetic basis for 3-ADON and 15-ADON trichothecene chemotypes in *Fusarium*. Fungal Genet Biol 48:485–495

Anderson JP, Badruzsaufari E, Schenk PM, Manners JM, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K (2004) Antagonistic interaction between abscisic acid and jasmonate-ethylene signaling pathways modulates defense gene expression and disease resistance in *Arabidopsis*. Plant Cell 16:3460–3479

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Arias SL, Theumer MG, Mary VS, Rubinstein HR (2012) Fumonisins: probable role as effectors in the complex interaction of susceptible and resistant maize hybrids and *Fusarium verticillioides*. J Agric Food Chem 60:5667–5675

Arie T, Kaneko I, Yoshida T, Noguchi M, Nomura Y, Yamaguchi I (2000) Mating-type genes from asexual phytopathogneic ascomycetes *Fusarium oxysporum* and *Alternaria alternate*. Mol Plant-Microbe Interact 13:1330–1339

Armstrong GM, Armstrong JK (1981) Formae speciales and races of *Fusarium oxysporum* causing wilt diseases. Pennsylvania State University Press, University Park, pp 392–399

Bai GH, Desjardins AE, Plattner RD (2002) Deoxynivalenol non-producing *Fusarium graminearum* causes initial infection, but does not cause disease spread in wheat spikes. Mycopathologia 153:91–98

Barley Genome Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 491:711–716

Bayram O, Krappmann S, Ni M, Bok J-W, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon N-J, Keller NP, Yu J-H, Braus GH (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science 320:1504–1506

Beacham AM, Antoniw J, Hammond-Kosack KE (2009) A genomic fungal foray. Biologist 56:98

Bentley AR, Summerell BA, Burgess LW (2008) Sexual compatibility in *Fusarium pseudograminearum* (*Gibberella coronicola*). Mycol Res 112:1102–1106

Bluhm BH, Zhao X, Flaherty JE, Xu JR, Dunkle LD (2007) RAS2 regulates growth and pathogenesis in *Fusarium graminearum*. Mol Plant-Microbe Interact 20:627–636

Boenisch M, Schaefer W (2011) *Fusarium graminearum* forms mycotoxin producing infection structures on wheat. BMC Plant Biol 11:110

Bömke C, Tudzynski B (2009) Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. Phytochemistry 70:1876–1893

Bottalico A, Perrone G (2002) Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe. Eur J Plant Pathol 108:611–624

Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, Kay S, Waite D, Trick M, Bancroft I, Gu Y, Huo N, Luo M-C, Sehga S, Gill B, Kianian S, Anderson O, Kersey P, Dvorak J, McCombie WR, Hall A, Mayer KFX, Edwards KJ, Bevan MW, Hall N (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature 491:705–710

Brown DW, Butchko RAE, Busman M, Proctor RH (2012a) Identification of gene clusters associated with fusaric acid, fusarin, and perithecial pigment production in *Fusarium verticillioides*. Fungal Genet Biol 49:521–532

Brown NA, Antoniw J, Hammond-Kosack KE (2012b) The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: a refined comparative analysis. PLoS ONE 7:e33731

Bruce SA, Saville BJ, Emery RJN (2011) *Ustilago maydis* produces cytokinins and abscisic acid for potential regulation of tumor formation in maize. J Plant Growth Regul 30:51–63

Bryden WL, Logrieco A, Abbas HK, Porter JK, Vesonder RF, Richard JL, Cole RJ (2001) Other significant *Fusarium* mycotoxins. In: Summerell BA, Leslie JF, Backhouse D, Bryden WL, Burgess LW (eds) *Fusarium*: Paul E. Nelson memorial symposium. APS Press, St. Paul, pp 360–392

Burgess LW, Backhouse D, Summerell BA, Swan LJ (2001) Crown rot of wheat. In: Summerell BA, Leslie JF, Backhouse D, Bryden WL, Burgess LW (eds) *Fusarium*: Paul E. Nelson memorial symposium. APS Press, St. Paul, pp 271–294

Bushley KE, Ripoll DR, Turgeon BG (2008) Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis. BMC Evol Biol 8:328

Butler G (2007) The evolution of MAT: the Ascomycetes. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA (eds) Sex in fungi: molecular determination and evolutionary implications. ASM Press, Washington, pp 3–18

Cardoza RE, Malmierca MG, Hermosa MR, Alexander NJ, McCormick SP, Proctor RH, Tijerino AM, Rumbero A, Monte E, Gutiérrez S (2011) Identification of loci and functional dharacterization of trichothecene biosynthesis genes in filamentous fungi of the genus *Trichoderma*. Appl Environ Microbiol 77:4867–4877

CAST Report (2003) Mycotoxins: risks in plant, animal, and human systems. Council for Agricultural Science and Technology, Ames

Cavinder B, Trail F (2012) The role of Fig1, a component of the low affinity calcium uptake system (LACS), in growth and sexual development of filamentous fungi. Eukaryot Cell 11:987–988

Chakraborty S, Liu CJ, Mitter V, Scott JB, Akinsanmi OA, Ali S, Dill-Macky R, Nicol J, Backhouse D, Simpfendorfer S (2006) Pathogen population structure and epidemiology are keys to wheat crown rot and *Fusarium* head blight management. Australas Plant Pathol 35:643–655

Clutterbuck AJ (1969) A mutational analysis of conidial development in *Aspergillus nidulans*. Genetics 63:317–327

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S, Schwartz DC, Freitag F, Ma L-J, Danchin EGJ, Henrissat B, Coutinho PM, Nelson DR, Straney D, Napolil CA, Barker BM, Gribskov M, Rep M, Kroken S, Molna I, Rensing C, Kennell JC, Zamora J, Farman ML, Selker EU, Salamov A, Shapiro H, Pangilinan J, Lindquist E, Lamers C, Grigoriev IV, Geiser DM, Covert SF, Temporini E, VanEtten HD (2009) The genome of Nectria haematococca: contribution of supernumerary chromosomes to gene expansion. PLoS Genet 5: e1000618

Coleman JJ, White GJ, Rodriguez-Carres M, VanEtten H (2011) An ABC transporter and a cytochrome P450 of *Nectria haematococca* MPVI are virulence factors on pea and are the major tolerance mechanisms to the phytoalexin pisatin. Mol Plant-Microbe Interact 24:368–376

Cuomo CA, Güldener U, Xu J-R, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma L-J, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, DeCaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Munsterkotter M, Nelson D, O'Donnell K, Ouellet T, Qi WH, Quesneville H, Roncero MIG, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao JQ, Birren BW, Kistler HC (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317:1400–1402

Dahlberg KR, Van Etten JL (1982) Physiology and biochemistry of fungal sporulation. Annu Rev Phytopathol 20:218–301

de Jonge R, van Esse HP, Maruthachalam K, Bolton MD, Santhanam P, Saber MK, Zhang Z, Usami T, Lievens B, Subbarao KV, Thomma BPHJ (2012) Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing. Proc Natl Acad Sci U S A 109:5110–5115

Deacon JW (2006) Fungal biology. Blackwell, Malden

Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu J-R, Pan HQ, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeyer C, Li WX, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434:980–986

Desjardins AE, Plattner RD (2000) Fumonisin B1-nonproducing strains of *Fusarium verticillioides* cause maize (*Zea mays*) ear infection and ear rot. J Agric Food Chem 48:5773–5780

Desjardins AE, Hohn TM, McCormick SP (1992) Effect of gene disruption of trichodiene synthetase on the virulence of *Gibberella pulicaris*. Mol Plant-Microbe Interact 5:214–222

Desjardins AE, Plattner RD, Nelsen TC, Leslie JF (1995) Genetic analysis of fumonisin production and virulence of *Gibberella fujikuroi* mating population A (*Fusarium moniliforme*) on maize (*Zea mays*) seedlings. Appl Environ Microbiol 61:79–86

Desjardins AE, Munkvold GP, Plattner RD, Proctor RH (2002) *FUM1* – a gene required for fumonisin biosynthesis but not for maize ear rot and ear

infection by *Gibberella moniliformis* in field tests. Mol Plant-Microbe Interact 15:1157–1164

Desjardins AE, Plattner RD, Shaner G, Brown DW, Buechley G, Proctor RH, Turgeon GG (2006) Field release of *Gibberella zeae* genetically modified to lack ascospores. In: Canty SM, Clark A, Van Sanford D (eds) *Fusarium* head blight forum. University of Kentucky, Research Triangle Park

Desmond OJ, Manners JM, Stephens AE, Maclean DJ, Schenk PM, Gardiner DM, Munn AL, Kazan K (2008) The *Fusarium* mycotoxin deoxynivalenol elicits hydrogen peroxide production, programmed cell death and defence responses in wheat. Mol Plant Pathol 9:435–445

Di Pietro A, Garcia-MacEira FI, Meglecz E, Roncero MI (2001) A MAP kinase of the vascular wilt fungus *Fusarium oxysporum* is essential for root penetration and pathogenesis. Mol Microbiol 39:1140–1152

Djamei A, Schipper K, Rabe F, Ghosh A, Vincon V, Kahnt J, Osorio S, Tohge T, Fernie AR, Feussner I, Feussner K, Meinicke P, Stierhof YD, Schwarz H, Macek B, Mann M, Kahmann R (2011) Metabolic priming by a secreted fungal effector. Nature 478:395–398

Dodge BO (1930) Breeding albinistic strains of the monilia bread mold. Mycologia 22:9–38

Dombrecht B, Xue GP, Sprague SJ, Kirkegaard JA, Ross JJ, Reid JB, Fitt GP, Sewelam N, Schenk PM, Manners JM, Kazan K (2007) MYC2 differentially modulates diverse jasmonate-dependent functions in *Arabidopsis*. Plant Cell 19:2225–2245

Doohan FM, Brennan J, Cooke BM (2003) Influence of climatic factors on *Fusarium* species pathogenic to cereals. Eur J Plant Pathol 109:755–768

Dutton M (2009) The African *Fusarium*/maize disease. Mycotoxin Res 25:29–39

Estrada AF, Avalos J (2007) The White Collar protein WcoA of *Fusarium fujikuroi* is not essential for photocarotenogenesis, but is involved in the regulation of secondary metabolism and conidiation. Fungal Genet Biol 45:705–718

Etxebeste O, Garzia A, Espeso EA, Ugalde U (2010) *Aspergillus nidulans* asexual development: making the most of cellular modules. Trends Microbiol 18:569–576

Fanelli F, Schmidt-Heydt M, Haidukowski M, Susca A, Geisen R, Logrieco A, Mulè G (2012) Influence of light on growth, conidiation and fumonisin production by *Fusarium verticillioides*. Fungal Biol 116:241–248

Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemiewx S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, White OR, Whitty BR, Youngman P, Wolfe KH, Golman GH, Wortman JR, Jiang B, Denning DW, Nierman WC (2008) Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet 4: e1000046

Feng F, Yang F, Rong W, Wu X, Zhang J, Chen S, He C, Zhou J-M (2012) A *Xanthomonas* uridine 5′-monophosphate transferase inhibits plant immune kinases. Nature 485:114–118

Fleißner A, Sopalla C, Weltring K-M (2002) An ATP-binding cassette multidrug-resistance transporter is necessary for tolerance of *Gibberella pulicaris* to phytoalexins and virulence on potato tubers. Mol Plant-Microbe Interact 15:102–108

Foroud NA, McCormick SP, MacMillan T, Badea A, Kendra DF, Ellis BE, Eudes F (2012) Greenhouse studies reveal increased aggressiveness of emergent Canadian *Fusarium graminearum* chemotypes in wheat. Plant Dis 96:1271–1279

Friebe A, Vilich V, Hennig L, Kluge M, Sicker D (1998) Detoxification of benzoxazolinone allelochemicals from wheat by *Gaeumannomyces graminis* var. *tritici*, *G. graminis* var. *graminis*, *G. graminis* var. *avenae*, and *Fusarium culmorum*. Appl Environ Microbiol 64:2386–2391

Gaffoor I, Trail F (2006) Characterization of two polyketide synthetase genes involved in zearalenone biosynthesis in *Gibberella zeae*. Appl Environ Microbiol 72:1793–1799

Gaffoor I, Brown DW, Plattner R, Proctor RH, Qi W, Trail F (2005) Functional analysis of the polyketide synthetase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). Eukaryot Cell 4:1926–1933

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Brau EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CPC, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden Y, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868

Gale LR, Bryant JD, Calvo S, Giese H, Katan T, O'Donnell K, Suga H, Taga M, Usgaard TR, Ward TJ, Kistler HC (2005) Chromosome complement of the fungal plant pathogen *Fusarium graminearum*

based on genetic and physical mapping and cytological observations. Genetics 171:985–1001

Gardiner DM, Kazan K, Manners JM (2009) Novel genes of *Fusarium graminearum* that negatively regulate deoxynivalenol production and virulence. Mol Plant-Microbe Interact 22:1588–1600

Gardiner DM, McDonald MC, Covarelli L, Solomon PS, Rusu AG, Marshall M, Kazan K, Chakraborty S, McDonald BA, Manners JM (2012) Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. PLoS Pathog 8:e1002952

Glass NL, Vollmer SJ, Staben C, Grotelueschen J, Metzenberg RL, Yanofsky C (1988) DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. Science 241:570–573

Glenn AE, Bacon CW (2009) *FDB2* encodes a member of the arylamine N-acetyltransferase family and is necessary for biotransformation of benzoxazolinones by *Fusarium verticillioides*. J Appl Microbiol 107:657–671

Glenn AE, Gold SE, Bacon CW (2002) Fdb1 and Fdb2, *Fusarium verticillioides* loci necessary for detoxification of preformed antimicrobials from corn. Mol Plant-Microbe Interact 15:91–101

Glenn AE, Zitomer NC, Zimeri AM, Williams LD, Riley RT, Proctor RH (2008) Transformation-mediated complementation of a *FUM* gene cluster deletion in *Fusarium verticillioides* restores both fumonisin production and pathogenicity on maize seedlings. Mol Plant-Microbe Interact 21:87–97

Graziani S, Vasnier C, Daboussi MJ (2004) Novel polyketide synthetase from *Nectria haematococca*. Appl Environ Microbiol 70:2984–2988

Guenther J, Trail F (2005) The development and differentiation of *Gibberella zeae* (anamorph: *Fusarium graminearum*) during colonization of wheat. Mycologia 97:232–240

Guenther JC, Hallen-Adams HE, Bücking H, Shachar-Hill Y, Trail F (2009) Triacylglyceride metabolism by *Fusarium graminearum* during colonization and sexual development on wheat. Mol Plant-Microbe Interact 22:1492–1503

Güldener U, Seong K-Y, Boddu J, Cho S, Trail F, Xu J-R, Adam G, Mewes H-W, Muehlbauer GJ, Kistler HC (2006) Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression *in vitro* and *in planta*. Fungal Genet Biol 43:316–325

Hallen H, Huebner M, Shiu S-H, Guldener U, Trail F (2007) Gene expression shifts during perithecium development in *Gibberella zeae* (anamorph *Fusarium graminearum*), with particular emphasis on ion transport proteins. Fungal Genet Biol 44:1146–1156

Han Y, Liu X, Benny U, Kistler HC, VanEtten HD (2001) Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. Plant J 25:305–314

Hansen FT, Sørensen JL, Giese H, Sondergaard TE, Frandsen RJN (2012) Quick guide to polyketide synthetase and nonribosomal synthetase genes in *Fusarium*. Int J Food Microbiol 155:128–136

Heitman J, Kronstad JW, Taylor JW, Casselton LA (eds) (2007) Sex in fungi: molecular determination and evolutionary implications. ASM Press, Washington

Herrmann M, Zocher R, Haese A (1996) Effect of disruption of the enniatin synthetase gene on the virulence of *Fusarium avenaceum*. Mol Plant-Microbe Interact 9:226–232

Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, Proctor M, St.Onge RP, Tyers M, Koller D, Altman RB, Davis RW, Nislow C, Giaever G (2008) The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science 320:362–365

Hohn TM, Beremand PD (1989) Isolation and nucleotide sequence of a sesquiterpene cyclase gene from the trichothecene-producing fungus *Fusarium sporotrichioides*. Gene 79:131–138

Hohn TM, McCormick SP, Desjardins AE (1993) Evidence for a gene cluster involving trichothecene-pathway biosynthetic genes in *Fusarium sporotrichioides*. Curr Genet 24:291–295

Horevaj P, Bluhm BH (2012) BDM1, a phosducin-like gene of *Fusarium graminearum*, is involved in virulence during infection of wheat and maize. Mol Plant Pathol 13:431–444

Hou ZM, Xue CY, Peng YL, Katan T, Kistler HC, Xu JR (2002) A mitogen-activated protein kinase gene (MGV1) in *Fusarium graminearum* is required for female fertility, heterokaryon formation, and plant infection. Mol Plant-Microbe Interact 15:1119–1127

Houterman PM, Speijer D, Dekker HL, De Koster CG, Cornelissen BJC, Rep M (2007) The mixed xylem sap proteome of *Fusarium oxysporum*-infected tomato plants. Mol Plant Pathol 8:215–221

Houterman PM, Cornelissen BJC, Rep M (2008) Suppression of plant resistance gene-based immunity by a fungal effector. PLoS Pathog 4:e1000061

Houterman PM, Ma L, Van Ooijen G, De Vroomen MJ, Cornelissen BJC, Takken FLW, Rep M (2009) The effector protein Avr2 of the xylem-colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. Plant J 58:970–978

Iida Y, Ohara T, Tsuge T (2006) Identification of genes up-regulated during conidiation of *Fusarium oxysporum* through expressed sequence tag analysis. Fungal Genet Biol 43:179–189

Iida Y, Ohara T, Tsuge T (2007) Identificaiton of genes with changes in transcription levels caused by mutations in conidiation regulator genes *REN1*

and *FoSTUA* in *Fusarium oxysporum*. J Gen Plant Pathol 73:158–167

Iida Y, Kurata T, Harimoto Y, Tsuge T (2008) Nitrite reductase gene upregulated during conidiation is involved in macroconidium formation in *Fusarium oxysporum*. Phytopathology 98:1099–1106

Ikeda K (2010) Role of perithecia as an inoculum source for stem rot type of pepper root rot caused by *Fusarium solani* f. sp. *piperis* (teleomorph: *Nectria haematococca* f. sp. *piperis*). J Gen Plant Pathol 76:241–246

Ilgen P, Maier F, Schäfer W (2008) Trichothecenes and lipases are host-induced and secreted virulence factors of *Fusarium graminearum*. Cereal Res Commun 36:421–428

Jain S, Akiyama K, Mae K, Ohguchi T, Takata R (2002) Targeted disruption of a G protein alpha subunit gene results in reduced pathogenicity in *Fusarium oxysporum*. Curr Genet 41:407–413

Jain S, Akiyama K, Kan T, Ohguchi T, Takata R (2003) The G protein beta subunit FGB1 regulates development and pathogenicity in *Fusarium oxysporum*. Curr Genet 43:79–86

Jain S, Akiyama K, Takata R, Ohguchi T (2005) Signaling via the G protein alpha subunit FGA2 is necessary for pathogenesis in *Fusarium oxysporum*. FEMS Microbiol Lett 243:165–172

Jansen C, von Wettstein D, Schafer W, Kogel K-H, Felk A, Maier FJ (2005) Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthetase gene disrupted *Fusarium graminearum*. Proc Natl Acad Sci U S A 102:16892–16897

Jenczmionka N, Schäfer W (2005) The Gpmk1 MAP kinase of *Fusarium graminearum* regulates the induction of specific secreted enzymes. Curr Genet 47:29–36

Jenczmionka N, Maier F, Lösch A, Schäfer W (2003) Mating, conidiation and pathogenicity of *Fusarium graminearum*, the main causal agent of the head-blight disease of wheat, are regulated by the MAP kinase gpmk1. Curr Genet 43:87–95

Jiang J, Liu X, Yin Y, Ma Z (2011a) Involvement of a velvet protein *FgVeA* in the regulation of asexual development, lipid and secondary metabolisms and virulence in *Fusarium graminearum*. PLoS ONE 6:328291

Jiang J, Yun Y, Yang Q, Shim W-B Wang MaZ (2011b) A type 2C protein phosphatase FgPtc3 is involved in a cell wall integrity, lipid metabolism and virulence in *Fusarium graminearum*. PLoS ONE 6:325311

Jonkers W, Dong Y, Broz K, Corby Kistler H (2012a) The Wor1-like protein Fgp1 regulates pathogenicity, toxin synthesis and reproduction in the phytopathogenic fungus *Fusarium graminearum*. PLoS Pathog 8:e1002724

Jonkers W, Rodriguez Estrada AE, Lee K, Breakspear A, May G, Kistler HC (2012b) Metabolome and transcriptome of the interaction between *Ustilago maydis* and *Fusarium verticillioides in vitro*. Appl Environ Microbiol 78:3656–3667

Kazan K, Gardiner DM, Manners JM (2012) On the trail of a cereal killer: recent advances in *Fusarium graminearum* pathogenomics and host resistance. Mol Plant Pathol 13:399–413

Kidd BN, Kadoo NY, Dombrecht B, Tekeoglu M, Gardiner DM, Thatcher LF, Aitken EAB, Schenk PM, Manners JM, Kazan K (2011) Auxin signaling and transport promote susceptibility to the root-infecting fungal pathogen *Fusarium oxysporum* in *Arabidopsis*. Mol Plant-Microbe Interact 24:733–748

Kim Y-T, Han K-H, Jin J, Kim H, Kim J-C, Yun S-H, Lee Y-W (2005a) Putative polyketide synthetase and laccase genes for biosynthesis of aurofusarin in *Gibberella zeae*. Appl Environ Microbiol 71:1701–1708

Kim Y-T, Lee Y-R, Jin J, Han K-H, Kim H, Kim J-C, Lee T, Yun S-H, Lee Y-W (2005b) Two different polyketide synthetase genes are required for synthesis of zearalenone in *Gibberella zeae*. Mol Microbiol 58:1102–1113

Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, Thomma BPHJ, Chen Z, Henrissat B, Lee Y-H, Park J, Garcia-Pedrajas MC, Barbara DJ, Anchieta A, de Jonge R, Santhanam P, Maruthachalam K, Atallah Z, Amyotte SG, Paz Z, Inderbitzin P, Hayes RJ, Heiman DI, Young S, Zeng QD, Engels R, Galagan J, Cuomo CA, Dobinson KF, Ma LJ (2011) Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathog 7:e1002137

Kpodo KA, Bankole SA (2008) Mycotoxin contamination in foods in West and Central Africa. In: Leslie JF, Bandyopadhyay R, Visconti A (eds) Mycotoxins: detection methods, management, public health and agricultural trade. CABI, Wallingford, pp 103–116

Lee T, Han Y-K, Kim K-H, Yun S-H, Lee Y-W (2002) *Tri13* and *Tri7* determine deoxynivalenol- and nivalenol-producing chemotypes of *Gibberella zeae*. Appl Environ Microbiol 68:2148–2154

Lee J, Lee T, Lee Y-W, Yun S-H, Turgeon BG (2003) Shifting fungal reproductive mode by manipulation of mating type genes: obligatory heterothallism of *Gibberella zeae*. Mol Microbiol 50:145–152

Lee S-H, Lee S, Choi D, Lee Y-W, Yun S-H (2006) Identifiction of the down-regulated genes in a mat1-2-deleted strain of *Gibberella zeae*, using cDNA subtraction and microarray analysis. Fungal Genet Biol 43:295–310

Lee J, Myong K, Kim J-E, Kim H-K, Yun S-H, Lee Y-W (2012) FgVelB globally regulates sexual reproduction, mycotoxin production and pathogenicity in the cereal pathogen *Fusarium graminearum*. Microbiology 158:1723–1733

Lepage P, Leclerc MC, Joossens M, Mondot S, Blottière HM, Raes J, Ehrlich D, Doré J (2013) A metagenomic insight into our gut's microbiome. Gut 62:146–158

Leslie JF, Summerell B (2006) The *Fusarium* laboratory manual. Blackwell, Ames

Li S, Myung K, Guse D, Donkin B, Proctor RH, Grayburn WS, Calvo AM (2006) *FvVE1* regulated filamentous growth, the ratio of microconidia to macroconidia and cell wall formation in *Fusarium verticillioides*. Mol Microbiol 62:1418–1432

Löffler HJM, Mouris JR (1992) Fusaric acid: phytotoxicity and in vitro production by *Fusarium oxysporum* f.sp. *lilii*, the causal agent of basal rot in lilies. Neth J Plant Pathol 98:107–115

López-Berges MS, Hera C, Sulyok M, Schäfer K, Capilla J, Guarro J, Di Pietro A (2013) The velvet complex governs mycotoxin production and virulence of *Fusarium oxysporum* on plant and mammalian hosts. Mol Microbiol 87:49–65

Lund ST, Stall RE, Klee HJ (1998) Ethylene regulates the susceptible response to pathogen infection in tomato. Plant Cell 10:371–382

Lysøe E, Klemsdal SS, Bone KR, Frandsen RJN, Johansen T, Thrane U, Giese H (2006) The *PKS4* gene of *Fusarium graminearum* is essential for zearalenone production. Appl Environ Microbiol 72:3924–3932

Lysøe E, Seong K-Y, Kistler HC (2011a) The transcriptome of *Fusarium graminearum* during the infection of wheat. Mol Plant-Microbe Interact 24:995–1000

Lysøe E, Pasquali M, Breakspear A, Kistler HC (2011b) The transcription factor FgStuAp influences spore development, pathogenicity, and secondary metabolism in *Fusarium graminearum*. Mol Plant-Microbe Interact 24:54–67

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, Woloshuk C, Xie XH, Xu JR, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee YH, Li LD, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng QD, Zhou SG, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373

Maheshwari R (1999) Microconidia of *Neurospora crassa*. Fungal Genet Biol 26:1–18

Malonek S, Bomke C, Bornberg-Bauer E, Rojas MC, Hedden P, Hopkins P, Tudzynski B (2005) Distribution of gibberellin biosynthetic genes and gibberellin production in the *Gibberella fujikuroi* species complex. Phytochemistry 66:1296–1311

Malz S, Grell MN, Trane C, Maier FJ, Rosager P, Felk A, Albertsen KS, Salomon S, Bohn L, Schäfer W, Giese H (2005) Identification of a gene cluster responsible for the biosynthesis of aurofusarin in the *Fusarium graminearum* species complex. Fungal Genet Biol 42:420–433

Marin S, Magan N, Ramos AJ, Sanchis V (2004) Fumonisin-producing strains of *Fusarium*: a review of their ecophysiology. J Food Prot 67:1792–1805

McCormick SP, Harris LJ, Alexander NJ, Ouellet T, Saparno A, Allard S, Desjardins AE (2004) *Tri1* in *Fusarium graminearum* encodes a P450 oxygenase. Appl Environ Microbiol 70:2044–2051

McCormick SP, Alexander NJ, Harris LJ (2010) *CLM1* of *Fusarium graminearum* encodes a longiborneol synthetase required for culmorin production. Appl Environ Microbiol 76:136–141

McMullen M, Bergstrom G, DeWolf E, Dill-Macky R, Hershman D, Shaner G, Van Sanford D (2012) A unified effort to fight an enemy of wheat and barley: *Fusarium* head blight. Plant Dis 96:1712–1728

Medentsev AG, Kotik AN, Trufanova VA, Akimenko VK (1993) Identification of aurofusarin in *Fusarium graminearum* isolates, causing a syndrome of worsening of egg quality in chickens. Prikl Biokhim Mikrobiol 29:542–546

Merhej J, Urban M, Dufresne M, Hammond-Kosack KE, Richard-Forget F, Barreau C (2012) The velvet gene, FgVe1, affects fungal development and positively regulates trichothecene biosynthesis and pathogenicity in *Fusarium graminearum*. Mol Plant Pathol 13:363–374

Metzenberg RL, Glass NL (1990) Mating type and mating strategies in *Neurospora*. Bioessays 12:53–59

Miao VPW, Vanetten HD (1992) Three genes for metabolism of the phytoalexin maackiain in the plant pathogen *Nectria haematococca*: meiotic instability and relationship to a new gene for pisatin demethylase. Appl Environ Microbiol 58:801–808

Michielse CB, Rep M (2009) Pathogen profile update: *Fusarium oxysporum*. Mol Plant Pathol 10:311–324

Michielse CB, Reijnen L, Olivain C, Alabouvette C, Rep M (2012) Degradation of aromatic compounds through the β-ketoadipate pathway is required for pathogenicity of the tomato wilt pathogen *Fusarium oxysporum* f. sp. *lycopersici*. Mol Plant Pathol 13:1089–1100

Miedaner T, Cumagun CJR, Chakraborty C (2008) Population genetics of three important head blight pathogens *Fusarium graminearum*, *F. pseudograminearum*, and *F. culmorum*. J Phytopathol 156:129–139

Miersch O, Bohlmann H, Wasternack C (1999) Jasmonates and related compounds from *Fusarium oxysporum*. Phytochemistry 50:517–523

Milani NA, Lawrence DP, Elizabeth Arnold A, VanEtten HD (2012) Origin of pisatin demethylase (PDA) in the genus *Fusarium*. Fungal Genet Biol 49:933–942

Mitter V, Francl L, Ali S, Simpfendorfer S, Chakraborty S (2006) Ascosporic and conidial inoculum of *Gibberella zeae* play different roles in *Fusarium* head blight and crown rot of wheat in Australia and the USA. Australas Plant Pathol 35:441–452

Morgan XC, Segata N, Huttenhower C (2013) Biodiversity and functional genomics in the human microbiome. Trends Genet 29:51–58

Mukherjee M, Kim J-E, Park Y-S, Kolomiets MV, Shim W-B (2011) Regulators of G-protein signaling in *Fusarium verticillioides* mediate differential host-pathogen responses on nonviable versus viable maize kernels. Mol Plant Pathol 12:479–491

Myung K, Zitomer NC, Duvall M, Glenn AE, Riley RT, Calvo AM (2012) The conserved global regulator VeA is necessary for symptom production and mycotoxin synthesis in maize seedlings by *Fusarium verticillioides*. Plant Pathol 61:152–160

Nelson RR (1971) Hormonal involvement in sexual reproduction in the fungi with special reference to F-2, a fungal estrogen. In: Adai S, Ouchi S (eds) Morphological and biochemical events in plant-parasite interaction. Phytopathological Society of Japan, Tokyo, pp 181–200

Niemeyer HM (1988) Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defence chemicals in the gramineae. Phytochemistry 27:3349–3358

Nishiuchi T, Masuda D, Nakashita H, Ichimura K, Shinozaki K, Yoshida S, Kimura M, Yamaguchi I, Yamaguchi K (2006) *Fusarium* phytotoxin trichothecenes have an elicitor-like activity in *Arabidopsis thaliana*, but the activity differed significantly among their molecular species. Mol Plant-Microbe Interact 19:512–520

O'Donnell K, Kistler HC, Tacke BK, Casper HH (2000) Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. Proc Natl Acad Sci U S A 97:7905–7910

O'Donnell K, Rooney AP, Proctor RH, Brown DW, McCormick SP, Ward TJ, Frandsen RJN, Lysøe E, Rehner SA, Aoki T, Robert VARG, Crous PW, Groenewald JZ, Kang S, Geiser DM (2013) Phylogenetic analyses of RPB1 and RPB2 support a middle Cretaceous origin for a clade comprising all agriculturally and medically important fusaria. Fungal Genet Biol 52:20–31

Ohara T, Inoue I, Namiki F, Kunoh H, Tsuge T (2004) *REN1* is required for development of microconidia and macroconidia, but not of chlamydospores, in the plant pathogenic fungus *Fusarium oxysporum*. Genetics 166:113–124

Oide S, Krasnoff SB, Gibson DM, Turgeon BG (2007) Intracellular siderophores are essential for ascomycete sexual development in heterothallic *Cochliobolus heterostrophus* and homothallic *Gibberella zeae*. Eukaryot Cell 6:1339–1353

Pandit A, Maheshwari R (1996) Life history of *Neurospora intermedia* in a sugar cane field. J Biosci 21:57–79

Paper JM (2011) Secreted proteins from the plant pathogenic fungus *Fusarium graminearum*. Ph.D. thesis, Michigan State University, p 171

Paper JM, Scott-Craig JS, Adhikari ND, Cuomo CA, Walton JD (2007) Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus *Fusarium graminearum*. Proteomics 7:3171–3183

Pareja-Jaime Y, Roncero MIG, Ruiz-Roldn MC (2008) Tomatinase from *Fusarium oxysporum* f. sp. *lycopersici* is required for full virulence on tomato plants. Mol Plant-Microbe Interact 21:728–736

Park H-S, Yu JH (2012) Genetic control of asexual sporulation in filamentous fungi. Curr Opin Microbiol 15:669–677, Epub before print

Park AR, Cho A-R, Seo J-A, Min K, Son H, Lee J, Choi GJ, Kim J-C, Lee Y-W (2012) Functional analyses of regulators of G-protein signaling in *Gibberella zeae*. Fungal Genet Biol 49:511–520

Paz Z, García-Pedrajas MD, Anderws DL, Klosterman S, Baeza-Montañez GSE (2011) One Step Construction of *Agrobacterium*-Recombinant-ready-plasmids (OSCAR), an efficient and robust tool for ATMT based gene deletion construction in fungi. Fungal Genet Biol 48:677–684

Peplow AW, Meek IB, Wiles MC, Phillips TD, Beremand MN (2003) Tri16 is required for esterification of position C-8 during trichothecene mycotoxin production by *Fusarium sporotrichioides*. Appl Environ Microbiol 69:5935–5940

Peraldi A, Beccari G, Steed A, Nicholson P (2011) *Brachypodium distachyon*: a new pathosystem to study *Fusarium* head blight and other *Fusarium* diseases of wheat. BMC Plant Biol 11:100

Pestka JJ (2007) Deoxynivalenol: toxicity, mechanisms and animal health risks. Anim Feed Sci Technol 137:283–298

Prados-Rosales RC, Serena C, Delgado-Jarana J, Guarro J, Di Pietro A (2006) Distinct signalling pathways coordinately contribute to virulence of *Fusarium oxysporum* on mammalian hosts. Microb Infect 8:2825–2831

Proctor RH, Hohn TM, McCormick SP (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene. Mol Plant-Microbe Interact 8:593–601

Proctor RH, Desjardins AE, Plattner RD, Hohn TM (1999) A polyketide synthetase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. Fungal Genet Biol 27:100–112

Proctor RH, Brown DW, Plattner RD, Desjardins AE (2003) Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. Fungal Genet Biol 38:237–249

Proctor RH, Butchko RAE, Brown DW, Moretti A (2007) Functional characterization, sequence comparisons and distribution of a polyketide synthetase gene required for perithecial pigmentation in some *Fusarium* species. Food Addit Contam 24:1076–1087

Proctor RH, Busman M, Seo J-A, Lee YW, Plattner RD (2008) A fumonisin biosynthetic gene cluster in *Fusarium oxysporum* strain O-1890 and the genetic basis for B versus C fumonisin production. Fungal Genet Biol 45:1016–1026

Proctor RH, McCormick SP, Alexander NJ, Desjardins AE (2009) Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus *Fusarium*. Mol Microbiol 74:1128–1142

Rep M, van der Does HC, Meijer M, van Wijk R, Houterman PM, Dekker HL, de Koster CG, Cornelissen BJC (2004) A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for I-3-mediated resistance in tomato. Mol Microbiol 53:1373–1383

Rep M, Meijer M, Houterman PM, van der Does HC, Cornelissen BJC (2005) *Fusarium oxysporum* evades I-3-mediated resistance without altering the matching avirulence gene. Mol Plant-Microbe Interact 18:15–23

Rodriguez-Carres M, White G, Tsuchiya D, Taga M, VanEtten HD (2008) The supernumerary chromosome of *Nectria haematococca* that carries pea-pathogenicity-related genes also carries a trait for pea rhizosphere competitiveness. Appl Environ Microbiol 74:3849–3856

Sagaram US, Shim W-B (2007) *Fusarium verticillioides GBB1*, a gene encoding heterotrimeric G protein β subunit, is associated with fumonisin B1 biosynthesis and hyphal development but not with fungal virulence. Mol Plant Pathol 8:375–384

Salomon S, Gcser A, Frerichmann S, Kröger C, Schäfer W, Voigt C (2012) The secreted lipase FGL1 is sufficient to restore the initial infection step to the apathogenic *Fusarium graminearum* MAP kinase disruption mutant Δgpmk1. Eur J Plant Pathol 134:23–37

Scherm B, Balmas V, Spanu F, Pani G, Delogu G, Pasquali M, Migheli Q (2013) *Fusarium culmorum*: causal agent of foot and root rot and head blight on wheat. Mol Plant Pathol 14:323–341

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du FY, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen WZ, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He RF, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin JK, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren LY, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han YJ, Lee H, Li PH, Lisch DR, Liu SZ, Liu ZJ, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang LX, Yu Y, Zhang LF, Zhou SG, Zhu Q, Bennetzen JL, Dawe RK, Jiang JM, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Seong K-Y, Zhao X, Xu J-R, Güldener U, Kistler HC (2008) Conidial germination in the filamentous fungus *Fusarium graminearum*. Fungal Genet Biol 45:389–399

Seong K-Y, Pasquali M, Zhou X, Song J, Hilburn K, McCormick S, Dong Y, Xu J-R, Kistler HC (2009) Global gene regulation by *Fusarium* transcription factors *Tri6* and *Tri10* reveals adaptations for toxin biosynthesis. Mol Microbiol 72:354–367

Shaner GE (2003) Epidemiology of *Fusarium* head blight of small grain cereals in North America. In: Leonard KJ, Bushnell WR (eds) *Fusarium* head blight of wheat and barley. American Phytopathological Society Press, St. Paul, pp 84–119

Short DPG, O'Donnell K, Zhang N, Juba JH, Geiser DM (2011) Widespread occurrence of diverse pathogenic types of the fungus *Fusarium* in bathroom plumbing drains. J Clin Microbiol 49:4264–4272

Siame BA, Nawa IN (2008) Mycotoxin contamination in food systems in Eastern and Southern Africa. In: Leslie JF, Bandyopadhyay R, Visconti A (eds) Mycotoxins: detection methods, management, public health and agricultural trade. CABI, Wallingford, pp 117–125

Sikhakolli UR, López-Girldez F, Li N, Common R, Townsend JP, Trail F (2012) Transcriptome analyses during fruiting body formation in *Fusarium graminearum* and *F. verticillioides* reflect species

life history and ecology. Fungal Genet Biol 49:663–673

Smýkal P, Aubert G, Burstin J, Coyne CJ, Ellis NTH, Flavell AJ, Ford R, Hýbl M, Macas J, Neumann P, McPhee KE, Redden RJ, Rubiales D, Weller JL, Warkentin TC (2012) Pea (*Pisum sativum* L.) in the genomic era. Agronomy 2:74–115

Son H, Seo Y-S, Min K, Park AR, Lee J, Jin J-M, Lin Y, Cao P, Hong S-Y, Kim E-K, Lee S-H, Cho A, Lee S, Kim M-G, Kim Y, Kim J-E, Kim J-C, Choi GJ, Yun S-H, Lim JY, Kim M, Lee Y-H, Choi Y-D, Lee Y-W (2011) A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog 7: e1002310

Springer ML, Yanofsky C (1992) Expression of *con* genes along the three sporulation pathways of *Neurospora crassa*. Genes Dev 6:1052–1057

Starkey DE, Ward TJ, Aoki T, Gale LR, Kistler HC, Geiser DM, Suga H, Tóth B, Varga J, O'Donnell K (2007) Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity. Fungal Genet Biol 44:1191–1204

