Evolutionary studies of ectomycorrhizal fungi: recent advances and future directions[1]

Thomas D. Bruns and Richard P. Shefferson

Abstract: The three biggest advances in fungal molecular phylogenetics in the last few years have been (1) the huge expansion in data sets, (2) the development of nonribosomal loci for phylogenetic analysis, and (3) the use of increasingly sophisticated types of analyses. In addition, advances in parallel computing hold great promise for dramatic increases in speed of analysis. These changes have had, or will have, a direct impact on mycorrhizal ecology through the use of sequence-based identification and an indirect impact through the conclusions drawn from such studies. One problem in the field has been the accidental addition of erroneous sequences to the public databases through a variety of means, including polymerase change reaction contamination. We discuss several examples, suggest ways to identify errors, and argue the case for third-party annotations of sequences. Multiple studies have produced compelling evidence that the ectomycorrhizal habit has developed convergently in multiple lineages of fungi and plants. We reexamine the case for loss of the ectomycorrhizal habit in fungi and show that the results are model dependent.

Key words: internal transcribed spacer (ITS) region, peroxidase genes, likelihood models, erroneous data, ectomycorrhizal habit.

Résumé : Les trois avancées les plus marquantes de la phylogénie moléculaire des champignons au cours des dernières années sont (1) l'énorme expansion des bases de données, (2) l'analyse de loci non ribosomiques et (3) l'utilisation de méthodes de reconstruction phylogénétique de plus en plus sophistiquées. Les progrès réalisés dans le calcul parallèle vont d'autre part accroître considérablement la vitesse d'analyse. Toutes ces améliorations ont eu, ou vont avoir un impact majeur en écologie mycorhizienne, sur les méthodes d'identification de séquences. Ces études phylogénétiques permettent également d'obtenir un certain nombre d'informations à caractère écologique. Cependant l'addition accidentelle de séquences erronées (p. ex., des contaminations à l'étape de la réaction de polymérisation en chaîne) dans les bases de données publiques est un problème sérieux. Nous présentons plusieurs exemples de ce type de problème dans la littérature mycorhizienne. Nous suggérons également des solutions pour identifier les erreurs, et nous proposons l'établissement d'un système d'annotation par un troisième intervenant. De nombreuses études phylogénétiques ont permis de montrer que le mode de vie ectomycorhizien est apparu de façon convergente dans de multiples clades de champignons. Nous examinons à nouveau les conclusions de ces études sur la perte du mode de vie ectomycorhizien, et nous montrons que les résultats sont en fait étroitement dépendants du modèle d'analyse.

Mots clés : région de l'espaceur interne transcrit (ITS), gènes de la peroxydase, modèle de probabilité, données erronées, comportement ectomycorhizien.

Large data sets are now the rule

Between 1990 and 2000, the number of molecular systematic studies doubled roughly every 2 years (Pagel 1999*a*), and the number of taxa per study probably increased at a greater rate. This huge expansion in molecular data has been enabled by the dramatic increase in speed of sequence acquisition and has been driven by a similarly dramatic expansion of people in the field. Within the agarics, the recent papers by Moncalvo et al. (2000, 2002) are stunning examples that included over 300 and 800 taxa, respectively. Binder and Hibbett's (2002) paper is equally impressive; it includes 93 species of basidiomycetes and presents data from four genes for 82 of these. The number of sequences deposited in Genbank, especially for the rDNA loci, is growing rapidly for all types of fungi. This trend will clearly continue, both because of the interests of individual investigators and because of the efforts of larger, organized groups such as the Fungal Tree of Life project (http://ocid.nacse.org/research/deephyphae/htmls/AFTOL.html).

Targeted loci: Why the switch to protein coding regions?

The internal transcribed spacer region (ITS) is a locus that needs no introduction to a mycorrhizal audience. A strong

T.D. Bruns.[2] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-140, USA.
R.P. Shefferson. Department of Integrative Biology, University of California, Berkeley, CA 94720-3140, USA.

[1]This article is one of a selection of papers published in the Special Issue on Mycorrhizae and was presented at the Fourth International Conference on Mycorrhizae.

[2]Corresponding author (e-mail: boletus@socrates.berkeley.edu).

case can be made for it as the locus of choice for identification of environmental samples. This is because it typically shows variation around the species level, and its multicopy arrangement and highly conserved priming sites make it easy to amplify from virtually all fungi, even when the material is marginal in quantity or quality. The current and growing sequence data available for ITS further enhance its value. For these reasons it is an important locus for phylogenetics and especially ecology.

However, ITS has some serious deficiencies, particularly if the primary goal is to produce a robust estimate of the phylogeny. Alignment is clearly the biggest problem. Because the spacer regions are prone to frequent IDELs (i.e., insertion and (or) deletions), alignment becomes more difficult, and finally arbitrary, as divergence increases. This problem effectively narrows the range of utility for ITS data sets, and it makes useful outgroups difficult to find. ITS can produce a well-supported phylogeny, as in the case of *Pisolithus* (Martin et al. 2002), but in many cases, particularly those involving larger genera, it is typical to see ITS trees that group closely related species at high confidence levels, but do a poor job of discerning relationships among these clades. Examples include: *Cortinarius*, *Oidiodendron*, *Russula*, *Rhizopogon*, *Tomentella*, *Tricholoma*, and *Suillus* (Kretzer et al. 1996; Hambleton et al. 1998; Mankel et al. 1999; Hoiland and Holst-Jensen 2000; Koljalg et al. 2000; Peintner et al. 2001; Bidartondo and Bruns 2002; Grubisha et al. 2002; Miller and Buyck 2002). Outgroups, if they are alignable, often attach to the longest branches in the tree, but without any confidence. An additional problem is that very closely related species may not have accumulated many differences within the ITS; thus, an analysis based only on ITS may fail to separate all species.

The structural ribosomal genes, such as the nuclear small subunit rRNA (*nuc-SSU*) and nuclear large subunit rRNA genes (*nuc-LSU*), have been very popular for higher level systematics. They are now the basis for the overwhelming majority of our phylogenetic estimates in the fungi, and they constitute a huge and growing data set that is useful for identification at higher levels. These genes share some of the positive and negative features of the ITS region, although their utility spans a much broader phylogenetic range. Similar to the ITS region, alignment of the more variable regions becomes a problem as more distantly related fungi are added to the comparison. This problem is particularly evident in the mitochondrial genes. In addition, it is very common to see poor resolution at the basal branches and at the tips of trees. Taking the recent Moncalvo et al. (2002) tree as an example, one can see that the *nuc-LSU* data often grouped closely related species and genera together with high confidence, but relationships among these groups were often unresolved or resolved with low confidence. Berbee et al. (2000) identified a similar problem with *nuc-SSU* resolution of the basal branches within the Ascomycota and used jackknife resampling to show that up to 15-fold more data would be needed to confidently resolve some of these branches. Binder and Hibbett (2002) showed that when all four of the larger rRNA genes (mitochondrial (*mt*) and *nuc*, *LSU* and *SSU*) are combined, many additional branches are resolved within the Basidiomycota; nonetheless, many basal branches still remain unsupported.

