Mycologia, 98(6), 2006, pp. 838–849.

Dating divergences in the Fungal Tree of Life: review and new analyses

John W. Taylor[1]
Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, California 94720-3102

Mary L. Berbee
Department of Botany, University of British Columbia, Vancouver, British Columbia, V6T 1Z4 Canada

Abstract: The collection of papers in this issue of *Mycologia* documents considerable improvements in taxon sampling and phylogenetic resolution regarding the Fungal Tree of Life. The new data will stimulate new attempts to date divergences and correlate events in fungal evolution with those of other organisms. Here, we review the history of dating fungal divergences by nucleic acid variation and then use a dataset of 50 genes for 25 selected fungi, plants and animals to investigate divergence times in kingdom Fungi. In particular, we test the choice of fossil calibration points on dating divergences in fungi. At the scale of our analysis, substitution rates varied without showing significant within-lineage correlation, so we used the Langley-Fitch method in the R8S package of computer programs to estimate node ages. Different calibration points had a dramatic effect on estimated divergence dates. The estimate for the age of the Ascomycota/Basidiomycota split was 1 808 000 000 y ago when calibrated assuming that mammals and birds diverged 300 000 000 y ago, 1 489 000 000 y ago when calibrated assuming that the 400 000 000 y old fungal fossil *Paleopyrenomycites devonicus* represents Sordariomycetes and ~400 000 000 y ago when calibrated assuming 206 000 000 y ago for the plant eudicot/monocot divergence. An advantage of a date of ~400 000 000 y ago for the Ascomycota/Basidiomycota divergence is that the radiation of fungi associated with land plants would not greatly precede the earliest land plant fossils. Acceptance of ~400 000 000 y ago for the Ascomycota /Basidiomycota split would require that *P. devonicus* be considered a deeply branching Ascomycota. To improve on current estimates of divergence times, mycologists will require calibration points from within groups of fungi that share similar substitution rates. The most useful calibration is likely to depend on the discovery and description of continuous records of fossil fungi, or their spores, that show recognizable shifts in morphology.

Key words: dating divergences, fossil fungi, molecular phylogeny

INTRODUCTION AND REVIEW

Evolution is change over time and time is the common currency that lets evolutionary biologists compare historical events in their group with events in other groups. Dating evolutionary events began with evaluation of the geologic record and fossils, but when genetic variation could be conveniently detected, initially by protein electrophoresis, it also was employed to date evolutionary events (Zuckerkandl and Pauling 1965, Sarich and Wilson 1967). Access to nucleic acid sequence made the genetic contribution only stronger and now it is with a combination of paleontology and molecular evolution that biologists attempt to date divergences in the Tree of Life from near the present to the period of the last common ancestor of life. Estimates of dates for events that occurred as many or more than 1 000 000 000 y ago cannot be exact and it is not surprising that there are competing estimates that differ greatly in time. Here we review the history of dating fungal divergences and possible reasons for discrepancies in the process in the hope that recent advances in methods of estimating divergences from molecular data and critical evaluation of fossils might help resolve this controversy. After this review we examine substitution rate heterogeneity and the effect of fossil calibration on dating fungal divergences with sequence for 50 genes sampled from available genome sequences of animals, plants and a broad sampling of fungi.

Mycologists were quick to use nucleic acid sequence variation to date important events in the history of fungal evolution and to use those dates to compare events in fungal evolution to those of plants and animals. Fungi are good subjects for such studies because they are for the most part haploid, have relatively few multigene families and many taxa are large enough, particularly compared to other microbes, to leave useful fossils, most often associated with plants (Taylor et al 2005). The field was inaugurated by Simon et al (1993), who asked if the radiation of arbuscular mycorrhizal fungi coincided with the appearance of green plants on land. Their analysis of 12 Glomeromycota small subunit ribosomal DNA (SSU rDNA) sequences, using parsimony analysis and calibration points of 200 000 000 y ago

Accepted for publication 25 September 2006.
[1] Corresponding author. E-mail: jtaylor@nature.berkeley.edu

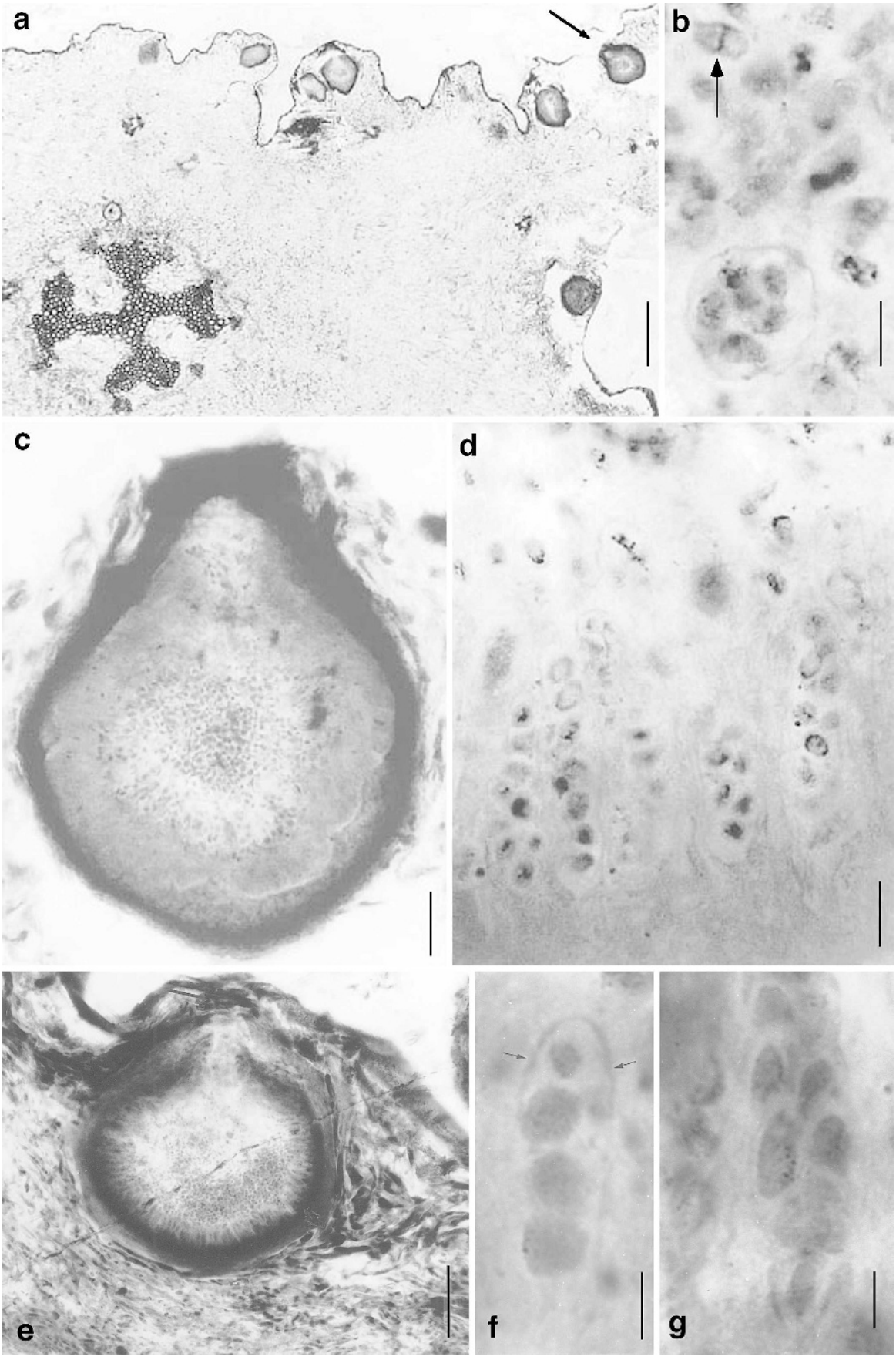
for the origin of monocots and 1 000 000 000 y ago for the divergence of fungi and plants, dated the radiation of Glomeromycota at 462 000 000–353 000 000 y ago. These dates were compared to fossils of arbuscular mycorrhizae in the rhizomes of plants in the Rhynie chert (Pirozynski and Dalpe 1989, Remy et al 1994) to support the hypothesis that mycorrhizal mutualisms of plants and arbuscular mycorrhizal fungi conquered land together.