Staskawicz BJ, Mudgett MB, Dangl JL, Galan JE (2001) Common and contrasting themes of plant and animal diseases. Science 292:2285–2289

Stephens AE, Gardiner DM, White RG, Munn AL, Manners JM (2008) Phases of infection and gene expression of *Fusarium graminearum* during crown rot disease of wheat. Mol Plant-Microbe Interact 21:1571–1581

Studt L, Wiemann P, Kleigrewe K, Humpf H-U, Tudzynski B (2012) Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. Appl Environ Microbiol 78:4468–4480

Suga H, Karugia GW, Ward T, Gale LR, Tomimura K, Nakajima T, Miyasaka A, Koizumi S, Kageyama K, Hyakumachi M (2008) Molecular characterization of the *Fusarium graminearum* species complex in Japan. Phytopathology 98:159–166

Tatum JH, Baker RA, Berry RE (1987) Naphthofurans produced by *Fusarium oxysporum* isolated from citrus. Phytochemistry 26:2499–2500

Temporini E, VanEtten H (2004) An analysis of the phylogenetic distribution of the pea pathogenicity genes of *Nectria haematococca* MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: *Neocosmospora boniensis*. Curr Genet 46:29–36

Thatcher LF, Manners JM, Kazan K (2009) *Fusarium oxysporum* hijacks COI1-mediated jasmonate signaling to promote disease development in *Arabidopsis*. Plant J 58:927–939

Thatcher LF, Gardiner DM, Kazan K, Manners JM (2012) A highly conserved effector in *Fusarium oxysporum* is required for full virulence on *Arabidopsis*. Mol Plant-Microbe Interact 25:180–190

Tobiasen C, Aahman J, Ravnholt KS, Bjerrum MJ, Grell MN, Giese H (2007) Nonribosomal peptide synthetase (NPS) genes in *Fusarium graminearum*, *F. culmorum* and *F. pseudograminearium* and identification of NPS2 as the producer of ferricrocin. Curr Genet 51:43–58

Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Trail F (2013) Sex and fruiting in Fusarium. In: Brown DW, Proctor RH (eds) Fusarium: genomics, molecular and cellular biology. Horizon Scientific Press, Norwich, pp 11–29

Tsavkelova E, Oeser B, Oren-Young L, Israeli M, Sasson Y, Tudzynski B, Sharon A (2012) Identification and functional characterization of indole-3-acetamide-mediated IAA biosynthesis in plant-associated *Fusarium* species. Fungal Genet Biol 49:48–57

Tudzynski B, Hölter K (1998) Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: evidence for a gene cluster. Fungal Genet Biol 25:157–170

Urban M, Mott E, Farley T, Hammond-Kosack K (2003) The *Fusarium graminearum* MAP1 gene is essential for pathogenicity and development of perithecia. Mol Plant Pathol 4:347–359

Voigt CA, Schäfer W, Salomon S (2005) A secreted lipase of *Fusarium graminearum* is a virulence factor required for infection of cereals. Plant J 42:364–375

Waalwijk C, Kastelein P, deVries I, Kereny Z, van der Lee T, Hesselink T, Kohl J, Kema G (2003) Major changes in *Fusarium* spp. in wheat in the Netherlands. Eur J Plant Pathol 109:743–754

Waalwijk C, van der Lee T, de Vries I, Hesselink T, Arts J, Kema GHJ (2004) Synteny in toxigenic *Fusarium* species: the fumonisin gene cluster and the mating type region as examples. Eur J Plant Pathol 110:533–544

Walter S, Nicholson P, Doohan FM (2010) Action and reaction of host and pathogen during *Fusarium* head blight disease. New Phytol 185:54–66

Wang Q, Xu L (2012) Beauvericin, a bioactive compound produced by fungi: a short review. Molecules 17:2367–2377

Wang C, Zhang S, Hou R, Zhao Z, Zheng Q, Xu Q, Zheng D, Wang G, Liu H, Ma J-W, Kistler HC, Kang ZS, Xu JR (2011) Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. PLoS Pathog 7:e1002460

Ward TJ, Bielawski JP, Kistler HC, Sullivan E, O'Donnell K (2002) Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. Proc Natl Acad Sci U S A 99:9278–9283

Ward TJ, Clear RM, Rooney AP, O'Donnell D, Gaba D, Patrick S, Starkey DE, Gilbert J, Geiser DM, Nowicki TW (2008) An adaptive evolutionary shift in *Fusarium* head blight pathogen populations is

driving the rapid spread of more toxigenic *Fusarium graminearum* in North America. Fungal Genet Biol 45:473–484

Wassman CC, VanEtten HD (1996) Transformation-mediated chromosome loss and disruption of a gene for pisatin demethylase decrease the virulence of *Nectria haematococca* on pea. Mol Plant-Microbe Interact 9:793–803

Wiemann P, Willmann A, Straeten M, Kleigrewe K, Beyer M, Humpf HU, Tudzynski B (2009) Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. Mol Microbiol 72:931–946

Wiemann P, Brown DW, Kleigrewe K, Bok JW, Keller NP, Humpf HU, Tudzynski B (2010) FfVel1 and FfLae1, components of a *velvet*-like complex in *Fusarium fujikuroi*, affect differentiation, secondary metabolism and virulence. Mol Microbiol 77:972–994

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Cu AM, Connelly C, Davis K, Dietrich F, Dow SW, EL Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles I, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Veronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu KX, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285:901–906

Wolf JC, Mirocha CJ (1973) Regulation of sexual reproduction in *Gibberella zeae* (*Fusarium roseum graminearum*) by F-2 (zearalenone). Can J Microbiol 19:725–734

Xu X, Nicholson P (2009) Community ecology of fungal pathogens causing wheat head blight. Annu Rev Phytopathol 47:83–103

Yang L, van der Lee T, Yang X, Yu D, Waalwijk C (2008) *Fusarium* populations on Chinese barley show dramatic gradient in mycotoxin profiles. Phytopathology 98:719–727

Yang FEN, Jensen JD, Svensson B, Jørgensen HJL, Collinge DB, Finnie C (2012) Secretomics identifies *Fusarium graminearum* proteins involved in the interaction with barley and wheat. Mol Plant Pathol 13:445–453

Yu H-Y, Seo J-A, Kim J-E, Han K-H, Shim W-B, Yun S-H, Lee Y-W (2008) Functional analyses of heterotrimeric G protein Gα and Gβ subunits in *Gibberella zeae*. Microbiology 154:392–401

Yun S-H, Arie T, Kaneko I, Yoder OC, Turgeon BG (2000) Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. Fungal Genet Biol 31:7–20

Zhang H, Zhang Z, van der Lee T, Chen WQ, Xu J, Xu JS, Yang L, Yu D, Waalwijk C, Feng C (2010) Population genetic analyses of *Fusarium asiaticum* populations from barley suggest a recent shift favoring 3ADON producers in southern China. Phytopathology 100:328–336

Zhang X-W, Jia L-J, Zhang Y, Jiang G, Li X, Zhang D, Tang W-H (2012) *In planta* stage-specific fungal gene profiling elucidates the molecular strategies of *Fusarium graminearum* growing inside wheat coleoptiles. Plant Cell 24:5159–5176

Zheng W, Zhao X, Xie Q, Huang Q, Zhang C, Zhai H, Xu L, Lu G, Shim W-B, Wang Z (2012) A conserved homeobox transcription factor Htf1 is required for phialide development and conidiogenesis in *Fusarium* species. PLoS ONE 7:e45432

12 Metabolomics and Proteomics to Dissect Fungal Phytopathogenicity

Kar-Chun Tan[1], Richard P. Oliver[1]

CONTENTS

I. Introduction

Since 2004, improvements in DNA-sequencing technology have resulted in a 40,000-fold increase in sequencing throughput capacities. As a consequence, the overall costs of DNA sequencing have been dramatically reduced (Mardis 2006; Stahl and Lundeberg 2012). Many laboratories now have the means to carry out high-throughput genome sequencing of their organism of choice. In phytopathology, this revolution has produced a plethora of sequenced fungal genomes with the aim of attaining a deeper understanding of phytopathogenicity at the biochemical level (Egan and Talbot 2008; Schmidt and Panstruga 2011). Prior to the genome era, the identification of genes for phytopathogenicity studies was a comparatively laborious process that relied on the generation of transcript data (Talbot et al. 1993), cross-species gene homology (Solomon et al. 2004) and random mutagenesis (Sweigard et al. 1998). The availability of genome data has greatly assisted in the identification of candidate genes for large-scale pathogenicity studies (Wang et al. 2011a). Furthermore, the availability of sequenced genomes has opened the door for the dissection of phytopathogenicity via systems biology using sophisticated high-throughput technologies. Genome microarrays are a popular platform commonly used to provide a physiological overview of a biological system by measuring the transcript abundance of annotated genes (Güldener et al. 2006; Ipcho et al. 2012). However, a number of studies have observed that the correlation between transcript and protein abundances is poor (Gygi et al. 1999; Tan et al. 2008). This may be attributed to translational lag, protein/transcript stability and posttranscriptional processing (Mann and Jensen 2003; Kozak 2005).

The availability of genome data has also enabled high-throughput analysis of the proteome, which can be defined as the protein complement within a biological system (Wasinger et al. 1995; Zhu et al. 2003). Proteomics is a field that attempts to understand the proteome. Outputs from proteomic studies can provide key metabolic information, such as posttranslational modifications (PTMs; Mann and Jensen

[1]Australian Centre for Necrotrophic Fungal Pathogens, Department of Environment and Agriculture, Curtin University, Bentley, WA 6102, Australia; e-mail: Richard.Oliver@curtin.edu.au

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

2003) and subcellular localisation of proteins in fungal pathogens (Tan et al. 2008, 2009a). Metabolites are intermediates or products of metabolic reactions. Metabolomics is a field that attempts to identify and quantify the complement of low molecular weight metabolites, or metabolome, present in a biological system (Oliver et al. 1998). The metabolome is directly related to the phenotype of the fungus as it is the end product of gene expression. When used in a complementary fashion, these '-omics' techniques can provide a detailed biochemical snapshot of fungal phytopathogens and assist in identifying pathogenicity genes (Van De Wouw and Howlett 2011). In 2009, we presented a comprehensive review on the impact of -omics technology on phytopathogenicity (Tan et al. 2009b). This chapter presents an update of recent literature that significantly contributed to the current impact and future outlook of proteomics and metabolomics on fungal phytopathology.

II. Proteomics

A. Proteomic Methodologies

Two-dimensional gel electrophoresis (2D-E) has long been a method of choice for separating complex protein mixtures because of its high resolving power. A typical 2D-E procedure involves the separation of a complex protein mixture by isoelectric focusing in the first dimension followed by molecular mass in the second. Following this, proteins within the gel are visualised using stains such as Coomassie, silver and SYPRO Ruby. Quantification and comparative analysis of protein spots between treatments are performed with commercially available densitometry analysis software such as Progenesis and PDQuest. However, 2D-E has been limited by its ability to separate proteins of extreme pI/molecular mass (Issaq and Veenstra 2008). Furthermore, limitations in current staining technology meant that low-abundant proteins are rarely seen on gel (Herbert et al. 2001).

Proteomics based on two-dimensional **gel-free liquid chromatography** (2D-LC) was developed to overcome limitations that are associated with the gel-based method and to provide an unbiased approach in proteome-wide profiling (Washburn et al. 2001). 2D-LC involves the separation of peptides derived from a crude digested protein mixture with strong cation exchange, followed by reverse-phase chromatography prior to mass spectrometric (MS) analysis (Washburn et al. 2001). In addition to proteome profiling, quantitative proteomics can be performed using 2D-LC via **isobaric labelling strategies**, such as isobaric tag for relative and absolute quantification (iTRAQ) (Ross et al. 2004). With iTRAQ, trypsin-digested peptides of different treatments or samples are derivatised with amine-based isotope-coded tags. Individual labelled samples are then pooled and subjected to 2D-LC prior to MS. The generation of reporter ions from attached tags via tandem MS (MS/MS) facilitates peptide identification and quantification. Matrix-assisted laser desorption/ionisation (MALDI) and LC coupled to electrospray MS or MS/MS are commonly used peptide analytical platforms. **Gene identification** through pattern matching of the mass spectra to a translated protein database is facilitated by specialised pattern-matching software such as Mascot and Sequest. Alternatively, peptides can be directly mapped onto the genome for gene annotation purposes using tools such as Cdsmapper (https://sourceforge.net/projects/cdsmapper/files/) (Bringans et al. 2009). Protein identification for 2D-E follows a similar workflow. Target protein spots are excised from the gel, subjected to in-gel digestion, and identification is usually achieved by MS as described previously. Detailed experimental protocols for 2D-E, 2D-LC and fungal subproteome extractions are available (Pitarch et al. 2008; Vincent et al. 2009, 2012b). These will provide an excellent reference source in carrying out fungal proteomics.

B. Identification of Signalling Pathway Targets

The role of signal transduction in fungal phytopathogenicity has been the subject of intense investigations in the last decade, for example,

mitogen-activated protein kinase (MAPK), cyclic AMP (cAMP) and two-component signalling pathways. In most cases, perturbation of these pathways resulted in a wide range of phenotypic responses, such as sporulation, development, mating, host penetration, mycotoxin production and virulence on the host plant. A key question arising is what these pathways regulate that is so critical in fungal development and virulence.

Stagonospora nodorum is a major cereal pathogen and causes *S. nodorum* blotch of wheat. Genome sequencing has identified components of the cAMP pathway, such as the **heterotrimeric G-proteins** (Hane et al. 2007). Of these, *Gna1* has been shown to play a key role in *S. nodorum* virulence (Solomon et al. 2004). *Gna1* encodes a Gα inhibitory class subunit and is required for sporulation, development and pathogenicity (Solomon et al. 2004). 2D-E was previously used to identify intracellular targets of cAMP signalling. Six differentially abundant protein spots were identified from 475 unique protein spots. One of these targets is *Sch1* involved in sporulation and mycotoxin production (Tan et al. 2008, 2009a, 2009c). Casey et al. (2010) complemented the 2D-E study with a 2D-LC MALDI-time-of-flight (TOF) MS/MS comparative analysis of wild-type and *gna1* intracellular proteomes via ITRAQ labelling. The study identified 1,336 proteins common to both strains. Of these, 49 showed significant changes in abundance, including Sch1. Analysis of these differentially abundant proteins revealed that the pentose phosphate and mannitol metabolism pathways were perturbed in *gna1*. In addition, the study identified short-chain dehydrogenase, Sch3, downregulated in *gna1* to play a role in asexual sporulation.

C. Response to Nutrition and Plant Elicitors

Fungi will encounter a dynamic environment during infection when nutrient sources will vary greatly (Solomon et al. 2003). Nutritional status can influence the expression of key fungal pathogenicity determinants (Brito et al. 2006; Thomma et al. 2005). The explosion of fungal genome sequences and the refinement of proteomic tools provide a comprehensive opportunity to understand biochemical pathways that regulate how pathogenic fungi respond to their environment and are critical in developing strategies to combat plant diseases.

Much research has focussed on the identification of effectors - pathogen molecules that produce a specific effect on the host and affect the outcome of the interaction. *Botrytis cinerea* causes grey mould disease on a broad range of agricultural crops. The fungus is a versatile necrotroph that utilises secreted depolymerases such as polygalacturonase (*Bcpg1*) and endo-β-1,4-xylanase (*Xyn11A*) for pathogenicity (ten Have et al. 1998; Brito et al. 2006). *Botrytis cinerea* possesses a sequenced genome and hence is an attractive system for the identification of secreted pathogenicity factors through the use of proteomics. 2D-E MALDI-TOF MS/MS was used to profile the **secreted proteome** of *B. cinerea* from growth on a variety of carbon sources commonly found during infection (Fernandez-Acero et al. 2010). A number of hydrolases, such as pectin methyl esterases, xylanases and proteases, were identified. Some of these molecules were previously implicated in fungal virulence. The secreted proteome of *B. cinerea* was also examined in detail using 1D-SDS-PAGE after growth on pectin. MS analysis of protein bands identified nine pectin-induced pectinases that included polygalacturonases, pectate lyases and methylesterases. These proteins were not observed after growth on sucrose (Shah et al. 2009b).

It is likely that *B. cinerea* will encounter a hydrophobic cutin layer on the host surface prior to penetration. To mimic this condition for secreted proteomics, Shah et al. (2009a) grew *B. cinerea* on a hydrophobic cellophane membrane-placed media containing strawberry, tomato or *Arabidopsis* extracts. The secreted proteome was trypsinated and directly analysed by LC-MS/MS, leading to the identification of 89 proteins. Of these, only seven were common to all nutritional treatments. These included two pectin methyl esterases, β-glucanase, glycosidate, ceratoplatanin and two unknown proteins. Spectral count analysis indicated differential protein abundance across

all growth conditions. In a similar study, Espino et al. (2010) exposed *B. cinerea* to glucose and plant elicitors (tomato and kiwifruit extracts) and collected the secreted proteome after 16 h of growth to reflect the beginning of plant infection. The secreted proteome was subjected to 2D-E MALDI-MS and 1D-E LC-MS/MS analyses. There were 105 proteins identified via the *B. cinerea* genome from all treatments. When supplemented with glucose, proteins that are associated with pectin degradation and proteolysis accounted for 7 % and 38 % abundances, respectively. These abundances were significantly larger when supplemented with plant elicitors. Proteins that are associated with hemicellulose degradation were only identified from *B. cinerea* supplemented with tomato extracts.

The soluble mycelial proteome of *B. cinerea* was also profiled from growth on cellulose via 2D-E MALDI-TOF MS/MS, which identified 303 unique proteins from 276 protein spots (Fernandez-Acero et al. 2009). Protein analysis through evolutionary relationships (PANTHER) classification revealed that proteins associated with carbohydrate, protein and amino acid metabolism made up 45 % of the unique identifications.

Different plant tissues differ greatly in pH (Manteau et al. 2003). Recently, Li et al. (2012) examined the effect of pH on the secreted proteome profile of *B. cinerea* using 2D-E MALDI-TOF MS/MS. At acidic pH, proteolytic enzymes predominated, whereas under near-neutral pH conditions, cell-wall-degrading enzyme predominated. Transcript analysis suggested that these genes are regulated at the transcriptional level.

Thielaviopsis basicola is a host-specific pathogen that causes black root rot disease in cotton. To identify metabolic components associated with host specificity, the fungus was grown on extracts from susceptible host and non-host plants to study the mycelial proteome (Coumans et al. 2010). There were 50 differentially abundant protein spots identified from 750 protein spots. Despite not possessing a sequenced genome, cross-species database searching and de novo sequencing positively identified 41 protein spots. The majority of the identified proteins are associated with primary metabolism. Interestingly, several proteins associated with vitamin B synthesis were more abundant in the presence of the non-host extract. The biological reason behind this is unclear.

The rice blast fungus *Magnaporthe oryzae* (syn. *grisea*) uses appressoria to penetrate the host plant and establish infection. It is highly likely that the fungus encounters a nutrient-deprived environment prior to host entry. Nitrogen is a key nutrient source that is required for many key metabolic processes. Microarray analysis of the *M. oryzae* transcriptome showed massive changes in the gene expression profile during **nitrogen starvation** (Donofrio et al. 2006). To mimic the prepenetration phase of *M. oryzae*, Wang et al. (2011b) examined the secreted proteome of the fungus in response to nitrogen starvation using 2D-E MALDI-TOF MS/MS and LC-MS/MS. The study identified 89 differently abundant proteins spots. Interestingly, the majority of these protein spots were more abundant during nitrogen starvation. Proteins that are associated with reactive oxygen species (ROS) detoxification, cell wall modification, proteo- and lipidolysis were predominantly identified in upregulated protein spots. Analysis of the *M. oryzae* secretome produced under in vitro conditions that mimicked early infection, such as germination on glass, polyvinylidene fluoride (PVDF) membrane and liquid medium, have yielded a similar finding (Jung et al. 2012).

Mycotoxins are toxic secondary metabolites produced by fungi that can potentially contaminate food sources. This poses a significant safety risk and a health threat. The production of mycotoxins is a process tightly regulated by nutritional factors such as nitrogen availability (Yu and Keller 2005). The maize pathogen *Fusarium verticilloides* is a renowned producer of fumonisin B_1, a carcinogenic polyketide-based mycotoxin. The biosynthesis of fumonisin B_1 is subjected to nitrogen catabolite repression; however, the regulatory mechanism involved in its synthesis is complex and not fully deciphered. 2D-E was used to examine the intracellular proteome of *Gibberella moniliformis* under a nitrogen-limiting condition (Choi et al. 2012). MS analysis identified seven differentially abundant proteins, such as a putative nitroreductase that was implicated as a

negative regulator of fatty acid metabolism. Acetyl-CoA (coenzyme A) is a derivative of fatty acid metabolism and is a substrate for the biosynthesis of polyketide-based compounds.

Fusarium graminearum is a serious pathogen that causes fusarium head blight of cereals and is notoriously known as a producer of the trichothecene mycotoxin deoxynivalenol (DON). DON production is induced by a number of nutritional factors, including nitrogen starvation. It has been shown that phosphorylation events are linked to the activation of the DON biosynthesis pathway (Rampitsch et al. 2010). However, detailed knowledge of signalling events that govern this is limited. Rampitsch et al. (2010) examined the soluble mycelial phosphoproteome of *F. graminearum* grown in nitrogen-poor media using a combination of 2D-E Pro-Q Diamond phospho staining, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and gel-free 1D SCX LC-MS analyses of TiO_2/immobilised metal affinity chromatography (IMAC) enriched phosphoproteins. In combination, a total of 301 putative phosphopeptides from 241 proteins were identified. Quantification via 2D-E identified nine protein spots that were altered in abundance during prolonged growth under nitrogen starvation. In a later study, the phosphoproteome of *F. graminearum* was profiled under non-limiting in vitro conditions using enriched phosphoproteins via gel-free 2D-LC Orbitrap MS/MS (Rampitsch et al. 2012). Analysis of the *F. graminearum* genome protein database with the resulting mass spectra identified 2,902 putative phosphopeptides from 1,496 proteins. Functional analysis of these proteins via Gene Ontology revealed that nuclear and intracellular proteins associated with ATP and nucleotide binding predominated.

The proteome of *F. graminearum*-infected barley was investigated in a series of studies by Yang et al. (2010a, b). 2D-E analysis of infected barley spikelets heavily contaminated with *F. graminearum* identified fungal superoxide dismutases, xylulose reductase, peptidyl-prolyl *cis-trans* isomerase, triosephosphate isomerase and proteins of unknown function. Following this, the extracellular proteome of *F. graminearum* was examined in detail using 2D-E MALDI-TOF MS/MS. The fungus was grown on barley and wheat flour as the sole nutrient source to mimic the nutritional condition during infection. The study was able to resolve 170 protein spots. Of these, 72 protein spots were differentially abundant between growth on barley and wheat flour. Most of these proteins were associated with putative fungal cell wall remodelling and plant-cell-degrading enzymes (Yang et al. 2012). To complement the in vitro study, selected genes associated with trichothecene biosynthesis, plant cell wall degradation, cell wall remodelling and oxidative stress were monitored for expression during infection. In all cases, differential gene expression patterns were observed.

D. Development of Infection Structures and Virulence

Attachment of the fungal spore to the host is the first step towards establishing infection. Gokce et al. (2012) performed an in-depth investigation into the **conidial proteome** diversity of *M. oryzae* using rapid filter-aided sample preparation and anion fractionation using a StageTip format that was developed in the Mann laboratory (Wisniewski et al. 2009). The fractionated peptides were then subjected to MS analysis via an LTQ Orbitrap. Matches against the *M. oryzae* predicted protein database identified 2,912 conidial proteins. Interestingly, proteins associated with translation, energy and metabolism were overrepresented when compared to the whole-genome protein prediction. On host penetration, *M. oryzae* resides within the apoplastic space of the host before initiating a necrotrophic phase. To address this, the intercellular proteome of rice infected with *M. oryzae* was extracted for 2D-E LC-MS/MS analysis (Shenton et al. 2012). Several proteins that changed in abundance between compatible and incompatible pathogen interactions were identified. This included a cyclophilin (Cyp1), which was previously identified as a pathogenicity factor in *M. oryzae* (Viaud et al. 2002). Cyp1 was previously implicated in appressorium turgor gener-

ation, lipid biosynthesis and the development of asexual spores. However, the role of Cyp1 as a secreted product in the host apoplast is unknown at this stage.

A similar study to examine the proteomic composition of the **apoplastic space** during *M. oryzae* infection was recently performed by Kim et al. (2012b) using 2D-E difference gel electrophoresis (DIGE) and multidimensional protein identification technology (MuDPIT). There were 732 unique proteins identified, predominantly through gel-free MuDPIT. Of these, 291 proteins were plant derived and were predominantly implicated in stress response, ROS and energy metabolism. Furthermore, 441 pathogen-derived proteins were predominantly associated with plant cell wall degradation. RT-PCR (Reverse transcription polymerase chain reaction) analysis on selected plant cell-wall-degrading enzyme-encoding genes revealed evidence of increased expression during nutrient starvation and compatible host interaction. In addition, the study identified Cyp1 from the apoplast proteome derived from the compatible interaction.

The **fungal cell wall** protects the organism from a variety of environmental stresses (De Groot et al. 2005). As an exterior part of the fungus, the cell wall also mediates interaction with the host. Efforts to decipher the protein diversity through proteomics have been hampered by the difficulty in extracting proteins that are covalently attached to the β-1,6-glucan and chitin cell wall matrix. Using trypsin to liberate exposed peptides from purified cell walls of the vascular wilt pathogen *Fusarium oxysporum*, Prados-Rosales et al. (2009) were able to identify 174 proteins. Of these, 19 proteins possessed a predicted signal peptide. Of these 19, there were 10 with a predicted glycosylphosphatidyl-inositol motif that is associated with cell wall attachment. Two proteins with a cysteine-rich CFEM domain were identified. Similarly, two proteins with a carbohydrate-binding WSC domain were identified. In addition, a tyrosinase-like protein was identified with a probable role in the formation of melanin. Melanin has been linked to tolerance to environmental stress and virulence (Perpetua et al. 1996).

The white mould fungus *Sclerotinia sclerotiorum* possesses a broad host range and causes significant crop loss. Through hyphal aggregation, the pathogen produces highly melanised structures called **sclerotia**. These structures permit the pathogen to survive environmental and biotic stresses and function as a disease inoculum. Liang et al. (2010a) compared the proteome of sclerotia from early development until maturity using 2D-E. The study identified 719 unique proteins. Of these, 88 showed differential abundance throughout development. LC-MS/MS followed by Mascot was used to identify these proteins through matches with the National Centre for Biotechnology Information (NCBI) non-redundant database, which includes the *S. sclerotiorum* genome. Most proteins implicated in amino acid, lipid and secondary metabolism decreased in abundance during the course of sclerotia development. Not surprisingly, proteins that have implicated functions in dihydroxynapththalene (DHN) melanin biosynthesis (scytalone dehydratase and trihydroxynapththalene [THN] reductase) increased in abundance during maturation.

During maturity, the sclerotium produces an exudate. However, little is known about the composition of the sclerotial exudate. Gel-based proteomic analyses identified 56 unique proteins. Most identified proteins from the exudate proteome were implicated in lipid/secondary and carbohydrate metabolism (plant cell wall degradation) (Liang et al. 2010b).

Mycosphaerella graminicola (anamorph *Septoria tritici*) is the causal agent of *Septoria tritici* blotch of wheat. *Mycosphaerella graminicola* exhibits dimorphism and forms hyphae from yeast-like structures to facilitate infection. Motteram et al. (2011) were able to disrupt the **yeast-to-hyphae transition** through the deletion of a gene that encodes an α-1,2-mannosyltransferase (*MgAlg2*). Mutants defective in *MgAlg2* cannot form hyphae and failed to establish infection. Analysis of the proteome with ProQ-Emerald glycoprotein-specific staining of 1D-E identified major alterations in the glycoproteome of *mgalg2* mutants. Two proteins with a molecular mass of 75 and 77 kDa stained strongly in *mgalg2*. In yeast, Alg2 is associated with *N*-glycosylation. Treatment of *mgalg2* mutants with Peptide -N-Glycosidase F (PNGaseF), an enzyme that removes N-glycans,

did not affect the staining intensity of both proteins. Although these alterations have not been identified specifically, the study has demonstrated the use of proteomics for the analysis of PTMs in fungal dimorphism.

The soilborne fungus *Verticillium dahliae* causes vascular wilting on a wide range of plants worldwide. The pathogen is a necrotroph with a high variability in aggressiveness. El-Bebany et al. (2010) used 2D-E LC MS/MS to compare the mycelial proteome of two isolates that differ in virulence. There were 25 protein spots consistently different between both isolates. Proteins that were more abundant in the highly virulent isolate included a THN reductase that is associated with DHN melanin biosynthesis, a thioredoxin putatively involved in oxidative stress protection and an isochorismatase hydrolase that may function as a plant defence suppressor.

Biotrophic fungi are notoriously difficult to study in isolation because of their host-dependent lifestyle. Once host penetration is achieved, the pathogen seeks to acquire nutrients from the host while causing minimal damage. These fungi establish cell membrane invaginated biotrophic feeding structures called **haustoria** that direct nutrients away from the host and into the pathogen (Mendgen and Hahn 2002). Recently, proteomics has been used to profile the haustorial proteome of two economically important biotrophic fungal pathogens of cereals (Godfrey et al. 2009; Song et al. 2011). Haustoria of the barley powdery mildew pathogen *Blumeria graminis* were extracted from heavily infected barley via gradient centrifugation. Proteins from the enriched haustoria were subjected to SDS extraction. Proteins were trypsin digested and directly analysed with LC-MS/MS. As the *B. graminis* annotated genome was not available at the time of study, ESTs (expressed sequence tags) and the NCBI non-redundant (NR) database were used for protein identification (Godfrey et al. 2009). The majority of the proteins identified were implicated in protein metabolism and folding. Interestingly, pyruvate decarboxylase (PDC) was observed in neither the protein fraction nor the preliminary *B. graminis* genome sequence. PDC is involved in alcohol fermentation and found throughout the fungal taxa. This indicates that *B. graminis* is not subjected to significant hypoxia during infection. In addition, the study was unable to find membrane-bound carbohydrate transporters that are critical for nutrient acquisition despite using a protein extraction method that targets hydrophobic membrane proteins. Despite this, a plasma membrane H^+-ATPase was identified.

Haustoria of the wheat leaf rust fungus *Puccinia triticina* were extracted in a similar fashion, and the proteome was profiled using gel-based fractionation and LC-MS/MS (Song et al. 2011). There were 260 proteins collectively identified through ESTs and a partially annotated *P. triticina* genome. As expected, most identified proteins are implicated in general metabolism, energy, transport and translation. However, the study identified 6 potential effectors with no functional homology match, consisting of a predicted signal peptide and less than 300 amino acids. Genes that code for these proteins are highly expressed during advanced infection that coincides with haustoria development.

A proteome map of the urediniospores from the soybean rust fungus *Phakopsora pachyrhizi* was established by Luster et al. (2010) using 2D-E MALDI-TOF MS/MS. Whilst the organism lacked a sequenced genome, 117 proteins were identified through a non-redundant database search with high confidence. Of these, 101 proteins matched to *P. pachyrhizi* ESTs. As expected, the majority of the proteins identified are involved in primary metabolism, energy transduction, stress, cellular regulation and signalling. Only 35 identified proteins were common to an earlier shotgun proteomic study by Cooper et al. (2006) on ungerminated urediniospores of a related rust pathogen *Uromyces appendiculatis*. This is not surprising considering that both studies were performed with different proteomic analytical platforms.

E. Biocontrol

Cryphonectria parasitica is a serious fungal pathogen of chestnuts (Kim et al. 2012a). However, the pathogen is susceptible to a family of double-stranded RNA viruses named **Hypoviridae**. *Cryphonectria hypovirus* 1 (CHV1) is a

member of this family that causes hypovirulence as well as perturbation in sporulation, fertility and pigmentation. The mechanism of hypovirulence in *C. parasitica* is a subject of intense investigations with molecular genetics (Dawe and Nuss 2001) and transcriptomics with cDNA arrays (Allen et al. 2003; Dawe et al. 2004). The recent availability of a sequenced genome allows a global and high-throughput analysis of virulence using proteomics. Using 2D-E LC MS/MS, Kim et al. (2012a) compared the intracellular proteome of a virus-free strain (EP155/2) and a CHV1-harbouring hypovirulent strain (UEP1) supplemented with tannic acid (TA), which is present in the bark of chestnut trees and was used in the analysis to mimic host contact. The study identified 30 proteins from 704 protein spots that were differentially abundant under TA-inducing and hypovirus-infected conditions. Proteins implicated in stress response, signalling, structure and metabolism were observed in differently abundant protein spots.

F. Proteomics to Assist Genome Annotation

Most modern proteomic studies are performed by matching spectral data against one or more genome-derived protein data sets of the organism of interest. However, this approach is not without its drawbacks. Genes are often annotated using homology-based methods or ab initio prediction using sophisticated software with little experimental support (Wall et al. 2011). Such an approach can often lead to gene prediction errors, such as false-positive/-negative genes and incorrect Open Reading Frames. **Proteogenomics** uses peptide-derived spectral data generated from MS to verify gene annotations as well as identify new genes or errors (Jaffe et al. 2004). This is an emerging high-throughput technique that is a direct 'to-genome' mapping technique by which the mass spectra from protein analyses are mapped to current translated genome annotation and 'stop-codon-to-stop-codon' Open Reading Frame frame translations of the genome sequence in all six frames (Fig. 12.1). In conjunction with transcriptomics, proteogenomics can be used to verify and to identify new genes. In addition, proteogenomics offers advantages over transcript data as it allows the determination of subcellular protein localisations and provides accurate information on the correct Open Reading Frame of a gene. The use of proteomic data for genome annotation was first demonstrated on the budding yeast *Saccharomyces cerevisiae* by the Mann laboratory in the mid-1990s (Shevchenko et al. 1996). Using 2D-E combined with MALDI-MS and LC-MS, the study was able to verify 128 proteins from 150 protein spots via 2D-E. In addition, de novo peptide sequencing of protein spots resulted in the identification at least 32 new proteins from uncharacterised Open Reading Frames (Shevchenko et al. 1996).

The decline in the cost of genome sequencing has resulted in the genomes of many fungal pathogens being sequence over since 2007. This has led to the need for high-throughput and accurate experimental methods to assist in genome annotation. Several publications in phytopathology have recently emerged demonstrating the full potential of proteogenomics in genome annotation. A draft genome sequence of *S. nodorum* was published in 2007 (Hane et al. 2007). Gene prediction based on EST-supported gene training resulted in a confident prediction of 10,762 version 2 genes, in addition to 5,345 less-reliably predicted version 1 genes that lacked experimental verification (Hane et al. 2007). A large-scale proteomic study using an offline 2D-LC MALDI-TOF MS/MS analysis of the *S. nodorum* intracellular proteome has resulted in the verification of 2,134 genes, including 188 version 1 genes (Bringans et al. 2009). Furthermore, interrogation of the mass spectra against a six-frame whole-genome translation yielded evidence of 68 genes with frameshift errors and 604 genes requiring untranslated region (UTR)/exon extensions. Also, 47 potential new genes were identified. We anticipate that this study can be complemented with future proteogenomic data generated from the analysis of other *S. nodorum* subproteomes.

The sequencing of the 120-Mb *B. graminis* genome resulted in the identification of 5,848 manually curated genes (Spanu et al. 2010).

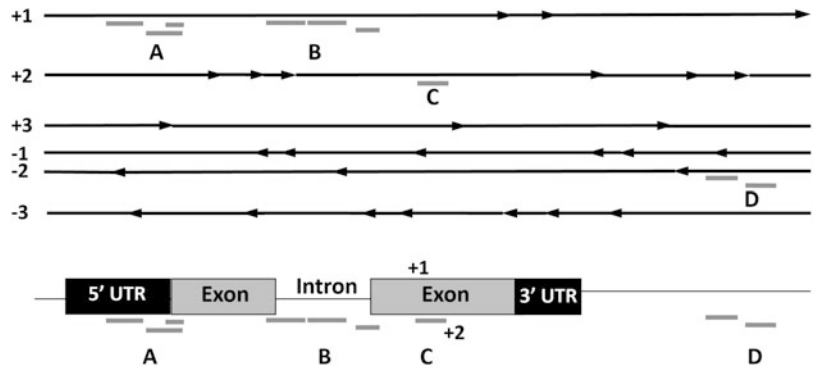


Fig. 12.1. Proteogenomic gene annotation using a six-frame 'stop-codon-to-stop-codon' whole-genome translation (+1 to −3). Spectral data generated from tryptic peptides are matched to the six-frame database. This method allows the identification of ***A*** UTR extension, ***B*** false intron and ***C*** frameshift on predicted genes or the discovery of ***D*** potential new genes

With the aim of validating these genes, Bindschedler et al. (2009) performed proteogenomics on mycelial proteins derived from the haustoria, conidia and sporulating hyphae. LC-MS/MS was used to analyse the resulting tryptic peptides from excised 1D SDS-PAGE gel slices. The resulting mass spectra were matched against the predicted protein data set and an extensive collection of 17,869 ESTs. As a result, a total of 827 proteins were identified. Surprisingly, 24 of these proteins were only observed in the EST collection and not in the annotated gene set. This indicates that the *B. graminis* genome possesses gaps within the genome assembly where genes are split between contigs (Bindschedler et al. 2009). In addition, haustorial-specific proteins that may play a specific role in the pathogen interaction with the host were identified and structurally modelled in a later study (Bindschedler et al. 2011).

III. Metabolomics

A. Metabolomic Methodologies

Metabolites are often obtained from cell-free supernatant or from fungal mycelia. Polar organic solvents such as methanol, acetonitrile and acetone are commonly used to extract hydrophilic metabolites, whereas chloroform is used to extract non-polar metabolites. Gas chromatography (GC)-MS, LC-MS, flow-injection/direct-infusion electrospray ionisation (ESI)-MS and nuclear magnetic resonance (NMR) spectroscopy are major analytical platforms commonly used for metabolomics in fungal phytopaththology. Of these, GC-MS is perhaps the most widely used technique for fungal metabolomics because of low running cost, high sensitivity and reproducibility. It is particularly suited for the analysis of volatile and thermally stable metabolites (Watson and Sparkman 2007). Furthermore, an extensive collection of GC-MS databases is available that can facilitate metabolite identification. However, GC-MS is limited in its capacity to analyse large and non-volatile metabolites.

LC-MS is a versatile technique that can be used to separate metabolites with a wide range of polarity and requires minimal sample preparation. However, LC-MS possesses some inherent disadvantages, such as that the detection of different metabolite classes is column dependent and can suffer from retention time drifts. With ESI-MS, complex metabolite samples are resolved through direct injection without any forms of chromatographic separations (Koulman et al. 2007). It has a short run time and requires little sample preparation.

NMR uses a different detection method compared to MS-based platforms. It is a spectroscopic technique that analyses the spin properties of the nucleus of atoms (Moco et al. 2007). NMR is a non-destructive and non-biased platform that requires minimal sample

preparation. Furthermore, NMR can provide structural information. NMR suffers from relatively low sensitivity and requires a large sample size. However, the introduction of powerful field magnets and cryoprobes has addressed some of the sensitivity issues (Wishart 2008). It should be noted that other analytical platforms, such as capillary electrophoresis-MS and high-resolution Fourier transform (FT)-MS have been used with some success in medical sciences but have seen limited usage in fungal phytopathology because of high costs and technically demanding setups. Detailed experimental protocols for induction, extraction, quantification of fungal metabolites and the use of MS/spectroscopic analytical platforms are available (Nielsen and Smedsgaard 2003; Forseth and Schroeder 2012; Frisvad 2012; Gummer et al. 2012). These provide an excellent reference source for carrying out fungal metabolomics.

