Mycology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tmyc20

Fueling the future with fungal genomics

Igor V. Grigoriev [a], Daniel Cullen [b], Stephen B. Goodwin [c], David Hibbett [d], Thomas W. Jeffries [b], Christian P. Kubicek [e], Cheryl Kuske [f], Jon K. Magnuson [g], Francis Martin [h], Joseph W. Spatafora [i], Adrian Tsang [j] & Scott E. Baker [a g]

[a] US DOE Joint Genome Institute, Walnut Creek, CA, USA

[b] USDA, Forest Service, Forest Products Laboratory, Madison, WI, USA

[c] USDA-Agricultural Research Service, Purdue University, West Lafayette, IN, USA

[d] Biology Department, Clark University, Worcester, MA, USA

[e] Institute of Chemical Engineering, University of Technology, Vienna, Austria

[f] Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA

[g] Chemical and Biological Process Development Group, Pacific Northwest National Laboratory, Richland, WA, USA

[h] Tree-Microbe Interactions Joint Unit, INRA-Nancy Champenoux, France

[i] Department of Botany and Plant Pathology, Oregon State University, OR, USA

[j] Centre for Structural and Functional Genomics, Concordia University, Montreal, QC, Canada

Available online: 25 Jul 2011

To cite this article: Igor V. Grigoriev, Daniel Cullen, Stephen B. Goodwin, David Hibbett, Thomas W. Jeffries, Christian P. Kubicek, Cheryl Kuske, Jon K. Magnuson, Francis Martin, Joseph W. Spatafora, Adrian Tsang & Scott E. Baker (2011): Fueling the future with fungal genomics, Mycology, DOI:10.1080/21501203.2011.584577

To link to this article: http://dx.doi.org/10.1080/21501203.2011.584577

Mycology
2011, iFirst 1–18

Taylor & Francis
Taylor & Francis Group

Fueling the future with fungal genomics

Igor V. Grigoriev[a*], Daniel Cullen[b], Stephen B. Goodwin[c], David Hibbett[d], Thomas W. Jeffries[b], Christian P. Kubicek[e], Cheryl Kuske[f], Jon K. Magnuson[g], Francis Martin[h], Joseph W. Spatafora[i], Adrian Tsang[j] and Scott E. Baker[a,g]

[a]*US DOE Joint Genome Institute, Walnut Creek, CA, USA;* [b]*USDA, Forest Service, Forest Products Laboratory, Madison, WI, USA;* [c]*USDA-Agricultural Research Service, Purdue University, West Lafayette, IN, USA;* [d]*Biology Department, Clark University, Worcester, MA, USA;* [e]*Institute of Chemical Engineering, University of Technology, Vienna, Austria;* [f]*Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA;* [g]*Chemical and Biological Process Development Group, Pacific Northwest National Laboratory, Richland, WA, USA;* [h]*Tree-Microbe Interactions Joint Unit, INRA-Nancy, Champenoux, France;* [i]*Department of Botany and Plant Pathology, Oregon State University, OR, USA;* [j]*Centre for Structural and Functional Genomics, Concordia University, Montreal, QC, Canada*

(*Received 7 April 2011; final version received 19 April 2011*)

Fungi play important roles across the range of current and future biofuel production processes. From crop/feedstock health to plant biomass saccharification, enzyme production to bioprocesses for producing ethanol, higher alcohols, or future hydrocarbon biofuels, fungi are involved. Research and development are underway to understand the underlying biological processes and improve them to make bioenergy production efficient on an industrial scale. Genomics is the foundation of the systems biology approach that is being used to accelerate the research and development efforts across the spectrum of topic areas that impact biofuels production. In this review, we discuss past, current, and future advances made possible by genomic analyses of the fungi that impact plant/feedstock health, degradation of lignocellulosic biomass, and fermentation of sugars to ethanol, hydrocarbon biofuels, and renewable chemicals.

Keywords: biofuels; genomics; fungi

Introduction

Around the world, rapidly growing populations increase the demand for energy. Limited fossil resources and negative ecological impacts of petroleum exploration, extraction, transport, and consumption dictate the need to explore and develop alternative renewable sources of energy. One of the alternatives, cellulosic biofuels, is based on the combined efficiency of plants to produce and store carbon in the form of cellulose and other biopolymers such as hemicelluloses and lignin in their cell walls and of microbes to decompose these biopolymers and convert the resulting sugar-building blocks to renewable fuels, fuel precursors, and chemicals. Genomics offers the tools to explore the molecular mechanisms of these natural processes and to engineer industrial analogs. Several bioprocesses, which currently produce biofuels, biofuel precursors, and renewable chemicals, are fungal based.

The first major fungal genomics milestone was the publication of the whole genome sequence of the yeast *Saccharomyces cerevisiae* (Goffeau et al. 1996), an organism which has played an exceptional role in expanding our basic knowledge of eukaryotic cell physiology, genetics, and biochemistry as well as brewing, baking, and industrial ethanol production. The latter is one of the final steps in converting biomass to a biofuel – sugar fermentation to alcohol. Another industrially utilized and sequenced fungus, *Trichoderma reesei* (Martinez et al. 2008), produces large amounts of enzymes employed in breaking cellulose into simple sugars (glucose and xylose). A significant number of fungal genomes sequenced so far as well as many of those in progress are of medical importance (Cuomo and Birren 2010). Many fungi that are important to bioenergy and the environment still do not have sequenced genomes. To generate and understand the biology underlying fungal bioprocesses for producing enzymes, biofuels, and chemicals we need to start with decoding the genomic blueprint of these organisms, identify parts lists, enzymes, pathways, and fungal hosts, and design industrial processes utilizing these parts (Baker et al. 2008).

The US Department of Energy (DOE) Joint Genome Institute (JGI) has launched a Fungal Genomics Program (FGP) aiming to scale up sequencing and analysis of fungal genomes to explore their diversity and applications for energy and environmental science. We started with several successful single-genome sequencing projects (e.g., Martinez et al. 2004, 2008, 2009; Jeffries et al. 2007; Martin et al. 2008) and are moving toward larger-scale system-level genomics. Genomes

*Corresponding author. Email: ivgrigoriev@lbl.gov

ISSN 2150-1203 print/ISSN 2150-1211 online

DOI: 10.1080/21501203.2011.584577
http://www.informaworld.com

and transcriptomes to be sequenced are selected from user community proposals to annual calls of the JGI Community Sequencing Program (e.g., Martin et al. 2011). The sequenced and annotated genomes are deposited to MycoCosm (http://www.jgi.doe.gov/fungi), a web-based fungal genomics resource. This resource integrates genomics data and tools for comparative genomics and genome-centric analysis and promotes user community participation in data submission, annotation, and analysis.

The key project of FGP, the *Genomic Encyclopedia of Fungi*, focuses on three areas of research connected to bioenergy: (1) plant feedstock health, encompassing symbiosis, plant pathogenicity, and biocontrol; (2) biorefinery, analyzing lignocellulose degradation, sugar fermentation, and industrial organisms; and (3) fungal diversity. Sustainable production of feedstock plants such as switchgrass or poplar depends on plant interactions with fungal symbionts and pathogens. Converting biomass into biofuel requires a detailed inventory of enzymes and processes developed in filamentous fungi for saccharifying plant material and producing hydrocarbon biofuels and in yeasts that ferment sugars to ethanol, higher alcohols, and hydrocarbon biofuels. Knowledge of the molecular biology of organisms serving as industrial hosts in biorefineries is critical. Developing robust industrial processes and their optimization requires organisms and enzymes that are tolerant to a range of temperatures, physical, and chemical conditions, with lower maintenance and higher yield of desired product. These could come from a broader sampling of the phylogenetic and ecological diversity of fungi. Here we summarize current and planned genomics efforts in each chapter of the *Genomic Encyclopedia of Fungi*.

Feedstock

Energy crops, such as *Miscanthus*, switchgrass, cottonwoods (hybrid poplars), hybrid willows, and sugarcane, are grown specifically for their ability to generate energy. In addition, corn and sorghum can also be grown for fuel, while the leftover by-products serve other purposes such as feed or food. Plant health maintenance is critical for sustainable growth of biofuel feedstocks and fungi, as symbionts, pathogens, or biocontrol agents, play a very important role. Symbionts such as mycorrhizae can increase productivity of bioenergy feedstock plants. Pathogens can have dramatic negative effects on yields, quality, and production of bioenergy crops (Ullstrup 1972). Feedstock protection also can be achieved by biocontrol fungi, which are attractive alternatives to the chemical treatments used currently to kill fungi, nematodes, and insects that are pathogenic to plants. Optimizing feedstock plant growth therefore is dependent on understanding the molecular mechanisms of interactions between plants and fungi. In addition, reference genomes of mycorrhizae and other soil-inhabiting fungi also will facilitate comprehensive metagenomics studies of soil and the rhizosphere, studies which until now have been mostly limited to bacterial communities (Tringe et al. 2005).

Symbiosis

Belowground is a complex and often hidden world filled with microbes, such as rhizospheric fungi and bacteria that can interact with the plant roots and make a vital contribution to the major terrestrial ecosystems. In addition, lichens represent another phototroph–heterotroph symbiosis that has yet to be fully explored at genomic level. Approximately 80% of vascular plant species form mycorrhizal associations between their roots and soil-borne fungi (Wang and Qiu 2006), the most prevalent form of plant mutualistic symbiosis known to date. Elucidating the processes that regulate the establishment and maintenance of these symbiotic relationships is essential for understanding the flow of nutrients and carbon within ecosystems. The major types of mycorrhizal interactions include ectomycorrhizal (ECM), arbuscular mycorrhizal (AM) Glomeromycota, and ectendomycorrhizal types, including orchid and ericoid mycorrhizas (Smith and Read 2008). ECM fungi colonize the dominant tree species in temperate and tropical forests where they play a key role in nutrient cycling and carbon sequestration. ECM fungi are not a phylogenetically distinct group, but they are all part of the Dikarya (Ascomycota and Basidiomycota). Ancestors to the Basidiomycetes and Ascomycetes were thought to be saprotrophic (Hibbett and Matheny 2009). From these saprobes, a number of fungal lineages evolved the ability to colonize living root tissues without causing disease. This change, from a saprotroph that decays plant tissues to a mutualistic lifestyle involving evasion of plant defenses to obtain plant-derived sugars, has likely developed a number of times through convergent evolution. Hibbett and Matheny (2009) estimated that the symbiotic habit of certain Basidiomycetes evolved at least eight times with angiosperms and had between six and eight independent origins with gymnosperms. Thus, certain clades (e.g., Boletales, Sebacinales) have both symbiotic and saprotrophic members while others are purely saprotrophic (e.g., Polyporales) or symbiotic (e.g., Thelephorales) (Hibbett and Matheny 2009).

A common symbiotic genome?

Sequencing genomes of saprotrophic and pathogenic fungi has provided an unprecedented opportunity to decipher the key components determining their various lifestyles (Arvas et al. 2007). In contrast, the first ectomycorrhizal genome to be sequenced, that of *Laccaria bicolor* in 2008 (Martin et al. 2008), was found to be very large, at 64.9 Mbp with ~19,000 predicted protein-coding genes, of which 85% have been verified by transcript profiling (Martin et al.