Also in 1993, Berbee and Taylor fit a broad phylogenetic analysis of fungi to the geologic timescale. They also used SSU rDNA sequences and analysis by parsimony and neighbor joining, the latter with maximum likelihood distances, to determine a phylogeny for fungi. They sorted fungi into nine lineages and showed, via relative rate tests, that the lineages had different rates of nucleotide substitution. A mean rate was calculated and the rates on each lineage were normalized to make a tree with one global rate of nucleotide substitution (1.0% per 100 000 000, for SSU rDNA). This tree was fit to the geologic timescale using fungal fossils, fossils of organisms associated with fungi (beetle galleries) or radiations of organisms mutualistic with fungi (ruminates). From these calibrations, the divergence of fungi and animals was estimated to have occurred at 600 000 000 y ago.

The publications of 1993 have been influential, judging from their citations (Simon et al [1993] at 224 and Berbee and Taylor [1993] at 200), and these numbers show that dating evolutionary events has become an integral part of fungal phylogenetic studies, nowhere more important than with symbioses whether they involve mycorrhizae, insects or humans. In addition to the application of molecular analyses, new fossil finds have been reported; especially noteworthy were two fossils, one identified as Glomeromycota from 460 000 000 y ago (Redecker et al 2000) and another identified as a perithecial ascomycete and classified as a pyrenomycete (Sordariomycetes) that was found in the stem and rhizomes of the fossil plant *Asteroxylon mackie* from 400 000 000 y ago (Taylor et al 1999), which later was described as *Paleopyrenomycites devonicus* (Taylor et al 2005) (FIG. 1).

In terms of molecular analyses, it was not until 2001 that the field shifted direction as much as in 1993, by which time phylogenetic methods employing likelihood had been developed that could enforce a constant molecular clock. Berbee and Taylor (2001) revisited fitting the fungal SSU rDNA tree to geologic time with maximum likelihood analysis and calibrating dates on this tree using an earlier date for the divergence of animals and fungi, 965 000 000 y ago, which had been determined in an independent study (Doolittle et al 1996). This calibration increased the gap between molecular and fossil estimates of divergences. However, as noted by the authors, this analysis failed to adequately account for the aforementioned *Paleopyrenomycites* fossil as a sordariomycete from 400 000 000 y ago (Berbee and Taylor 2001). Heckman et al (2001) took the most innovative step in molecular analysis in 2001 when they revisited the evolution of land plants and associated arbuscular mycorrhizae by expanding the nucleic acid database to include multiple protein coding genes and estimating divergence times either by averaging the dates estimated by the different protein genes or by concatenating them before averaging. The topology estimated by the proteins was not in conflict with that estimated from SSU rDNA, but the divergence times were much older, further increasing the gap between divergences estimated from nucleic acid variation and those estimated from fossils. For example the radiation of Glomeromycota (ca. 1 200 000 000 y ago) now would predate the appearance of arbuscular mycorrhizal fossils (400 000 000 y ago) by ca. 800 000 000 y; however the new molecular divergence dates accommodated the 400 000 000 y old *Paleopyrenomycites* fossil. The key reason that the divergence times of Heckman et al were much older is that a previous study by the same research group, using the same multiprotein approach and fossil evidence for the divergence of mammals and birds, estimated the divergence of animals and fungi at ca. 1 576 000 000 y ago (Wang et al 1999).

Fossils define minimum ages of divergence.—Some of the gap between molecular and fossil dates certainly can be attributed to the fact that fossils usually provide only a minimum age for divergences. There are several well understood reasons for this observation, which have been explored in depth for plants (Magallon 2004). Simply put, all divergences begin when an ancestral species splits to form progeny species; as a result, until morphological differences characteristic of the descendent groups develop, fossils of the different clades will not be recognizable as such. In addition neither fossilization nor fossil discovery favor preservation and recovery of the earliest fossils for any lineage. The problem is compounded when the organisms of interest are microbes, but the problems remain even for charismatic macrofauna. For example there is controversy even about the vertebrate fossils that document the divergence of birds and mammals, which provide the geologic calibration for the 1 576 000 000 y old estimate of animal-fungal divergence (Wang et al 1999). Graur and Martin (2004) argue that the earliest fossils on the two branches, synapsid (leading


a
b
c
d
e
f
g

to mammals) and diapsid (leading to birds and lizards), cannot be identified indisputably and that their divergence might have occurred earlier or later than 310 000 000 y ago (i.e. 338 000 000–288 000 000 y ago). Reisz and Müller (2004) also question the 310 000 000 y date for the divergence of the progenitors of the mammal and bird lineages and advocate using the divergence between crocodiles and lizards. However in both cases the discrepancy is small, less than 10%, and would hardly affect the date for the animal-fungal divergence. Furthermore Hedges and Kumar (2004) defend their use of the mammal-bird divergence at 310 000 000 y ago as well documented and note that molecular data for mammals and birds are superior to those for crocodiles.

Bias in estimating a mean global evolution rate.—Rodriguez-Trelles et al (2002) noted that frequency distributions of divergence times calculated for multiple genes showed a skew toward older times because recent times are constrained by the present, whereas older times have no such constraint. They demonstrated the effects of this phenomenon using simulations of divergence times for proteins of four lengths evolving at three different rates of evolution (with correlated variation in rates among sites). They used four-taxon trees consisting of three ingroup taxa and the root; in these trees the recent divergence was constrained to 300 000 000 y ago while the older divergence was placed at one of three dates (600 000 000, 1 200 000 000 and 3 000 000 000 y ago). The most extreme disparity between the age of the divergence used to simulate the data and the average rate estimated from 1000 simulations was a 25% overestimate for short sequences evolving at slow rates over long periods. If, for example, this extreme correction were applied to the mean age of the animal-fungal divergence estimated by Wang et al (1999) it would be moved forward from 1 600 000 000 to 1 200 000 000 y ago, a significant change.