B. Mycotoxin Discovery

Stagonospora nodorum was considered a classical necrotrophic fungal pathogen with no known postharvest capabilities until recently. This changed with the analysis of a deletion mutant of *Sch1* that resulted in severely reduced asexual sporulation (Tan et al. 2008). A non-targeted metabolomic approach used for analysis of the *sch1* mutant metabolome resolved 223 metabolites. One metabolite eluting at a retention time of 45.57 min accumulated 200-fold higher in the mutant than in the wild type (Tan et al. 2009c). Purification of the metabolite via LC followed by a component match from the National Institute of Standards and Technology (NIST) database identified the compound as the carcinogenic mycotoxin **alternariol** (AOH). AOH was previously associated with the closely related *Alternaria* spp. (Scott 2001), and its presence in *S. nodorum* came as a surprise. In a later study, the putative ASPES transcriptional factor gene *SnStu1* was inactivated through gene deletion in *S. nodorum*. The resulting mutants were non-sporulating. GC-MS analysis of the hyphal metabolome revealed that the AOH level in all *snstu1* mutants was significantly reduced (IpCho et al. 2010). It is unlikely that AOH production is linked to sporulation in *S. nodorum*. With the identification of AOH in *S. nodorum*, the pathogen can now be considered a potential human threat.

Alternaria spp. are avid producers of post-harvest mycotoxins and phytotoxins (Magan et al. 1984; Wolpert et al. 2002). *Alternaria brassicicola* is the causal agent of dark leaf spot on Brassicaceae. Transcriptome analysis of *A. brassicicola-Brassica* interaction identified several genes associated with the biosynthesis of phytotoxins. To investigate this further, Pedras et al. (2009) fractionated the secreted metabolome of phytotoxic *A. brassicicola* culture filtrate using flash column chromatography followed by structural determination via NMR. The study identified nine secondary metabolites. Among these, **brassicicolin A** was obtained from a highly phytotoxic fraction. Brassicicolin A displays host-specific effector activity as leaves from highly susceptible *Brassica juncea* treated with the metabolite underwent necrosis. This was not observed for other Brassicaceae species, including a resistant cultivar of *Sinapis alba*.

In *F. verticillioides*, enzymes that are involved in the biosynthesis of **fumonisins** are located within a biosynthetic *Fum* gene cluster. The pathway is well characterised, and most pathway intermediates were identified. However, intermediates leading to the formation of fumonisin-like precursors have not been observed. *Fum6* is predicted to play an early role in fuminosin biosynthesis. Uhlig et al. (2012) analysed the metabolome of *fum6* mutants via LC-MS for accumulated metabolites followed by structural verification via NMR. This led to the identification of the fumonisin-like metabolite 2-amino-12, 16-dimethylicosane-3, 10-diol. Subsequent feeding assay experiments with the upstream *Fum1* polyketide synthase mutant converted the novel metabolite into fuminosin B1 through C14 and C15 hydroxylation. Hence, it is likely that *Fum6* encodes a hydroxylase.

In addition to being a pathogen, *F. verticillioides* can exist within the host as an **endophyte**. Endophytic *F. verticillioides* possesses

antagonistic properties and can reduce the severity of *Ustilago maydis* infection on maize (Lee et al. 2009). However, not much is known about the physiology of the interaction. To elucidate this, *U. maydis* and *F. verticillioides* colonies were co-cultured in vitro. Metabolites were extracted from agar plugs that were derived from single cultures and the interaction zone between the two species (Rodriguez Estrada et al. 2011; Jonkers et al. 2012). LC-MS analysis of the extracted metabolites identified ustilagic acids, mannosylerythritol lipid and a number of unknown compounds derived from *U. maydis*. Fumonisins, bikaverin, fusaric acid (FA) and two unknown compounds were observed from *F. verticillioides*. FA accumulated significantly when cocultivated with *U. maydis*. Upregulation of FA biosynthesis was further confirmed via RT-PCR analysis of putative FA biosynthetic genes. FA seems to function as an antibiotic as it inhibits the growth of *U. maydis* in vitro (Jonkers et al. 2012).

The *Tri5* gene of *F. graminearum* encodes a trichodiene synthase, the first key enzyme in the biosynthesis of **trichothecene** mycotoxins (Proctor et al. 1995). Inactivation of this gene abolishes the production of trichothecenes and reduces pathogenicity (Proctor et al. 1995). Chen et al. (2011) performed a comprehensive non-targeted metabolomic analysis of the *tri5* mutant using 600-MHz NMR equipped with a cryoprobe. When combined with gene expression analysis via qunatitative RT-PCR, it was observed that *tri5* deletion caused extensive alterations in primary metabolic pathways, such as the TCA (tricarboxylic acid) cycle, amino acid biosynthesis and lipid metabolism. It was previously thought that trichothecene biosynthesis contributes to the pathogenicity of *F. gramimearum* (Proctor et al. 1995). It is possible that alterations in primary metabolism may also be a factor that contributes to the reduction of pathogenic fitness. In addition to the trichothecene mycotoxin DON, *F. gramineaum* possesses the genetic machineries to produce polyketide-based metabolites such as the estrogenic mycotoxin zearalenone (Kim et al. 2005). The fungus possesses 15 putative polyketide synthase (PKS) genes, with 10 of these not linked to a product. These include the *Fls1* polyketide synthase gene, which is located within a putative gene cluster. To identify the Fsl1 product, Sorensen et al. (2012) constitutively activated genes within the cluster through the overexpression of the activating transcription factor gene *Fsl7*. A combination of LC-MS followed by NMR identified three novel metabolites that belong to the **fusarielin** family. The deletion of *Fsl1* abolished fusarielin production. Furthermore, these novel fusarielin inhibited the growth of mammalian cell lines. Whether these metabolites are produced under agricultural settings remains to be determined. However, findings from this study have further enhanced the notoriety of *F. graminearum* as a postharvest threat.

C. Profiling the Fungal Metabolome

Metabolomics is a powerful tool that can be used to generate a metabolome map. Lowe et al. (2010) elegantly demonstrated this by utilising a combination of NMR and ESI-MS to screen the basal metabolism of various isolates of DON-producing *Fusarium spp*. in vitro. A common metabolic response to DON-inducive conditions includes an increased abundance of mannitol, trehalose, γ-aminobutyric acid, glycerol and glycine-betaine. Furthermore, principal components analysis was able to differentiate between different *Fusarium* spp. as well as strains of the same species, thus highlighting metabolomics as a potential tool that can be used for fungal **genotyping**.

Rhizoctonia solani is an important soilborne pathogen of a large number of economically important crops. The pathogen forms sclerotia under adverse environmental conditions. Exudates produced by sclerotia possess phytotoxic and antifungal activities (Aliferis and Jabaji 2010). The metabolite composition of sclerotial exudates was recently elucidated using a multitude of analytical platforms that include GC-MS, NMR and FT-MS (Aliferis and Jabaji 2010). More than 90 metabolites were

identified; predominant among these were carboxylic acids and phenolic metabolites. Phenylacetic acid and ferulic acid are two of the most abundant metabolites of the carboxylic acids and phenolics, respectively. Both metabolites possess phytotoxic properties.

D. Defining the Role of Sugar Alcohol Metabolism in Fungal Phytopathogens

Sugar alcohols are widely distributed in fungi. Sugar alcohols possess a multitude of physiological roles in fungi that include stress tolerance (Son et al. 2012), signalling (White et al. 1991) and carbohydrate storage (Jennings 1995). In *S. nodorum*, a combination of functional genomics and metabolomics identified a role for **mannitol** in the development of asexual spores (Solomon et al. 2005, 2006). Further metabolomic profiling *S. nodorum* via GC-MS revealed the accumulation of arabitol and glycerol under osmotic stress (Lowe et al. 2008). Subsequent inactivation of the arabitol pathway via targeted gene deletion resulted in mutants that showed reduced tolerance to high salt. Intriguingly, these mutants retained the ability to synthesise arabitol under osmotic stress, albeit at a significantly lower abundance than the wild type. It was hypothesised that *S. nodorum* possesses a secondary pathway for arabitol biosynthesis. Furthermore, glycerol was accumulated under osmotic stress and may function to compensate for arabitol reduction.

E. Metabolomics in Plant-Fungal Interactions

For a successful infection to occur, fungal phytopathogens must possess the mechanism to either cope or manipulate the host defence in its favour. An example of this is *Cmu1*, which encodes a secreted **chorismate mutase** in *U. maydis*. Cmu1 enters the plant cell and has been shown to function as an effector. Metabolic profiling via LC/UPLC (ultra-performance liquid chromatography) TOF-MS demonstrated that Cmu1 manipulates the host metabolome by suppressing the biosynthesis of salicylic acid, phenylpropanoid and lignin biosynthesis metabolites. Consequently, *U. maydis* deleted in *Cmu1* showed a significant reduction in tumour formation (Djamei et al. 2011). It was proposed that Cmu1 functions to deplete chorismate from the host plastid, thereby reducing substrate availability for the biosynthesis of plant defense-related compounds.

Magnaporthe oryzae is capable of living within the host as a biotroph prior to establishing the rice blast disease. However, the mechanism that governs this hemibiotrophic lifestyle is poorly understood. Parker et al. (2009) examined the metabolome of susceptible host plant species via ESI-MS and GC-MS to identify a common pattern of metabolite regulation. Prior to symptom development, accumulation of the TCA intermediates, polyamines, quinate, phenylalanine and non-polymerised lignin precursors was observed. *Magnaporthe oryzae* seems to suppress the host defence by manipulating the metabolite content of the phenylpropanoid pathway and restricting the production of lignin in the susceptible host. Accumulation of photosynthates and amino acids was observed around the site of infection during the rice blast. The fungus assimilated the photosynthate as indicated by rapid mannitol and glycerol accumulation. This provides a clear indication that the fungus has established a nutritional sink at the site of infection.

Manipulation of the host carbon metabolism was also observed in sunflower infected with *S. sclerotiorum*. GC-MS analysis revealed that several TCA intermediates were less abundant in susceptible sunflower compared to the resistant genotype undergoing fungal infection (Peluffo et al. 2010).

Stagonospora nodorum uses proteinaceous necrotrophic effectors to disable the host during infection. **ToxA** is one of three major effectors identified in *S. nodorum*. ToxA is a virulence factor that causes necrosis on wheat carrying the dominant susceptibility gene *Tsn1*. ToxA is also present in the tan (syn. yellow) spot fungus *Pyrenophora tritici-repentis*. Experimental evidence suggests that *PtrToxA* may have been acquired from *S. nodorum* via horizontal gene transfer (Friesen et al. 2006). Like SnToxA, PtrToxA also causes necrosis in a *Tsn1* background. The wheat response to SnToxA

was recently investigated at the proteome and metalome levels by Vincent et al. (2012a). Analysis of SnToxA-treated *Tsn1*-wheat via 2D-E proteomics and GC-MS metabolomics revealed substantial reprogramming of the host TCA cycle and the collapse of photosynthesis. This is expected as PtrToxA was previously shown to uncouple photosystems I and II (Manning et al. 2009). In addition, pathogenesis-related proteins were more abundant, possibly from oxidative stress generated from the increased abundance of proteins implicated in ROS production. Furthermore, findings from this study strongly correlated with a number of transcriptomic studies on PtrToxA-exposed wheat (Adhikari et al. 2009; Pandelova et al. 2009). This study demonstrated a clear role of SnToxA in tailoring the host metabolism for necrotrophic infection.

Verticillium longisporum causes vascular wilt in canola and other Brassicaceae. The fungus infects the host through the root system and subsequently colonises the xylem (Eynck et al. 2007). It is likely that the pathogen acquires most of its nutrients within the xylem, thereby exerting its detrimental influence on other parts of the plant. Floerl et al. (2012) undertook a 2D-E approach to analyse the apoplastic proteome of *Arabidopsis thaliana* infected with *V. longisporum*. Six plant-derived proteins implicated in defense, cell wall modifications and development were increased during infection, whereas a lectin-like protein implicated in plant-fungal interaction was reduced. Furthermore, LC-MS metabolomics revealed an increased abundance of defense-related compounds such as jasmonic and salicylic acids in the apoplast metabolome, thus indicating an attempt by the host to mount a defense response.

A similar approach to identify potential metabolite markers that are associated with a fusarium head blight resistance in barley was performed using a highly sensitive LC-LTQ (linear trap quadrupole) Orbitrap-MS platform (Bollina et al. 2010, 2011; Kumaraswamy et al. 2011). Metabolites associated with the phenylpropanoid, fatty acid and flavonoid pathways were consistently more abundant in resistant than in susceptible barley varieties.

Botrytis cinerea is responsible for grey rot in grapes, which leads to a reduction wine quality. Metabolomics have been used extensively to investigate the influence of *B. cinerea* on volatile aromatic compounds in wine (Sarrazin et al. 2007); however, the effect on primary metabolism is poorly understood. NMR profiling revealed a substantial reduction in fermentative products generated from alcohol fermentation, such as succinate, glycerol, 2,3-butanediol, tyrosine and valine (Hong et al. 2011).

IV. Outlook on Proteomics and Metabolomics in Phytopathology

We have performed a comprehensive review of recent publications that demonstrated the contribution of proteomics and metabolomics in phytopathology since Tan et al. (2009b). It is immediately apparent that fungal proteomics have greatly benefited from the recent influx of sequenced fungal genomes in comparison to pre-2009 studies (Tan et al. 2009b). Refinement of gene annotation will greatly assist obtaining accurate protein identification. In addition, we have observed several laboratories applying proteogenomics to accurately refine in silico gene prediction. We anticipate that proteogenomics will be applied to an increasing number of fungal genomes as peptide mapping tools are now readily available (Bringans et al. 2009). This approach is complementary to transcript-based gene validation methods such as microarrays and RNA-seq (RNA sequencing).

Proteome-wide analysis of **posttranslational modifications** in fungal virulence via proteomics has seen limited application in phytopathology. PTMs in higher eukaryotes are complex as there over 200 known covalent modifications (Walsh 2005). Many of these modifications are labile and may be lost during processing or require enrichment prior to detection. The Rampitsch laboratory (Rampitsch et al. 2010, 2012) has demonstrated that it is possible to apply phosphosproteomics in a fungal phytopathogen. Thus, we anticipate that phosphoproteomics will be increasingly used to

dissect fungal virulence, such as phosphorylation in key fungal signalling pathways that are associated with virulence.

The discovery of fungal effectors is a key topic of research in phytopathology. Identification of these effectors often requires purification through chromatographic techniques and quantitative trait loci mapping of resistance/susceptibility genes in the host. To complement this approach, the availability of genome sequences allows computational proteomics to screen genes and their predicted polypeptide sequences for effector hallmarks, such as the presence of secretory signals, cysteine patterns and localisation in genome 'hot spots' (Mueller et al. 2008; Brown et al. 2012; Morais do Amaral et al. 2012). Application of this predictive approach in effector discovery has led to the identification of the Tox1 effector in *S. nodorum* (Liu et al. 2012).

The distribution of small molecules during fungal infection is of particular interest. MALDI-MS imaging is an emerging technique that is used to monitor the spatial distribution of small molecules in a biological system in situ (Fournier et al. 2008; Lee et al. 2012). To our knowledge, there are no published reports on the use of MALDI-MS imaging in phytopathology. However, it has been used to monitor metabolite exchange in microorganisms (Gonzalez et al. 2012) and for protein profiling of barley grain (Kaspar et al. 2011). The adaptation of MALDI-MS imaging in phytopathology can be a revolutionary step towards monitoring the spatial distribution of small molecules in infected plant tissues.

References

Adhikari TB, Bai J, Meinhardt SW, Gurung S, Myrfield M, Patel J, Ali S, Gudmestad NC, Rasmussen JB (2009) *Tsn1*-mediated host responses to ToxA from *Pyrenophora tritici-repentis*. Mol Plant Microbe Interact 22:1056–1068

Aliferis KA, Jabaji S (2010) Metabolite composition and bioactivity of *Rhizoctonia solani* sclerotial exudates. J Agric Food Chem 58:7604–7615

Allen TD, Dawe AL, Nuss DL (2003) Use of cDNA microarrays to monitor transcriptional responses of the chestnut blight fungus *Cryphonectria parasitica* to infection by virulence-attenuating hypoviruses. Eukaryot Cell 2:1253–1265

Bindschedler LV, Burgis TA, Mills DJ, Ho JT, Cramer R, Spanu PD (2009) In planta proteomics and proteogenomics of the biotrophic barley fungal pathogen *Blumeria graminis f. sp. hordei*. Mol Cell Proteomics 8:2368–2381

Bindschedler LV, McGuffin LJ, Burgis TA, Spanu PD, Cramer R (2011) Proteogenomics and in silico structural and functional annotation of the barley powdery mildew *Blumeria graminis f. sp. hordei*. Methods 54:432–441

Bollina V, Kumaraswamy GK, Kushalappa AC, Choo TM, Dion Y, Rioux S, Faubert D, Hamzehzarghani H (2010) Mass spectrometry-based metabolomics application to identify quantitative resistance-related metabolites in barley against *Fusarium* head blight. Mol Plant Pathol 11:769–782

Bollina V, Kushalappa AC, Choo TM, Dion Y, Rioux S (2011) Identification of metabolites related to mechanisms of resistance in barley against *Fusarium graminearum*, based on mass spectrometry. Plant Mol Biol 77:355–370

Bringans S, Hane JK, Casey T, Tan KC, Lipscombe R, Solomon PS, Oliver RP (2009) Deep proteogenomics; high throughput gene validation by multidimensional liquid chromatography and mass spectrometry of proteins from the fungal wheat pathogen *Stagonospora nodorum*. BMC Bioinformatics 10:301

Brito N, Espino JJ, Gonzalez C (2006) The endo-beta-1,4-xylanase Xyn11A is required for virulence in *Botrytis cinerea*. Mol Plant Microbe Interact 19:25–32

Brown NA, Antoniw J, Hammond-Kosack KE (2012) The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: a refined comparative analysis. PLoS One 7:e33731

Casey T, Solomon PS, Bringans S, Tan KC, Oliver RP, Lipscombe R (2010) Quantitative proteomic analysis of G-protein signalling in *Stagonospora nodorum* using isobaric tags for relative and absolute quantification. Proteomics 10:38–47

Chen F, Zhang J, Song X, Yang J, Li H, Tang H, Liao YC (2011) Combined metabonomic and quantitative real-time PCR analyses reveal systems metabolic changes of *Fusarium graminearum* induced by *Tri5* gene deletion. J Proteome Res 10:2273–2285

Choi YE, Butchko RA, Shim WB (2012) Proteomic comparison of *Gibberella moniliformis* in limited-nitrogen (fumonisin-inducing) and excess-nitrogen (fumonisin-repressing) conditions. J Microbiol Biotechnol 22:780–787

Cooper B, Garrett WM, Campbell KB (2006) Shotgun identification of proteins from uredospores of the bean rust *Uromyces appendiculatus*. Proteomics 6:2477–2484

Coumans JV, Moens PD, Poljak A, Al-Jaaidi S, Pereg L, Raftery MJ (2010) Plant-extract-induced changes in the proteome of the soil-borne pathogenic fungus *Thielaviopsis basicola*. Proteomics 10:1573–1591

Dawe AL, Nuss DL (2001) Hypoviruses and chestnut blight: exploiting viruses to understand and modulate fungal pathogenesis. Annu Rev Genet 35:1–29

Dawe AL, Segers GC, Allen TD, McMains VC, Nuss DL (2004) Microarray analysis of *Cryphonectria parasitica* Ga- and Gbg-signalling pathways reveals extensive modulation by hypovirus infection. Microbiology 150:4033–4043

De Groot PW, Ram AF, Klis FM (2005) Features and functions of covalently linked proteins in fungal cell walls. Fungal Genet Biol 42:657–675

Djamei A, Schipper K, Rabe F, Ghosh A, Vincon V, Kahnt J, Osorio S, Tohge T, Fernie AR, Feussner I, Feussner K, Meinicke P, Stierhof YD, Schwarz H, Macek B, Mann M, Kahmann R (2011) Metabolic priming by a secreted fungal effector. Nature 478:395–398

Donofrio NM, Oh Y, Lundy R, Pan H, Brown DE, Jeong JS, Coughlan S, Mitchell TK, Dean RA (2006) Global gene expression during nitrogen starvation in the rice blast fungus, *Magnaporthe grisea*. Fungal Genet Biol 43:605–617

Egan MJ, Talbot NJ (2008) Genomes, free radicals and plant cell invasion: recent developments in plant pathogenic fungi. Curr Opin Plant Biol 11:367–372

El-Bebany AF, Rampitsch C, Daayf F (2010) Proteomic analysis of the phytopathogenic soilborne fungus *Verticillium dahliae* reveals differential protein expression in isolates that differ in aggressiveness. Proteomics 10:289–303

Espino JJ, Gutierrez-Sanchez G, Brito N, Shah P, Orlando R, Gonzalez C (2010) The *Botrytis cinerea* early secretome. Proteomics 10:3020–3034

Eynck C, Koopmann B, Grunewaldt-Stoecker G, Karlovsky P, von Tiedemann A (2007) Differential interactions of *Verticillium longisporum* and *V. dahliae* with *Brassica napus* detected with molecular and histological techniques. Eur J Plant Pathol 118:259–274

Fernandez-Acero FJ, Colby T, Harzen A, Cantoral JM, Schmidt J (2009) Proteomic analysis of the phytopathogenic fungus *Botrytis cinerea* during cellulose degradation. Proteomics 9:2892–2902

Fernandez-Acero FJ, Colby T, Harzen A, Carbu M, Wieneke U, Cantoral JM, Schmidt J (2010) 2-DE proteomic approach to the *Botrytis cinerea* secretome induced with different carbon sources and plant-based elicitors. Proteomics 10:2270–2280

Floerl S, Majcherczyk A, Possienke M, Feussner K, Tappe H, Gatz C, Feussner I, Kues U, Polle A (2012) *Verticillium longisporum* infection affects the leaf apoplastic proteome, metabolome, and cell wall properties in *Arabidopsis thaliana*. PLoS One 7:e31435

Forseth RR, Schroeder FC (2012) Correlating secondary metabolite production with genetic changes using differential analysis of 2D NMR spectra. Methods Mol Biol 944:207–219

Fournier I, Wisztorski M, Salzet M (2008) Tissue imaging using MALDI-MS: a new frontier of histopathology proteomics. Expert Rev Proteomics 5:413–424

Friesen TL, Stukenbrock EH, Liu ZH, Meinhardt S, Ling H, Faris JD, Rasmussen JB, Solomon PS, McDonald BA, Oliver RP (2006) Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38:953–956

Frisvad JC (2012) Media and growth conditions for induction of secondary metabolite production. Methods Mol Biol 944:47–58

Godfrey D, Zhang Z, Saalbach G, Thordal-Christensen H (2009) A proteomics study of barley powdery mildew haustoria. Proteomics 9:3222–3232

Gokce E, Franck WL, Oh Y, Dean RA, Muddiman DC (2012) In-depth analysis of the *Magnaporthe oryzae* conidial proteome. J Proteome Res 11:5827–5835

Gonzalez DJ, Xu Y, Yang YL, Esquenazi E, Liu WT, Edlund A, Duong T, Du L, Molnar I, Gerwick WH, Jensen PR, Fischbach M, Liaw CC, Straight P, Nizet V, Dorrestein PC (2012) Observing the invisible through imaging mass spectrometry, a window into the metabolic exchange patterns of microbes. J Proteomics 75:5069–5076

Güldener U, Seong KY, Boddu J, Cho S, Trail F, Xu JR, Adam G, Mewes HW, Muehlbauer GJ, Kistler HC (2006) Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression in vitro and in planta. Fungal Genet Biol 43:316–325

Gummer JP, Krill C, Du Fall L, Waters OD, Trengove RD, Oliver RP, Solomon PS (2012) Metabolomics protocols for filamentous fungi. Methods Mol Biol 835:237–254

Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19:1720–1730

Hane JK, Lowe RG, Solomon PS, Tan KC, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE, Torriani SF, McDonald BA, Oliver RP (2007) Dothideomycete plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. Plant Cell 19:3347–3368

Herbert BR, Harry JL, Packer NH, Gooley AA, Pedersen SK, Williams KL (2001) What place for polyacrylamide in proteomics? Trends Biotechnol 19:S3–S9

Hong YS, Cilindre C, Liger-Belair G, Jeandet P, Hertkorn N, Schmitt-Kopplin P (2011) Metabolic influence of *Botrytis cinerea* infection in champagne base wine. J Agric Food Chem 59:7237–7245

Ipcho SV, Tan KC, Koh G, Gummer J, Oliver RP, Trengove RD, Solomon PS (2010) The transcription factor StuA regulates central carbon metabolism, mycotoxin production, and effector gene expression in the wheat pathogen *Stagonospora nodorum*. Eukaryot Cell 9:1100–1108

Ipcho SV, Hane JK, Antoni EA, Ahren D, Henrissat B, Friesen TL, Solomon PS, Oliver RP (2012) Transcriptome analysis of *Stagonospora nodorum*: gene models, effectors, metabolism and pantothenate dispensability. Mol Plant Pathol 13:531–545

Issaq H, Veenstra T (2008) Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. Biotechniques 44:697–698, 700

Jaffe JD, Berg HC, Church GM (2004) Proteogenomic mapping as a complementary method to perform genome annotation. Proteomics 4:59–77

Jennings DH (1995) The physiology of fungal nutrition. Cambridge University Press, Cambridge

Jonkers W, Rodriguez Estrada AE, Lee K, Breakspear A, May G, Kistler HC (2012) Metabolome and transcriptome of the interaction between *Ustilago maydis* and *Fusarium verticillioides in vitro*. Appl Environ Microbiol 78:3656–3667

Jung YH, Jeong SH, Kim SH, Singh R, Lee JE, Cho YS, Agrawal GK, Rakwal R, Jwa NS (2012) Secretome analysis of *Magnaporthe oryzae* using in vitro systems. Proteomics 12:878–900

Kaspar S, Peukert M, Svatos A, Matros A, Mock HP (2011) MALDI-imaging mass spectrometry—an emerging technique in plant biology. Proteomics 11:1840–1850

Kim YT, Lee YR, Jin J, Han KH, Kim H, Kim JC, Lee T, Yun SH, Lee YW (2005) Two different polyketide synthase genes are required for synthesis of zearalenone in *Gibberella zeae*. Mol Microbiol 58:1102–1113

Kim JM, Park JA, Kim DH (2012a) Comparative proteomic analysis of chestnut blight fungus, *Cryphonectria parasitica*, under tannic-acid-inducing and hypovirus-regulating conditions. Can J Microbiol 58:863–871

Kim SG, Wang Y, Lee KH, Park ZY, Park J, Wu J, Kwon SJ, Lee YH, Agrawal GK, Rakwal R, Kim ST, Kang KY (2012b) In-depth insight into in vivo apoplastic secretome of rice-*Magnaporthe oryzae* interaction. J Proteomics 78C:58–71

Koulman A, Tapper BA, Fraser K, Cao M, Lane GA, Rasmussen S (2007) High-throughput direct-infusion ion trap mass spectrometry: a new method for metabolomics. Rapid Commun Mass Spectrom 21:421–428

Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene 361:13–37

Kumaraswamy KG, Kushalappa AC, Choo TM, Dion Y, Rioux S (2011) Mass spectrometry based metabolomics to identify potential biomarkers for resistance in barley against fusarium head blight (*Fusarium graminearum*). J Chem Ecol 37:846–856

Lee K, Pan JJ, May G (2009) Endophytic *Fusarium verticillioides* reduces disease severity caused by *Ustilago maydis* on maize. FEMS Microbiol Lett 299:31–37

Lee YJ, Perdian DC, Song Z, Yeung ES, Nikolau BJ (2012) Use of mass spectrometry for imaging metabolites in plants. Plant J 70:81–95

Li B, Wang W, Zong Y, Qin G, Tian S (2012) Exploring pathogenic mechanisms of *Botrytis cinerea* secretome under different ambient pH based on comparative proteomic analysis. J Proteome Res 11:4249–4260

Liang Y, Rahman MH, Strelkov SE, Kav NN (2010a) Developmentally induced changes in the sclerotial proteome of *Sclerotinia sclerotiorum*. Fungal Biol 114:619–627

Liang Y, Strelkov SE, Kav NN (2010b) The proteome of liquid sclerotial exudates from *Sclerotinia sclerotiorum*. J Proteome Res 9:3290–3298

Liu Z, Zhang Z, Faris JD, Oliver RP, Syme R, McDonald MC, McDonald BA, Solomon PS, Lu S, Shelver WL, Xu S, Friesen TL (2012) The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. PLoS Pathog 8:e1002467

Lowe RG, Lord M, Rybak K, Trengove RD, Oliver RP, Solomon PS (2008) A metabolomic approach to dissecting osmotic stress in the wheat pathogen *Stagonospora nodorum*. Fungal Genet Biol 45:1479–1486

Lowe RG, Allwood JW, Galster AM, Urban M, Daudi A, Canning G, Ward JL, Beale MH, Hammond-Kosack KE (2010) A combined $^{(1)}$H nuclear magnetic resonance and electrospray ionization-mass spectrometry analysis to understand the basal metabolism of plant-pathogenic *Fusarium* spp. Mol Plant Microbe Interact 23:1605–1618

Luster DG, McMahon MB, Carter ML, Fortis LL, Nunez A (2010) Proteomic analysis of germinating urediniospores of *Phakopsora pachyrhizi*, causal agent of Asian soybean rust. Proteomics 10:3549–3557

Magan N, Cayley GR, Lacey J (1984) Effect of water activity and temperature on mycotoxin production by *Alternaria alternata* in culture and on wheat grain. Appl Environ Microbiol 47:1113–1117

Mann M, Jensen ON (2003) Proteomic analysis of post-translational modifications. Nat Biotechnol 21:255–261

Manning VA, Chu AL, Steeves JE, Wolpert TJ, Ciuffetti LM (2009) A host-selective toxin of *Pyrenophora tritici-repentis*, Ptr ToxA, induces photosystem changes and reactive oxygen species accumulation in sensitive wheat. Mol Plant Microbe Interact 22:665–676

Manteau S, Abouna S, Lambert B, Legendre L (2003) Differential regulation by ambient pH of putative

virulence factor secretion by the phytopathogenic fungus *Botrytis cinerea*. FEMS Microbiol Ecol 43:359–366

Mardis ER (2006) Anticipating the 1,000 dollar genome. Genome Biol 7:112

Mendgen K, Hahn M (2002) Plant infection and the establishment of fungal biotrophy. Trends Plant Sci 7:352–356

Moco S, Bino RJ, De Vos RCH, Vervoort J (2007) Metabolomics technologies and metabolite identification. Trends Anal Chem 26:855–866

Morais do Amaral A, Antoniw J, Rudd JJ, Hammond-Kosack KE (2012) Defining the predicted protein secretome of the fungal wheat leaf pathogen *Mycosphaerella graminicola*. PLoS One 7:e49904

Motteram J, Lovegrove A, Pirie E, Marsh J, Devonshire J, van de Meene A, Hammond-Kosack K, Rudd JJ (2011) Aberrant protein *N*-glycosylation impacts upon infection-related growth transitions of the haploid plant-pathogenic fungus *Mycosphaerella graminicola*. Mol Microbiol 81:415–433

Mueller O, Kahmann R, Aguilar G, Trejo-Aguilar B, Wu A, de Vries RP (2008) The secretome of the maize pathogen *Ustilago maydis*. Fungal Genet Biol 45 (Suppl 1):S63–S70

Nielsen KF, Smedsgaard J (2003) Fungal metabolite screening: database of 474 mycotoxins and fungal metabolites for dereplication by standardised liquid chromatography-UV-mass spectrometry methodology. J Chromatogr A 1002:111–136

Oliver SG, Winson MK, Kell DB, Baganz F (1998) Systematic functional analysis of the yeast genome. Trends Biotechnol 16:373–378

Pandelova I, Betts MF, Manning VA, Wilhelm LJ, Mockler TC, Ciuffetti LM (2009) Analysis of transcriptome changes induced by Ptr ToxA in wheat provides insights into the mechanisms of plant susceptibility. Mol Plant 2:1067–1083

Parker D, Beckmann M, Zubair H, Enot DP, Caracuel-Rios Z, Overy DP, Snowdon S, Talbot NJ, Draper J (2009) Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by *Magnaporthe grisea*. Plant J 59:723–737

Pedras MS, Chumala PB, Jin W, Islam MS, Hauck DW (2009) The phytopathogenic fungus *Alternaria brassicicola*: phytotoxin production and phytoalexin elicitation. Phytochemistry 70:394–402

Peluffo L, Lia V, Troglia C, Maringolo C, Norma P, Escande A, Esteban Hopp H, Lytovchenko A, Fernie AR, Heinz R, Carrari F (2010) Metabolic profiles of sunflower genotypes with contrasting response to *Sclerotinia sclerotiorum* infection. Phytochemistry 71:70–80

Perpetua NS, Kubo Y, Yasuda N, Takano Y, Furusawa I (1996) Cloning and characterization of a melanin biosynthetic *THR1* reductase gene essential for appressorial penetration of *Colletotrichum lagenarium*. Mol Plant Microbe Interact 9:323–329

Pitarch A, Nombela C, Gil C (2008) Cell wall fractionation for yeast and fungal proteomics. Methods Mol Biol 425:217–239

Prados-Rosales R, Luque-Garcia JL, Martinez-Lopez R, Gil C, Di Pietro A (2009) The *Fusarium oxysporum* cell wall proteome under adhesion-inducing conditions. Proteomics 9:4755–4769

Proctor RH, Hohn TM, McCormick SP (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene. Mol Plant Microbe Interact 8:593–601

Rampitsch C, Subramaniam R, Djuric-Ciganovic S, Bykova NV (2010) The phosphoproteome of *Fusarium graminearum* at the onset of nitrogen starvation. Proteomics 10:124–140

Rampitsch C, Tinker NA, Subramaniam R, Barkow-Oesterreicher S, Laczko E (2012) Phosphoproteome profile of *Fusarium graminearum* grown in vitro under nonlimiting conditions. Proteomics 12:1002–1005

Rodriguez Estrada AE, Hegeman A, Kistler HC, May G (2011) In vitro interactions between *Fusarium verticillioides* and *Ustilago maydis* through real-time PCR and metabolic profiling. Fungal Genet Biol 48:874–885

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169

Sarrazin E, Dubourdieu D, Darriet P (2007) Characterization of key aroma compounds of botrytized wines, influence of grape botrytization. Food Chem 59:7237–7425

Schmidt SM, Panstruga R (2011) Pathogenomics of fungal plant parasites: what have we learnt about pathogenesis? Curr Opin Plant Biol 14:392–399

Scott PM (2001) Analysis of agricultural commodities and foods for *Alternaria* mycotoxins. J AOAC Int 84:1809–1817

Shah P, Atwood JA, Orlando R, El Mubarek H, Podila GK, Davis MR (2009a) Comparative proteomic analysis of *Botrytis cinerea* secretome. J Proteome Res 8:1123–1130

Shah P, Gutierrez-Sanchez G, Orlando R, Bergmann C (2009b) A proteomic study of pectin-degrading enzymes secreted by *Botrytis cinerea* grown in liquid culture. Proteomics 9:3126–3135

Shenton MR, Berberich T, Kamo M, Yamashita T, Taira H, Terauchi R (2012) Use of intercellular washing fluid to investigate the secreted proteome of the rice-*Magnaporthe* interaction. J Plant Res 125:311–316

Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Boucherie H, Mann M (1996) Linking genome and proteome by

mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci U S A 93:14440–14445

Solomon PS, Tan KC, Oliver RP (2003) The nutrient supply of pathogenic fungi; a fertile field for study. Mol Plant Pathol 4:203–210

Solomon PS, Tan KC, Sanchez P, Cooper RM, Oliver RP (2004) The disruption of a Ga subunit sheds new light on the pathogenicity of *Stagonospora nodorum* on wheat. Mol Plant Microbe Interact 17:456–466

Solomon PS, Tan KC, Oliver RP (2005) Mannitol 1-phosphate metabolism is required for sporulation in planta of the wheat pathogen *Stagonospora nodorum*. Mol Plant Microbe Interact 18:110–115

Solomon PS, Waters OD, Jorgens CI, Lowe RG, Rechberger J, Trengove RD, Oliver RP (2006) Mannitol is required for asexual sporulation in the wheat pathogen *Stagonospora nodorum* (glume blotch). Biochem J 399:231–239

Son H, Lee J, Lee YW (2012) Mannitol induces the conversion of conidia to chlamydospore-like structures that confer enhanced tolerance to heat, drought, and UV in *Gibberella zeae*. Microbiol Res 167:608–615

Song X, Rampitsch C, Soltani B, Mauthe W, Linning R, Banks T, McCallum B, Bakkeren G (2011) Proteome analysis of wheat leaf rust fungus, *Puccinia triticina*, infection structures enriched for haustoria. Proteomics 11:944–963

Sorensen JL, Hansen FT, Sondergaard TE, Staerk D, Lee TV, Wimmer R, Klitgaard LG, Purup S, Giese H, Frandsen RJ (2012) Production of novel fusarielins by ectopic activation of the polyketide synthase 9 cluster in *Fusarium graminearum*. Environ Microbiol 14:1159–1170

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stuber K, Ver Loren, van Themaat E, Brown JK, Butcher SA, Gurr SJ, Lebrun MH, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjdia M, Bidzinski P, Bindschedler LV, Both M, Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, Lopez-Ruiz FJ, Lu X, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H, Reinhardt R, Rott M, Sacristan S, Schmidt SM, Schon M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M, Wessling R, Wicker T, Panstruga R (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543–1546

Stahl PL, Lundeberg J (2012) Toward the single-hour high-quality genome. Annu Rev Biochem 81:359–378

Sweigard JA, Carroll AM, Farrall L, Chumley FG, Valent B (1998) *Magnaporthe grisea* pathogenicity genes obtained through insertional mutagenesis. Mol Plant Microbe Interact 11:404–412

Talbot NJ, Ebbole DJ, Hamer JE (1993) Identification and characterization of *MPG1*, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. Plant Cell 5:1575–1590

Tan KC, Heazlewood JL, Millar AH, Thomson G, Oliver RP, Solomon PS (2008) A signaling-regulated, short-chain dehydrogenase of *Stagonospora nodorum* regulates asexual development. Eukaryot Cell 7:1916–1929

Tan KC, Heazlewood JL, Millar AH, Oliver RP, Solomon PS (2009a) Proteomic identification of extracellular proteins regulated by the Gna1 Ga subunit in *Stagonospora nodorum*. Mycol Res 113:523–531

Tan KC, Ipcho SV, Trengove RD, Oliver RP, Solomon PS (2009b) Assessing the impact of transcriptomics, proteomics and metabolomics on fungal phytopathology. Mol Plant Pathol 10:703–715

Tan KC, Trengove RD, Maker GL, Oliver RP, Solomon PS (2009c) Metabolite profiling identifies the mycotoxin alternariol in the pathogen *Stagonospora nodorum*. Metabolomics 5:330–335

ten Have A, Mulder W, Visser J, van Kan JA (1998) The endopolygalacturonase gene *Bcpg1* is required for full virulence of *Botrytis cinerea*. Mol Plant Microbe Interact 11:1009–1016

Thomma BPHJ, Van Esse HP, Crous PW, De Wit PJGM (2005) Cladosporium fulvum (syn. Passalora fulva), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. Mol Plant Pathol 6:379–393

Uhlig S, Busman M, Shane DS, Ronning H, Rise F, Proctor R (2012) Identification of early fumonisin biosynthetic intermediates by inactivation of the *FUM6* gene in *Fusarium verticillioides*. J Agric Food Chem 60:10293–10301

Van De Wouw AP, Howlett BJ (2011) Fungal pathogenicity genes in the age of 'omics'. Mol Plant Pathol 12:507–514

Viaud MC, Balhadere PV, Talbot NJ (2002) A *Magnaporthe grisea* cyclophilin acts as a virulence determinant during plant infection. Plant Cell 14:917–930

Vincent D, Balesdent MH, Gibon J, Claverol S, Lapaillerie D, Lomenech AM, Blaise F, Rouxel T, Martin F, Bonneu M, Amselem J, Dominguez V, Howlett BJ, Wincker P, Joets J, Lebrun MH, Plomion C (2009) Hunting down fungal secretomes using liquid-phase IEF prior to high resolution 2-DE. Electrophoresis 30:4118–4136

Vincent D, Du Fall LA, Livk A, Mathesius U, Lipscombe RJ, Oliver RP, Friesen TL, Solomon PS (2012a) A functional genomics approach to dissect the mode of action of the *Stagonospora nodorum* effector protein SnToxA in wheat. Mol Plant Pathol 13:467–482

Vincent D, Tan KC, Cassidy L, Solomon PS, Oliver RP (2012b) Proteomic techniques for plant-fungal interactions. Methods Mol Biol 835:75–96

Wall ME, Raghavan S, Cohn JD, Dunbar J (2011) Genome majority vote improves gene predictions. PLoS Comput Biol 7:e1002284

Walsh C (2005) Postranslational modification of proteins: expanding nature's inventory. Roberts, Greenwood Village

Wang C, Zhang S, Hou R, Zhao Z, Zheng Q, Xu Q, Zheng D, Wang G, Liu H, Gao X, Ma JW, Kistler HC, Kang Z, Xu JR (2011a) Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. PLoS Pathog 7:e1002460

Wang Y, Wu J, Park ZY, Kim SG, Rakwal R, Agrawal GK, Kim ST, Kang KY (2011b) Comparative secretome investigation of *Magnaporthe oryzae* proteins responsive to nitrogen starvation. J Proteome Res 10:3136–3148

Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I (1995) Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. Electrophoresis 16:1090–1094

Watson JT, Sparkman OD (2007) Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation. Wiley, Chichester

White MJ, Lopes JM, Henry SA (1991) Inositol metabolism in yeasts. Adv Microb Physiol 32:1–51

Wishart DS (2008) Quantitative metabolomics using NMR. Trends Anal Chem 27:228–236

Wisniewski JR, Zougman A, Mann M (2009) Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J Proteome Res 8:5674–5678

Wolpert TJ, Dunkle LD, Ciuffetti LM (2002) Host-selective toxins and avirulence determinants: what's in a name? Annu Rev Phytopathol 40:251–285

Yang F, Jensen JD, Spliid NH, Svensson B, Jacobsen S, Jorgensen LN, Jorgensen HJ, Collinge DB, Finnie C (2010a) Investigation of the effect of nitrogen on severity of *Fusarium* head blight in barley. J Proteomics 73:743–752

Yang F, Jensen JD, Svensson B, Jorgensen HJ, Collinge DB, Finnie C (2010b) Analysis of early events in the interaction between *Fusarium graminearum* and the susceptible barley (*Hordeum vulgare*) cultivar Scarlett. Proteomics 10:3748–3755

Yang F, Jensen JD, Svensson B, Jorgensen HJ, Collinge DB, Finnie C (2012) Secretomics identifies *Fusarium graminearum* proteins involved in the interaction with barley and wheat. Mol Plant Pathol 13:445–453

Yu JH, Keller N (2005) Regulation of secondary metabolism in filamentous fungi. Annu Rev Phytopathol 43:437–458

Zhu H, Bilgin M, Snyder M (2003) Proteomics. Annu Rev Biochem 72:783–812

13 Functional Genomics to Characterize Opportunistic Pathogens

Thorsten Heinekamp[1,2], Vito Valiante[1,2], Nora Koester-Eiserfunke[1,2], Axel A. Brakhage[1,2]

CONTENTS

I. Introduction: Functional Genomics and Opportunistic Fungal Pathogens

More than 60 years after the structure of the DNA helix was published, functional genomics became an everyday tool for scientists. In the beginning of the 2000s, revolutions in sequencing technologies and other high-throughput applications paved the way toward experimental approaches that seemed futuristic even a few years ago.

