Chapter 3
Future Perspectives and Challenges of Fungal Systematics in the Age of Big Data

Zheng Wang, R. Henrik Nilsson, Timothy Y. James, Yucheng Dai, and Jeffrey P. Townsend

Fungal Diversity and Systematics

The beginning of wisdom is to call things by their proper name (Chinese proverb: 名正言顺)

Z. Wang, Ph.D. (✉)
Department of Biostatistics, Yale School of Public Health, Yale University,
135 College Street, New Haven, CT 06510, USA
e-mail: wangzhengff@gmail.com; wang.zheng@yale.edu

R.H. Nilsson, Ph.D.
Department of Biological and Environmental Sciences, University of Gothenburg,
Box 461, Göteborg 405 30, Sweden
e-mail: henrik.nilsson@bioenv.gu.se

T.Y. James, Ph.D.
Department of Ecology and Evolutionary Biology, University of Michigan,
1147 Kraus Natural Science Building, 830 North University, Ann Arbor,
MI 48109-1048, USA
e-mail: tyjames@umich.edu

Y. Dai, Ph.D.
Institute of Microbiology, Beijing Forestry University, Beijing 100083, China
e-mail: yuchengd@yahoo.com

J.P. Townsend, Ph.D. (✉)
Department of Biostatistics, Yale School of Public Health, Yale University,
135 College Street, New Haven, CT 06510, USA

Department of Ecology and Evolutionary Biology, Yale University,
165 Prospect Street, New Haven, CT 06520, USA

Program in Computational Biology and Bioinformatics, Yale University,
165 Prospect Street, New Haven, CT 06520, USA
e-mail: jeffrey.townsend@yale.edu

The mission and agenda of fungal systematists are to discover, describe, and inventory the global species diversity of one of the most diverse groups on earth. The circumscription of the fungi has evolved over time. Fungi are most closely related to animals and share a more recent common ancestor with them than with all other major groups of eukaryotes. The majority of the fungal kingdom is composed of heterotrophic, non-photosynthetic eukaryotes with cell walls containing chitin and β-glucans and, when present, a single flagellum (Stajich et al. 2009). Fungi can occur both as single-celled and multicelled organisms and can reach sizes typically associated with plants and animals. For example, the largest single fungal fruiting body on record was found to be nearly 500 kg in weight (Dai and Cui 2011), and the oldest and largest mycelium described covers 15 ha of area and is over 1500 years old (Smith et al. 1992). Life cycles of many fungi include a vegetative growth phase that spreads throughout its environment by extension of hyphae and/or release of a large number of asexual spores from simple structures and by a more complex, transient sexual phase producing smaller numbers of resistant sexual spores from well-developed fruiting bodies. Fungal diversity is estimated to comprise 1.5–7.1 million species. An increasing number of new taxa continue to be reported worldwide (Blackwell 2011; Bass and Richards 2011), and fungi have been isolated from almost all kinds of ecosystems on Earth (Stajich et al. 2009). This fungal diversity is described by systematics, which is the science not only of naming fungi but also of positioning the species among other existing names to represent their evolutionary relationships. To properly describe the substantial diversity of the Kingdom Fungi, mycologists have been updating its classification and systematics, based on accumulated knowledge of fungal biology interpreted within new concepts and approaches that are emerging from evolutionary biology.

In contrast to large aboveground organisms that can be easily spotted and counted, fungi are major components of underground diversity. Their study is often made difficult by their microscopic structures and shortage of discriminatory morphological characters. Traditional biological information used for classifying fungi into major groups includes morphology, ultrastructure, physiology, tissue biochemistry, and ecological traits. Early synthesis of this information yielded major fungal groups that have remained comparatively stable over a very long time period in the twentieth century. Some morphological and ecological traits, such as the structure of the cell wall and hyphal septa, sexual reproduction and meiotic spores, nutritional modes, as well as geographic distribution, have proven to be relatively conserved and informative, especially for high-level classification. However, phenotypic plasticity of traits and fast-evolving traits have caused considerable uncertainty regarding lower-level phylogenies based on morphology and ecology (Lutzoni et al. 2004). Starting in the 1970s, but gaining momentum in the late 1990s, the use of DNA sequence data to infer phylogenetic relationships among fungal lineages brought about a revolution in terms of taxonomic resolution and scientific reproducibility (de Bertoldi et al. 1973; Bruns et al. 1991; Bridge et al. 2005; Blackwell et al. 2006). Initial molecular studies, typically based on a single gene region, were followed by a wave of multilocus phylogenetic studies including all major fungal groups. The new phylogenies facilitated several major taxonomic revisions,

including new lineages at the phylum and class level (James et al. 2006a; Hibbett et al. 2007; Kirk et al. 2008; Schoch et al. 2009a, b; Rosling et al. 2011; James and Berbee 2012; Matheny et al. 2007). More changes and many new taxa were added to lower-level fungal groups. In addition, much novel diversity was revealed in sequence data collected from environmental samples and identified as operational taxonomic units (OTUs) (Blaxter et al. 2005). The quantity of novel OTUs in most environmental samples hints at a massive, inconspicuous, undescribed, and thriving fungal diversity (Hibbett et al. 2011). Classifying and naming this huge fungal diversity is a necessary step toward understanding the functions of these fungi in the ecosystems. Thus, finding ways to take full advantage of the power afforded by next-generation sequencing approaches to integrate environmental DNA sequences has become one of the major challenges for fungal systematics. Simultaneously, a complementary aspect of the future of the fungal systematics is the integration of systematics, the evolution of complex traits, and functional genomics to understand the comparative biology of fungi and to create a holistic view of the fungus and how it evolves.

Currently, efficient communication regarding fungal species rests upon on the use of scientific names constructed based on a system of hierarchical ranks. Within this system, one of the major purposes of fungal taxonomy and systematics is to create and position nomenclatural units unique for each fungal species. While a stable name as a symbol for communication is always appreciated by researchers—especially for the widely used industrial, medical, plant pathogenic, and model species—systematics must also continue to refine and revise the application of names to reflect continual gains in knowledge about the evolutionary histories of all taxa. We make no attempt here to cite all papers on development of fungal taxonomy and systematics nor to summarize recent systematic progress within and among the major fungal phyla. Instead, we have chosen to highlight recent research that enables us to illustrate specific points about perspectives and challenges of fungal systematics in the age of big data.

Integrative Taxonomy and Current Fungal Systematics

Traditionally, morphological and sometimes ecological traits have been used to classify fungi into hierarchical ranks and groups. However, evolutionary relationships derived from these traits, whose ontology is often inferred from a phylogenetic hypothesis, can be problematic, especially for lower-level phylogeny, where diverse fungal groups can have plesiomorphic or convergent morphologies. One problem in reconstructing fungal evolutionary history is a lack of paleontological information due to the scarcity of well-preserved fungal fossils (Bidartondo et al. 2011). This scarcity makes it extremely difficult to evaluate the evolutionary history of morphological traits for fungal systematics, especially for morphologically simple groups. The meager fossil record also makes it difficult to precisely calibrate molecular phylogenies. Nevertheless, information on molecular evolutionary

events, such as mutation and gain or loss of nucleotide characters, has been well preserved in gene sequences. Molecular phylogenies using single genetic markers or multilocus data have led to dramatic advances in the systematics of a range of taxonomic levels within the fungal kingdom over the past three decades. However, systematic hypotheses based on molecular phylogenetic data alone can be questioned, especially when in conflict with morphological evidence. Ideally, evidence from different lines, such as morphology, ecology, and molecular data, can be evaluated jointly to robustly define taxa at all ranks. This approach has been called integrative taxonomy and has been advocated by Will et al. (2005) and Pante et al. (2015).