Global substitution rates are unrealistic.—The studies of Berbee and Taylor (2001) and Heckman et al (2001) both used a single global rate of nucleotide substitution to estimate branch lengths and divergence times, even though it was clear that substitution rates differed among fungal lineages (Berbee and Taylor 1993), as had earlier been shown for mammals (Wu and Li 1985). At the time of these studies methods had not been developed for accounting for substitution rate variation on different branches, but the development of such methods now makes it possible to do so. One approach to detecting rate variation depends on the availability of calibration points such as fossils that can be used to date nodes. Substitution rate changes can be localized on a phylogeny by taking the calibrations into account when optimizing the likelihood of the data. Another approach involves phylogenetic prediction of the pattern of variation. A key assumption here is that, although each branch can have its own rate, the rates in progeny species after a divergence likely are correlated to that of the ancestral species and to each other, a phenomenon known as autocorrelation. In the analyses presented below we use methods that allow for calibration to the fossil record and we test for autocorrelation of rates.

In recent studies examining the fit of phylogenetic trees to geologic time, three related methods have proved most popular, Langley-Fitch (LF), penalized likelihood (PL) and the Bayesian relaxed clock (BRC). Unlike older programs such as PHYLIP's DNAMLK (Felsenstein 2005), LF, PL and BRC all take into account the fit of calibration points to nodes in the tree. LF and PL are implemented by the R8S package of programs (available from M. Sanderson at http://ginger.ucdavis.edu/r8s/). LF uses maximum likelihood to fit the data to a molecular clock with a single global substitution rate. PL also seeks to maximize the likelihood of the sequence data given the evolutionary model while minimizing the penalty for rate variation (lack of autocorrelation) after divergence. For datasets where rate variation is correlated with the phylogeny, the fit of the data to the clock tree will be better with PL than with LF. BRC methods (e.g. MULTIDIVTIME Thorne and Kishino 2002) produce a probability distribution for substitution rates and divergence dates, both also based on likelihood. Useful reviews of these methods and

←

FIG. 1. *Paleopyrenomycites devonicus*, Lower Devonian ascomycete. a. Section of stem with ascocarps (arrow) in cortex. b. Cross-section of ascus (lower left) with ascospores, some of which are bicelled (arrow). c. Perithecium in longitudinal section. d. Asci containing ascospores arising from inner wall of perithecium. e. Medial longitudinal section of perithecium showing central cavity and ostiole. f. Four-nucleate (?) stage in immature ascus. Note slight invagination at the ascus tip (arrows) that might represent an early stage in the development of the collar. g. Mature ascus showing biseriate arrangement of ascospores. Bars: a = 500 μm; b = 20 μm; c = 50 μm; d = 10 μm; e = 100 μm; f = 5 μm; g = 5 μm. Figures and legends a–d are from Taylor et al (1999) and e–g are from Taylor et al. (2005).

others recently have been published by Sanderson (2004), Rutschmann (2006) and Welch and Bromham (2005). Users of LF and PL must specify a tree topology with branch lengths. PL requires an additional rate smoothing parameter that reflects the amount of autocorrelation. At its extremes, this parameter could specify one global rate or as many different rates as there are branches in the tree. Users also must specify either the substitution rate on the branch from the ancestral node to the ingroup or the date of this node. To fit the results to geologic time, a date for at least one divergence must be specified. Depending on the patterns of rate variation, satisfying this requirement may involve fixing the age of one or more nodes or specifying a combination of minimum and maximum ages for nodes. Users of Bayesian methods must specify a tree topology, an autocorrelation parameter (such as the PL rate smoothing parameter) and prior distributions of rates and dates. Again, to correlate the divergences with geologic time, maximum and minimum dates for at least one divergence must be specified. Needless to say, a priori these values might not be known and their estimates can affect the outcome, a point that we will return to later.

Rate heterogeneity and calibration of molecular trees to geologic time.—We will consider two aspects of substitution rate heterogeneity that can affect the dates of divergences; the first is the use of fossils from one kingdom to calibrate dates in another, and the second is the use of fossils from one kingdom to calibrate dates in the same kingdom. Implicit in the use of calibration dates from one kingdom (e.g. Animalia) for analysis in another (e.g. Fungi) is that rates of substitution are similar in both phyla. If they are not similar the resulting dates will be inaccurate in proportion to the disparity. This problem first was investigated by Peterson et al (2004), who hypothesized that molecular clock estimates for dipteran ages have been far older than the corresponding dipteran fossils because the clock estimates were calibrated with vertebrate fossils. Employing sequences of multiple protein-coding genes, they demonstrated rate heterogeneity and showed that the rate of substitution on the vertebrate lineage was slow compared to that on the several invertebrate lineages, including dipterans. When they then used the LF method calibrated with invertebrate fossils, they found that estimated divergences among invertebrates were much more recent that those calibrated by vertebrate fossils (e.g. the last common ancestor of bilatera moved forward in time from 1 576 000 000 y ago to 900 000 000 y ago. In the analyses that follow, we investigate the effect of using calibrations outside and inside the fungal kingdom.

If calibration using a sister kingdom is misleading, the alternative is to calibrate with fossils from within the kingdom of interest. Much of this effort has focused on plants. Sanderson (2003), who developed PL, applied it to a dataset of plastid protein genes sequenced from 10 plants with a green algal outgroup. The input tree was made with parsimony analysis of inferred amino acid sequences and branch lengths were estimated using maximum likelihood (PAML Yang 1997). For calibration to geologic time, Sanderson used 330 000 000 y ago for the node of crown seed plants and secondarily 125 000 000 y ago for the node of crown group eudicot angiosperms. With the seed plant calibration and a global substitution rate, the crown group of land plants was placed at 435 000 000 y; with PL, which accommodates different rates, the date was 483 000 000 y ago. Adding the secondary calibration hardly affected the dates. With or without a global substitution rate (i.e. with or without a uniform molecular clock) dates for the crown group of land plants were much more similar to one another than either was to the date from the study of Heckman et al (2001), which was 700 000 000 y ago. Although the dataset of this study and that of Heckman et al were different, the most important difference in the studies is likely to be calibration to geologic time, which was based on plant fossils in Sanderson's study and based on vertebrates in the study of Heckman et al (2001). This difference in divergence times is similar to that found for vertebrates and invertebrates (Peterson et al 2004) and it raises a question about the effects of different calibration times for fungi that we address in our analyses. Also noteworthy is that Sanderson did not find a significant difference in divergence times when rates were allowed to vary over the tree or were made global.

