



Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits

FRANÇOIS LUTZONI, FRANK KAUFF, CYMON J. COX, DAVID MC LAUGHLIN, GAIL CELIO, BRYN DENTINGER, MAHABEEN PADAMSEE, DAVID HIBBETT, TIMOTHY Y. JAMES, ELISABETH BALOCHI, MARTIN GRUFE, VALÉRIE REEB, VALÉRIE HOFLSTETTER, CONRAD SCHOCH, ANGELA ZARNOLD, JOLANTA MIADLIKOWSKA, JOSEPH SPATAFORA, DESIREE JOHNSON, SARA HAMBLETON, MICHAEL CROCKETT, ROBERT SHOEMAKER, GI-HO SUNG, ROBERT LÜCKING, THORSTEN LUMBSCH, KERRY O'DONNELL, MANFRED BINDER, PAUL DIEDERICH, DAMIEN ERTZ, CÉCILE GUEIDAN, BENJAMIN HALL, KAREN HANSEN, RICHARD C. HARRIS, KENTARO HOSAKA, YOUNG-WOON LIM, YAJUAN LIU, BRANDON MATHENY, HIROMI NISHIDA, DON PFISTER, JACK ROGERS, AMY ROSSMAN, IMKE SCHMITT, HARRIE SIPMAN, JEFFREY STONE, JUNTA SUGIYAMA, REBECCA YAHR, AND RYTAS VILGALYS.

**Two-locus Bayesian MCMCMC tree
nucSSU+nuLSU**
558 species, 430 genera, 68 orders, 5 phyla

Bayesian posterior probability
NJ bootstrap proportion
0.1 substitutions/site

Acknowledgment

We thank Molly McMullen and Connie Robertson (Duke University Herbarium) for curating lichen specimens used for this study; Josée Béliste for graphic design of phylogenetic trees, as well as Snehal Sarvate and Lindsay Higgins for assistance in preparing appendices. We are very thankful to Dr. Trevor Bill Rankin, and John Pormann for providing access to the Duke C.S.I.M. computer cluster; the Supercomputer (Information Technology Division at North Carolina State University) and Sudhakar Pamidighantam (National Center for Supercomputing Applications) for installing software to enable us to conduct many of the analyses essential for this study. Thanks to Joyce Longcore, and John Taylor for providing fungal strains. This publication resulted in part from the Assembling the Fungal Tree of Life (AFTOL) project, which is supported by NSF Assembling the Tree of Life (AFTOL) awards DEB-0228725 to JS, DEB-0228657 to DH, DEB-0228671 to DJM, and DEB-0228668 to FL and RV, as well as, DEB-9615542 and DEB-0133891 to FL, DEB-0105194 to FL and VR, DEB-9903835 to DH, DEB-0128925 to DH and MB, DEB-0129212 to JS, DEB-0306578 and DEB-9318232 to DJM, DEB-9813304 to JR, FWF P16410 and P19052 to MG. We also gratefully acknowledge support of the NSF Research Coordination Networks in Biological Sciences: A Phylogeny for Kingdom Fungi (NSF 0090301).

Introduction and objectives
Despite the numerous technological advancements available to fungal systematists, progress in understanding the deepest nodes in the fungal tree of life will be limited without a new approach to conducting large-scale multilocus phylogenetic studies and phenotype-based comparative studies on Fungi. This novel approach will require concerted data acquisition by focusing sequencing efforts on specific loci and fungal taxa, by conducting phenotypic studies on specific fungal traits, by improving interaction among fungal systematists, and by the automation of data acquisition and analysis coupled with data bases accessible through the World Wide Web. These goals form the framework of the Assembling the Fungal Tree of Life Project (AFTOL), which seeks to infer the phylogenetic relationships among 1500 species representing all fungal phyla based on eight loci (~10 kb). Here, we report phylogenetic studies for the maximal number of species across all known fungal phyla for which DNA sequence data from two, three, and four loci are available. The resulting phylogenetic trees are based on sequences available in GenBank and unpublished sequences generated by various laboratories or by the AFTOL project. This study provides a phylogenetic synthesis for the Fungi and a framework for future phylogenetic studies on fungi.

Other presentations at this meeting associated with this poster:

1) Valerie Hofstetter, J. Miadlikowska, C. Gueidan, V. Reeb, A. E. Arnold, F. Kauff, C. Cox, R. Yahr & F. Lutzoni. What do protein-coding genes (ATP6, EF 1-alpha and RNA polymerase II) bring to molecular systematics of lichens? Tuesday August 17, Poster Session 1: Systematics and Evolution, 18:00-19:00.