2008). Approximately 25% of the predicted gene models have no known homolog and among these orphan genes there will be a number that is important, or unique, to the symbiotic lifestyle of *L. bicolor* (Martin and Selosse 2008). The genome of the ectomycorrhizal ascomycete fungus *Tuber melanosporum* gave a very different impression as, despite the largest fungal genome to date of 125 Mbp, it was fairly gene poor with only ~7500 predicted protein-coding regions (Martin et al. 2010). There are few similarities between the "symbiosis molecular toolboxes" used by *T. melanosporum* and *L. bicolor* to interact with their hosts. While both ECM species show gene loss in many cell wall-degrading families such as cellulases, and both species have symbiosis-specific gene expression, in neither case are the genes expressed during symbiosis the same, except for a few membrane transporters and a GH5 glycosyl hydrolase (Martin et al. 2010). Thus, based on our present limited genomic view of ECM fungi it appears that the "wheel" has been re-invented each time the symbiotic lifestyle has arisen in the tree of life, potentially through gene duplication and neofunctionalization (Teshima and Innan 2008), selected gene loss as well as genome reshuffling. It remains to be seen whether ECM symbiosis genomes are truly unique to each genus or if there are more similarities between ECM fungi than those found to date. The detailed analysis of symbiotic fungal

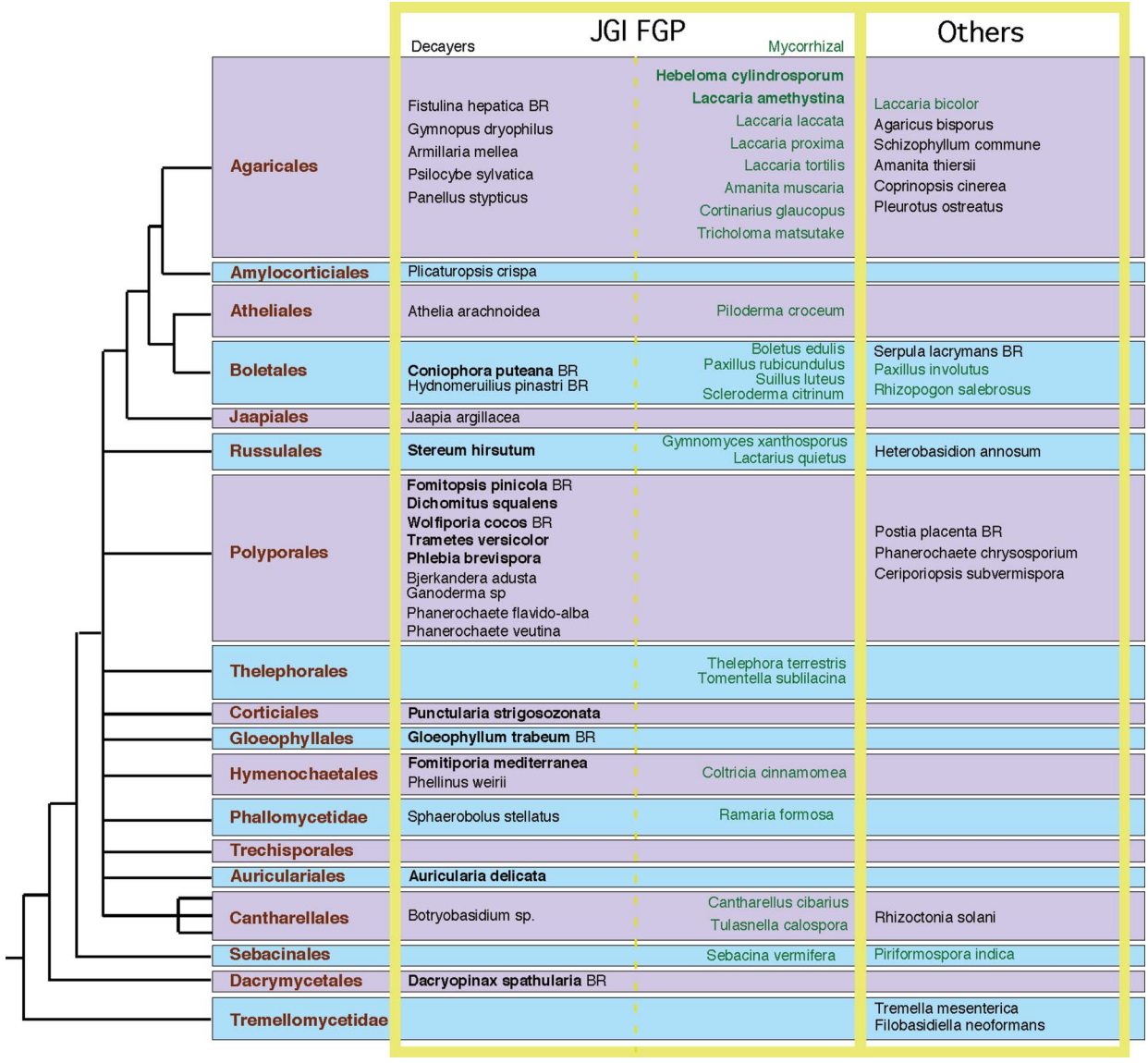

Figure 1. Phylogenetic distribution of genome targets in Agaricomycotina. *Black font* indicates saprotrophs. *Gray font* indicates ectomycorrhiza formers. The "JGI FGP" box contains the taxa being targeted by the JGI Fungal Genomics Program. Saprotrophic taxa proposed here are *left* of the *dashed line*; ectomycorrhizal taxa are *right* of the *dashed line*. Decayers in the JGI FGP box in *bold font* are currently in production; those in *plain font* are targeted for sequencing.

genomes forming different types of mycorrhizae (Figure 1) will shed light on what the symbiosis genes may be as well as determine the true similarities and differences in the strategies developed by different ECM lineages. This will provide unique additional opportunities to test the theory of "tailor-made symbiotic toolboxes" (Martin et al. 2008, 2010; Martin and Selosse 2008).

A phylogenetically driven sequencing of mycorrhizal species

An overall lack of broad phylogenetic considerations in the selection of mycorrhizal genomes for sequencing has led to a strongly biased representation of recognized symbiotic diversity. For a more systematic effort we initiated a large-scale project to sequence 25 genomes of mycorrhizal species selected for (1) their ability to promote plant growth and health; (2) their phylogenetic novelty; (3) their ability to establish different types of mycorrhizal symbiosis (ectomycorrhizas, ericoid and orchid endomycorrhizas); and (4) their taxonomic relationships with already sequenced mycorrhizal genomes to explore the intra-clade variability in symbiosis gene repertoire. The proposed taxa include representatives of the major clades (orders or subclasses) of culturable Mycotina that contain mycorrhizal taxa (Figure 1). This phylogenetically based sample of genomes would propel the field forward and allow us to answer fundamental questions about the evolution of this mutualism and the variation in function and interaction across the phylogenetic depth occupied by these organisms. This project will yield important insights into the evolution of mycorrhizal symbiosis, which has been a driving force in the global development of terrestrial ecosystems.

Plant pathogens

Plant pathogenic fungi pose a major threat to bioenergy production. Availability of a stable supply of plant biomass will be threatened by plant diseases as acreages of bioenergy crops increase. The southern corn leaf blight epidemic caused by a host-specific toxin (T toxin) produced by the ascomycete *Cochliobolus heterostrophus* (genome sequenced by JGI) during the early 1970s destroyed 15% of the US maize crop and caused a major shift in maize breeding and seed production practices (Ullstrup 1972; reviewed in Turgeon and Baker 2007). Black leaf streak or Sigatoka of banana, caused by *Mycosphaerella fijiensis*, requires weekly year-round fungicide applications for disease control in crop monocultures, with a huge impact on precarious tropical environments and the health of field workers. Introduction of the chestnut blight pathogen to North America from Asia eliminated chestnut as one of the dominant tree species of the eastern deciduous forest in only a few years. These are just three of many examples of fungal pathogens causing massive destruction to uniformly susceptible populations of host plants.

Of more recent interest for biomass production is *Septoria musiva*, the asexual stage of the Dothideomycete *Mycosphaerella populorum*, which causes *Septoria* leaf spot and canker of poplar (*Populus* spp.). On native poplars the disease can cause occasional defoliation but usually is a minor problem. However, on the hybrid poplars that have been selected for rapid biomass production the pathogen causes stem cankers that weaken the trunks and lead to breakage within 1–2 m of ground level, with an almost complete loss of the crop (Ostry et al. 1989). This disease makes the production of hybrid poplars for biomass uneconomical in the eastern half of North America, and it has spread recently to new areas in western North America and to South America.

Some potential bioenergy crops, such as the North American grass *Panicum virgatum* (switchgrass) or the Asian native *Miscanthus giganteus*, currently have few major pest or disease problems. However, this is unlikely to last once these crops are grown in pure stands over extensive acreages. Recent estimates of speciation times for fungal pathogens of crop plants have shown that they often occurred within the past 1000–10,000 years, possibly concomitant with domestication of the crop. For example, the wheat pathogen *Mycosphaerella graminicola* was estimated to have diverged from a wild grass pathogen approximately 10,500 years ago (Stukenbrock et al. 2007). The barley pathogen *Rhynchosporium secalis* appears to have evolved on that host from 2500 to 5000 years ago (Brunner et al. 2007). Other pathogens may have evolved within recorded history; the wheat pathogen *Pyrenophora tritici-repentis* may have evolved into a major pathogen by transfer of a toxin gene from another wheat pathogen (*Stagonospora nodorum*) into a minor leaf nibbler possibly as recently as the late 1930s (Friesen et al. 2006). These dynamic, rapidly evolving pathogen populations pose a continual threat to the sustainable production of bioenergy crops. Protecting crops from diseases requires a detailed understanding of host–pathogen interactions to elucidate the mechanisms of fungal pathogenicity and of host resistance. Availability of both host and pathogen genome sequences ultimately will allow interacting pairs of genes to be identified and will permit an unprecedented analysis of host–pathogen interactions.

Fortunately, due to their high economic importance the genomes of many plant pathogenic fungi infecting bioenergy crops have been sequenced or are in progress (Table 1). Primary emphasis of the Plant Pathogens section of the JGI Fungal Genomics Program has been on pathogens of bioenergy crops with sequenced host genomes and on Dothideomycetes. The first to be done was the wheat pathogen *M. graminicola*, which was sequenced to completion during 2005 (Goodwin et al. 2011). This genome is just under 40 Mbp and has 21 chromosomes, 8 of which are dispensable (Wittenberg et al. 2009) and appear to have originated by horizontal transfer from an

Table 1. Genomes of fungal pathogens of bioenergy crops with sequenced host genomes and other primarily Dothideomycetes pathogens.