In another study within the plant kingdom, Bell et al (2005) estimated the age of the angiosperms with four genes (two chloroplast, one mitochondrial and one nuclear), while comparing PL and BRC methods using calibration dates from the plant fossil record. Their comparison provided several results that are of general interest. PL gave older molecular divergences than BRC (i.e. the age of the angiosperm crown group for BRC was 122 000 000–202 000 000 y ago and for PL it was 150 000 000–275 000 000 y ago). With both methods, analysis of combined data gave more reliable results than each gene separately and rDNA gave older divergences than protein-coding loci. The rate of substitution specified for the root of the ingroup had a large influence on divergence times, larger than specification of the rate variation (autocorrelation) parameter. Finally, and not surprisingly, fixing the age of eudicots or specifying minimum or maximum ages for four divergences

within angiosperms provided the best agreement with the fossil record, again showing that calibration to the geologic record has a strong effect on dates of divergences. Bell et al (2005) point out that the origin of eudicot plants is ''one of the firmest dates from the fossil record because of the numerous reports of fossil tricolpate pollen, with no tricolpate pollen appearing before this time point.''

Recent fungal studies.—The first study to include fungi and to use methods that can accommodate rate variation was the broad analysis of eukaryote divergence dates with BRC by Douzery et al (2004). Included among 36 eukaryotes were one representative of Basidiomycota and three of Ascomycota, and for each taxon as many as 129 protein amino acid sequences were used (albeit 25% of the character states were missing). In this study the number and composition of minimum and maximum dates for six divergences had a stronger effect on the posterior probabilities for divergences than did prior probabilities of tree topology (i.e. root placement), substitution rate at the root, time from root to tip or the rate autocorrelation parameter. Among the six fossils used to calibrate divergences was *Paleopyrenomycites,* which provided a minimum age of 400 000 000 y for Ascomycota (Taylor et al 1999). Using these calibration dates BRC analysis dated the divergence of animals and fungi at 984 000 000 y ago, the divergence of Ascomycota and Basidiomycota at 727 000 000 y ago and the divergence of *Saccharomyces cerevisiae* and *Candida albicans* at 235 000 000 y ago. By way of comparison, the mean dates from Wang et al (1999) and Heckman et al (2001) for these divergences were 1 564 000 000, 1 208 000 000 and 841 000 000 y ago.

Just a year later Padovan et al (2005) used PL to analyze divergence times for 166 fungal SSU rDNA sequences. As a geologic calibration they used *Paleopyrenomycites* at 400 000 000 y as a minimum date for Sordariomycetes, and they compared results when the animal-fungal divergence was specified as either 1 576 000 000 y ago (Wang et al 1999) or 965 000 000 y ago (Doolittle et al 1996). When 1 576 000 000 y was used for the animal-fungal divergence, the Ascomycota-Basidiomycota divergence was estimated at 1 206 000 000 y ago, with the animal-fungal divergence date at 965 000 000 y it was 786 000 000 y ago. Both dates are substantially older than the ca. 560 000 000 y estimated by Berbee and Taylor (2001), who also used 965 000 000 y for the animal-fungal divergence and SSU rDNA sequences. Padovan et al (2005) note that the most important difference between their PL analysis and Berbee's and Taylor's ML analysis is the calibration of a minimum age for Sordariomycetes of 400 000 000 y ago. As they state, acceptance of dates from Berbee and Taylor (2001) would ''… impose a reclassification of the fossil record.''

Taking the statement of Padovan et al (2005) to heart, and following the approach of Peterson et al (2004) for investigating calibration from outside and inside a kingdom, we augmented a previously published 50 amino acid dataset (Rokas et al 2005) and used it to test the effect of different fossil calibrations on fungal divergence times.

MATERIALS AND METHODS

Sequence data.—We began with an alignment of 50 concatenated amino acid sequence regions from Rokas et al (2005). The initial alignment included 15 fungal species. We excluded 12 animal sequences from this alignment because our emphasis was on fungi but, to provide calibration points, we then added sequences from rice (*Oryza sativa*), *Arabidopsis thaliana*, chicken (*Gallus gallus*) and mosquito (*Anopheles gambiae*). We also added *Rhizopus oryzae* as an example of an early diverging fungus and *Plasmodium falciparum* to serve as outgroup. Sequences and annotations from rice were from the Institute for Genomic Research (TIGR), from the Jan 2006 annotation release 4, available at http://www.tigr.org/tdb/e2k1/osa1/index.shtml. We retrieved published sequences and annotations of the *Arabidopsis thaliana* (Arabidopsis Genome 2000) the chicken (Hillier et al 2004) the mosquito *Anopheles gambiae* (Holt et al 2002) and *Plasmodium falciparum* (Gardner et al 2002) from genomic sequences in GenBank. The *R. oryzae* sequences and annotations are from the Broad Institute Fungal Genome Initiative Website http://www.broad.mit.edu/annotation/fgi/, from the Nov 2005 assembly version 3 (RO3) strain RA99-880).

Selecting putative orthologues.—We used reciprocal orthology to select genes from the six additional taxa for inclusion in the matrix of Rokas et al (2005), which is similar to the strategy used to assemble the original dataset. To begin the search we extracted three sequences from Rokas et al (2005) to serve as representatives for Basidiomycota (*Cryptococcus neoformans*), Ascomycota (yeast, *Saccharomyces cerevisiae*) and animals (the fish *Danio rerio*). Using standalone BLASTALL from the program BLAST-2.2.13 (available from National Center for Biotechnology Information, http://130.14.29.110/BLAST/download.shtml) we first queried each of the six additional genomes with each of the 50 genes from the three representatives of the Rokas et al (2005) dataset. If the gene sequence from the basidiomycete, the ascomycete and the fish all retrieved the same gene as the top blast hit from a subject genome, then that gene qualified for further analysis. We next retrieved the sequence for the gene corresponding to the top blast hit from the subject genome and used it to query the basidiomycete, yeast and fish genomes. When a gene from an additional genome returned the original gene from each of three taxa, it was added to the alignment.

We modified this approach slightly for three genomes that had 2–6 almost equally good and almost identical ''top blast hits'' for every query sequence, possibly due to recent polyploidization, *Arabidopsis*, rice and *Rhizopus*. Where the top blast hits to queries by the basidiomycete, yeast and fish were identical, or differed by blast match scores of less than 30, we considered the variants to be recent paralogues and arbitrarily selected one for inclusion in the dataset. That is, for purposes of this analysis, such paralogues will be considered orthologous with other sequences in the dataset. As a final check for orthology between the two plant species, we tested all candidate genes from *Arabidopsis* and the rice genome against one another for reciprocal orthology. In the end, of the 50 regions used in the Rokas et al (2005) alignment, we found 36 orthologous sequences from the mosquito, 24 from *Arabidopsis*, 29 from rice, 31 from chicken, 33 from *Plasmodium falciparum* and 39 from *Rhizopus oryzae*. We aligned each of the 50 new sets of six orthologous sequences (allowing for some missing data) to the corresponding gene sequences from the alignment from Rokas et al (2005) using the profile alignment feature of Clustal X ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/. The original amino acid alignment of Rokas et al (2005) was conservative, and we similarly pruned areas of ambiguity from the added sequences. We then concatenated the individual gene alignments.