These problems are the main reasons that people have turned to protein coding sequences as additional sources of data. One of the primary advantages of protein coding sequences is that they are easy to align. This ease of alignment is due to the fact that they are constrained by reading frames, and thus IDELs are much less common. However, protein coding sequences are not as conserved at the nucleotide level as structural RNA genes. This is because it is the translated gene product rather than the gene that is primarily under selection. This difference is useful at lower taxonomic levels, because it means that third-base positions and sites within introns, which are typically under little selection, often provide informative characters among recently diverged taxa. On the negative side, these same characters are often saturated at higher levels. For this reason, when protein genes are analyzed for higher level questions they are often translated to amino acid sequences or codon positions are differentially weighted. Another negative aspect of the variable third bases is that it makes it more difficult to design "universal" primer sites, because a completely invariant 20-bp stretch of sequence often does not exist. Degenerate primers and lower annealing temperatures allow one to span a wider range of sequence variations, but this approach often results in less specific amplification that requires cloning prior to sequencing.

Common protein-coding loci that have been used for phylogenetics in fungi include β-tubulin, elongation factor 1-α, ribosomal polymerase B, and mitochondrial ATPase 6. These particular loci have been selected for use by the ongoing Fungal Tree of Life project. They were targeted because they are usually single-copy genes and thus avoid the pitfalls of paralogous comparisons, and because smaller scale studies in various fungal groups have shown these loci to be useful for phylogenetic estimates at several different taxonomic levels (O'Donnell et al. 1998*a*, 1998*b*, 2001; Hirt et al. 1999; Kretzer and Bruns 1999; Liu et al. 1999; O'Donnell 2000; Craven et al. 2001; Landvik et al. 2001; Cruse et al. 2002; Matheny et al. 2002; Chaverri et al. 2003; Keeling 2003). However, it should be noted that Tanabe et al. (2002, 2004) have found evidence for saturation and bizarre phylogenetic patterns in both elongation factor 1-α and ribosomal polymerase B. Although we suspect that limited taxon sampling may be part of the problem with their phylogenetic estimates, the point about saturation is worth noting. Links for primer sequences for all of these genes, except β-tubulin, are posted on the Fungal Tree of Life Web site (http://ocid.nacse.org/research/deephyphae/htmls/AFTOL.html); primers for β-tubulin are given in Landvik et al. (2001). Many other protein-coding loci have also been used or are being developed. Among these, mating-type genes in ascomycetes (Turgeon 1998; Witthuhn et al. 2000) and chitin synthase genes (Mehmann et al. 1994; Hintz 1999; Cruse et al. 2002) hold great promise, particularly at lower taxonomic levels.

Advances in analysis: a necessary ingredient for large data sets

Large data sets pose difficulties for analysis. This is particularly true if tree-building methods such as parsimony and maximum likelihood are used, because they evaluate trees by optimality criteria and must search through "tree space"

(i.e., all potential trees) to find the best tree or trees that meet the specified criterion. Tree space becomes astronomically large as the number of taxa increases. Approximately 8.9×10^{23} different bifurcating rooted trees exist for a 20-taxon data set (Felsenstein 1978); this once seemed like a large data set. For a 100-taxon tree, which is almost nine times smaller than the Moncalvo et al. (2002) data set, the total number of trees is 3.4×10^{184} (Felsenstein 1978). Thus, as the number of taxa increases, searching thoroughly through tree space becomes impossible, and the "best" tree may not be found.

In addition to being large, tree space is structured into "islands" of suboptimal trees. Because a search is conducted by rearranging branching of existing trees, searches can become trapped in larger islands, and even with extended times they may never find better trees, even if they exist. Nixon (1999) developed a clever search strategy to push the search away from islands. His method, the "parsimony ratchet", is based on temporarily weighting a random subset of the characters. He shows that even with large data sets, search times can be reduced by 1000-fold, and better trees can often be found. His methods have recently been adapted for use by the PAUP (Sikes and Lewis 2001) and have also been applied to likelihood analyses (Vos 2003).

Searching for the best tree is usually not the main goal, because the best tree may not be correct. For that reason, one usually wants to estimate confidence in clades. Until recently, the most common methods for assessing confidence involved resampling procedures, such as bootstrapping (Felsenstein 1985) and more recently, jack-knifing (Farris 1997). These methods have quickened over time by the use of more efficient but less thorough search algorithms (Salamin et al. 2003). Fortunately, estimates derived from these less rigorous search algorithms usually correlate well with those derived from more thorough searches. However, fast bootstrap or jack-knife algorithms tend to yield slightly lower estimates of confidence with increased variance (Salamin et al. 2003). Nevertheless, estimates of confidence above 95% tend to be well supported by all methods (Salamin et al. 2003).

Distance methods, such as neighbor joining, build trees based on a formula and are very fast ways to analyze even large data sets. However, they are not as popular, because they are thought to be less accurate in reconstructing the correct phylogeny (Farris 1997), although some simulations contradict this commonly espoused belief (Hillis 1996).

Large data sets do seem to have one major advantage for analysis: they recover trees that are better estimates of the true phylogeny than do smaller taxon samples (Hillis 1996). This counter-intuitive result may be due in part to the way phylogenetic noise is distributed nonrandomly across the tree and to the tendency of more complete samples to break up long branches. The net result of this property is that increased taxon sample is at least as important as increased sequence, if the data are evolving at the appropriate rate for the question at hand (Pollock et al. 2002).

The application of Bayesian methods to tree construction and evaluation is a relatively new phenomenon. These methods have become increasingly popular and perform well with large data sets. An excellent recent review by Huelsenbeck et al. (2002) provides details of this approach, and a more general review by Lewis (2001) has some good explanations and figures. Free software is available on the Web (Mr.Bayes: http://morphbank.ebc.uu.se/mrbayes3/info.php; BAMBE: http://www.mathcs.duq.edu/larget/bambe.html).

The underlying model for a Bayesian analysis is a likelihood function. This means that parameters, such as transition/transversion ratios, rates of substitution, variation in rates across positions, and base composition biases, are specified a priori and used to calculate the likelihood for a given tree. However, in a Bayesian analysis, the initial model parameters and tree topology, termed "priors", are adjusted during the run, and the program searches for values of these that maximize the posterior probabilities, which is essentially the likelihood conditioned by the priors.

The Bayesian search algorithm is based on the Markov chain Monte Carlo (MCMC) method. Basically, models (i.e., trees plus parameters) are sampled randomly and retained based on their probability. The method searches through tree space by moving from one tree to another in a random manner. Higher probability trees are always accepted and lower probability trees are accepted with a probability that is weighted by the ratio of its probability compared with that of the current tree. There is an initial "burn-in period" in which the searched tree space is composed of trees with significantly lower probabilities; with time the chains climb to more probable trees, and the search ideally converges on a sample of interest that includes trees with highest probabilities. After convergence, the number of times a tree is sampled and retained in the search is a direct measure of its posterior probability. Typically, a 95% "credibility interval" is produced by starting with the best tree and adding lower probability trees until the accumulated probability is 95%. As discussed previously, the tree space is too large to actually sample all trees, but the MCMC method samples this in much the same way one would integrate the area under a complex distribution.

