The Ascomycota Tree of Life: A Phylum-wide Phylogeny Clarifies the Origin and Evolution of Fundamental Reproductive and Ecological Traits

CONRAD L. SCHOCH1, a G-HO SUNG1, FRANCESC LOPEZ-GIRALDEZ2, JEFFREY P. TOWSEND2,
JOLANTA MIADLIKOWSKA3, VALERIE HOFSTETTER4, BARBARA ROBBERTSE5, P. BRANDON MATHENY5,6,
FRANK KAUFF7, ZHENG WANG2, CECILE GUEIDAN8, RACHAEL M. ANDRIE1,
KRISTIN TRIPPE5, LINDA M. GUFETTI1, ANJA WYNN8, EMILY FRAKER3,
BRENDAN P. HODKINSON5, GREGORY BONITO3, JOHANNES Z. GRONEWALD8,
MAHDI ARZANLOU8, PEDRO W. CROUS8, DAVID HEWITT9,
DONALD H. PFISTER10, KRISTIN PETERSON10, MARIEKA GRYZENHOUT11, MICHAEL J. WINGFIELD11,
ANDRE APTROROT12, SUNG-OUI SUI-I13, MEREDITH BLACKWELL14, DAVID M. HILLIS15,
GARETH W. GRIFFITH6, LISA A. CASTLEBURY17, AMY Y. ROSSMAN17, H. THORSTEN LUMSCH8,
ROBERT LUCKING18, BURKHARD BUDEL19, ALEXANDRA RAUHUT19, PAUL DIEDERICH20,
DAMIEN ERTZ21, DAVID M. GEISER22, KENTARO HOSAKA23, PATRIK INDERBITZIN24,
JAN KOHLMEYER25, BRIGHTJE VOLKMANN-KOHLMEYER25, LIXZEL MOSTERT25,
HARRIE SIPMAN25, JACEK D. ROGERS26, ROBERT A. SHOEMAKER30,
RICHARD C. SUMMERBELL32, WENDY UNTEREINER33, PETER R. JOHNSTON34,
ALGA ZUCCARO36, PAUL S. DYER37, PETER D. CRITTENDEN37, MARIE T. COLE38,
KAREN HANSEN39, JAMES M. TRAPPE40, REBECCA YAHRI41, FRANCOIS LUTZONI42, and JOSEPH W. SPATAFORA1.

1Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA;
2Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, CT 06520, USA;
3Department of Biology, Duke University, Durham, NC 27708, USA;
4Swiss Federal Research Station for Plant Production of Changes (RAC), PO Box 254, CH-1260 Nyon 1, Switzerland;
5Biological Department, Clark University, Worcester, MA 01610, USA;
6Department of Ecology and Evolutionary Biology, University of Tennessee, 569 Dailey Hall, Knoxville, TN 37996, USA;
7Molecular Phylogenetics, FB Biologie, 13/276, TU Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany;
8Centralbureau voor Schimmelcultures, Fungal Biodiversity Centre, PO Box 85167, 3508 AD, Utrecht, the Netherlands;
9Department of Ecology, Copenhagen University Institute for Ecology, Thorvaldsvej 40, 1871 Frederiksberg, Copenhagen, Denmark;
10Department of Organismic and Evolutionary Biology, Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA;
11Forestry & Agricultural Biotechnology Institute, University of Pretoria, Pretoria 0002, South Africa;
12Adviesbureau voor Bryologie en Lichenologie, G. v. d. Veensweg 1, 1890 XE Hilvarenbeek, Netherlands;
13Mycology Program, American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110, USA;
14Department of Botanical Sciences, Louisiana State University, Baton Rouge, LA 70803, USA;
15Section of Integrative Biology, University of Texas, Austin, TX 78712, USA;
16Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, Ceredigion, SY23 3DQ, Wales, UK;
17United States Department of Agriculture, Agricultural Research Service, Systematic Mycology and Microbiology Laboratory, 10300 Baltimore Avenue, Beltsville, MD 20705, USA;
18Department of Botany, The Field Museum, 1400 South Lake Shore Drive, Chicago, IL 60605, USA;
19Fachbereich Biologie, Abteilung Pilzenzooologie und Systematik, 67653 Kaiserslautern, Germany;
20Musée National d’Histoire Naturelle, L-2160 Luxembourg, Luxembourg;
21Department of Cryptograms (BT), National Botanical Garden of Belgium, Domaine de Bouchout, B-1860 Meise, Belgium;
22Department of Plant Pathology Pennsylvania State University, University Park, PA 16802, USA;
23Department of Botany, National Museum of Nature and Science, Tsukuba-shi, Ibaraki 305-0005, Japan;
24University of California, Kearney Agricultural Center, 9240 South Riverbend Avenue, Parlier, CA 93648, USA;
25Institute of Marine Sciences, University of North Carolina at Chapel Hill, Morehead City, NC 28557, USA;
26Department of Plant Pathology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa;
27MicrobialGnomics Research Unit, National Center for Agricultural Utilization Research, U.S. Department of Agriculture, Agricultural Research Service, Peoria, IL 61604, USA;
28Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, Konigin-Luise-Straße, 6-8, D-14195 Berlin, Germany;
29Department of Plant Pathology, Washington State University, Pullman, WA 99164, USA;
30Biodiversity (Mycology and Botany), Agriculture and Agri-Food Canada, Ottawa, Ontario K1A OC6, Canada;
31TechnoSuruga Laboratory Co. US, Tokyo Office, Chiyoda-ku, Tokyo 101-0052, Japan;
3227 Hillcrest Park Toronto, Ontario MAX IE9, Canada;
33Department of Biology, Brandon University, Brandon, Manitoba R7A 6G9, Canada;
34HerbariumPOD, Landcare Research, Private Bag 92170, Auckland, New Zealand;
35Botanical Museum, Finnish Museum of Natural History, PO Box 7, FI-00014 University of Helsinki, Finland.
Ascomycota, with approximately 64000 known species (Kirk et al. 2008), is the largest phylum of Fungi and one of the most diverse and ubiquitous phyla of eukaryotes. Its species occur in numerous ecological niches and virtually all terrestrial and aquatic ecosystems. They function in the decay of organic substrates (e.g., wood, leaf litter, and dung) and act as mutualists, parasites, and pathogens of animals, plants, and other fungi. More than 40% of all named Ascomycota are lichenized, covering approximately 8% of the Earth’s landmasses (Brodo et al. 2001). In addition to their presence in most natural, industrial, and agricultural settings, they have been isolated from some of the most extreme environments on earth—from inside rocks on the frozen plains of Antarctica (Selbmann et al. 2005) to deep-sea wood (Kohlmeyer 1977) and sediments (Raghukumar et al. 2004). Notably, many Ascomycota are known only from asexual reproductive states, often making their phylogenetic placement on morphological criteria alone challenging.