Phylogenetic analysis and divergence time estimates.—To estimate phylogeny and branch lengths we used MrBayes 3.1.2. (Ronquist and Huelsenbeck 2003). We let the program estimate substitution parameters, including the proportion of invariant sites and the gamma shape parameter, separately for each of the 50 sequence regions and select a model of amino acid substitution. We used four independent runs of 500 000 generations each, sampling trees every 100 generations. To estimate posterior probabilities for a consensus topology, and for consensus branch lengths, we used a set of 16 000 trees, 4000 per run, sampled after 100 000 generations, at which time the log likelihoods for each run had converged and reached a plateau. We also used 500 parsimony bootstrap replicates for an additional estimate of support for clades.

We used calibration points from earlier studies, alone or in various combinations. Calibration points included the divergence of the chicken from human, 300 000 000 y ago (Peterson et al 2004); the divergence of the fruit fly from the mosquito, with a minimum age of 235 000 000 y, based on fossils from derived flies (Blagoderov et al 2002, Peterson et al 2004) and a maximum of 417 000 000 y, the age of the first primitive hexapod fossils (Gaunt and Miles 2002, Douzery et al 2004); and the divergence of eudicots within the angiosperms, with a minimum age of 144 000 000 and maximum of 206 000 000 y (Sanderson and Doyle 2001, Douzery et al 2004). We also considered the effect on calibration of several morphological and taxonomic interpretations of the 400 000 000 y old ascomycetous fossil *Paleopyrenomycites* (Taylor et al 1999).

Divergence times were estimated by the LF method with a truncated Newton algorithm in R8S version 1.70, using as an input tree the consensus topology and branch lengths from MrBayes (FIG. 2). Although the closest sister group of the plants and animals plus fungi is not yet resolved phylogenetically, an outgroup was needed to root the basal trichotomy and estimate the basal branch lengths. Therefore we chose *Plasmodium falciparum* because it does not appear within the plant, animal and fungal lineages. After Bayesian analysis *Plasmodium* was removed from the resulting tree, leaving a tree with a basal trichotomy and basal branch lengths. We tested for possible alternative equally likely estimates for node ages with the LF program's CHECKGRADIENT option and set the analysis for 10 replicates under different starting conditions, using ''set num_time_guesses''.

Penalized likelihood can compensate for substitution rate variation among lineages only when substitution rates are correlated statistically within lineages. With PL we tried to optimize ''smoothing'' to correct for autocorrelation (R8S, v1.70, user's manual, available from M. Sanderson at http://ginger.ucdavis.edu/r8s/). The optimization involved varying the value of the smoothing parameter, at increments of 0.5, from −2 to 9 on a $\log_{10}$ scale and, at each value for the smoothing parameter, having the program estimate the prediction error that resulted from pruning each terminal branch in turn. In our dataset, the prediction error did not show a minimum; instead it continued to decrease even when the value of the smoothing parameter became high. This lack of a minimum indicated that substitution rates were not correlated by lineage, but they could have been either clocklike or chaotic (R8S, v1.70, user's manual, available from M. Sanderson at http://ginger.ucdavis.edu/r8s/). Because the horizontal root-to-tip branch lengths in the tree (FIG. 2) varied, we knew that the rates had not been clock-like. This led us to conclude that rates had been chaotic. In the absence of predictable, lineage-specific rate correlations, we followed the recommendation of the R8S program documentation and analyzed the data with LF, assuming a global substitution rate.

To test whether rates might be correlated given a different multigene dataset, we applied the approach outlined above a second time, beginning with a dataset originally from Peterson et al (2004). The Peterson et al (2004) alignment of seven genes consisted of sequences from animals, with rice and *Arabidopsis* as outgroups. Using a series of blast searches we looked for corresponding sequences from eight fungal species for each gene and used phylogenetics to identify orthologues. Asterisks on the tree (FIG. 2) indicate taxa that were included. We discarded two of the original seven genes; the gene for aldolase because it was missing from the fungi and the gene for catalase because fungi had paralogues that confounded phylogenetic analysis. Each of the taxa in the modified alignment had each of the remaining five genes, which coded for S-adenosylmethionine synthetase, elongation factor 1-alpha, the ATP synthase beta chain, triosephosphate isomerase, and the beta subunit of phosphofructokinase. Our alignment of ~1700 amino acid sites for 15 taxa is available on request. After Bayesian analysis we used R8S and PL to examine autocorrelation levels. Again we were unable to optimize smoothing values, and again the Bayesian tree showed great variation in horizontal branch lengths (results not shown).