A few features of MCMC analyses, which were described as "pitfalls" by Huelsenbeck et al. (2002), are worth reviewing. The sensitivity of the result to the priors is an issue that tends to worry users, but it can and should be tested by starting chains with different priors and seeing if they converge to the same outcome. More difficult issues involve determining when convergence has occurred and whether the sampling of parameter space is efficient. Sampling may not be efficient if high probability trees are on many small separate islands in parameter space. Finally, the lack of correspondence between bootstrap values and posterior probabilities is reason for concern (Suzuki et al. 2002; Alfaro et al. 2003; Douady et al. 2003). Posterior probabilities tend to give higher estimates than bootstrap values. Many may gladly embrace this apparent increase in support, as bootstrap values are often low for complex data. However, interpretation needs to be cautious, because the two estimates address slightly different uncertainties. Bootstrapping addresses the uncertainty of the data, viewing the current set as a small sample from a universe of data that has a similar distribution of sites, while Bayesian posterior probabilities address uncertainties of the model, given the data. This point is underlined by the observation that bootstrap and Bayesian estimates for nodal support correlate well if Bayesian esti-

mates are made on bootstrapped data sets (Douady et al. 2003).

Another relatively recent development in phylogenetic analysis is the extension of likelihood methods to morphological and behavioral traits. Here, we only cover a few salient points that are necessary for our later discussion of the evolution of the ectomycorrhizal (EM) habit; for a detailed discussion of these methods, see Pagel (1999*b*). Likelihood models of morphology are typically coupled with molecular systematic estimates, such that the morphological trait of interest is not used to construct the tree. Instead, the tree is used as part of the "model" to examine the evolution of the trait. A common goal is to estimate ancestral states, or states of internal nodes. The tree topology and branch lengths are assumed to be known, and the probability of character state change is proportional to branch length, which is assumed to be proportional to time. The relative rates of transitions between characters are assumed. Rates in the forward and reverse direction do not need to be equal, but are often initially assumed to be so. With these relatively simple assumptions, a likelihood can be calculated that describes the probability of the observed data (i.e., the current states of the extant taxa), given the model (i.e., the tree, the assumed character transition rates, and various reconstructions of the trait on the tree). Likelihoods of different models can be compared by a likelihood ratio test, but this requires that the two models are nested (i.e., one is a special case of the other). When non-nested models are compared, the likelihood ratio does not have a χ^2 distribution, and the test does not yield a true probability, although the results are often interpreted as an index of support (Pagel 1999*b*).

Hardware advances on the horizon

The previous discussion takes the current, generally available computer technology as a given, but the advances in computer speed have clearly been a necessary ingredient for phylogenetic analysis of large data sets with complex models. Although increases in speed of individual processors will continue to have an important impact, parallel processing is destined to have an even greater effect on the field of molecular phylogenetics. The potential was demonstrated by Brauer et al. (2002), who showed that an increase in speed is almost linear with processor number. Although Braurer et al. (2002) used an advanced parallel supercomputer, other options for cheaper parallel machines, or distributing analyses across existing networked computers, will make parallel processing of phylogenetic analyses a reality at a lower budget (Grimshaw et al. 1997; Becker et al. 1999; Snell et al. 2001; Brauer et al. 2002).

Using phylogenetic analyses to check the validity of molecular data

An important use of phylogenetic analysis is simply to check the validity of molecular data. This is particularly relevant when sequences are acquired from field-collected material or from cultures that cannot be independently verified by morphology. Unfortunately, it is also relevant for sequences deposited in public databases. A recent paper by

Bridge et al. (2003) shows that misidentified sequences are common. They examined deposited sequences of ITS for *Amanita* and *Phoma* and SSU sequences for genera of the Helotiales, and they estimated that roughly 20% of these sequences were unreliable for all three taxa.

ITS data provided several other impressive examples of misidentified sequences. Three classic examples illustrate the problem and show the solution. The first involved purported spruce sequences that were derived from field-collected needle samples (Smith and Klein 1994). These were ultimately discovered to be sequences of fungal endophytes that had been preferentially amplified (Camacho et al. 1997). A similar problem occurred with "bamboo" sequences that turned out to be primarily basidiomycetes (Zhang et al. 1997). The third example relates to sequences that were purported to be from the Glomeromycota, but which turned out to be from the Ascomycota (Redecker et al. 1999). In all three cases, the 5.8S rRNA gene and the addition of outgroups were the key to testing the validity of the sequences. Although the 5.8S gene is too conserved to provide much fine-scale resolution, it is usually good enough to place sequences within broad groups such as the Ascomycota, Basidiomycota, and higher plants (Cullings and Vogler 1998; Redecker et al. 1999), and if greater resolution is needed, the flanking SSU and LSU genes can be sequenced (Bidartondo et al. 2003). As the sequence database for ITS expands, a BLAST search with the spacer regions themselves will probably be sufficient to identify many erroneous sequences; currently, however, as pointed out by Schüßler et al. (2003), BLAST searches alone risk perpetuating and expanding misidentification problems by applying incorrect names to new unknowns. In addition, a BLAST search may sometimes not retrieve the closest sequence, because differences in sequence length can interact with the score. At least for the near future, a phylogenetic analysis will remain the safest way to indentify unknown sequences.

The previous examples involved sequences that were generated for phylogenetic analyses or ecological studies, and so testing the validity of the data should be a fairly natural outcome of a thorough analysis. However, any time molecular data are generated for multiple taxa, the expectations from molecular evolutionary models can be used to examine the authenticity of the sequences.

A recent paper by Chen et al. (2001), in which polymerase chain reaction (PCR) was used to search for the presence of lignin and magnesium peroxidase genes in EM basidiomycetes, serves as an example. This paper reported an important result, because if these genes are widespread and functional it adds support to the idea that the EM fungi retain a significant saprobic ability. However, after looking more closely at the data with phylogenetics, the authors have now retracted their claim (Cairney et al. 2003); we review the reasoning for their decision in the following section, as it is instructive.

The results presented in the original paper were unusual, because almost no sequence variation was reported among several distantly related fungi. For example, no variation was found within *Phanerochaete*, *Paxillus*, *Tylospora*, and *Piloderma* in the *Lip 2* gene. Within 29 of the 30 comparisons summarized in Table 3 of Chen et al. (2001), the mean iden-

tity to *Phanerochaete* is 96.6%, and all but one sequence showed greater than 92% identity. The taxon sample is fairly diverse and includes representatives of groups such as the euagarics (*Tricholoma*, *Cortinarius*, *Rosites*), suilloid lineage (*Chroogomphus*), gyrodontoid lineage (*Gyroporus*, *Pisolithus*), boletoid lineage (*Tylopilus*), cantharelloid lineage (*Hydnum*), Russulales (*Lactarius*), and resupinate taxa in the athelioid lineage (*Piloderma* and *Tylospora*). All of these groups are distantly related to each other according to mitochondrial and nuclear LSU and SSU rRNA genes and *ATPase 6* genes (Bruns et al. 1998; Kretzer and Bruns 1999; Binder and Hibbett 2002; Moncalvo et al. 2002). Thus, the extraordinary conservation of the peroxidase genes across distantly related taxa of EM fungi is certainly unexpected, even if the gene product is highly conserved by strong stabilizing selection. Two explanations are possible: (1) the gene is extraordinarily conserved both at the amino acid level and the nucleotide level and (2) the similarity is due to PCR carry-over contamination. To discriminate among these possibilities, one can examine the data to see if they fulfill the expectations for an evolving protein gene.