The defining synapomorphy for Ascomycota is a specialized sac-like structure (ascus) in which meiotic spores (ascospores) are produced (Fig. 1M–Q). Although a large number of species occur only in a single-celled yeast phase, historical classifications emphasized the morphology of the multicellular ascus bearing sporocarps, ascus morphology, and ascospore release (ascus dehiscence). Based on these characters, 4 classes of sporocarp-producing fungi were traditionally recognized including Discomycetes, Pyrenomycetes, Plectomycetes, and Loculoascomycetes (Ainsworth et al. 1971). Discomycetes were defined by the production of exposed asci organized in a disc, cup, or club-shaped sporocarp, an ascopore (Fig. 1A–F). Discomycete asci possessed a single-functional wall layer (unitunicate) and released ascospores through an ascus tip that either possessed a lid-like structure or operculum (operculate discomycetes; Fig. 1J) or lacked it (moperculate discomycetes). Pyrenomycetes produced inoperculate, unitunicate asci (Fig. 1D) that were enclosed in a flask-shaped sporocarp, a perithecium (Fig. 1L). Plectomycetes possessed thin-walled prototunicate asci (Fig. 1M), which typically dissolved at maturity within a completely enclosed sporocarp, a cleistothecium. Loculoascomycetes was differentiated based on the production of specialized thick-walled asci in preformed openings in stromatic tissue (pseudothecia, Fig. 1K). The ascus wall layers separate in a “jack-in-the-box” or fissitunicate manner (Fig. 1P). Additional variants of sporocarp and asci also occur such as various types of unrelated enclosed sporocarps (Fig. 1G, H) and thick-walled asci (Fig. 1O). However, these can generally be considered variants of the main morphologies mentioned above, and more detailed morphological depictions of these character states can be found in McLaughlin et al. (2001) and Kirk et al. (2008).

The first phylogenies using DNA sequence from the nuclear ribosomal genes indicated that fungal groups separated by the aforementioned sporocarp and ascus characters did not always correspond to phylogenetic clades (Berbee and Taylor 1992; Spatafora et al. 1995). Phylogenetic resolution increased considerably with the utilization of slowly evolving protein-coding genes especially suited to the investigation of deep divergences (Liu et al. 1999; Liu and Hall 2004). The availability of multigene sequence data from the research consortium "Assembling the Fungal Tree of Life" (AFTOL) and numerous genome sequences resulted in several recent phylogenies with comprehensive molecular character sampling across Ascomycota (Fitzpatrick et al. 2012).
Materials and Methods

Gene Amplification and DNA Sequencing

We collected molecular data for 6 different loci—nSSU and nLSU, mSSU, and fragments from 3 protein-coding genes: the largest and second largest subunits of RNA polymerase II (RPB1 and RPB2) and transcription elongation factor (TEF1). Sequences were generated in various laboratories associated with the AFTOL project using primers and polymerase chain reaction amplification conditions as described in Lutzoni et al. (2004). The most commonly used primers were as follows—nSSU: NS1 and NS24; nLSU: LROR and LR7; RPB1: RPB1-Ac, RPB1-Cr, and R1-1DDR; RPB2: RPB2-SF, RPB2-7cR, and RRPB2-11a; and TEF1: 983 and 2218R. Sequences for primers, together with the taxa they amplified are listed in the WASABI database (http://www.aftol.org). Both DNA strands were sequenced with BigDye Terminator v3.1 cycle sequencing (Applied Biosystems, Foster City, CA) on an ABI Prism 3730 or 3730-1 Genetic Analyzer by either the core labs sequencing facility at Oregon State University, the Duke Sequencing Facility, or Macrogen, Inc. (Seoul, Korea). Contigs were assembled using SeqMere in the GCG 11 software suite (Accelrys, San Diego, CA). Previously published data were obtained from GenBank and WASABI.

Data Sets

Strains of 52 taxa were sequenced for the first time, and when combined with preexisting unpublished data, resulted in 394 newly determined sequence submissions to GenBank (indicated by bold numbers in online Appendix 1, Supplementary Table 2, http://www.sysbio.oxfordjournals.org/). To test whether phylogenies were significantly impacted by the amount of missing data due to incomplete or absent sequences, we used 5 different data sets that represent a continuum of maximizing characters versus maximizing taxa. First, we assembled the 2-gene, 434-taxon data set (2G434T) in which most taxa possessed 2 of the 6 genes sampled. A representative set of 13 Basidiomycota species and 1 member of Mucoromycotina was included as out-groups, leaving 420 members of Ascomycota in the analysis. Of these, 24 taxa were represented by only 2 genes and Pyxidiophora sp. represented by only 1 gene in order to verify clade representation. The total 2G434T data set comprised 7297 characters, of which 36% were missing after exclusion of ambiguous characters. Second, more focused data sets were constructed, consisting of matrices where at least 3 genes and 409 taxa (3G409T), 4 genes and 335 taxa (4G335T) and 5 genes and 234 taxa (5G234T) were present. These reconstructions resulted in missing character percentages of 34%, 28%, and 22%, respectively. A fifth data set had all 6 genes present and 126 taxa (6G126T) with 17% of the characters missing. Specific data on all taxa used are listed in Supplementary Tables 2 and 3 (online Appendix 1, Supplementary Table 2, http://www.sysbio.oxfordjournals.org/).
protein-coding genes. Improvements of variable region alignments were performed using MUSCLE 3.6 (Edgar 2004) (web interface hosted by the Centre for Genome Research and Biocomputing at Oregon State University). Exclusions obtained from WASABI and previous AFTOL publications (Lutzoni et al. 2004; James et al. 2006) were used initially and expanded after further inspection. This was done by viewing the concatenated alignment in BioEdit v.7.0.1 (Hall 1999) with a shade threshold set to 40%. After visual inspection, stretches containing more than 5 bases of non-highlighted, ambiguous regions were excluded. Final alignments and trees were released to TreeBASE, accession number S2141, and were also uploaded to WASABI.
Tests for Conflict