A major driver of new advances in the molecular phylogeny of fungi was the Assembling the Fungal Tree of Life (AFTOL) project, funded by the National Science Foundation (NSF) of the United States and organized by mycologists at several leading laboratories. This project sprung out of an NSF-funded research coordination network known as Deep Hypha and culminated in significant gains in the study of evolution and molecular phylogenetics of fungi (Lutzoni et al. 2004; Blackwell et al. 2006; James et al. 2006a; Hibbett et al. 2007; Schoch et al. 2009a). Among the very first multilocus phylogenies targeting the majority of major fungal lineages, Lutzoni et al. (2004) highlighted two major challenges in fungal systematics in the molecular age. One is achieving a balanced sampling of taxa and genetic markers. The other is identifying and interpreting inconsistency between the evolution of morphology and molecular phylogeny. When standard PCR using degenerate primers and Sanger sequencing were the major tools for recovering DNA sequences from fungal tissue, loci such as nuclear and mitochondrial rRNAs and several widely used protein-coding genes, including subunits of elongation factors and RNA polymerases, were selected by the AFTOL project. A six-gene phylogeny using these markers, including data from 52 sequenced fungal genomes, was generated to assess early evolution of fungi, and ecological characters were mapped onto the tree.

Groups recognized in the six-gene phylogeny were generally consistent with traditional views of fungal systematics prior to the molecular systematic age, but only for the fungi in *Dikarya* (James et al. 2006a). Non-monophyly of two of the six recognized phyla led to the abandonment of one (*Zygomycota*) and the description of two new phyla *Blastocladiomycota*, by James et al. (2006b), and *Neocallimastigomycota*, by Hibbett et al. (2007). Similar sequence datasets, which were often incomplete with missing sequences, were generated for a more inclusive taxon sampling within each major fungal group at class level, and the resulting phylogenetic classifications were collected in the special Deep Hypha issue of Mycologia in 2006 (Blackwell et al. 2006). A comprehensive phylogenetic classification of the fungi kingdom was later proposed by Hibbett et al. (2007), featuring 16 new taxa above the level of order. This classification was adopted by the latest version of the Dictionary of the Fungi (Kirk et al. 2008). A more taxonomically complete six-gene dataset for 420 ascomycetes was subsequently assembled and analyzed. Key morphological and ecological characters were evaluated for usefulness in ascomycete systematics, and a new class was differentiated for two

earthtongue genera: *Geoglossum* and *Trichoglossum* (Schoch et al. 2009a, b). This dataset made it possible to quantify phylogenetic informativeness (Townsend et al. 2008; Townsend and Lopez-Giraldez 2010; Lopez-Giraldez and Townsend 2011) for several widely used genetic markers (Schoch et al. 2009a, b). With the release of more fungal genome sequences and the ever-growing availability of data from additional genetic markers, several multilocus phylogenies inferred using partially or solely from genomic data (phylogenomics) have been published (Ebersberger et al. 2012; Binder et al. 2013; Ortiz-Santana et al. 2013; Dutilh et al. 2007). Updated classifications for major fungal groups were collected in Mycota VII—Systematics and Evolution (McLaughlin et al. 2014).

The vast majority of molecular systematic studies of fungi have been based on annotated (voucher) specimens of primarily sexual (teleomorphic) but also asexual (anamorphic) collections. The accuracy of voucher specimens is particularly important now, because in many modern studies, only molecular data are shared and examined: fungal herbaria thus play important roles in keeping records for well-annotated specimens (Bidartondo 2008; Schoch et al. 2014). Best-practice guidelines on how to appropriately use molecular data in mycology are readily available (Lindahl et al. 2013; Hyde et al. 2013; Nilsson et al. 2012). Nevertheless, these guidelines are not sufficiently frequently adhered to fungal molecular phylogeneticists. Well-preserved and annotated collections are now mandatorily required by journals for newly published morphological and molecular data (Seifert and Rossman 2010). However, there has been no guarantee of accurate identification of fungal collections, especially for microfungi, partially due to the problematic outcomes of applying species concepts in fungi.

Morphological, biological, or phylogenetic species concepts all have limitations when they are applied to fungal species (Taylor et al. 2000, 2006). In particular, different mycologists often have different quantitative or qualitative interpretations of data used to define species boundaries. For example, using several genetic markers, multiple species were identified within the single morphological and biological species commonly known as the "turkey tail" fungus *Trametes versicolor* (Carlson et al. 2014), and two species were recognized for North American *Heterobasidion annosum*, which has been considered one of the most important forest pathogens in the world (Otrosina and Garbelotto 2010). Another extraordinary and exciting example would be that of the morel fungi, for which tens, if not hundreds, of new species have been recognized within several original common names (Du et al. 2012; Richard et al. 2014). An increasing number of low-level classifications are based on integrative approaches using both morphological and molecular data. These approaches have been applied to solve identification issues of several commercially important fungi (Cao et al. 2012; Wu et al. 2014; Zhang et al. 2005).

In many cases, the reference molecular data are directly downloaded from various databases, assuming accurate identification without checking the resource specimens. Cryptic species complexes are particularly likely for many species of microfungi, in which case, dense samples from accurately annotated specimens will be especially critical for proper species taxonomy. However, phylogenetic recognition of fungal species has proved to be reliable, reproducible, and increasingly

widely applicable, facilitating convenient naming of species or strains, especially for microfungi. The huge undisclosed fungal diversity and the difficulty of reconciling species concepts in fungi can make the application of the International Code of Nomenclature (McNeill and Turland 2012) very challenging—to the extent that it can ironically slow down, rather than speed up, mycological progress. Recently, for instance, instead of following the code to use the teleomorph genus name for monophyletic groups, mycologists advocated recognizing the genus *Fusarium* as the sole name for groups that have been studied under that name but are not monophyletic (Geiser et al. 2013). Such challenges will become more significant as more invisible diversity is discovered within diverse environmental samples. These challenges should aid the community in pushing for the development of standards for sequence-based classification (Hibbett and Taylor 2013). A recent review of the impacts of the nomenclatural code on the scientific names that have been adopted is available for plant pathogenic fungi (Zhang et al. 2013).

Systematics and Classification for Invisible Diversity

Fungi are widely distributed in all terrestrial and aquatic ecosystems. About 100,000 fungal species have been discovered and documented. They play critical roles in inorganic and organic nutrition, nutrient cycling, and especially in the decay of carbon compounds that were fixed and integrated into complex compounds by plants. Furthermore, fungi are frequently intimate partners in coevolving biotic and trophic relationships with other organisms, notably through mycorrhizal associations with plants; almost all land plants form symbiotic associations with mycorrhizal fungi (Stajich et al. 2009; van der Heijden et al. 2015). However, only a small portion of the total fungal diversity has been documented based on specimens/strains deposited in herbaria, culture collections, or in personal collections all over the world. Indeed, a modest ~1000 new species are described per year (Hibbett et al. 2011), which would require 5000 years of cataloging at this rate, should the 5.1 million estimate of species diversity hold.

The challenges to description of this undescribed fungal diversity are threefold. First, there are few mycological researchers and little research to study this undescribed diversity. Second, many of these undescribed species whose morphology can be characterized are actually cryptic species hidden within species previously described on the basis of morphological characters; morphological characters might not separate the genetic species, as discussed for *Trametes versicolor* and *Morchella* spp. above. Third, the majority of the extant fungal diversity produces no distinguishing morphological structures that are visible or describable, e.g., these fungi carry out their lives mostly or entirely as unculturable and morphologically indistinguishable yeasts or vegetative hyphae that cannot be described formally. If these fungi are unculturable as well as morphologically and biochemically indistinguishable, only can molecular identification be used as a tool to classify this potentially huge diversity. This kind of molecular-only identification leads to the absurd

situation where next-generation sequencing efforts of environmental substrates reveal the existence of thousands upon thousands of new species of very high relevance to phylogenetic and ecological characterization of the fungal kingdom—and yet this huge diversity of species cannot be described. This inability to describe these species effectively excludes them from further scientific scrutiny. Such sequences are typically submitted to sequence databases labeled as "uncultured fungus," making unambiguous reference to those species across datasets and studies problematic at best. This lack of linkage across studies in turn makes it difficult to assemble data for these species; what countries, hosts, and substrates these individual, unnamed species are known from cannot easily be compiled. In turn, this lack of synthetic inferential power further complicates the eventual formal description of these species.