The *Lip 5* genes, which were the most commonly amplified by Chen et al. (2001), have 28 total differences across the 12 taxa (11 mycorrhizal plus *Phanerochaete*) sampled. One would expect most of these changes to occur at the third positions within codons or within the two introns. However, 15 of the 28 changes occur within the reading frame, and 10 of the 15 changes occur at first or second positions within codons. All 10 result in nonsynonymous amino acid substitutions, and many of these are nonconservative substitutions. In addition, the percentage of variable positions within the reading frame (9%) and introns (9.6%) are almost identical. Taken at face value, these data would suggest that the gene product is not conserved, yet the high level of overall sequence similarity contradicts this.

Another unexpected feature of these data is that all of the changes are unique to particular taxa; even closely related pairs of taxa such as the two *Paxillus* species, the two *Pisolithus* species, and the two *Tricholoma* species share no substitutions. If we make a phylogenetic tree from these data, it yields a perfect "star phylogeny" (Fig. 1), which is indicative of a rapid radiation. Yet we know from other genes, including *mt-LSU*, *mt-SSU*, *mt-ATPase 6*, *nuc-LSU*, and *nuc-SSU*, that these taxa do not appear to be a single radiation (Bruns et al. 1998; Kretzer and Bruns 1999; Binder and Hibbett 2002; Moncalvo et al. 2002). This complete lack of synapomorphies in only these genes is nearly impossible to explain by any evolutionary process, because it means that all changes must have occurred only at the finest tips of the tree. Again, it is not what one would expect for a highly conserved gene.

For these reasons, we can reject the first hypothesis; the gene is not behaving like a highly conserved locus. Thus, the most likely explanation is that these sequences are all minor PCR modifications of a previously amplified *Phanerochaete* template. This carry-over contaminant hypothesis would explain the striking similarity of Chen et al.'s (2001) sequences and would also explain the star phylogeny. The observed autapomorphic differences could have occurred during amplification or may simply be sequence errors. Carry-over

Fig. 1. Perfect star phylogeny from LIP 5 data of Chen et al. (2001). Unrooted parsimony tree shows that the LIP 5 data have none of the phylogenetic structure seen in other genes. This pattern, coupled with distribution of substitutions across first, second, and third codon positions and introns, suggests that the data are erroneous (see text).
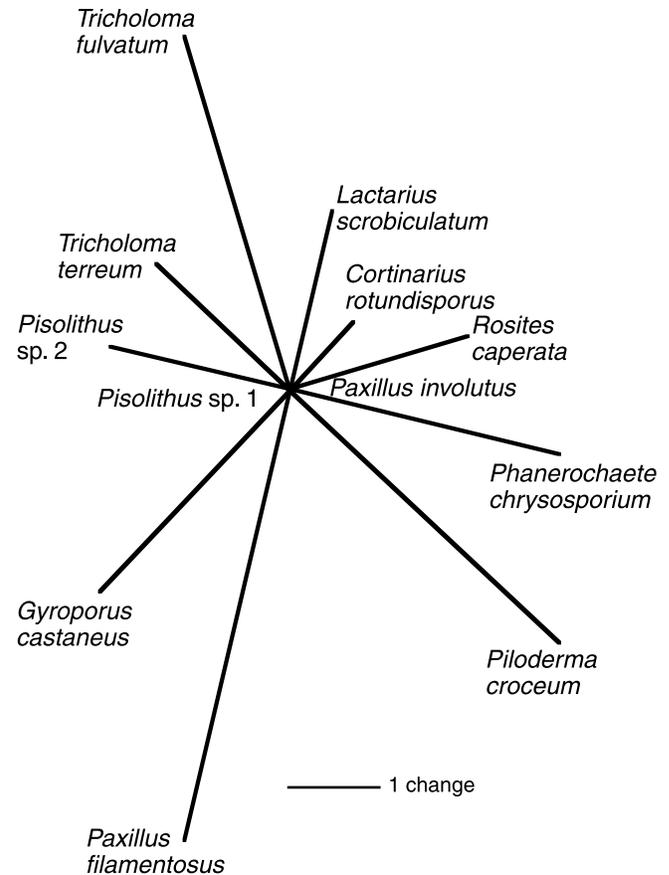
contamination is a common problem that involves previously amplified products that contaminate laboratory equipment and reagents; these amplicons are then reamplified when new samples are tested (Kwok and Higuchi 1989). It is a greater problem with large numbers of cycles or nested PCR reactions, or when no legitimate template exists in the sample.

Errors in identification are certainly common in culture collections and herbaria, and the examples of incorrect sequences that we discussed show us that sequence databases are not really different. In fact, they may be worse in one sense, because sequences are often derived from cultures and collections that could themselves be misidentified. Thus, users need to view taxonomic names associated with sequences with some skepticism and use phylogenetics and other comparative methods to confirm their validity. In addition, sequences that are identified as suspect need to be annotated, just as one would annotate a herbarium collection. Currently, this can only be done by the original submitter of the sequence, and so when one discovers an error, it is necessary to work with the original authors to correct the problem. Direct third-party annotation would certainly accelerate

the rate of error correction. Ultimately, we will need to persuade public database managers to allow this.

Using molecular trees to understand the evolution of the ectomycorrhizal habit

The investigation of the evolution of the EM habit is one of the most relevant uses of phylogenetics to the field of mycorrhizal ecology. A surprising, but inescapable, conclusion has resulted from these studies: EM interactions have evolved independently among multiple lineages of fungi. It is inescapable because the trait is found in the Zygomycota (*Endogone*), multiple lineages of Ascomycota, and multiple lineages of Basidiomycota (Gargas et al. 1995; Bruns et al. 1998; Hibbett et al. 2000). For all of these to have had a common EM ancestor would mean that the trait must have evolved earlier than the Glomeromycota, at a time prior to land plant evolution. It would also require many multiple losses of the EM habit, with maintenance of the trait until the evolution of the holobasidiomycetes. These requirements seem unlikely enough to force us to believe that multiple evolution of the EM fungi has occurred. This conclusion has been reached independently by multiple researchers using different data sets (Gargas et al. 1995; Bruns et al. 1998; Hibbett et al. 2000).