Under the term *functional genomics*, all the state-of-the-art high-throughput and "-omics" technologies, such as genomics, transcriptomics, proteomics, and metabolomics, are combined with bioinformatic analyses. Functional genomics not only aims at assigning a function to pieces of DNA but also includes the elucidation of the organization and control of genetic pathways that together determine the physiology of an organism (Fig. 13.1). An appropriate definition for functional genomics was given by the European Bioinformatics Institute (EBI): "Functional genomics is a field of molecular biology that enables you to explore genes, protein functions and interactions on a global scale. Functional genomic experiments typically utilize large-scale assays to measure and track many genes or proteins in parallel under different environmental conditions. This approach allows the DNA function to be discovered by combining information from genes, transcripts and proteins" (http://www.ebi.ac.uk). All the high-throughput technologies connected to functional genomics are under constant development, ensuring the long-term relevance of this approach.

At the time the human genome sequence was published, in an editorial introduction into the "Functional Genomics" issue of *Nature* Insights (Patterson and Dant 2000), it was claimed that "An inventory of genes will impact molecular medicine the greatest, leading to improved diagnosis of disease and eventually to custom-made drugs tailored to the individual." However, this concept proved to be too simple for humans and human diseases, which

[1]Department of Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knoell Institute (HKI), Jena, Germany; e-mail: axel.brakhage@hki-jena.de
[2]Department of Molecular and Applied Microbiology, Friedrich Schiller University, Jena, Germany; e-mail: axel.brakhage@hki-jena.de

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

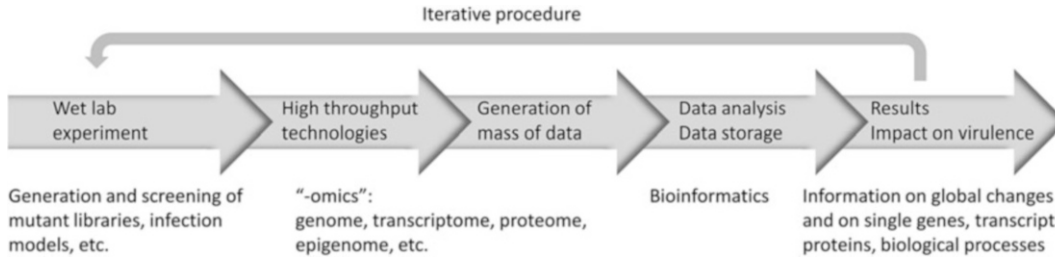


Fig. 13.1. Flowchart for a functional genomics approach on how information derived from several high-throughput experiments can be combined to gain knowledge of impact of both single genes and global biological processes on fungal pathogenicity

are often associated with several gene defects. Nevertheless, the concept of a functional genomics approach can be transferred to other organisms, especially to (pathogenic) microorganisms. Knowledge of the DNA sequence of microorganisms and the application of functional genomics is a promising tool for understanding the interplay between pathogens and host cells, improving diagnosis, identifying new drug targets, and developing tailor-made drugs.

Although in the last decades infections with pathogenic fungi, resulting in considerable morbidity and mortality, remarkably increased, **human-pathogenic fungi** are still an "underestimated" problem (Brown et al. 2012). Today, it is generally accepted that the total number of fungal species amounts to more than 1.5 million. Although the vast majority still await discovery, not to mention characterization, currently only a minority of about 100 fungal species are regarded as human pathogenic. Fungal pathogens cause a wide range of diseases, from allergies and superficial infections to life-threatening invasive mycoses. Often, the outcome of a fungal infection depends on the **immune status of the host organism**. Especially, humans with a compromised immune system represent a high-risk group for developing a fatal fungal infection. The continuous progress in medicine (e.g., in chemotherapy and organ or bone marrow transplantation) contributes to the ongoing increasing number of patients with impaired immune status.

The most important causes of life-threatening invasive mycoses are infections caused by the yeast ***Candida albicans*** (see also Chap. 14 in this volume) and the filamentous fungus ***Aspergillus fumigatus***. Besides *A. fumigatus*, only a few other members of the genus *Aspergillus*, which has more than 200 species, are considered opportunistic human pathogens; among them are, for example, *A. terreus*, *A. flavus*, and *A. niger* (Brakhage 2005). Among the *Candida* species, *C. albicans* prevails in more than 50 % of all clinical *Candida* samples, followed by infections with *C. glabrata*. Other pathogenic *Candida* species, such as *C. tropicalis*, *C. dubliniensis*, *C. krusei*, and *C. parapsilosis*, cause far fewer infections (Pfaller and Diekema 2007). The most common fungal infection among AIDS patients, cryptococcal meningitis, is caused by the basidiomycete *Cryptococcus neoformans*, which represents another opportunistic human-pathogenic fungus of clinical relevance (La Hoz and Pappas 2013). Furthermore, other fungal species, such as Zygomycetes and *Fusarium* species, recently emerged as causal agents of invasive mycoses (Park et al. 2011).

To date, neither reliable diagnostic tools nor effective treatment options are available, resulting in unacceptably high mortality rates of patients suffering from invasive fungal infections. In general, research on pathogenic fungi includes the following aims:

1. Identification of virulence determinants for each pathogen
2. Elucidation of the role of components of the host's immune system (epithelial barriers, innate immunity, adaptive immunity)
3. Identification of immune evasion strategies
4. Deciphering the complex mechanisms of the infection process

5. Identification of common principles of fungal pathogenesis

The following mainly focuses on infections caused by *A. fumigatus* as this is to date the number one airborne fungal pathogen. We will, of course, also include other common fungal pathogens to show how functional genomics approaches can aid achieving these aims and help find new therapeutic options.

II. Genomics

A first global approach of functional genomics is genome sequencing. Next-generation sequencing (NGS) now allows sequencing the genomes of many fungi, including a remarkable number of pathogenic fungi. This broad approach gives insight into the genomic basis underlying the diverse mechanisms by which pathogenic fungi infect their host. NGS provides large amounts of sequence data in a fast and affordable manner. The resulting genomes can be screened for virulence-associated genes by various approaches. For example, genomes of members of a group of fungi that can cause disease can be compared to nonpathogenic representatives. By this comparative approach, genes can be identified that contribute to host or disease specificity, whereas other genes are attributed as of key importance for general virulence functions. With regard to the aspergilli, the results of comparative genomics are discussed in detail next.

A. Comparative Genomics of Aspergilli

Among all systemic *Aspergillus* infections, *A. fumigatus* causes 90 %, indicating that this species produces virulence determinants in favor of becoming an opportunistic human pathogen. From the year 2005 to date, a total of **15 *Aspergillus* genomes** were sequenced and made publicly available via Internet-based resources, such as the Central *Aspergillus* Data Repository (CADRE; http://www.cadre-genomes.org.uk) (Mabey et al. 2004) or the *Aspergillus* Genome Database (*Asp*GD) (Arnaud et al. 2010). In the past few years, the *Asp*GD (http://www.aspergillusgenome.org) developed the most comprehensive data collection. Among sequenced aspergilli are the model organism *A. nidulans*, the industrially important *A. niger* and *A. oryzae*, and the pathogenic fungi *A. flavus*, *A. terreus*, and of course, *A. fumigatus*. Similar to other genome databases (yeast genome database, *Candida* genome database, etc.), *Asp*GD not only contains genome sequences but also all available information on genes (e.g., exon-intron structure) and proteins (deduced function, cellular localization, links to relevant literature). The online browser tool Sybil (accessible via http://www.aspgd.org) allows multispecies comparative genomics to determine orthology and synteny across multiple sequenced *Aspergillus* species.

The first comparative study of the *A. fumigatus* genome with the genomes of its distant relatives *A. nidulans* and *A. oryzae* showed an unexpected genetic variability of these species that is reflected most notably by a broad reorganization of the genome structure (Galagan et al. 2005). An average amino acid identity of less than 70 % between each species pair implies that these species are evolutionarily as distant from each other as mammals and fish, which diverged 450 million years ago.

In another study, two clinical isolates of *A. fumigatus*, Af293 and A1183, were compared with their close relatives *Neosartorya fischeri* and *A. clavatus* (Fedorova et al. 2008). The genomic comparison of the two *A. fumigatus* isolates revealed that 98 % of the genomes can be aligned. The nucleotide sequence is nearly identical in the shared regions. However, 143 and 218 genes were assigned to be isolate specific for Af293 and A1163, respectively (Fedorova et al. 2008). These isolate-specific genes were preferentially located within 300 kb from chromosome ends, and many of these species-specific genes are clustered in highly variable segments containing repeat elements and pseudogenes (Fedorova et al. 2008). It is hypothesized that species-specific genes represent, at least in part, the underlying cause for isolate-specific differences in antifungal susceptibility.

For more than 800 *A. fumigatus* genes, no orthologs were identifiable in *A. clavatus* and *N. fischeri*, therefore representing species-specific genes. Remarkably, these genes, which

are predominantly located in subtelomeric regions, have only bacterial or archaeal homologs and may contribute to the adaptation of the fungus to environmental niches, such as the human lung (Fedorova et al. 2008). It is intriguing to speculate that the subtelomeric diversity constitutes the potential driving force in the evolution of pathogenicity in *A. fumigatus*. This hypothesis is corroborated by the finding that transcription of one third of the *A. fumigatus* subtelomeric and species-specific genes is induced during initiation of invasive aspergillosis (McDonagh et al. 2008).

A good example for conservation of genes between aspergilli is given by those genes encoding components of signal transduction cascades. The cell wall integrity regulating mitogen-activated protein kinase MpkA proved to be highly conserved (Valiante et al. 2008, 2009; Rispail et al. 2009). The same applies to cyclic AMP (cAMP) signaling components, such as adenylyl cyclase and protein kinase A catalytic subunit (Liebmann et al. 2003), as well as genes involved in calcium signaling via calcineurin (Steinbach et al. 2006; da Silva Ferreira et al. 2007).

In addition to the comprehensive analysis of the genomes of different aspergilli, there is an an increasing number of comparative genomics approaches to other opportunistic pathogenic fungi, such as *C. neoformans* and *Candida* spp. *Cryptococcus neoformans* is an opportunistic pathogen of the immunosuppressed host; it can infect the central nervous system, leading to fatal meningitis if untreated. In AIDS patients, *C. neoformans* is a clinically relevant cause of mortality. The fungus is especially known to cause relapse in patients surviving an initial cryptococcal meningoencephalitis. Comparative genomics of serial isolates of *C. neoformans* was performed to unravel the underlying mechanisms, especially the impact of in-host microevolution on persistence (Ormerod et al. 2013). Sequencing the whole genome of these serial isolates allowed deducing the phenotypical differences observed with regard to metabolic profiles, nutrient acquisition, and virulence in a murine model to only a few genomic differences. Two identified key mutations, loss of an interaction domain protein and changes in the copy number of the arms of chromosome 12, apparently explain the phenotype of the relapse isolate, which is hypothesized to have evolved in-host where antifungal treatment selected for the ability to persist (Ormerod et al. 2013).

B. Genomic Mining to Identify Secondary Metabolite Gene Clusters

Fungi produce a remarkable variety of secondary metabolites. These natural products comprise a broad range of effective drugs (i.e., antibiotics and immunosuppressive agents). However, several secondary metabolites also display strong toxic activities. For the majority of all known secondary metabolites, their function in the producing organism is not known yet (Brakhage 2013). As biologically active compounds, they might protect the fungus against other (micro-)organisms in the same habitat and may also contribute to weaken the host's immune system (Askew 2008; Brakhage et al. 2008, 2009). Consistently, several fungal secondary metabolites (e.g., some toxins and pigments) contribute to pathogenicity.

As genes for biosynthesis of secondary metabolites are usually organized in **clusters**, the availability of complete fungal genome sequences allows screening for the presence of such clusters. Besides characteristic core genes (nonribosomal peptide synthetases, NRPSs; polyketide synthases, PKSs; hybrid PKS/NRPS; dimethylallyl tryptophan synthases, DMATs) relevant for metabolite production, these clusters often contain genes with a putative role in self-resistance and metabolite tailoring (Brakhage 2013). Often, genes involved in transcriptional regulation are also part of the gene cluster. For **genome mining** and the prediction of putative secondary metabolite gene clusters, web-based tools are available, such as SMURF (Secondary Metabolite Unique Regions Finder; http://jcvi.org/smurf) (Khaldi et al. 2010). For example, SMURF predicts an overall number of 30 secondary metabolite gene clusters in *A. fumigatus* Af293. For the model organism *A. nidulans*, a gene expression compendium was developed by applying a newly designed DNA expression array in combination with previously collected data, thereby allowing accurate

determination of secondary metabolite gene cluster members. As this approach is based on -omics information (annotated genome sequences and a catalog of gene expression) that are easily available, it is also applicable to other fungi (Andersen et al. 2013).

One of the best-studied secondary metabolites produced by *A. fumigatus* is **gliotoxin**. This infamous virulence determinant belongs to the class of epipolythiodioxopiperazine (ETP) toxins (Gardiner and Howlett 2005). Putative ETP toxin biosynthesis clusters are widely distributed in ascomycete fungi (Patron et al. 2007). Gliotoxin is produced during the infection process and exerts immune modulatory functions and cell-damaging activity (Scharf et al. 2012). Gliotoxin also is highly toxic to human immune cells. Identification of the gliotoxin biosynthesis gene cluster was possible because the genome of *A. fumigatus* was available. Recently, based on the genome sequence, the biosynthesis of gliotoxin was elucidated in detail (Scharf et al. 2010, 2011, 2012; Schrettl et al. 2010; Davis et al. 2011; Gallagher et al. 2012). Remarkably, the fungus was found to be self-protected against gliotoxin by function of a specific gliotoxin oxidase, GliT, and by production of a gliotoxin transporter. These proteins might represent targets for new drugs to render the fungus sensitive against its own toxin.

Another secondary metabolite decisive for virulence of *A. fumigatus* is the conidial gray-green pigment **dihydroxynaphthalene (DHN) melanin**. The key enzyme of DHN melanin biosynthesis is the polyketide synthase PksP. PksP-deficient mutants producing white conidia are attenuated in virulence (Langfelder et al. 1998). Expression of the gene occurs predominantly during conidiogenesis but was also found in hyphae isolated from infected mice lung tissue (Langfelder et al. 2001). DHN melanin protects *A. fumigatus* against host-derived reactive oxygen intermediates (ROIs) (Jahn et al. 2000; Langfelder et al. 2003) and was recently identified as crucial for inhibition of acidification of phagolysosomes of phagocytes and thereby killing of *A. fumigatus* conidia (Thywißen et al. 2011). Melanin is also responsible for interfering with macrophage apoptosis (Volling et al. 2011). Therefore, melanin provides the fungus a universal tool to survive the residual host immune response to enable a fatal infection.

In general, the repertoire of secondary metabolite biosynthesis gene clusters differs remarkably from species to species. Many clusters were found that are unique to only one *Aspergillus* species. Although they are evolutionarily closely related, *A. fumigatus* shares only 14 orthologous clusters with *N. fischeri* (including the gliotoxin cluster) and 5 with *A. clavatus*.

To date, a wide variety of structurally different secondary metabolites has been isolated from different fungi. However, the actual assignment of a specific metabolite to its biosynthesis gene cluster is still the exception. This also applies to the detailed analysis of the metabolite biosynthesis steps and the role of each enzyme predicted to be part of the cluster. To close this gap, that is, to assign gene clusters to their corresponding metabolite, several strategies were implemented only recently (Chiang et al. 2009; Brakhage 2013). Simple knockout strategies and subsequent metabolite analysis turned out to be hardly successful as, under normal laboratory cultivation conditions, the vast majority of secondary metabolite gene cluster genes are silent per se; that is, no detectable amounts of transcripts are produced (Bergmann et al. 2007; Brakhage 2013). However, as genes encoding transcription factors often are part of the predicted gene cluster, an efficient approach was developed by (inducible) overproduction of the respective transcription factor and thereby induced transcription of all genes of the cluster (Bergmann et al. 2007). Only recently, genome mining-based genetic modification of aspergilli resulted in the identification of new secondary metabolites, pheofungins and neosartoricins (Scherlach et al. 2011; Chooi et al. 2013). Activation of the originally silent gene cluster enabled elucidation of the structure and biosynthesis of these polyketides.

A more global strategy to interfere with the activity of many secondary metabolite gene clusters is based on the **coregulation at the chromatin level**. Therefore, another approach to activate silent gene clusters is the reorganization of eu- or heterochromatin regions (Bok

and Keller 2004; Lee et al. 2009; Brakhage 2013). Albeit mostly analyzed in *A. nidulans*, an impact of chromatin regulation on secondary metabolite cluster gene expression has been highlighted in several studies directly investigating secondary metabolite regulation in aspergilli (Palmer and Keller 2010; Nutzmann et al. 2011; Gacek and Strauss 2012; Brakhage 2013). Normally in a repressed state, several of such clusters become derepressed under particular, often poorly defined, culturing conditions. Many cluster genes are, for instance, expressed in *A. fumigatus* under infectious conditions in a mouse infection model (McDonagh et al. 2008), suggesting their impact on virulence. Interestingly, for studied examples, the neighboring genes remained unaffected on activation of normally silenced cluster genes, which imposes the suggestion of cluster arrangements into functionally confined chromatin domain entities. Investigating *A. fumigatus* deletion mutants of genes encoding the histone deacetylase HdaA or a subunit (CclA) of the H3K4 methylating Set1/COMPASS complex revealed in both cases deregulation of a couple of secondary metabolite cluster genes, which could also be demonstrated for the respective mutants in *A. nidulans* (Shwab et al. 2007; Bok et al. 2009; Lee et al. 2009; Palmer et al. 2013). In addition, the effector protein HepA and the histone methyl transferase ClrD have been shown to be important factors for silencing of several secondary metabolite clusters in *A. nidulans* (Reyes-Dominguez et al. 2010). In line with the general suggestion that many histone deacetylases enforce the repression of genes because of the removal of acetyl moieties on lysines, treatment of growing fungal cultures with histone deacetylase inhibitors caused derepression of several secondary metabolite cluster genes (Cichewicz 2010).

Positive effects on cluster gene regulation could be adjudged to the *A. nidulans* histone acetyl transferase EsaA as part of the NuA4 complex and corresponding acetylation sites on histone H4 (Soukup et al. 2012). Likewise, two members of the well-conserved SAGA/Ada acetylation complex, the histone acetyl transferase GcnE (Gcn5 in other organisms) and its complex partner AdaB were shown to be responsible for expression of a gene cluster required for orsellinic acid production as well as further gene clusters under induction conditions (Nutzmann et al. 2011). In a mutant screen, Δ*gcnE* was identified as the sole deletion from a library of all viable potential histone acetyl transferase mutants in *A. nidulans* (36 of 40) to be essential for orsellinic acid production on coincubation with *Streptomyces rapamycinicus* (Nutzmann et al. 2011). Further, the authors showed GcnE-dependent increased acetylation levels of H3K14 on coincubation with the streptomycete at all tested promoter regions but an increase of H3K9ac restricted to promoter regions of induced cluster genes. Again for *A. nidulans*, Shimizu et al. observed elevated H4K16ac levels at selected secondary metabolite gene promoters concurrent with derepression of the corresponding genes under hypoxia. This phenotype might be explained by a low cellular NAD^+ level under hypoxia and inhibition of NAD^+-dependent histone deacetylase activity of the *A. nidulans* Sir2 homolog SirA (Shimizu et al. 2012).

These findings in the nonpathogenic model organism *A. nidulans* are currently the basis for further investigation of the role of epigenetics in regulating secondary metabolite production and the impact on pathogenicity of *A. fumigatus* and other pathogenic fungi.

III. Generation of Fungal Mutant Strains to Study Gene Function

The basal approach in genetics to validate gene functions in general and to study pathogenesis in particular is the analysis of mutants. Today, genomic technologies facilitate, on the one hand, the identification of the mutated locus for mutants that were generated by random mutagenesis. On the other hand, based on the available genome sequence, targeted gene deletion is now a straightforward approach to create mutant strains of a gene of interest. A general approach is **random mutagenesis** by treating the fungus with DNA-mutagenic agents (UV [ultraviolet] irradiation, chemicals). Another possibility is **insertional mutagenesis**,

in which a DNA fragment is inserted randomly into the genome. A more sophisticated approach is **signature-tagged mutagenesis** by random insertion of a DNA fragment, harboring a 40-bp unique sequence tag. Such tagged mutants can be pooled and analyzed for pathogenicity in animal studies. The identification of mutations in genes essential for virulence is done by recovering the mutants from the host and quantifying the amount of each sequence tag. Underrepresented tags represent mutations in virulence-associated genes. Using sequence-tagged mutant pools therefore offers the possibility for an in vivo virulence screen. For *A. fumigatus*, this method was successfully adopted, and by screening a pool of 4,648 strains in a mouse model of invasive pulmonary aspergillosis, the *pabaA* gene was identified as essential for pathogenicity (Brown et al. 2000). Therefore, the folate biosynthesis pathway in which the *pabaA*-encoded *para*-amino-benzoic acid synthetase is involved is of importance for in vivo survival of the fungus during infection.

For the yeast pathogen *C. glabrata*, a pool of 4,800 sequence-tagged mutants was screened in vitro for their ability to adhere to human epithelial cells (Cormack et al. 1999). Several mutants were identified because of their reduction or loss in adherence; in most cases, insertion into the EPA1 gene, encoding a cell surface protein, was found. However, although the EPA1 mutant was nonadherent in vitro, it was fully virulent in mice, suggesting that in vivo additional proteins mediate adhesion (Cormack et al. 1999).

Screening a library of 18,350 *C. glabrata* Tn*7* insertion mutants for their interaction with human THP-1 macrophages, 56 genes were identified that were essential for survival (Rai et al. 2012). In a functional genomics approach, the authors characterized chromatin remodeling as a central mechanism for regulating survival strategies of phagozytosed *C. glabrata*. In a response to exposure to the intracellular milieu of the macrophage, yeast cells thereby facilitate a reprogramming of cellular energy metabolism and provide protection against DNA damage.

In contrast to random mutagenesis strategies, **targeted gene deletion** by homologues recombination allows a more systematic approach to construct mutants or even comprehensive mutant libraries. However, targeted gene deletion in fungi is often hampered by their tendency to integrate foreign DNA ectopically into the genome. Even with flanking regions comprising more than 1 kb when *A. fumigatus* protoplasts are transformed, the rate of homologues integration of the transferred construct is often below 5 % of all transformed strains. This frequency can be increased by using a bipartite marker approach by which the selection marker is split into two separately nonfunctional fragments that recombine after transformation via homologues recombination (Nielsen et al. 2006). Employing only polymerase chain reaction (PCR) for generation of such bipartite gene deletion substrates further enhances efficacy as no time-consuming cloning steps are required.

The largest leap forward in increasing targeted gene deletion frequency was made by generation of mutants deficient for nonhomologues end-joining DNA repair of *A. fumigatus* (da Silva Ferreira et al. 2006; Krappmann et al. 2006), facilitating the generation of targeted gene deletion mutants in up to 100 % of all transformants obtained.

Worth mentioning is that for many fungi only a limited repertoire of dominant selection markers is available. This is a limitation with regard to generation of multiple knockout mutants. To overcome this problem, several strategies based on **marker recycling cassettes** have been developed. For example, employing the beta-rec/six site-specific recombination system to generate a self-excising selection marker cassette (Hartmann et al. 2010), an *A. fumigatus* octuple deletion mutant of genes encoding oligopeptide transporters was generated (Hartmann et al. 2011).

The improvements mentioned allowed significantly speeding up the generation of mutants for pathogenic fungi. Where a decade ago only a few mutants were under investigation, today there are several hundreds of mutants for *A. fumigatus*, *C. albicans*, and other human pathogenic fungi available. Even the generation of **comprehensive gene deletion libraries** is now a realistic option. Such whole-genome gene deletion libraries already exist for

Saccharomyces cerevisiae and *Neurospora crassa*, just to name two prominent examples of model organisms (Winzeler et al. 1999; Colot et al. 2006). For at least some pathogenic fungi, gene deletion libraries are on the way; for example, for *C. albicans* a library in which 165 transcription factors were deleted is available for the *Candida* community (Homann et al. 2009). And, a *C. albicans* collection with 674 gene deletions was already successfully screened for virulence in mice and for morphogenic switching (Noble et al. 2010).

Taken together, gene deletion libraries offer an immense opportunity to screen for virulence-related phenotypes and thereby to unveil mechanisms of pathogenesis.

IV. Transcriptomics

Genome sequencing can easily highlight differences among different organisms or related species. However, to identify the dynamics of genes inside a distinct organism, analysis of differentially regulated genes still remains the most powerful tool. Transcriptomics is largely used to study growth phases, developmental stages, stress responses, mutagenesis effects, and every physiological variation that can be found in an organism. Since 1977, scientists have had the possibility to study gene expression using Northern blot analysis (Alwine et al. 1977). This simple technique introduced the possibility to detect gene expression in certain conditions. Of course, the biggest limitation was to repeat the experiment for every single gene studied. However, at this time, scientists flirted with the idea of investigating total gene expression changes.

After the introduction of the PCR (Mullis et al. 1986), new techniques were developed. **Differential display** was introduced in 1992 (Liang and Pardee 1992). This methodology was based on hybridization using DNA probes. Even if this produced a high rate of false positives, for the first time it was possible to investigate differentially expressed genes in a hundreds scale. The first real transcriptomics approach was given by the introduction of sequencing of **expressed sequence tags** (ESTs) (Adams et al. 1993). cDNA fragments were cloned randomly and then sequenced. Even if this technique was more qualitative than quantitative, this approach introduced the possibility of discovering a high number of genes related to defined physiological stages and to analyze differentially expressed genes in the thousands scale.

The first genome-wide transcriptomics approach based on **microarrays** was described in 1995 (Schena et al. 1995). This chip-based technology gave for the first time the possibility to study whole-transcriptome changes in a single experimental approach. Furthermore, the possibility of using customized supports gave the chance to increase the number of analyzed samples. At this stage, bioinformaticians joined lab scientists, starting a bond that becomes increasingly stronger. Analysis of different experimental conditions (several cultivation conditions, time course experiments, etc.) became routine, and computational biology provided major contributions. Because of the availability of fungal genome data, it was possible to design microarrays for genome-wide expression analysis for many pathogenic fungi, making transcriptomics a prevailing tool in infection biology studies. In the beginning of the 2000s, different microarrays were developed for the major pathogenic fungi, such as *A. fumigatus*, *C. albicans*, and *C. neoformans* (De Backer et al. 2001; Kraus et al. 2004; Nierman et al. 2005). Transcriptomics was mainly used to investigate antifungal stress responses. In particular, microarrays have been largely used to study the mode of action of orphan molecules, which potentially can be used for clinical interventions. However, the possibility to genetically manipulate all these species also contributed to the expansion of this technique, promoting high investments in the realization of platforms to have a powerful tool that could be used to discover genes differentially regulated in mutant strains.

Candida spp. represent a group of well-studied pathogens. For over a decade, data on different transcriptome studies of the commensal pathogen *C. albicans* have been publicly accessible. The majority of these experiments

focused on studying gene expression in response to antifungal agents (e.g., azole derivatives, amphotericin B, echinocandins; De Backer et al. 2001; Barker et al. 2004; Liu et al. 2005) or were focused on studying different developmental stages and physiological changes such as biofilm formation (Doedt et al. 2004; Murillo et al. 2005). The high number of available transcriptomics data for *C. albicans* were already used for a comparative gene expression analysis; many of the published data were pooled together and normalized to have a wide expression analysis for this organism (Ihmels et al. 2005). Using a differential clustering approach, the authors were able to compare almost 250 arrays, obtaining regulatory information about individual *C. albicans* Open Reading Frames (ORFs). Within this study, it was possible to check differential expression of a single gene or a group of genes in response to different stimuli. Moreover, they compared this data set with available data sets created to study model organisms such as *S. cerevisiae* and *Schizosaccharomyces pombe* (Alter et al. 2003; Chen et al. 2003). Differential expression of common orthologous genes was analyzed. These kinds of studies are extremely useful because they can highlight important genes that can be used as targets for therapeutic intervention, in particular genes that are conserved and seem to be fundamental for fungal growth.

A. Microarray Versus RNA Sequencing

All studies on functional genomics apply an -omics approach. Most of these studies start from a broader analysis and then turn to a less-complex level. The produced data sets represent just a small part of the biological system. The difficulty is to have a smaller group or pool of genes that can be used to develop a broad theory that could be representative for the system. The standardization of data that need to be applied prior to the comparison still remains a limiting factor. The first problem encountered for data integration concerns the differences among methods. We have access to various protocols that have been created and adapted for different purposes, which contain many variations, such as media composition, used strains or isolates, or incubation temperature. All these differences create barriers that complicate a direct comparison. In addition, there are many manufacturing-derived differences in the employed technology (e.g., microarray chips), which decreases the comparability between the analyzed experiments. One example is given by the different microarray platforms developed to date. Even for the same organism, several microarray platforms made using different probe sequences might exist; that is, different platforms could eventually give different outputs for the same experimental setting.

Transcriptomics studies of *A. fumigatus* are a typical example of the difficulty with the comparison and normalization of different studies. The first *A. fumigatus* transcriptome analysis was published together with the release of the genome sequence (Nierman et al. 2005). In the last few years, several transcriptome studies were performed that focused on iron starvation, biofilm formation, and response to antifungals (da Silva Ferreira et al. 2006; Schrettl et al. 2008; Bruns et al. 2010; Jain et al. 2011). Many of the published studies used different microarray platforms, which were developed independently using different sets of oligonucleotides. A recent work demonstrated that the results of whole-genome transcriptome analyses based on microarrays were highly dependent on the platform used (Muller et al. 2012). The low consistency of microarray gene expression data can have multiple reasons, such as platform effects, different media, RNA extraction protocols, or biological variability of the strain used. Hybridization seems to have the most significant impact. Hybridization strongly depends on the probes that were used to develop the microarrays. However, the hybridization bias seems to be highly conserved within the same platform. One solution would be to have standardized conditions among different experiments, although it seems highly ambitious to commit the research community to only one array platform and to adopt a standardized operation procedure.

However, these discussions have become (more or less) obsolete, as all efforts to develop

and improve microarray technology have been overtaken already by the rise of alternative techniques based on **deep RNA sequencing (RNA-seq)**. RNA-seq is based on reverse transcription of RNA in cDNA, subsequent amplification using customized primers as adapters, and sequencing of the obtained libraries (Bentley et al. 2008). Every sequenced fragment, named "read" (normally ranging between 50 and 200 nucleotides), can be subsequently mapped to a reference genome. The number of reads spanning a single gene determines its expression level. For comparative analyses, all the reads are normalized before calculating the fold change. Statistic tests are then used to determine the validity of the analysis (Wang et al. 2009).

There are several advantages of this new sequencing technology. As this technology is not based on hybridization, the comparability among different experiments is easier. The only critical point remains the gene expression normalization, but running protocols can be used for this purpose. Moreover, by applying a high-coverage setup, it is possible to identify lowly expressed transcripts that could not be detected using microarrays or even quantitative real time (qRT)-PCR. One of the biggest advantages of this new technique is that it potentially can be used to study pathogens during infection. Host-pathogen interaction studies using microarray analyses were always limited. The separation of organism-specific RNA prior to hybridization was hard to achieve, and the number of genes that could be analyzed was restricted. The simultaneous analysis of both interaction partners is becoming more feasible with the recent innovations of RNA-seq analyses and single-cell measurements (Westermann et al. 2012).

During host-pathogen interaction, the easiest would be to extract total RNA from the samples that include both organisms. The total RNA from different species can be analyzed as a pool. The data obtained can be subsequently separated during the analysis by aligning raw sequence data to each genome in separate steps. This approach is possible because the programs that map RNAs-eq reads can flag sequences that do not match with the reference genome. The RNA-seq technique could potentially provide the possibility of monitoring gene expression profiles from the pathogen and the host simultaneously.

However, even if this approach looks promising, it also shows some technical limitations. One of the most important parameters to be considered during an RNA-seq experiment is the applied coverage. To have good statistical analysis of the differentially regulated genes, it is important to choose the length and the number of reads suitable for the analyzed genome. Studies of the number of reads required for a specific genome were performed in *Caenorhabditis elegans* (Hillier et al. 2009), but a general formula is still lacking.

Deep RNA sequencing is going to be a leading approach in transcriptomics profiling. Recently, first studies investigating fungal pathogens with the help of RNA-seq data have been published. Gibbons and colleagues studied differentially expressed genes obtained by planktonic and biofilm growth in *A. fumigatus* (Gibbons et al. 2012). However, in addition to the characterization of genes putatively involved in biofilm formation, it was interesting to notice that the authors were able to identify differentially regulated gene clusters, which are commonly difficult to analyze in a transcriptomics approach.

Another example was given by Muller and coworkers, who investigated the regulatory role of the mitogen-activated protein kinase A encoded by the *mpkA* gene with the help of knockout mutants (Muller et al. 2012). In this study, the data sets obtained were compared with that obtained by Gibbons et al. (2012). Global comparison of transcriptomics data, which included also microarray data, demonstrated a higher correlation between the two RNA-seq data sets. Surprisingly, the authors found that the correlation between these two distinct data sets, which were obtained using two different experimental approaches, was higher than the correlation observed by comparing transcriptomics data obtained by different techniques (microarray vs. RNA-seq) applied to the same experiment. This finding demonstrates that RNA-seq provides the highest level of comparability in terms of gene expression levels in *A. fumigatus*, and that the

possibility of data integration will grow in the future.

V. Proteomics

Proteomic approaches are important tools in several aspects of fungal research. Indeed, occurrence and level of proteins within a cell are to some extent correlated with their corresponding transcript levels, which can be measured by transcriptome analysis, but other dependencies are given (e.g., by protein turnover or translational control). Further emphasizing proteomic approaches, protein localization, or posttranslational modifications cannot be detected on the messenger RNA (mRNA) level.

Two main technologies were established in proteomic research: gel-based and gel-free proteomics. Provided the genome information is available, both routes can be routinely used for mapping and identifying large parts of a microorganism's proteome, which is defined as the complete complement of proteins present within a cell at a given time point and condition (Wilkins et al. 1996). Depending on the question, respective subproteomes can also be analyzed.

Notably, comparative proteomics is often applied. Such studies are used to obtain first insights into factors and potential regulatory interconnections of proteins important for cellular/organismal integrity, function, or adaptation with respect to a particular developmental state, morphotype, or culturing condition. In addition, proteomes of particular deletion or overexpression mutants are analyzed for identification of redundant proteins, regulation circuits, or (alternative) regulation pathways. In a smaller scale, proteomics is, for instance, executed to identify protein complex and interaction partners or particular posttranslational modifications on proteins.

A. Gel-Based Proteomics

During **two-dimensional gel electrophoresis (2D-E)**, complex protein mixtures become resolved by separation of the proteome using two successional electrophoresis steps. For the first dimension, proteins are separated according to their intrinsic isoelectric point (pI) by **isoelectric focusing (IEF)**, generally using acrylamide-based gel strips with embedded immobilized pH gradient (IPG). Protein extract is loaded onto the strips either during rehydration of dry strips or during the process of IEF. IEF is achieved by applying strong electric current, forcing proteins to move within the IPG in the direction of either the cathode or the anode as a function of their net charge, which is directly dependent on protein sequence and environmental pH. Once the proteins reach a position within the IPG that equals their pI, they stop migrating because of loss of net charge (Rabilloud and Lelong 2011). Following focusing and in preparation of the second dimension, strips are equilibrated in buffer containing sodium dodecyl sulfate (SDS).