The UNITE database for molecular identification of fungi recently presented a solution to this problem (Kõljalg et al. 2013). All fungal ITS sequences are clustered to approximately the species level based on sequence similarity, and each such OTU—called a *species hypothesis*—is assigned a unique, stable name of the accession number. Thus, regardless of whether the OTU has a formal Latin name or not, unambiguous reference across publications—as well as data assembly for individual species—is possible and even automated. A recent study, based on 365 soil samples collected from across the globe, identified 80,486 fungal OTUs and used the UNITE species hypothesis system to analyze them. Although a very modest 4353 of the OTUs could be linked to highly similar reference sequences from herbarium specimens or described culture collections, the underlying sequences of the full results of the study are now integrated in UNITE for standardized reference (Tedersoo et al. 2014; Wardle and Lindahl 2014). At the time of this writing, GenBank has a collection of more than 600,000 fungal sequences from environmental samples, chiefly the nuclear ribosomal internal transcribed spacer (ITS) region. Among these, there are about 200,000 that have been identified as stemming from an "uncultured fungus," without an affiliation to any existing ranks.

It is hard to estimate how inclusion of this huge invisible diversity would affect the fungal systematics that so far encompassed only just over 100,000 accepted fungal species. Despite the challenges, it is clear that not including these extant but unnamed species in molecular studies of fungi and fungal communities is detrimental to mycology. Nilsson et al. (2011) examined the topological effects of including such environmental sequences in phylogenetic analyses that featured only sequences from vouchered fruiting bodies and cultures. Their inclusion made a significant difference to the inferred topology and to the support of internodes. Similarly, the relatively recent realization that aquatic ecosystems abound with uncharted fungal diversity, particularly in the *Chytridiomycota* and *Cryptomycota*, could provide taxonomic sampling that might provide resolution of this part of the fungal tree of life, which has been plagued by low resolution and poor branch support (Wurzbacher and Grossart 2012; Ishii et al. 2015). Recently a whole new class, *Archaeorhizomycetes*, comprising hundreds of cryptically reproducing culturable filamentous fungi of poorly understood ecology, has been discovered from soil samples (Rosling et al. 2011). Using multilocus analyses, they have been phylogenetically placed into the

species-poor group *Taphrinomycotina* of the *Ascomycota*. The recognition of the *Archaeorhizomycetes* represents a major step forward in our understanding of soil fungi, as these fungi seem to be common in soil samples throughout the world (Porter et al. 2008; Rosling et al. 2013). At an even higher rank, the new fungal phylum *Cryptomycota*, rich particularly in aquatic environments, is also known almost exclusively from environmental DNAs (James and Berbee 2012; Jones et al. 2011). The systematics of the *Archaeorhizomycetes* and *Cryptomycota* will remain hindered by the absence of complete genome sequences, which will be challenging to obtain from these minute fungi. On the other hand, recent advances in obtaining near-complete genome sequences from single cells hold promise for both placing uncultured fungal lineages on the tree of life and for inferring their ecological roles (Rinke et al. 2013).

For the majority of fungal lineages, ITS sequences provide a powerful and efficient means of identification. Therefore, the ITS has been proposed and accepted as a universal DNA barcode marker for fungi (Schoch et al. 2012). A DNA barcode, however, is nothing more than a sequence that can be unambiguously linked to a taxonomic label for a species. DNA barcodes do not promise a solution for nomenclatural classification of diversity. Such a solution might arise from digital codes such as PhyloCode (de Queiroz and Gauthier 1994). However, this concept still lacks a standardized real-life implementation (de Queiroz and Gauthier 1994; De Oliveira Martins et al. 2014; Money 2013). While ITS is generally considered as only informative for species recognition and low-lever phylogenetic analysis, classification of the environmental diversity typically relies on observations of high sequence similarity to reference sequences from annotated specimens (Schoch et al. 2014). However, with the use of new tools to address some serious alignment issues regarding the ITS region (Liu et al. 2009, 2012), ITS alignments have shown promise in use for intermediate-level phylogeny (Koetschan et al. 2010), providing comparable classification accuracy to some other frequently used gene markers, such as the large subunit of rRNA sequence (Wang et al. 2011). Including proper reference sequences would provide insights into evolutionary history and ecology for these so-called invisible fungi (Wang et al. 2011; Porras-Alfaro et al. 2014; Del Olmo-Ruiz and Arnold 2014). Automatic phylogenetic approaches, such as those implemented in MOR (Hibbett et al. 2005) and WASABI (Kauff et al. 2007) would be able to efficiently filter and classify environmental sequence data. Still, there might be many environmental species that have no comparable characterized lineages, such that they cannot be morphologically defined or easily systematically positioned. Moreover, the absence of barcodes of the ITS region associated with this phylum is also an impediment, as many barcodes that cannot be assigned to a phylum may belong to these poorly sampled basal lineages, which exist in databases primarily as 18S rDNA sequences. To incorporate these taxa into fungal systematics requires developing methods for gathering informative sequence data that link barcodes to darker regions of the fungal phylogeny and performing efficient phylogenetic analysis on large datasets.

Given the deep divergence of the major fungal lineages, plodding through taxa using PCR with degenerate primers to fish for loci is a challenging, if not impossible,

approach toward recovering an effective diversity of protein-coding genes that will prove informative for deep phylogeny. Moreover, establishing linkages among multiple independent genes that derive from the same OTU defined from environmental DNA is nearly impossible at present. Thus, with the development of single-cell genome sequencing, phylogenomic approaches might provide an alternative and more powerful means to reconstruct a systematics of both the visible and the invisible fungal diversity.

Fungal Genomes, Phylogenomics, and Phylotranscriptomics

The very first sequenced fungal genome was also the first sequenced eukaryotic genome: that of the wine yeast *Saccharomyces cerevisiae*, an important genetic model and an industrial workhorse. This comparatively small genome was published in 1996 (Goffeau et al. 1996). Since then, following the technical progress in genome sequencing, fungal genomes have been released at an ever-accelerating rate. The number of available fungal genome sequences has increased by another order of magnitude (Galagan et al. 2005). In GenBank (http://www.ncbi.nlm.nih.gov) alone, there are currently fungal genomes representing 451 species. The recently launched 1000 Fungal Genomes (1KFG) project (http://1000.fungalgenomes.org) plans to sequence representatives from more than 650 recognized families of fungi (Kirk et al. 2008; Hibbett et al. 2013). The released genomes facilitate assembly of closely related genomes against the reference genomes even in small laboratories, and the sampled genomes of closely related organisms are designed to enable comparative studies. Comparative genomics of closely related organisms can provide a powerful approach to ascertain the genetic basis of diverse phenotypes, such as fungi-host associations, secondary metabolic pathways, morphological development, and fungal responses to environmental signals (Galagan et al. 2005; Hibbett et al. 2013; Sikhakolli et al. 2012; Andersen et al. 2011; Lehr et al. 2014; Nishant et al. 2010; Rodriguez-Romero et al. 2010; Heitman 2007). Many comparative genomic studies focus on the biology and evolution of model fungi to make inferences about basic biological processes in all eukaryotes. Studies that analyze genomes in the context of their phylogenetic and evolutionary relationships are accelerating research into the fundamental aspects of eukaryotic biology. As stated in Delsuc et al. (2005) "…nothing in genomics makes sense except in the light of evolution."—large numbers of genomes alone do not provide much insight into organismal biology, however. Many features of genomes need to be related to organismal knowledge and understood in the context of their evolutionary history.