2) Frank Kauff, Cymon Cox & François Lutzoni. Bioinformatics of AFTOL (Assembling the Fungal Tree of Life). Tuesday August 17, Poster Session 1: Systematics and Evolution, 18:00-19:00.

3) Cymon Cox & Frank Kauff. A data management framework for the AFTOL project. Saturday, August 21, Discussion session 3: Phylogenetic methods 15:00-17:50.

4) François Lutzoni et al. Where are we in assembling the fungal tree of life, classifying the Fungi and understanding the evolution of their subcellular traits? Saturday, August 21, Discussion session 3: Phylogenetic methods 15:00-17:50.

Interpretation of support values

	Posterior Probabilities	
At least one BP	≥ 95%	< 95%
≥ 70%	High	Medium
70% > BP ≥ 50%	Medium	Low
< 50%	Low	No support

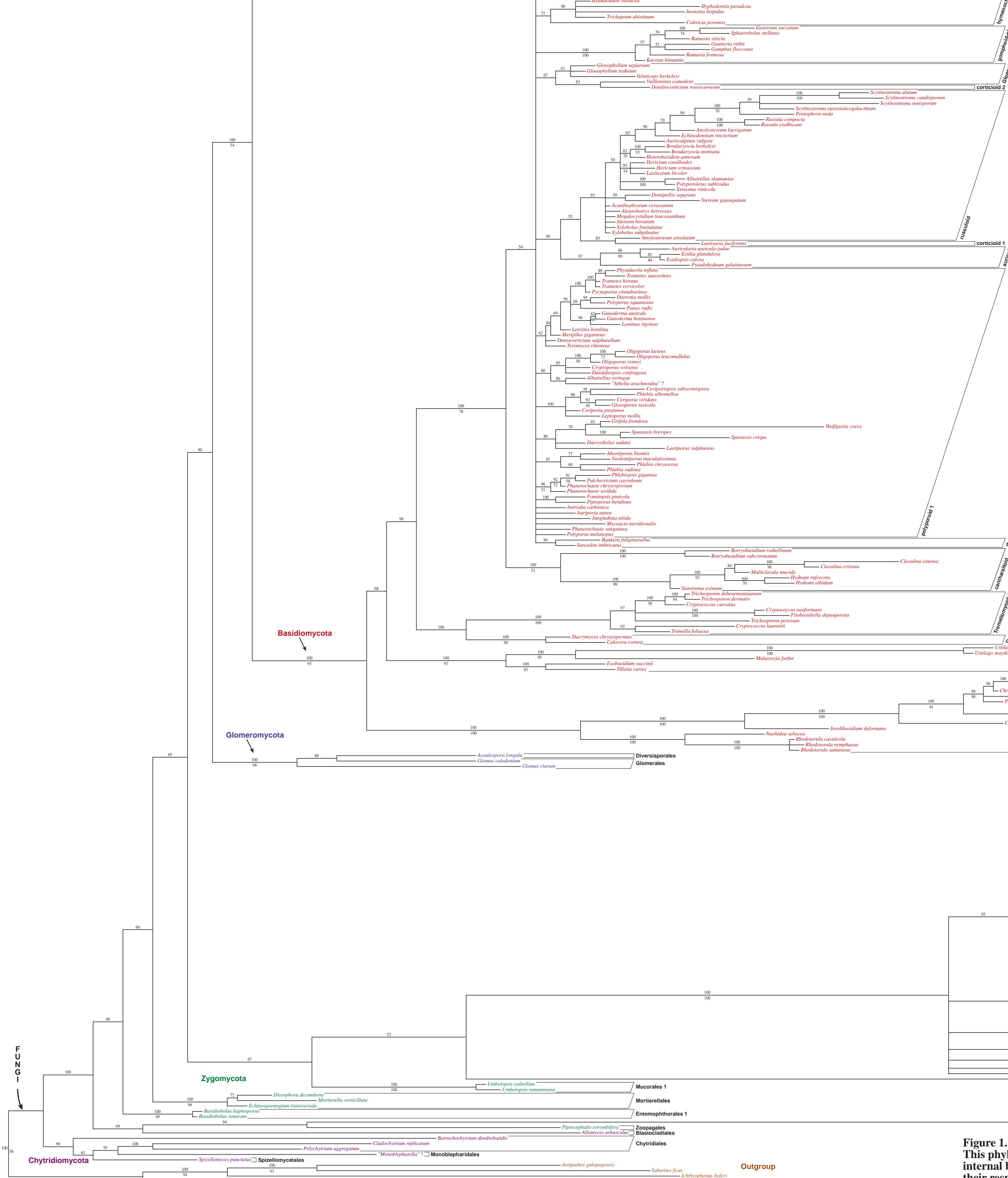


Figure 1. Two-locus (nucSSU + nuLSU) Bayesian MCMCMC fungal tree of life depicting phylogenetic relationships among 558 taxa in 430 genera, 68 orders and five phyla. This phylogeny resulted from a 50% majority rule consensus of 24,000 trees sampled with Bayesian MCMCMC. The resulting posterior probabilities (PP) are shown above internal branches; NJ bootstrap proportions (NJPB) with ML distances (1000 bootstrap replicates) are shown below internal branches. Species names are colored according to their respective phylum.

Materials and Methods

Taxon sampling — For the nucSSU + nuLSU (~2.4kb, Fig. 1) data set, a total of 13 467 GenBank sequences were considered, of which 1010 unique taxa had no sequences available. Sequences that were selected incorrectly due to inconsistencies in the GenBank record ("Definition Line" were discarded, as were sequences whose length was < 600 base pairs or whose overlap with other taxa was < 600 base pairs. Unpublished sequences available directly from the AFTOL project laboratories associated with them prior to submission with those available in GenBank were included in the reference data set. For the nucSSU + nuLSU + RPB2 (~4.6kb, Fig. 2) data set, RPB2 sequences for 19 taxa were obtained from the AFTOL project laboratories and taxon names were included in the reference data set. For the nucSSU + nuLSU + RPB2 (~5.4kb, Fig. 3) data set, miSSU sequences for 103 taxa were obtained from the AFTOL project laboratories and taxon names were included in the reference data set. For the nucSSU + nuLSU + RPB2 (~6.0kb, Fig. 4) data set, miSSU sequences for 103 taxa were obtained from the AFTOL project laboratories and taxon names were included in the reference data set. One hundred forty-two taxa were retrieved from GenBank, such that the nucSSU + nuLSU + RPB2 data set consisted of 161 taxa. For the nucSSU + nuLSU + miSSU + RPB2 (~5.4kb, Fig. 5) data set, miSSU sequences for 103 