10 Fungal Molecular Evolution: Gene Trees and Geologic Time

MARY L. BERBEE and JOHN W. TAYLOR

CONTENTS

I. Introduction ................................................................. 229
II. Genes for Phylogenies: Congruence and Conflict .......... 229
   A. Nuclear Ribosomal Repeat Regions ..................... 229
   B. When are Ribosomal Genes not Enough? ............. 230
   C. Protein-Coding Genes ............................................. 231
   A. Fungal Time Tree Under Molecular Clock Assumptions ........ 233
   B. Origin of the Fungi .................................................. 239
      1. Flagella and the Basal Fungi .............................. 240
      2. Colonization of Land ........................................ 240
   C. Radiation of the Terrestrial Fungi ...................... 241
      1. Euascomycetes .................................................... 241
   D. Basidiomycete Radiation ....................................... 243
IV. Summary ................................................................. 243
    References ............................................................. 243

I. Introduction

Fungal phylogenetics has always been based on characters, but technological and intellectual advances are introducing new kinds of characters and new ways of thinking about them. First light microscopy, then electron microscopy, and now DNA sequencing successively upset previous views of fungal relationships. Phenetics, cladistics, and computerized data analysis and phylogenetic tree generation are now changing the intellectual rules for taxonomy and phylogenetics. The combination of new characters and new analytical tools have supported some taxonomic groups, established some new ones, and demolished a few old ones.

In this chapter, we do not discuss methods of phylogenetic analysis, but we will discuss some of the ribosomal and protein-coding DNA characters that are increasingly important in fungal phylogenetics. For a discussion of phylogenetic methods applied to fungi, see Berbee and Taylor (1999). To show the strengths and limitations of 18S rRNA sequences for phylogenetic analysis, we analyze and discuss trees from 49 diverse species of fungi and 3 animal outgroups. We consider the implications of the tree and nucleotide substitution rates for the timing of fungal divergences.

II. Genes for Phylogenies: Congruence and Conflict

A. Nuclear Ribosomal Repeat Regions

The most commonly used DNA for fungal phylogenetic studies is located in the ribosomal DNA repeat regions of the nuclei. The ribosomal repeats have been regions of choice for several good reasons. Fungal ribosomal gene clusters are arranged in roughly 200 tandem repeats (Butler and Metzenberg 1989). Because each nucleus contains 200 or so identical copies of the region, at least one intact copy for molecular analysis can usually be recovered, even from low-quality DNA preparations. The rDNA regions permit phylogenetic comparisons at many taxonomic levels (Bruns et al. 1991). Within the rDNA repeats are three coding genes (18S, 5.8S, and 28S) and two spacers (ITS 1 and ITS 2) that are transcribed together as a single unit and then cleaved into separate rRNA products. The rRNAs encoded by the 18S, 5.8S, and 28S genes become structural parts of the ribosomes and are essential for protein synthesis. The genes are highly conserved and are universally present, permitting phylogenetic comparisons among domains and kingdoms as well as within the more restricted group, the Eumycota. Compared with the genes,
the spacers ITS1 and ITS2 are much freer to vary because their transcripts are excised from the rRNA and discarded instead of becoming part of the ribosome. The ITS regions are useful for taxonomic comparisons of closely related species and genera. A third spacer, the intergenic spacer or IGS, is not transcribed. It lies between pairs of repeats and is highly variable. In basidiomycetes, an additional RNA gene, the 5S rRNA gene, is within the IGS. In other fungi, the 5S gene is usually elsewhere in the genome (Metzenberg et al. 1985).

Why have phylogenetic studies of major fungal taxa based on rRNA been successful? Part of the reason lies in the ease of primer design for the rRNA regions. Nucleotides defining rRNA function seem to be distributed in highly conserved patches. Patches of 20 or more consecutive, highly conserved nucleotides are ideal for designing primers that will amplify DNAs from diverse organisms. Universal 18S rRNA gene primers (White et al. 1990), for example, amplify DNAs from insects as well as from red algae.

Also in favor of ribosomal genes for tree building is the fact that gene trees and species trees are likely to be the same, at least for distantly related species. Heterogeneity within species is usually slight. Among the ITS regions of 25 clinical isolates of *Coccidioides immitis*, for example, Burt et al. (1996) found a single polymorphic site. The ribosomal genes undergo concerted evolution so that sequence heterogeneity among repeats within a nucleus is rare. Horizontal transfer of ribosomal genes is unknown.

Sometimes the most compelling reason for choosing rRNA genes for a phylogenetic study is that the many sequences already in the molecular genetic databases can be reused to answer new questions. Sequencing ribosomal genes for just a few ingroup taxa while drawing the remaining ingroups and outgroups from national databases can answer many questions with minimum effort. In October 1997, GenBank contained over 5400 fungal rRNA sequences, accessible through the National Center for Biotechnology Computing (http://www.ncbi.nlm.nih.gov/). We took advantage of 52 of these sequences, from the 1997 GenBank pool to infer the trees in Figs. 2–5, discussed later in this chapter.

**B. When are Ribosomal Genes not Enough?**

Resolution of relationships among the fungi might be better if complete sequences were available from all three rRNA genes from the ribosomal repeats. However, this situation is likely to remain hypothetical for some time because DNA sequencing is time-consuming and expensive, so that complete sequences are generally available for the shorter 5.8S and 18S genes, but not the longer 28S genes.