In addition to the parallelism on the fungal side of the interaction, ectomycorrhizae appear to have evolved in parallel within the plants. This also appears to be an inescapable conclusion, because the Pinaceae do not share a unique common ancestor with the angiosperms, and the most basal angiosperms, such as the Magnoliids, do not contain EM taxa. Fitter and Moyersoen (1997) conservatively estimated that three independent origins of the trait are necessary in plants, and they concluded that multiple losses must have occurred within the Angiosperms. Their conclusions were based on the Chase et al. (1993) analysis of *rbcL* sequences in flowering plants. If we examine the more current analysis of flowering plant evolution by Soltis et al. (2000), which is based on three genes, and compare it with the EM plants listed in Smith and Read (1997; Table 6.2), we can identify a minimum of 12 independent groups of angiosperms in which the EM trait occurs (Table 1). Some of these groups, such as the Rosid I clade, themselves contain multiple lineages that may involve independent losses or gains.

The reason we call parallel evolution of ectomycorrhizae shocking is that it means the complex morphology of ectomycorrhizae (i.e., Hartig net, mantel, and root branching patterns) has been reinvented multiple times. Canalization by the plants might help explain some of this morphological convergence, if one lineage of closely related plants was involved, but as discussed previously, the plants themselves are diverse and have also evolved this interaction in parallel. In addition, the fact that many of the fungi involved in forming ectomycorrhizae can form morphologically different interactions with plants such as orchids and liverworts (Warcup 1985; Bidartondo et al. 2003; Taylor and Bruns 1997; Zelmer and Currah 1995) means that EM morphology is not the only possible way for these fungi to interact with plants. Thus, the evolutionary mechanism by which this par-

allelism has evolved remains unclear, even though the pattern of independent gain is inescapable.

One way to lessen the number of convergent gains is to partially explain the pattern by losses. As argued previously, this would not eliminate the need for convergent gains on either the plant or the fungal sides, but it might substantially reduce the number of such events. Parallel losses of complex morphologies and behaviors are generally easier to envision than parallel gains, but in this case loss of the habit would require the plant or fungus to address its nutrient or carbon needs in other ways. On the plant side this is easy to imagine, as several examples exist of plants that can switch between abuscular mycorrhizae and ectomycorrhizae. Thus, loss of the ability to be ectomycorrhizal might be relatively easy for at least some plants to accomplish over evolutionary time, because many may have other options for obtaining their mineral nutrients. Loss of the EM habit on the fungal side might be more difficult, because fungi must also gain (or regain) the ability to obtain fixed carbon in other ways. Although EM fungi often exhibit some saprobic abilities, their abilities appear to be very limited compared with saprobes (Colpaert and Van Laere 1996), and the saprobic abilities they do express may be directed primarily toward extracting mineral nutrients from organic detritus at a net carbon loss, ultimately resulting in reduced rates of decomposition (Gadgil and Gadgil 1971, 1975).

The question of the possible loss of EM habit by fungi was examined by Hibbett et al. (2000) by using parsimony reconstructions of the states, and by testing internal node reconstructions with a likelihood model for the evolution of the mycorrhizal trait (Pagel 1999*b*). In the latter case, the authors compared a model in which gains and losses of the EM habit were equally probable to one in which the EM habit, once gained, was not reversible. The nonreversible model was found to be significantly worse than the reversible model. They concluded that both gains and losses are necessary and estimated that as many as nine losses of the EM habit have occurred. This likelihood approach assumes (1) the phylogenetic tree used is correct in both topology and branch lengths; (2) there is a uniform process of evolution occurring throughout the tree; and (3) there is no correlation between rates of speciation and the presence of a particular state (i.e., EM or nonEM). The first assumption is likely to be violated, because confidence estimates in many of the branches within their displayed tree are low. However, model parameters are usually violated, and the real question is whether or not the violations affect the conclusions. Hibbett et al. (2000) examined alternative trees under various constraints and found that their conclusions were robust to the topologies examined.

To explore this result, we returned to a simple parsimony model and asked the following question: How biased would the process have to be to explain the pattern only by gains of mycorrhizal habit? To examine this question, we accepted their tree as a given and tried different integer weights for the relative cost of gains versus losses of the EM habit. We found that with a 1:2 weighting (i.e., losses are assumed to be twice as hard as gains), all transformations except one were converted to gains. The only remaining loss is that of *Lentaria byssoides*, which in the Hibbett et al. (2000) analy-

Table 1. Independent major clades of angiosperms that contain one or more ectomycorrhizal members.

Lineages[a] and taxa	Comments
1 Eudicots, Rosids, Eurosids I	Several gains or losses within and between orders are likely
Fagales	
Betulaceae,[b] Casuarinaceae,[b] Fagaceae[b], Juglandaceae[b]	
Fabales	
Mimosoideae[b], Caesalpinioideae[b], Papilionoideae	
Rosales	
Rosaceae[b], Ulmaceae, Elaeagnaceae, Ramnaceae	
2 Eudicots, Rosids, Eurosids I	
Malpighiales	Distinct from above groups
Salicaceae[b]	
3 Eudicots, Rosids, Eurosids II	
Malvales	Two gains or multiple losses likely within group
Dipterocarpaceae[b], Cistaceae[b], Tiliaceae[b], Thymeliaceae, Sterculiaceae, (Aceraceae)	
4 Eudicots, Rosids	
Myrtales	
Myrtaceae[b]	From Tree of Life Web site
5 Eudicots	
Euphorbiaceae, (Porantera), Uapacaceae[b]	Unplaced distinct lineage; Euphorabiacae may be polyphyletic
6 Eudicots	
Proteales	
Platanaceae	
7 Eudicots	
Caryophyllales	
Nyctaginaceae,[b] Polygonaceae[b]	Probably represents two independent lineages
8 Eudicots	
Saxifragales	
Hamamelidaceae, Parrotia	Genus unsampled in Soltis et al. (2000), but family was sampled
9 Euasterid II	
Asterales	
Stylidiaceae, Goodeniaceae	
10 Euasterid II	
Dipsacales	
Caprifoliaceae	
11 Eudicotes, asterids	
Ericales	
Ericaceae[b], Epacridaceae, Sapotaceae	Sapotaceae is likely an independent origin (Tree of Life)
12 Eumagnoliids, Monocots, Commelinoids	
Poales	
Cyperaceae, Kobresia[b]	

Note: Monotropoid and Arbutoid mycorrhizae are combined with ectomycorrhizae in this table.

[a]Numbered lineages are separated from other listed lineages by nonmycorrhizal groups and are either well supported by the Soltis et al. (2000) analyses or are present in that analysis and supported by other characteristics (e.g., morphology and chemistry).

[b]Taxa with indisputable ectomycorrhizal members; those in parentheses are less certain.

sis is depicted as derived from within the predominantly mycorrhizal gomphoid lineage. Hibbett et al. (2000) refer to this as the one "unambiguous reversal" within the six lineages of EM fungi examined. The assumption that *Lentaria byssoides* is nonmycorrhizal is based on the fact that it fruits on wood; given that known mycorrhizal taxa such as *Tomentella* are also "lignicolous", this assumption is suspect. How-

ever, even if we accept it as being nonmycorrhizal, its derived condition within the gomphoid lineage appears to be an artifact of limited sampling within that clade, coupled with a locus that is too conserved (i.e., *nuc-LSU*). A larger sample of the Gomphales studied with ITS sequences by Humpert et al. (2001) shows that the section that includes *L. byssoides* is basal to the lineage, as is the ligni-

colous habit. Thus, the EM habit would now look like a gain within the gomphoid clade rather than a loss. This inference is based on simple parsimony, with equal weights to gains and losses.