taxa were obtained from the AFTOL project laboratories and taxon names were included in the reference data set. For the nucSSU + nuLSU + miSSU + RPB2 (~6.0kb, Fig. 6) data set, miSSU sequences for 103 taxa were obtained from the AFTOL project laboratories and taxon names were included in the reference data set. One hundred forty-eight taxa were retrieved, such that the final nucSSU+nuLSU+miSSU data set consisted of 253 unique taxa. Taxa present in these three preceding data sets were combined, resulting in 107 unique taxa forming this four-taxon clade.

Phylogenetic analyses — By far the most common MCMCMC analyses were conducted with MrBayes Metropolis coupled Markov Chain Monte Carlo (B-MCMCMC) analyses were conducted using four chains, and a gamma distribution, if applicable, was approximated with four categories. In addition to posterior probabilities (PP), phylogenetic confidence was estimated with weighted maximum parsimony (MPBP), neighbor joining (bootstrap proportion), and by analyzing bootstrapped data sets with NJ distance (1000 bootstrap replicates) (Swofford 2002). Step matrices for weighted parsimony analyses were generated using stepmatrix.py (written by F. Kauff and available upon request from FK or FL). Uninformative characters were excluded from all bootstrapped data sets analyzed with MP. Ratchet search strategies (PAUPRat; Nixon, 1999; Sikes and Lewis, 2001, <http://www.calgary.ca/~sikes/software/2.htm>) were implemented in PAUP*. Bootstrapped data sets subjected to B-MCMCMC were converted to NJ distance matrices using PAUP* (Swofford 2002). For each partition, the NJ distance matrix was calculated using the NJ algorithm, and the NJ bootstrap proportion was used to determine the appropriate model (nucleotide substitution and rate heterogeneity parameters). For each NJ analysis, parameter values were fixed to the optimal values calculated for the optimal model. For the RPB2 data set, each codon position was subjected to a separate model in the B-MCMCMC analysis. Following the recommendation in Reeb et al. (2004), we used NJPBP (500 replicates) to detect topological conflicts among data partitions. A conflict was assumed to be significant if a different relationship (one monophyletic, the other non-monophyletic) for the same taxon was found in at least 50% of the NJPBP replicates. The program compair.py (written by F. Kauff and available upon request from FK or FL) was used to detect such topological incongruences. Taxa causing conflicts were removed, and the test was reimplemented until no conflicts were detected. Each locus in the combined data sets was subjected to this incongruence test for all possible pairwise comparisons prior to inclusion.