Even if complete rRNA sequences were available for all fungi, the ribosomal repeats would not answer all questions about relationships. Although the number of nucleotides in the ribosomal repeats is substantial (about 9000), the number of possible characters useful for a specific question is often small. Imagine that the repeats are being used for a question about an ancient divergence in the fungi. Of the 9000 nucleotides in a repeat, the 18S gene accounts for about 1800bp, the 5.8S gene for about 120bp, and the 28S gene for about 3200bp. The remaining nucleotides in the IGS and ITS are highly variable and change too quickly to record ancient divergences. That leaves about 5100 conserved positions in the RNA genes. Of the 5100 or so conserved positions, most do not vary among the fungi, presumably because they are functionally important and mutations in these positions are lethal. From the 18S data set we describe later in this chapter, PAUP 4.0d55 estimated that 56% of the nucleotide positions were not free to vary and so could not have undergone substitution-useful in phylogenetic reconstruction. Assuming that the 28S and 5.8S genes have similar substitution rates, this leaves about 2225 nucleotides in the rRNA genes that could potentially record ancient fungal divergences. Of course, over short periods of time between divergences in a rapid radiation, only a few of these 2225 positions will record the event by substituting nucleotides (Philippe et al. 1994). Therefore, there is every reason to look beyond the ribosomal genes and spacers to the rest of the genome. Fortunately, the whole genome of a fungus like *Neurospora crassa* contains about 47,000,000 nucleotides (Orbach et al. 1988). If lack of resolution of branching order is due to insufficient data, there is a good chance of finding a record of radiation outside the ribosomal repeats.

Worse than lacking information, genes including the rRNA genes can be sources of phylogenetic misinformation. Occasionally, ribosomal gene sequences evolve at very different rates in different species. If the rate differences is extreme, long branches "attract" one another and quickly evolving taxa group together or with distantly related outgroups rather than with their slowly evolving relatives (Felsenstein 1978). A clear
demonstration of the phenomenon comes from the 18S sequences of the insect world. Flies are close relatives of fleas and scorpionflies based on strong evidence from comparative morphology. However, because their sequences evolved unusually rapidly, flies group with outgroups or with long branched ingroups instead of with their flea and scorpionfly relatives (Carmean and Crespie 1995; Carmean et al. 1992; Huelsenbeck 1997).

Long branch attraction might also explain conflicting positions of slime molds and microsporidia in 18S rRNA and protein coding gene trees. Is it long branch attraction or the true phylogeny that positions the divergence of the slime molds down among the first protists in 18S rRNA trees (Cavalier-Smith 1993)? Actin and tubulin gene trees point to a more recent origin for the slime molds, near the time of origin of animals and plants (Baldauf and Palmer 1993). Although in the 18S gene trees the microsporidia appeared to be among the first eukaryotes to diverge, this was probably an artifact of long branch attraction. Trees from three other genes, from α- and β-tubulin (Keeling and Doolittle 1996; Keeling et al. 2000) and from the large subunit of RNA polymerase II (Hirt et al. 1999) all link microsporidia to the fungi.

As in morphological phylogenetics, comparing homologous characters is important in molecular phylogenetics. Because the rRNA genes are prone to insertions and deletions, the homologous patches can be difficult to recognize in distantly related species. The secondary structure of the rRNA is more highly conserved than the nucleotide sequence and can be a useful guide to identifying and aligning homologous regions. The rRNA World Wide Web (WWW) server (http://rrna.uia.ac.be/index.html) provides a universal alignment, based on secondary structure, for all available complete 18S and 28S sequences (Van de Peer et al. 1997). As an example from the universal alignment showing the effect of the history of added and subtracted nucleotides, the 18S sequence of Alternaria alternata required twice as many gaps as nucleotides to fit in with the universal alignment of diverse prokaryotes and eukaryotes.

Even with secondary structure as a guide, decisions about alignment can be subjective. Authors should make their alignments available to other researchers who can then challenge the initial results with a new alignment and fresh analysis. TreeBASE (http://www.herbaria.harvard.edu/treebase/index.html), not limited to rRNA, is being developed as a repository for alignments and trees from published reports. Since 1998, the journal Mycologia has required that authors of phylogenetic studies submit alignments and trees to TreeBASE before publication.

In general, one of the advantages to working with ribosomal genes over protein genes is that all the copies of a ribosomal gene from the same nucleus are identical or nearly identical. However, as occasionally reported for 5S and ITS regions, copies of the repeat can differ significantly in the same organism (O’Donnell 1992; O’Donnell and Cigelnik 1997). If the genes are polymorphic within an individual, then the same complications could arise as when inferring phylogeny from paralogous protein-coding genes.

C. Protein-Coding Genes

Protein sequences offer several advantages over rRNA gene sequences for phylogenetic analysis. Homology and convergence are easier to recognize in protein sequences made up of the 20 amino acids than in DNA sequences of the four nucleotides. Length changes are infrequent in protein-coding genes because insertions and deletions often lead to fatal frame shifts and elimination through natural selection. Eukaryotic genomes offer a wide range of protein-coding genes for phylogenetic analysis. However, studies of fungal phylogenetics based on protein-coding genes must surmount some minor obstacles beginning with primer design. Because selection acts mainly at the amino acid level in protein-coding genes, different DNA triplets can code for the same amino acid. A primer that works for one fungus may fail with its close relatives due to substitutions that destroy the priming site in the DNA without changing the amino acid sequence. The priming problem can sometimes be solved by using degenerate primers consisting of sequence mixtures that complement several different target template sequences, and by designing the primers so that their 3’ ends match the more conservative first and second codon positions.