How can these fungal genomes empower fungal systematic research? The genome itself comprises all informative genetic markers that could be sampled for any individual. Access to this scale of genomic data for phylogenetic purposes could potentially alleviate previous and present problems of phylogenetics that arise from insufficient or biased sampling of genetic markers. With this massive increase of potentially useful characters, the focus of phylogenetic inference must shift

toward development of new methodologies that can efficiently, accurately, and reliably handle big data and toward approaches that facilitate a powerful sampling of taxa (Philippe et al. 2011). Basic approaches and future challenges in phylogenomics toward reconstruction of the larger tree of life were addressed 10 years ago (Delsuc et al. 2005), and phylogenomic approaches and tree reconstruction methods have been tested using different sets of fungal genomic data (Ebersberger et al. 2012; Dutilh et al. 2007; Medina et al. 2011). Development of phylogenomic approaches for fungal phylogenetic inference has been addressed recently (Hibbett et al. 2013; Taylor and Berbee 2014) and is beyond the scope of this review. Current genome projects have sampled representative taxa in major lineages across fungal kingdom, providing extensive datasets for resolving relationships between major lineages of higher fungi. The current genomic projects might provide sufficient taxon sampling to resolve some of the unsolved polytomies within *Basidiomycota* and *Ascomycota*, as summarized in Hibbett et al. (2007). However, to resolve the phylogeny of the earliest fungal lineages, it is already clear that densely sampled genomes and the development of novel culture-independent methods will be critical. Recent phylogenomic analyses support the supergroup *Opisthosporidia* (*Microsporidia + Cryptomycota + Aphelida*) as the basal branch of all sequenced fungi (Capella-Gutierrez et al. 2012; Haag et al. 2014; James et al. 2013; Karpov et al. 2014). This group is known to be highly diverse on the basis of environmental DNA studies (Jones et al. 2011; Karpov et al. 2014) and also completely unculturable in the absence of a host. Sufficient sampling of genomes is also important for understanding divergence and recent adaptation among very closely related species, especially to reveal cryptic species and enable genome-wide population studies (Ellison et al. 2011; Park et al. 2011; Padamsee et al. 2012; Neafsey et al. 2010). Taking advantage of next-generation sequencing techniques, genome-wide expressed mRNA sequences can be easily generated without previous knowledge of genome sequence or of specific gene regions. Phylotranscriptomics, the use sequences of expressed messenger RNA sequences to infer phylogeny, has been shown to be a promising approach to infer phylogenies in several non-fungal groups (Breinholt and Kawahara 2013; Wickett et al. 2014). Similar applications in the fungal kingdom are certainly looming on the horizon.

Despite increasing sequencing capacity, it remains the case that for the majority of fungal species, genome-scale sequence is unlikely to be available soon. In most of these cases, a multilocus phylogeny is now realistically affordable and is expected to be informative enough for most systematic questions about these taxa. However, previously used genetic markers for phylogenetic analysis were originally identified by a trial and error process based on very limited data and often subsequently sequenced in other taxa solely motivated by the desire for completion of particular datasets. Thus, the phylogenetic usefulness of some genetic markers can be far from optimal (Robert et al. 2011). Sequenced genomes make it possible to assess the potential phylogenetic utility of many genetic markers as well as to enable more successful primer design and PCR efficiency (Ye et al. 2012). Knowledge regarding gene ontology and substitution rates is also critical for selecting proper markers for resolution of divergences occurring on diverse time scales during disparate epochs.

Approaches for selecting robust sets of phylogenetic markers based on sequenced genomes are starting to emerge and are urgently needed. For example, ranking genes for their usefulness in phylogenetic inferences showed promise as a means of solving phylogenies for some problematic fungal groups (Schoch et al. 2009a; Binder et al. 2013; Robert et al. 2011; Hyde et al. 2014; Capella-Gutierrez et al. 2014).

## Experimental Design and Analysis for Systematics Using Genome Data

Phylogenetic inference can be improved either by use of better models or by obtaining better data. For phylogenetic problems corresponding to short, deep internodes, quality of data is often the limitation to successful resolution (Townsend et al. 2008; Philippe et al. 2011; Su et al. 2014). Early fungal phylogenetic research expanded the repertoire of genetic markers beyond the common rRNA markers by testing and developing gene markers that had been found useful in other organisms. The first AFTOL project selected six markers to sample from major fungal groups after attempting to widely amplify more than 10 markers (Lutzoni et al. 2004; James et al. 2006a; Hibbett et al. 2007; Matheny et al. 2007; Liu and Hall 2004). Testing a small number of genetic markers on a small number of taxa using degenerate PCR amplification is laborious but feasible; however, its use for evaluating a genome-scale pool of genes for diverse taxonomic sampling would be infeasible. Identifying the most informative candidate loci across the genome in advance can provide a prioritized list for identification by degenerate PCR of novel promising markers or for use in deciding on reference gene sets for genome-scale targeted capture methodologies (e.g., Li et al. 2013). By adopting relaxed assumptions regarding the model of molecular evolution and deriving theory based on asymptotic interest in resolving short deep internodes of four taxon trees, a method for profiling phylogenetic informativeness over time of diverse gene markers was developed (Townsend 2007) and applied to the task of identifying better markers during the second AFTOL project (Schoch et al. 2009a; Townsend et al. 2008).

This theory was generalized to resolve nodes based on rates of evolution of individual characters or sets of characters onto the molecular evolutionary or chronological time scale of interest, weighing the accumulation of signal with internode length versus the loss of signal on subtending branches of the phylogenetic tree (Taylor and Berbee 2014; Su et al. 2014; Townsend et al. 2012; Feau et al. 2011; Miadlikowska et al. 2014; Walker et al. 2012). Binder et al. (2013) performed a thorough analysis of candidate loci to identify optimal experimental design for resolution of phylogenetic hypotheses. In this comprehensive study, among 356 single-copy genes, 25 markers ranked at the top for phylogenetic informativeness and probability to resolve key epochs were selected to resolve the problematic phylogeny of wood-decay fungi. As demonstrated in that study, gene markers selected

from sequenced genomes should be evaluated for their site rate distributions, phylogenetic informativeness, and predicted signal and noise. Markers then can be quantified for predicted utility compared to the worst possible performance or random sampling of taxa and genes. For a given phylogenetic hypothesis, the process should rank additional taxa whose genome sequences would provide the most power for resolving these nodes and then predict which taxon-gene elements of a presumed data matrix would provide the most power for resolving these nodes. The result minimizes the effort for resolving the given nodes (and simultaneously minimizes the probability of error) by assessing phylogenetic performance for top taxon-gene combination until a robust phylogeny is reached.