Relationships among fungal phyla (Fig. 1) — The Fungi were resolved in a clade with a 100% Bayesian posterior probability (PP) with respect to the animalia protista. Both the Ascomycota and Basidiomycota formed clades supported by a 100% NJPBP and 67% NJBP, respectively, in the nucSSU + nuLSU analysis. Furthermore, a sister relationship between the Basidiomycota and Ascomycota (the "Dikaryomycota") received medium support (PP = 100% and NJBP = 54%). The Glomeromycota formed a clade (PP = 100% and NJBP = 98%) sister to the "Dikaryomycota". This clade has often been recovered in nuclear rDNA phylogenetic analyses (Sugiyama, 1998; Schüller et al., 2001; Tehler et al., 2003). It has recently been given the informal name "Symbiomycota" because most of its members form symbioses (Tehler et al., 2003), but statistical support for this clade in this (PP = 90%) and other studies, has never been achieved.

Relationships among and within Chytridiomycota, Zygomycota, and Glomeromycota (Fig. 1) — The nucSSU + nuLSU phylogeny resolves a chytrid clade sister to all fungi and a paraphyletic assemblage of zygomycete lineages + Allomycetidae, which form a grade leading to more than a single taxon, only one of these is monophyletic (Moniliellales). The nucSSU + nuLSU analyses (Fig. 2) provide strong support for the monophyly of the Moniliellales, which includes the earliest branching fungi in the tree, and provide new insights into few critical branching events. The divergence of the two Basidiobolus spp. within the Zygomycota + Allomycetidae group is more consistent with the ecological and morphological traits of these fungi than the placement of the fungus within the chytrid lineage. Moniliophorales spp. (representing the Moniliophorales) groups within the basal-most fungal clade of chytrid fungi (though this position is unsupported), that also includes the orders Chytridiales and Spizellomycetales. In contrast, the Blastocladiales, represented by Allomyces arbusculus, group with the remainder of the Fungi rather than with the other Chytridiomycetes.

Relationships within Ascomycota and Basidiomycota (Figs. 1-3) — As a consequence of the lack of complementarity among single-locus data sets, the three- and four-locus analyses included only members of the Ascomycota and Basidiomycota. The four-locus analysis resolved multiple deep relationships within the Ascomycota and Basidiomycota that were not revealed previously or that received only weak support values in previous studies.

Publication

The authors have already discussed phylogenetic structure on supradordinal classifications, the results and reanalysis of subcellular data, a synthesis of current knowledge regarding the evolution of septal features of fungal hyphae, and a preliminary reassessment of ascospore evolution will be published in the October (2004) issue of the American Journal of Botany.

Conclusion

The authors have already discussed phylogenetic structure on supradordinal classifications, the results and reanalysis of subcellular data, a synthesis of current knowledge regarding the evolution of septal features of fungal hyphae, and a preliminary reassessment of ascospore evolution will be published in the October (2004) issue of the American Journal of Botany.

**Three-locus Bayesian MCMCMC tree
nucSSU+nuLSU+RPB2**
157 species, 134 genera, 34 orders, 2 phyla

Bayesian posterior probability
NJ bootstrap proportion / MP bootstrap proportion
0.1 substitutions/site