Like nuclear rRNA sequences, mitochondria and their protein-coding genes are present in multiple copies in most cells. Given the right primers, mitochondrial genes can be easy to amplify, even from low-quality starting DNA. However, to date, relatively few fungal phylogenetic studies use mitochondrial protein-coding genes. Pagquin et al. (1995a, b) inferred plausible trees from the mito-
ochondrial genes coding for subunits 1 and 3 of cytochrome oxidase and for subunit 5 of NADH dehydrogenase. More recently, Paquin et al. (1997) used subunits 1–3 and cytochrome b in a phylogenetic analysis to infer strong support for animals and fungi as sister taxa. Within the fungi, their analysis shows Spizellomycesales rather than Blasto cladiales at the base of the fungi, and relationships at the base of the Ascomycota are no better resolved with the mtDNA genes than those inferred from ribosomal genes. Other analyses by Paquin et al. (1997) of entire mtDNA sequences give a fascinating view into the evolution of mt introns, codon usage, and the interplay of nuclear and mitochondrial genes.

Nuclear protein-coding genes are usually present in one copy per nucleus, so recovering enough intact copies to amplify requires good quality genomic DNA. Only a few studies of fungal relationships used single copy nuclear genes (Smith 1989; Radford 1993; Keeling et al. 2000; Liu et al. 1999). Baldauf and Palmer (1993) used highly conserved nuclear genes for actin, α- and β-tubulin, histones, and for the translation elongation factor Ef-1α to show that animals and fungi appear as sister groups (Baldauf and Palmer 1993). Analyzing a total of 53 different enzyme encoding genes from the international genetic data bases, Doolittle et al. (1996) inferred that fungi diverged from animals about 965 Ma ago.

The genes used in the above studies were chosen in part because they did not appear to have undergone horizontal transfer or frequent gene duplication. Genes that do duplicate to form gene families require extra sequencing effort. The phylogenetic analysis of chitin synthase genes by Bowen et al. (1992) provides a good example of the technical steps needed to survey members of a gene family from several fungi. A number of genes code for chitin synthase in the Euascomycetes (Bowen et al. 1992; Din et al. 1996). Bowen et al. (1992) designed degenerate primers to amplify chitin synthase genes belonging to three gene families from a wide range of Ascomycota. The amplification product from a single fungus was usually a mixture of chitin synthases from two or three of the families. To separate genes from the three families, the polymerase chain reaction (PCR) products were cloned and sorted based on similarity. Clones representing each family were sequenced and a tree of the sequences was inferred using UPGMA. The chitin synthase gene must have duplicated twice before the Euascomycetes radiated, because several of the Euascomycetes species had three chitin synthase genes and, within each gene family, the same phylogeny was repeated (Fig. 1). In this case, interpreting organismal evolution from the gene tree simply involved following the branching pattern within each chitin synthase gene family. Interpreting organismal evolution from genes trees can be much more difficult when the pattern of gene duplication and loss is more complex.

Gene trees and organismal trees conflict when horizontal transfer occurs and unrelated organisms exchange genetic material. As an example of the consequences of horizontal transfer, flowering plant genes for the beta subunit of the F1-ATPase are polyphyletic even though flowering plants clearly have a single origin (Iwabe et al. 1989). In the F1-ATPase tree, the carrot clusters with the bread mold *Neurospora* but the tobacco and spinach cluster with bacteria. The explanation? The *Neurospora* and carrot genes were from the nucleus and had been vertically transmitted. The tobacco and spinach genes were from chloroplasts. Plants gained their chloroplast genomes through a horizontal transfer event involving symbiosis. An ancestral plant captured and enslaved the cyanobacterium that became the first chloroplast. The chloroplast genes retain the traces of their cyanobacterial origin, clustering with *E. coli* in gene trees rather than grouping with eukaryotes. The phylogeny showing the polyphyletic plants is correct for the F1-ATPase gene but would be wrong if generalized to the whole organisms.

III. Inferring Phylogeny and Timing of Fungal Divergences Using 18S rRNA Gene Sequences

Which fungal relationships can be resolved with confidence using 18S rRNA sequence data? Which regions of the phylogeny of major taxa might benefit most from investigation with other genes? To provide a sense of the strengths and limits of tree building from the 18S genes, we inferred the phylogeny for 49 fungi and 3 outgroups (Table 1) using parsimony and the computer program PAUP 4.0d55. We added sequences from a broad selection of species from the Ascomycota and Basidiomycota, from the few chytrids available in the databases and from zygomycetes including
endomycorrhizal species and animal-associated trichomycetes and Entomophthorales (Table 1). The outgroups include one collar flagellate (Choanoflagellate) (Wainwright et al. 1993) and two members of a group of fish parasitic protists (Ragan et al. 1996; Spanggaard et al. 1996). Each of our 50 replicated heuristic searches with random addition of taxa and tree bisection and reconnection branch swapping found the same 45 equally parsimonious trees of length 1209 (Fig. 2). The trees differed in branching order mainly in the Ascomycota. The branching order of the Archaeascomycetes (see Kurtzman and Sugiyama, Chap. Vol. VII, Part A) varied from tree to tree and the branching order of taxa within the Eusascomycetes was not consistent (not illustrated). For a rough estimate of the support from the data for the branching order, we performed 500 neighbor-joining bootstrap replicates, calculating the distance matrix using a Kimura correction for multiple hits, assuming that half of the sites in the alignment were not variable and that the transition to transversion ratio was two. The bootstrap numbers over 50% are shown in Fig. 3. In general, branches with more than 50% bootstrap support were also present in all 45 of the equally parsimonious trees (not illustrated). Weakly supported branches can also be discovered by determining which branches are lost in the consensus of trees that are one step longer than the most parsimonious tree (Bremer 1988). We next used PAUP 4.0d55 to find all trees one step longer than the most parsimonious tree in a single heuristic search with tree bisection and reconnection; PAUP found 531 trees. The consensus of the 531 trees is shown in Fig. 3.