Major incongruences between data sets were detected as in Miadlikowska et al. (2006) using the program compat.py (Kauff and Lutzoni 2002) from http://www.lutzonilab.net. Topological incongruences were detected in single-gene data sets by comparing reciprocal 70% neighbor-joining bootstrap support for all gene combinations. Conflicts were considered Significant when a taxon was differentially resolved between 2 gene trees with greater than 70% bootstrap support. The separate Single-gene trees were obtained by using neighbor-joining with maximum-likelihood distances calculated under default settings (equal rates, Hasegawa-Kishino-Yano [HKY] model) in PAUP v4b10 (Swofford 2002) with 500 bootstrap replications (Miadlikowska et al. 2006). The individual gene data sets were inspected, and taxa with no data or sequences that were too short to produce reliable distance values were removed. In order to reduce any conflicts due to the use of erroneous sequences, but still allow for variability between genes and the different taxon sets, we only defined sequences as being in conflict when 2 sequences from 1 species were resolved within different orders with greater than 70% bootstrap support.

Phylogenetic Analyses

Comparative analyses of large, but more focused data sets (Miadlikowska et al. 2006; Schoch et al. 2006; Spatafora et al. 2006) prompted us to select RAxML (Randomized Accelerated Maximum Likelihood) due to a combination of speed, accuracy, and scalability over numerous independent processors (Stamatakis 2006). Data set 2G434T was analyzed using RAxML MPI v7.0 (Stamatakis 2006) on 2 Sun Enterprise Servers (4 processors) and a Microway Linux cluster (8 dual alpha processor nodes). Combined data sets consisted of ribosomal DNA (rDNA) genes as nucleic acid and protein-coding genes coded as amino acids, with combined analyses partitioned by genes.

Models of evolution were selected for each partition of the 3 ribosomal genes by using the Akaike Information Criterion as implemented in Modeltest v3.7 (Posada and Crandall 1998). This criterion resulted in the choice of the general time-reversible model incorporating a proportion of invariable sites and a discretized gamma distribution of rates at variable sites (GTR + I + Γ). The same model was obtained for the DNA sequences of protein-coding genes for the DNA-only phylogeny. The 3 protein genes were also individually subjected to model testing with ProtTest v.1.2.6 (Abascal et al. 2005) under the same criteria. TheRTREV model with a discrete gamma distribution and 4 rate classes plus an estimation of invariable sites was selected as the most likely model in all 3 cases. In addition, protein models for RPB2 and TEF1 incorporated a parameter to estimate amino acid frequencies. The most likely RAxML tree from 100 randomized starting trees using the GTRGAMM1 setting and appropriate protein settings was selected for use as a template in Figure 2 and represents the most complete phylogeny of Ascomycota to date. A more detailed phylogeny is shown in online Appendix 2, as well as a tree obtained from DNA data only. For comparison, a most likely tree was obtained from a separate series of 500 RAxML trees using the GTRMIX and PROTMIX model settings with 25 rate categories (relying on GTRCAT and PROTCAT model approximations) and compared with the previous tree. Class- and ordinal-level nodes were congruent with the previous tree, and this phylogeny was not used in further analyses. Bootstrap values from 1000 repetitions with the GTRCAT and PROTCAT model approximations and 25 rate categories were used to assess nodal support (using the "exhaustive" setting "-f i"). A second set of analyses was performed on the same set of data with DNA alignments only. Similar to the combined protein and DNA analysis, the GTR + I + Γ model was used for each separate partition. Independent parameters for invariable sites and the gamma shape parameter were applied according to partitions by codon position (where applicable).

In addition to the above-mentioned complete data set, the following restrictive data sets were also analyzed with RAxML under the same model settings: 3G409T, 4G335T, 5G234T, and 6G126T for both DNA-only and a combination of DNA and amino acid data. Relevant trees were visually compressed in some figures by opening files generated by RAxML in TreeExplorer (Kumar et al. 2001) and collapsing clades of interest (Figs. 2 and 3). The bootstraps obtained from the DNA-only and combined DNA and amino acid alignments were plotted onto the nodes of the most likely tree obtained (online Appendix 2). A comparison of the results obtained from various matrices using DNA and combined DNA and amino acid alignments is shown in Table 1.

Phylogenetic Informativeness

A measure of phylogenetic informativeness, as proposed by Townsend (2007), was applied to the data at hand (Fig. 4). This analysis uses per-site rate estimates to project the utility of a gene for resolving phylogeny across historical epochs. A calculation of the informativeness per base pair allows for a comparison of different genes and loci used, by providing an estimation of the cost-effectiveness of character sampling for specific time periods. The informativeness was calculated using the most likely tree from the matrix 6G126T, partitioned by gene and codon position. Using this phylogeny, an ultrametric tree was obtained using PATHd8 (Britton et al. 2007), which uses mean path lengths from the leaves of a tree but corrects for deviation from a molecular clock by smoothing substitution rates locally. This ultrametric tree and alignments for the single genes were then analyzed as in Townsend (2007) to obtain the profiles of phylogenetic informativeness for all genes with regard to specific epochs on the phylogenetic tree. All genes were analyzed as DNA for comparative purposes (e.g., nSSU vs. RPB1). The dates used are referred to here
as time units (TUs) in order to emphasize that these are relative time periods and to reflect the uncertainty in the absolute scale of the evolutionary history. Profiles of the informativeness of each gene were plotted with reference to the ultrametric tree in Figure 4.