The second dimension consists of a standard SDS-PAGE (polyacrylamide gel electrophoresis) for separation of proteins according to their molecular weight. To preserve the information of the first dimension, the strip is positioned full length on top of a polyacrylamide SDS gel. The protein pattern on such gels is made visible by protein-sensitive staining; general analysis and comparison of distinct proteomes on gels with respect to presence or absence or relative amount of particular protein spots is accomplished using corresponding computer software packages for assistance (e.g., Delta 2-D, DeCyder, Melanie).

Colloidal Coomassie blue is commonly used for staining of gels because of its high linear dynamic range and reproducibility, but variants of silver-staining protocols that are compatible with downstream mass spectrometric (MS) analysis are also applied. Next to visible stains, fluorescent dyes like SYPRO ruby are popular alternatives (Miller et al. 2006).

In recent years, **difference gel electrophoresis (DIGE)** has been developed (Unlu et al. 1997). Here, protein samples are covalently labeled before 2D-E with the highly sensitive fluorescent dyes Cy2, Cy3, or Cy5. For direct comparison of spot patterns, two differently labeled samples (Cy3 and Cy5, respectively)

and an internal standard (equal mixture of both samples that is labeled by Cy2) are processed together on one gel (Viswanathan et al. 2006).

In general, protein identification is carried out using **matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) MS**. Before MS analysis, selected protein spots have to be excised from gels and enzymatically cleaved into peptides. Then, each peptide sample has to be cocrystallized together with a special light-absorbing matrix solution. Irradiation by a laser beam causes desorption not only of matrix but also of peptides that become protonated and ionized by this procedure. Such ions are accelerated and transferred to a TOF MS analyzer. Depending on the time the peptide ions need to reach the detector, their masses (in daltons, Da) are calculated by the analyzer. These masses are generally recorded as a mass-to-charge (m/z) ratio in relation to signal intensity. Often, for instance when using tandem TOF instruments (TOF-TOF), selected peptide ions (precursor ion) can be further fragmented into several product ions by collision-induced dissociation (CID). Then, the fragments are analyzed by a second MS step that gives MS spectra from which the peptide sequence can be deduced (MS/MS). The obtained MS and MS/MS spectra are used for protein identification by matching of these experimentally obtained data with calculated data from theoretical peptides deduced from protein databases (e.g., National Center for Biotechnology Information, NCBI) using search algorithms like Mascot, Sequest, or PeptideSearch (Steen and Mann 2004).

Depending on length and size of gels of the first and second dimension, resolution of protein spots on gels can vary. Higher resolution is also obtained by using IPGs with differing but narrow pH ranges on multiple gel runs. Although reproducibility in 2D-E has been improved, especially by the use of IPG strips, gel-to-gel variations should be balanced by running of gels in parallel and by analyzing and comparing a number of technical replicates for each biological replicate.

When performing 2D-E, sample preparation is of particular importance. Because of the high heterogeneity of protein characteristics and their large dynamic range within a cell, it is in general and for all proteomic workflows a challenging and virtually impossible task to prepare protein extracts qualitatively and quantitatively representing the entire proteome of interest. All the more important are simple protocols ensuring high reproducibility. Further, 2D-E demands special considerations in sample preparation (Rabilloud and Lelong 2011). To prevent disturbing the IEF step, proteins have to be denatured and kept from precipitation without changing a protein's charge, and high ionic strengths have to be avoided. Further, densely charged molecules like nucleic acids or polysaccharides and other disturbing compounds have to be removed from extracts. Next to general protocols, extract preparation should be adapted and optimized in consideration of the organism of interest and the purpose of the study. For example, optimization of sample preparation for 2D-E of mycelial *A. fumigatus* protein extracts was depicted by Kniemeyer et al. (2006). Further, particular subsets of proteins are inappropriate for resolution by 2D-E, like hydrophobic proteins that especially tend to precipitate during IEF or either very big or small proteins and proteins with very low or high pI (Garbis et al. 2005). When interpreting proteome data in general but notably in 2D-E-based proteomics, it is crucial to consider that the most highly abundant proteins within a proteome will consist of most of the detectable protein spots, whereas a majority of proteins will remain undetected.

B. Gel-Free Proteomics

Gel-free proteomics is in most cases conducted by bottom-up approaches, meaning that no intact proteins are analyzed, but instead their complement of peptides is. Therefore, protein extracts that have been prepared considering high quality and reproducibility attributes have to be **enzymatically digested into peptides** before proceeding with analysis. For digestion, the protease trypsin is used in most cases because of its high efficiency and specificity of cleavage site (peptide bonds C-terminal to lysine or arginine). Digestion of total protein

extract into peptides will result in a sample of much higher complexity. But, next to advantages like enhanced mass accuracy and feasibility of sequencing in later MS steps, bottom-up approaches are especially suitable for proteins that resist analysis by other methods (Steen and Mann 2004). For instance, solubility problems of hydrophobic proteins are generally solved when such proteins are cleaved into shorter fragments.

One main principle in gel-free proteomics is that the complex peptide mixtures become resolved by **high-performance liquid chromatography (HPLC)** prior to protein identification by MS analysis (liquid chromatography [LC]-MS).

As one exception, complex protein mixtures can eventually be separated by SDS-PAGE (one dimensional, 1D) in preparation of the gel-free proteomic procedure. To reduce complexity and hence achieve higher resolution in subsequent analysis, such a complete gel lane is cut into several slices. After in-gel digestion, peptide pools are provided that can be processed individually (GeLC-MS).

Usually, a reversed phase (RP)-HPLC in nanoscale capillary format is set upstream of MS analysis. In RP-HPLC, peptides interact with the stationary phase to various extents according to their hydrophobicity and become eluted by increasing degrees of organic solvent within the mobile phase (Steen and Mann 2004).

When performing a **MudPIT (multidimensional protein identification technology)** approach, more chromatographic steps with different separation modes like a strong cation exchange LC (SCX) are included (Liu et al. 2002). Although it is possible to analyze eluted LC fractions by MALDI-TOF MS, the LC steps are generally directly connected with an MS instrument. These instruments use the technique of electron spray ionization (ESI) in high electric fields for successive liquid evaporation and ionization of enclosed peptides while the peptides elute from RP-HPLC. Once ionized, the peptides are kept in a vacuum and guided into a mass analyzer (Steen and Mann 2004).

Several types of mass analyzers and combinations thereof, functioning by different principles of operation and with differing performances, do exist (e.g., TOF, quadrupole [Q], ion trap [IT], Orbitrap) (Domon and Aebersold 2006; Yates et al. 2009). As output, all instruments will generate mass spectra as depicted for TOF-MS. In proteomics, all systems used are competent for MS/MS using CID or another fragmentation technique.

An important limitation of LC-MS is the fact that definitely not all peptides will be detected. Because of the high number of peptides within such an analysis and because some peptides do not ionize well, a number of peptides become lost for analysis. Also, MS/MS is often only measured for the most abundant peptide ions within a given time window. These constraints result in limitations in sequence coverage of proteins and reduction in reproducibility of technical replicates because selected peptides must not be the same between different runs. Because of missing peptides, isoforms or posttranslational modifications of proteins might not be distinguished or identified.

One predominant scope of application in proteomics is not only the identification of a great many proteins of a particular proteome, but also the identification and quantification of proteins that are differently regulated within two or more different proteomes (e.g., extracts from cells cultured under different conditions or from cells at varying developmental stages). Because of the mentioned limitations in MS analysis, direct comparison of standard MS results is not an easy task for reliable quantitative evaluation of proteins from two different samples. One applied quantification method using standard MS is, for instance, based on counting MS spectra of peptides belonging to a particular protein (spectral counting) (Ishihama et al. 2005; Ong and Mann 2005).

In recent times, more definite **quantitative MS methodologies** became an important field in gel-free proteomics (Ong and Mann 2005; Nikolov et al. 2012). Indeed, absolute protein quantification can be performed for preselected proteins if standards are available, but relative quantification methods are sufficient in most cases. In preparation of relative quantification, protein complements to be analyzed have to be

labeled differently by implementation of dissimilar stable isotopes. Implementation can occur either metabolically by feeding cells of one condition exclusively with "heavy" lysines or arginines (^{13}C or ^{15}N labeled) and with the "light" counterparts in the other condition during culturing (stable isotope labeling with amino acid in cell culture, SILAC) or by chemical ligation of extracted proteins to isotopically labeled molecules (isotope-coded affinity tag, ICAT). Alternatively, labeling can be performed on a peptide level using a strategy called iTRAQ (isobaric tag for relative and absolute quantification). Independent of labeling strategy, all samples have to be mixed in equal proportions as early in the workflow as possible. Subsequently, they are processed simultaneously. Because differently labeled peptides will behave identically in the same MS run, their signal intensity ratio can be directly quantified. Comparison is made possible by the introduced distinct mass shift between incorporated heavy and light isotopes.

C. Proteome Analysis of Opportunistic Fungi

Although opportunistic fungi form a rather heterogeneous group, they have in common the capacity to colonize human tissues followed by tissue invasion or dissemination of the human body. While some fungal species (e.g., *C. albicans*) are typically found in distinct human tissues (e.g., gastrointestinal tract, mucous membranes, skin) as commensals, a number of other species naturally inhabit completely different environments (e.g., decaying organic matter). Albeit immunosuppressed individuals are in many cases the only cohort of individuals that can be severely infected by particular opportunistic fungi, these fungi have to nonetheless withstand adverse living conditions on infection (e.g., high temperature, nutrient and trace element limitation, residual immune response, hypoxia). To elucidate fungal survival and adaptation mechanisms on human infection, proteomic approaches are used complementary with transcriptomics. For such large-scale analyses, particular aspects of infection conditions are typically simulated in vitro. Recently reviewed by Kroll, Kniemeyer, and colleagues and depicted in some examples that follow, such analyses are considered to provide first insights into the comprehension of fungal adaptation and virulence mechanisms (Kniemeyer et al. 2011; Kroll et al. 2013). With these kinds of -omic approaches, protein interaction networks can be deduced using bioinformatics. Besides defining relations between factors, one general way of analyzing these data is the classification of identified factors into distinct functional groups (e.g., particular metabolic pathways). Further, after identification of interesting factors, these large-scale methodologies can serve as starting points for small-scale functional characterization of potentially relevant and virulence-determining proteins.

An initial starting point of proteomic approaches is the generation of **reference proteome maps** of the fungal species of interest, such as *A. fumigatus* (Vodisch et al. 2009; Teutschbein et al. 2010), *A. flavus* (Pechanova et al. 2013), and *C. albicans* (Hernandez et al. 2004; Kusch et al. 2008). Such maps reflect major functional characteristics of respective fungi at a distinct developmental or morphological state under defined culturing conditions. Generally, proteins participating in primary metabolism, respiration, translation, or stress response are highly represented on such maps (Kniemeyer et al. 2011).

1. Cell Wall Proteome

Next to total cell proteome analysis, the possibility to focus on certain subproteomes is one strong point of proteomic approaches over transcriptomics. Cell wall proteomics has the capacity to reveal protein candidates that act at the **fungus-host interface**.

For identification of covalently bound cell wall proteins from various opportunistic fungi, harsh extraction conditions (high salt, SDS, β-mercaptoethanol, high temperature) have been applied in a number of studies to isolate the cell walls and to remove noncovalently bound proteins and intracellular protein contaminants. Cell wall proteins can be directly digested with

trypsin in preparation for LC-MS analysis (Sosinska et al. 2008; Heilmann et al. 2011; Sosinska et al. 2011). Alternatively, cell walls have been treated with hydrogen fluoride (HF)-pyridine or β-1,6-glucanase to specifically solubilize glycosylphosphatidylinositol (GPI)-anchored proteins (de Groot et al. 2004; Weig et al. 2004; Castillo et al. 2008; Heddergott et al. 2012). Alternatively, alkali treatment has been used to liberate covalently bound proteins with no GPI anchor, or β-1,3-glucanase treatment was conducted for overall release of cell wall proteins (de Groot et al. 2004; Weig et al. 2004). In gentler approaches, cell surface proteomes have been analyzed by direct trypsin digestion of proteins from intact cells, followed by RP-HPLC and MS (Hernaez et al. 2010), or by slight alkali treatment in combination with 1,3-β-glucanase digestion omitting the cell wall isolation step (Asif et al. 2006).

2. Secretome

As one important (sub)proteome, secreted proteins (secretomes) should not be neglected (Sorgo et al. 2013). A fungal secretome has the capacity to alter extensively depending on culture conditions (Schwienbacher et al. 2005; Sorgo et al. 2010; Wartenberg et al. 2011). One major task of secreted proteins is the **processing of complex nutrient sources** like polysaccharides (C sources), lipids (C sources), or proteins (N and C sources) into smaller entities that can be assimilated by the fungus. Wartenberg and coworkers identified several enzymes potentially involved in nutrient acquisition when they generated secretome maps for *A. fumigatus* using 2D-E (Wartenberg et al. 2011). Further, compared to the secretome obtained after standard cultivation in AMM (*Aspergillus* minimal medium, containing $NaNO_3$ and glucose as sole N and C source, respectively), the proportion of particular proteases was strongly elevated in the secretome when culturing of *A. fumigatus* was performed using media with elastin, collagen, or keratin as the major N and C source. Interestingly, a number of proteins associated with cell rescue and defense mechanisms can be found in the secretome of *A. fumigatus*. Next to proteins that contain a signal peptide for conventional secretion using the route of endoplasmic reticulum and Golgi, several more proteins have been found in secretomes (Wartenberg et al. 2011). Their secretion cannot be deduced by DNA sequence analysis. One of the alternative routes for their secretion is their packaging into extracellular vesicles (EVs). Proteome analysis (LC-MS) of EVs from a number of fungi (*C. neoformans*, *Histoplasma capsulatum*, *S. cerevisiae*, *C. albicans*, *Paracoccidioides brasiliensis*) revealed sets of proteins belonging to diverse categories (Albuquerque et al. 2008; Rodrigues et al. 2008, 2013; Vallejo et al. 2012). Because such EVs have been shown to harbor a number of virulence-associated factors, one attributed function of such vesicles might be to serve as "virulence bags" (Rodrigues et al. 2013). For example, the moonlight protein Enolase was identified from isolated EV fractions of all tested species (Rodrigues et al. 2013). In *C. albicans*, this protein has been shown to interact with human plasminogen (Jong et al. 2003).

3. Adaptation to Host Niches

a) Morphotypes and Growth Forms

Both the yeast and the hyphal form of *C. albicans* are relevant for invasive infections, as each morphotype seems to be adapted to appropriate virulence-associated functions and host niches (Jacobsen et al. 2012). Functional heterogeneity is also reflected by proteins differently regulated in both morphotypes, and it is beyond question that morphotype switching necessitates major cellular remodeling. A number of triggers for **yeast-to-hyphae transition** have been identified (e.g., serum, nitrogen limitation, N-acetylglucosamine, ambient pH, high temperature) and have been used for in vitro cultivation of both states. Few comparative proteome studies regarding intracellular proteins of both morphotypes have been conducted (Monteoliva et al. 2011; Kamthan et al. 2012). Generally, identified proteins indicate global metabolic and cellular differences between the two states. Using quantitative gel-free proteomics, Heilmann and colleagues analyzed the cell wall fractions of *C. albicans* after culturing

under various hypha-induction conditions. Importantly, they could identify a common pattern of cell wall proteins that showed good consistency between the different hypha-induction conditions. Further, clear distinctions were detected between the proteomes of yeast or hyphal cell walls (Heilmann et al. 2011).

Realized as a clinically relevant growth form, **fungal biofilms** have come into the focus of research (Seneviratne et al. 2008). Next to transcriptome analysis, proteome analysis has been conducted on *C. albicans* and *A. fumgatus* biofilms (Thomas et al. 2006; Seneviratne et al. 2008; Bruns et al. 2010; Muszkieta et al. 2013). Interestingly, it has been shown by Bruns et al. that among a number of differently regulated proteins, proteins involved in gliotoxin (mycotoxin) biosynthesis became enriched during biofilm maturation (Bruns et al. 2010).

b) Nutrient and Iron Limitation

During an infection and dissemination process, opportunistic fungi are capable of adapting to several microenvironments. While nutrient acquisition within such host niches is in general seen to be far from optimal, available food supply also varies between microenvironments.

In vitro testing of a number of potentially infection-relevant N and C sources for their usability as a nutrient source, and the eventual nutrient-conditioned reprogramming of the fungus is one important task that can be investigated using proteomics. Intriguingly, it has been shown by LC-MS proteomics (quantification via spectral counts) that cultivation of *C. albicans* on lactate as the sole C source has a significant impact on cell wall and secretome composition (Ene et al. 2012). Such remodeling resulted in increased stress resistance and increased adherence but less antifungal drug susceptibility.

Besides nutrient uptake, the acquisition of iron is a demanding task in environments with poor availability of this metal, like the human body. Accordingly, tight regulation of iron acquisition and consumption pathways have been shown to be crucial for adaptation to high- or low-iron environments (Haas 2012). **Adaptation to low-iron availability** has also been analyzed using proteomics (Hortschansky et al. 2007). By combining MudPIT (quantification via spectral counts) and 2D-E, Crestani et al. analyzed, for instance, how *C. gatii* (VGIIa subtype) handled iron deficiency on a proteome level (Crestani et al. 2012).

c) Hypoxia, pH, and Temperature

Next to the challenges mentioned, surviving in host tissues takes competence to adapt to **low oxygen levels (hypoxia)**, as oxygen concentration is supposed to be rather low in tissues but especially low at inflammatory and necrotic sites (Grahl et al. 2012). In two proteomic studies (2D-E), the short-term response and long-term adaptation to low oxygen (0.2 %) have been analyzed for *A. fumigatus* (Vodisch et al. 2011; Barker et al. 2012). Next to variations in up- and downregulated proteins between these studies that are probably reflected by the differing length of hypoxia treatment (up to 24 h vs. 10 days) and by some other divergences between cultivation conditions (e.g., glucose availability), it is noticeable that factors involved in not only glycolysis but also respiration were upregulated under both conditions. By contrast, ethanol fermentation could only be detected as a short-term response to hypoxia, and the production of the secondary metabolite pseurotin A was detectable only after long-term hypoxic cultivation.

Further host-specific environmental conditions for an opportunistic fungus during (systemic) infection are temperatures of 37 °C or higher (fever) and an alkaline pH environment.

Temperature-dependent protein up- and downregulation have been investigated for *A. flavus* using a SILAC geLC-MS approach (Collier et al. 2008; Georgianna et al. 2008). Culturing medium was supplemented with isotopically labeled arginine. After culturing at 28 °C or 37 °C, protein extracts were generated, and both the $^{13}C_6$-arginine-labeled (28 °C) and $^{12}C_6$-arginine-labeled samples (37 °C) were processed together. Direct comparison of both conditions revealed a temperature-dependent downregulation of proteins involved in aflatoxin production at higher temperature. In total, 49

proteins showed differing abundancies between the growth conditions.

While proteome changes were only small between *A. fumigatus* cellular extracts after cultivation at 30 °C or 37 °C, a clear heat shock response was recognizable when the culturing temperature was further shifted to 48 °C (compared to 30 °C). DIGE could recognize 64 proteins as differing in abundance, including many chaperones. Also, proteins belonging to other functional groups, like the oxidative stress response, were differentially regulated (Albrecht et al. 2010).

Regarding **pH tolerance**, in a geLC-MS approach, it has been shown that *A. fumigatus* has the capability to secrete different sets of proteolytic enzymes depending on the pH of the culture, indicating an adaptation mechanism by which such enzymes tend to be secreted if their intrinsic pH optimum matches a given environmental pH (Sriranganadane et al. 2010). For *C. albicans*, it has been elucidated that the composition of the cell wall proteome is pH dependent. Of 22 identifiable cell wall proteins, 16 showed significant differences in abundance when *C. albicans* was cultivated either at pH 4 or at pH 7 under conditions that were established to mimic mucosal surface colonization (Sosinska et al. 2011). In this study, quantitative MS has been conducted by including isotypically ^{15}N-labeled reference cultures in the analysis, which were obtained by cultivation with [^{15}N] ammonium sulfate as the sole N source. Interestingly, it turned out that six of the identified proteins, three adhesion molecules (Als1, Als3, Hyr1) and three further proteins (Phr1, Hwp1, Sod5), could exclusively be found in the cell wall proteome of the pH 7 culture, and the transglucosidase Phr2 was extremely upregulated in the pH 4 protein sample.

d) Reactive Oxygen and Nitrogen Species

One challenge during infection of human hosts is to resist the attacks of phagocytes (macrophages, neutrophils) (Seider et al. 2010). Besides other defense mechanisms, immune cells produce **reactive oxygen intermediates**. Therefore, understanding the mechanisms by which opportunistic fungi can cope with elevated levels of ROIs in their environment will give insights into whether ROIs contribute to killing.

In *A. fumigatus*, 28 mycelial proteins were detected by 2D-E to be differently regulated after exposure to 2 mM H_2O_2 for 15 or 45 min. Here, stress-associated proteins involved in detoxification of ROIs were among the highest upregulated ones: The Cu/Zn superoxide dismutase SOD, the peroxiredoxin Prx1, a cytochrome c peroxidase, and the putative thioredoxin reductase AspF3 (Lessing et al. 2007).

Similarly, a number of enzymes functionally classified as ROI detoxification proteins and oxidoreductases were shown to be upregulated in *C. albicans* yeast cells because of incubation with H_2O_2 (1 mM) or the thiol-oxidizing agent diamide (5 mM) (Kusch et al. 2007). Several chaperones appeared in the list of upregulated proteins in this study. The experimental procedure used was sequential pulse labeling of fungal cultures at several time points by ^{35}S-methionine, followed by 2D-E and phosphoscreen detection; this enabled Kusch et al. to fix the main level of protein up- or downregulation to take place in a time window of 5–30 min after initial contact with the stressing reagent. This time-dependent protein regulation could be one reason why only a little overlap of regulated factors was detected with a proteome study using *C. albicans* harvested after 1 h of incubation in 5 mM H_2O_2 (Yin et al. 2009).

Also, fungi have to cope with **reactive nitrogen intermediates (RNIs)**, when confronted with phagocytic cells. Like ROIs, RNIs are highly reactive, causing serious damage on proteins, DNA, and lipids (Missall et al. 2004). On a proteomic level (2D-E), cellular response to nitric oxide has been investigated in *C. neoformans* (Missall et al. 2006). Similar to a stress response on treatment with ROIs, the authors showed several enzymes with antioxidant activity were upregulated by NO stress.

VI. Functional Genomics and Bioinformatics

Today, high-throughput data can be obtained with reasonable efforts and costs. The bottleneck for these new techniques remains the analysis of the generated amounts of data. Analysis of high-

throughput data can be described by the operational protocol of data-driven systems biology. Different steps make up this protocol: experimental design, data preprocessing, feature selection, computational modeling, and biological evaluation. The goal of the analysis is to obtain models and, if these models are still incomplete, to obtain variables and parameters that can be used to generate new hypotheses for further experiments. Currently, the real challenge is to integrate all the different levels of -omics data, considering multiple molecular levels, and the integration of different technologies. Analysis of data can be only carried out using a holistic approach combining life sciences (e.g., biology, medicine, biochemistry); mathematics; informatics; and engineering sciences, thus bridging the gap between computational modeling techniques and biomolecular experimental techniques. This approach is termed ***systems biology***.

Every functional genomics approach is facilitated when the genome sequence of the object of study is available. If host-pathogen interaction is considered, the genomic information from the host and the pathogen forms the basis for all further molecular and computational analyses. Currently, there is access to about 45 genomes corresponding to pathogenic fungi (Horn et al. 2012). Recent technical advances drastically decreased the costs for DNA sequencing, increasing the number of sequenced fungal species. There are actually about 100 genome projects in progress, such as the sequencing of several *C. albicans* strains to evaluate *Candida* genome plasticity. Moreover, the 1000 Fungal Genomes Project (1KFG) was initiated at the Department of Energy (DOE) Joint Genome Institute (JGI), which will provide broad genome coverage of many different fungal species (http://1000.fungalgenomes.org).

After genome sequencing, the second step would be to have fairly good prediction of potentially expressed genes. Gene prediction provides information concerning genes, including alternative splicing variants, localization of proteins, and putative interaction partners. Many methodologies have been developed to improve the quality of predicted genes. One of these proposed using block motifs from reviewed protein sequences to help intron/exon prediction (Keller et al. 2011). Another way to predict splicing signatures is by integrating genome-sequencing data with RNA-sequencing data. This approach was successfully used to predict new misannotated transcripts in the *A. fumigatus* genome (Muller et al. 2012).

The genomic sequences and their respective annotations are normally provided through **genome portals**, such as NCBI (http://www.ncbi.nlm.nih.gov) and Ensembl (http://www.ensembl.org); by the sequencing institution, such as the Broad institute (http://www.broadinstitute.org/) and JGI (http://genome.jgi-psf.org); or by community data portals, such as AspGD (http://www.aspergillusgenome.org), CGD (*Candida* Genome Database; http://www.candidagenome.org), and FungiDB (http://fungidb.org). Besides deepening taxonomic knowledge, the availability of many released genome sequences allowed the extensive use of comparative genome studies.

Using this information, it was possible to establish in silico analyses of regulatory pathways. The investigation of **gene regulatory networks** is crucial for functional genomic studies because it can suggest possible interactions with partner molecules. In this regard, several tools have been developed to understand predicted gene function of protein families. The tools mainly applied are Gene Ontology (GO; Ashburner et al. 2000), the Functional Catalogue (FunCat; Ruepp et al. 2004), and the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa and Goto 2000). However, all these analyses are based on predicted protein signatures, which suggest the category or family to which the predicted genes belong. The categorization is based on empirical knowledge obtained by bibliography (i.e., the better a pathway is already studied, the better it is recognized and predicted).

The generation of experimental data in molecular biology is no longer considered a limitation. Now, the number of available data sets is exponentially increasing, and a big challenge is the storage of the data produced. However, accumulating data do not necessarily mean that new biological insight is gained. Sometimes, the

increase of information enlarges the model but does not make it more precise. This is basically because of a small community managing the experimental setup, while the analysis of data is subsequently given to a third party.

The experimental setup is important for providing high-quality data that can easily be analyzed and integrated. The first stage is the experimental design, by which the most appropriate technology is chosen, and the experiment is defined in terms of the number of replicates and time points. During this stage, it also important to analyze prior knowledge. After carrying out the experiments, the data need to be preprocessed to remove systematic biases introduced by the technology or the experimental design. In extreme cases, the quality of the measurements obtained may lead to a complete rejection of the data set and replanning the experiment. The next step is to choose which features should be applied to highlight the results obtained and to have sufficient information to build a theory. If the quality of the data was high and the experimental design arranged accordingly, a mathematical modeling stage would be subsequently applied. The main goals of modeling are to describe the observed experimental data dynamically, to further prioritize candidates, or to predict new, so-far-unknown interactions within the data. Models are used to describe the system at the molecular, the cell, the tissue, the organ, the systemic, or the population levels (Horn et al. 2012). Concerning infection biology, there are many studies using modeling techniques on the human host side, mainly focused on viruses and bacteria (Vodovotz et al. 2010). Concerning the molecular mechanisms of fungal pathogenicity, there are some studies that applied modeling techniques to the fungal perspective of an infection (Altwasser et al. 2012; Linde et al. 2012). These reported studies were basically focused on the elucidation of key virulence traits and their molecular characterization.

The last step of the analysis should be the **building of possible network models**. Network modeling attempts to describe biological systems and their interactions with the help of a graph. In this graph, the nodes usually correspond to molecular entities (e.g., genes or proteins), and the edges between them depict a causal relationship (e.g., regulation). These types of networks help to understand the biological system and to put the biological information into a unique context. One goal of systems biology is to infer such networks. Networking can be applied to different levels, which include gene regulatory networks (nodes represent genes, and edges represent an interaction between them); protein interaction networks (nodes represent proteins, whereas the edges between represent an interaction, which may directly imply a possible binding between two proteins); signaling networks (nodes indicate proteins, which may have a certain activation level); and metabolic networks (nodes represent which metabolites are consumed and produced during one reaction) (Hecker et al. 2009).

At the end of the analysis and the modeling, all hypotheses made should be experimentally validated. At this stage, wet lab practices are commonly applied using classical experimental technologies such as Western blotting, chromatin immunoprecipitation, flow cytometry, enzyme-linked immunosorbent assay, immunohistochemistry, small-interfering RNA (siRNA), quantitative real time PCR (qRT-PCR) qRT-PCR, and eventually the generation and analysis of gene knockout mutants that can validate the driving hypothesis. This stage not only confirms the aforementioned stages in data-driven systems biology but also helps to assess the quality of applied computational methods and their according parameter settings. Ideally, the results obtained are transferred from the in vivo confirmation and, in the infection biology field, to clinical applications.

VII. Conclusions

The tremendous advances in NGS and other high-throughput -omics technologies related to functional genomics taken together with the rapidly advancing computational resources offer enormous possibilities to unravel the mechanisms by which opportunistic fungal pathogens infect their host.

To further speed up deciphering the genetic and molecular basis of pathogenicity, current efforts focus on the development of automated phenotype screens. As many mutations do not result in a tractable phenotype using standard tests, multiwell tests are desirable to address the challenge of detecting phenotypical changes in mutants. Finally, in-depth analysis of candidate genes identified in high-throughput screens will enable breakthroughs in decoding the mechanisms underlying pathogenicity of opportunistic fungal pathogens.

References

Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC (1993) Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. Nat Genet 4:373–380

Albrecht D, Guthke R, Brakhage AA, Kniemeyer O (2010) Integrative analysis of the heat shock response in *Aspergillus fumigatus*. BMC Genomics 11:32

Albuquerque PC, Nakayasu ES, Rodrigues ML, Frases S, Casadevall A, Zancope-Oliveira RM, Almeida IC, Nosanchuk JD (2008) Vesicular transport in *Histoplasma capsulatum*: an effective mechanism for trans-cell wall transfer of proteins and lipids in ascomycetes. Cell Microbiol 10:1695–1710

Alter O, Brown PO, Botstein D (2003) Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms. Proc Natl Acad Sci U S A 100:3351–3356

Altwasser R, Linde J, Buyko E, Hahn U, Guthke R (2012) Genome-wide scale-free network inference for *Candida albicans*. Front Microbiol 3:51

Alwine JC, Kemp DJ, Stark GR (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci U S A 74:5350–5354

Andersen MR, Nielsen JB, Klitgaard A, Petersen LM, Zachariasen M, Hansen TJ, Blicher LH, Gotfredsen CH, Larsen TO, Nielsen KF, Mortensen UH (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. Proc Natl Acad Sci U S A 110:E99–E107

Arnaud MB, Chibucos MC, Costanzo MC, Crabtree J, Inglis DO, Lotia A, Orvis J, Shah P, Skrzypek MS, Binkley G, Miyasato SR, Wortman JR, Sherlock G (2010) The *Aspergillus* Genome Database, a curated comparative genomics resource for gene, protein and sequence information for the *Aspergillus* research community. Nucleic Acids Res 38: D420–D427

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 25:25–29

Asif AR, Oellerich M, Amstrong VW, Riemenschneider B, Monod M, Reichard U (2006) Proteome of conidial surface associated proteins of *Aspergillus fumigatus* reflecting potential vaccine candidates and allergens. J Proteome Res 5:954–962

Askew DS (2008) *Aspergillus fumigatus*: virulence genes in a street-smart mold. Curr Opin Microbiol 11:331–337

Barker KS, Crisp S, Wiederhold N, Lewis RE, Bareither B, Eckstein J, Barbuch R, Bard M, Rogers PD (2004) Genome-wide expression profiling reveals genes associated with amphotericin B and fluconazole resistance in experimentally induced antifungal resistant isolates of *Candida albicans*. J Antimicrob Chemother 54:376–385

Barker BM, Kroll K, Vodisch M, Mazurie A, Kniemeyer O, Cramer RA (2012) Transcriptomic and proteomic analyses of the *Aspergillus fumigatus* hypoxia response using an oxygen-controlled fermenter. BMC Genomics 13:62

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T,

Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59

Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat Chem Biol 3:213–217

Bok JW, Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell 3:527–535

Bok JW, Chiang YM, Szewczyk E, Reyes-Dominguez Y, Davidson AD, Sanchez JF, Lo HC, Watanabe K, Strauss J, Oakley BR, Wang CC, Keller NP (2009) Chromatin-level regulation of biosynthetic gene clusters. Nat Chem Biol 5:462–464

Brakhage AA (2005) Systemic fungal infections caused by *Aspergillus* species: epidemiology, infection process and virulence determinants. Curr Drug Targets 6:875–886

Brakhage AA (2013) Regulation of fungal secondary metabolism. Nat Rev Microbiol 11:21–32

Brakhage AA, Schuemann J, Bergmann S, Scherlach K, Schroeckh V, Hertweck C (2008) Activation of fungal silent gene clusters: a new avenue to drug discovery. Prog Drug Res 66(1):3–12

Brakhage AA, Thon M, Sprote P, Scharf DH, Al-Abdallah Q, Wolke SM, Hortschansky P (2009) Aspects on evolution of fungal beta-lactam biosynthesis gene clusters and recruitment of trans-acting factors. Phytochemistry 70:1801–1811

Brown JS, Aufauvre-Brown A, Brown J, Jennings JM, Arst H Jr, Holden DW (2000) Signature-tagged and directed mutagenesis identify PABA synthetase as essential for *Aspergillus fumigatus* pathogenicity. Mol Microbiol 36:1371–1380

Brown GD, Denning DW, Gow NA, Levitz SM, Netea MG, White TC (2012) Hidden killers: human fungal infections. Sci Transl Med 4:165rv13

Bruns S, Seidler M, Albrecht D, Salvenmoser S, Remme N, Hertweck C, Brakhage AA, Kniemeyer O, Muller FM (2010) Functional genomic profiling of *Aspergillus fumigatus* biofilm reveals enhanced production of the mycotoxin gliotoxin. Proteomics 10:3097–3107

Castillo L, Calvo E, Martinez AI, Ruiz-Herrera J, Valentin E, Lopez JA, Sentandreu R (2008) A study of the *Candida albicans* cell wall proteome. Proteomics 8:3871–3881

Chen D, Toone WM, Mata J, Lyne R, Burns G, Kivinen K, Brazma A, Jones N, Bahler J (2003) Global transcriptional responses of fission yeast to environmental stress. Mol Biol Cell 14:214–229

Chiang YM, Lee KH, Sanchez JF, Keller NP, Wang CC (2009) Unlocking fungal cryptic natural products. Nat Prod Commun 4:1505–1510

Chooi YH, Fang J, Liu H, Filler SG, Wang P, Tang Y (2013) Genome mining of a prenylated and immunosuppressive polyketide from pathogenic fungi. Org Lett 15:780–783

Cichewicz RH (2010) Epigenome manipulation as a pathway to new natural product scaffolds and their congeners. Nat Prod Rep 27:11–22

Collier TS, Hawkridge AM, Georgianna DR, Payne GA, Muddiman DC (2008) Top-down identification and quantification of stable isotope labeled proteins from *Aspergillus flavus* using online nanoflow reversed-phase liquid chromatography coupled to a LTQ-FTICR mass spectrometer. Anal Chem 80:4994–5001

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103:10352–10357

Cormack BP, Ghori N, Falkow S (1999) An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. Science 285:578–582

Crestani J, Carvalho PC, Han X, Seixas A, Broetto L, Fischer Jde S, Staats CC, Schrank A, Yates JR 3rd, Vainstein MH (2012) Proteomic profiling of the influence of iron availability on *Cryptococcus gattii*. J Proteome Res 11:189–205

da Silva Ferreira ME, Malavazi I, Savoldi M, Brakhage AA, Goldman MH, Kim HS, Nierman WC, Goldman GH (2006) Transcriptome analysis of *Aspergillus fumigatus* exposed to voriconazole. Curr Genet 50:32–44

da Silva Ferreira ME, Heinekamp T, Hartl A, Brakhage AA, Semighini CP, Harris SD, Savoldi M, de Gouvea PF, de Souza Goldman MH, Goldman GH (2007) Functional characterization of the *Aspergillus fumigatus* calcineurin. Fungal Genet Biol 44:219–230

Davis C, Carberry S, Schrettl M, Singh I, Stephens JC, Barry SM, Kavanagh K, Challis GL, Brougham D, Doyle S (2011) The role of glutathione S-transferase GliG in gliotoxin biosynthesis in *Aspergillus fumigatus*. Chem Biol 18:542–552

De Backer MD, Ilyina T, Ma XJ, Vandoninck S, Luyten WH, Vanden Bossche H (2001) Genomic profiling of the response of *Candida albicans* to itraconazole

treatment using a DNA microarray. Antimicrob Agents Chemother 45:1660–1670

de Groot PW, de Boer AD, Cunningham J, Dekker HL, de Jong L, Hellingwerf KJ, de Koster C, Klis FM (2004) Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. Eukaryot Cell 3:955–965

Doedt T, Krishnamurthy S, Bockmuhl DP, Tebarth B, Stempel C, Russell CL, Brown AJ, Ernst JF (2004) APSES proteins regulate morphogenesis and metabolism in *Candida albicans*. Mol Biol Cell 15:3167–3180

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312:212–217

Ene IV, Heilmann CJ, Sorgo AG, Walker LA, de Koster CG, Munro CA, Klis FM, Brown AJ (2012) Carbon source-induced reprogramming of the cell wall proteome and secretome modulates the adherence and drug resistance of the fungal pathogen *Candida albicans*. Proteomics 12:3164–3179

Fedorova ND, Nierman WC, Turner G, Joardar V, Maiti R, Anderson MJ, Denning DW, Wortman JR (2008) A comparative view of the genome of *Aspergillus fumigatus*. In: Osmani SA, Goldman GH (eds) The aspergilli: genomics, medical aspects, biotechnology, and research methods. CRC Press, Boca Raton, pp 25–42

Gacek A, Strauss J (2012) The chromatin code of fungal secondary metabolite gene clusters. Appl Microbiol Biotechnol 95:1389–1404

Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Penalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115

Gallagher L, Owens RA, Dolan SK, O'Keeffe G, Schrettl M, Kavanagh K, Jones GW, Doyle S (2012) The *Aspergillus fumigatus* protein GliK protects against oxidative stress and is essential for gliotoxin biosynthesis. Eukaryot Cell 11:1226–1238

Garbis S, Lubec G, Fountoulakis M (2005) Limitations of current proteomics technologies. J Chromatogr A 1077:1–18

Gardiner DM, Howlett BJ (2005) Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of *Aspergillus fumigatus*. FEMS Microbiol Lett 248:241–248

Georgianna DR, Hawkridge AM, Muddiman DC, Payne GA (2008) Temperature-dependent regulation of proteins in *Aspergillus flavus*: whole organism stable isotope labeling by amino acids. J Proteome Res 7:2973–2979

Gibbons JG, Beauvais A, Beau R, McGary KL, Latge JP, Rokas A (2012) Global transcriptome changes underlying colony growth in the opportunistic human pathogen *Aspergillus fumigatus*. Eukaryot Cell 11:68–78

Grahl N, Shepardson KM, Chung D, Cramer RA (2012) Hypoxia and fungal pathogenesis: to air or not to air? Eukaryot Cell 11:560–570

Haas H (2012) Iron – a key nexus in the virulence of *Aspergillus fumigatus*. Front Microbiol 3:28