A. Fungal Time Tree Under Molecular Clock Assumptions

If the rate of nucleotide substitution is approximately constant for all lineages and if fossil evi-
Table 1. Species used for phylogenetic analysis and the accession number of their 18S rRNA genes

<table>
<thead>
<tr>
<th>Group</th>
<th>Name</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascomycota</td>
<td>Blumeria graminis</td>
<td>L26253</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Candida albicans</td>
<td>X33497</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Capronia pilosella</td>
<td>U42473</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Dipodascopsis unimicella</td>
<td>U00969</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Eremascus albus</td>
<td>M83258</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Eurotium rubrum</td>
<td>U00970</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Galactomyces geotrichum</td>
<td>U00974</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Herpotrichia juniperi</td>
<td>U42483</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Hypomyces chrysospermus</td>
<td>M89993</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Leccanora dispersa</td>
<td>L37734</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Morchella elata</td>
<td>L37537</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Neurospora crassa</td>
<td>X04971</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Ophiostoma ulmi</td>
<td>M83261</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Peziza badia</td>
<td>L37539</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Pleospora radiis</td>
<td>U00975</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Pneumocystis carinii</td>
<td>X12708</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Porpidia crustulata</td>
<td>L37735</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Saccharomyces cerevisiae</td>
<td>J01353</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Schizosaccharomyces pombe</td>
<td>X54866</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Sclerotinia sclerotiorum</td>
<td>X69850</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Talaromyces flavus</td>
<td>M83262</td>
</tr>
<tr>
<td>Ascomycota</td>
<td>Taphrina deformans</td>
<td>U00971</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Boletus satanas</td>
<td>M94337</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Crogoophus vinicolor</td>
<td>M90822</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Coprinus cinereus</td>
<td>M92991</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Cronartium ribicola</td>
<td>M94338</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Leucosporidium scottii</td>
<td>X53499</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Russula compacta</td>
<td>U59093</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Spongipellis unicolor</td>
<td>M59760</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Suillus cavipes</td>
<td>M90828</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Tilletia caries</td>
<td>U00972</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Tremella globospora</td>
<td>U00976</td>
</tr>
<tr>
<td>Basidiomycota</td>
<td>Ustilago hordii</td>
<td>U00973</td>
</tr>
<tr>
<td>Chytridiomycota (Chytridiomycota)</td>
<td>Chytridium confervae</td>
<td>M59758</td>
</tr>
<tr>
<td>Blastocladiomycota (Chytridiomycota)</td>
<td>Blastocladiella emersonii</td>
<td>M54937</td>
</tr>
<tr>
<td>Neocallimasticales (Chytridiomycota)</td>
<td>Caecomyces (Shaeromonas) communis</td>
<td>M62707</td>
</tr>
<tr>
<td>Neocallimasticales (Chytridiomycota)</td>
<td>Neocallimastix frontalis</td>
<td>X80341</td>
</tr>
<tr>
<td>Neocallimasticales (Chytridiomycota)</td>
<td>Neocallimastix joynii</td>
<td>M62705</td>
</tr>
<tr>
<td>Neocallimasticales (Chytridiomycota)</td>
<td>Neocallimastix sp.</td>
<td>M59761</td>
</tr>
<tr>
<td>Neocallimasticales (Chytridiomycota)</td>
<td>Piromyces (Piromonas) communis</td>
<td>M62706</td>
</tr>
<tr>
<td>Spizellomycteales (Chytridiomycota)</td>
<td>Spizellomyces acuminatus</td>
<td>M59759</td>
</tr>
<tr>
<td>Endogonales (Zygomycota)</td>
<td>Endogone pisiformis</td>
<td>X85724</td>
</tr>
<tr>
<td>Entomophthorales (Zygomycota)</td>
<td>Basidiobolus ranarum</td>
<td>D29946</td>
</tr>
<tr>
<td>Entomophthorales (Zygomycota)</td>
<td>Conidioholus coronatus</td>
<td>D29947</td>
</tr>
<tr>
<td>Entomophthorales (Zygomycota)</td>
<td>Enomophthora muceae</td>
<td>D29948</td>
</tr>
<tr>
<td>Glomales (Zygomycota)</td>
<td>Gigaspora margarita</td>
<td>X58726</td>
</tr>
<tr>
<td>Glomales (Zygomycota)</td>
<td>Glomus intraradices</td>
<td>X58725</td>
</tr>
<tr>
<td>Glomales (Zygomycota)</td>
<td>Glomus mosseae</td>
<td>Z14007</td>
</tr>
<tr>
<td>Trichomycetes (Zygomycota)</td>
<td>Smittium culisetae</td>
<td>D29950</td>
</tr>
<tr>
<td>Basal animal</td>
<td>Dermocystidium salmonis</td>
<td>U21337</td>
</tr>
<tr>
<td>Basal animal</td>
<td>Ichthyophonous hoferi</td>
<td>U25637</td>
</tr>
<tr>
<td>Collar flagellate</td>
<td>Diaphanoeca grandis</td>
<td>L10824</td>
</tr>
</tbody>
</table>