Character State Coding and Analyses

Character states used in Figure 2 were selected based on constraints in available information and computing resources. The following 4 character groups were chosen.

Ecology.-a) Saprobe, b) animal pathogen, c) insect commensal, d) lichen, e) mycoparasite, f) plant associated: mycorrhiza/plant endophyte/plant pathogen. This coding follows that of James et al. (2006). All 3 specific niches listed under F are listed in online Appendix 3; however, for the purpose of ancestral reconstruction we simplified these character states into "plant associated." This simplification eased the ancestral reconstruction analysis and certainly does not capture the variety of fungal attributes necessary for an interaction with plants.

Lichenization.-a) Lichen and b) nonlichen. In order to investigate whether more resolution was possible with a binary character state reconstruction, all nonlichenized states were combined.

Mature sporocarp.-Coding of sporocarp character states was based on the development and morphology of the mature sporocarp and included Loculoascomycetes which may form open, partially or completely enclosed sporocarps. a) Open (apothecioid): mature ascii presented in an exposed manner on a disc, cup, or clublike sporocarp (Fig. 1A-F); b) partially enclosed (peri-the cioid): mature ascii produced in a partially enclosed,
Figure 3. A complete RAxML Ascomycota tree from 434 taxa with all lineages collapsed to ordinal level where possible. The tree is the same as Figure 2. Classes and subphyla are indicated where supported (subclasses not shown due to space constraints). Suppnode nodes are indicated as in Figure 1. Bootstrap values greater than 50% from 1000 replicates are shown above the nodes. Where no node is supported single species or genera are listed. Nodes with more than 75% bootstrap support were named and numbered in the Leotiomycetes where current ordinal concepts are still poorly supported by molecular data. All shaded areas correspond to single figures in Supplementary data (online Appendix 2).
Ascus dehiscence. · a) Operculate: uniloculate with an operculum (Fig. 1N); b) poricidal: uniloculate with a pore or canal (Fig. 1Q); c) non-poricidal: uniloculate, inoperculate without a pore or canal; d) deliquescent: non forcible discharge (Fig. 1M); e) fissitunicate: bitunicate with "jack-in-the-box" dehiscence (Fig. 1P); f) rostrate: thickened apically, breaking through an outer wall (Fig. 1O). We have largely followed Eriksson’s (1981) coding of ascus types. Operculate, poricidal and fissitunicate follow these definitions closely. Rostrate, semifissitunicate, pseudofissitunicate and bilabiate were combined into one character state, "rostrate." "Non-poricidal" refers to uniloculate asci with no special-ized canal or pore in the ascus tip. "Deliquescent" is applied to all taxa that do not forcibly eject their as-cospores due to the breakdown of the ascus walls, regardless of whether deliquescence occurs early or late in development.

Character states for individual taxa are listed in the Supplementary data in online Appendix 3. Ances-tral reconstructions were performed in Mesquite v2.0 (Maddison w.P. and Maddison D.R. 2008). Character states were traced over 1000 bootstrap trees obtained with RAxML MPI v7.0 (Stamatakis 2006). This reconstruction was performed with a maximum-likelihood criterion using the single-parameter Mk1 model. In order to capture as much of the fungal variation as possible, we chose to use the matrix that contained the most diverse set of taxa (2G434T). In order to be able to use the fast computational advantages offered by using the GTRCAT model approximation, branch lengths were not incorporated into bootstrap trees used for ancestral character state reconstructions. A small subset was subsequently run with 100 bootstrap trees using the GTRGAMMA setting containing branch lengths and yielded congruent results to those presented here. Ancestral states were assigned to a node if the raw likelihood was higher by at least 2 log units than the likelihood value of the other ancestral states according to default settings. Character states traced over 1000 bootstrap trees are shown on well-supported (> 70% bootstrap support) superclass and class nodes in Figure 2. Character states were also mapped using the tree viewing program TreeDyn (Chevenet et al. 2006), shown in Figure 5.

RESULTS

Conflict among Data Sets

We relied on combining and incorporating selections of data sets that were already tested for conflicts in previous analyses (Geiser et al. 2006; James et al. 2006; Miadlikowska et al. 2006; Spatafora et al. 2006; Schoch et al. 2007). Despite this prior filtering, an additional small number of sequences were detected with incon-gruent resolutions. The extensive taxon set used in this study allowed for the increased detection of possible conflicting sequences that could not be achieved with more limited taxon sampling. The following gene se-quences (with AFTOL numbers given in parentheses) were removed-nSSU: Lepidocephalia nicoiiae (1576), Gyromitra esculenta (1534), and Opegrapha varia (881); nLSU: Pyrenula cruenta (386) and Lichinella iodopulchra (896); EF1: Eupenicillium joaanicum (429); RPB1: Pythia cupressita cf (69) and Roccella fuciformis (126); RPB2: Spiromasis varcupii 430; mSSU: Saitoella complicata (229), Cylindrocladus lacinia (1033), Ceramothyrium carniolicum (1063), Usnea sphacelaia (816), Ophiostoma piliferum (910), Capnodi um salicinum (937), Fyxdiphoraavernensis (560), and Neolecta vitellina (1362).

Phylogenetic Analyses

The summarized tree shown in Figure 2 represents the most likely tree obtained from RAxML MPI v7.0 under conditions of combining rDNA and amino acid se-quences. A more detailed tree showing details down to the ordinal level is shown in Figure 3, and the full tree is available in online Appendix 2. The clearest topological difference between the DNA and the combined analyses involved the placement of Geoglossaceae and Candela-lariales (online Appendix 2). DNA analyses placed Geoglossaceae as a sister group to the Lichinomycetes with 85% bootstrap support, whereas the combined analyses placed it as a distinct lineage with no clear affinity to any of the currently defined classes in Ascomycota. This latter finding is consistent with other studies (Wang et al. 2006; Guéidan et al. 2008). Candelariales is placed as a sister group to all other Lecanoromycetes in the DNA analyses (with poor support), but as a separate lin-eage of "Leotiomyeta" in the combined analyses (also with poor support). A previous paper (Miadlikowska
et al. 2006) placed the order as an early diverging lineage within Lecanoromycetes.