Species	Order	Host	Disease	Genome size	# Genes	Source
Alternaria brassicicola	Pleosporales	Canola	Black spot	30.3	10, 688	Washington University
Botryosphaeria dothidea	Dothideomycetes incertae sedis	Many	Many	–[a]	–[a]	JGI
Cercospora zeae-maydis	Capnodiales	Maize	Gray leaf spot	–[a]	–[a]	JGI
Cochliobolus heterostrophus	Pleosporales	Maize	Southern corn leaf blight	34.9	9633	JGI
Cochliobolus sativus	Pleosporales	Wheat, barley	Root rot, spot blotch	34.6	–[a]	JGI
Dothistroma septosporum	Capnodiales	Pine	*Dothistroma* needle blight	30.2	12, 580	JGI
Fusarium graminearum	Hypocreales	Maize[b]	Ear rot	36.5	13, 332	Broad Institute
Fusarium verticillioides	Hypocreales	Maize[b]	Kernel and ear rot	41.8	14, 179	Broad Institute
Leptosphaeria biglobosa	Pleosporales	Canola	Blackleg	–[a]	–[a]	JGI
Leptosphaeria maculans	Pleosporales	Canola	Blackleg and stem canker	45.1	12, 469	GenoScope
Melampsora larici-populina	Pucciniales	Poplar	Poplar leaf rust	101	16, 694	JGI
Mycosphaerella fijiensis	Capnodiales	Banana	Black leaf streak or Sigatoka	73.4	10, 952	JGI
Mycosphaerella graminicola	Capnodiales	Wheat	*Septoria tritici* blotch	39.7	10, 952	JGI
Pyrenophora tritici-repentis	Pleosporales	Wheat	Tan spot	37.8	12, 141	Broad Institute, JGI
Rhynchosporium secalis	Helotiales	Barley	Barley scald	54	∼13, 000	Wolfgang Knogge
Sclerotinia sclerotiorum	Helotiales	Canola[b]	Sclerotinia stem rot	38.3	14, 522	Broad Institute
Septoria musiva (*Mycosphaerella populorum*)	Capnodiales	Poplar	*Septoria* leaf spot and canker	29.5	–[a]	JGI
Setosphaeria turcica	Pleosporales	Maize	Northern leaf blight	–[a]	–[a]	JGI
Stagonospora nodorum	Pleosporales	Wheat	Leaf and glume blotch	37.1	12, 380	Broad Institute
Teratosphaeria (*Mycosphaerella*) *nubilosa*	Capnodiales	Eucalyptus	*Mycosphaerella* leaf disease	–[a]	–[a]	JGI
Ustilago maydis	Ustilaginales	Maize	Corn smut	19.7	6522	Broad Institute
Venturia inaequalis	Dothideomycetes incertae sedis	Apple	Apple scab	73	–[a]	GenoScope

[a] In process.
[b] These species have broader host ranges and infect additional hosts.

unknown donor (Goodwin et al. 2011). The genome has a reduced set of genes for cell wall-degrading enzymes, possibly an adaptation to evade host defenses during a long latent phase on the host (Goodwin et al. 2011). Genomes of other *Mycosphaerella* species have 13 chromosomes of about 1 Mbp or larger, equal to the core set for *M. graminicola* (unpublished). The species that are done, in process or targeted for sequencing, provide good coverage for many bioenergy crops (Table 1), but it is impossible to predict which diseases will be important in the future and many more remain to be sequenced.

Effectors are not only used by fungi to cause disease but also can be recognized by the hosts to trigger resistance responses that can be mined from genome sequences in order to help identify new sources of resistance. For example, comparative genomics analyses between the tomato pathogen *Cladosporium fulvum* and banana-infecting *M. fijiensis* identified homologs of genes for effector proteins in each species. Effectors from *M. fijiensis* were functional homologs of those in *C. fulvum* and were recognized by the corresponding resistance gene in tomato (Stergiopoulos et al. 2010), which could provide a new source of resistance against *M. fijiensis* if it can be successfully transferred into banana (Stergiopoulos et al. 2010). Effector-based identification of non-host resistance genes may provide a virtually unlimited source of genetic resources for plant improvement that could greatly increase yields of many crops including those used for bioenergy. This potential, unlocked by the availability of genome sequences, is only beginning to be tapped.

Another important benefit from genomic sequencing projects could be the identification of enzymes that potentially could be harnessed for improved production of bioenergy. Pathogens have adapted to their hosts through thousands of years of evolution. Those that are successful have been selected for efficient decomposition of host biomass. Comparisons of multiple pathogens of grass hosts (type II cell walls) with pathogens of non-grass monocots and dicots (type I cell walls) will help to identify characteristics of enzymes that are adapted to breaking down the cell walls of each plant type. Furthermore, by including saprophytes and other fungi with unusual lifestyles in comparative genomics analyses there is the potential to identify enzymes for degrading biomass that is quite different from those encountered in pathogens. Because they are the result of thousands of years of evolution of ability to break down the specific chemical components of their host cell walls, some of these enzymes could be utilized, for example, to improve fermentation of cellulosic ethanol.

In addition to helping provide a stable source of biomass feedstocks, controlling plant diseases can yield several incidental benefits to the environment. First among these is a reduction of greenhouse gases, both by lowering fertilizer use and by increasing plant health. Without chemical controls, pathogens make production of both bioenergy and agricultural crops uneconomic in many areas. Increased use of fossil fuel-derived nitrogen fertilizers to overcome low yields leads to the release of nitrous oxide, a greenhouse gas that also causes ozone depletion. Better health of crop plants, particularly of nitrogen fixers such as soybean, will reduce greenhouse gases and help preserve ozone. A second environmental benefit is reduction of greenhouse gases by improving forest health. Several recently sequenced pathogens (*Melampsora larici-populina* (Duplessis et al. 2011), *Dothistroma septosporum*, and *S. musiva*) infect trees that are major components of temperate forests. Recent epidemics of *Dothistroma* needle blight have devastated pine forests in both the northern and southern hemispheres and may be exacerbated by climate change (Woods et al. 2006). Disease epidemics that destroy natural and cultivated forests greatly reduce their ability to fix and sequester carbon. Attaining the detailed understanding of host–pathogen interactions that is needed to mitigate plant diseases will be greatly facilitated by availability of genomic sequences for both hosts and pathogens and will be essential for the future success of the bioenergy industry.

Biocontrol

Besides plant pathogenic fungi and bacteria that negatively impact feedstock production, some diseases are caused by larger organisms such as insects or pathogenic nematodes (particularly on *Miscanthus* and switchgrass). Due to the adverse environmental effects of pesticides which create health hazards for human and other non-target organisms, they have been the object of substantial criticism in recent years. The development of safer, environmentally feasible alternatives for plant protection has therefore become a top priority. In this context, biological control ("biocontrol") has become an urgently needed component of agriculture.

Mycoparasites. The excellent ability of species of the fungal genus *Trichoderma* to suppress diseases and to stimulate the growth and development of plants is well known (Baker 1987). The beneficial effects of *Trichoderma* spp. on plants have been attributed to their capability to antagonize soil-borne pathogens by a combination of mycoparasitism, secretion of antibiotics, and competition for space and substrates (Howell 2003). However, it is now known that these biocontrol agents are also able to interact intimately with plant roots, colonizing the outer epidermal layers, and to act as opportunistic, avirulent plant symbionts (Harman et al. 2004). To expand our understanding of the biochemical and genetic mechanisms that form the basis of the biocontrol abilities, the genomes of *Trichoderma atroviride* IMI 206040 and *Trichoderma virens* Gv29-8 were sequenced recently (Kubicek et al. 2011). Their genome sizes were in the range of those of other multicellular

Ascomycetes (e.g., 36.1 Mbp for *T. atroviride* and 38.8 Mbp for *T. virens*). Gene modeling and verification yielded 11,863 gene models for *T. atroviride* and 12,427 gene models for *T. virens*, among which 1856 and 2215 genes, respectively, were unique to each species. Gene clusters that, in comparison to other fungi, have expanded the most in *T. virens* and *T. atroviride* include the richest known arsenal of polyketide synthases, non-ribosomal peptide synthases, and hybrids of both. Similarly, the enzymes acting on chitin and chitosan as well as proteins mediating chitin binding have significantly proliferated in these two species. Ankyrins (a family of adaptor proteins that mediate the anchoring of ion channels or transporters in the plasma membrane), proteins with WD40 that act as hubs in cellular networks, or HET (heteroincompatibility) domain proteins and sugar transporters also were significantly expanded in *T. virens* and *T. atroviride*. These features could help these species to compete more successfully for limited substrates and to respond faster to their presence. A microarray analysis of the expression of their genes when confronted with a potential host (*Rhizoctonia solani*) revealed a number of significant differences between the two mycoparasites, illustrating that different *Trichoderma* spp. may use different strategies to combat their host or prey. Genome sequencing and analysis of additional *Trichoderma* biocontrol agents, e.g., those that can act as plant endophytes, may lead to further discoveries related to biocontrol. The availability of genome sequences from several pathogenic fungi and potential host plants provides a challenging opportunity to develop a deeper understanding of the processes by which *Trichoderma* affects plant pathogens in nature.

Entomopathogens. Fungal pathogens of insects may provide a possible new generation of safe, selective insecticides. Over a thousand pathogens have been isolated from insects, and many of these are potential candidates for development into microbial insecticides. To date, 13 species or subspecies of insect pathogenic fungi have been formulated and registered as mycoinsecticides, mostly consisting of the Ascomycetes *Metarhizium anisopliae* and *Beauveria bassiana* (de Faria and Wraight 2007). Formulations of *M. anisopliae* have been developed commercially for control of various beetles, termites, flies, gnats, thrips, and ticks. In other countries, several strains of *M. anisopliae* have been developed as biological control agents for grasshoppers, locusts, cockchafers, spittlebugs, grubs, and borers and have been released on a large scale (St. Leger and Wang 2009). The employment of various molecular biological techniques has identified adhesins, extracellular enzymes, and systems for evading host immunity that may be used to create novel combinations of insect specificity and virulence. Currently, the complete sequencing of four insect pathogenic fungi (*B. bassiana*, *M. anisopliae* var. *acridium*, *M. anisopliae* var. *anisopliae*, and *Cordyceps militaris*) is scheduled to be completed during 2011. Sequencing of these genomes will help determine the identity, origin, and evolution of traits needed for diverse lifestyles and host switching. Such knowledge will allow multiple pathogens with different strategies to be used for different ecosystems and mitigates the possibility of the host developing resistance.

Nematophagous fungi. Fungal antagonists of nematodes include nematode-trapping fungi, predacious fungi, endoparasitic fungi, egg-parasitic fungi, cyst-parasitic fungi, and fungi that produce nematotoxic metabolites (Mankau 1980). *Arthrobotrys oligospora* was the first nematode-killing fungus described (Zopf 1888), but this trait is also exhibited by other taxa such as *Clonostachys rosea* (formerly *Gliocladium roseum*) and several *Trichoderma* spp. Some of these fungi produce diverse structures to capture their prey such as hyphal traps or adhesive-trapping nets, whereas others simply poison the prey by a combination of antibiotics and proteases. *Trichoderma*, as an example of the latter, colonizes eggs and second-stage juveniles and sometimes also penetrates the egg masses (Sharon et al. 2001). Our understanding of the host–fungus interaction in this process at the molecular level is nevertheless poorly understood. As with mycoparasitism, lytic enzymes and toxins are important virulence factors in the infection process, but more research is necessary to better understand the mechanisms underlying the infection of nematodes by fungi. A genome sequencing effort on *A. oligospora* (ATCC 24927) is currently being in process (An et al. 2010). However, the availability of *Trichoderma* spp. that can efficiently control nematodes would be attractive additional candidates, given the availability already of three sequenced genomes that span a wide phylogenetic distance. In fact, *Trichoderma longibrachiatum* (which is found frequently as an antagonist of nematodes in soil; I.S. Druzhinina and C.P. Kubicek, unpublished data) is phylogenetically close to the already sequenced *T. reesei* (Martinez et al. 2008) and would allow a straightforward identification of traits associated with interaction with nematodes.