The justification for using a 1:2 weight for gains versus losses of the EM habit is admittedly arbitrary, although it is certainly no less arbitrary than the assumption of equal weighting, given that we know nothing about the underlying process of gain or loss of EM habit. A more extensive analysis of weighting effects could be conducted (e.g., Ree and Donoghue 1998), but it is clear from our simplified test that a relatively small change in the model can change the outcome of the result and that weights somewhere between 1:1 and 1:2 will cause a shift from numerous losses to none.

In any case, our result brings up the question of why one might want to assume gains of the EM habit are easier than losses for fungi. First, as argued previously, convergent gains have unambiguously occurred based on the broader pattern across all fungi. In contrast, there are no unambiguous examples of losses, such as nonmycorrhizal taxon nested well within a strongly supported clade of mycorrhizal taxa. Second, gaining the EM habit would likely have been a tremendous ecological advantage during key periods of geologic time, whereas regaining saprotrophic abilities would not. More specifically, development of the EM habit would have lead to radiations at the time in history when EM host plants were evolving and expanding their geographic range, while losing the trait would not be expected to have a similar result, as the saprotrophy was already a well-developed strategy with many competitors. Two geological time frames fit such a scenario: the Cretaceous, when the fossil record for the Pinaceae and many of the Rosid lineages first became common, and the Eocene–Oligocene transition, when the climate cooled and Northern Hemisphere forests became dominated by the latter plant groups. The senior author and other colleagues (Bruns et al. 1998) previously observed that many independent lineages of EM fungi appear to radiate at approximately the same time when one examines the mitochondrial LSU data. When this time was estimated using molecular clock analysis of *nuc-SSU* data, the radiations seem to fit within the Eocene–Oligocene time frame (Bruns et al. 1998). This time estimate needs to be revisited now that older estimates for the origin of fungi have been hypothesized (Heckman et al. 2001), newer methods for molecular clock estimates are available (Huelsenbeck et al. 2000), and biases in molecular clock approaches have been identified (Benton and Ayala 2003). However, even if the time estimates change, the point about radiations will remain. In contrast, the loss of EM habit, which is equivalent to the reacquisition of the saprobic habit in Hibbett et al.'s (2000) analysis, would not have provided species with an open niche, as saprobic basidiomycetes had been well established prior to the evolution of the EM habit (Berbee and Taylor 1993, 2001). Thus, losses of EM habit, if they occurred, would not be expected to cause radiations. Note that this hypothesized asymmetry in the effect of gains and losses would be a major violation of the likelihood model used by Hibbett et al. (2000), which assumes there is no correlation between rates of speciation and the presence of either state.

We are not arguing that losses of EM habit among fungi could not have happened, only that Hibbett et al.'s (2000)

conclusion remain equivocal and highly model dependent. The real issue is that modeling morphological and behavioral evolution is a difficult problem, and although one can make likelihood models for these data (Pagel 1999*b*), such models lack the number of observations and the mechanistic knowledge that underlie molecular likelihood models. Similarly, parsimony approaches can produce erroneous conclusions about ancestral states. We think that better data, rather than more sophisticated models, are needed. At a minimum, better resolved trees, better taxon sampling, and improved knowledge of the habits of the sampled fungi will be needed to make these approaches convincing. Perhaps as molecular developmental studies progress, the underlying mechanisms for EM habit convergence will be clarified, and those lineages that share common pathways may look very different from those that have developed the trait independently.

Final thoughts

Molecular systematics is a dynamic and expanding field, and it has changed the way the evolution and ecology of fungi are studied. As large taxon samples and multigene trees accumulate, our evolutionary estimates are improving, and the accumulated data are providing an unparalleled resource for identification of the fungi from environmental samples. In many ways we are still in the golden age of molecular systematics where new and unexpected results are still commonly reported. However, the field has also reached a maturity that now requires massive data and complex analyses to drive the most important advances. Coupled with these advances is the need for users and producers of these studies to be able to independently evaluate both the data and the conclusions.

Acknowledgements

References

Alfaro, M.E., Zoller, S., and Lutzoni, F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. **20**: 255–266.

Becker, D., Ligon, W., Merkey, P., and Ross, R. 1999. Beowulf: low-cost supercomputing using Linux. IEEE Softw. **16**: 79.

Benton, M.J., and Ayala, F.J. 2003. Dating the tree of life. Science (Washington, D.C.), **300**: 1698–1700.

Berbee, M.L., and Taylor, J.W. 1993. Dating the evolutionary radiations of the true fungi. Can. J. Bot. **71**: 1114–1127.

Berbee, M.L., and Taylor, J.W. 2001. Fungal molecular evolution: gene trees and geologic time. *In* The mycota: a comprehensive

treatise on fungi as experimental systems for basic and applied research. Systematics and evolution. Part B. *Edited by* D.J. McLaughlin, E.G. McLaughlin, and P.A. Lemke. Springer-Verlag, Berlin. pp. 229–245.

Berbee, M.L., Carmean, D.A., and Winka, K. 2000. Ribosomal DNA and resolution of branching order among the ascomycota: How many nucleotides are enough? Mol. Phylogenet. Evol. **17**: 337–344.

Bidartondo, M.I., and Bruns, T.D. 2002. Fine-level mycorrhizal specificity in the Monotropoideae (Ericaceae): specificity for fungal species groups. Mol. Ecol. **11**: 557–569.

Bidartondo, M.I., Bruns, T.D., Weiss, M., Sergio, C., and Read, D.J. 2003. Specialized cheating of the ectomycorrhizal symbiosis by an epiparasitic liverwort. Proc. R. Soc. Lond. Ser. B. Biol. Sci. **270**: 835–842.

Binder, M., and Hibbett, D.S. 2002. Higher-level phylogenetic relationships of homobasidiomycetes (mushroom-forming fungi) inferred from four rDNA regions. Mol. Phylogenet. Evol. **22**: 76–90.

Brauer, M.J., Holder, M.T., Dries, L.A., Zwickl, D.J., Lewis, P.O., and Hillis, D.M. 2002. Genetic algorithms and parallel processing in maximum-likelihood phylogeny inference. Mol. Biol. Evol. **19**: 1717–1726.

Bridge, P.D., Roberts, P.J., Spooner, B.M., and Panchal, G. 2003. On the unreliability of published DNA sequences. New Phytol. **160**: 43–48.

Bruns, T.D., Szaro, T.M., Gardes, M., Cullings, K.W., Pan, J.J., Taylor, D.L., Horton, T.R., Kretzer, A., Garbelotto, M., and Li, Y. 1998. A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. Mol. Ecol. **7**: 257–272.

Cairney, J.W.G., Taylor, A.F.S., and Burke, R.M. 2003. No evidence for lignin peroxidase genes in ectomycorrhizal fungi. New Phytol. **160**: 461–462.

Camacho, F.J., Gernandt, D.S., Liston, A., Stone, J.K., and Klein, A.S. 1997. Endophytic fungal DNA, the source of contamination in spruce needle DNA. Mol. Ecol. **6**: 983–987.