To determine the role played by missing data and taxon sampling on superclass and class nodes, the other more focused data sets (3G409T, 4G335T, 5G234T, and 6G126T), which contained fewer taxa and a lower percentage of missing characters, were run in RAxML with 1000 bootstraps for both DNA and combined (DNA and amino acid) alignments. The bootstrap values for the superclass and class nodes highlighted in Figures 2 and 3 are compared in Table 1. Most nodes produced similar patterns of support, regardless of the data set analyzed, but 3 clades in particular—Taphrinomycotina, Eurotiomycetes, and Lecanoromycetes—received differing patterns of support across the data sets (Table 1). Taphrinomycotina received strong support in all analyses except for 6G126T, which has the fewest taxa and least amount of missing data. The remaining Taphrinomycotina taxa in 6G126T were Schizosaccharomyces pambe and 2 members of Taphrinomycetes (Taphrinales), which is also resolved poorly in the larger matrices suggesting a significant role for taxon sampling. Similarly, Eurotiomycetes displayed higher levels of bootstrap support in data sets with increased taxon sampling.
although these data sets were characterized by higher amounts of missing data. A significant decrease in bootstrap support between matrix 4G335T and 5G234T is noted for Lecanoromycetes (node 5, Fig. 2; Fig. 3). This decrease is mainly due to the instability for a single taxon representing Rhizocarpales, *Rhizocarpon oederi*. This shift is likely caused by the removal of the second taxon representing this order, *Rhizocarpon superficiale* in matrix 5G234T. The same trend is noted in the DNA-only analysis. The bootstrap values for both alignments again show an increase in the last matrix, 6G126T, due the absence of *R. oederi*. A difference in behavior is noted for Leotiomycetes (node 6, Fig. 2; Fig. 3), where combined alignments (DNA plus amino acids) have low bootstrap proportions, but the DNA-only analysis yielded 100% bootstraps in all cases. A few other examples, not related to the nodes shown in Table I, did receive higher values when fewer taxa were analyzed.

**Phylogenetic Informativeness**

To visualize the power applied here toward resolution of phylogeny in specific time spans, we profiled the phylogenetic informativeness of each gene as described in Townsend (2007). The profiles are derived from the rates of evolution of sites within genes. On a per-site basis, the gene fragments of *RPB1*, *RPB2*, and *TEF* all produced a pulse of high informativeness across relatively older dates more recent than 3.5 TUs (Fig. 4). The informativeness of these genes diminishes across relatively older dates than 5 TUs. The genes mSSU and nLSU demonstrate steady levels of informativeness over this history, whereas nSSU has a slow increase in utility with the depth of analysis. Net phylogenetic informativeness, which takes into account sequence length, produced a very similar profile for the 6 genes with the exception that, due to its longer length in this data set, *RPB2* was identified as the most informative gene. Regardless, protein-coding data were inferred to be more phylogenetically informative than rDNA genes across much, if not all, of the tree space analyzed here.

**Overview of Superclass Nodes**

The phylograms in Figures 2 and 3, as well as online Appendix 2, produced from data set 2G434T represent the most diverse and complete Ascomycota phylogenetic analysis performed to date. These span all 15 currently recognized classes and include representatives for 90% of the currently recognized orders in the AFTOL classification (Hibbett et al. 2007). As in previous papers, we continue to find support for numerous backbone nodes in Ascomycota. The early diverging lineages in the subclass Taphrinomycota continue to be supported with bootstrap values >70%. Support is also found for the subclasses Saccharomycotina and Pezizomycotina, as well as their sister-group relationship. The superclass Leotiomycetes (Eriksson and Winka 1997) that was not accepted as part of the recently proposed classification (Hibbett et al. 2007) is well supported (node E, Fig. 2; Fig. 3), as is the relationship between Arthoniomycetes and Dothideomycetes (Fig. 3). For the first time, the insect symbiont class Laboulbeniomycetes can be placed with bootstrap support as a sister to Sordariomycetes. Leotiomycetes is well supported in the DNA analyses, but only weakly supported in the combined analyses. We find support for a clade of poricidal, unitunicate taxa that includes Sordariomycetes, Laboulbeniomycetes, and Leotiomycetes, which is in agreement with phylogenomic studies and previous analyses (Lumbsch et al. 2005; Robbertse et al. 2006; Spatafora et al. 2006).

**Ancestral Character State Reconstructions**

Ancestral characters are shown on class and superclass nodes and are presented as pie charts representing the character state for which a certain percentage of bootstrap trees were recovered with that state (Fig. 2). The total values of pie charts are the number of bootstrap trees with that node present out of a possible 1000 trees. A number of nodes that were significantly supported by bootstraps are highlighted, but small classes with only a few representative taxa are not indicated on Figure 2 for purposes of clarity. In order
to indicate character state changes occurring below the class level, we also plotted the same character set on a set of circular trees (Fig. 5; online Appendix 2, Supplementary Fig. 8). The ecological characters showed support for a saprobie or nonlichenized character state as ancestral to the majority of superclass nodes. This support was also significant for the ancestral node of Eurotiomycetes and “Dothideomycetes,” which contain multiple lichenized and nonlichenized lineages. Can- delariales, Lecanoromycetes, and Lichinomycetes are predicted to have a lichenized ancestor. We chose to use the well-supported node subtending Acarosporales to denote Lecanoromycetes because Candelariales was unstable and its placement was not supported in our analyses. A more simplified binary coding where all taxa were coded as either lichenized or nonlichenized yielded the same conclusion as above. These analyses reconstructed a total of 4-7 gains and 1-2 losses of lich- enization across Pezizomycotina (Fig. 5).