The biorefinery

Biorefinery methods convert biopolymers such as cellulose into simple sugars (e.g., glucose and xylose) and then into biofuels employing fungal strains optimized for large-scale industrial processes. Knowing the enzymes and processes employed by diverse fungi in lignocellulose degradation and sugar fermentation as well as understanding the molecular biology of strains adopted by industry is essential for development of robust platforms for biomass-to-biofuel production on an industrial scale. Genome sequencing in this area should provide a comprehensive parts list: a catalog of enzymes, metabolic processes, and regulatory and secretory mechanisms. Resequencing of industrial strains

should help to map desirable properties such as morphology, hyperproductivity, and thermostability to their genomic blueprints.

Yeast fermentation and pentose conversion

Economical bioconversion of lignocellulosic sugars to ethanol requires rapid, high-yield fermentations of both cellulosic and hemicellulosic sugars (chiefly D-xylose, D-glucose, D-mannose, D-galactose, and L-arabinose) in the presence of acetic acid and a number of other hemicellulose and lignin degradation products such as furfural, hydroxymethyl furfural, ferulic, and hydroxycinnamic acids (Himmel et al. 1997; Morrison et al. 1998; Saha 2003). Yeasts conventionally used for industrial ethanol production such as *S. cerevisiae*, *Saccharomyces bayanus*, and various hybrids, such as *Saccharomyces carlsbergensis*, produce ethanol rapidly from glucose, mannose, or sucrose, but they are not capable of fermenting xylose, arabinose, or cellobiose, and they metabolize galactose relatively slowly. Because of these deficiencies, a great deal of research has gone into metabolic engineering of *S. cerevisiae* for improved xylose and cellobiose metabolism (Galazka et al. 2010; Hahn-Hägerdal et al. 2007a, b; Jeffries and Jin 2004).

In approaching this problem from a fungal genomics perspective, researchers have examined a number of unconventional yeasts. These include the native xylose-fermenting species, *Pachysolen tannophilus* (Schneider et al. 1981; Slininger et al. 1982), *Candida shehatae* (Dupree and Vanderwalt 1983), *Scheffersomyces (Pichia) stipitis* (Jeffries et al. 2007, Jeffries and Van Vleet 2009; Smith et al. 2008; Kurtzman and Suzuki 2010), and the related species, *Spathaspora passalidarum* (Nguyen et al. 2006) and *Spathaspora arborariae* (Nguyen et al. 2006). These "unconventional" yeasts are highly divergent from *S. cerevisiae* (Dujon 2006, 2010) and are commonly found in association with wood-boring beetles or rotting wood rather than fruits, plant exudates, flowers, and pollinating insects. As such, the beetle-associated yeasts are capable of breaking down and fermenting a wide range of sugars including xylose, cellobiose, arabinose, galactose, and even some lignin-related compounds (Jeffries and Van Vleet 2009).

One interesting discovery to emerge by sequencing the complete genome of *S. stipitis* is that this yeast possesses numerous genes for the rapid metabolism of cellobiose. Moreover, they are organized into functional, co-regulated clusters that include β-glucosidases, *endo*-glucanases, and several sugar transporters that enable the direct uptake of cellobiose and cellodextrins (Galazka et al. 2010; Jeffries and Van Vleet 2009). A number of genes from *S. stipitis*, such as those for xylose reductase (*XYL1*), xylitol dehydrogenase (*XYL2*), D-xylulokinase (*XYL3*), transaldolase (*TAL1*), and various sugar transporters, have been engineered into *S. cerevisiae* to improve its capacity for xylose metabolism.

Beyond the fermentation of xylose and cellobiose, however, unconventional yeasts have many traits that could improve biomass conversion. These include the capacity to ferment L-arabinose, as in *Candida arabinofermentans* (Kurtzman and Dien 1998), the ability to rapidly utilize hemicellulosic sugars in the presence of sulfite, as in *Candida utilis* (Simard and Cameron 1974), and a high capacity for starch utilization (*Hyphopichia burtonii*) (Kato et al. 2007) or casein degradation (*Candida caseinolytica*) (Buzzini and Martini 2002). Some yeasts, such as *Pichia membranifaciens*, have a high tolerance for brine and acetic acid (Arroyo-Lopez et al. 2008). Others, such as *Lipomyces starkeyi*, have a high capacity for lipid production (Zhao et al. 2008; Wild et al. 2010).

The genomes of all of the unconventional yeasts listed here have either been sequenced or are part of an ongoing effort by the JGI to sequence and annotate a wide range of yeast species that can contribute to our efforts for the efficient bioconversion of lignocellulosics. Given the strong traditional role that yeast fermentations have played in bioconversion thus far, we can expect these efforts to be highly productive. By expanding the range of sequenced genomes to encompass the entire taxonomic and physiological diversity of both conventional and unconventional yeasts, we expect to capture many more traits useful for bioconversion.

Wood decay in Agaricomycotina

Lignocellulosic plant materials comprise the most abundant biopolymers in terrestrial ecosystems. Fungal-mediated degradation of lignocellulose is a critical link in the carbon cycle and is of great interest for its applications in the production of biomaterials, including biofuels, enzymes, and other commodities. Diverse fungi attack plant cell wall components, including Dikarya (Ascomycota and Basidiomycota) and multiple lineages of "basal fungi" (traditionally classified as chytridiomycetes and zygomycetes; James et al. 2006a). The Agaricomycotina, one of the three subphyla of Basidiomycota (Hibbett et al. 2007), is of particular interest as it contains active and abundant degraders of all classes of plant tissues and the majority of taxa capable of attacking woody substrates.

Two principal modes of wood decay occur in the Agaricomycotina, white rot and brown rot. White rot species are capable of efficiently degrading all components of plant cell walls, including the highly recalcitrant lignin fraction. Brown rot species modify but do not appreciably remove the lignin, which remains as a polymeric residue (Blanchette et al. 1995). Thus, brown rot fungi have evolved efficient mechanisms for circumventing

lignin while obtaining carbon nutrition from cellulose and hemicellulose (Worrall et al. 1997).

As of this writing, complete genome sequences have been published for three wood-decaying Agaricomycotina, *Phanerochaete chrysosporium, Postia placenta,* and *Schizophyllum commune* (Martinez et al. 2004, 2009; Ohm et al. 2010). Even this very-restricted sampling reveals tremendous diversity of wood-decay mechanisms. For example, *P. chrysosporium* has 15 class II fungal peroxidases, which function in lignin degradation (Hofrichter et al. 2010; Lundell et al. 2010), whereas *P. placenta* has only one low-redox potential peroxidase, of unknown function. In addition, *Postia* lacks exocellobiohydrolases and cellulose-binding modules, which are previously known from all cellulolytic fungi (Baldrian and Valaskova 2008). *S. commune* is regarded as a white rot species but it lacks class II peroxidases and in this regard is more similar to the brown rot *P. placenta* as well as the ectomycorrhizal *L. bicolor* (Agaricales; Martin et al. 2008). These findings from just three species in two orders suggest that there has been extensive diversification in the decay mechanisms in Agaricomycotina, warranting a broad sampling of the clade.

The Agaricomycotina contains approximately 18 major independent clades (Hibbett 2006; Figure 1). The most intensively studied white rot species is *P. chrysosporium* (Polyporales), although other white rot taxa, such as *Trametes versicolor* and *Pleurotus ostreatus*, have also been investigated. Brown rot fungi are nested within at least five major clades of Agaricomycotina. Research on the chemistry of brown rot has focused on *P. placenta* (Polyporales), *Gloeophyllum trabeum* (and *Gloeophyllum sepiarium*; Gloeophyllales), and *Coniophora puteana* and *Serpula lacrymans* (Boletales). The brown rot mechanisms in these taxa are independently evolved and should not be assumed to have common genetic or biochemical bases. Moreover, the Agaricomycotina has been estimated to be roughly 400 million years old (Taylor and Berbee 2006), and it is therefore likely that the ancestral white rot decay apparatus has been extensively modified in independent lineages.

As part of the JGI Fungal Genomics Program, 12 species of wood-decaying Agaricomycotina are now being analyzed, including 7 white rot species and 5 brown rot species. The targets represent eight orders, of which seven have no published genomes (although other projects are underway in several orders). They include representatives of two early diverging "jelly fungus" clades, Dacrymycetales and Auriculariales, which may provide clues to the early evolution of wood-decay mechanisms in Agaricomycotina (Matheny et al. 2007). Other target species represent ecologically important groups for which there are not yet genome sequences, such as Hymenochaetales, Gloeophyllales, and Corticiales. Additional targets will further expand the diversity of wood-decaying Agaricomycotina genomes, including representatives of several small clades (e.g., Atheliales) that are closely related to the hyperdiverse Agaricales and Boletales (Binder and Raines 2010). Collectively these genomes, and those from other independent projects (many supported by the JGI), will illuminate the diversity and evolution of wood-decay mechanisms in Agaricomycotina and provide new tools for biotechnological applications.

Hydrocarbons

Recently, there has been a dramatic increase in interest with regard to discovering and exploiting biosynthetic routes to hydrocarbon biofuels or biofuel precursors. Fungi are a logical source for known and novel hydrocarbons and an excellent platform for their production. They are known to produce at least four families of compounds with hydrocarbon backbones that start from acetyl-CoA (Table 2): first, the fatty acids that accumulate in many species as a storage compound; second, the polyketides, which have tremendous diversity in carbon number, degree of saturation, number of hydroxyl or carbonyl groups, etc.; third, the terpenes or isoprenoids, which are another diverse family of compounds assembled from five carbon units; and fourth, the alkanes or alkenes that may be derived by decarbonylation of fatty acids or other novel mechanisms.

As the number of sequenced fungal genomes increases, the known diversity of genes encoding enzymes and pathways that produce these compounds has increased multiplicatively since every fungal genome encodes large numbers of fatty acid synthases, polyketide synthases, terpene synthases, and modifying enzymes associated with these pathways. These pathways need to be screened for the variety of structures that they produce, which can be solved by mass spectroscopy and nuclear magnetic resonance

Table 2. Five families of hydrocarbons produced by fungi.

Compound	No. of carbons	Key enzyme(s)
Fatty acids, triglycerides	C_{12}–C_{22}	Fatty acid synthases
Fatty alcohols	C_{12}–C_{22}	Acyl-reductases
Polyketides	$C_{(2n)+1}$; C_9–C_{65}	Polyketide synthases, oxidases, modifying enzymes
Terpenes (isoprenoids)	C_{5n}; C_{10}–C_{40}	Terpene synthases, modifying enzymes
Alkanes, cycloalkanes	C_4–C_{36}	Acyl-reductases, thioesterases, decarbonylase

spectroscopy. Optimization of metabolic flux from various carbon sources to the primary metabolite-building blocks (e.g., acetyl-CoA) and pathways that are used to synthesize these hydrocarbon molecules requires genome-scale metabolic models such as those for *S. cerevisiae* and *Aspergillus* spp. (Herrgard et al. 2008; Andersen et al. 2008) to predict genes whose deletion or overexpression increases carbon flux toward building block molecules.