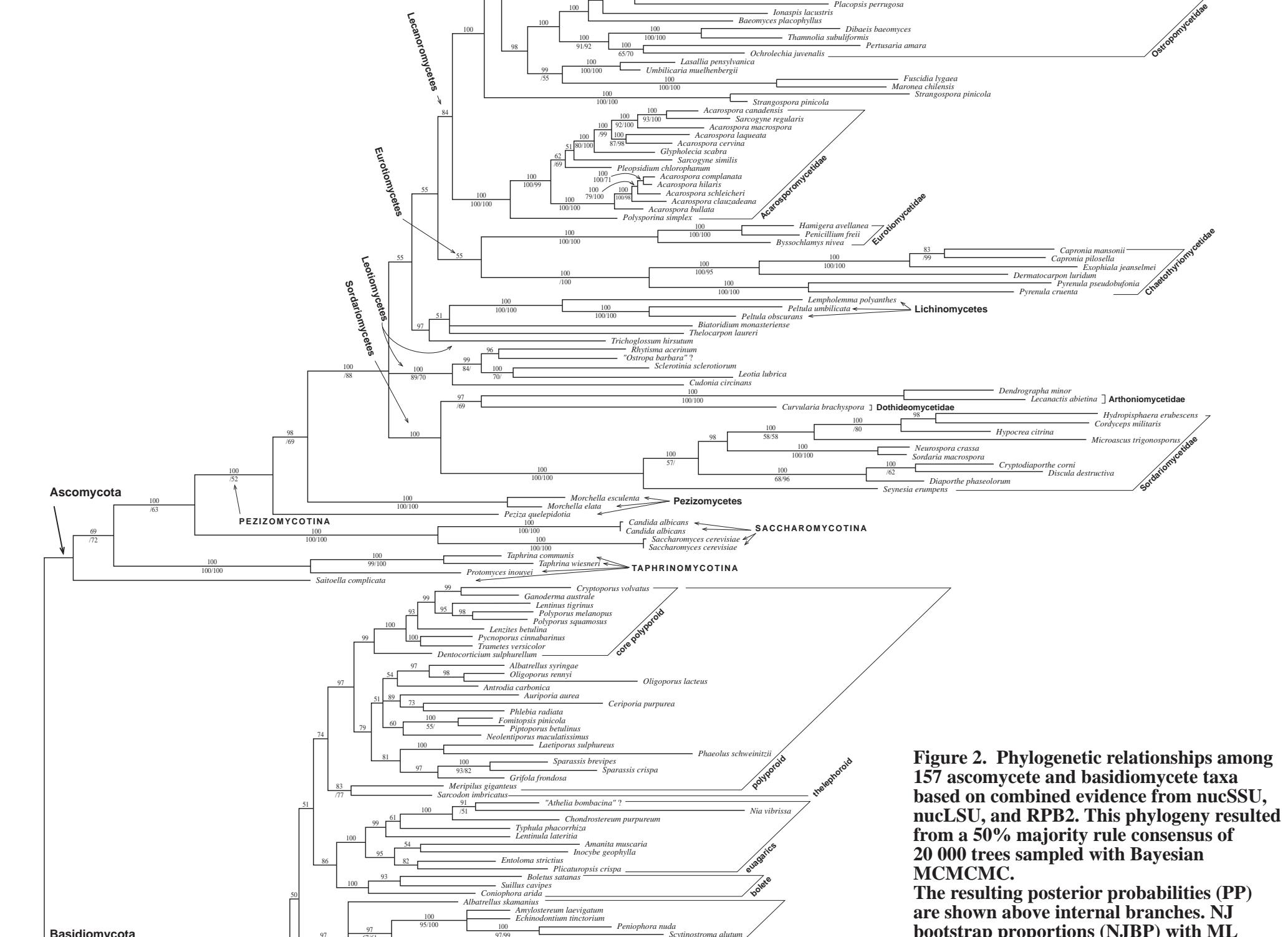


Figure 2. Phylogenetic relationships among 157 ascomycete and basidiomycete taxa based on combined evidence from nucSSU, nuLSU, and RPB2. This phylogeny resulted from a 50% majority rule consensus of 20,000 trees sampled with Bayesian MCMCMC.

The resulting posterior probabilities (PP) are shown above internal branches; NJ bootstrap proportions (NJPB) with ML distance (1000 bootstrap replicates) are shown below internal branches before the slash sign. Weighted MP bootstrap proportions (MPBP) are shown below internal branches after the slash sign.