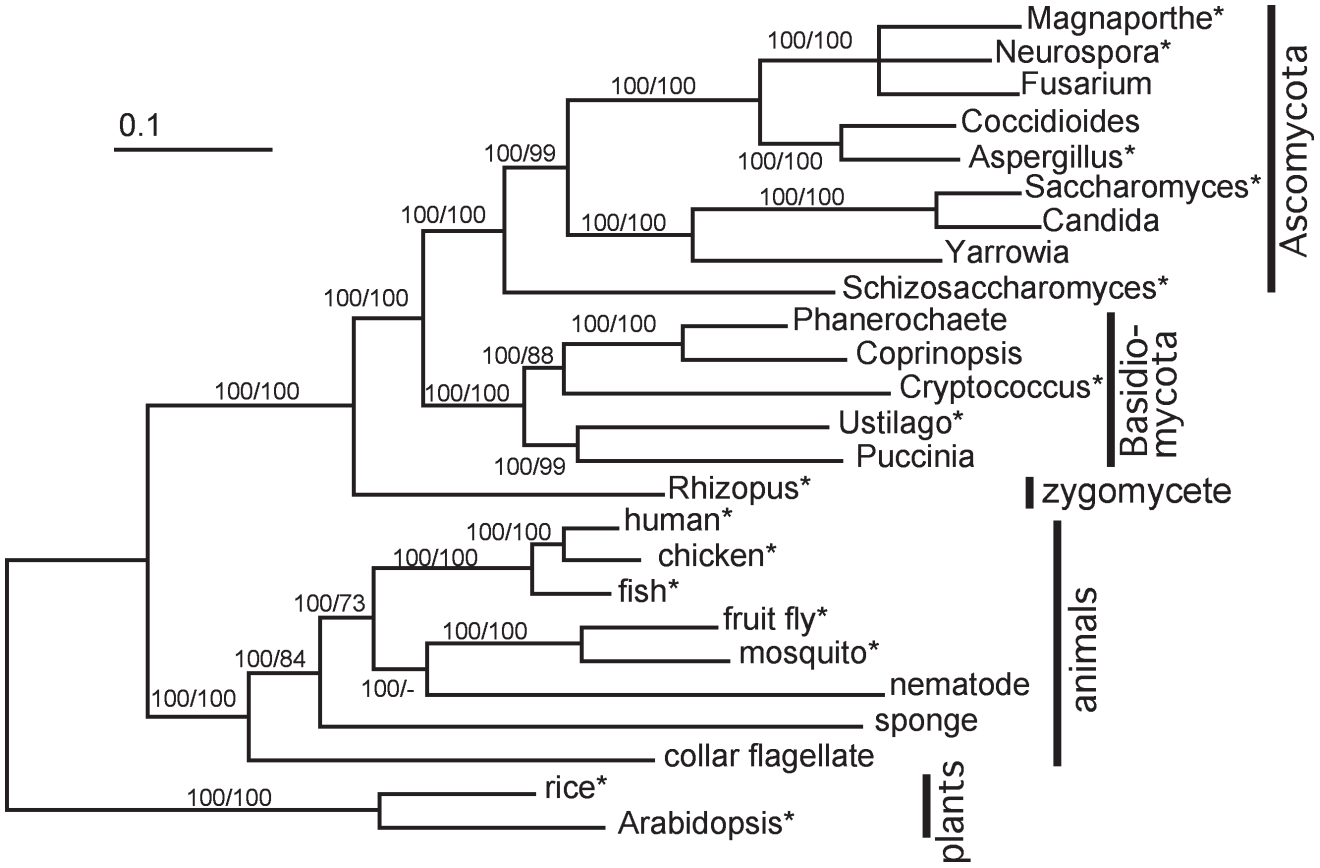


FIG. 2. Tree showing variation in distance, from root to tip, among 25 fungi, animals and plants. This Bayesian consensus phylogram is based on amino acid sequences of 50 concatenated gene regions and 12 089 aligned sites. The first number at the branch is the Bayesian posterior probability for the node and is based on 16 000 trees sampled from a total of 2 000 000 Bayesian generations. The second number is the parsimony bootstrap proportion. Asterisks indicate taxa also used in the second of our two analyses.

RESULTS

As reviewed above, the most important factor in fitting a phylogeny to the geologic timescale is the fossil information used to set minimum and maximum dates for reference divergences. The Bayesian tree (FIG. 2) shows that rates of substitution vary among lineages. The total length of the branches from the tree root to the vertebrates and rice is short, indicating below average substitution rates, while the length to the ascomycetes is generally long, indicating higher than average substitution rates. Following the model from Peterson et al (2004) we tested the consequences of applying single animal or plant calibrations to fungal divergence times using the phylogeny for 25 taxa and 50 genes (FIG. 2) and R8S (LF method and truncated Newton algorithm). As can be seen (TABLE I), using the calibration of 300 000 000 y for the divergence of birds from mammals, the estimate for the divergence of fungi from animals was 2 635 000 000 y ago, which would imply that the oxygen-consuming animals and fungi originated when the earth still had a reducing atmosphere. Using the calibration of 235 000 000 y for the age of the divergence of the fly from the mosquito reduced the age estimate of the fungus/animal split to 944 000 000 y ago. Using the split of the eudicots from monocots for calibration, the fungus/animal split was reduced still further, to 579 000 000 y ago, but the age of the bird/mammal split was estimated at 66 000 000 y ago, far too recent given the good fossil evidence that this divergence took place around 300 000 000 y ago.

We next evaluated the effect on the Fungal Tree of Life of varying the minimum divergence dates referenced to the fossil fungus, *Paleopyrenomycites* (FIG. 1). This fossil fungal fruiting body is considered to be a sordariomycete and is found in rhizomes of *Asteroxylon mackei* in the Rhynie chert, dated at 400 000 000 y ago (Taylor et al 1999, Taylor et al 2005). The key event that we evaluated is the correspondence of the appearance of the Glomeromycota and the appearance of fossil land plants. We used the divergence of Ascomycota from Basidiomy-

TABLE I. Improbable age estimates (in boldface) resulted when molecular clocks used rates estimated from calibration points from one phylum and then applied the rates to other groups

	Estimates of the age, in millions of years, of clade divergences*											
Calibration point choice; constraint age in millions of years	Plant vs opistho -kont	Fungi vs animals	Eudicots vs monocots	Fish vs mam-mal	Chicken vs human	Mos-quito vs fruit fly	Fungi crown	Asco- vs Basidio-mycota	Basidio-mycota crown	Asco-mycota crown	Pezizo-mycotina crown	Sordario-mycetes crown
Chicken vs. human = 300	**3119**	**2635**	**937**	450	300	**656**	1979	1808	1361	1598	858	486
Mosquito vs. fly = 235	1118	944	**336**	**161**	**108**	235	709	648	488	573	308	174
Eudicots vs. monocots = 206	686	579	206	**91**	**66**	144	435	398	299	351	189	107
Sordariomycetes crown = 400	**2568**	**2170**	**771**	**370**	**247**	540	1630	1489	1120	1316	707	400
Pezizomycotina crown = 400	1454	1228	437	**210**	**140**	306	923	843	634	745	400	226
Ascomycota crown = 400	780	659	**234**	**113**	**75**	**164**	495	452	340	400	215	122
mosquito vs. fruit fly, min 235 max 417; eudicots vs monocots min 144 max 300; Pezizomycotina min 400	1212	1036	206	**178**	**119**	259	792	729	544	652	400	211

*Calibratated ages are in double boxes. Where calibrations were provided as minima and maxima, the ages in the boxes were estimates limited by a possible range. ''Crown'' is used to specify the first available internal node within a group (e.g. the age of crown fungi refers to the age of the first internal node in the fungal clade in our study, which was the split of *Rhizopus oryzae* from the ancestor to the Ascomycota and Basidiomycota).

cota as a surrogate for the divergence within Glomeromycota. Although *Paleopyrenomycites* is considered to be a member of the Sordariomycetes, other taxonomic interpretations can be evaluated by using it to provide a minimum date for the earliest divergence in Sordariomycetes, the earliest divergence in Pezizomycotina, or the earliest divergence in Ascomycota.