Two independent origins of sporocarp production, 1 for Neoleotiomycetes and 1 for Pezizomycotina, were predicted by these analyses. An apotheciod sporocarp with an exposed hymenium is reconstructed as the ancestral state for Pezizomycotina with multiple derivations of perithecioid and cleistothecioid sporocarp (online Appendix 2, Supplementary Fig. 8). As- cusc dehiscence was resolved for class-level nodes, but analyses resulted in equivocal reconstructions for most superclass nodes. Several character state transitions below class level are not shown in Figure 2. Eurotiomycetes and, to a lesser extent, Dothideomycetes and Lecanoromycetes underwent several transitions in these characters (online Appendix 2, Supplementary Fig. 6C, 8).

DISCUSSION

Phylogenetic Informativeness of the 6 Core AFTOL Genes

Profiling of the phylogenetic informativeness of individual genes (Townsend 2007) has enabled us to assess the overall utility in Ascomycota systematics of genes used in the AFTOL combined matrices. These analyses support a conclusion that greater levels of phylogenetic informativeness are provided by protein-coding genes RPB1, RPB2, and TEF1. This finding is consistent with recent studies where the introduction of protein data provided critical support for numerous nodes that were either not resolved or not robustly supported by ribosomal data (Geiser et al. 2006; Schoch et al. 2006; Spatafora et al. 2006; Hofstetter et al. 2007). As can be seen in Figure 4, in our data set, RPB2 is the highest contributor to net informativeness, partly due its length (average 1400 bp) in comparison with other protein genes. How- ever, RPB1 possesses the highest amount of per-site informativeness (net informativeness divided by gene length). The per-site informativeness of mSSU in the deepest nodes of the tree (more than 5 TU) ranks first in these analyses, but its net informativeness is somewhat diminished due to its short length (average 500 bp). Importantly, the net and per-site informativeness of mSSU is greater than that of both nSSU and nLSU, 2 standard loci used in analyses of the molecular phylogenetics of fungi. The informativeness of nSSU was the lowest of all 6 loci, a finding that is especially illuminating given the fact that it has been a major component of large-scale fungal phylogenies during the past 2 decades. The same may be said for nLSU, albeit with a slightly higher informativeness profile. These results support those of a previous study by Hofstetter et al. (2007), which demonstrated that a combination of RPB1, RPB2, and mSSU provided the best definition for most nodes in a large-scale tree of Lecanoromycetes.

The slopes of curves representing relative informativeness over time were different for rDNA compared with protein-coding genes. The 3 protein-coding genes, especially RPB1 and RPB2, exhibited marked decreases in informativeness past relatively ancient projected dates. Conversely, the rDNA genes showed a steady, more uniform level of informativeness through time. These results allow us to predict that although expanded sampling of RPB1, RPB2, TEF1 and other protein-coding genes in fungal systematics will have significant positive impacts on our ability to resolve problematic relationships, phylogenetic resolution of more ancient phylogenetic relationships among the Kingdom Fungi will likely require large-scale analyses of considerably more genes than those employed here. Also, we do not interpret these results as reflecting a wholesale dismissal of using rDNA gene sequences in fungal systematics. In fact, these genes have already contributed much to fungal systematics through the application of universal primers and their relative ease of amplification (Blackwell et al. 2006). They continue to remain important as potential sources of data in environmental sampling (Schadt et al. 2003; Porter et al. 2008) and for bar coding in fungi (Summerbell et al. 2005) and may yet contribute significant, valuable information for resolution of ancient relationships where all genes show diminished levels of phylogenetic informativeness.

Ancestral Character-State Reconstruction

The mapping of ecological and morphological charac- ters on a set of RAxML bootstrap trees allowed us to take uncertainty into account when predicting ancestral state reconstruction. Also, the Ascomycota taxon set analyzed here provides an opportunity to conduct phylum-wide ancestral character state reconstructions on the largest taxon set to date. We extend established ecological defi- nitions in order to provide continuity with the previous paper of James et al. (2006). We also include the morphological concepts of sporocarps that have formed the cornerstone of most hypotheses in higher-level As- comycota systematics (Nannfeldt 1932; Luttrell 1955; Henssen and Jahns 1974), often together with distinc- tions of ascus morphology and dehiscence (Luttrell 1951; Eriksson 1981; Hafellner 1988). The support for ancestral reconstruction on the 4 sets of character states in Figure 2 is conservative, given the use of maximum-likelihood bootstraps. As expected, a number of deep
nodes were recovered with equivocal character states, highlighting that certain classes (e.g., Eurotiomycetes and Dothideomycetes) are especially prone to shifts in ecological and morphological character states (Geiser et al. 2006; Schoch et al. 2006).

To facilitate the discussion of results, we follow Eriksson and Winka (1997) and Lumbsch et al. (2005) in proposing informal names to nodes above the class level with superclass extensions. We propose these as "rankless taxa" within the current hierarchical ranked classification. Thus, the node combining Dothideomycetes and Arthoniomycetes is proposed as Dothideomyceta (node F, Fig. 2). Leotiomycetes, Sordariomycetes, and Laboulbeniomycetes comprise "Sordariomyceta" (node G, Fig. 2), and Saccharomycotina and Pezizomycotina constitute "Saccharomyceta" (node C, Fig. 2). The remaining superclass nodes are either reported here for the first time with support or remain unresolved, and thus unnamed.