Although much remains to be explored and learned, significant progress on these fronts has been made. For example, production of biodiesel from *Mucor circinelloides* cultures has been demonstrated (Vicente et al. 2009, 2010). The genome of this oleaginous zygomycete fungus has been sequenced and can be utilized to accelerate the generation of strains with increased levels of lipid accumulation. Other examples of oleaginous or carotenogenic yeasts are *Lipomyces*, *Debaryomyces*, *Rhodotorula*, *Candida*, *Yarrowia*, and *Hansenula*; and filamentous fungi *Cunninghamella*, *Mortierella*, *Blakeslea*, and *Phycomyces* (all zygomycetes). Many of these genera (i.e., *Debaryomyces*, *Candida*, *Yarrowia* and *Phycomyces*) now have sequenced representatives and genetic tools are available for most of the yeasts. Using these oleaginous fungi as platforms would be potentially beneficial for two reasons. First, many fungal-derived hydrocarbons are generated by a few additional steps after fatty acid synthesis; thus, it follows that fatty acid anabolism will be crucial for generating large quantities of the hydrocarbons. Second, these oleaginous fungi already have mechanisms for accumulating and tolerating high concentrations of hydrophobic compounds, so they might be well equipped to accumulate high concentrations of hydrocarbons. This is important since high concentrations of hydrocarbons may be toxic to the native fungal producers of the hydrocarbons.

Even more intriguing, given the interest in alternative biofuels, are the examples of fungi producing one or more alkanes. These may arise from the fatty acid decarbonylation pathway, via the terpene pathway, or some other undetermined route of biosynthesis. The decarbonylation route is thought to involve multiple enzymes, including acyl-coenzyme A (acyl-CoA) reductases, thioesterases, and decarbonylase (Schneider-Belhaddad and Kolattukudy 2000). The result of this is the production of a hydrocarbon that is one carbon shorter than the starting fatty acid. *Hormoconis resinae* (aka *Cladosporium resinae*) is an example of a fungus capable of producing alkanes from C7 to C36 while growing heterotrophically on glucose (Walker and Cooney 1973). This fungus is now the subject of a genome sequencing project in a collaboration involving Pacific Northwest National Laboratory, the DOE Office of Science (DOE-SC)-funded Joint BioEnergy Institute (JBEI), and JGI. Volatile compounds are produced by a variety of fungi, including species of the cosmopolitan genera *Penicillium* and *Aspergillus*. They produce a range of volatiles, including hydrocarbons, short-chain alcohols, ketones, ethers, esters, and terpenes (Sunesson et al. 1995). All of these fungi accumulate low levels of the desirable hydrocarbons, and the biochemistry and genetics underlying the production of these compounds remain poorly understood. The availability of genome sequences for these organisms will aid the identification of key metabolic pathways and the generation of strains needed for production of hydrocarbons.

Resequencing industrial host strains

Selection is a powerful tool that has been employed in genetics labs across a number of systems for decades. Production of biofuels and bioproducts is no exception; improved pentose utilization, higher enzyme secretion, decreased carbon catabolite repression, and increased organic acid production are just a few examples of desirable industrial phenotypes.

As of this writing, only two publications detail high-throughput sequencing technologies to address biofuel production-related phenotypes in fungi. A study of *Pichia stipitis* compared sequencing technologies (454, Illumina and SoLid) on a strain improved for utilization of xylose for ethanol production to identify the 14 mutations in the improved strain and 3 errors in the reference sequence (Smith et al. 2008). A second study focused on a mutant lineage of the enzyme-producing fungus, *T. reesei*, identified 15 small indels, 18 large deletions, and 223 single-nucleotide changes (Le Crom et al. 2009). This study compared 2 out of 3 strains in a mutant lineage and found that only 42 genes contained nucleotide changes in RUT-C30, the end strain of the lineage. The functional categorization of the mutated genes was relatively narrow and the majority was involved in transcription control, RNA stability, the vacuolar and secretion systems, metabolism, and nuclear–cytoplasmic transport (Le Crom et al. 2009). A major lesson learned from the limited amount of mutated genes identified is that selection is a powerful "force" that acts with a surprisingly high level of precision.

What does the future of resequencing hold for fungal genomics and bioenergy? One possibility is that decreased cost and increased throughput of genome resequencing will make forward-genetic strain improvement an attractive method for determining genes whose mutation (and manipulation) leads to phenotypes that include improved productivity, substrate utilization, and increased resistance to growth inhibitors.

Search for novel enzymes

Over a million species in the Kingdom Fungi have evolved in the past billion years to occupy diverse ecological niches and have accumulated an enormous but yet undiscovered natural arsenal of innovations potentially useful for humans. While the number of fungal genome

sequencing projects continues to increase, the phylogenetic breadth of current sequencing targets is extremely limited. Exploration of phylogenetic and ecological diversity of fungi by genome sequencing is therefore a potentially rich source of valuable metabolic pathways and enzyme activities that will remain undiscovered and unexploited until a systematic survey of phylogenetically diverse genome sequences is undertaken.

Extremophiles

Many fungi that grow optimally under moderate conditions can survive extreme environmental conditions. They are classified as extreme-tolerant fungi. Relatively, few fungi are true extremophiles, i.e., species that grow best in extreme environments. The mechanisms underlying their abilities to thrive under extraordinary conditions are in general poorly understood. Nonetheless fungal extremophiles are thought to have evolved unique defenses as well as unusual cellular mechanisms to support their lifestyle. They are therefore fertile grounds to discover novel compounds and enzymes. The ability to perform reactions under extreme conditions can provide extremophiles an advantage in the production of fuels and chemicals. Deciphering the genomes of these organisms would lay the foundation for uncovering new processes, developing novel applications and products, and understanding the mechanisms by which organisms overcome abiotic stresses.

Thermophilic fungi are defined as those which grow optimally between 40 and 50 °C (Maheshwari et al. 2000). Industrial processes operating at elevated temperature offer advantages over moderate temperatures such as increased reaction rates, reduced risk of microbial contamination, and enhanced mass transfer. For these reasons, it is highly desirable to obtain thermophilic yeasts which are proficient in fuel production. In addition, prevailing wisdom warrants that enzymes from thermophilic fungi are likely to be thermostable, affording them utility in industry. Some 30 species of filamentous fungi have been characterized to be thermophilic (reviewed in Maheshwari et al. 2000; Crisan 1973). These thermophilic fungi have been primarily isolated from soil, compost, and decaying wood. Most of the characterized species fall into three lineages: Eurotiomycetes (e.g., *Thermomyces lanuginosus* and *Talaromyces thermophilus*), Sordariomycetes (e.g., *Thielavia terrestris* and *Myceliophthora thermophila*), and Mucoromycotina (e.g., *Rhizomucor miehei* and *Rhizomucor pusillus*). The genomes of *T. terrestris* and *M. thermophila* have been sequenced by the JGI. Comparative genomics and transcriptomics revealed that these two fungi use similar enzymatic strategies in cellulose and xylan degradation, but employ enzymes with different functionalities in the breakdown of pectin. The genome sequences of many other identified thermophilic fungi are being determined by a Genome Canada project (http://fungalgenomics.ca). Some yeasts are capable of growing at 50°C, but their optimum temperature for growth remains in the 30–35°C range (Shin et al. 2001; Gellissen et al. 2005). The characterized species may represent only a subgroup of thermophilic fungi found in nature. In addition to the characterized species, two unidentified thermophilic Basidiomycetes had been isolated from compost (Straatsma et al. 1994). Moreover, a search for thermophilic fungi in semi-arid grasslands resulted in over 450 isolates, some of which represent species that have not been characterized previously as thermophilic (Bustamante 2006).

Psychrophiles have an optimum growth temperature below 16°C and cease to grow at about 20°C. The tundra, subarctic regions, and oceans are habitats for psychrophilic fungi and yeasts. Many mitosporic fungi, such as *Sclerotinia boreali*, and species of *Typhula* are examples of psychrophiles (reviewed in Magan 2007). Among other mechanisms used to protect against low temperature, psychrophilic fungi employ unusual antifreeze proteins (Hoshino et al. 2003).

Acidophiles can grow as low as pH 1.0 and have optimum growth below pH 4.0 (Ingledew 1990). Several species of Basidiomycetes and Ascomycetes, especially those belonging to Dothideomycetes and Eurotiomycetes, have been identified from acidic mine wastes (Baker et al. 2004; Lopez-Archilla et al. 2004). Molecular classification indicated that the genetic diversity of acidophiles isolated from these environments is limited. Moreover, these acidophiles have close relatives in the neutrophils suggesting that their acidic existence is a recent adaptation (Baker et al. 2009). Lignocellulosic biomass is typically treated with acids prior to enzymatic conversion into fermentable sugars. Enzymes acting at low pH would therefore be useful for the conversion step.

Alkalophiles grow optimally at pH > 9.0 and/or are unable to grow below pH 8.5 (Kroll 1990). A strain of *Chrysosporium tropicum* isolated as a keratin decomposer has been described as alkalophilic (Hubalek 1976). *Acremonium alcalophilum*, isolated from pig manure sludge (Okada et al. 1993), which grows optimally at pH > 9.0, is considered a true alkalophile. The genome sequence of this cellulolytic fungus has recently been completed by the JGI. Besides manure sludge, desert soils and soda lakes are promising locales to isolate alkalophilic fungi.

Halophiles inhabit hypersaline environments such as saline lakes, solar salterns, and subglacial ice. In addition to high salt concentrations, these settings are exposed to a wide range of temperature, pH, light intensity, and nutrient

conditions. Hence halophiles often have to acquire tolerance to other aspects of environmental extremes. Fungi commonly found in hypersaline water are black yeasts and species of the Ascomycetes *Eurotium* as well as some *Cladosporium* species (Butinar et al. 2005a, b). Basidiomycete species belonging to the genus *Wallemia* are also halophilic (Zalar et al. 2005). Accumulation and/or increased synthesis of glycerol or polyols has been a hallmark feature of osmoregulation. The mitogen-activated protein kinase (MAPK) plays a key role in osmoregulation. Expression of the *MAPK* gene of the halophilic *Eurotium herbariorum* can confer tolerance to hypersalinity in budding yeast (Jin et al. 2005). Salt tolerance can have a major impact on agriculture and the development of energy crops.

Irradiation and fungi. Soil fungi are exposed to non-ionizing (UV) and ionizing (natural and manmade) radiation. Pigmented fungi predominate in radiation-contaminated soil and materials, even in and around the Chernobyl Nuclear Power Plant after the meltdown (Zhdanova et al. 2004). *Cladosporium cladosporioides*, *Cladosporium sphaerospermum*, and *Paecilomyces lilacinus* are commonly found in contaminated sites (Zhdanova et al. 2004). It has been generally assumed that dark pigmentation provides protection against radiation. One intriguing idea put forward suggests that the pigmented fungi are using ionizing radiation as an energy source (Dadachova et al. 2007), implying that the genomes of these organisms encode novel mechanisms for molecular energy transfer. If they exist, understanding these mechanisms could have a major impact on the development of renewable energy and environmental remediation.