Hartmann T, Dumig M, Jaber BM, Szewczyk E, Olbermann P, Morschhauser J, Krappmann S (2010) Validation of a self-excising marker in the human pathogen *Aspergillus fumigatus* by employing the beta-rec/six site-specific recombination system. Appl Environ Microbiol 76:6313–6317

Hartmann T, Cairns TC, Olbermann P, Morschhauser J, Bignell EM, Krappmann S (2011) Oligopeptide transport and regulation of extracellular proteolysis are required for growth of *Aspergillus fumigatus* on complex substrates but not for virulence. Mol Microbiol 82:917–935

Hecker M, Lambeck S, Toepfer S, van Someren E, Guthke R (2009) Gene regulatory network inference: data integration in dynamic models-a review. Biosystems 96:86–103

Heddergott C, Bruns S, Nietzsche S, Leonhardt I, Kurzai O, Kniemeyer O, Brakhage AA (2012) The *Arthroderma benhamiae* hydrophobin HypA mediates hydrophobicity and influences recognition by human immune effector cells. Eukaryot Cell 11:673–682

Heilmann CJ, Sorgo AG, Siliakus AR, Dekker HL, Brul S, de Koster CG, de Koning LJ, Klis FM (2011) Hyphal induction in the human fungal pathogen *Candida albicans* reveals a characteristic wall protein profile. Microbiology 157:2297–2307

Hernaez ML, Ximenez-Embun P, Martinez-Gomariz M, Gutierrez-Blazquez MD, Nombela C, Gil C (2010) Identification of *Candida albicans* exposed surface proteins in vivo by a rapid proteomic approach. J Proteomics 73:1404–1409

Hernandez R, Nombela C, Diez-Orejas R, Gil C (2004) Two-dimensional reference map of *Candida albicans* hyphal forms. Proteomics 4:374–382

Hillier LW, Reinke V, Green P, Hirst M, Marra MA, Waterston RH (2009) Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. Genome Res 19:657–666

Homann OR, Dea J, Noble SM, Johnson AD (2009) A phenotypic profile of the *Candida albicans* regulatory network. PLoS Genet 5:e1000783

Horn F, Heinekamp T, Kniemeyer O, Pollmacher J, Valiante V, Brakhage AA (2012) Systems biology of fungal infection. Front Microbiol 3:108

Hortschansky P, Eisendle M, Al-Abdallah Q, Schmidt AD, Bergmann S, Thon M, Kniemeyer O, Abt B, Seeber B, Werner ER, Kato M, Brakhage AA, Haas H (2007) Interaction of HapX with the CCAAT-binding complex – a novel mechanism of gene regulation by iron. Embo J 26:3157–3168

Ihmels J, Bergmann S, Berman J, Barkai N (2005) Comparative gene expression analysis by differential clustering approach: application to the *Candida albicans* transcription program. PLoS Genet 1:e39

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteom 4:1265–1272

Jacobsen ID, Wilson D, Wachtler B, Brunke S, Naglik JR, Hube B (2012) *Candida albicans* dimorphism as a therapeutic target. Expert Rev Anti Infect Ther 10:85–93

Jahn B, Boukhallouk F, Lotz J, Langfelder K, Wanner G, Brakhage AA (2000) Interaction of human phagocytes with pigmentless *Aspergillus conidia*. Infect Immun 68:3736–3739

Jain R, Valiante V, Remme N, Docimo T, Heinekamp T, Hertweck C, Gershenzon J, Haas H, Brakhage AA (2011) The MAP kinase MpkA controls cell wall integrity, oxidative stress response, gliotoxin production and iron adaptation in *Aspergillus fumigatus*. Mol Microbiol 82:39–53

Jong AY, Chen SH, Stins MF, Kim KS, Tuan TL, Huang SH (2003) Binding of *Candida albicans* enolase to plasmin(ogen) results in enhanced invasion of human brain microvascular endothelial cells. J Med Microbiol 52:615–622

Kamthan M, Mukhopadhyay G, Chakraborty N, Chakraborty S, Datta A (2012) Quantitative proteomics and metabolomics approaches to demonstrate N-acetyl-D-glucosamine inducible amino acid deprivation response as morphological switch in *Candida albicans*. Fungal Genet Biol 49:369–378

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30

Keller O, Kollmar M, Stanke M, Waack S (2011) A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics 27:757–763

Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND (2010) SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol 47:736–741

Kniemeyer O, Lessing F, Scheibner O, Hertweck C, Brakhage AA (2006) Optimisation of a 2-D gel electrophoresis protocol for the human-pathogenic fungus *Aspergillus fumigatus*. Curr Genet 49:178–189

Kniemeyer O, Schmidt AD, Vodisch M, Wartenberg D, Brakhage AA (2011) Identification of virulence determinants of the human pathogenic fungi *Aspergillus fumigatus* and *Candida albicans* by proteomics. Int J Med Microbiol 301:368–377

Krappmann S, Sasse C, Braus GH (2006) Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end-joining-deficient genetic background. Eukaryot Cell 5:212–215

Kraus PR, Boily MJ, Giles SS, Stajich JE, Allen A, Cox GM, Dietrich FS, Perfect JR, Heitman J (2004) Identification of *Cryptococcus neoformans* temperature-regulated genes with a genomic-DNA microarray. Eukaryot Cell 3:1249–1260

Kroll K, Pahtz V, Kniemeyer O (2013) Elucidating the fungal stress response by proteomics. J Proteom pii:S1874–3919(13)00315–1. doi:10.1016/j.jprot.2013.06.001. [Epub ahead of print]

Kusch H, Engelmann S, Albrecht D, Morschhauser J, Hecker M (2007) Proteomic analysis of the oxidative stress response in *Candida albicans*. Proteomics 7:686–697

Kusch H, Engelmann S, Bode R, Albrecht D, Morschhauser J, Hecker M (2008) A proteomic view of *Candida albicans* yeast cell metabolism in exponential and stationary growth phases. Int J Med Microbiol 298:291–318

La Hoz RM, Pappas PG (2013) Cryptococcal infections: changing epidemiology and implications for therapy. Drugs 73:495–504

Langfelder K, Jahn B, Gehringer H, Schmidt A, Wanner G, Brakhage AA (1998) Identification of a polyketide synthase gene (*pksP*) of *Aspergillus fumigatus* involved in conidial pigment biosynthesis and virulence. Med Microbiol Immun (Berl) 187:79–89

Langfelder K, Philippe B, Jahn B, Latge JP, Brakhage AA (2001) Differential expression of the *Aspergillus fumigatus pksP* gene detected in vitro and in vivo with green fluorescent protein. Infect Immun 69:6411–6418

Langfelder K, Streibel M, Jahn B, Haase G, Brakhage AA (2003) Biosynthesis of fungal melanins and their importance for human pathogenic fungi. Fungal Genet Biol 38:143–158

Lee I, Oh JH, Shwab EK, Dagenais TR, Andes D, Keller NP (2009) HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. Fungal Genet Biol 46:782–790

Lessing F, Kniemeyer O, Wozniok I, Loeffler J, Kurzai O, Haertl A, Brakhage AA (2007) The *Aspergillus fumigatus* transcriptional regulator AfYap1 represents the major regulator for defense against reactive oxygen intermediates but is dispensable for pathogenicity in an intranasal mouse infection model. Eukaryot Cell 6:2290–2302

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971

Liebmann B, Gattung S, Jahn B, Brakhage AA (2003) cAMP signaling in *Aspergillus fumigatus* is involved in the regulation of the virulence gene *pksP* and in defense against killing by macrophages. Mol Genet Genomics 269:420–435

Linde J, Hortschansky P, Fazius E, Brakhage AA, Guthke R, Haas H (2012) Regulatory interactions for iron homeostasis in *Aspergillus fumigatus* inferred by a systems biology approach. BMC Syst Biol 6:6

Liu H, Lin D, Yates JR, IIIrd (2002) Multidimensional separations for protein/peptide analysis in the post-genomic era. Biotechniques 32:898, 900, 902 passim

Liu TT, Lee RE, Barker KS, Lee RE, Wei L, Homayouni R, Rogers PD (2005) Genome-wide expression profiling of the response to azole, polyene, echinocandin, and pyrimidine antifungal agents in *Candida albicans*. Antimicrob Agents Chemother 49:2226–2236

Mabey JE, Anderson MJ, Giles PF, Miller CJ, Attwood TK, Paton NW, Bornberg-Bauer E, Robson GD, Oliver SG, Denning DW (2004) CADRE: the central *Aspergillus* data REpository. Nucleic Acids Res 32: D401–D405

McDonagh A, Fedorova ND, Crabtree J, Yu Y, Kim S, Chen D, Loss O, Cairns T, Goldman G, Armstrong-James D, Haynes K, Haas H, Schrettl M, May G, Nierman WC, Bignell E (2008) Sub-telomere directed gene expression during initiation of invasive aspergillosis. PLoS Pathog 4:e1000154

Miller I, Crawford J, Gianazza E (2006) Protein stains for proteomic applications: which, when, why? Proteomics 6:5385–5408

Missall TA, Lodge JK, McEwen JE (2004) Mechanisms of resistance to oxidative and nitrosative stress: implications for fungal survival in mammalian hosts. Eukaryot Cell 3:835–846

Missall TA, Pusateri ME, Donlin MJ, Chambers KT, Corbett JA, Lodge JK (2006) Posttranslational, translational, and transcriptional responses to nitric oxide stress in *Cryptococcus neoformans*: implications for virulence. Eukaryot Cell 5:518–529

Monteoliva L, Martinez-Lopez R, Pitarch A, Hernaez ML, Serna A, Nombela C, Albar JP, Gil C (2011) Quantitative proteome and acidic subproteome profiling of *Candida albicans* yeast-to-hypha transition. J Proteome Res 10:502–517

Muller S, Baldin C, Groth M, Guthke R, Kniemeyer O, Brakhage AA, Valiante V (2012) Comparison of transcriptome technologies in the pathogenic fungus *Aspergillus fumigatus* reveals novel insights into the genome and MpkA dependent gene expression. BMC Genomics 13:519

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 51(Pt 1):263–273

Murillo LA, Newport G, Lan CY, Habelitz S, Dungan J, Agabian NM (2005) Genome-wide transcription profiling of the early phase of biofilm formation by *Candida albicans*. Eukaryot Cell 4:1562–1573

Muszkieta L, Beauvais A, Pahtz V, Gibbons JG, Anton Leberre V, Beau R, Shibuya K, Rokas A, Francois JM, Kniemeyer O, Brakhage AA, Latge JP (2013) Investigation of *Aspergillus fumigatus* biofilm formation by various "omics" approaches. Front Microbiol 4:13

Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH (2006) Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet Biol 43:54–64

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafon A, Latge JP, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Pertea M, Price C, Pritchard BL, Quail MA, Rabbinowitsch E, Rawlins N, Rajandream MA, Reichard U, Renauld H, Robson GD, Rodriguez de Cordoba S, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, Vazquez de Aldana CR, Weidman J, White O, Woodward J, Yu JH, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438:1151–1156

Nikolov M, Schmidt C, Urlaub H (2012) Quantitative mass spectrometry-based proteomics: an overview. Methods Mol Biol 893:85–100

Noble SM, French S, Kohn LA, Chen V, Johnson AD (2010) Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. Nat Genet 42:590–598

Nutzmann HW, Reyes-Dominguez Y, Scherlach K, Schroeckh V, Horn F, Gacek A, Schumann J, Hertweck C, Strauss J, Brakhage AA (2011) Bacteria-induced natural product formation in the fungus *Aspergillus nidulans* requires Saga/Ada-mediated histone acetylation. Proc Natl Acad Sci U S A 108:14282–14287

Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 1:252–262

Ormerod KL, Morrow CA, Chow EW, Lee IR, Arras SD, Schirra HJ, Cox GM, Fries BC, Fraser JA (2013) Comparative genomics of serial isolates of Cryptococcus neoformans reveals gene associated with carbon utilization and virulence. G3 (Bethesda) pii:g.113.005660v1. doi:10.1534/g3.113005660. [Epub ahead of print]

Palmer JM, Keller NP (2010) Secondary metabolism in fungi: does chromosomal location matter? Curr Opin Microbiol 13:431–436

Palmer JM, Bok JW, Lee S, Dagenais TR, Andes DR, Kontoyiannis DP, Keller NP (2013) Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in *Aspergillus fumigatus*. Peer J 1:e4

Park BJ, Pappas PG, Wannemuehler KA, Alexander BD, Anaissie EJ, Andes DR, Baddley JW, Brown JM, Brumble LM, Freifeld AG, Hadley S, Herwaldt L, Ito JI, Kauffman CA, Lyon GM, Marr KA, Morrison VA, Papanicolaou G, Patterson TF, Perl TM, Schuster MG, Walker R, Wingard JR, Walsh TJ, Kontoyiannis DP (2011) Invasive non-*Aspergillus* mold infections in transplant recipients, United States, 2001–2006. Emerg Infect Dis 17:1855–1864

Patron NJ, Waller RF, Cozijnsen AJ, Straney DC, Gardiner DM, Nierman WC, Howlett BJ (2007) Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol Biol 7:174

Patterson M, Dhand R (2000) Nature insight: functional genomics. Nature 405:819

Pechanova O, Pechan T, Rodriguez JM, Williams WP, Brown AE (2013) A two-dimensional proteome map of the aflatoxigenic fungus *Aspergillus flavus*. Proteomics 13:1513–1518

Pfaller MA, Diekema DJ (2007) Epidemiology of invasive candidiasis: a persistent public health problem. Clin Microbiol Rev 20:133–163

Rabilloud T, Lelong C (2011) Two-dimensional gel electrophoresis in proteomics: a tutorial. J Proteomics 74:1829–1841

Rai MN, Balusu S, Gorityala N, Dandu L, Kaur R (2012) Functional genomic analysis of *Candida glabrata*-macrophage interaction: role of chromatin remodeling in virulence. PLoS Pathog 8:e1002863

Reyes-Dominguez Y, Bok JW, Berger H, Shwab EK, Basheer A, Gallmetzer A, Scazzocchio C, Keller N, Strauss J (2010) Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. Mol Microbiol 76:1376–1386

Rispail N, Soanes DM, Ant C, Czajkowski R, Grunler A, Huguet R, Perez-Nadales E, Poli A, Sartorel E, Valiante V, Yang M, Beffa R, Brakhage AA, Gow NA, Kahmann R, Lebrun MH, Lenasi H, Perez-Martin J, Talbot NJ, Wendland J, Di Pietro A (2009) Comparative genomics of MAP kinase and calcium-calcineurin signalling components in plant and human pathogenic fungi. Fungal Genet Biol 46:287–298

Rodrigues ML, Nakayasu ES, Oliveira DL, Nimrichter L, Nosanchuk JD, Almeida IC, Casadevall A (2008) Extracellular vesicles produced by *Cryptococcus neoformans* contain protein components associated with virulence. Eukaryot Cell 7:58–67

Rodrigues ML, Nakayasu ES, Almeida IC, Nimrichter L (2013) The impact of proteomics on the understanding of functions and biogenesis of fungal extracellular vesicles. J Proteomics pii:S1874–3919(13)00182–6. doi:10.1016/j.jprot.2013.04.001. [Epub ahead of print]

Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res 32:5539–5545

Scharf DH, Remme N, Heinekamp T, Hortschansky P, Brakhage AA, Hertweck C (2010) Transannular disulfide formation in gliotoxin biosynthesis and its role in self-resistance of the human pathogen *Aspergillus fumigatus*. J Am Chem Soc 132:10136–10141

Scharf DH, Remme N, Habel A, Chankhamjon P, Scherlach K, Heinekamp T, Hortschansky P, Brakhage AA, Hertweck C (2011) A dedicated glutathione S-transferase mediates carbon-sulfur bond formation in gliotoxin biosynthesis. J Am Chem Soc 133:12322–12325

Scharf DH, Heinekamp T, Remme N, Hortschansky P, Brakhage AA, Hertweck C (2012) Biosynthesis and function of gliotoxin in *Aspergillus fumigatus*. Appl Microbiol Biotechnol 93:467–472

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Scherlach K, Nutzmann HW, Schroeckh V, Dahse HM, Brakhage AA, Hertweck C (2011) Cytotoxic pheofungins from an engineered fungus impaired in posttranslational protein modification. Angew Chem Int Ed Engl 50:9843–9847

Schrettl M, Kim HS, Eisendle M, Kragl C, Nierman WC, Heinekamp T, Werner ER, Jacobsen I, Illmer P, Yi H, Brakhage AA, Haas H (2008) SreA-mediated iron regulation in *Aspergillus fumigatus*. Mol Microbiol 70:27–43

Schrettl M, Carberry S, Kavanagh K, Haas H, Jones GW, O'Brien J, Nolan A, Stephens J, Fenelon O, Doyle S (2010) Self-protection against gliotoxin – a component of the gliotoxin biosynthetic cluster, GliT, completely protects *Aspergillus fumigatus* against exogenous gliotoxin. PLoS Pathog 6:e1000952

Schwienbacher M, Weig M, Thies S, Regula JT, Heesemann J, Ebel F (2005) Analysis of the major pro-

teins secreted by the human opportunistic pathogen *Aspergillus fumigatus* under in vitro conditions. Med Mycol 43:623–630

Seider K, Heyken A, Luttich A, Miramon P, Hube B (2010) Interaction of pathogenic yeasts with phagocytes: survival, persistence and escape. Curr Opin Microbiol 13:392–400

Seneviratne CJ, Jin L, Samaranayake LP (2008) Biofilm lifestyle of *Candida*: a mini review. Oral Dis 14:582–590

Shimizu M, Masuo S, Fujita T, Doi Y, Kamimura Y, Takaya N (2012) Hydrolase controls cellular NAD, sirtuin, and secondary metabolites. Mol Cell Biol 32:3743–3755

Shwab EK, Bok JW, Tribus M, Galehr J, Graessle S, Keller NP (2007) Histone deacetylase activity regulates chemical diversity in *Aspergillus*. Eukaryot Cell 6:1656–1664

Sorgo AG, Heilmann CJ, Dekker HL, Brul S, de Koster CG, Klis FM (2010) Mass spectrometric analysis of the secretome of *Candida albicans*. Yeast 27:661–672

Sorgo AG, Heilmann CJ, Brul S, de Koster CG, Klis FM (2013) Beyond the wall: *Candida albicans* secret(e)s to survive. FEMS Microbiol Lett 338:10–17

Sosinska GJ, de Groot PW, Teixeira de Mattos MJ, Dekker HL, de Koster CG, Hellingwerf KJ, Klis FM (2008) Hypoxic conditions and iron restriction affect the cell-wall proteome of *Candida albicans* grown under vagina-simulative conditions. Microbiology 154:510–520

Sosinska GJ, de Koning LJ, de Groot PW, Manders EM, Dekker HL, Hellingwerf KJ, de Koster CG, Klis FM (2011) Mass spectrometric quantification of the adaptations in the wall proteome of *Candida albicans* in response to ambient pH. Microbiology 157:136–146

Soukup AA, Chiang YM, Bok JW, Reyes-Dominguez Y, Oakley BR, Wang CC, Strauss J, Keller NP (2012) Overexpression of the *Aspergillus nidulans* histone 4 acetyltransferase EsaA increases activation of secondary metabolite production. Mol Microbiol 86:314–330

Sriranganadane D, Waridel P, Salamin K, Reichard U, Grouzmann E, Neuhaus JM, Quadroni M, Monod M (2010) *Aspergillus* protein degradation pathways with different secreted protease sets at neutral and acidic pH. J Proteome Res 9:3511–3519

Steen H, Mann M (2004) The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol 5:699–711

Steinbach WJ, Cramer RA Jr, Perfect BZ, Asfaw YG, Sauer TC, Najvar LK, Kirkpatrick WR, Patterson TF, Benjamin DK Jr, Heitman J, Perfect JR (2006) Calcineurin controls growth, morphology, and pathogenicity in *Aspergillus fumigatus*. Eukaryot Cell 5:1091–1103

Teutschbein J, Albrecht D, Potsch M, Guthke R, Aimanianda V, Clavaud C, Latge JP, Brakhage AA, Kniemeyer O (2010) Proteome profiling and functional classification of intracellular proteins from conidia of the human-pathogenic mold *Aspergillus fumigatus*. J Proteome Res 9:3427–3442

Thomas DP, Bachmann SP, Lopez-Ribot JL (2006) Proteomics for the analysis of the *Candida albicans* biofilm lifestyle. Proteomics 6:5795–5804

Thywißen A, Heinekamp T, Dahse H-M, Schmaler-Ripcke J, Nietsche S, Zipfel PF, Brakhage AA (2011) Front Microbiol 2. doi:10.3389/fmicb201100096

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18:2071–2077

Valiante V, Heinekamp T, Jain R, Hartl A, Brakhage AA (2008) The mitogen-activated protein kinase MpkA of *Aspergillus fumigatus* regulates cell wall signaling and oxidative stress response. Fungal Genet Biol 45:618–627

Valiante V, Jain R, Heinekamp T, Brakhage AA (2009) The MpkA MAP kinase module regulates cell wall integrity signaling and pyomelanin formation in *Aspergillus fumigatus*. Fungal Genet Biol 46:909–918

Vallejo MC, Nakayasu ES, Matsuo AL, Sobreira TJ, Longo LV, Ganiko L, Almeida IC, Puccia R (2012) Vesicle and vesicle-free extracellular proteome of *Paracoccidioides brasiliensis*: comparative analysis with other pathogenic fungi. J Proteome Res 11:1676–1685

Viswanathan S, Unlu M, Minden JS (2006) Two-dimensional difference gel electrophoresis. Nat Protoc 1:1351–1358

Vodisch M, Albrecht D, Lessing F, Schmidt AD, Winkler R, Guthke R, Brakhage AA, Kniemeyer O (2009) Two-dimensional proteome reference maps for the human pathogenic filamentous fungus *Aspergillus fumigatus*. Proteomics 9:1407–1415

Vodisch M, Scherlach K, Winkler R, Hertweck C, Braun HP, Roth M, Haas H, Werner ER, Brakhage AA, Kniemeyer O (2011) Analysis of the *Aspergillus fumigatus* proteome reveals metabolic changes and the activation of the pseurotin A biosynthesis gene cluster in response to hypoxia. J Proteome Res 10:2508–2524

Vodovotz Y, Constantine G, Faeder J, Mi Q, Rubin J, Bartels J, Sarkar J, Squires RH Jr, Okonkwo DO, Gerlach J, Zamora R, Luckhart S, Ermentrout B, An G (2010) Translational systems approaches to the biology of inflammation and healing. Immunopharmacol Immunotoxicol 32:181–195

Volling K, Thywissen A, Brakhage AA, Saluz HP (2011) Phagocytosis of melanized *Aspergillus* conidia by macrophages exerts cytoprotective effects by sustained PI3K/Akt signalling. Cell Microbiol 13:1130–1148

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63

Wartenberg D, Lapp K, Jacobsen ID, Dahse HM, Kniemeyer O, Heinekamp T, Brakhage AA (2011) Secretome analysis of *Aspergillus fumigatus* reveals Asp-hemolysin as a major secreted protein. Int J Med Microbiol 301:602–611

Weig M, Jansch L, Gross U, De Koster CG, Klis FM, De Groot PW (2004) Systematic identification in silico of covalently bound cell wall proteins and analysis of protein-polysaccharide linkages of the human pathogen *Candida glabrata*. Microbiology 150:3129–3144

Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-seq of pathogen and host. Nat Rev Microbiol 10:618–630

Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, Williams KL (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev 13:19–50

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Veronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285:901–906

Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11:49–79

Yin Z, Stead D, Walker J, Selway L, Smith DA, Brown AJ, Quinn J (2009) A proteomic analysis of the salt, cadmium and peroxide stress responses in *Candida albicans* and the role of the Hog1 stress-activated MAPK in regulating the stress-induced proteome. Proteomics 9:4686–4703

14 Integration of Metabolism with Virulence in *Candida albicans*

Iuliana V. Ene[1,2], Alistair J.P. Brown[1]

CONTENTS

Abbreviations

3AT	3-aminotriazole
bHLH	β-helix loop helix domain
GCN response	General amino acid control
ROS	Reactive oxygen species
GCRE	GCN Response Element
uORF	upstream Open Reading Frame

I. Introduction

Candida albicans is a major systemic pathogen of humans (Odds 1988; Calderone 2002; Calderone and Clancy 2011). This fungus exists as a relatively harmless commensal organism in the gut, urogenital tract and skin of at least half of all individuals. When host immune defences become compromised or when there is an imbalance of the competing bacterial microflora, *C. albicans* can cause mucocutaneous infections such as oral or vaginal candidiasis. About 75 % of women suffer at least one vaginal *Candida* infection in their lifetime, and a significant proportion of these infections are recurrent (Kim and Sudbery 2011). In severely immunocompromised individuals, *C. albicans* can establish deep-seated systemic infections. For example, in patients undergoing chemotherapy or organ transplant surgery, the fungus is able to disseminate via the bloodstream and colonise internal organs. Systemic candidiasis is currently the third most common cause of nosocomial bloodstream infections and is fatal in over half of these cases (Perlroth et al. 2007). Disseminated candidiasis may affect the kidneys, liver, spleen, brain, eyes, heart and other tissues (Pfaller and Diekema 2007; Calderone and Clancy 2011).

Risk factors that are thought to predispose patients to candidemia include protracted neutropenia (reflecting the importance of neutrophils in host defences); diabetes; the use of broad-spectrum antibiotics (which eliminate the endogenous bacterial flora); iatrogenic procedures such as surgery; and the application of indwelling intravenous devices (e.g. catheters) (Harris et al. 1999; Koh et al. 2008; Calderone and Clancy 2011). The significance of catheters relates to the ability of *C. albicans* to form biofilms on the surfaces of these devices (Blankenship and Mitchell 2006), with the biofilms becoming a source of recurrent infection. Furthermore, cells in these microbial communities exhibit phenotypes that are distinct from those

[1]Aberdeen Fungal Group, School of Medical Sciences, Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK; e-mail: al.brown@abdn.ac.uk
[2]Department of Molecular Microbiology and Immunology, Brown University, Providence, RI 02912, USA

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

of free-living cells and that often make them drug resistant (Tobudic et al. 2012). Hence, biofilms represent another medically relevant microenvironment in which *C. albicans* can thrive. Of particular relevance to this review is the observation that the fungus adjusts its metabolic activities during growth in biofilms (Garcia-Sanchez et al. 2004; Zhu et al. 2012), and that inappropriate regulation of blood glucose levels in patients predisposes them to candidiasis (Harris et al. 1999).

Several factors are thought to promote the virulence of *C. albicans* (Calderone and Clancy 2011; Gow et al. 2011). Yeast-hypha morphogenesis is thought to promote fungal dissemination and tissue invasion by active penetration and by inducing endocytosis by the host cells (Phan et al. 2007; Brand 2012). Adhesins are thought to promote adherence to host tissue, biofilm formation and colonisation (Hoyer et al. 2008; Nobile et al. 2008). Secreted aspartyl proteinases and lipases may promote invasion, counteract host defences, and provide nutrients (Naglik et al. 2004). High-frequency switching between different phenotypic forms probably helps the fungus to evade host defences and colonise certain anatomical niches (Lachke et al. 2003; Lohse and Johnson 2008). These virulence attributes are thought to be required to differing extents during disease establishment and progression (Brown and Gow 1999; Huang 2012).

Some years ago, we predicted that the regulation of metabolism and virulence might be mechanistically linked in *C. albicans* (Brown and Gow 1999; Brown et al. 2000). We reasoned that these links might allow this pathogen to adjust its metabolic programme in parallel with its portfolio of virulence attributes in response to the new microenvironments it encounters during disease establishment and progression. In this chapter, we review recent studies that address this prediction. Our theme is the integration of metabolic regulation and virulence, rather than simply the assimilation of nutrients. Therefore, for information on the assimilation of essential micronutrients such as iron and zinc, we refer readers elsewhere (Almeida et al. 2009; Citiulo et al. 2012).

It is worth noting that an improved molecular toolbox has had a significant impact on our understanding of this field because *C. albicans* was considered to be an experimentally intransigent fungus. In particular, genomics has uncovered unexpected links between metabolism and virulence in *C. albicans* that were not exposed by pregenomic experimentation. One of the greatest advances has been the development of the *C. albicans* postgenomics era, which has facilitated global examinations of the roles of specific genes and proteins in a relatively unbiased fashion. As described in the following material, this led to the finding that the regulation of metabolic and virulence functions seems to be integrated in this pathogen. It is apparent that we are just beginning to understand host-pathogen interactions and the importance of metabolic pathways during pathogenesis.

II. Overview of Metabolism in *Candida albicans*

The ability of pathogenic microbes to assimilate and utilise essential nutrients from their host niches is a fundamental aspect of infection. Pathogens such as *C. albicans*, which thrive within diverse niches such as the skin, mucosal surfaces, blood, and internal organs and in biofilms (Calderone and Clancy 2011), must display sufficient metabolic flexibility to assimilate and utilise the available nutrients in these contrasting niches.

Investigations of *C. albicans* metabolism began over half a century ago (Van Neil and Cohen 1942), and enzymes of central carbon metabolism were characterised as long ago as 1960 by Rao et al. (1960). In general, studies of *C. albicans* metabolism have focused either directly or indirectly on virulence attributes or the mode of action of antifungal drugs (Odds 1988). For example, considerable attention has been paid to the pathways involved in ergosterol biosynthesis and cell wall biogenesis, both of which are antifungal targets (Sanglard et al. 2003; Shapiro et al. 2011). Pathways of central and amino acid metabolism have been highlighted as potential antifungal targets by large-scale gene deletion studies in *C. albicans* (Roemer et al. 2003) and by transcript-profiling studies (Lorenz and Fink 2001, 2002). With respect to antifungal therapy, the priority is to kill the

fungus rather than to inhibit its virulence. Hence, metabolic functions that are essential for growth and survival in the host represent potential antifungal targets (Brock 2009). Of course, the validity of such targets and the efficacy of drugs that hit these targets will depend on the degree of cross-reactivity with the host. However, *C. albicans* also requires considerable metabolic flexibility to adapt to the differential nutritional availabilities in diverse host microenvironments. Therefore, an increased understanding of *C. albicans* metabolism in such niches should facilitate the improvement of existing therapies and possibly the development of new therapeutic approaches. Indeed, over the last decade, considerable efforts have been made to understand the metabolic requirements of *C. albicans* during infection. However, additional studies are required to appreciate the broad metabolic toolbox that contributes to the fitness of this pathogen and the extent to which metabolic adaptation is coordinated with the regulation of virulence attributes. Even now, *C. albicans* metabolism is not addressed directly in the latest definitive text on *Candida* and candidiasis (Calderone and Clancy 2011).

To date, *Saccharomyces cerevisiae* has provided a reasonable metabolic paradigm for *C. albicans*. The pathways of central carbon metabolism are conserved in this fungus, including the glycolytic, gluconeogenic and pentose phosphate pathways and the tricarboxylic and glyoxylate cycles (Odds 1988; Jones et al. 2004). Pathways for the generation of storage and cell wall carbohydrates are conserved, while pathways of amino acid, lipid and nucleotide assimilation and anabolism also seem to be conserved. However, significant metabolic differences do exist between *C. albicans* and *S. cerevisiae*, the most obvious of which relates to their patterns of sugar utilisation (Askew et al. 2009; Brown et al. 2009). Indeed, differences in the patterns of carbohydrate assimilation are used routinely to distinguish *C. albicans* from other microbes in the clinical setting (Williamson et al. 1986).

A whole-genome comparison between *S. cerevisiae* and *C. albicans* has revealed further differences in their central carbon metabolism (Jones et al. 2004). Unlike *S. cerevisiae*, *C. albicans* possesses both mitochondrial- and nuclear-encoded subunits of electron transport complex I, a pyruvate dehydrogenase kinase (which regulates the flow from glycolysis into the TCA [tricarboxylic acid] cycle), a large lipase gene family and additional enzymes involved in fatty acid and amino acid catabolism. On this basis, the authors proposed an increased role for respiration and oxidative metabolism in *C. albicans* (Jones et al. 2004). This view is consistent with the observation that *C. albicans* converts a lower proportion of glucose to ethanol in comparison to *S. cerevisiae* (Bertram et al. 1996).

With regard to the *regulation* of metabolism, the *S. cerevisiae* paradigm is less robust (Brown 2005; Sabina and Brown 2009). This is hardly surprising given that the pathogen *C. albicans* and the relatively benign *S. cerevisiae* have evolved in fundamentally different niches. In *S. cerevisiae*, the balance between fermentation and respiratory metabolism is modulated by glucose concentration, oxygen availability and growth rate (Gancedo 1998). *S. cerevisiae* predominantly ferments sugars to ethanol even under aerobic conditions instead of using respiration (Johnston 1999), a phenomenon known as crabtree effect. When fermentable carbon sources become depleted, *S. cerevisiae* switches to respiration during the diauxic shift. A set of transcriptional repressors, including Mig1, Rgt1 and Snf1, control the glucose repression circuitry in this yeast (Johnston 1999). *Candida albicans* also ferments glucose to ethanol and displays glucose repression at the transcriptional level (Lorenz et al. 2004; Rodaki et al. 2009). However, *C. albicans* has been classified as a **glucose-negative yeast** because it retains respiratory activity even in the presence of glucose, and carbon source metabolism occurs in response to oxygen availability (Johnston 1999; Niimi et al. 1988).

In *C. albicans*, at least some glucose-repressible functions seem to be regulated by the transcriptional repressor Mig1 in a fashion analogous to that of *S. cerevisiae* (Murad et al. 2001a, b). Furthermore, *C. albicans* contains a homolog of *S. cerevisiae* Snf1 (Petter et al. 1997), which is required for the derepression of glucose-repressed genes in budding yeast. However, *C. albicans* Mig1 displays several differences

vis-à-vis *S. cerevisiae* Mig1, including the lack of a putative Snf1 phosphorylation site (Zaragoza et al. 2000). Also, *S. cerevisiae snf1* mutants are viable, whereas Snf1 seems to be essential for viability in *C. albicans* (Petter et al. 1997), suggesting that Snf1 might execute additional functions in this yeast. Therefore, although many of the main players are conserved in *C. albicans* and *S. cerevisiae*, significant differences exist between these yeasts with respect to glucose repression mechanisms (Sabina and Brown 2009).

The differences extend to the regulation of hypoxic gene expression in *C. albicans* and *S. cerevisiae*. For example, in *S. cerevisiae*, the transcriptional regulator Rox1 controls hypoxic gene expression in conjunction with the Ssn6/Tup1 general repression complex and in a haem-dependent fashion (Kastaniotis and Zitomer 2000). In contrast, its ortholog in *C. albicans*, Rfg1, controls filamentous growth and virulence but not hypoxic gene expression (Khalaf and Zitomer 2001; Kadosh and Johnson 2001).

Bioinformatic analyses of microarray data sets have provided global perspectives of metabolic regulation and suggested significant differences between *S. cerevisiae* and *C. albicans* (Ihmels et al. 2005; Martchenko et al. 2007; Lorenz et al. 2004). The ability to grow anaerobically in yeast has been associated with **global transcriptional rewiring**, involving changes in promoter regions and loss of regulatory motifs (Ihmels et al. 2005). Ihmels and coworkers reported that the expression of MRP (mitochondrial ribosomal protein) genes correlated strongly with the expression of ribosome assembly genes in *C. albicans*, but not in *S. cerevisiae*, and pointed out that this correlation reflects the respective physiologies of these divergent yeasts (Ihmels et al. 2005). Indeed, recent work has emphasized that significant transcriptional rewiring and divergence have occurred among these ascomycetes, and that the regulatory circuits controlling central carbon metabolism are distinct between these two species, some of these findings supporting the connection between pathogenicity and metabolism (Martchenko et al. 2007; Askew et al. 2009; Lavoie et al. 2009). For instance, Leloir pathway genes (*GAL*), which encode galactose utilisation enzymes, are syntenically arranged in both yeasts. However, their upstream regulator regions are completely distinct (Martchenko et al. 2007). This is significant because dietary galactose is thought to be an important source of carbon for *C. albicans* (Sabina and Brown 2009). *C. albicans* also lacks homologs for *GCR1* and *GCR2*, which are the main regulators of glycolysis in *S. cerevisiae* (Askew et al. 2009). Instead, in *C. albicans*, Tye7 and Gal4 are the key regulators in a fundamentally distinct glycolytic transcriptional circuit and are required for full virulence in the host (Askew et al. 2009).

Recent work has revealed that this regulatory rewiring extends to posttranscriptional circuitry. The rewiring of ubiquitination targets amongst central metabolic enzymes (Leach et al. 2011; Sandai et al. 2012) has led to dramatic differences between *C. albicans* and *S. cerevisiae* with regard to the stability of the glyoxylate cycle and gluconeogenic enzymes following exposure to glucose. Whilst these enzymes are rapidly degraded in *S. cerevisiae* cells following glucose exposure, they are retained by *C. albicans* (Sandai et al. 2012). As a result, when glucose becomes available, *S. cerevisiae* rapidly switches from carbon anabolism to fermentation. In contrast, *C. albicans* can simultaneously assimilate glucose and alternative carbon sources (Sandai et al. 2012), which probably contributes to its fitness in complex host niches (Sect. IV).

To summarise, whilst many metabolic pathways are conserved between *S. cerevisiae* and *C. albicans*, these organisms display significant differences in their metabolic programmes. It is likely that the high level of plasticity in the regulatory circuits that control the metabolic machinery reflects the evolutionary adaptation of yeasts to their respective niches and, in the case of *C. albicans*, to the diverse microenvironments it encounters in the host during colonisation and disease progression.

III. Integration of Amino Acid Metabolism with Virulence

A. Amino Acid Starvation and Virulence

Amino acids represent an important source of both carbon and nitrogen. In *S. cerevisiae*, starvation for a single amino acid triggers the

expression of almost all amino acid biosynthetic pathways, and this is termed **general amino acid control** or the **GCN response** (Hinnebusch 1988). The GCN response has been well characterised at the molecular level in *S. cerevisiae* and relies on the key regulators Gcn4 and Gcn2 (Hinnebusch 1988; Natarajan et al. 2001; Hinnebusch and Natarajan 2002). Briefly, amino acid starvation leads to the intracellular accumulation of uncharged transfer RNAs (tRNAs), which bind to the histidyl tRNA synthetase-like domain of Gcn2 (Wek et al. 1995). This activates the protein kinase activity of Gcn2, which then phosphorylates the α-subunit of the translation initiation factor eIF2 (Dever et al. 1992). eIF2α phosphorylation increases the affinity of eIF2 for eIF2B, its guanine nucleotide exchange factor, and inhibits the regeneration of recharged eIF2-GTP required for translation initiation (Krishnamoorthy et al. 2001). This leads not only to a decrease in the overall rate of translation initiation in the yeast cell but also to an increase in the translation of the *GCN4* messenger RNA (mRNA).

The translation of the *GCN4* mRNA is regulated by four short upstream Open Reading Frames (uORFs), which lie in its unusually long 5′-leader region (Mueller and Hinnebusch 1986). Essentially, this 5′-leader normally acts to repress the translation of the main Gcn4-encoding ORF on the *GCN4* mRNA under non-starvation conditions. However, in response to amino acid starvation, this repression is released when eIF2-GTP levels are reduced (Mueller and Hinnebusch 1986). The resultant increase in *GCN4* mRNA translation leads to an elevation in the abundance of the Gcn4 protein. The transcription factor Gcn4 binds as a dimer to GCRE elements [5′-TGA(C/G)TCA] located in the promoters of target genes and activates their transcription (Oliphant et al. 1989; Ellenberger et al. 1992). These target genes include amino acid biosynthetic enzymes on all amino acid biosynthetic pathways, with the exception of the cysteine pathway (Natarajan et al. 2001). Therefore, amino acid biosynthesis is induced in response to amino acid starvation via a signalling pathway involving the key regulators Gcn2 and Gcn4.