When *Paleopyrenomycites* is considered to be in the Sordariomycetes, the minimum age for the divergence of Ascomycota and Basidiomycota is 1 489 000 000 y ago, when the fossil is considered

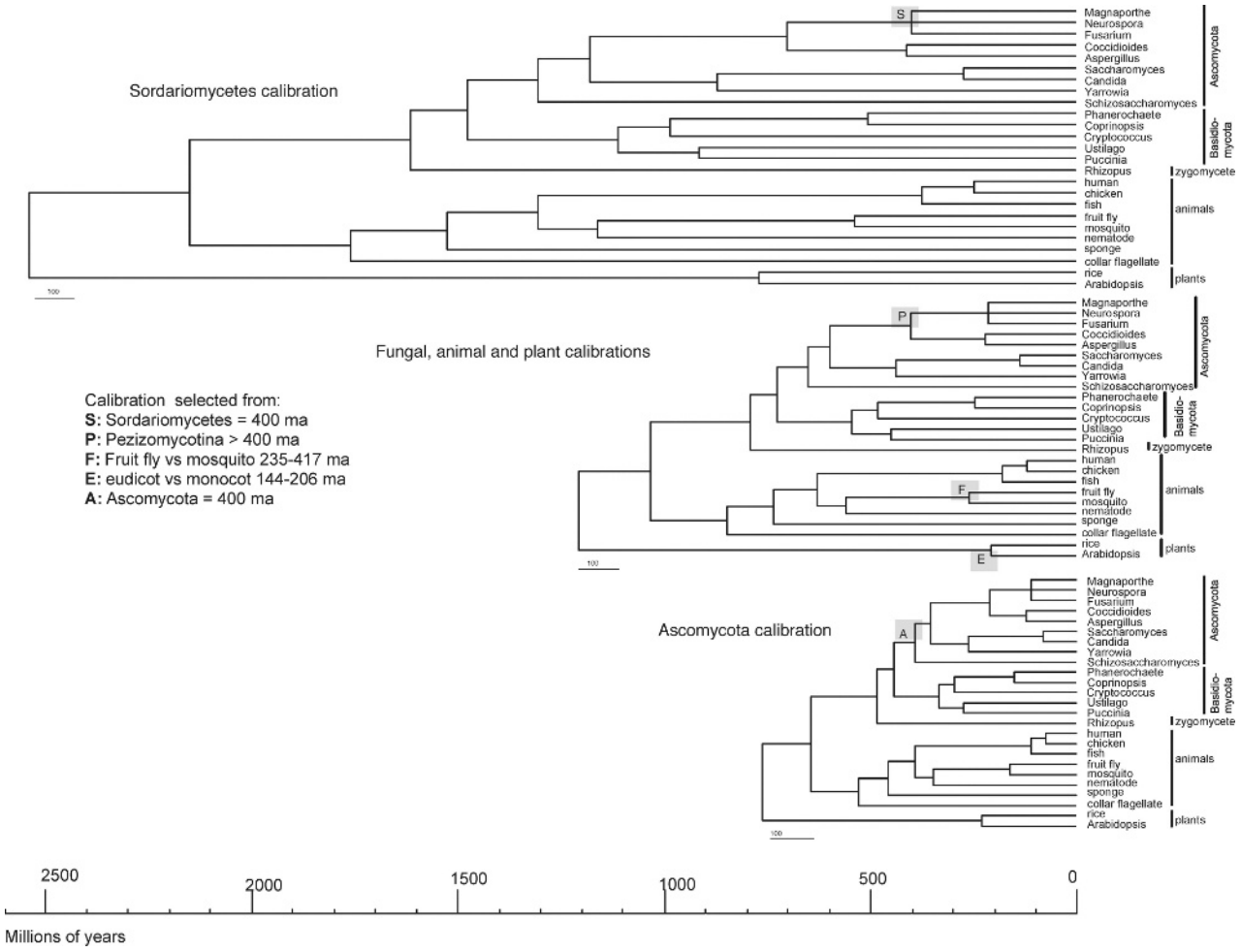


FIG. 3. Different calibration points change estimated divergence dates. Ascomycota split from Basidiomycota after the origin of two phyla not shown, the Chytridiomycota, and Glomeromycota. In the top diagram we assumed that the 400 000 000 y old fossil *Paleopyrenomycites devonicus* represents Sordariomycetes as indicated by the ''S''. This pushes the minimum age for the origin of the stem lineages of all five fungal phyla, including the Glomeromycota, to 1 489 000 000 y ago, more than three times the age of the first fossil evidence of land plants. In the middle diagram we assumed that *P. devonicus* represents Pezizomycotina but not necessarily Sordariomycetes. Applying calibrations from plant and animal fossils in addition to assuming that *P. devonicus* represents Pezizomycotina gave an estimate of 792 000 000 y ago for the origin of the fungal phyla, still almost twice the age of the first fossil evidence for vascular plants. In the bottom diagram, assuming that *P. devonicus* provided a minimum age for the Ascomycota, indicated by the ''A'', at 400 000 000, and resulted in the estimate that fungal phyla had been established by 452 000 000 y ago, roughly the age of the first land plants. In this scenario, however, the ages for the divergences among vertebrates are far too recent given fossil data. This discrepancy might result from using fossils from one kingdom (Fungi) to calibrate events in the same kingdom and another kingdom (Animalia), as discussed in the text.

to be in the Pezizomycotina, the minimum age for the Ascomycota-Basidiomycota divergence is 843 000 000 y ago, and when it is considered to be in the Ascomycota, the minimum age for the Ascomycota-Basidiomycota divergence is estimated to be 452 000 000 y ago (FIG. 3, TABLE I). As noted above the estimate for the radiation of crown land plants based on PL and 27 plastid genes from 10 taxa (Sanderson 2003) ranged from 435 000 000 to 480 000 000 y ago. Only the analysis that treats the fossil as a member of the Ascomycota estimates a divergence time for Glomeromycota that is consistent with the radiation of land plants. The other two interpretations of the fossil, that it is a member of Sordariomycetes or the Pezizomycotina, would have the Glomeromycota diverging far in advance of the radiation of crown land plants. In fact, the other two interpretations would have the major divisions of Ascomycota and Basidiomycota, the modern members of which live by parasitism or decay of land plants, becoming established before the appearance of land plants. On the other hand, when the fungal

and plant events are in the best alignment, the divergence of birds and mammals is inconveniently recent (TABLE I) at less than 100 000 000 y ago.

DISCUSSION

Substitution rates clearly vary among lineages and we, like Peterson et al (2004), found no justification for using rates calibrated by fossils from one kingdom to estimate dates in another. We had hoped that, given enough genes, rate variation would average out to be comparable for all lineages or else that variation would follow a pattern that could be accommodated with an autocorrelation parameter. Instead, even with multigene phylogenies from two different datasets, substitution rates not only varied but the variation was not correlated by lineage. Our taxon sampling was limited by available genomes and by computational limits. If more diverse taxa were sampled from the fungal phyla, rate autocorrelation might be evident. However ages still might be distorted by undetected rate change when the nodes of interest are the old ones and the rate changes have accumulated on the long branches of the stem lineages. To improve on earlier molecular clock estimates for divergence times, it appears as if fossil calibration points will be needed within the lineages of taxa that have relatively constant substitution rates.