Sampling phylogenetic and ecological diversity of fungi

While tremendous progress has been made in the sequencing of fungal genomes, many of the major branches of the fungal tree of life (FTOL) lack representation in current sequencing efforts (Figure 2). These gaps in sampling limit a more complete biological understanding of fungi and their ultimate application in human, industrial, and scientific endeavors (e.g., biofuels and carbon sequestration). Very little is known about the metabolic and enzymatic diversity across the FTOL. Genomic sequencing efforts have focused on the phyla Ascomycota and, to a lesser degree, Basidiomycota. Furthermore, lineages

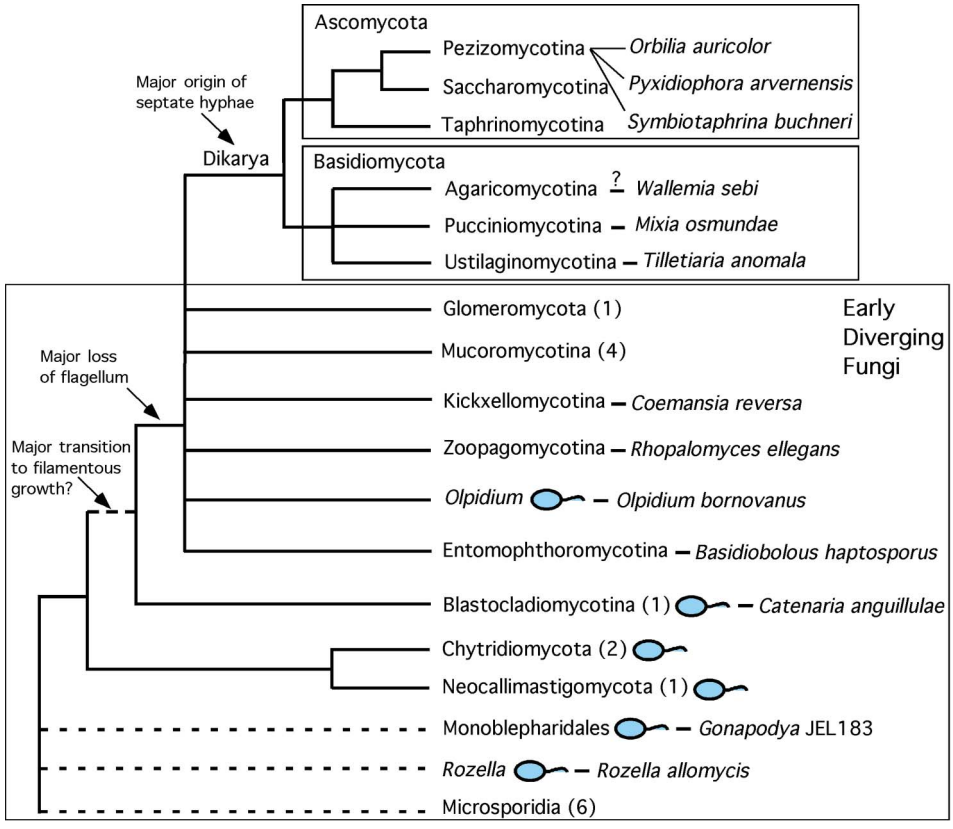

Figure 2. The fungal tree of life. Phylogenetic tree of the current phyla, subphyla, and incertae sedis lineages of the Kingdom Fungi. Numbers in parentheses are approximate number of genome projects completed or in progress (http://fungalgenomes.org/wiki/Fungal_Genome_Links). Species names designate target taxa for the early diverging lineages that will be sequenced within the JGI Fungal Genomics Program.

containing well-studied model organisms (e.g., *Neurospora* of Sordariomycetes, *Saccharomyces* of Saccharomycetes) constitute the majority of the currently available genomes. The unexplored classes, subphyla, and phyla of the FTOL are populated by pathogens, beneficial symbionts, potential biocontrol agents, and inhabitants of unique ecological niches, and their genomes contain potentially useful enzymes and metabolic pathways. Many of the lineages represent divergences that span more than 500 million years of evolution. Thus, by sequencing the genomes of these fungi, large number of novel enzymes may be uncovered. Populating the unexplored regions of fungal diversity with reference genomic sequences will build a platform for enzyme and metabolite discovery as well as provide a basis for future complex metagenomic studies.

Molecular phylogenetics has radically changed our view of fungal evolution and classification in the last 20 years, and only recently has the mycological community been able to have a glimpse at the "true" FTOL. Until recently (Blackwell et al. 2006; Hibbett et al. 2007), the Kingdom Fungi was classified into four major phyla – Chytridiomycota, Zygomycota, Basidiomycota, and Ascomycota – based primarily on morphological traits associated with hyphae and reproductive structures (Alexopoulos et al. 1996). Although this classification was relatively simple to communicate and encompassed observable morphologies, it did not accurately reflect the evolutionary history of the fungi (Blackwell et al. 2006; Hibbett et al. 2007). Numerous molecular phylogenetic investigations did not support the monophyly of either Chytridiomycota or Zygomycota (James et al. 2000, 2006a; Tanabe et al. 2004; White et al. 2006). Fungi that were formerly classified in Chytridiomycota comprise at least four clades that represent a paraphyletic grade of zoospore-producing lineages along the base of the FTOL (Figure 2). The order Blastocladiales, traditionally characterized within Chytridiomycota, is now recognized in its own fungal phylum, Blastocladiomycota, based on the ultrastructural distinctiveness of the zoospore and phylogenetic placement inferred from the rDNA operon (James et al. 2006b). Members of Blastocladiomycota may exhibit filamentous/rhizoidal or unicellular growth and some species display a true alternation of generation with free-living diploid and haploid thalli. In the cases of filamentous growth, hyphae are mostly aseptate. Neocallimastigales, which includes the only obligate anaerobic fungi, were also traditionally classified with Chytridiomycota but have been elevated to their own phylum, Neocallimastigomycota (Hibbett et al. 2007). Of considerable interest are the zoosporic genera *Rozella* and *Olpidium*, which do not group with any of these phyla of flagellated fungi (James et al. 2006b). *Olpidium* may be the closest flagellated relative to the remaining terrestrial, non-flagellated taxa and *Rozella* may be among the first lineages to diverge from the common ancestor of all fungi.

Some of the more poorly explored early diverging lineages of the FTOL occupy terrestrial habitats and include taxa traditionally classified in Zygomycota: the subphyla Entomophthoromycotina, Mucoromycotina, Kickxellomycotina, and Zoopagomycotina and the arbuscular mycorrhizae of Glomeromycota. Of these, only Mucoromycotina and Glomeromycota are currently represented in genomic studies (Figure 2). Sampling of the remaining lineages holds keys for understanding the evolution of unique ecologies (e.g., pathogens of invertebrates) and related metabolisms of the early terrestrial fungi. Fungi are truly remarkable in the diversity of niches they exploit and their ability to survive in these unique environments is directly related to the enzymes they produce and excrete into their surrounding habitat. Conspicuous unsampled regions of the FTOL are not limited to the early diverging lineages of the FTOL, however. Numerous classes of Ascomycota (e.g., Orbiliomycetes) and Basidiomycota (e.g., Wallemiomycetes) necessary for a more comprehensive understanding of fungal metabolic diversity currently have no current genomic information. As such, sequencing exemplars from the remaining subphylum and class-level lineages will provide a census of the gene diversity of the early terrestrial fungi and identify genes with potential industrial applications.

While traditional phylogenetics has informed the sampling strategy of fungal genomics, and access to genomic sequence data has proven to be a watershed event in genome-scale phylogenetic analyses of fungi (Fitzpatrick et al. 2006; Robbertse et al. 2006; Stajich et al. 2009), the majority of the early diverging lineages lack any genome data. Sequencing unsampled early diverging lineages of the FTOL is critical to understanding early evolutionary events in the Kingdom Fungi and will likely discover a diversity of metabolic potential far greater than that characterized by currently sequenced fungal genomes.

Metagenomics of fungal ecosystems

Many of the fungi whose genomes are being sequenced are residents of soil. They have very different lifestyles and diverse metabolic capabilities that provide essential starting material to identify novel enzymes for plant biomass conversion to biofuels. Equally important, they provide baseline genomic information that enables us to explore the genomes and functions of thousands of soil fungal species we are not able to culture and sequence directly.

With the development of high-throughput sequencing technologies, it is now feasible to explore the metabolic potential of soil microbes directly, through metagenomic, or cross-genome, approaches. Such approaches bypass the need to culture members of the microbial community and focus on the collective DNA pool in the soil. Targeted metagenomic approaches provide an intensive survey of target genes that encode a phylogenetic or taxonomic

marker (e.g., rRNA gene; Jumponnen and Jones 2009) or a functional trait (Zak et al. 2006). Shotgun metagenomic approaches sequence random fragments of the total soil DNA, to obtain a profile of the total community DNA (Tringe et al. 2005). A more directed shotgun approach, termed metatranscriptomics, provides a survey of the genes currently being expressed in a soil community by sequencing fragments from the pool of soil RNA (Bailly et al. 2007).

Whole genome sequencing of model fungal species provides foundational information that enables target and shotgun metagenomic studies in two ways. First, genes from the cultured, sequenced fungi that are subsequently identified and biochemically validated to encode enzymatic functions of interest (e.g., cellulose-degrading and lignin-degrading activities) become useful reference points for target or shotgun metagenome surveys of soil nucleic acids. Second, the model fungal genomes become scaffolds for recruiting the random sequence fragments from metagenomes and metatranscriptomes. They provide essential information on genome structure, gene coding, and intergenic regions that are required to identify the random sequences to recognizable organisms (at least at some level of phylogeny) or functional categories. Through targeted or shotgun surveys, novel genes related to those known to encode functions of interest (e.g., cellobiohydrolase, chitinase, laccase genes, and others involved in plant carbon decomposition) are identified and can be studied further in cultured hosts using molecular techniques. In this way enzymes with broader or altered activities can be "mined" directly from the soil even if they are present in an uncultured species.

Ecosystem roles of soil fungi

In addition to direct mining of genomes or metagenomes of soil fungi for novel traits, fungal genomics enables studies of the roles and activities of fungi in soils (Martin and Nehls 2009), which directly impact the ability to grow biofuel crops and to understand the ecological consequences of carbon fuel use. The need to understand fungal activities in ecosystems is driven by the increasing urgency to maximize productivity of managed lands for food and fuels while maintaining sustainable ecosystems and to mitigate greenhouse gas emissions that affect climate and air quality. Such needs require greater knowledge of how fungal biomass, species composition, activities, and interactions shape ecosystem-scale processes.

In natural and agricultural ecosystems, activities of soil fungi impact ecosystem stability, productivity, and the global flux of carbon and other nutrients. Soil fungi are best known for their roles as plant associates and decomposers. The vast majority of plant species on Earth require close mycorrhizal associations with soil fungi for optimal growth and survival. In parallel, the vast majority of plant species are plagued by fungal pathogens that can cause serious and economically important diseases. Endophytic fungi are also widespread, but their contribution to plant health remains poorly understood. Soil fungi are important contributors to global carbon cycling, through decomposition of plant carbon, both in the short term as decomposers of leaf litter and fine root biomass and in the longer term as decomposers of cellulosic biomass from annual and perennial plant systems. In addition, competition for patchy and limited resources is an important feature between a vast diversity of fungi, bacteria, and microfauna in a complex soil community that may often trigger fungal capabilities not usually observed in culture.