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., et al. 1993. Phylogenetics of seed plants — an analysis of nucleotide sequences from the plastid gene *rbcl*. Ann. Mo. Bot. Gard. **80**: 528–580.

Chaverri, P., Castlebury, L.A., Samuels, G.J., and Geiser, D.M. 2003. Multilocus phylogenetic structure within the *Trichoderma harzianum/Hypocrea lixii* complex. Mol. Phylogenet. Evol. **27**: 302–313.

Chen, D.M., Taylor, A.F.S., Burke, R.M., and Cairney, J.W.G. 2001. Identification of genes for lignin peroxidases and manganese peroxidases in ectomycorrhizal fungi. New Phytol. **152**: 151–158.

Colpaert, J.V., and Van Laere, A. 1996. A comparison of the extracellular enzyme activities of two ectomycorrhizal and leaf-saprotrophic basidiomycete colonizing beech leaf litter. New Phytol. **134**: 133–141.

Craven, K.D., Hsiau, P.T.W., Leuchtmann, A., Hollin, W., and Schardl, C.L. 2001. Multigene phylogeny of *Epichloe* species, fungal symbionts of grasses. Ann. Mo. Bot. Gard. **88**: 14–34.

Cruse, M., Telerant, R., Gallagher, T., Lee, T., and Taylor, J.W. 2002. Cryptic species in *Stachybotrys chartarum*. Mycologia, **94**: 814–822.

Cullings, K.W., and Vogler, D.R. 1998. A 5.8S nuclear ribosomal RNA gene sequence database: applications to ecology and evolution. Mol. Ecol. **7**: 919–923.

Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F., and Douzery, E.J.P. 2003. Comparison of Bayesian and maximum

likelihood bootstrap measures of phylogenetic reliability. Mol. Biol. Evol. **20**: 248–254.

Farris, J.S. 1997. The future of phylogeny reconstruction. Zool. Scr. **26**: 303–311.

Felsenstein, J. 1978. The number of evolutionary trees. Syst. Zool. **27**: 27–33.

Felsenstein, J. 1985. Confidence intervals on phylogenies: an approach using the bootstrap. Evolution, **39**: 783–791.

Fitter, A.H., and Moyersoen, B. 1997. Evolutionary trends in root–microbe symbioses. *In* Plant life histories: ecology, phylogeny and evolution. *Edited by* J.F. Silvertown, M. Franco, and J.L. Harper. Cambridge University Press, Cambridge. pp. 265–282.

Gadgil, R.L., and Gadgil, P.D. 1971. Mycorrhiza and litter decomposition. Nature (Lond.), **233**: 133.

Gadgil, R.L., and Gadgil, P.D. 1975. Suppression of litter decomposition by mycorrhizal roots of *Pinus radiata*. N.Z. J. For. Sci. **5**: 33–41.

Gargas, A., Depriest, P.T., Grube, M., and Tehler, A. 1995. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. Science (Washington, D.C.), **268**: 1492–1495.

Grimshaw, A.S., Nguyentuong, A., Lewis, M.J., and Hyett, M. 1997. Campus-wide computing — early results using legion at the University of Virginia. Int. J. Supercomp. Appl. **11**: 129–143.

Grubisha, L.C., Trappe, J.M., Molina, R., and Spatafora, J.W. 2002. Biology of the ectomycorrhizal genus *Rhizopogon*. VI. Re-examination of infrageneric relationships inferred from phylogenetic analyses of ITS sequences. Mycologia, **94**: 607–619.

Hambleton, S., Egger, K.N., and Currah R.S. 1998. The genus *Oidiodendron*: species delimitation and phylogenetic relationships based on nuclear ribosomal DNA analysis. Mycologia, **90**: 854–869.

Heckman, D.S., Geiser, D.M., Eidell, B.R., Stauffer, R.L., Kardos, N.L., and Hedges, S.B. 2001. Molecular evidence for the early colonization of land by fungi and plants. Science (Washington, D.C.), **293**: 1129–1133.

Hibbett, D.S., Gilbert, L.-B., and Donoghue, M.J. 2000. Evolutionary instability of ectomycorrhizal symbioses in basidiomycetes. Nature (Lond.), **407**: 506–508.

Hillis, D.M. 1996. Inferring complex phylogenies. Nature (Lond.), **383**: 130–131.

Hintz, W.E. 1999. Sequence analysis of the chitin synthase A gene of the Dutch elm pathogen *Ophiostoma novo-ulmi* indicates a close association with the human pathogen *Sporothrix schenckii*. Gene (Amst.), **237**: 215–221.

Hirt, R.P., Logsdon, J.M., Healy, B., Dorey, M.W., Doolittle, W.F., and Embley, T.M. 1999. Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. U.S.A. **96**: 580–585.

Hoiland, K., and Holst-Jensen, A. 2000. *Cortinarius* phylogeny and possible taxonomic implications of ITS rDNA sequences. Mycologia, **92**: 694–710.

Huelsenbeck, J.P., Larget, B., and Swofford, D. 2000. A compound Poisson process for relaxing the molecular clock. Genetics, **154**: 1879–1892.

Huelsenbeck, J.P, Larget, B., Miller R.E., and Ronquist, F. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. **51**: 673–688.

Humpert, A.J., Muench, E.L., Giachini, A.J., Castellano, M.A., and Spatafora, J.W. 2001. Molecular phylogenetics of *Ramaria* and related genera: evidence from nuclear large subunit and mitochondrial small subunit rDNA sequences. Mycologia, **93**: 465–477.

Keeling, P.J. 2003. Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet. Biol. **38**: 298–309.

Koljalg, U., Dahlberg, A., Taylor, A.F.S., Larsson, E., Hallenberg, N., Stenlid, J., Larsson, K.H., Fransson, P.M., Karen, O., and Jonsson, L. 2000. Diversity and abundance of resupinate thelephoroid fungi as ectomycorrhizal symbionts in Swedish boreal forests. Mol. Ecol. **9**: 1985–1996.

Kretzer, A., and Bruns, T.D. 1999. Use of *atp*6 in fungal phylogenetics: an example from the Boletales. Mol. Phylogenet. Evol. **13**: 483–492.

Kretzer, A., Li, Y., Szaro, T., and Bruns, T.D. 1996. Internal transcribed spacer sequences from 38 recognized species of *Suillus* sensu lato: phylogenetic and taxonomic implications. Mycologia, **88**: 776–785.

Kwok, S., and Higuchi, R. 1989. Avoiding false positives with PCR. Nature (Lond.), **339**: 237–238.

Landvik, S., Eriksson, O.E., and Berbee, M.L. 2001. *Neolecta* — a fungal dinosaur? Evidence from beta-tubulin amino acid sequences. Mycologia, **93**: 1151–1163.

Lewis, P.O. 2001. Phylogenetic systematics turns over a new leaf. Trends Ecol. Evol. **16**: 30–37.

Liu, Y.J.J., Whelen, S., and Benjamin, D.H. 1999. Phylogenetic relationships among ascomycetes: Evidence from an RNA polymerase II subunit. Mol. Biol. Evol. **16**: 1799–1808.