Sporocarp morphology.—Certain sporocarp morphologies (e.g., cleistothecia) have been demonstrated as homoplastic for Ascomycota (Berbee and Taylor 1992; Suh and Blackwell 1999; Stchigel and Guarro 2007). Other studies have postulated an apothecial ancestor for Ascomycota, but without explicit ancestral character state reconstructions (Gernandt et al. 2001; Spatafora et al. 2006). Here, an open sporocarp with an exposed hymenium (apothecial sporocarp) is recovered in more than 70% of bootstrap trees as the ancestral state for the most recent common ancestor of Pezizomycotina, as well as the "Leotiomyceta", "Sordariomyceta", and "Dothideomyceta", but with more uncertainty for the latter (Fig. 2). The reconstruction in Figure 2 and online Appendix 2 supports multiple transitions from apothecial sporocarps to partially enclosed (perithecioid) and completely enclosed (cleistothecioid) sporocarps. Independent origins of perithecioid sporocarps include common ancestors of Dothideomycetes, Sordariomycetes plus Laboulbeniomycetes, Chaetothyriomycetidae, and Eurotiomycetidae (Eurotiomycetes) (online Appendix 2, Supplementary Fig. 8), as well as Thelenniaceae, Thrombiaceae, and Protothelenniaceae in Leconomycetes (Schmitt et al. 2005) and the Orbicipa group in Pezizomycetes (Hansen et al. 2005). Importantly, this is the first study to confidently place Laboulbeniomycetes as an enigmatic lineage of insect symbionts and mycoparasites that have long proved problematic with respect to placement in higher-level classification schemes. The data and analyses presented here strongly support Laboulbeniomycetes as being a sister group to Sordariomycetes. We can now confidently interpret the reduced sporocarp of Pyxidiophorales as homologous to the perithecium of Sordariomycetes and the sporocarp of Laboulbeniales (Fig. 1A) as a highly derived perithecium that develops directly from an ascospore, a finding that conforms with the terminology originally applied to this group (Thaxter 1896). Independent origins of cleistothecioid sporocarps include Eurotiales and Onygenales of Eurotiomycetes, Pseudozum and Leuconeurospora and Erysiphales of Leotiomycetes, Preussia of Dothideomycetes, and multiple derivations within the mainly perithecioid Sordariomycetes.

Neoleut, the sole known sporocarp-producing taxon of Taphrinomyomycota, also produces apothecioid sporocarps. Taphrinomyomycota is characterized by a wide range of biochemical and ecological variation. Most species have a yeast growth phase (e.g., Schizosaccharomyces and Pneumocystis; Sugiyama 1998), but mycelial growth is also known for some species (e.g., Taphrina), and with the exception of Neoleuca (Landvik 1996), they lack sporocarps. Although Neoleuca produces club-shaped apothecia (Fig. 1A), it also has several presumably ancestral features, such as repeated branching of hyphae that produce and bear simplified nonporicidal asci and the absence of distinctive ascus precursor cells (croziers) in which meiosis occurs (Redhead 1976; Landvik et al. 2003). The monophyly of Taphrinomyomycota has been controversial with most previous studies based on rDNA genes resolving the group as a paraphyletic assemblage at the base of the Ascomycota tree (Sugiyama 1998; Sugiyama et al. 2006). Recent studies that incorporated multiple protein-coding genes, however, resolved the group as monophyletic with varying levels of support (James et al. 2006; Liu et al. 2006; Spatafora et al. 2006). Here, with additional sampling of both taxa and genes we find support for the monophyly of Taphrinomyomycota, and thus demonstrate that the earliest diverging clade of Ascomycota includes both filamentous, sporocarp-producing species and yeast growth forms.

These findings provide statistical support for one of the older hypotheses of sporocarp evolution (Nannfeldt 1932), where perithecioid and cleistothecioid sporocarps were thought to be derived through modifications and reductions of an apothecioid ancestor. Furthermore, we expand this hypothesis by proposing that the ability to produce sporocarps has arisen twice during the evolution of Ascomycota, once in the common ancestor of Pezizomycotina and once in Neoleuca.
asci. The remaining classes of Pezizomycotina correspond to "Leotiomyceta" and collectively represent a diversity of inoperculate asci (e.g., fissitunicate, poricidal, and deliquescent). Therefore, the nodes immediate to the common ancestor of Pezizomycotina, including Saccharomycetes reconstructed as deliquescent asci, are characterized by different ascus types. This finding is consistent with ascus evolution being a more dynamic process early in Ascomycota evolution, where the invention of the major ascus types preceded the diversification of most modern classes.

"Leotiomyceta" contains the majority of species within Pezizomycotina, and these analyses shed light on the evolution of fissitunicate and unitunicate asci, 2 of the major inoperculate ascus types present within the taxon. Jack-in-the-box or fissitunicate asci are restricted to "Dothideomyceta" and Eurotiomycetes, and whereas "Dothideomyceta" is exclusively fissitunicate, the Eurotiomycetes contain a diversity of ascus types including fissitunicate, rostrate and deliquescent. Although ancestral character states could not be assigned for the most basal nodes of the Eurotiomycetes, the ancestor of the Eurotiomycetidae and Chaethothyriomycetidae is resolved as fissitunicate (Fig. 2). Within the Eurotiomycetes, Coryneliales produce unique fissitunicate asci in which the outer ascus wall layer breaks early in development, leaving only a remnant near the base of the ascs; the inner wall deliquesces at ascospore maturity (Johnston and Minter 1989). Eurotiales and Onygenales, which includes a number of important human pathogens (e.g., <em>Aspergillus</em> and <em>Coccidioides</em>), produce thin-walled, deliquescent asci. The close relationship between these orders and the ancestral character state reconstruction of the subclass support the derivation of eurotiallean prototunicate, deliquescent asci (Fig. 1M) from a fissitunicate ancestor (Fig. 1P) (Geiser et al. 2006) and not from the modification of a single-walled, unitunicate ancestor (online Appendix 2, Supplementary Fig.6C).

"Sordariomyceta" includes important plant pathogens (e.g., <em>Sclerotinia</em>, Leotiomycetes) and model eukaryotic systems (e.g., <em>Neurospora</em>, Sordariomycetes). Although Leotiomycetes is characterized by the production of apothecial sporocarps and Sordariomycetes by perithecioid sporocarps, "Sordariomyceta" classes collectively contain the majority of species that produce unitunicate, relatively thin-walled, picroidial asci. Examples of derived deliquescent asci are also present in both classes and are typically correlated with the independent derivations of cleistothecia from either apothecia (Leotiomycetes) or perithecia (Sordariomycetes), or the ecologies of insect dispersal of ascospores (Laboulbeniomycetes and Sordariomycetes), or fruiting in a marine environment (Sordariomycetes). Although polarities in character state transitions were not imposed in the reconstructions performed here, it is recognized in several distantly related lineages that deliquescent asci are derived from asci with persistent cell walls and forcible discharge of ascospores (e.g., Berbee and Taylor 1992; Blackwell 1994). The sampling of numerous lineages with deliquescent asci, and strictly asexual species with no ascus production, affected ascus character state reconstruction and resulted in a large number of bootstrap trees with equivocal reconstructions.