Although plate culture and baiting/trapping techniques, sporocarp surveys, and microscopic spore counts have long been used to monitor the abundance and composition of specific fungal groups (e.g., a mycorrhizal associate, a plant pathogen), nucleic acid surveys can provide a more comprehensive, deeper-coverage assessment of the in situ fungal community and its dynamics under changing conditions. In theory, gene fragments selected as phylogenetic or functional markers can be designed to survey a specific functional assemblage or taxonomic group or the entire community. While such surveys have been widely used for soil bacteria (Tringe and Hugenholtz 2008), the genomic tools needed for implementation of these targeted surveys are just emerging for fungi (Buee et al. 2009; Jumponnen and Jones 2009) and they require information from the current whole genome sequencing efforts. Fragments of the fungal ribosomal RNA operon encompassing the internal transcribed spacer (ITS) region or the *LSU* gene are becoming standard molecular markers for fungal communities in soils but suffer from inadequate reference databases where the taxonomy is not consistent or reliable (Vilgalys 2003) and the representation of environmental species is minimal. Guided by whole genome sequence identification, continued study of the biochemistry of fungal metabolism will provide a broader suite of phylogenetic and functional markers to track fungal functions in soils. Establishment of valid databases for many fungal genes (ITS, LSU, cellobiohydrolase) is underway by several research groups (e.g., for ITS, Koljalg et al. 2005). By combining whole genome sequences of known fungi with rigorous biochemical and metabolic studies and landscape-scale experiments of soil fungal dynamics in situ, we are entering a time where linking the presence, composition, and abundance of the soil fungal community with important soil processes at an ecosystem scale is possible. Such studies promise to provide the understanding needed to sustainably manage agricultural systems for food and fuel and to predict the consequences of our activities on long-term ecosystem processes and health.

In conclusion, the genomics revolution of the last century and recent breakthroughs in new sequencing technologies have set the stage for large-scaled genomic analyses

of fungi. Different lifestyles in this kingdom – pathogens, symbionts, and saprobes – are poorly understood. Very little is known about their interactions between themselves, with hosts, and with competitors, especially in complex environments like soil. Abundance, long evolutionary history, and diversity of fungi carry enormous potential to be explored. By pursuing more complete genomic coverage of the fungi, a more accurate census of gene diversity, regulatory elements, and genome organization will be accomplished. This increase in knowledge will positively affect translational science (e.g., biological engineering) in fungal biology and ultimately the successful application of fungi to solve some of the crucial energy and climate change challenges faced by human civilization. Like the human genome project, which brought new tools for improving human health, fungal genomics has an even higher potential to positively affect the planet's energy and environmental health.

Acknowledgements

The work is conducted by the US Department of Energy Joint Genome Institute and supported by the Office of Science of the US Department of Energy under contract no. DE-AC02-05CH11231.

References

Alexopoulos CJ, Mims CW, Blackwell M. 1996. Introductory Mycology. 4th ed. New York (NY): John Wiley and Sons Inc.

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

Andersen MR, Nielsen ML, Nielsen J. 2008. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol. 4:178.

Arroyo-Lopez FN, Querol A, Bautista-Gallego J, Garrido-Fernandez A. 2008. Role of yeasts in table olive production. Int J Food Microbiol. 128:189–196.

Arvas MM, Kivioja TT, Mitchell AA, Saloheimo MM, Ussery DD, Penttila MM, Oliver SS. 2007. Comparison of protein coding gene contents of the fungal phyla Pezizomycotina and Saccharomycotina. BMC Genomics. 8:325.

Bailly J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, Wésolowski-Louvel M, Marmeisse R. 2007. Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J. 1:632–642.

Baker R. 1987. Mycoparasitism: ecology and physiology. Can J Plant Patho. 19:370–379.

Baker BJ, Lutz MA, Dawson SC, Bond PL, Banfield JF. 2004. Metabolically active eukaryotic communities in extremely acidic mine drainage. Appl Environ Microbiol. 70:6264–6271.

Baker BJ, Tyson GW, Goosherst L, Banfield JF. 2009. Insights into the diversity of eukaryotes in acid mine drainage biofilm communities. Appl Environ Microbiol. 75: 2192–2199.

Baker SE, Thykaer J, Adney WS, Brettin T, Brockman F, D'haeseleer P, Martinez AD, Miller RM, Rokhsar D, Schadt CW, et al. 2008. Fungal genome sequencing and bioenergy. Fungal Biol Rev. 22:1–5.

Baldrian P, Valášková V. 2008. Degradation of cellulose by basidiomycetous fungi. FEMS Microbiology Reviews 32(3): 501–521.

Binder JB, Raines RT. 2010. Fermentable sugars by chemical hydrolysis of biomass. Proc Natl Acad Sci USA. 107:4516–4521.

Blackwell M, Hibbett DS, Taylor JW, Spatafora JW. 2006. Research coordination networks: a phylogeny for Kingdom Fungi Deep Hypha. Mycologia. 98:829–837.

Blanchette R. 1995. Degradation of the lignocellulose complex in wood. Can J Bot. 73(Suppl. 1):999–1010.

Brunner PC, Schurch S, McDonald BA. 2007. The origin and colonization history of the barley scald pathogen *Rhynchosporium secalis*. J Evol Biol. 20:1–12.

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.

Bustamante JM 2006. Thermophilic fungi on the Sevilleta National Wildlife Refuge [dissertation] [Albuquerque (NM)]: University of New Mexico.

Butinar L, Santos S, Spencer-Martins I, Oren A, Gunde-Cimerman N. 2005b. Yeast diversity in hypersaline habitats. FEMS Microbiol Lett. 244(2):229–234.

Butinar L, Zalar P, Frisvad JC, Gunde-Cimerman N. 2005a. The genus *Eurotium* – members of indigenous fungal community in hypersaline waters of salterns. FEMS Microbiol Ecol. 512:155–166.

Buzzini P, Martini A. 2002. Extracellular enzymatic activity profiles in yeast and yeast-like strains isolated from tropical environments. J Appl Microbiol. 93:1020–1025.

Crisan EV. 1973. Current concepts of thermophilism and the thermophilic fungi. Mycologia. 655:1171–1198.

Cuomo CA, Birren BW. 2010. The fungal genome initiative and lessons learned from genome sequencing. Methods Enzymol. 470:833–855.

Dadachova E, Bryan RA, Huang X, Moadel T, Schweitzer AD, Aisen P, Nosanchuk JD, Casadevall A. 2007. Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. PLoS One. 25:e457.

de Faria MR, Wraight SP. 2007. Mycoinsecticides and mycoacaricides: a comprehensive list with worldwide coverage and international classification of formulation types. Biol Control. 43:237–256.

Dujon B. 2006. Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution. Trends Genet. 22:375–387.

Dujon B. 2010. Yeast evolutionary genomics. Nat Rev Genet. 11:512–524.

Duplessis S, Cuomo C, Lin Y-C, Aerts A, Tisserant E, Veneault-Fourrey C, Joly D, Hacquard S, Amselem J, et al. 2011. Obligate biotrophy features unraveled by the genomic analysis of the rust fungi, *Melampsora larici-populina* and *Puccinia graminis* f. sp. *tritici*. Proc Natl Acad Sci U S A. Forthcoming.

Dupree JC, Vanderwalt JP. 1983. Fermentation of dxylose to ethanol by a strain of *Candida shehatae*. Biotech Lett. 5:357–362.

Fitzpatrick DA, MaryELogue 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.

Friesen TL, Stukenbrock EH, Liu Z, Meinhardt SN, 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.

Galazka JM, Tian C, Beeson WT, Martinez B, Glass N, Cate JHD. 2010. Cellodextrin transport in yeast for improved biofuel production. Science. 330:84–86.

Gellissen G, Kunze G, Gaillardin C, Cregg JM, Berardi E, Veenhuis M, van der Klei I. 2005. New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica* – a comparison. FEMS Yeast Res. 511:1079–1096.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, et al. 1996. Life with 6000 genes. Science. 274(5287):546.

Goodwin SB, Ben M'Barek S, Dhillon B, Wittenberg AHJ, Crane CF, Van der Lee TAJ, Grimwood J, Aerts A, Antoniw J, Bailey A, et al. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity and stealth pathogenesis. PLoS Genet. Forthcoming.

Hahn-Hägerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, Gorwa-Grauslund MF. 2007a. Towards industrial pentose-fermenting yeast strains. Appl Microbiol Biotech. 74:937–953.

Hahn-Hägerdal B, Karhumaa K, Jeppsson M, Gorwa-Grauslund MF. 2007b. Metabolic engineering for pentose utilization in *Saccharomyces cerevisiae*. Biofuels. 108:147–177.

Harman GE, Howell CR, Viterbo A, Chet I, Lorito M. 2004. *Trichoderma* species – opportunistic, avirulent plant symbionts. Nat Rev Microbiol. 2:43–56.

Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Bluthgen N, Borger S, Costenoble R, Heinemann M, et al. 2008. A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. Nat Biotechnol. 26(10):1155–1160.

Hibbett DS. 2006. A phylogenetic overview of the Agaricomycotina. Mycologia 98(6):917–925.

Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, et al. 2007. A higher-level phylogenetic classification of the fungi. Myco Res. 111:509–547.

Hibbett DS, Matheny PB. 2009. The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. BMC Biol. 7:13.

Himmel ME, Adney WS, Baker JO, Elander R, McMillan JD, Nieves RA, Sheehan JJ, Thomas SR, Vinzant TB, Zhang M. 1997. Advanced bioethanol production technologies: a perspective. In: Saha BC, Woodward J. Fuels and Chemicals from Biomass. vol. 666. p. 2–45.

Hofrichter M, Ullrich R, Pecyna MJ, Liers C, Lundell T. 2010. New and classic families of secreted fungal heme peroxidases. Appl Microbiol Biotechnol. 87:871–897.

Hoshino T, Kiriaki M, Ohgiya S, Fujiwara M, Kondo H, Nishimiya Y, Yumoto I, Tsuda S. 2003. Antifreeze proteins from snow mold fungi. Can J Bot. 81:1175–1181.

Howell CR. 2003. Mechanisms employed by *Trichoderma* species in the biological control of plant diseases: the history and evolution of current concepts. *Plant Dis*. 87:4–10.

Hubalek Z. 1976. Interspecific affinity among keratinolytic fungi associated with birds. Folia Parasitol (Praha). 23:267–272.

Ingledew W. 1990. Acidophiles. In: Edwards C. Microbiology of extreme environments. Milton Keynes (UK): Open University Press. p. 33–54.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, et al. 2006a. Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature. 443:818–822.

James TY, Letcher PM, Longcore JE, Mozley-Standridge 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.

James TY, Porter D, Leander CA, Vilgalys R, Longcore JE. 2000. Molecular phylogenetics of the Chytridiomycota support the utility of ultrastructural data in chytrid systematics. Can J Bot. 78:336–350.

Jeffries TW, Grigoriev IV, Grimwood J, Laplaza JM, Aerts A, Salamov A, Schmutz J, Lindquist E, Dehal P, Shapiro H, et al. 2007. Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. Nat Biotech. 25:319–326.

Jeffries TW, Jin YS. 2004. Metabolic engineering for improved fermentation of pentoses by yeasts. Appl Microbiol Biotech. 63:495–509.

Jeffries TW, Van Vleet JRH. 2009. *Pichia stipitis* genomics, transcriptomics, and gene clusters. FEMS Yeast Res. 9:793–807.

Jin Y, Weining S, Nevo E. 2005. A *MAPK* gene from Dead Sea fungus confers stress tolerance to lithium salt and freezing–thawing: prospects for saline agriculture. Proc Natl Acad Sci U S A. 102(52):18992–18997.

Jumponnen A, Jones KL. 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. New Phytol. 184:438–448.

Kato S, Shimizu-Ibuka A, Mura K, Takeuchi A, Tokue C, Arai S. 2007. Molecular cloning and characterization of an alpha-amylase from *Pichia burtonii* 15–1. Biosci Biotech Biochem. 71:3007–3013.

Koljalg U, Larsson KH, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, et al. 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytol. 166:1063–1068.

Kroll RG. 1990. Alkalophiles. In: Edwards C. Microbiology of extreme environments. Milton Keynes (UK): Open University Press. p. 55–92.