Transcript profiling of amino acid starvation in *S. cerevisiae* has revealed that the GCN response is much broader than was initially expected (Natarajan et al. 2001). In these experiments, the histidine analog 3-aminotriazole (3AT) was used to activate the GCN response. Exposure to 3AT, which causes histidine starvation, is a classic means of activating the GCN response (Hinnebusch 1988). The transcript-profiling experiments showed that, in addition to inducing the expression of amino acid biosynthetic genes, Gcn4 induces the expression of aminoacyl tRNA synthetases, amino acid transporters, vitamin biosynthetic pathways and glycogen biosynthetic functions and represses the expression of ribosomal proteins (Natarajan et al. 2001). The transcript profiling also confirmed the earlier observation that Gcn4 regulates purine biosynthetic functions (*ADE* genes; Rolfes and Hinnebusch 1993), while controlling the expression of numerous transcription factors, including Arg80, Leu3, Lys14, Met4, Met28, Bas1, Gln3, Rtg3, Pip2, Gat1, Uga3, Mal13, Cup9 and Rim101. This led Natarajan et al. (2001) to propose Gcn4 as a master regulator of gene expression in *S. cerevisiae*.

Gcn4-like proteins are conserved in other fungi (Paluh et al. 1988; Wanke et al. 1997; Hoffmann et al. 2001; Tripathi et al. 2002; Tournu et al. 2005). Furthermore, the mRNAs encoding these proteins carry unusually long 5′-leader regions with multiple short uORFs. This suggested that these mRNAs might be regulated at the translational level, like the *S. cerevisiae GCN4* mRNA. *Saccharomyces cerevisiae* Gcn4 levels are also regulated at the transcriptional level and via accelerated protein turnover (Kornitzer et al. 1994; Albrecht et al. 1998), but most control seems to be executed at the translational level (Hinnebusch and Natarajan 2002). The expression of *GCN4*-like genes in *C. albicans* and *Aspergillus nidulans* seems to be regulated primarily at the transcriptional level (Hoffmann et al. 2001; Tournu et al. 2005). Therefore, subtle differences do exist between fungi with respect to the regulation of their GCN responses. Transcript-profiling and proteomic studies have revealed that the global GCN response itself, as defined in *S. cerevisiae*, has been conserved in *C. albicans*, but that some significant differences have emerged during evolution (Natarajan et al. 2001; Tripathi et al. 2002; Yin et al. 2004; Tournu et al. 2005).

A significant proportion of the *S. cerevisiae* proteome (11 %) responds to 3AT (Yin et al. 2004), which is roughly analogous to the corresponding *S. cerevisiae* transcriptome (Natarajan et al. 2001). Forty-five percent of the Gcn4-dependent changes identified in the proteome correspond to amino acid biosynthetic enzymes on the arginine, asparagine, cysteine, chorismate, glutamine, glycine, histidine, isoleucine-valine, leucine, lysine, methionine, serine, threonine and tryptophan pathways (Yin et al. 2004), confirming that histidine starvation (using 3AT) invokes a broad response across most (if not all) amino acid biosynthetic pathways. In addition, *S. cerevisiae* purine biosynthetic enzymes (Ade proteins), vitamin and cofactor biosynthetic enzymes, and proteins involved in carbon metabolism and energy generation are induced by 3AT in a Gcn4-dependent fashion. Similarly, about 30 % of 3AT-induced changes in the *C. albicans* proteome correspond to amino acid biosynthetic enzymes on the arginine, chorismate, isoleucine-valine, leucine, lysine, methionine and serine pathways (Yin et al. 2004). *Candida albicans* proteins involved in carbon metabolism and energy generation are also elevated by 3AT in a Gcn4-dependent fashion. Therefore, the *C. albicans* and *S. cerevisiae* GCN proteomes display a high degree of similarity. However, in *C. albicans ADE* genes and proteins are not induced in response to 3AT, unlike *S. cerevisiae* (Yin et al. 2004; Tournu et al. 2005).

Molecular studies have revealed another significant difference between the GCN responses of *S. cerevisiae* and *C. albicans*. Amino acid starvation induces cellular morphogenesis in the pathogenic fungus via interaction of the master regulator Gcn4 but does not do so in *S. cerevisiae* (Tripathi et al. 2002). The exact molecular mechanism by which Gcn4 stimulates morphogenesis in *C. albicans* is not known. However, Gcn4 is known to activate morphogenesis specifically via the Ras-cAMP signalling pathway (Tripathi et al. 2002). Mutations on the mitogen-activated protein kinase (MAPK) pathway do not prevent the ability of ectopically expressed Gcn4 to stimulate morphogenesis, but the inactivation of Efg1 blocks this response. This indicates that the formation of *C. albicans* pseudohyphae in response to amino acid starvation is regulated by the transcriptional activator Gcn4 in an Efg1-dependent fashion. Hence, Gcn4 seems to coordinate morphogenesis with amino acid metabolism by connecting specific signalling pathways.

Is the GCN response required for the virulence of *C. albicans*? The inactivation of Gcn4 does not attenuate the virulence of this fungus in the mouse model of systemic candidiasis (Brand et al. 2004). However, Gcn4 is required for *C. albicans* to form normal biofilms (Garcia-Sanchez et al. 2004). Hence, the GCN response does seem to be required for survival in at least one medically important niche. It remains to be seen whether *C. albicans* depends on the GCN response in other infection sites, for example, during oral candidiasis, vaginitis or during commensalism.

B. Amino Acid Metabolism and Virulence

As described in Sect. III.A, a detailed examination of the GCN response in *C. albicans* has revealed a well-defined molecular link between a morphogenetic and metabolic response (Tripathi et al. 2002). Amino acid starvation in *C. albicans* stimulates pseudohyphal development as well as the induction of amino acid biosynthetic genes and other metabolic functions. Both of these responses depend on the transcription factor Gcn4 (Tripathi et al. 2002). Hence, through Gcn4, the regulation of a virulence attribute is integrated with the control of amino acid metabolism in *C. albicans*. This section reviews further experimental observations that lend weight to the view that the control of virulence and metabolism is integrated in this pathogen.

Further mechanistic links between morphogenesis and amino acid metabolism involve an amino acid sensor (Csy1) and an ammonium transporter (Mep2). Csy1 is thought to be the main amino acid sensor in *C. albicans*. Its inactivation blocks the transcriptional induction of amino acid permease genes in response to the presence of amino acids in the growth medium (Brega et al. 2004). *Candida albicans csy1* mutants display morphogenetic defects in response to serum and pH induction. However, the molecular mechanisms by which

Csy1 links amino acid sensing with cellular morphogenesis remain to be defined. It is attractive to speculate that, like Gcn4, Csy1 might mediate its morphogenetic effects via the Ras-cAMP pathway. In addition, Mep2 mediates ammonium uptake, and its expression is downregulated at high ammonium concentrations (Biswas and Morschhauser 2005). When nitrogen sources limit growth, the ammonium permeases Mep1 and Mep2 are responsible for uptake of ammonium and its utilization as a nitrogen source. Of these transporters, only Mep2 is also required for the induction of filamentous growth (Biswas and Morschhauser 2005). Its function is similar to that of its *S. cerevisiae* ortholog (Boeckstaens et al. 2007). Both the Cph1 MAPK pathway and the cAMP-PKA pathway are required for filamentous growth under nitrogen starvation, and Mep2 stimulates both, possibly via Ras1. However, the mechanisms and the regulation of the signalling activity of Mep2 are not fully understood.

Expression profiling of *C. albicans* cells growing in biofilms has revealed another unexpected link between amino acid metabolism and virulence (Garcia-Sanchez et al. 2004). In this study, cells growing in biofilms were compared with those in planktonic cultures under a variety of different growth conditions. Amino acid biosynthetic genes were amongst those that were consistently upregulated in biofilm cells, compared to planktonic cells. These included *ARO*, *CYS*, *HIS*, *ILV*, *MET*, *SER* and *TRP* genes, suggesting that many amino acid biosynthetic pathways are induced during biofilm formation. Furthermore, *C. albicans gcn4* mutants displayed defects in biofilm formation (Garcia-Sanchez et al. 2004), suggesting that the GCN response is required for efficient biofilm formation. The activation of sulphur–amino acid biosynthesis, in particular, is a feature of *C. albicans* biofilms (Garcia-Sanchez et al. 2004).

As described in Sect. I, *C. albicans* biofilms that form on the catheters of hospital patients seem to be a relatively frequent source of bloodstream infections in these patients. The regulation of amino acid metabolism still seems to be important when *C. albicans* cells enter the bloodstream. This view was supported by a global study of fungal gene expression during phagocytosis by human neutrophils (Rubin-Bejerano et al. 2003). Neutrophils represent an important weapon in the host's defences against disseminated candidiasis. Following phagocytosis by neutrophils, *C. albicans* cells remain in the yeast form and are killed. However, on phagocytosis by cultured macrophages, the fungus can form hyphae inside the macrophage and escape from these cells (Lo et al. 1997; Rubin-Bejerano et al. 2003). The transcript-profiling study of Rubin-Bejerano et al. (2003) showed that methionine and arginine biosynthetic genes are induced when *C. albicans* is exposed to human neutrophils. *Candida albicans* genes on other amino acid biosynthetic pathways were generally not affected (Rubin-Bejerano et al. 2003), suggesting that the neutrophil response is distinct from the GCN response described previously. This might explain why blocking the GCN response does not inhibit disseminated candidiasis (Brand et al. 2004). Instead, the neutrophil response seems more reminiscent of stress responses, which also induce the expression of genes on the methionine and arginine biosynthetic pathways (supplementary data in Enjalbert et al. 2003). This view was reinforced by a more recent analysis of fungal phagocytosis (Jiménez-López et al. 2013). These authors showed that arginine biosynthesis is induced specifically in cells that are phagocytosed by macrophages, and this induction is driven by exposure to reactive oxygen species (ROS). This induction was not observed in macrophages deficient in the gp91 (phox) subunit of the phagocyte oxidase, thereby underlining the dependence of this response on ROS exposure. This response may be critical for hypha formation because *C. albicans* arginine pathway mutants display defects in germ tube and hypha formation (Jiménez-López et al. 2013). Taken together, these studies highlight mechanisms by which metabolic regulation might contribute to the fitness and virulence of *C. albicans* cells during their interactions with cells of the immune system.

Further links between amino acid metabolism and virulence have been highlighted by a recent study by Vylkova et al. (2011). The authors demonstrated that *C. albicans* utilises

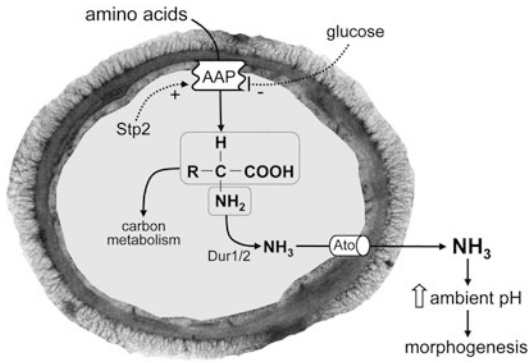


Fig. 14.1. Through amino acid metabolism, *Candida albicans* alkalinises the ambient pH of its microenvironment, thereby promoting hyphal development. As described by Vylkova et al. (2011), in the absence of glucose, *C. albicans* can assimilate amino acids by importing them via amino acid permeases (AAPs) that are induced by the transcription factor Stp2. After import, the amino acids are deaminated by Dur1/2. The carbon skeletons are utilised via carbon metabolism, and the excess nitrogen is excreted in the form of ammonia (NH_3) via the Ato transporters. The accumulation of ammonia outside the cell leads to an increase in the ambient pH of the microenvironment, and this promotes hyphal development (Vylkova et al. 2011), which is a well-known virulence factor

amino acids as a carbon source when glucose is limiting, excreting the excess nitrogen in the form of ammonia, thereby causing alkalinisation of the local environment (Fig. 14.1). This alkalinisation of the microenvironment is sufficient to promote the morphogenetic switch from yeast to hyphal growth. This phenomenon is dependent on Stp2 (a transcription factor that regulates amino acid permeases) and Ato5 (a putative ammonia transporter) and involves Dur1,2 (urea amidolyase).

Consistent with the importance of this mechanism in pathogenesis, a related function encoded by *DUR31* is required for morphogenesis, epithelial damage, survival against neutrophil attack and virulence in vivo, having been termed a multistage pathogenicity factor (Mayer et al. 2012).

Hence, *C. albicans* can modulate its local microenvironment through its metabolic activity, thereby influencing the expression of specific virulence traits (Vylkova et al. 2011). This study provides an elegant example of the dynamism of fungus-host interactions during disease progression, and the integral part that fungal metabolism plays in these interactions.

IV. Integration of Carbon Metabolism with Virulence

Carbon assimilation is essential for the generation of new biomass (i.e. growth). Therefore, the rapid growth of *C. albicans* in the immunocompromised host depends on the efficient assimilation of available carbon sources in vivo. Furthermore, as described in Sect. III.B, the fungus must adjust its metabolic programme as it encounters new microenvironments in the host. Various studies have indicated that the regulation of carbon metabolism, like amino acid metabolism, is intimately linked to the control of virulence in *C. albicans*. This section focuses mainly on pathways of **central carbon metabolism**: glycolysis, gluconeogenesis and the TCA and glyoxylate cycles.

The levels of glycolytic mRNAs change during yeast-hypha morphogenesis in *C. albicans* (Swoboda et al. 1994). The observed changes in *PYK1*, *ADH1*, *PGK1* and *GPM1* mRNA levels reflected the underlying physiological changes that accompany yeast-hypha morphogenesis rather than bone fide morphogenetic regulation of these genes (Swoboda et al. 1994). While a strict definition of morphogenetic regulation was helpful at the time, in terms of identifying hypha-specific genes (Bailey et al. 1996), this shifted attention away from potentially interesting links between *C. albicans* physiology and development. For example, subsequent reports showed that yeast-hypha morphogenesis is influenced by glucose and amino acids (Hudson et al. 2004; Maidan et al. 2005; Vylkova et al. 2011). Attention only returned to these links following the publication of several transcript-profiling studies that reinforced and extended these observations.

Nantel et al. (2002) published a transcript-profiling study of *C. albicans* morphogenesis in which the behaviour of wild-type and mutant cells was compared during serum-induced morphogenesis. The article focussed to a large

extent on the behaviour of hypha-specific and signalling genes, and as expected, hypha-specific genes such as *ECE1*, *HWP1*, *RBT1* and *SAP4,5,6* were induced during hyphal development. However, numerous genes encoding glycolytic enzymes were also regulated during hyphal development, including *ENO1*, *FBA1*, *PYK1* (*CDC19*), *TPI1* and *PGI1* (Nantel et al. 2002). Clearly, the regulation of these genes might be indirect, as implied in the early study by Swoboda et al. (1994), and these genes do seem to respond to physiological stimuli, such as ambient growth temperature (Nantel et al. 2002). However, Nantel and coworkers also showed that these glycolytic genes respond to the central morphogenetic regulator Efg1. This observation was confirmed by an independent transcript profiling study (Doedt et al. 2004). Efg1 not only regulates morphogenesis, but also induces glycolytic genes while repressing genes for oxidative metabolism (Doedt et al. 2004). Hence, the link between morphogenetic and glycolytic regulation might be more direct than was originally anticipated.

Ernst and coworkers first identified Efg1 as an important regulator of yeast-hypha morphogenesis in *C. albicans* (Stoldt et al. 1997). The inactivation or depletion of Efg1 attenuates hyphal development (Lo et al. 1997; Stoldt et al. 1997). In addition, Efg1 controls chlamydospore formation and regulates phenotypic switching (Stoldt et al. 1997; Sonneborn et al. 1999a, b; Srikantha et al. 2000). Efg1 controls the expression of adhesin genes such as *ALS3* and *HWP1* (Sharkey et al. 1999; Lane et al. 2001; Leng et al. 2001; Nantel et al. 2002). Hence, Efg1 is a key developmental regulator that regulates the expression of virulence attributes in *C. albicans*, including yeast-hypha morphogenesis, phenotypic switching and adhesins. The inactivation of Efg1 (albeit in combination with a *cph1* mutation) attenuates the virulence of *C. albicans* in the mouse model of systemic infection, reduces invasion in epithelial infection models, and inhibits biofilm formation (Lo et al. 1997; Gow et al. 2003; Korting et al. 2003; Garcia-Sanchez et al. 2004).

Efg1 acts in concert with the structurally and functionally related β-helix loop helix domain (bHLH) factor Efh1 (Doedt et al. 2004). Whereas Efh1 is a transcriptional activator, Efg1 seems to act mechanistically as a transcriptional repressor in *C. albicans* (Tebarth et al. 2003; Doedt et al. 2004). A detailed analysis of the Efg1 and Efh1 regulons in *C. albicans* (Doedt et al. 2004) showed that central metabolic pathways are regulated mainly by Efg1 and, to a lesser extent, by Efh1. The glycolytic genes *GLK1*, *HXK2*, *PGI1*, *PFK1,2*, *FBA1*, *TPI1*, *GAP1*, *PGK1*, *GPM1* and *ENO1* were all upregulated by Efg1. In contrast, gluconeogenic (*FBP1*), TCA cycle (*CIT1*, *MDH1*, *FUM12*, *SDH12*, *KGD1*) and respiratory functions (*ATP1*, *ATP17*, *PET9*) were downregulated. Therefore, Efg1 stimulates fermentative metabolism and represses respiratory metabolism. It is not clear whether Efg1 regulates these metabolic genes by acting directly at their promoters or indirectly by controlling the activities of other regulatory factors. Nevertheless, this metabolic regulation by Efg1 is physiologically significant because *efg1* cells are more sensitive to the respiratory inhibitor antimycin A (Doedt et al. 2004). Hence, Efg1 regulates metabolism as well as virulence attributes.

The link between central carbon metabolism and virulence attributes in *C. albicans* is further strengthened by another transcript profiling study, the focus of which was phenotypic switching (Lan et al. 2002). *Candida albicans* strain WO1 switches at high frequency between white and opaque forms. White cells are slightly ellipsoidal, and their cell walls have a relatively smooth surface when examined by scanning electron microscopy. In contrast, opaque cells are larger than white cells, are more elongated, contain a large vacuole and display pimples on their cell surface (Calderone and Clancy 2011). Significantly, clinical isolates of *C. albicans* from diseased patients undergo phenotypic switching at higher rates than do commensal strains (Hellstein et al. 1993). The white-opaque switch is thought to play an important role in host-pathogen interactions (Lohse and Johnson 2009). While white cells are better adapted for systemic infection, opaque cells are more stable at low temperatures, favouring skin colonisation (Lachke et al. 2003). White and opaque cells also differ in their interactions with the host immune system, as opaque cells are less susceptible to phagocytosis by macrophages (Lohse and Johnson 2008). Thus, phenotypic switching

favours *C. albicans* adaptation to specific host sites while avoiding detection by the innate immune system; therefore, it is closely associated with *C. albicans* virulence.

Lan et al. (2002) compared the global expression patterns of white and opaque cells. Well-defined sets of genes were found to be up- or downregulated in opaque cells, compared with white cells during growth on rich or defined media. Adhesion, stress response, drug resistance and signalling genes were regulated with cell type. Interestingly, about one third of the regulated genes encoded metabolic functions. Glycolytic genes were downregulated in opaque cells (*HXT3*, *HXT4*, *HXK1*, *PFK2* and *PYK1* (*CDC19*)), whereas TCA genes were upregulated (*IDP2*, *MDH1*, *MLS1*). Also, genes involved in fatty acid β-oxidation were upregulated in opaque cells (*FAA2*, *POX1*, *ECI1*, *FOX2*, *FOX3*). Therefore, white and opaque cells seem to express different metabolic programmes (Lan et al. 2002), which are presumably associated with their preference for different anatomical niches in the host. White cells favour fermentative metabolism, whereas opaque cells favour respiratory metabolism.

The findings of Lan et al. (2002) relate closely to those of Doedt et al. (2004). The inactivation of Efg1 leads to the generation of opaque cells, and *EFG1* is expressed at low levels in opaque cells. As described previously, fermentative metabolism is downregulated and oxidative metabolism is upregulated in *efg1* and opaque cells. The opposite is true in *EFG1* overexpressing cells and in white cells. Hence, Efg1 seems to link mechanistically central carbon metabolism with phenotypic switching, as well as with yeast-hypha morphogenesis. Efg1 activity is regulated by the Ras-cAMP signalling pathway (Bockmuhl and Ernst 2001). In *S. cerevisiae*, the Ras-cAMP pathway is known to regulate growth and carbon metabolism (Thevelein and de Winde 1999; Rolland et al. 2001).

A recent report highlighted the significant influence of Efg1 during gastrointestinal (GI) colonisation (Pierce et al. 2013). The authors compared *C. albicans* gene expression during colonization of the different organs in the GI tract, showing that *efg1* cells, rather than wild-type *C. albicans* cells, induced metabolic pathways that promoted the hypercolonisation of mice. Genes involved in lipid catabolism, carnitine biosynthesis and carnitine utilisation were upregulated in *efg1* cells colonising the GI tract, thereby promoting fatty acid metabolism and potentially increasing fitness. Efg1 was also shown to regulate expression of *SOD5*, a superoxide dismutase with protective roles against oxidative stress (Martchenko et al. 2004; Pierce et al. 2013).

Global analyses of another regulator, Tup1, further reinforced the view that the regulation of carbon metabolism and virulence attributes is mechanistically linked in *C. albicans*. Tup1 represses hyphal development in *C. albicans* and influences white-opaque phenotypic switching (Braun and Johnson 1997; Zhao et al. 2002), and *tup1* mutants display attenuated virulence (Murad et al. 2001a). Like its *S. cerevisiae* homolog (Smith and Johnson 2000), *C. albicans* Tup1 operates as a global transcriptional repressor that is targeted to specific promoters via interactions with specific DNA-binding proteins, one of which is Nrg1. Nrg1 interacts with Nrg1 response elements (NREs) in the promoters of hyphal genes to repress their transcription in a Tup1-dependent fashion (Braun et al. 2001; Murad et al. 2001b; Russell and Brown 2005). Transcript profiling confirmed that Tup1 and Nrg1 combine to represses hypha-specific genes such as *ECE1* and the adhesin genes *ALS3* and *HWP1* (Murad et al. 2001a; Kadosh and Johnson 2005; García-Snchez et al. 2005). These experiments also revealed that Tup1 represses numerous additional functions, including pathways of central and peripheral carbon metabolism. Gluconeogenic genes (*PCK1*, *FBP1*), TCA and glyoxylate cycle genes (*IDP2*, *MDH11*, *ICL1*) and genes required for the catabolism of non-fermentative carbon sources were derepressed in *tup1* cells (*GAL1*,*10*, *ADH3*,*4*,*5*, *POX4*). Tup1-mediated regulation of at least some of these metabolic functions might operate via the repressors Mig1 and possibly Mig2 (Murad et al. 2001a; Sabina and Brown 2009). Therefore, like Efg1, Tup1 provides another mechanistic link between carbon metabolism and morphogenesis and phenotypic switching.

The importance of carbon metabolism for fungal pathogenicity has been further emphasised by expression profiling of fungal cells in infection models. The first such observation arose through microarray analyses of *S. cerevisiae* cells following phagocytosis by macrophages (Lorenz and Fink 2001) (Fig. 14.2). The glyoxylate cycle genes *CIT2*, *ICL1*, *MDH2* and *MLS1* were amongst the most strongly induced genes, but the TCA cycle genes *FUM1*, *IDH1*, *KGD1*, and *SDH1*,*2*,*3*,*4* were also induced. Lorenz and Fink (2001, 2002) noted that glyoxylate cycle genes are required for the virulence of other microbial pathogens; hence, they tested whether this was the case for *C. albicans*. They confirmed that *ICL1* inactivation attenuates the virulence of *C. albicans* (Lorenz and Fink 2001), highlighting the importance of central metabolic pathways for the survival of this pathogenic fungus in vivo. This work was confirmed and extended by Barelle and coworkers, who showed that glycolytic (Pyk1) and gluconeogenic functions (Pck1) are also required for full virulence (Barelle et al. 2006).

These observations were then reinforced by transcriptional profiling of *C. albicans* cells exposed to macrophages (Lorenz et al. 2004). This confirmed that in *C. albicans*, glycoxylate cycle genes are induced following phagocytosis, and it revealed that fatty acid β-oxidation genes are also induced (*FOX2*, *POX1*, *POX2*) (Fig. 14.2). The inactivation of fatty acid β-oxidation does not block the development of systemic candidiasis, but this pathway does contribute to the full virulence of *C. albicans* (Piekarska et al. 2006; Ramirez and Lorenz 2007). On internalisation, *C. albicans* cells display a starvation response, triggering gluconeogenesis and fatty acid degradation (Lorenz et al. 2004). At later stages, filamentation promotes escape from macrophages, and *C. albicans* cells quickly resume glycolysis. The authors suggested that, following phagocytosis by macrophages, *C. albicans* reprograms its metabolism to generate glucose via lipid catabolism, the glyoxylate cycle and gluconeogenesis (Lorenz et al. 2004). This transcriptional reprogramming is not observed in the benign yeast *S. cerevisiae* (Lorenz and Fink 2001), suggesting that *C. albicans* has undergone evolutionary adaptation to survive phagocytic attack by macrophages and thrive as a pathogen.

Additional microarray studies have been performed on *C. albicans* cells exposed to neutrophils (Rubin-Bejerano et al. 2003), human blood and blood fractions (Fradin et al. 2003, 2005) (Fig. 14.2). These studies provided important clues about the regulation of numerous virulence attributes during the early stages of disseminated candidiasis and revealed that, of the various blood fractions studied, granulocytes impose the greatest effects on the fungal transcriptome. Intriguingly, both glycolytic (*PGI1*, *PFK2*, *FBA1*, *GAP1*, *PGK1*, *ENO1*, *PYK1*) and glyoxylate cycle genes (*ICL1*, *MLS1*, *MDH1*, *ACS1*) were induced in *C. albicans* populations exposed to human blood (Fradin et al. 2003, 2005). Yet, in both *C. albicans* and *S. cerevisiae*, glycolytic and gluconeogenic mRNAs display opposing regulation. Indeed, in both yeasts gluconeogenic and glyoxylate cycle genes are exquisitely sensitive to glucose, being repressed at concentrations as low as 0.01 % (i.e. well below the glucose concentrations encountered in the bloodstream: 4 mM, 0.072 %) (Yin et al. 2003; Rodaki et al. 2009). Yet, as discussed, the glyoxylate cycle is required for the full development of systemic candidiasis (Lorenz and Fink 2001).

This conundrum – the simultaneous activation of competing pathways of central metabolism – was often compounded by further transcript profiling of *C. albicans* cells isolated from other infection models. These include cells from the ex vivo reconstituted human oral epithelium model of mucosal infection (Zakikhany et al. 2007), from intraperitoneal infections (Thewes et al. 2007), and from the kidneys of mice and rabbits suffering systemic candidiasis (Andes et al. 2005; Walker et al. 2009) (Fig. 14.2). For example, both glycolytic and gluconeogenic or TCA cycle genes were upregulated in the mucosal and intraperitoneal infection models (Thewes et al. 2007; Wilson et al. 2009; Zakikhany et al. 2007). The simultaneous induction of glycolytic plus gluconeogenic and glyoxylate cycle functions might

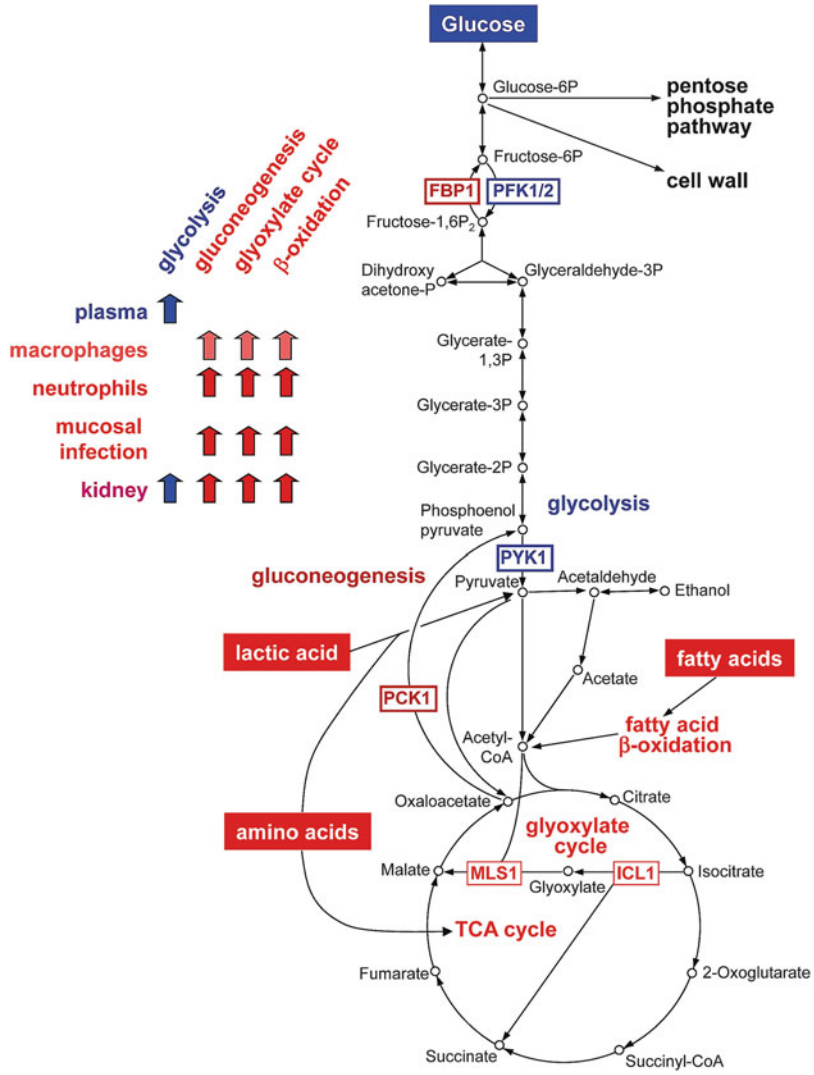


Fig. 14.2. Central carbon metabolism in *Candida albicans* cells occupying different host niches. The *right side* of the figure indicates central carbon metabolism, with the main pathways highlighted. Glycolysis and glucose are highlighted in *blue*. Gluconeogenesis, the glyoxylate and tricarboxylic acid (TCA) cycles, and fatty acid β-oxidation are highlighted in *red*, as are the alternative carbon sources lactate, amino acids and fatty acids. A small number of key enzymes that characterise glycolysis (Pfk1/2, Pyk1), gluconeogenesis (Pck1, Fbp1) and the glyoxylate cycle (Icl1, Mls1) are also highlighted because these are targets of single-cell profiling experiments described in the text. The *left side* of the figure summarises the main central metabolic activities of *C. albicans* cells in blood plasma (Fradin et al. 2005), macrophages (Barelle et al. 2006), neutrophils (Barelle et al. 2006), mucosal infection (Zakikhany et al. 2007), and kidney infection (Barelle et al. 2006); the tissues are colour coded according to their main metabolic activities: *blue*, mainly glycolytic; *red*, mainly gluconeogenic; *purple*, heterogeneous population of glycolytic and gluconeogenic cells

seem metabolically counterintuitive. However, the probable explanation is that these experiments examined heterogeneous populations of fungal cells that were derived from different microenvironments (Brown et al. 2007; Wilson et al. 2009). For example, in the ex vivo blood models, glyoxylate cycle genes were presumably induced in phagocytosed *C. albicans* cells,

whereas glycolytic genes were probably induced in those fungal cells remaining in the plasma and with access to glucose.

Single-cell profiling has been adopted to address this apparent paradox. Green fluorescent protein (GFP) fusions were constructed for diagnostic metabolic genes in *C. albicans*: two glycolysis-specific genes *PFK2* and *PYK1*; the gluconeogenic-specific gene *PCK1*; and the glyoxylate cycle genes *ICL1* and *MLS1* (Barelle et al. 2006; Miramon et al. 2012). By examining the green fluorescence emitted by individual cells, it was confirmed that before phagocytosis by macrophages or neutrophils, *C. albicans* cells express glycolytic functions, whereas they switch to the glyoxylate cycle and gluconeogenesis following phagocytosis. Also, as predicted, a heterogeneous population of *C. albicans* cells was observed in the complex microenvironment of the mammalian kidney, some cells expressing gluconeogenic and glyoxylate cycle genes and others showing no detectable expression (Barelle et al. 2006) (Fig. 14.2). As mentioned, all three pathways contribute to the virulence of *C. albicans* in a murine model of systemic infection (Lorenz and Fink 2001; Barelle et al. 2006). These observations highlight the importance of central carbon metabolism for the establishment of *C. albicans* infections, the metabolic flexibility of *C. albicans*, and the complexity of the niches that this pathogen occupies.

These single-cell profiling studies raised a further paradox. About one third to a half of *C. albicans* cells infecting the kidney expressed gluconeogenic and glyoxylate cycle genes, and over 80 % of cells expressed glycolytic genes, suggesting that some *C. albicans* cells might be expressing both catabolic and anabolic pathways (Barelle et al. 2006). There were feasible technical explanations for this apparent paradox, based on the stability of the GFP reporter. However, recent data have indeed confirmed that *C. albicans* can express glycolytic, gluconeogenic and glyoxylate cycle enzymes at the same time, allowing them to assimilate glucose and alternative carbon sources simultaneously (Sandai et al. 2012) (Fig. 14.3). In *S. cerevisiae*, glucose triggers the rapid degradation of Icl1 and Pck1, but in *C. albicans*, these glyoxylate cycle and gluconeogenic enzymes remain stable and are retained by the cell (Sandai et al. 2012). The differential stability of the *S. cerevisiae* and *C. albicans* Icl1 and ScPck1 enzymes relates to the **evolutionary rewiring of ubiquitination targets** in these yeasts. For example, the lack of ubiquitination sites in *C. albicans* Icl1 prevents this enzyme from being targeted for glucose-accelerated degradation (Sandai et al. 2012). This allows *C. albicans* cells growing on alternative carbon sources to continue to assimilate these nutrients even if glucose becomes available (Sandai et al. 2012) (Fig. 14.3). This is particularly relevant in vivo because host niches are complex; in some cases, sugar availability might be limited or transient.

In some niches, **alternative carbon sources** are essential for fungal growth (Piekarska et al. 2006; Ueno et al. 2011). For example, lactate assimilation is essential for the proliferation of *Candida glabrata* in the mouse intestine (Ueno et al. 2011). Growth on lactate rather than glucose modulates the architecture of the cell wall, the expression of cell wall and secreted proteins, and hence host recognition (Ene et al. 2012a, b, 2013). This cell wall remodelling also has an impact on adherence, biofilm formation, sensitivity to stress, and resistance to antifungal drugs (Ene et al. 2012a). For *C. albicans* cells growing on lactate, exposure to sugars leads to increased resistance to ROS and to an azole antifungal agent (Rodaki et al. 2009). Hence, changes in carbon source can have dramatic effects on the virulence of *C. albicans* (Ene et al. 2012a) and the susceptibility of this pathogen to therapeutic intervention. Taken together, the available data suggest that *C. albicans* has evolved such that the regulation of central carbon metabolism and virulence are intimately linked. This presents significant clinical and experimental challenges given the metabolic heterogeneity of *C. albicans* cell populations during colonisation and disease progression (described previously). This highlights the importance of developing technologies that allow one to analyse the molecular behaviour of individual fungal cells within the

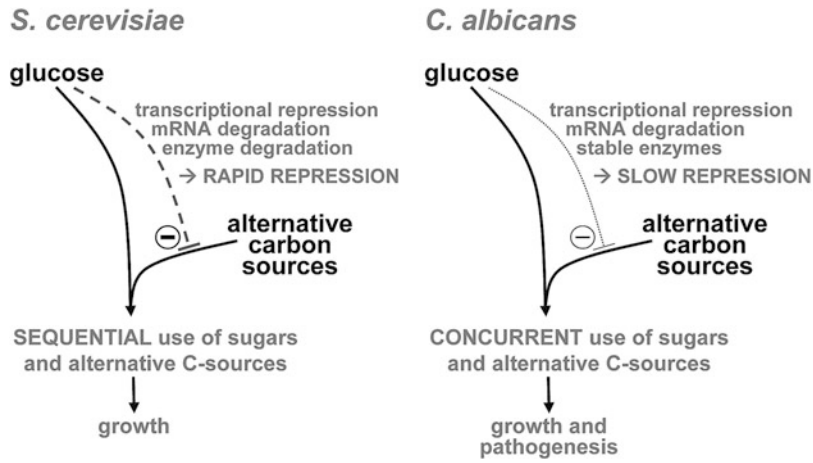


Fig. 14.3. Carbon catabolite repression is relaxed in *Candida albicans*. In *Saccharomyces cerevisiae*, the addition of glucose to cells growing on alternative carbon sources such as lactate or fatty acids triggers rapid transcriptional repression of genes involved in the assimilation of these alternative carbon sources, accelerated degradation of the corresponding mRNAs, and rapid degradation of the corresponding enzymes. As a result, the assimilation of alternative carbon sources is rapidly downregulated. *Saccharomyces cerevisiae* then utilises the glucose, exploiting the alternative carbon sources only once the glucose is exhausted. As described by Sandai et al. (2012), there has been evolutionary rewiring of ubiquitination targets in *C. albicans* such that key enzymes involved in the assimilation of alternative carbon sources are not subject to ubiquitin-dependent, glucose-accelerated degradation. These enzymes are stable, and the pathways remain active, allowing *C. albicans* to continue assimilating alternative carbon sources simultaneously with glucose. However, the glucose triggers transcriptional repression and rapid degradation of the corresponding genes and mRNAs, leading to the downregulation of these pathways in the long term (Sandai et al. 2012)

complex populations that occupy host microenvironments (Barelle et al. 2004, 2006; Miramon et al. 2012).

V. Coevolution of Virulence and Fitness

Candida albicans is an opportunistic pathogen (Odds 1988; Calderone 2002; Calderone and Clancy 2011). The fungus causes systemic infections relatively rarely, and the development of such infections is probably more dependent on the immune status of the host than on the virulence status of the fungus. Hence, it could be argued that the evolutionary pressures exerted on *C. albicans* have been directed towards its survival as a commensal or mucocutaneous pathogen rather than as a systemic pathogen.

Many *C. albicans* genes, including those encoding metabolic enzymes, are required for virulence (reviewed by Navarro-Garcia et al. 2001). This sparked debate in the field about the definition of a **virulence gene**. *Virulence genes* are now defined as those that encode factors that interact directly with host components, such as adhesins or secreted aspartyl proteinases (Odds et al. 2003; Klis et al. 2009; Calderone and Clancy 2011). Factors such as Efg1 and Tup1 that control the expression of virulence genes are termed *virulence regulators*. Functions that are required for microbial growth but do not interact directly with the host should not be termed virulence factors even though they might be required for virulence (Navarro-Garcia et al. 2001; Roemer et al. 2003). These functions should be called **fitness attributes**.