dence is available to calibrate the rate of nucleotide substitutions, then the percentage substitution between pairs of species can be used to estimate their times of divergence. This approach has been used for bacteria (Ochman and Wilson 1987), animals including humans (Hasegawa et al. 1993), plants (Wolfe et al. 1989), and fungi (Berbee and Taylor 1993; Simon et al. 1993). Our initial estimate of the divergence times for major groups of fungi was based on 18S rRNA gene sequences.
Fig. 2A, B. Comparison of the distribution of animal and plant association mapped onto an 18S rRNA gene tree. Tree A shows branches leading to animal parasites or commensals in bold. Tree B shows branches leading to plant associated fungi, including endosymbionts, pathogens and saprophytes, in bold. Animal association appears at the base of the tree while association with plants characterizes most members of the monophyletic group including the Ascomycota, the Basidiomycota, and endomycorrhizal zygomycetes (Glomales). The tree, with a length of 1209 steps, was one of 45 equally parsimonious trees found in 50 replicated heuristic parsimony searches with PAUP 4.0a55. Species names are abbreviated with the first five letters of the genus followed by the first three letters of the specific epithet. Table 1 gives full names.
(Berbee and Taylor 1993). Although based on single gene and although the fossil data available for calibration of the tree are few, our estimates have been generally supported by fossil finds over the past few years. We estimated, for example, that basidiomycetous mushrooms radiated after the Cretaceous. Hibbett et al. (1995) reported finding the oldest mushroom known from amber dated at 90–94 Ma. We suggested that ectomycorrhizae in the fossil record should be good indicators for the radiation of the mushroom-forming basidiomycetes. LePage et al. (1997) found 50 Ma fossilized ectomycorrhizae typical of fungi in the suillus group. Based on our analysis, Blastocladiales should have diverged well before the Devonian, and one could expect to find their fossils in the Devonian. Taylor et al. (1994) subsequently described a well-preserved blastocladiaceous fossil from the Devonian, showing how static fungal morphology can be over even 400 Ma.

An intriguing exception to the predictions of our hypothesis was Taylor’s (1999) discovery of a complex, fruiting body producing, ascomycetous fossil fungus from the 400 Ma Rhynie chert. We estimated that the Ascomycota diverged from the Basidiomycota about 400 Ma, and that the Euascomycetes, with complex fruiting bodies, did not appear for another 100 million years or so.

Since our initial estimate of the timing of fungal origins, Nagahama et al. (1995) have released 18S rRNA gene sequences for some of the zygomycetes that radiated early in fungal history. Choanoflagellates and the fish parasitic protists Dermocystidium and Ichthyophonous were...
among the first organisms on the branch leading to animals to diverge and their 18S sequences have recently been released. These sequences allowed us to update our time tree, looking for evidence of how the very first fungi colonized land. We repeated our 1993 estimate using the phylogenetic data set described above and one of the 45 equally parsimonious trees from that data set (Fig. 2).

We then used maximum likelihood in PAUP 4.0d55 to calculate branch lengths for the tree (Fig. 4) assuming that (1) that a molecular clock was operating and (2) the rate of change varied from site to site in the data set. We assumed that the substitution rates showed a gamma distribution with shape parameter 0.5. An estimated 56% of sites were invariable. The variable sites were distributed into three classes of substitution rates.

To calibrate the molecular clock, we established a correlation between points in the geological time scale and percentage substitutions for fungal divergences. As can be seen in Fig. 5, we estimated the overall relationship between geological time and nucleotide substitution to be 1.26% per 100 Ma, compared to the 1.0% estimated previously (Berbee and Taylor 1993). The new, higher substitution rate probably results from using a less conservative alignment with a higher proportion of variable characters. In our current estimate, age estimates for divergences are generally higher. The Glomales, for example, diverged from the Ascomycota and Basidiomycota 600 vs. 500 Ma in our previous estimate. The older age estimate results in part from using the 965 Ma date for the divergence of animals from fungi as a calibration point. Molecular clock estimates are approximations and the differences among estimates may also be due to differences in taxon sampling or to variations in DNA substitution rates among lineages.

Our calibration points include:

1. The divergence of animals from fungi was estimated at 965 Ma, plus or minus about 140 Ma, by Doolittle et al. (1996). Based on 18S rRNA sequence data, some of the first “animals” to diverge are protists, including the collar flagellate, *Dermocystidium*, and *Ichthyophon* (Wainwright et al. 1993; Ragan et al. 1996; Spanggaard et al. 1996). The corresponding percentage substitution per lineage was 11.1% for ca. 1 billion years, or 1.1% per 100 million years.

2. Most Glomales, Endogonales, ascomycetes, and basidiomycetes are associated with terrestrial plants. The most parsimonious assumption is that radiation of these fungi followed the origin of land plants. While the date of origin of the first terrestrial plants is uncertain, microfossils from 460 Ma (Gray 1985) have been attributed to terrestrial plants. Conservatively placing the origin of land plants at 600 Ma, 140 Ma earlier than their appearance as fossils, provides an earliest possible date for terrestrial fungus radiation. The corresponding percentage substitution for the first terrestrial fungus radiation (*Endogone* from the rest) is 7.6 or 1.26% per 100 Ma.

3. Fossil spores and arbuscules from about 390 Ma (Kidston and Lang 1921; Remy et al. 1994) represent the most recent possible date for the origin of the Glomales. The divergence of the Glomales from the ascomycete/basidiomycete clade, corresponding to 7.4% substitution, must have happened before the 390-Ma fossils were produced. The divergence of the Ascomycota from the Basidiomycota, corresponding to 6.9% substitution, came after the origin of the Glomales.

4. Regularly septate hyphae have also been reported from 390-Ma Rhynie chert (Kidston and Lang 1921), and most likely originated earlier. Most ascomycetes and basidiomycetes have regularly septate hyphae, but most of the zygomycetes and chytrids do not. Compared with the origin of regularly septate hyphae, the divergence of Glomales from ascomycete/basidiomycete stem species, again corresponding to 7.4% substitution, was earlier, while the divergence of the Ascomycota from the Basidiomycota, corresponding to 6.9% substitution, was later.

5. A 290-Ma clamp connection (Dennis 1970) provides a most recent possible date for Basidiomycota. Compared with clamp origin, the split of Ascomycota from Basidiomycota, corresponding to 6.9% substitution, must have been earlier, while the radiation of Basidiomycota, corresponding to 6.2% substitution, was probably later.