The advent of big data in phylogenomics has brought renewed attention not only to issues of phylogenetic signal but also to issues of phylogenetic noise and bias (Townsend and Lopez-Giraldez 2010; Lopez-Giraldez and Townsend 2011; Lopez-Giraldez et al. 2013). In data-limiting analyses, it was always possible to quiet concerns about the relative efficacy of some data over other data with a plaintive call for more data. In the genomic era, with the availability of big data, due to known issues such as inconsistency of substitution rates, horizontal gene transfer, and unclear gene ontology, it has become clear that big data results bulwarked by the traditional hallmarks of strong support are sometimes in conflict with each other (Salichos and Rokas 2013). The resolution of this conflict requires rigorous thought about the sources of noise and consequently the relative power of data to address phylogenetic hypotheses. At the same time, the growing resource of publicly available sequenced genes and genomes should in principle provide some guidance as to how to optimally design a phylogenetic sequencing study. For example, genes can be chosen from sequenced genomes of known phylogeny and then ranked for their performance in accurately inferring phylogenetic relationships—this approach is an extension of the practice of traditional marker selection facilitated by automatic computer programs (Capella-Gutierrez et al. 2014). Performance of these analyses is facilitated by the web application PhyDesign (Fig. 3.1) (Lopez-Giraldez and Townsend 2011). PhyDesign evaluates gene performance based on sequence alignment and a chronogram to predict signal and noise and the best-possible performance, where the metric of interest is the amount of support provided for the given nodes. Providing a means for prioritizing gene sequencing and taxon sampling and for sorting the "wheat from the chaff" in large phylogenomic studies, this application of the theory for phylogenetic study design would robustly improve the scope of data collection and analysis, the overall cost-effectiveness, and the probability of correct inference of a phylogenetic study. In addition, phylogenetic inferences are increasingly required to be robust to differential gene divergence under the multispecies coalescent, necessitating informed choices not only on what genetic markers to employ but also on what analysis approaches to take (Hyde et al. 2013).

Theoretical tools are still needed to address long-standing controversies in experimental design that have occasionally engendered contentious academic debate, including (1) the power of different genetic markers, (2) the relative utility of taxon sampling versus gene sampling, (3) the differentiation between soft and hard polytomies, and (4) the design of taxonomically dense phylogenetic studies
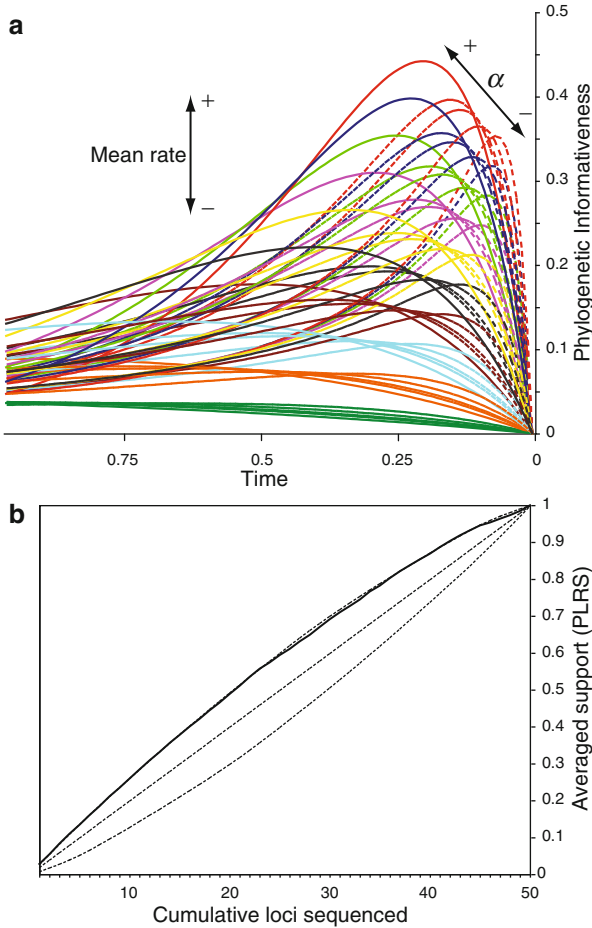
Fig. 3.1 Utility to phylogenetics of extraction of informative genes from genome sequence data. (**a**) Phylogenetic informativeness can be estimated and compared among different genes for given epochs (modified from López-Giráldez et al. 2013). Phylogenetic informativeness profiles for simulated sequence alignments on a single molecular evolutionary unit. Each of the ten different colors represents a different mean rate, from 0.0001 (slowest, bottom) to 0.001 (fastest, top) substitutions per site per time unit. *Dashed lines* are profiles from alignments simulated with gamma rate heterogeneity ($\alpha = 0.5$, 1, 2, and 3). (**b**) Cumulative proportionate likelihood-ratio support (PLRS), averaged across nodes for simulated amino acid datasets. Genes are ranked by differential phylogenetic informativeness encompassing all branches in the tree. The *upper dashed line* and the *lower dashed line* separately represent cumulative PLRS when loci are prioritized, post hoc, from highest to lowest PLRS values and from lowest to highest PLRS values. The *intermediate dashed line* is the hypothetical average one would achieve sampling at random from loci available. The *solid line* is the performance when genes are selected by their phylogenetic informativeness based on inferred rates of sites only

optimized by taxonomically sparse genome-scale data (Lopez-Giraldez et al. 2013; Moeller and Townsend 2013). A robust fungal phylogeny would provide a solid framework for fungal systematics that would, in turn, be of increasing significance in modern mycological research.

Fungal Systematics in the Future: Integration of Fungal Systematics and Fungal Evolutionary Biology

Systematics is fundamental to organismal biology and is the discipline that synthesizes achievements from all of biology and ultimately underlies all research in evolutionary biology. Arising in part from systematics, the theory of evolution is the basis of modern biology. A robust phylogeny and reliable classification is the first step for the development of fungal systematics, and systematists should not be satisfied only with describing the evolutionary history of fungal lineages (Hibbett and Taylor 2013). More importantly it is our responsibility to qualitatively and quantitatively explain how this history led to the diversity we observe today, a question that brings us to the integration of systematics and evolutionary biology. In fact, from taxonomy, diversity, molecular phylogeny, to the tree of life, the study of systematics of all organism groups has itself been evolving, and new contents from evolutionary biology have been continually if controversially incorporated into modern systematics (Losos et al. 2013).

Fungal systematics is critical for understanding the evolution of genes and their functions in fungal genetics, and multigene analysis provides an opportunity to avoid the pitfalls associated with assuming a single-gene phylogeny represents a true species phylogeny. Genetics has long focused on gene behavior and function within species, especially for model organisms, until recently the availability of sequenced genomes and robust fungal phylogeny made data available to trace gene ontology among different lineages within a long evolutionary history. Like many other eukaryotic organisms, horizontal gene transfer and gene/genome duplication are main contributors for new genes and gene functions in many fungal species (Bruto et al. 2014; Cohen-Gihon et al. 2011; Fitzpatrick 2012; Wapinski et al. 2007), and horizontal transfer of toxic gene clusters among fungal species was discovered based on sequenced fungal genomes across lineages of fungal tree of life (Slot and Rokas 2011; Wisecaver et al. 2014). For many fungi, the dominant form of their life history is haploid, and mitotic and meiotic recombination can happen via parasexual and sexual reproductions in fungal species (Schoustra et al. 2007; van de Vondervoort et al. 2007). Thus, the reconciliation of gene phylogeny and species phylogeny in low-level species taxonomy in fungi could provide insights into the modeling of speciation events (Taylor et al. 2000).

Fungal systematics and genome-enabled mycology are linked through evolutionary biology. Sequenced genomes provide a huge amount of data that can be brought to bear on all branches of fungal research. Recent progress has been espe-

cially interesting in efficiently addressing the genetic basis of various phenotypes (Hibbett et al. 2013; Taylor and Berbee 2014). Genomic research based on fungal models, such as *S. cerevisiae*, *Neurospora*, and *Aspergillus* species, has been focused on fundamental biology with implications that extend toward many non-fungal branches of the tree of life, including meiosis, cell cycle, and internal oscillation (Galagan et al. 2005). In contrast, with an increase of released fungal genomes, genome-enabled mycology has emerged: early studies have focused either on specific ecology or on metabolic pathways or functional gene families and their evolution (Spanu et al. 2010; Vogel and Moran 2013; O'Connell et al. 2012; Stajich et al. 2010; Pel et al. 2007; Martin et al. 2010; Morin et al. 2012). In most of these early studies, fungal systematics generally serves not only as a guide for what taxa to sample and study independently but also as a reference for tracking gene history. With the expected robust phylogeny and well-sampled genomes that could come as an outcome from the 1000 Fungal Genome project, a reliable gene ontology should be inferred that would facilitate inference of how fungal morphologies and ecologies have evolved, knowledge of which is one of the overarching goals of fungal systematics. For example, one-celled (yeast) stages are distributed throughout the fungal kingdom, and comparative genomics has revealed that yeast forms arose early and independently in multiple fungal clades via parallel diversification of a fungal-specific transcription factor family involved in regulating yeast-filamentous switches (Nagy et al. 2014). Reliable gene ontology is critical to the reconstruction of gene networks and the assessment of gene functions, especially for systems biology investigation that attempt to answer how complexity can be developed while essential functions are maintained.