Paleopyrenomycites is beautifully preserved and has been carefully described as a perithecial ascomycete with an ostiole, paraphyses and elongate asci that contain as many as 32 ascospores (Taylor et al 1999, Taylor et al 2005). However can we rule out its belonging to the Taphrinomycotina or an even earlier diverging, but extinct, group of Ascomycota? Although most described Taphrinomycotina lack ascomata, species of the Taphrinomycotina genus *Neolecta* do have open, apothecial ascomata. Although no partially closed, flask-shaped ascoma are known among extant Taphrinomycotina, this type of ascoma is found in several classes of Pezizomycotina: Sordariomycetes, Dothideomycetes and Chaetothyriomycetes. It has evolved more than once and also could have evolved in Taphrinomycotina or in fungi extant before the divergence of Taphrinomycotina. Therefore we raise the possibility that this fossil is not a member of the Sordariomycetes but might represent a lineage that arose earlier in the history of Ascomycota than did Sordariomycetes. Our claim obviously cannot easily be tested by currently available fossils. To evaluate our proposal, the most informative fossils would be those that have a continuous record with a discernable shift in phenotype that correlates with the emergence of a new taxon (e.g. a fungal equivalent to the pollen record and the emergence of tricolpate pollen).

Where might mycologists look for a similar record, one where a shift in spore form could be recognized easily by microscopic analysis? Among thick-walled spores that could be expected to preserve well, perhaps the shift from mitospores without internal septations to longitudinally septate spores could provide a minimum age for Dothideomycetes? Or, if thin-walled spores preserve well, the shift from symmetrical spores to asymmetrical spores with hilar appendices could provide a minimum age for Basidiomycota? Are the answers to these questions to be found in microscope slides of fossil pollen already prepared by paleobotanists? Pursuing this question will require collaborations of mycologists and paleobotanists; that much is certain.

ACKNOWLEDGMENTS

We thank K. Peterson and A. Rokas for datasets and A. Rokas for advice on applying reciprocal orthology. M. Sanderson kindly checked our cross validation results with R8S. D. Carmean installed and tested software. B. Thomas assisted with running MrBayes on the UCB, CNR computing cluster. T.N. Taylor kindly provided photographic prints used in Figure 1. JWT is supported by NSF DEB-0316710 and DEB 0516511. MLB is supported by the National Science and Engineering Research Council of Canada. We also acknowledge support from NSF 0090301, Research Coordination Network: a phylogeny for kingdom Fungi to M. Blackwell, J.W. Spatafora and J.W. Taylor.

LITERATURE CITED

Arabidopsis Genome I. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815.

Bell CD, Soltis DE, Soltis PS. 2005. The age of the angiosperms: a molecular timescale without a clock. Evolution 59:1245–1258.

Berbee ML, Taylor JW. 1993. Dating the evolutionary radiations of the true fungi. Can J Bot 71:1114–1127.

———. 2001. Fungal molecular evolution: gene trees and geologic time. In: McLaughlin D, McLaughlin E, Lemke P, eds. The Mycota: systematics and evolution. Berlin: Springer-Verlag. p 229–245.

Blagoderov VA, Lukashevich ED, Mostovski MB. 2002. Order Diptera Linné, 1758. The true flies (=Muscida Laicharting, 1781). In: Rasnitsyn AP, Quicke DLJ, eds. History of Insects. Dordrecht, Boston, London: Kluwer Academic. p 227–240.

Doolittle RF, Feng D-F, Tsang S, Cho G, Little E. 1996. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271: 470–477.

Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H. 2004. The timing of eukaryotic evolution: Does a re-

laxed molecular clock reconcile proteins and fossils? Proc Nat Acad Sci USA 101:15386–15391.

Felsenstein J. 2005. PHYLIP (phylogeny inference package). Distributed by the author. Department of Genome Science, University of Washington, Seattle, Washington.

Gardner M, Hall N, Fung E, White O, Berriman M, et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419:498–511.

Gaunt MW, Miles MA. 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol Biol Evol 19: 748–761.

Graur D, Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trend Genet 20:80–86.

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB. 2001. Molecular evidence for the early colonization of land by fungi and plants. Science 293: 1129–1133.

Hedges SB, Kumar S. 2004. Precision of molecular time estimates. Trend Genet 20:242–247.

Hillier L, Miller W, Birney E, Warren W, Hardison R, et al. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

Holt R, Subramanian G, Halpern A, Sutton G, Charlab R, et al. 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298:129–149.

Magallon SA. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. Int J Plant Sci 165:S7–S21.

Padovan ACB, Sanson GFO, Brunstein A, Briones MRS. 2005. Fungi evolution revisited: application of the penalized likelihood method to a Bayesian fungal phylogeny provides a new perspective on phylogenetic relationships and divergence dates of ascomycota groups. J Mol Evol 60:726–735.

Peterson KJ, Lyons JB, Nowak KS, Takacs CM, Wargo MJ, McPeek MA. 2004. Estimating metazoan divergence times with a molecular clock. Proc Nat Acad Sci USA 101:6536–6541.

Pirozynski KA, Dalpe Y. 1989. Geological history of the Glomaceae with particular reference to mycorrhizal symbiosis. Symbiosis 7:1–36.

Redecker D, Kodner R, Graham LE. 2000. Glomalean fungi from the Ordovician. Science 289:1920–1921.

Reisz RR, Muller J. 2004. Molecular timescales and the fossil record: a paleontological perspective. Trend Genet 20: 237–241.

Remy W, Taylor TN, Hass H, Kerp H. 1994. Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Nat Acad Sci USA 91:11841–11843.

Rodriguez-Trelles F, Tarrio R, Ayala FJ. 2002. A methodological bias toward overestimation of molecular evolutionary time scales. Proc Nat Acad Sci USA 99:8112–8115.

Rokas A, Kruger D, Carroll SB. 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310:1933–1938.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rutschmann F. 2006. Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. Divers Distribut 12:35–48.

Sanderson MJ. 2003. Molecular data from 27 proteins do not support a Precambrian origin of land plants. Am J Bot 90:954–956.

———, Doyle JA. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from rbcL and 18S rDNA data. Am J Bot 88:1499–1516.

———, Thorne JL, Wikstrom N, Bremer K. 2004. Molecular evidence on plant divergence times. Am J Bot 91:1656–1665.

Sarich VM, Wilson AC. 1967. Rates of albumin evolution in primates. Proc Nat Acad Sci USA 58:142–148.

Simon L, Bousquet J, Levesque RC, Lalonde M. 1993. Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. Nature 363: 67–69.

Taylor TN, Hass H, Kerp H, Krings M, Hanlin RT. 2005. Perithecial ascomycetes from the 400 million year old Rhynie chert: an example of ancestral polymorphism (Vol. 96, p 1403, 2004). Mycologia 97:269–285.

———, Hass T, ———. 1999. The oldest fossil ascomycetes. Nature 399:648–648.

Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Systemat Biol 51:689–702.

Wang DYC, Kumar S, Hedges SB. 1999. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc Royal Soc London, Series B-Biol Sci 266:163–171.

Welch JJ, Bromham L. 2005. Molecular dating when rates vary. Trend Ecol Evol 20:320–327.

Wu CI, Li WH. 1985. Evidence for higher rates of nucleotide substitutions in rodents than in man. Proc Nat Acad Sci USA 82:1741–1745.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Cabios 13:555–556.

Zuckerkandl E, Pauling L. 1965. Molecules as documents of evolutionary history. J Theoret Biol 8:357–366.