6. *Caecomyces*, *Piromyces*, and *Neocallimastix* are genera of chytrids found in stomachs of mammals. Chytrid stomach inhabitants probably radiated after the first mammal radiations, following the split of placentals and marsupials, about 150–200 Ma. The divergence of the stomach chytrids from free-living chytrids, corresponding to 3.3% substitution, may have occurred after the host organisms
Fig. 4. The timing of fungal divergences. This tree is one of the 45 equally parsimonious trees from 50 replicated heuristic searches with PAUP 4.0d55. The branch lengths were calculated using maximum likelihood, assuming site-to-site rate variation and a molecular clock, and then absolute ages of divergences were estimated using the calibration of 1.26% substitution per lineage per 100 Ma. Species names are abbreviated with the first five letters of the genus followed by the first three letters of the specific epithet. Table 1 gives full names.
with fungal radiations, we estimated that the DNA substitution rate was about 1.26% per 100 Ma.

evolved, but before the split between placental and marsupial mammals, given that modern stomach chytrids inhabit both groups.

7. Chytrids in the genus *Neocallichlamastix* are found only in ruminants. The origin of ruminants, about 40 Ma, provides a most recent likely date for the radiation of these chytrids (associated with 0.2% substitution).

We had to drop one calibration that we used once before (Berbee and Taylor 1993). We used *Phellinutes*, originally reported as a 165-Ma fossil polypore (Singer and Archangelsky 1958), as a calibration point for the minimum age for homobasidiomycetes, but Hibbett et al. (1997) demonstrated convincingly that the microscopical anatomy of *Phellinutes* is typical for gymnosperm bark.

**B. Origin of the Fungi**

Trees from 18S rRNA gene sequences show animals as closest relatives to the fungi. Although fungi do not much resemble beetles and great apes, in habit and flagellation they do resemble some of the basal representatives of the animal kingdom. *Dermocystidium* and *Ichthyophonous* are protists that share with the basal fungi the habit of parasitizing fish and invertebrates (Ragan et al.
Collar flagellates are marine filter feeders that share with the chytrids the feature unusual among the protists of having a single posterior flagellum.

Based on our estimates of substitution rates, the most ancient of the fungal divergences occurred before plants colonized land (Fig. 4). Basal fungal lineages diverging before the likely origin of land plants include zygomycetes in the Trichomycetes and in the Entomophthorales. These fungi are usually obligate parasites or commensals of animals, most commonly of invertebrates (Fig. 2). The search for fungi parasitizing marine arthropods and other invertebrates has not been thorough. Additional search might reveal additional highly divergent fungi on marine hosts.

1. Flagella and the Basal Fungi

Among the fungi, only chytrids have flagella. The chytrid flagellum, being entirely typical for a eukaryote, is probably a feature retained from ancestral protists. Whether flagella were lost once or more than once during the evolution of terrestrial fungi is currently debated, and different gene trees show conflicting patterns. The flagellated chytrids do not appear at the base of the 18S tree. By inference, then, from the 18S data set all the groups of the non-flagellated fungi that cluster below chytrids or as sister taxa to chytrids came from the same flagellated ancestor that gave rise to chytrids. These nonflagellated fungi must have undergone convergent flagellar losses (Nagahama et al. 1995). The nonflagellated trichomycete Smittium calisea clusters below the chytrid Blastocladiella emersonii, requiring one loss of the flagellum. Entomophthora muscae and Conidiobolus coronatus constitute the sister taxon to B. emersonii, requiring a second flagellar loss. The zygomycete Basidiobolus ranarum is the sister taxon to the Chytridium confervae, requiring a third flagellar loss (Fig. 2). This phylogenetic pattern suggests that the zygomycetous fungi colonized the land and lost flagella at least three times in association with animal hosts. The Glomales represent the fourth loss, this time in association with plant hosts.

There are, however, two important qualifications to this interpretation of evolutionary events based on the tree from 18S sequence data. First, taxon sampling for the first fungi is incomplete. Possibly, the preference for association with animals would not remain basal if the diversity of chytrids and zygomycetes was more fully represented in the tree. Second, branch lengths among some of the zygomycetes, including some of the Entomophthorales, are unusually long (Fig. 2). A long branch also leads to the chytrid B. emersonii. The lack of monophyly for the chytrids and the zygomycetes in 18S rRNA gene trees may be an artifact of long branch attraction (Felsenstein 1978). The next step in assessing the monophyly of chytrids and zygomycetes involves assessing whether another gene without the long branch problem supports the same topology. The other available gene trees including zygomycetes and chytrids are from β-tubulin (Keeling et al. 2000). Interpretation of β-tubulin trees is complicated by historical gene duplication. One of the two copies of β-tubulin gene in the chytrid Spizellomyces punctatus clusters with the zygomycetes (Keeling et al. 2000), which supports convergent loss of flagella. However, other aspects of the zygomycete phylogeny are different in the β-tubulin and 18S gene trees, and more data are needed to resolve branching patterns and ancient fungal history. The mitochondrial protein-coding gene NAD5, for example, might be useful because branches leading to a chytrid, a zygomycete, and other fungi appear to be about the same length (Paquin et al. 1995b).