Lichenization.-One of the first Ascomycota molecular phylogenetic studies of nSSU and nLSU rDNA resulted in a general hypothesis of independent origins of lichenization (Gargas et al. 1995). More complete taxon sampling in Ascomycota (Lutzoni et al. 2001) provided support for lichenization occurring early in Ascomycota evolution and suggested that some extant non-lichenized lineages (e.g., Eurotiales) are derived from lichenized ancestors. The presence of a 600-million-year-old fossil lichen associated with cyanobacteria (Yuan et al. 2005) is consistent with an ancient origin for a lichenized state in fungi and is consistent with "Protolichenes hypothesis" (Eriksson 2005), which suggested that fungi existed as lichens in terrestrial palaeosystems prior to the diversification of land plants. It is noteworthy, however, that lichens remain unreported in extant lineages of Taphrinomycotina, Saccharomycotina, and the earliest diverging Pezizomycotina classes (Orbiliomycetes and Pezizomyctina).

Based on our taxon sampling, the ancestral ecologies of "Ascomycota", "Saccharomyceta", Saccharomycotina, and Pezizomycotina are reconstructed as saprophic or nonlichenized. All lichenized taxa are restricted to "Leotiomyceta" and are members of Candelariales, Lecanoromycetes, "Dothideomyceta", Eurotiomycetes, and Lichinomycetes. The ancestral character reconstruction of lichenization is complicated, however, by the lack of resolution among the major superclass nodes of "Leotiomyceta" and the coding of nonlichenized taxa. Although arguments can be made in favor of both multistate and binary codings, many species are difficult to be definitively coded in the multistate approach. For example, many species of Chaetothyriales (e.g., <em>Exophiala</em>) and Eurotiales (e.g., <em>Aspergillus</em>) are opportunistic animal pathogens but undoubtedly exist as soil saprobes. Regardless of coding, however, most class and some superclass-Level nodes were reconstructed with saprobic/nonlichenized ancestral character ecologies including Geoglossaceae, <em>Symbiophrina</em>, Eurotiomycetes, "Dothideomyceta", and "Sordariomyceta". Lichenized ancestors were predicted for Lecanoromycetes, Lichinomycetes, and Candelariales. A definitive estimate of the number of gains of lichenization is tenuous, due to the poorly resolved internal nodes of "Leotiomyceta" including Lichinomycetes and Candelariales. Despite these caveats, a conservative interpretation of our data remains that lichenization evolved multiple times in Ascomycota including once in the ancestor of Lecanoromycetes, 1-2 times within Eurotiomycetes, and at least twice within "Dothideomyceta". Lichinomycetes and Candelariales may either represent independent origins or part of Lecanoromycetes lichenization event (Miadlikowska et al. 2006). Unequivocal losses of lichenization occurred.
in Ostropales of Lecanoromycetes, and there was possibly one loss within Eurotiomycetes resulting in Chaetothyriales (Fig. 5).

Eurotiomycetes and "Dothideomyceta" were the only 2 nodes that showed different proportions of character state reconstruction based on the binary and multistate codings. The node subtending the Eurotiomycetes was absent in 19% of bootstrap trees, and multistate coding recovered ancestral states 13% saprobic and 70% equivocal. The binary coding overwhelmingly supported a nonlichenized state (83%). Similarly, for "Dothideomyceta", that node was absent in 16% of bootstrap trees. For the remainder, multistate coding recovered ratios of 28% saprobic, 27% lichenized, and 29% equivocal, whereas the binary coding recovered 50% nonlichenized and 29% equivocal. Furthermore, it is within Eurotiomycetes that additional taxon sampling detected 3 nonlichenized lineages (plant pathogenic Coryneliales, marine saprobic Dactylospora-clade, and lichen parasitic Mycocaliciomycetidae) and significantly supported the reconstruction of a nonlichenized ancestor when coded as nonlichenized. Within Chaetothyriomycetes, however, we could not distinguish between 2 parallel gains of lichenization (Pyrenulales and Verrucariales) and a single gain and one loss (Chaetothyriales) (online Appendix 2, Supplementary Fig. 6C), a finding in agreement with Gueidan et al. (2008). These data, then, do not support the hypothesis that lichenization represents an ancestral state for the majority of extant lineages of filamentous Ascomycota (Lutzoni et al. 2001).

Summary

We have presented a complete class-level tree of Ascomycota obtained from data representing 6 gene regions generated by AFTOL. An assessment of the various loci confirmed that the addition of protein-coding genes had a major impact on the phylogenetic informativeness in our data set and significantly improved rDNA phylogenies. The production of multigene data sets for this study, and those preceding it, has allowed for the resolution of several class-level nodes of Ascomycota and provided us with a reliable template to reassess fungal evolution. Ancestral character state reconstruction supports the hypothesis that the common ancestor of Pezizomycotina was characterized by production of apothecoid ascomata and saprobic/riparian lichenized ecology. Lichenization was gained multiple times within "Leotiomyceta" with losses occurring more rarely and generally restricted to more terminal clades. Importantly, these results are consistent with 2 independent origins of sporocarp production within Ascomycota, once in Pezizomycotina and once in Neoleotiomycotina.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at http://www.sysbio.oxfordjournals.org/.

FUNDING

We acknowledge funding from the National Science Foundation for "Assembling the Fungal Tree of Life" (DEB-0228725 and DEB-0732993) and the "Deep Hypha" research coordination network (DEB-0909301). Funding from the Academy of Finland (No.211172) to S.S. is also acknowledged.

ACKNOWLEDGMENTS

We thank J. K. Stone, O. E. Eriksson, C. Printzen, I. Brodo, S. Joneson, and M. Berbee for helpful comments and material provided.

REFERENCES