The importance of robust phylogenies to infer the evolution of fungal ecology is clear, but fungal systematics is also an essential component of any complete understanding of fungal ecology. Inorganic and organic components of the environment impose significant selection on fungal phenotypes (Tedersoo et al. 2014). Ecological factors, such as host types, nutrient resources, or geographic distribution, have long been considered characters that are important for fungal classification. With well-resolved molecular phylogenies, we could evaluate the role of ecology in fungal evolution, reconcile the ontology of specific gene function groups, and infer the genetic basis of ecological success. Recent discoveries on the evolution of wood decay among polypore species and mushroom-forming fungi have demonstrated how this strategy can work (Binder et al. 2013; Floudas et al. 2012; Eastwood et al. 2011). Applying principles of systematics to metagenomics makes it possible to monitor the dynamics of biological processes involving diverse fungal species on both spatial and temporal scales to understand the contributions of those fungal species to the process of interest. For instance, a study on global soil sampled by (Tedersoo et al. 2014) demonstrates direct and indirect effects of climatic and edaphic variables on plant and fungal richness. The National Science Foundation has launched a program called Dimensions of Biodiversity, which "takes a broad view of biodiversity and focuses on the intersection of genetic, phylogenetic, and functional dimensions of biodiversity."

Further extension of the broad impacts of fungal systematic research requires experienced mycologists with broad training in traditional fungal classification, molecular systematics, and bioinformatics/genomics. Mycologists have long been considered as naturalists. Training of fungal systematics has been provided in many institutes, especially in colleges or departments for plant pathology. Fungal classification and taxonomy training usually via monographic work require a lot of time in both field and laboratory, while molecular systematic training requires a decent facility for sequencing and/or computation. Significant computational needs are especially required for phylogenomics. Funding resources are heavily biased toward molecular research, leading to a scarcity of high-quality training in traditional fungal systematics, especially at the graduate level (Pearson et al. 2011). In the long run, the lack of well-trained mycological systematists would be a problem not only holding back the development of fungal systematics but also impeding many other research fields that rely on knowledge of fungal biodiversity and evolutionary biology. Well-trained mycologists are also critical for helping the public to understand the gaps between the quickly developing "omics" sciences and the long-developed traditional senses of fungi and fungal biology.

The greatest challenge for fungal systematics has always been to be able to take disparate pieces of knowledge from diverse kinds of studies of fungi to make synthetic biological inference, and only in this way will fungal systematics be of maximum benefit to the whole community conducting research on fungi and the scientific community at large.

References

Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJ, Culley D et al (2011) Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21:885–897

Bass D, Richards TA (2011) Three reasons to re-evaluate fungal diversity 'on Earth in the ocean'. Fungal Biol Rev 25:159–164

Bidartondo MI (2008) Preserving accuracy in GenBank. Science 319:1616

Bidartondo MI, Read DJ, Trappe JM, Merckx V, Ligrone R et al (2011) The dawn of symbiosis between plants and fungi. Biol Lett 7:574–577

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F et al (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350–1373

Blackwell M (2011) The fungi: 1, 2, 3 … 5.1 million species? Am J Bot 98:426–438

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

Blaxter M, Mann J, Chapman T, Thomas F, Whitton C et al (2005) Defining operational taxonomic units using DNA barcode data. Philos Trans R Soc Lond B Biol Sci 360:1935–1943

Breinholt JW, Kawahara AY (2013) Phylotranscriptomics: saturated third codon positions radically influence the estimation of trees based on next-gen data. Genome Biol Evol 5:2082–2092

Bridge P, Spooner B, Roberts P (2005) The impact of molecular data in fungal systematics. Adv Bot Res 42:33–67

Bruns T, White T, Taylor J (1991) Fungal molecular systematics. Annu Rev Ecol Syst 22:525–564

Bruto M, Prigent-Combaret C, Luis P, Moenne-Loccoz Y, Muller D (2014) Frequent, independent transfers of a catabolic gene from bacteria to contrasted filamentous eukaryotes. Proc Biol Sci 281:20140848

Cao Y, Wu SH, Dai YC (2012) Species clarification of the prize medicinal *Ganoderma* mushroom "Lingzhi". Fungal Divers 56:49–62

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

Capella-Gutierrez S, Kauff F, Gabaldon T (2014) A phylogenomics approach for selecting robust sets of phylogenetic markers. Nucleic Acids Res 42, e54

Carlson A, Justo A, Hibbett DS (2014) Species delimitation in *Trametes*: a comparison of ITS, RPB1, RPB2 and TEF1 gene phylogenies. Mycologia 106:735–745

Cohen-Gihon I, Sharan R, Nussinov R (2011) Processes of fungal proteome evolution and gain of function: gene duplication and domain rearrangement. Phys Biol 8:035009

Dai YC, Cui BK (2011) *Fomitiporia ellipsoidea* has the largest fruiting body among the fungi. Fungal Biol 115:813–814

de Bertoldi M, Lepidi AA, Nuti MP (1973) Significance of DNA base composition in classification of *Humicola* and related genera. Trans Br Mycol Soc 60:77–85

De Oliveira ML, Mallo D, Posada D (2014) A Bayesian supertree model for genome-wide species tree reconstruction. Syst Biol. doi:10.1093/sysbio/syu082

de Queiroz K, Gauthier J (1994) Toward a phylogenetic system of biological nomenclature. Trends Ecol Evol 9:27–31

Del Olmo-Ruiz M, Arnold AE (2014) Interannual variation and host affiliations of endophytic fungi associated with ferns at La Selva, Costa Rica. Mycologia 106:8–21

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

Du XH, Zhao Q, O'Donnell K, Rooney AP, Yang ZL (2012) Multigene molecular phylogenetics reveals true morels (*Morchella*) are especially species-rich in China. Fungal Genet Biol 49:455–469

Dutilh BE, van Noort V, van der Heijden RT, Boekhout T, Snel B et al (2007) Assessment of phylogenomic and orthology approaches for phylogenetic inference. Bioinformatics 23:815–824

Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P et al (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science 333:762–765

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

Ellison CE, Hall C, Kowbel D, Welch J, Brem RB et al (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. Proc Natl Acad Sci USA 108:2831–2836

Feau N, Decourcelle T, Husson C, Desprez-Loustau ML, Dutech C (2011) Finding single copy genes out of sequenced genomes for multilocus phylogenetics in non-model fungi. PLoS One 6, e18803

Fitzpatrick DA (2012) Horizontal gene transfer in fungi. FEMS Microbiol Lett 329:1–8

Floudas D, Binder M, Riley R, Barry K, Blanchette RA et al (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719

Galagan JE, Henn MR, Ma LJ, Cuomo CA, Birren B (2005) Genomics of the fungal kingdom: insights into eukaryotic biology. Genome Res 15:1620–1631

Geiser DM, Aoki T, Bacon CW, Baker SE, Bhattacharyya MK et al (2013) One fungus, one name: defining the genus *Fusarium* in a scientifically robust way that preserves longstanding use. Phytopathology 103:400–408