Kubicek CP, Herrera-Estrella A, Seidl V, Le Crom S, Martinez DA, Druzhinina IS, Zeilinger S, Casas-Flores S, Horwitz BA, Mukherjee PK, et al. 2011. Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. Genome Biol. 12(4):R40.

Kurtzman CP, Dien BS. 1998. *Candida arabinofermentans*, a new l-arabinose fermenting yeast. Antonie Van Leeuwenhoek Intl J Gen Mol Microbiol. 74:237–243.

Kurtzman CP, Suzuki M. 2010. Phylogenetic analysis of ascomycete yeasts that form coenzyme Q-9 and the proposal of the new genera *Babjeviella, Meyerozyma, Millerozyma, Priceomyces,* and *Scheffersomyces*. Mycoscience. 51:2–14.

Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, et al. 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(38):16151–16156.

Lopez-Archilla AI, González AE, Terrón MC, Amils R. 2004. Ecological study of the fungal populations of the acidic Tinto River in southwestern Spain. Can J Microbiol. 50(11):923–934.

Lundell TK, Makela MR, Hilden K. 2010. Lignin-modifying enzymes in filamentous basidiomycetes–ecological, functional and phylogenetic review. Journal of basic microbiology 50(1):5–20.

Magan N. 2007. Fungi in extreme environments. The Mycota. 4:85–103.

Maheshwari R, Bharadwaj G, Bhat MK. 2000. Thermophilic fungi: their physiology and enzymes. Microbiol Mol Biol Rev. 643:461–488.

Mankau R. 1980. Biocontrol: fungi as nematode control agents. J Nematol. 12:244–252.

Martin F, Aerts A, Ahrén D, Brun A, Duchaussoy F, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, et al. 2008. The genome sequence of the basidiomycete fungus *Laccaria bicolor* provides insights into the mycorrhizal symbiosis. Nature. 452:88–92.

Martin F, Cullen D, Hibbett D, Pisabarro A, Spatafora JW, Baker SE, Grigoriev IV. 2011. Sequencing the fungal tree of life. New Phytol. 190(4):818–821.

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R, et al. 2010. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature. 464:1033–1038.

Martin F, Nehls U. 2009. Harnessing ectomycorrhizal genomics for ecological insights. Curr Opin Plant Biol. 12:508–515.

Martin F, Selosse MA. 2008. The *Laccaria* genome: a symbiont blueprint decoded. New Phytol. 180:379–390.

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, et al. 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, et al. 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA. 106(6):1954–1959.

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, et al. 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol. 22:695–700.

Matheny PB, Wang Z, Binder M, Curtis JM, Lim YW, Nilsson RH, Hughes KW, Hofstetter V, Ammirati JF, Schoch CL, et al. 2007. Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). Mol Phylogenet Evol 43(2): 430–451.

Morrison TA, Jung HG, Buxton DR, Hatfield RD. 1998. Cell-wall composition of maize internodes of varying maturity. Crop Sci. 38:455–460.

Nguyen NH, Suh SO, Marshall CJ, Blackwell M. 2006. Morphological and ecological similarities: wood-boring beetles associated with novel xylose-fermenting yeasts, gen. sp nov and *Candida jeffriesii* sp nov. Myco Res. 110:1232–1241.

Ohm RA, de Jong JF, Lugones LG, Aerts A, Kothe E, Stajich JE, de Vries RP, Record E, Levasseur A, Baker SE, et al. 2010. Genome sequence of the model mushroom *Schizophyllum commune*. Nat Biotechnol. 28:957–963.

Okada G, Niimura Y, Sakata T, Uchimura T, Ohara N, Suzuki H, Kozaki M. 1993. *Acremonium alcalophilum*, a new alkalophilic cellulolytic hyphomycete. Trans Myco Soc Japan. 34:171–185.

Ostry ME, Wilson LF, McNabb HS 1989. Impact and control of *Septoria musiva* on hybrid poplars. St. Paul, MN: USDA, Forest Service, North Central Forest Experiment Station. Gen Tech Rep. NC-133.

Robbertse B, Reeves J, Schoch CL, Spatafora JW. 2006. A phylogenomic analysis of the Ascomycota. Fungal Gen Biol. 43:715–725.

Saha BC. 2003. Hemicellulose bioconversion. J Ind Microbiol Biotechnol. 30:279–291.

Schneider H, Wang PY, Chan YK, Maleszka R. 1981. Conversion of D-xylose into ethanol by the yeast *Pachysolen tannophilus*. Biotech Lett. 3:89–92.

Schneider-Belhaddad F, Kolattukudy P. 2000. Solubilization, partial purification, and characterization of a fatty aldehyde decarbonylase from a higher plant, *Pisum sativum*. Arch Biochem Biophys. 377(2):341–349.

Sharon E, Bar-Eyal M, Chet I, Herrera-Estrella A, Kleifeld O, Spiegel Y. 2001. Biological control of the root-knot nematode *Meloidogyne javanica* by *Trichoderma harzianum*. Phytopathology. 91:687–693.

Shin KS, Shin YK, Yoon JH, Park YH. 2001. *Candida thermophila* sp. nov, a novel thermophilic yeast isolated from soil. Int J Syst Evol Microbiol. 51(Pt 6):2167–2170.

Simard RE, Cameron A. 1974. Fermentation of spent sulfite liquor by *Candida utilis*. Pulp and Paper-Canada. 75:T117–T120.

Slininger PJ, Bothast RJ, Vancauwenberge JE, Kurtzman CP. 1982. Conversion of d-xylose to ethanol by the yeast, *Pachysolen tannophilus*. Biotech Bioeng. 24:371–384.

Smith DR, Quinlan AR, Peckham HE, Makowsky K, Tao W, Woolf B, Shen L, Donahue WF, Tusneem N, Stromberg MP, et al. 2008. Rapid whole-genome mutational profiling using next-generation sequencing technologies. Genome Res. 18:1638–1642.

Smith SE, Read DJ. 2008. Mycorrhizal Symbiosis. 3rd ed. London (UK): Academic Press.

St. Leger RJ, Wang C. 2009. Genetic engineering of fungal biocontrol agents to achieve greater efficacy against insect pests. Appl Microbiol Biotechnol. 85:901–907.

Stajich JE, Berbee ML, Blackwell M, Hibbett DS, James TY, Spatafora JW, Taylor JW. 2009. Primer: the fungi. Curr Biol. 19:R840–R845.

Stergiopoulos I, van den Burg HA, Okmen B, Beenen HG, van Liere S, Kema GHJ, de Wit PJGM. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. Proc Nat Acad Sci USA. 107:7610–7615.

Straatsma G, Samson RA, Olijnsma TW, Op Den Camp HJ, Gerrits JP, Van Griensven LJ. 1994. Ecology of thermophilic fungi in mushroom compost, with emphasis on *Scytalidium thermophilum* and growth stimulation of *Agaricus bisporus* mycelium. Appl Environ Microbiol. 602:454–458.

Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA. 2007. Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. Mol Biol Evol. 24:398–411.

Sunesson A, Vaes W, Nilsson C, Blomquist G, Andreson B, Carlson R. 1995. Identification of volatile metabolites from five fungal species cultivated on two media. Appl Environ Microbiol. 61:2911–2918.

Tanabe Y, Saikawa M, Watanabe MW, Sugiyama J. 2004. Molecular phylogeny of Zygomycota based on EF-1 alpha and RPB1 sequences: limitations and utility of alternative markers to rDNA. Mol Phylogenet Evol. 30: 438–449.

Taylor JW, Berbee ML. 2006. Dating divergences in the Fungal Tree of Life: review and new analyses. Mycologia 98(6): 838–849.

Teshima KM, Innan H. 2008. Neofunctionalization of duplicated genes under the pressure of gene conversion. Genetics. 178:1385–1398.

Tringe SG, Hugenholtz P. 2008. A renaissance for the pioneering 16S rRNA gene. Curr Opin Microbiol. 11:442–446.

Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC. 2005. Comparative metagenomics of microbial communities. Science. 308:554–556.

Turgeon BG, Baker SE. 2007. Genetic and genomic dissection of the *Cochliobolus heterostrophus* Tox1 locus controlling biosynthesis of the polyketide virulence factor T-toxin. Adv Genet. 57:219–261.

Ullstrup AJ. 1972. The impacts of the southern corn leaf blight epidemics of 1970–1971. Annu Rev Phytopathol. 10: 37–50.

Vicente G, Bautista LF, Gutierrez FJ, Rodriguez R, Martinez V, Rodriguez-Frometa RA, Ruiz-Vázquez RM, Torres-Martinez S, Garre V. 2010. Direct transformation of fungal biomass from submerged cultures into biodiesel. Energy Fuels. 24(5):3173–3178.

Vicente G, Bautista LF, Rodríguez R, Gutiérrez FJ, Sádaba I, Ruiz-Vázquez RM, Torres-Martínez S, Garre V. 2009. Biodiesel production from biomass of an oleaginous fungus. Biochem Eng J. 481:22–27.

Vilgalys R. 2003. Taxonomic misidentification in public DNA databases. New Phytol. 160:4–5.

Walker JD, Cooney JJ. 1973. Aliphatic hydrocarbons of Cladosporium resinae cultured on glucose, glutamic acid, and hydrocarbons. Appl Microbiol. 26(5):705–708.

Wang B, Qiu YL. 2006. Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza. 16:299–363.

White MM, James TY, O'Donnell K, Tanabe Y, Sugiyama J. 2006. Phylogeny of the Zygomycota based on nuclear ribosomal sequence data. Mycologia. 98:872–884.

Wild R, Patil S, Popovic M, Zappi M, Dufreche S, Bajpai R. 2010. Lipids from *Lipomyces starkeyi*. Food Technol Biotech. 48:329–335.

Wittenberg AHJ, van der Lee TAJ, Ben M'Barek S, Ware SB, Goodwin SB, Kilian A, Visser RGF, Kema GHJ, Schouten HJ. 2009. Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen *Mycosphaerella graminicola*. PLoS ONE. 4:e5863.

Woods A, Coates KD, Hamann A. 2006. Is an unprecedented epidemic of *Dothistroma* needle blight related to climate change? BioScience. 55:761–769.

Worrall JJ, Anagnostakis SE, Zabel RA. 1997. Comparison of wood decay among diverse lignicolous fungi. Mycologia 89:199–219.

Zak DR, Blackwood CB, Waldrop MP. 2006. A molecular dawn for biogeochemistry. Trends Ecol Evol. 21:288–295.

Zalar P, Sybren de Hoog G, Schroers HJ, Frank JM, Gunde-Cimerman N. 2005. Taxonomy and phylogeny of the xerophilic genus *Wallemia* (Wallemiomycetes and Wallemiales). Antonie Leeuwenhoek. 874:311–328.

Zhao X, Kong XL, Hua YY, Feng B, Zhao ZB. 2008. Medium optimization for lipid production through co-fermentation of glucose and xylose by the oleaginous yeast *Lipomyces starkeyi*. Eur J Lipid Sci Tech. 110:405–412.

Zhdanova NN, Tugay T, Dighton J, Zheltonozhsky V, McDermott P. 2004. Ionizing radiation attracts soil fungi. Mycol Res. 108(Pt 9):1089–1096.

Zopf W. 1888. Zur Kenntnis der Infektionskrankheiten niederer Thiere und Pflanzen. Nova Acta Leopold Carol. 52: 314–376.