2. Colonization of Land

While association with animals is particularly common at the base of the fungal tree, association with plants as mycorrhizal symbionts, as parasites, and as saprobes is the most common condition for the monophyletic group of terrestrial fungi that includes the Endogonales, Glomales, and Ascomycota and Basidiomycota (Fig. 2). The endomycorrhizal Endogonales and Glomales comprise the fourth group of zygomycetes that independently lost flagella, possibly during colonization of land in association with terrestrial plants; but were terrestrial plants available as food sources when these fungi were radiating? Some of the earliest fossil evidence for land plants comes in the form of dispersed microfossils of spores and tracheid-like structures, some of them from the Ordovician from 460Ma. Our most recent tree (Fig. 4) suggests that the first divergences among the terrestrial fungi, the divergences of Endogonales and Glomales, occurred about 600Ma or even earlier. This is about 140Ma before any evidence for vascular plants that are now hosts for these obligate endomycorrhizal fungi and is about 100Ma older than our estimate for the age of the divergences based on our earlier work. The difference between
the age estimate of fungi and hosts may reflect error because of rate variation in sequence evolution. It may reflect a greater age for the vascular plants than the fossil record suggests, or it may be that the fungi were initially associated with the ancestors of the vascular plants, following their hosts onto land.

C. Radiation of the Terrestrial Fungi: Origin of the Ascomycota

The tree from 18S rRNA sequence data shows Ascomycota and Basidiomycota diverging from one another in the Paleozoic, about 500Ma (Fig. 4).

In the Ascomycota, an early radiation established the Archiascomycetes. (see Kurtzman and Sugiyama, Chap. 9, Vol. VII, Part A). The Archiascomycetes is a diverse group of fungi with and without hyphal fruiting bodies. It includes the fission yeast Schizosaccharomyces pombe, the animal pathogen Pneumocystis carinii, the peach leaf curl fungus Taphrina deformans, and the discomycete Neolacta vitellina. No obvious morphological character unites these taxa, and they do not always cluster together based on 18Ss rRNA sequence data (Fig. 3). Inferring the sequence of evolutionary events leading to their origin will require better resolution of their branching order and an improved understanding of the patterns underlying their morphological diversity. A more complete fossil record would contribute greatly to reconstruction of the first ascomycetes. Possibly, the complex, 400Ma old “pyrenomycete” fossil (Taylor et al. 1999) represents an Archiascomycete, revealing morphological features of a very early ascomycete.

The two other primary groups in the Ascomycota, the yeasts or Hemiascomycetes (also known as Saccharomycetales) and the filamentous ascomycetes or Euascomycetes, are both clearly monophyletic. In our analysis, the Hemiascomycetes received 96% support and the Euascomycetes received 99% bootstrap support (Fig. 3). The two groups appear in trees from other genes as well, including, for example, orotidine 5'-monophosphate decarboxylase (Radford 1993), β-tubulin (Baldauf and Palmer 1993), and the second largest subunit of RNA polymerase II (Liu et al. 1999).

The Hemiascomycetes include budding yeasts like the baker’s yeast Saccharomyces cerevisiae as well as hyphal “yeasts” like Dipodascopsis uninguiculata and Galactomyces geotrichum. The hyphal yeasts appear to be basal in our trees, suggesting that the budding yeasts may have been derived from hyphal ancestors.

1. Euascomycetes

The Euascomycetes always appear as a monophyletic group (see Barr, Chap. 13, Vol. VII, Part A). However, the branching order along the backbone at the base of the Euascomycetes is unresolved with available data (Fig. 3). Conflicting phylogenies from different laboratories and low bootstrap support for any particular phylogeny characterize this region of the tree. In all of the 45 equally parsimonious trees from our data set, the pyrenomycetes including Neurospora crassa appeared at the base of the Euascomycetes but without bootstrap support (Fig. 2). However, the basal group is unresolved in a consensus of 531 trees found in a search for trees up to one step longer (Fig. 3). Using a similar method of analysis but including a different selection of taxa, Gargas et al. (1995) found Pezizales at the base of the Euascomycetes, although again without bootstrap support. Spatafora’s (1995) analysis showed a basal divergence between pyrenomycetes plus Pleosporales and the discomycetes plus plectomycetes and allies.

While the divergence order for the lineages of Euascomycetes is unclear, all groups of Euascomycetes have features in common which, by inference, should also be characters of the ancestor to the whole group. The basal members of Euascomycetes lineages share long, slender ascis that are arranged in a single layer in a fruiting body and which shoot their spores forcibly into the air. Most groups of Euascomycetes produce conidia (i.e., mitosporic). Conidia, along with pollen and other spores, are preserved in the fossil record. Finding evidence of conidial diversification in the fossil record would provide a most recent possible age for radiation of the Euascomycetes. Our time tree suggests that this radiation may have been taking place in the Mesozoic, about 240Ma.

Originating from the backbone of the Euascomycetes are several well-supported, monophyletic groups of taxa, some corresponding to traditional Ascomycota classes. Plectomycetes is a traditional class characterized by cleistothecial fruiting bodies containing thin-walled asci (see Geiser and LoBuglio, Chap. 10, Vol. VII, Part A). The plectomycetes consistently form a
monophyletic group in trees from 18S rRNA (Berbee and Taylor 1992a); (Fig. 3) and other genes (Fig. 1); (Bowen et al. 1992). They encompass fungi ranging from false truffles in the Elaphomycetales (Landvik et al. 1996; LoBuglio et al. 1996) through human pathogens in the Onygenales to *Penicillium* species in the Trichocomaceae (Berbee et al. 1995). Similarly, the pyrenomycetes is a traditional class strongly supported by sequence data in all studies (see Samuels and Blackwell, Chap. 11, Vol. VII, Part A). The typical pyrenomycete has long, slim, thin-walled asci arranged in a single layer at the base of a flask-like fruiting body. Ascospores are forcibly discharged at maturity. Suggesting that the fruiting body shape, ascus shape, and ascus disposition are all functionally correlated to forcible ascospore discharge, the fungi that lose forcible discharge often lose the other features too, so that they resemble plectomycetes. The 18S rRNA trees provide clear evidence that *Ophiostoma* and *Ceratocystis*, for example, are descendants of fungi with all the typical pyrenomycete characters (Berbee and Taylor 1992b; Spatafora and Blackwell 1994).