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B et al (1996) Life with 6000 genes. Science 274(546):563–567

Haag KL, James TY, Pombert JF, Larsson R, Schaer TM et al (2014) Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. Proc Natl Acad Sci USA 111:15480–15485

Heitman J (2007) Sex in fungi: molecular determination and evolutionary implication. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA (eds). ASM Press, Washington, DC

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

Hibbett DS, Nilsson RH, Snyder M, Fonseca M, Costanzo J et al (2005) Automated phylogenetic taxonomy: an example in the homobasidiomycetes (mushroom-forming fungi). Syst Biol 54:660–668

Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF et al (2007) A higher-level phyloge-netic classification of the fungi. Mycol Res 111:509–547

Hibbett DS, Ohman A, Glotzer D, Nuhn M, Kirk PM et al (2011) Progress in molecular and mor-phological taxon discovery in Fungi and options for formal classification of environmental sequences. Fungal Biol Rev 25:38–47

Hibbett DS, Stajich JE, Spatafora JW (2013) Toward genome-enabled mycology. Mycologia 105:1339–1349

Hyde KD, Udayanga D, Manamgoda D, Tedersoo L, Larsson E et al (2013) Incorporating molecu-lar data in fungal systematics: a guide for aspiring researchers. Currt Res Environ Appl Mycol 3:1–32

Hyde KD, Nilsson RH, Alias SA, Ariyawansa HA, Blair JE et al (2014) One stop shop: backbones trees for important phytopathogenic genera: I. Fungal Divers 67:21–125

Ishii N, Ishida S, Kagami M (2015) PCR primers for assessing community structure of aquatic fungi including Chytridiomycota and Cryptomycota. Fungal Ecol 13:33–43

James TY, Berbee ML (2012) No jacket required—new fungal lineage defies dress code: recently described zoosporic fungi lack a cell wall during trophic phase. Bioessays 34:94–102

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V 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 et al (2006b) A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). Mycologia 98:860–871

James TY, Pelin A, Bonen L, Ahrendt S, Sain D et al (2013) Shared signatures of parasitism and phylogenomics unite Cryptomycota and *microsporidia*. Curr Biol 23:1548–1553

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

Karpov SA, Mamkaeva MA, Aleoshin VV, Nassonova E, Lilje O et al (2014) Morphology, phy-logeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new super-phylum Opisthosporidia. Front Microbiol 5:112

Kauff F, Cox CJ, Lutzoni F (2007) WASABI: an automated sequence processing system for mul-tigene phylogenies. Syst Biol 56:523–531

Kirk PM, Cannon PF, Minter DW, Stalpers JA (eds) (2008) Ainsworth and Bisby's dictionary of the fungi. CAB International, Wallingford

Koetschan C, Forster F, Keller A, Schleicher T, Ruderisch B et al (2010) The ITS2 Database III—sequences and structures for phylogeny. Nucleic Acids Res 38:D275–D279

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M et al (2013) Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22:5271–5277

Lehr NA, Wang Z, Li N, Hewitt DA, Lopez-Giraldez F et al (2014) Gene expression differences among three *Neurospora* species reveal genes required for sexual reproduction in *Neurospora crassa*. PLoS One 9, e110398

Li C, Hofreiter M, Straube N, Corrigan S, Naylor GJ (2013) Capturing protein-coding genes across highly divergent species. Biotechniques 54:321–326

Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T et al (2013) Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. New Phytol 199:288–299

Liu YJ, Hall BD (2004) Body plan evolution of ascomycetes, as inferred from an RNA polymerase II phylogeny. Proc Natl Acad Sci USA 101:4507–4512

Liu K, Raghavan S, Nelesen S, Linder CR, Warnow T (2009) Rapid and accurate large-scale coestimation of sequence alignments and phylogenetic trees. Science 324:1561–1564

Liu K, Warnow TJ, Holder MT, Nelesen SM, Yu J et al (2012) SATe-II: very fast and accurate simultaneous estimation of multiple sequence alignments and phylogenetic trees. Syst Biol 61:90–106

Lopez-Giraldez F, Townsend JP (2011) PhyDesign: an online application for profiling phylogenetic informativeness. BMC Evol Biol 11:152

Lopez-Giraldez F, Moeller AH, Townsend JP (2013) Evaluating phylogenetic informativeness as a predictor of phylogenetic signal for metazoan, fungal, and mammalian phylogenomic data sets. Biomed Res Int 2013:621604

Losos JB, Arnold SJ, Bejerano G, Brodie ED 3rd, Hibbett D et al (2013) Evolutionary biology for the 21st century. PLoS Biol 11, e1001466

Lutzoni F, Kauff F, Cox CJ, McLaughlin D, Celio G et al (2004) Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. Am J Bot 91:1446–1480

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM et al (2010) Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature 464:1033–1038

Matheny PB, Wang Z, Binder M, Curtis JM, Lim YW et al (2007) Contributions of *rpb2* and *tef1* to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). Mol Phylogenet Evol 43:430–451

McLaughlin DJ, Spatafora JW, Esser K (eds) (2014) The Mycota. A comprehensive treatise on fungi as experimental systems for basic and applied research. Springer, Heidelberg

McNeill J, Turland NJ (2012) International code of nomenclature for algae, fungi, and plants (Melbourne Code). PREFACE. 154:Ix–Xxii

Medina EM, Jones GW, Fitzpatrick DA (2011) Reconstructing the fungal tree of life using phylogenomics and a preliminary investigation of the distribution of yeast prion-like proteins in the fungal kingdom. J Mol Evol 73:116–133

Miadlikowska J, Kauff F, Hognabba F, Oliver JC, Molnar K et al (2014) A multigene phylogenetic synthesis for the class Lecanoromycetes (Ascomycota): 1307 fungi representing 1139 infrageneric taxa, 317 genera and 66 families. Mol Phylogenet Evol 79:132–168

Moeller AH, Townsend JP (2013) Response to: the relative utility of sequence divergence and phylogenetic informativeness profiling in phylogenetic study design. Mol Phylogenet Evol 66:436

Money NP (2013) Against the naming of fungi. Fungal Biol 117:463–465

Morin E, Kohler A, Baker AR, Foulongne-Oriol M, Lombard V et al (2012) Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. Proc Natl Acad Sci USA 109:17501–17506

Nagy LG, Ohm RA, Kovacs GM, Floudas D, Riley R et al (2014) Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. Nat Commun 5:4471

Neafsey DE, Barker BM, Sharpton TJ, Stajich JE, Park DJ et al (2010) Population genomic sequencing of *Coccidioides* fungi reveals recent hybridization and transposon control. Genome Res 20:938–946

Nilsson RH, Ryberg M, Sjokvist E, Abarenkov K (2011) Rethinking taxon sampling in the light of environmental sequencing. Cladistics 27:197–203

Nilsson RH, Tedersoo L, Abarenkov K, Ryberg M, Kristiansson E et al (2012) Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. MycoKeys 4:37–63

Nishant KT, Wei W, Mancera E, Argueso JL, Schlattl A et al (2010) The baker's yeast diploid genome is remarkably stable in vegetative growth and meiosis. PLoS Genet 6, e1001109

O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J et al (2012) Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. Nat Genet 44:1060–1065

Ortiz-Santana B, Lindner DL, Miettinen O, Justo A, Hibbett DS (2013) A phylogenetic overview of the antrodia clade (Basidiomycota, Polyporales). Mycologia 105:1391–1411

Otrosina WJ, Garbelotto M (2010) *Heterobasidion occidentale* sp. nov. and Heterobasidion irregulare nom. nov.: a disposition of North American Heterobasidion biological species. Fungal Biol 114:16–25

Padamsee M, Kumar TK, Riley R, Binder M, Boyd A et al (2012) The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. Fungal Genet Biol 49:217–226