Mankel, A., Kost, G., and Kothe, E. 1999. Re-evaluation of the phylogenetic relationship among species of the genus *Tricholoma*. Microbiol. Res. **153**: 377–388.

Martin, F., Diez, J., Dell, B., and Delaruelle, C. 2002. Phylogeography of the ectomycorrhizal *Pisolithus* species as inferred from nuclear ribosomal DNA ITS sequences. New Phytol. **153**: 345–357.

Matheny, P.B., Liu, Y.J.J., Ammirati, J.F., and Hall, B.D. 2002. Using RPB1 sequences to improve phylogenetic inference among mushrooms (*Inocybe*, Agaricales). Am. J. Bot. **89**: 688–698.

Mehmann, B., Brunner, I., and Braus, G.H. 1994. Nucleotide sequence variation of chitin synthase genes among ectomycorrhizal fungi and its potential use in taxonomy. Appl. Environ. Microbiol. **60**: 3105–3111.

Miller, S.L., and Buyck, B. 2002. Molecular phylogeny of the genus *Russula* in Europe with a comparison of modern infrageneric classifications. Mycol. Res. **106**: 259–276.

Moncalvo, J.M., Lutzoni, F.M., Rehner, S.A., Johnson, J., and Vilgalys, R. 2000. Phylogenetic relationships of agaric fungi based on nuclear large subunit ribosomal DNA sequences. Syst. Biol. **49**: 278–305.

Moncalvo, J.M., Vilgalys, R., Redhead, S.A., Johnson, J.E., James, T.Y., Aime, M.C., Hofstetter, V., Verduin, S.J.W., Larsson, E., Baroni, T.J., Thorn, R.G., Jacobsson, S., Clemencon, H., and Miller, O.K. 2002. One hundred and seventeen clades of euagarics. Mol. Phylogenet. Evol. **23**: 357–400.

Nixon, K.C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics, **15**: 407–414.

O'Donnell, K. 2000. Molecular phylogeny of the *Nectria haematococca – Fusarium solani* species complex. Mycologia, **92**: 919–938.

O'Donnell, K., Cigelnik, E., and Casper, H.H. 1998*a*. Molecular phylogenetic, morphological, and mycotoxin data support re-identification of the quorn mycoprotein fungus as *Fusarium venenatum*. Fungal Genet. Biol. **23**: 57–67.

O'Donnell, K., Kistler, H.C., Cigelnik, E., and Ploetz, R.C. 1998*b*. Multiple evolutionary origins of the fungus causing panama-disease of banana — concordant evidence from nuclear and mi-tochondrial gene genealogies. Proc. Natl. Acad. Sci. U.S.A. **95**: 2044–2049.

O'Donnell, K., Lutzoni, F.M., Ward, T.J., and Benny, G.L. 2001. Evolutionary relationships among mucoralean fungi (Zygomycota): evidence for family polyphyly on a large scale. Mycologia, **93**: 286–297.

Pagel, M. 1999*a*. Inferring the historical patterns of biological evolution. Nature (Lond.), **401**: 877–884.

Pagel, M. 1999*b*. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst. Biol. **48**: 612–622.

Peintner, U., Bougher, N.L., Castellano, M.A., Moncalvo, J.-M., Moser, M.M., Trappe, J.M., and Vilgalys, R. 2001. Multiple origins of sequestrate fungi related to Cortinarius (Cortinariaceae). Am. J. Bot. **88**: 2168–2179.

Pollock, D.D., Zwickl, D.J., McGuire, J.A., and Hillis, D.M. 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. **51**: 664–671.

Redecker, D., Hijri, M., Dulieu, H., and Sanders, I.R. 1999. Phylogenetic analysis of a data set of fungal 5.8S rDNA sequences shows that highly divergent copies of internal transcribed spacers reported from *Scutellospora castanea* are of ascomycete origin. Fungal Genet. Biol. **28**: 238–244.

Ree, R.H., and Donoghue, M.J. 1998. Step matrices and the interpretation of homoplasy. Syst. Biol. **47**: 582–588.

Salamin, N., Chase, M.W., Hodkinson, T.R., and Savolainen, V. 2003. Assessing internal support with large phylogenetic DNA matrices. Mol. Phylogenet. Evol. **27**: 528–539.

Schüßler, A., Schwarzott, D., and Walker, C. 2003. Glomeromycota rRNA genes — the diversity of myths? Mycorrhiza, **13**: 233–236.

Sikes, D.S., and Lewis, P.O. 2001. Beta software, version 1, PAUPRatL PAUP* implementation of the parsimony racket [computer program]. Distributed by authors. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Conn.

Smith, D.E., and Klein, A.S. 1994. Phylogenetic inferences on the relationship of north American and European *Picea* species based on nuclear ribosomal 18S sequences and the internal transcribed spacer 1 region. Mol. Phylogenet. Evol. **3**: 17–26.

Smith, S.E., and Read, D.J. 1997. Mycorrhizal symbiosis. Academic Press, New York.

Snell, Q., Judd, G., and Clement, M. 2001. The DOGMA approach to parallel and distributed computing. *In* Cluster computing. *Edited by* R. Buyya and C. Szyperski. Nova Science Publishers, Huntington, N.Y. pp. 23–34.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., and Farris, J.S. 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. Bot. J. Linn. Soc. **133**: 381–461.

Suzuki, Y., Glazko, G.V., and Nei, M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. U.S.A. **99**: 16 138–16 143.

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

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

Taylor, D.L., and Bruns, T.D. 1997. Independent, specialized invasions of ectomycorrhizal mutualism by two non-photosynthetic orchids. Proc. Natl. Acad. Sci. U.S.A. **94**: 4510–4515.

Turgeon, B.G. 1998. Application of mating type gene technology to problems in fungal biology. Annu. Rev. Phytopathol. **36**: 115–137.

Vos, R.A. 2003. Accelerated likelihood surface exploration: the likelihood ratchet. Syst. Biol. **52**: 368–373.

Warcup, J.H. 1985. *Rhizanthella gardneri* (Orchidaceae), its *Rhizoctonia* endophyte and close association with *Melaleuca uncinata* (Myrtaceae) in western Australia. New Phytol. **99**: 273–280.

Witthuhn, R.C., Harrington, T.C., Steimel, J.P., Wingfield, B.D., and Wingfield, M.J. 2000. Comparison of isozymes, rDNA spacer regions and MAT-2 DNA sequences as phylogenetic characters in the analysis of the *Ceratocystis coerulescens* complex. Mycologia, **92**: 447–452.

Zelmer, C.D., and Currah, R.S. 1995. Evidence for a fungal liaison between *Corallorhiza trifida* (Orchidaceae) and *Pinus contorta* (Pinaceae). Can. J. Bot. **73**: 862–866.

Zhang, W., Wendel, J.F., and Clark, L.G. 1997. Bamboozled again. Inadvertent isolation of fungal rDNA sequences from bamboos (Poaceae: Bambusoideae). Mol. Phylogenet. Evol. **8**: 205–217.