In contrast to the above examples, the discomycetes (cup fungi) is a traditional class that lacks support from most 18S rRNA studies (see Pfister and Kimbrough, Chap. 12, Vol. VII, Part A). Two discomycetace orders, the Leotiales and the Pezizales sometimes cluster together (Fig. 2) but without bootstrap support (Fig. 3). The powdery mildews, or Erysiphales, often cluster within the Leotiales even though they had not previously been considered discomycetes, but again bootstrap support for the connection is weak (Saenz et al. 1994). Lichenized discomycetes in the Lecanorales occupy different positions, not necessarily with other discomycetes, in different phylogenetic analyses. As mentioned above, *Neoleotia vitellina* although a discomycete, clusters with the Archiascomycetes (Landvik et al. 1993).

Like the discomycetes, the loculoascomycetes is a traditional ascomycete class without strong support from 18S rRNA studies (see Barr and Huhndorf, Chap. 13, Vol. VII, Part A). Before molecular evidence contradicted morphological evidence, Chaetothyriales were always included in the loculoascomycetes. Like other loculoascomycetes, they have bitunicate asci arranged in a hymenial layer. From these two morphological characters, the Chaetothyriales have little in common with the plectomycetes. However, consistently in 18S rDNA sequence-based trees (Berbee and Taylor 1995; Haase et al. 1995; Spatafora et al. 1995; Silva-Hanlin and Hanlin 1999) and occasionally in trees from RNA polymerase subunit II (Liu et al. 1999), the Chaetothyriales cluster with the plectomycetes. This suggests that the cleistothecial form with irregularly distributed asci lacking forcible discharge in the plectomycetes may have been derived from fruiting bodies with asci arranged in a single layer having forcible ascospore discharge (Berbee 1996).

Although the taxonomic position of the groups of ascomycete lichens is uncertain, the phylogenies do show that ascomycete lichens are polyphyletic. Lichenization appears to have originated independently in at least two orders, the Arthoniales and the Lecanorales (Gargas et al. 1995; Stenroos and DePriest 1998), although uncertainty about the branching order along the backbone of the Euascomycetes makes a single origin difficult to rule out (F. Lutzoni, AIBS oral presentation, August 1997). In our trees, lichenization does not appear as the most ancient habit for the Ascomycota. Lecanorales appear as the earliest group of lichenized ascomycetes and the separation of *Porpidia crustulata* from *Lecanora dispersa* represents an old divergence for the group. If lichenization arose after the divergence of Lecanorales from other Euascomycetes, it could be as old as about 240 Ma and it most parsimoniously arose before the separation of *P. crustulata* from *L. dispersa* at about 180 Ma.

Until recently, Euascomycetes fungi could only be classified once their sexual state was observed and studied. No sexual state is known for over 5000 species of fungi (Rossmann 1993), a situation that has confounded classification. For the Ascomycota, sequencing of rDNA regions is becoming increasingly useful in integrating the asexual and sexual species into the same classification system by revealing similarities at the nucleotide level (Berbee and Taylor 1992c; Gaudet et al. 1989; LoBuglio et al. 1993; Reynolds and Taylor 1993; Taylor 1995). Seifert et al. (1997) used 18S and 28S ribosomal gene sequences to show that the asexual *Trichothecium roseum* is probably a member of the pyrenomycete order Hypocreales, and Kuhls et al. (1996) showed that asexual *Trichoderma reesei* is probably a clonal derivative of the sexual species *Hypocrea jecorina*. 
D. Basidiomycete Radiation

One of the first molecular phylogenetic analyses of fungi, from SS rRNA data, divided the Basidiomycota into two groups, one with inflated rings or dolipores surrounding the septa between adjacent hyphal cells, and the other lacking the inflated rings (Walker and Doolittle 1982). Partly because SS sequences include only about 120 base pairs of DNA, the exact branching order of the smuts, the rust group, and the mushrooms and jelly fungi differed in different analyses (Blanz and Unseld 1987; Walker 1984). 18S and 28S ribosomal RNA gene sequences provided more characters and resolving power. The trees from 18S and 28S sequences show three lineages in the Basidiomycota that diverged early in the history of the group but do not clearly resolve the branching order for the three (Fig. 3). These are the rusts or Uredinomyces (see Swann and Frieders, Chap. 2, this Vol.), the smuts or Ustilaginomyces (see Bauer et al., Chap. 3, this Vol.) and the Hymenomycetes (Swann and Taylor 1993; Begerow et al. 1997; Swann et al. 1999); see Wells et al., Chap. 4, this Vol.; Hibbett and Thorn, Chap. 5, this Vol.). The clamp connection at the septa that is found in at least some members of all three groups may be a primitive character of the phylum that evolved before the first basidiomycete radiation, by about 440 Ma. If this estimate of phylogeny and of timing is correct, then it may be possible to find fossilized clamp connections much older than the 290-Ma clamps found in Pennsylvanian coal (Dennis 1970).

IV. Summary

This chapter has emphasized many phylogenetic questions where sequence data provide clear resolution. Because of the evolutionary constraints limiting horizontal transfer and paralogous evolution that act on the ribosomal genes, they will probably remain valuable phylogenetic indicators for quite some time. In future studies, the ribosomal genes will continue to bring misplaced fungi into their correct phylogenetic positions. As we have emphasized in this chapter, however, for many phylogenetic questions ribosomal gene sequences do not provide clear answers. For further understanding of the deeper branches in fungal tree, sequences from more genes and characters of all other types will be welcome.

References

Molecular evidence that the assexual industrial fungus Trichoderma reesei is a clonal derivative of the ascomycete Hypocreales eurycotica. Proc Natl Acad Sci USA 93:7755–7760