There is now considerable evidence for the evolution of fitness attributes in *C. albicans*. For example, this pathogen has evolved specialised stress responses that differ significantly from stress responses in *S. cerevisiae* and *Schizosaccharomyces pombe* (Enjalbert et al. 2003;

Nicholls et al. 2004; Smith et al. 2004; Rodaki et al. 2009; Brown et al. 2011). These specialised stress responses seem to reflect the rapid and relatively recent evolutionary adaptation of *C. albicans* to its host (Nikolaou et al. 2009). Unlike budding and fission yeast, *C. albicans* does not seem to recognise an ambient temperature of 37 °C as a stress, and this fungus is relatively resistant to the oxidative stresses it encounters following phagocytosis (Smith et al. 2004; Nikolaou et al. 2009). The inactivation of Hog1, which is required to activate oxidative stress responses, attenuates the virulence of *C. albicans* (Alonso-Monge et al. 1999). Glucose enhances oxidative stress resistance in *C. albicans*, whereas it attenuates stress resistance on *S. cerevisiae* (Rodaki et al. 2009). Metabolic genes and pathways seem to have evolved to allow this fungus to assimilate those nutrients that are available to it within the various microenvironments it encounters during host colonisation and disease progression (Lorenz and Fink 2002; Barelle et al. 2006; Ramirez and Lorenz 2007; Vylkova et al. 2011; Sandai et al. 2012). Hence, fitness attributes seem to have evolved alongside virulence attributes in *C. albicans*.

Not only have fitness and virulence attributes themselves been evolving, but also the mechanisms that regulate their activity seem to have been evolving in *C. albicans*. As described, there is considerable evidence for niche-specific changes in virulence factors, metabolic functions and stress genes in *C. albicans* (e.g. Fradin et al. 2003, 2005; Rubin-Bejerano et al. 2003; Lorenz et al. 2004; Andes et al. 2005; Barelle et al. 2006; Enjalbert et al. 2007; Thewes et al. 2007; Wilson et al. 2009; Zakikhany et al. 2007; Rodaki et al. 2009; Walker et al. 2009; Miramon et al. 2012). Therefore, virulence and fitness attributes seem to be regulated in such a manner that it ensures they are expressed appropriately in particular microenvironments. This review has highlighted just some examples that indicate that metabolic reprogramming is mechanistically linked to the control of virulence via specific regulators (e.g. Efg1, Tup1 and Gcn4) (Sects. III and IV; Lan et al. 2002; Tripathi et al. 2002; Doedt et al. 2004; García-Snchez et al. 2005; Kadosh and Johnson 2005). Hence, *C. albicans* seems to have evolved specific signalling mechanisms that facilitate the coordinated regulation of the virulence and fitness attributes that it requires for survival in the host.

VI. Conclusions and Future Perspectives

In this chapter, we have argued that metabolism is intimately linked with virulence in *C. albicans*. We have also highlighted the critical role of *C. albicans* genomics in revealing these links. Unexpected links between metabolism and virulence have been revealed by global exploration via transcript profiling, and further links remain to be discovered. This task will be greatly facilitated by the availability of large mutant collections such as the GRACE collection (Roemer et al. 2003); the transposon insertion mutants generated by the Mitchell and Johnson groups (Nobile et al. 2003; Uhl et al. 2003); the library of null mutants created by Noble, Johnson, and coworkers (Noble et al. 2010); and the development of the *C. albicans* ORFeome (Chauvel et al. 2012). Antisense-based screens provide an alternative approach towards the identification of such pleiotropic factors (De Backer et al. 2001; Sellam et al. 2010). However, the next major goal will be the molecular dissection of the specific mechanisms that coordinate metabolic adaptation with the regulation of virulence factors. The work of Vylkova and coworkers, linking amino acid assimilation and pH modulation to morphogenesis (Vylkova et al. 2011), provides an elegant illustration of what is required.

It is becoming increasingly apparent that metabolic regulation is also intimately linked to stress adaptation, and this also contributes to the pathogenicity of *C. albicans* (Ene et al. 2012a). Cell wall remodelling has been shown to contribute significantly to these effects (Ene et al. 2012a, b). However, additional factors are almost certainly involved in the coordination of virulence and fitness attributes, and these remain to be identified.

How does *C. albicans* tune its fitness and virulence attributes within specific host micro-

environments? Addressing this important question is critical to our understanding of *C. albicans* pathogenicity. Obviously, transcript profiling of in vivo samples will help to address this question. However, this technology has been applied mainly to in vitro samples. There is only limited coverage of ex vivo infection models, and data on in vivo infections is sparse (Fradin et al. 2003, 2005; Rubin-Bejerano et al. 2003; Lorenz et al. 2004; Andes et al. 2005; Thewes et al. 2007; Wilson et al. 2009; Zakikhany et al. 2007; Walker et al. 2009; Nailis et al. 2010). Furthermore, increased sensitivity and spatial resolution is required to address the complexity and heterogeneity of fungal populations within host niches (Sect. IV). Advances in sensitivity will soon allow RNA-sequencing technologies to be applied to minute, spatially distinct tissue samples generated by laser capture dissection microscopy, for example. Also, this type of genomic exploration will have to be followed by detailed molecular dissection of the regulatory mechanisms that drive the observed expression patterns in host niches. This will involve a combination of in vitro dissection and in vivo validation (e.g. by single-cell profiling: Barelle et al. 2006; Enjalbert et al. 2007; Miramon et al. 2012). Only in this way will we be able to determine how *C. albicans* cells integrate metabolic regulation with other virulence-related functions during colonisation and disease progression.

What happens to metabolic reprogramming when fungal cells respond and adapt to antifungal drugs? Genomic studies of *C. albicans* cells as they acquire drug resistance have suggested that metabolism is affected (Cowen et al. 2002; Bruneau et al. 2003; Rogers and Barker 2003; Shapiro et al. 2011). These metabolism changes are not restricted to the antifungal targets themselves. For example, they are not limited to ergosterol biosynthesis following exposure to an azole antifungal. Is this metabolic reprogramming required for drug adaptation, or is it an indirect by-product of exposure to the antifungal agent?

What relevance has the issue of *C. albicans* fitness and virulence to the patient suffering from a *Candida* infection? To date, the academic community has focussed primarily on *C. albicans* virulence attributes and has paid relatively little attention to metabolism. Ironically, most virulence attributes may be poor targets for antifungal therapy because most are polygenic and not essential for survival in vivo. In contrast, most successful antifungal drug screens have focussed on the fitness of *C. albicans*, and many central metabolic functions are essential for fitness in vivo (Lorenz and Fink 2001; Roemer et al. 2003; Rodaki et al. 2006; Brock 2009). Clearly, a complete understanding of *C. albicans* pathogenicity is dependent on a characterisation of fitness as well as virulence attributes.

Acknowledgements We are grateful to many colleagues for stimulating debates about *Candida* genomics, especially our friends in the Aberdeen Fungal Group and the European FINSysB Consortium and Ken Haynes and Jan Quinn. IVE and AJPB were supported by a grant from the European Commission (PITN-GA-2008-214004). AJPB was also supported by the European Research Council (ERC-2009-AdG-249793), the U.K. Biotechnology and Biological Sciences Research Council (BBS/B/06679; BB/C510391/1; BB/D009308/1; BB/F000111/1; BB/F010826/1; BB/F00513X/1), and the Wellcome Trust (080088; 097377).

References

Albrecht G, Mosch HU, Hoffmann B, Reusser U, Braus GH (1998) Monitoring the Gcn4 protein-mediated response in the yeast *Saccharomyces cerevisiae*. J Biol Chem 273:12696–12702

Almeida RS, Wilson D, Hube B (2009) *Candida albicans* iron acquisition within the host. FEMS Yeast Res 9:1000–1012

Alonso-Monge R, Navarro-García F, Molero G, Diez-Orejas R, Gustin M, Pla J, Snchez M, Nombela C (1999) Role of the mitogen-activated protein kinase Hog1p in morphogenesis and virulence of *Candida albicans*. J Bacteriol 181:3058–3068

Andes D, Lepak A, Pitula A, Marchillo K, Clark J (2005) A simple approach for estimating gene expression in *Candida albicans* directly from a systemic infection site. J Infect Dis 192:893–900

Askew C, Sellam A, Epp E, Hogues H, Mullick A, Nantel A, Whiteway M (2009) Transcriptional regulation of carbohydrate metabolism in the human pathogen *Candida albicans*. PLoS Pathog 5:e1000612

Bailey DA, Feldmann PJ, Bovey M, Gow NA, Brown AJ (1996) The *Candida albicans HYR1* gene, which is activated in response to hyphal development,

belongs to a gene family encoding yeast cell wall proteins. J Bacteriol 178:5353–5360

Barelle CJ, Manson CL, MacCallum DM, Odds FC, Gow NA, Brown AJ (2004) GFP as a quantitative reporter of gene regulation in *Candida albicans*. Yeast 21:333–340

Barelle CJ, Priest CL, Maccallum DM, Gow NA, Odds FC, Brown AJ (2006) Niche-specific regulation of central metabolic pathways in a fungal pathogen. Cell Microbiol 8:961–971

Bertram G, Swoboda RK, Gooday GW, Gow NA, Brown AJ (1996) Structure and regulation of the *Candida albicans ADH1* gene encoding an immunogenic alcohol dehydrogenase. Yeast 12:115–127

Biswas K, Morschhauser J (2005) The Mep2p ammonium permease controls nitrogen starvation-induced filamentous growth in *Candida albicans*. Mol Microbiol 56:649–669

Blankenship JR, Mitchell AP (2006) How to build a biofilm: a fungal perspective. Curr Opin Microbiol 9:588–594

Bockmuhl DP, Ernst JF (2001) A potential phosphorylation site for an A-type kinase in the Efg1 regulator protein contributes to hyphal morphogenesis of *Candida albicans*. Genetics 157:1523–1530

Boeckstaens M, Andre B, Marini AM (2007) The yeast ammonium transport protein Mep2 and its positive regulator, the Npr1 kinase, play an important role in normal and pseudohyphal growth on various nitrogen media through retrieval of excreted ammonium. Mol Microbiol 64:534–546

Brand A (2012) Hyphal growth in human fungal pathogens and its role in virulence. Int J Microbiol 2012:517529

Brand A, MacCallum DM, Brown AJ, Gow NA, Odds FC (2004) Ectopic expression of *URA3* can influence the virulence phenotypes and proteome of *Candida albicans* but can be overcome by targeted reintegration of *URA3* at the *RPS10* locus. Eukaryot Cell 3:900–909

Braun BR, Johnson AD (1997) Control of filament formation in *Candida albicans* by the transcriptional repressor *TUP1*. Science 277:105–109

Braun BR, Kadosh D, Johnson AD (2001) *NRG1*, a repressor of filamentous growth in *C. albicans*, is down-regulated during filament induction. EMBO J 20:4753–4761

Brega E, Zufferey R, Mamoun CB (2004) *Candida albicans* Csy1p is a nutrient sensor important for activation of amino acid uptake and hyphal morphogenesis. Eukaryot Cell 3:135–143

Brock M (2009) Fungal metabolism in host niches. Curr Opin Microbiol 12:371–376

Brown AJ (2005) Integration of metabolism with virulence in *Candida albicans*. In: Brown AJ (ed) The Mycota, vol 13, Fungal genomics. Springer, Berlin, pp 185–203

Brown AJ, Gow NA (1999) Regulatory networks controlling *Candida albicans* morphogenesis. Trends Microbiol 7:333–338

Brown AJ, Barelle CJ, Budge S, Duncan J, Harris S, Lee PR, Leng P, Macaskill S, Abdul Murad AM, Ramsdale M, Wiltshire C, Wishart JA, Gow NA (2000) Gene regulation during morphogenesis in *Candida albicans*. Contrib Microbiol 5:112–125

Brown AJ, Odds FC, Gow NA (2007) Infection-related gene expression in *Candida albicans*. Curr Opin Microbiol 10:307–313

Brown V, Sabina J, Johnston M (2009) Specialized sugar sensing in diverse fungi. Curr Biol 19:436–441

Brown AJ, Haynes K, Gow NAR, Quinn J (2011) Stress responses in *Candida*. In: Clancy CJ, Calderone RA (eds) *Candida* and candidiasis, 2nd edn. ASM Press, Washington, DC, pp 225–242

Bruneau JM, Maillet I, Tagat E, Legrand R, Supatto F, Fudali C, Caer JP, Labas V, Lecaque D, Hodgson J (2003) Drug induced proteome changes in *Candida albicans*: comparison of the effect of beta(1,3) glucan synthase inhibitors and two triazoles, fluconazole and itraconazole. Proteomics 3:325–336

Calderone R (2002) *Candida* and candidiasis. ASM Press, Washington, DC

Calderone RA, Clancy CJ (2011) *Candida* and candidiasis. ASM Press, Washington, DC

Chauvel M, Nesseir A, Cabral V, Znaidi S, Goyard S, Bachellier-Bassi S, Firon A, Legrand M, Diogo D, Naulleau C, Rossignol T, d'Enfert C (2012) A versatile overexpression strategy in the pathogenic yeast *Candida albicans*: identification of regulators of morphogenesis and fitness. PLoS One 7:e45912

Citiulo F, Jacobsen ID, Miramón P, Schild L, Brunke S, Zipfel P, Brock M, Hube B, Wilson D (2012) *Candida albicans* scavenges host zinc via Pra1 during endothelial invasion. PLoS Pathog 8:e1002777

Cowen LE, Nantel A, Whiteway MS, Thomas DY, Tessier DC, Kohn LM, Anderson JB (2002) Population genomics of drug resistance in *Candida albicans*. Proc Natl Acad Sci U S A 99:9284–9289

De Backer MD, Nelissen B, Logghe M, Viaene J, Loonen I, Vandoninck S, de Hoogt R, Dewaele S, Simons FA, Verhasselt P, Vanhoof G, Contreras R, Luyten WH (2001) An antisense-based functional genomics approach for identification of genes critical for growth of *Candida albicans*. Nat Biotechnol 19:235–241

Dever TE, Feng L, Wek RC, Cigan AM, Donahue TF, Hinnebusch AG (1992) Phosphorylation of initiation factor 2 alpha by protein kinase *GCN2* mediates gene-specific translational control of *GCN4* in yeast. Cell 68:585–596

Doedt T, Krishnamurthy S, Bockmühl DP, Tebarth B, Stempel C, Russell CL, Brown AJ, Ernst JF (2004) APSES proteins regulate morphogenesis and metabolism in *Candida albicans*. Mol Biol Cell 15:3167–3180

Ellenberger TE, Brandl CJ, Struhl K, Harrison SC (1992) The *GCN4* basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. Cell 71:1223–1237

Ene IV, Adya AK, Wehmeier S, Brand AC, MacCallum DM, Gow NA, Brown AJ (2012a) Host carbon sources modulate cell wall architecture, drug resistance and virulence in a fungal pathogen. Cell Microbiol 14:1319–1335

Ene IV, Heilmann CJ, Sorgo AG, Walker LA, de Koster CG, Munro CA, Klis FM, Brown AJ (2012b) Carbon source-induced reprogramming of the cell wall proteome and secretome modulates the adherence and drug resistance of the fungal pathogen *Candida albicans*. Proteomics 12:3164–3179

Ene IV, Cheng SC, Netea MG, Brown AJ (2013) Growth of *Candida albicans* cells on the physiologically relevant carbon source lactate affects their recognition and phagocytosis by immune cells. Infect Immun 81:238–248

Enjalbert B, Nantel A, Whiteway M (2003) Stress-induced gene expression in *Candida albicans*: absence of a general stress response. Mol Biol Cell 14:1460–1467

Enjalbert B, MacCallum DM, Odds FC, Brown AJ (2007) Niche-specific activation of the oxidative stress response by the pathogenic fungus *Candida albicans*. Infect Immun 75:2143–2151

Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C, Hube B (2003) Stage-specific gene expression of *Candida albicans* in human blood. Mol Microbiol 47:1523–1543

Fradin C, De Groot P, MacCallum D, Schaller M, Klis F, Odds FC, Hube B (2005) Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood. Mol Microbiol 56:397–415

Gancedo JM (1998) Yeast carbon catabolite repression. Microbiol Mol Biol Rev 62:334–361

Garcia-Sanchez S, Aubert S, Iraqui I, Janbon G, Ghigo JM, d'Enfert C (2004) *Candida albicans* biofilms: a developmental state associated with specific and stable gene expression patterns. Eukaryot Cell 3:536–545

García-Snchez S, Mavor AL, Russell CL, Argimon S, Dennison P, Enjalbert B, Brown AJ (2005) Global roles of Ssn6 in Tup1- and Nrg1-dependent gene regulation in the fungal pathogen, *Candida albicans*. Mol Biol Cell 16:2913–2925

Gow NA, Knox Y, Munro CA, Thompson WD (2003) Infection of chick chorioallantoic membrane (CAM) as a model for invasive hyphal growth and pathogenesis of *Candida albicans*. Med Mycol 41:331–338

Gow NA, van de Veerdonk FL, Brown AJ, Netea MG (2011) *Candida albicans* morphogenesis and host defence: discriminating invasion from colonization. Nat Rev Microbiol 10:112–122

Harris AD, Castro J, Sheppard DC, Carmeli Y, Samore MH (1999) Risk factors for nosocomial candiduria due to *Candida glabrata* and *Candida albicans*. Clin Infect Dis 29:926–928

Hellstein J, Vawter-Hugart H, Fotos P, Schmid J, Soll DR (1993) Genetic similarity and phenotypic diversity of commensal and pathogenic strains of *Candida albicans* isolated from the oral cavity. J Clin Microbiol 31:3190–3199

Hinnebusch AG (1988) Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. Microbiol Rev 52:248–273

Hinnebusch AG, Natarajan K (2002) Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. Eukaryot Cell 1:22–32

Hoffmann B, Valerius O, Andermann M, Braus GH (2001) Transcriptional autoregulation and inhibition of mRNA translation of amino acid regulator gene cpcA of filamentous fungus *Aspergillus nidulans*. Mol Biol Cell 12:2846–2857

Hoyer LL, Green CB, Oh SH, Zhao X (2008) Discovering the secrets of the *Candida albicans* agglutinin-like sequence (*ALS*) gene family – a sticky pursuit. Med Mycol 46:1–15

Huang G (2012) Regulation of phenotypic transitions in the fungal pathogen *Candida albicans*. Virulence 3:251–261

Hudson DA, Sciascia QL, Sanders RJ, Norris GE, Edwards PJB, Sullivan PA, Farley PC (2004) Identification of the dialysable serum inducer of germ-tube formation in *Candida albicans*. Microbiology 150:3041–3049

Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, Barkai N (2005) Rewiring of the yeast transcriptional network through the evolution of motif usage. Science 309:938–940

Jiménez-López C, Collette JR, Brothers KM, Shepardson KM, Cramer RA, Wheeler RT, Lorenz MC (2013) *Candida albicans* induces arginine biosynthetic genes in response to host-derived reactive oxygen species. Eukaryot Cell 12:91–100

Johnston M (1999) Feasting, fasting and fermenting. Glucose sensing in yeast and other cells. Trends Genet 15:29–33

Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, Newport G, Thorstenson YR, Agabian N, Magee PT, Davis RW, Scherer S (2004) The diploid genome sequence of *Candida albicans*. Proc Natl Acad Sci U S A 101:7329–7334

Kadosh D, Johnson AD (2001) Rfg1, a protein related to the *Saccharomyces cerevisiae* hypoxic regulator Rox1, controls filamentous growth and virulence in *Candida albicans*. Mol Cell Biol 21:2496–2505

Kadosh D, Johnson AD (2005) Induction of the *Candida albicans* filamentous growth program by relief of transcriptional repression: a genome-wide analysis. Mol Biol Cell 16:2903–2912

Kastaniotis AJ, Zitomer RS (2000) Rox1 mediated repression. Oxygen dependent repression in yeast. Adv Exp Med Biol 475:185–195

Khalaf RA, Zitomer RS (2001) The DNA binding protein Rfg1 is a repressor of filamentation in *Candida albicans*. Genetics 157:1503–1512

Kim J, Sudbery P (2011) *Candida albicans*, a major human fungal pathogen. J Microbiol 49:171–177

Klis FM, Sosinska GJ, de Groot PW, Brul S (2009) Covalently linked cell wall proteins of *Candida albicans* and their role in fitness and virulence. FEMS Yeast Res 9:1013–1028

Koh AY, Kohler JR, Coggshall KT, Van Rooijen N, Pier GB (2008) Mucosal damage and neutropenia are required for *Candida albicans* dissemination. PLoS Pathog 4:e35

Kornitzer D, Raboy B, Kulka RG, Fink GR (1994) Regulated degradation of the transcription factor Gcn4. EMBO J 13:6021–6030

Korting HC, Hube B, Oberbauer S, Januschke E, Hamm G, Albrecht A, Borelli C, Schaller M (2003) Reduced expression of the hyphal-independent *Candida albicans* proteinase genes *SAP1* and *SAP3* in the *efg1* mutant is associated with attenuated virulence during infection of oral epithelium. J Med Microbiol 52:623–632

Krishnamoorthy T, Pavitt GD, Zhang F, Dever TE, Hinnebusch AG (2001) Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. Mol Cell Biol 21:5018–5030

Lachke SA, Lockhart SR, Daniels KJ, Soll DR (2003) Skin facilitates *Candida albicans* mating. Infect Immun 71:4970–4976

Lan CY, Newport G, Murillo LA, Jones T, Scherer S, Davis RW, Agabian N (2002) Metabolic specialization associated with phenotypic switching in *Candida albicans*. Proc Natl Acad Sci U S A 99:14907–14912

Lane S, Birse C, Zhou S, Matson R, Liu H (2001) DNA array studies demonstrate convergent regulation of virulence factors by Cph1, Cph2, and Efg1 in *Candida albicans*. J Biol Chem 276:48988–48996

Lavoie H, Hogues H, Whiteway M (2009) Rearrangements of the transcriptional regulatory networks of metabolic pathways in fungi. Curr Opin Microbiol 12:655–663

Leach MD, Stead DA, Argo E, MacCallum DM, Brown AJP (2011) Proteomic and molecular analyses highlight the importance of ubiquitination for stress resistance, metabolic adaptation, morphogenetic regulation and virulence of *Candida albicans*. Mol Microbiol 79:1574–1593

Leng P, Lee PR, Wu H, Brown AJ (2001) Efg1, a morphogenetic regulator in *Candida albicans*, is a sequence-specific DNA binding protein. J Bacteriol 183:4090–4093

Lo HJ, Kohler JR, DiDomenico B, Loebenberg D, Cacciapuoti A, Fink GR (1997) Nonfilamentous *C. albicans* mutants are avirulent. Cell 90:939–949

Lohse MB, Johnson AD (2008) Differential phagocytosis of white versus opaque *Candida albicans* by *Drosophila* and mouse phagocytes. PLoS One 3:e1473

Lohse MB, Johnson AD (2009) White-opaque switching in *Candida albicans*. Curr Opin Microbiol 12:650–654

Lorenz MC, Fink GR (2001) The glyoxylate cycle is required for fungal virulence. Nature 412:83–86

Lorenz MC, Fink GR (2002) Life and death in a macrophage: role of the glyoxylate cycle in virulence. Eukaryot Cell 1:657–662

Lorenz MC, Bender JA, Fink GR (2004) Transcriptional response of *Candida albicans* upon internalization by macrophages. Eukaryot Cell 3:1076–1087

Maidan MM, Thevelein JM, Van Dijck P (2005) Carbon source induced yeast-to-hypha transition in *Candida albicans* is dependent on the presence of amino acids and on the G-protein-coupled receptor Gpr1. Biochem Soc Trans 33:291–293

Martchenko M, Alarco AM, Harcus D, Whiteway M (2004) Superoxide dismutases in *Candida albicans*: transcriptional regulation and functional characterization of the hyphal-induced SOD5 gene. Mol Biol Cell 15:456–467

Martchenko M, Levitin A, Hogues H, Nantel A, Whiteway M (2007) Transcriptional rewiring of fungal galactose-metabolism circuitry. Curr Biol 17:1007–1013

Mayer FL, Wilson D, Jacobsen ID, Miramon P, Große K, Hube B (2012) The novel *Candida albicans* transporter Dur31 is a multi-stage pathogenicity factor. PLoS Pathog 8:e1002592

Miramon P, Dunker C, Windecker H, Bohovych IM, Brown AJP, Kurzai O, Hube B (2012) Cellular responses *of Candida albicans* to phagocytosis and the extracellular activities of neutrophils are critical to counteract carbohydrate starvation, oxidative and nitrosative stress. PLoS One 7:e52850

Mueller PP, Hinnebusch AG (1986) Multiple upstream AUG codons mediate translational control of *GCN4*. Cell 45:201–207

Murad AM, d'Enfert C, Gaillardin C, Tournu H, Tekaia F, Talibi D, Marechal D, Marchais V, Cottin J, Brown AJ (2001a) Transcript profiling in *Candida albicans* reveals new cellular functions for the transcriptional repressors CaTup1, CaMig1 and CaNrg1. Mol Microbiol 42:981–993

Murad AM, Leng P, Straffon M, Wishart J, Macaskill S, MacCallum D, Schnell N, Talibi D, Marechal D, Tekaia F, d'Enfert C, Gaillardin C, Odds FC, Brown AJ (2001b) *NRG1* represses yeast-hypha morphogenesis and hypha-specific gene expression in *Candida albicans*. EMBO J 20:4742–4752

Naglik J, Albrecht A, Bader O, Hube B (2004) *Candida albicans* proteinases and host/pathogen interactions. Cell Microbiol 6:915–926

Nailis H, Kucharíkov S, Ricicov M, Van Dijck P, Deforce D, Nelis H, Coenye T (2010) Real-time PCR expression profiling of genes encoding potential virulence factors in *Candida albicans* biofilms:

identification of model-dependent and -independent gene expression. BMC Microbiol 10:114

Nantel A, Dignard D, Bachewich C, Harcus D, Marcil A, Bouin AP, Sensen CW, Hogues H, van het Hoog M, Gordon P, Rigby T, Benoit F, Tessier DC, Thomas DY, Whiteway M (2002) Transcription profiling of *Candida albicans* cells undergoing the yeast-to-hyphal transition. Mol Biol Cell 13:3452–3465

Natarajan K, Meyer MR, Jackson BM, Slade D, Roberts C, Hinnebusch AG, Marton MJ (2001) Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. Mol Cell Biol 21:4347–4368

Navarro-Garcia F, Sanchez M, Nombela C, Pla J (2001) Virulence genes in the pathogenic yeast *Candida albicans*. FEMS Microbiol Rev 25:245–268

Nicholls S, Straffon M, Enjalbert B, Nantel A, Macaskill S, Whiteway M, Brown AJ (2004) Msn2- and Msn4-like transcription factors play no obvious roles in the stress responses of the fungal pathogen *Candida albicans*. Eukaryot Cell 3:1111–1123

Niimi M, Kamiyama A, Tokunaga M (1988) Respiration of medically important *Candida* species and *Saccharomyces cerevisiae* in relation to glucose effect. J Med Vet Mycol 26:195–198

Nikolaou E, Agrafioti I, Stumpf M, Quinn J, Stansfield I, Brown AJP (2009) Phylogenetic diversity of stress signalling pathways in fungi. BMC Evol Biol 9:44

Nobile CJ, Bruno VM, Richard ML, Davis DA, Mitchell AP (2003) Genetic control of chlamydospore formation in *Candida albicans*. Microbiology 149:3629–3637

Nobile CJ, Schneider HA, Nett JE, Sheppard DC, Filler SG, Andes DR, Mitchell AP (2008) Complementary adhesin function in *C. albicans* biofilm formation. Curr Biol 18:1017–1024

Noble SM, French S, Kohn LA, Chen V, Johnson AD (2010) Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. Nat Genet 42:590–598

Odds FC (1988) *Candida* and candidosis. Bailliere Tindall, London

Odds FC, Brown AJ, Gow NA (2003) Antifungal agents: mechanisms of action. Trends Microbiol 11:272–279

Oliphant AR, Brandl CJ, Struhl K (1989) Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast *GCN4* protein. Mol Cell Biol 9:2944–2949

Paluh JL, Orbach MJ, Legerton TL, Yanofsky C (1988) The cross-pathway control gene of *Neurospora crassa*, cpc-1, encodes a protein similar to *GCN4* of yeast and the DNA-binding domain of the oncogene v-jun-encoded protein. Proc Natl Acad Sci U S A 85:3728–3732

Perlroth J, Choi B, Spellberg B (2007) Nosocomial fungal infections: epidemiology, diagnosis, and treatment. Med Mycol 45:321–346

Petter R, Chang YC, Kwon-Chung KJ (1997) A gene homologous to *Saccharomyces cerevisiae SNF1* appears to be essential for the viability of *Candida albicans*. Infect Immun 65:4909–4917

Pfaller MA, Diekema DJ (2007) Epidemiology of invasive candidiasis: a persistent public health problem. Clin Microbiol Rev 20:133–163

Phan QT, Myers CL, Fu Y, Sheppard DC, Yeaman MR, Welch WH, Ibrahim AS, Edwards JE Jr, Filler SG (2007) Als3 is a *Candida albicans* invasin that binds to cadherins and induces endocytosis by host cells. PLoS Biol 5:e64

Piekarska K, Mol E, van den Berg M, Hardy G, van den Burg J, van Roermund C, MacCallum D, Odds F, Distel B (2006) Peroxisomal fatty acid beta-oxidation is not essential for virulence of *Candida albicans*. Eukaryot Cell 5:1847–1856

Pierce JV, Dignard D, Whiteway M, Kumamoto CA (2013) Normal adaptation of *Candida albicans* to the murine gastrointestinal tract requires Efg1p-dependent regulation of metabolic and host defense genes. Eukaryot Cell 12:37–49

Ramirez MA, Lorenz MC (2007) Mutations in alternative carbon utilization pathways in *Candida albicans* attenuate virulence and confer pleiotropic phenotypes. Eukaryot Cell 6:280–290

Rao GR, Ramakrishnan T, Sirsi M (1960) Enzymes in *Candida albicans*. I. Pathways of glucose dissimilation. J Bacteriol 80:654–658

Rodaki A, Young T, Brown AJP (2006) Effects of depleting the essential central metabolic enzyme, fructose-1,6-bisphosphate aldolase, upon the growth and viability of *Candida albicans*: implications for antifungal drug target discovery. Eukaryot Cell 5:1371–1377

Rodaki A, Bohovych IM, Enjalbert B, Young T, Odds FC, Gow NA, Brown AJ (2009) Glucose promotes stress resistance in the fungal pathogen *Candida albicans*. Mol Biol Cell 20:4845–4855

Roemer T, Jiang B, Davison J, Ketela T, Veillette K, Breton A, Tandia F, Linteau A, Sillaots S, Marta C, Martel N, Veronneau S, Lemieux S, Kauffman S, Becker J, Storms R, Boone C, Bussey H (2003) Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. Mol Microbiol 50:167–181

Rogers PD, Barker KS (2003) Genome-wide expression profile analysis reveals coordinately regulated genes associated with stepwise acquisition of azole resistance in *Candida albicans* clinical isolates. Antimicrob Agents Chemother 47:1220–1227

Rolfes RJ, Hinnebusch AG (1993) Translation of the yeast transcriptional activator *GCN4* is stimulated by purine limitation: implications for activation of the protein kinase *GCN2*. Mol Cell Biol 13:5099–5111

Rolland F, Winderickx J, Thevelein JM (2001) Glucose-sensing mechanisms in eukaryotic cells. Trends Biochem Sci 26:310–317

Rubin-Bejerano I, Fraser I, Grisafi P, Fink GR (2003) Phagocytosis by neutrophils induces an amino acid deprivation response in *Saccharomyces cerevisiae* and *Candida albicans*. Proc Natl Acad Sci U S A 100:11007–11012

Russell CL, Brown AJP (2005) Expression of one-hybrid fusions with *Staphylococcus aureus* lexA in *Candida albicans* confirms that Nrg1 is a transcriptional repressor and that Gcn4 is a transcriptional activator. Fungal Genet Biol 42:676–683

Sabina J, Brown V (2009) Glucose sensing network in *Candida albicans*: a sweet spot for fungal morphogenesis. Eukaryot Cell 8:1314–1320

Sandai D, Yin Z, Selway L, Stead D, Walker J, Leach MD, Bohovych I, Ene IV, Kastora S, Budge S, Munro CA, Odds FC, Gow NA, Brown AJ (2012) The evolutionary rewiring of ubiquitination targets has reprogrammed the regulation of carbon assimilation in the pathogenic yeast *Candida albicans*. MBio 3. doi:10.1128/mBio.00495-12

Sanglard D, Ischer F, Parkinson T, Falconer D, Bille J (2003) *Candida albicans* mutations in the ergosterol biosynthetic pathway and resistance to several antifungal agents. Antimicrob Agents Chemother 47:2404–2412

Sellam A, Hogues H, Askew C, Tebbji F, van Het Hoog M, Lavoie H, Kumamoto CA, Whiteway M, Nantel A (2010) Experimental annotation of the human pathogen *Candida albicans* coding and noncoding transcribed regions using high-resolution tiling arrays. Genome Biol 11:R71. doi:10.1186/gb-2010-11-7-r71

Shapiro RS, Robbins N, Cowen LE (2011) Regulatory circuitry governing fungal development, drug resistance, and disease. Microbiol Mol Biol Rev 75:213–267

Sharkey LL, McNemar MD, Saporito-Irwin SM, Sypherd PS, Fonzi WA (1999) *HWP1* functions in the morphological development of *Candida albicans* downstream of *EFG1*, *TUP1*, and *RBF1*. J Bacteriol 181:5273–5279

Smith RL, Johnson AD (2000) Turning genes off by Ssn6-Tup1: a conserved system of transcriptional repression in eukaryotes. Trends Biochem Sci 25:325–330

Smith DA, Nicholls S, Morgan BA, Brown AJ, Quinn J (2004) A conserved stress-activated protein kinase regulates a core stress response in the human pathogen *Candida albicans*. Mol Biol Cell 15:4179–4190

Sonneborn A, Bockmuhl DP, Ernst JF (1999a) Chlamydospore formation in *Candida albicans* requires the Efg1p morphogenetic regulator. Infect Immun 67:5514–5517

Sonneborn A, Tebarth B, Ernst JF (1999b) Control of white-opaque phenotypic switching in *Candida albicans* by the Efg1p morphogenetic regulator. Infect Immun 67:4655–4660

Srikantha T, Tsai LK, Daniels K, Soll DR (2000) EFG1 null mutants of *Candida albicans* switch but cannot express the complete phenotype of white-phase budding cells. J Bacteriol 182:1580–1591

Stoldt VR, Sonneborn A, Leuker CE, Ernst JF (1997) Efg1p, an essential regulator of morphogenesis of the human pathogen *Candida albicans*, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. EMBO J 16:1982–1991

Swoboda RK, Bertram G, Delbrück S, Ernst JF, Gow NA, Gooday GW, Brown AJ (1994) Fluctuations in glycolytic mRNA levels during morphogenesis in *Candida albicans* reflect underlying changes in growth and are not a response to cellular dimorphism. Mol Microbiol 13:663–672

Tebarth B, Doedt T, Krishnamurthy S, Weide M, Monterola F, Dominguez A, Ernst JF (2003) Adaptation of the Efg1p morphogenetic pathway in *Candida albicans* by negative autoregulation and PKA-dependent repression of the *EFG1* gene. J Mol Biol 329:949–962

Thevelein JM, de Winde JH (1999) Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast *Saccharomyces cerevisiae*. Mol Microbiol 33:904–918

Thewes S, Kretschmar M, Park H, Schaller M, Filler SG, Hube B (2007) *In vivo* and *ex vivo* comparative transcriptional profiling of invasive and non-invasive *Candida albicans* isolates identifies genes associated with tissue invasion. Mol Microbiol 63:1606–1628

Tobudic S, Kratzer C, Lassnigg A, Presterl E (2012) Antifungal susceptibility of *Candida albicans* in biofilms. Mycoses 55:199–204

Tournu H, Tripathi G, Bertram G, Macaskill S, Mavor A, Walker L, Odds FC, Gow NA, Brown AJ (2005) Global role of the protein kinase Gcn2 in the human pathogen *Candida albicans*. Eukaryot Cell 4:1687–1696

Tripathi G, Wiltshire C, Macaskill S, Tournu H, Budge S, Brown AJ (2002) Gcn4 co-ordinates morphogenetic and metabolic responses to amino acid starvation in *Candida albicans*. EMBO J 21:5448–5456

Ueno K, Matsumoto Y, Uno J, Sasamoto K, Sekimizu K, Kinjo Y, Chibana H (2011) Intestinal resident yeast *Candida glabrata* requires Cyb2p-mediated lactate assimilation to adapt in mouse intestine. PLoS One 6:e24759

Uhl MA, Biery M, Craig N, Johnson AD (2003) Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C. albicans*. EMBO J 22:2668–2678

Van Neil CB, Cohen AL (1942) On the metabolism of *Candida albicans*. J Cell Comp Physiol 20:95–112

Vylkova S, Carman AJ, Danhof HA, Collette JR, Zhou H, Lorenz MC (2011) The fungal pathogen *Candida albicans* autoinduces hyphal morphogenesis by raising extracellular pH. MBio 2:e00055–11

Walker LA, Maccallum DM, Bertram G, Gow NA, Odds FC, Brown AJ (2009) Genome-wide analysis

of *Candida albicans* gene expression patterns during infection of the mammalian kidney. Fungal Genet Biol 46:210–219

Wanke C, Eckert S, Albrecht G, van Hartingsveldt W, Punt PJ, van den Hondel CA, Braus GH (1997) The *Aspergillus niger GCN4* homologue, cpcA, is transcriptionally regulated and encodes an unusual leucine zipper. Mol Microbiol 23:23–33

Wek SA, Zhu S, Wek RC (1995) The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase *GCN2* interacts with tRNA and is required for activation in response to starvation for different amino acids. Mol Cell Biol 15:4497–4506

Williamson MI, Samaranayake LP, MacFarlane TW (1986) Biotypes of oral *Candida albicans* and *Candida tropicalis* isolates. J Med Vet Mycol 24:81–84

Wilson D, Thewes S, Zakikhany K, Fradin C, Albrecht A, Almeida R, Brunke S, Grosse K, Martin R, Mayer F, Leonhardt I, Schild L, Seider K, Skibbe M, Slesiona S, Waechtler B, Jacobsen I, Hube B (2009) Identifying infection-associated genes of *Candida albicans* in the postgenomic era. FEMS Yeast Res 9:688–700

Yin Z, Wilson S, Hauser NC, Tournu H, Hoheisel JD, Brown AJ (2003) Glucose triggers different global responses in yeast, depending on the strength of the signal, transiently stabilizes ribosomal protein mRNAs. Mol Microbiol 48:713–724

Yin Z, Stead D, Selway L, Walker J, Riba-Garcia I, McLnerney T, Gaskell S, Oliver SG, Cash P, Brown AJ (2004) Proteomic response to amino acid starvation in *Candida albicans* and *Saccharomyces cerevisiae*. Proteomics 4:2425–2436

Zakikhany K, Naglik JR, Schmidt-Westhausen A, Holland G, Schaller M, Hube B (2007) *In vivo* transcript profiling of *Candida albicans* identifies a gene essential for interepithelial dissemination. Cell Microbiol 9:2938–2954

Zaragoza O, Rodriguez C, Gancedo C (2000) Isolation of the *MIG1* gene from *Candida albicans* and effects of its disruption on catabolite repression. J Bacteriol 182:320–326

Zhao R, Lockhart SR, Daniels K, Soll DR (2002) Roles of *TUP1* in switching, phase maintenance, and phase-specific gene expression in *Candida albicans*. Eukaryot Cell 1:353–365

Zhu Z, Wang H, Shang Q, Jiang Y, Cao Y, Chai Y (2012) Time course analysis of *Candida albicans* metabolites during biofilm development. J Proteome Res 12:2375–2385

Bio-Systematic Index

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)

Subject Index

Fungal Genomics, 2nd Edition
The Mycota XIII
M. Nowrousian (Ed.)