Pante E, Schoelinck C, Puillandre N (2015) From integrative taxonomy to species description: one step beyond. Syst Biol 64:152–160

Park B, Park J, Cheong KC, Choi J, Jung K et al (2011) Cyber infrastructure for *Fusarium*: three integrated platforms supporting strain identification, phylogenetics, comparative genomics and knowledge sharing. Nucleic Acids Res 39:D640–D646

Pearson DL, Hamilton AL, Erwin TL (2011) Recovery plan for the endangered taxonomy profession. Bioscience 61:58–63

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G et al (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol 25:221–231

Philippe H, Brinkmann H, Lavrov DV, Littlewood DT, Manuel M et al (2011) Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol 9, e1000602

Porras-Alfaro A, Liu KL, Kuske CR, Xie G (2014) From genus to phylum: large-subunit and internal transcribed spacer rRNA operon regions show similar classification accuracies influenced by database composition. Appl Environ Microbiol 80:829–840

Porter TM, Schadt CW, Rizvi L, Martin AP, Schmidt SK et al (2008) Widespread occurrence and phylogenetic placement of a soil clone group adds a prominent new branch to the fungal tree of life. Mol Phylogenet Evol 46:635–644

Richard F, Bellanger JM, Clowez P, Courtecuisse R, Hansen K et al (2014) True morels (*Morchella*, Pezizales) of Europe and North America: evolutionary relationships inferred from multilocus data and a unified taxonomy. Mycologia 107:359–382.

Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ et al (2013) Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431–437

Robert V, Szoke S, Eberhardt U, Cardinali G, Meyer W et al (2011) The quest for a general and reliable fungal DNA barcode. Open Appl Inf J 5(Suppl 1-M6):45–61

Rodriguez-Romero J, Hedtke M, Kastner C, Muller S, Fischer R (2010) Fungi, hidden in soil or up in the air: light makes a difference. Annu Rev Microbiol 64:585–610

Rosling A, Cox F, Cruz-Martinez K, Ihrmark K, Grelet GA et al (2011) Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. Science 333:876–879

Rosling A, Timling I, Taylor DL (2013) Distribution and abundance of Archaeorhizomycetes. In: Horwitz A, Mukherjee PK, Mukherjee M, Kubicek CP (eds) Genomics of soil- and plant-associated fungi. Springer, Heidelberg, pp 333–349

Salichos L, Rokas A (2013) Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497:327–331

Schoch CL, Sung GH, Lopez-Giraldez F, Townsend JP, Miadlikowska J et al (2009a) The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst Biol 58:224–239

Schoch CL, Wang Z, Townsend JP, Spatafora JW (2009b) Geoglossomycetes cl. nov., Geoglossales ord. nov. and taxa above class rank in the Ascomycota tree of life. Persoonia 22:129–138

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL et al (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci USA 109:6241–6246

Schoch CL, Robbertse B, Robert V, Vu D, Cardinali G et al (2014) Finding needles in haystacks: linking scientific names, reference specimens and molecular data for fungi. Database (Oxford). doi:10.1093/database/bau061

Schoustra SE, Debets AJ, Slakhorst M, Hoekstra RF (2007) Mitotic recombination accelerates adaptation in the fungus *Aspergillus nidulans*. PLoS Genet 3, e68

Seifert KA, Rossman AY (2010) How to describe a new fungal species. IMA Fungus 1:109–116

Sikhakolli UR, Lopez-Giraldez F, Li N, Common R, Townsend JP et al (2012) Transcriptome analyses during fruiting body formation in *Fusarium graminearum* and *Fusarium verticillioides* reflect species life history and ecology. Fungal Genet Biol 49:663–673

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

Smith ML, Bruhn JN, Anderson JB (1992) The fungus *Armillaria bulbosa* is among the largest and oldest living organisms. Nature 356:428–431

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM et al (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330:1543–1546

Stajich J, Berbee M, Blackwell M, Hibbett D, James T et al (2009) The fungi. Curr Biol 19:R840–R845

Stajich JE, Wilke SK, Ahren D, Au CH, Birren BW et al (2010) Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). Proc Natl Acad Sci USA 107:11889–11894

Su Z, Wang Z, Lopez-Giraldez F, Townsend JP (2014) The impact of incorporating molecular evolutionary model into predictions of phylogenetic signal and noise. Front Ecol Evol 2:1–11

Taylor JW, Berbee ML (2014) Fungi from PCR to Genomics: the spreading revolution in evolutionary biology. In: Esser K, McLaughlin D, Spatafora J (eds) Mycota VII. A comprehensive treatise on fungi as experimental systems for basic and applied research: systematics and evolution, 2nd edn. Springer, Berlin, pp 1–18

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM et al (2000) Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol 31:21–32

Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006) Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. Philos Trans R Soc Lond B Biol Sci 361:1947–1963

Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS et al (2014) Global diversity and geography of soil fungi. Science 346(6213):1256688

Townsend JP (2007) Profiling phylogenetic informativeness. Syst Biol 56:222–231

Townsend JP, Lopez-Giraldez F (2010) Optimal selection of gene and ingroup taxon sampling for resolving phylogenetic relationships. Syst Biol 59:446–457

Townsend JP, Lopez-Giraldez F, Friedman R (2008) The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. J Mol Evol 67:437–447

Townsend JP, Su Z, Tekle YI (2012) Phylogenetic signal and noise: predicting the power of a data set to resolve phylogeny. Syst Biol 61:835–849

van de Vondervoort PJ, Langeveld SM, Visser J, van Peij NN, Pel HJ et al (2007) Identification of a mitotic recombination hotspot on chromosome III of the asexual fungus *Aspergillus niger* and its possible correlation with [corrected] elevated basal transcription. Curr Genet 52:107–114

van der Heijden MG, Martin FM, Selosse MA, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytol 205:1406–1423

Vogel KJ, Moran NA (2013) Functional and evolutionary analysis of the genome of an obligate fungal symbiont. Genome Biol Evol 5:891–904

Walker DM, Castlebury LA, Rossman AY, White JF Jr (2012) New molecular markers for fungal phylogenetics: two genes for species-level systematics in the Sordariomycetes (Ascomycota). Mol Phylogenet Evol 64:500–512

Wang Z, Nilsson RH, Lopez-Giraldez F, Zhuang WY, Dai YC et al (2011) Tasting soil fungal diversity with earth tongues: phylogenetic test of SATe alignments for environmental ITS data. PLoS One 6, e19039

Wapinski I, Pfeffer A, Friedman N, Regev A (2007) Natural history and evolutionary principles of gene duplication in fungi. Nature 449:54–61

Wardle DA, Lindahl BD (2014) Disentangling global soil fungal diversity. Science 346:1052–1053

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E et al (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci USA 111:E4859–E4868

Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. Syst Biol 54:844–851

Wisecaver JH, Slot JC, Rokas A (2014) The evolution of fungal metabolic pathways. PLoS Genet 10, e1004816

Wu F, Yuan Y, Malysheva VF, Du P, Dai YC (2014) Species clarification of the most important and cultivated *Auricularia* mushroom "Heimuer": evidence from morphological and molecular data. Phytotaxa 186:241–253

Wurzbacher C, Grossart HP (2012) Improved detection and identification of aquatic fungi and chitin in aquatic environments. Mycologia 104:1267–1271

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S et al (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinf 13:134

Zhang LF, Yang ZL, Song DS (2005) A phylogenetic study of commercial Chinese truffles and their allies: taxonomic implications. FEMS Microbiol Lett 245:85–92

Zhang N, Rossman AY, Seifert K, Bennett JW, Cai G et al (2013) Impacts of the international code of nomenclature for algae, fungi, and plants (Melbourne Code) on the scientific names of plant pathogenic fungi. Online, APSnet Feature: American Phytopathoglogical Society, St. Paul, MN