**Four-locus Bayesian MCMCMC tree
nucSSU+nuLSU+mitSSU+RPB2**
103 species, 94 genera, 30 orders, 2 phyla

Bayesian posterior probability/Bayesian bootstrap proportion
NJ bootstrap proportion/MP bootstrap proportion
0.1 substitutions/site

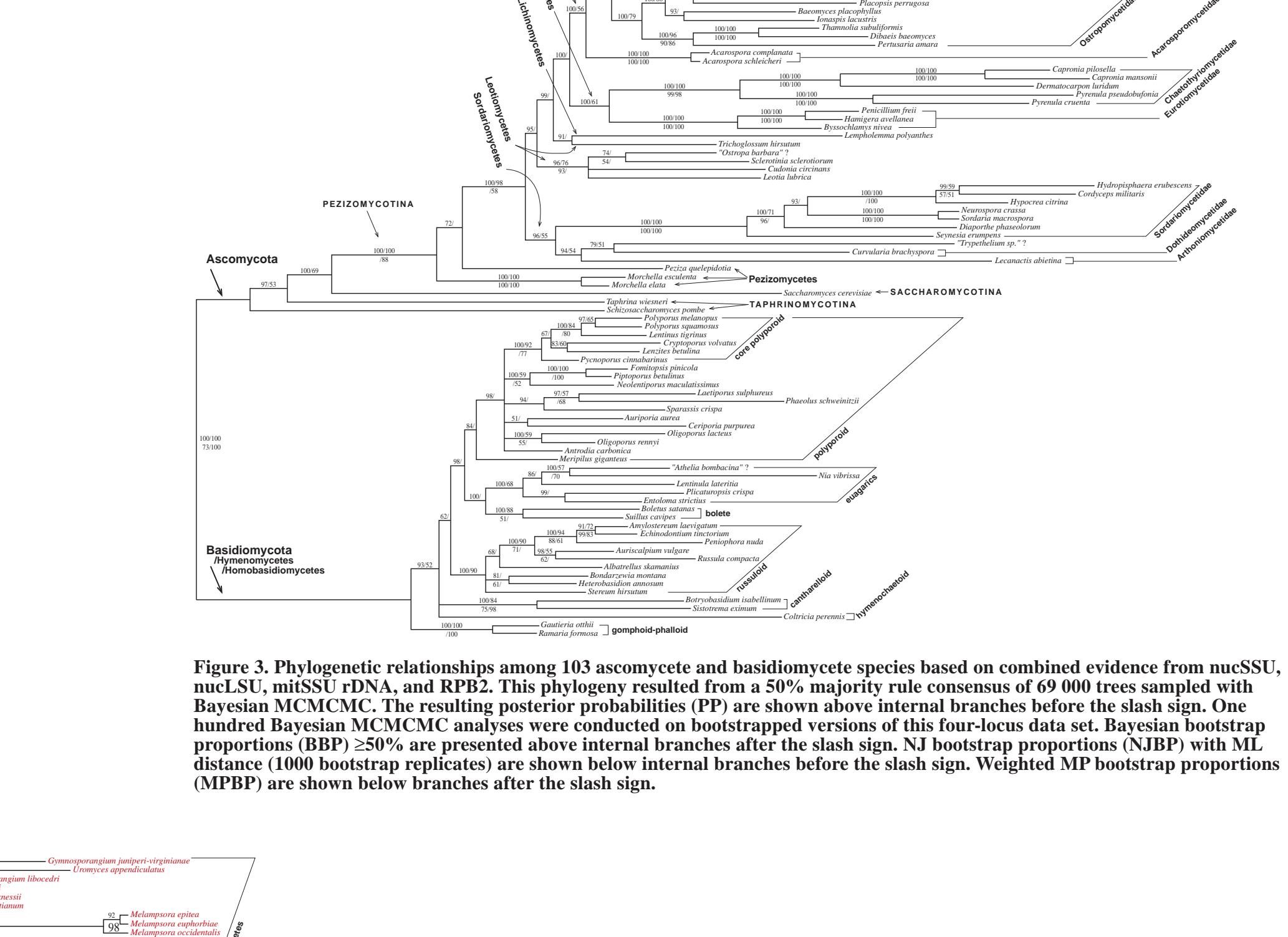


Figure 3. Phylogenetic relationships among 103 ascomycete and basidiomycete taxa based on combined evidence from nucSSU, nuLSU, mitSSU, and RPB2. This phylogeny resulted from a 50% majority rule consensus of 69,000 trees sampled with Bayesian MCMCMC.

One hundred Bayesian MCMCMC analyses were conducted on bootstrapped versions of this four-locus data set. Bayesian bootstrap proportions (BBP) ≥ 50% are presented above internal branches after the slash sign. NJ bootstrap proportions (NJPB) with ML distance (1000 bootstrap replicates) are shown below internal branches before the slash sign. Weighted MP bootstrap proportions (MPBP) are shown below internal branches after the slash sign.

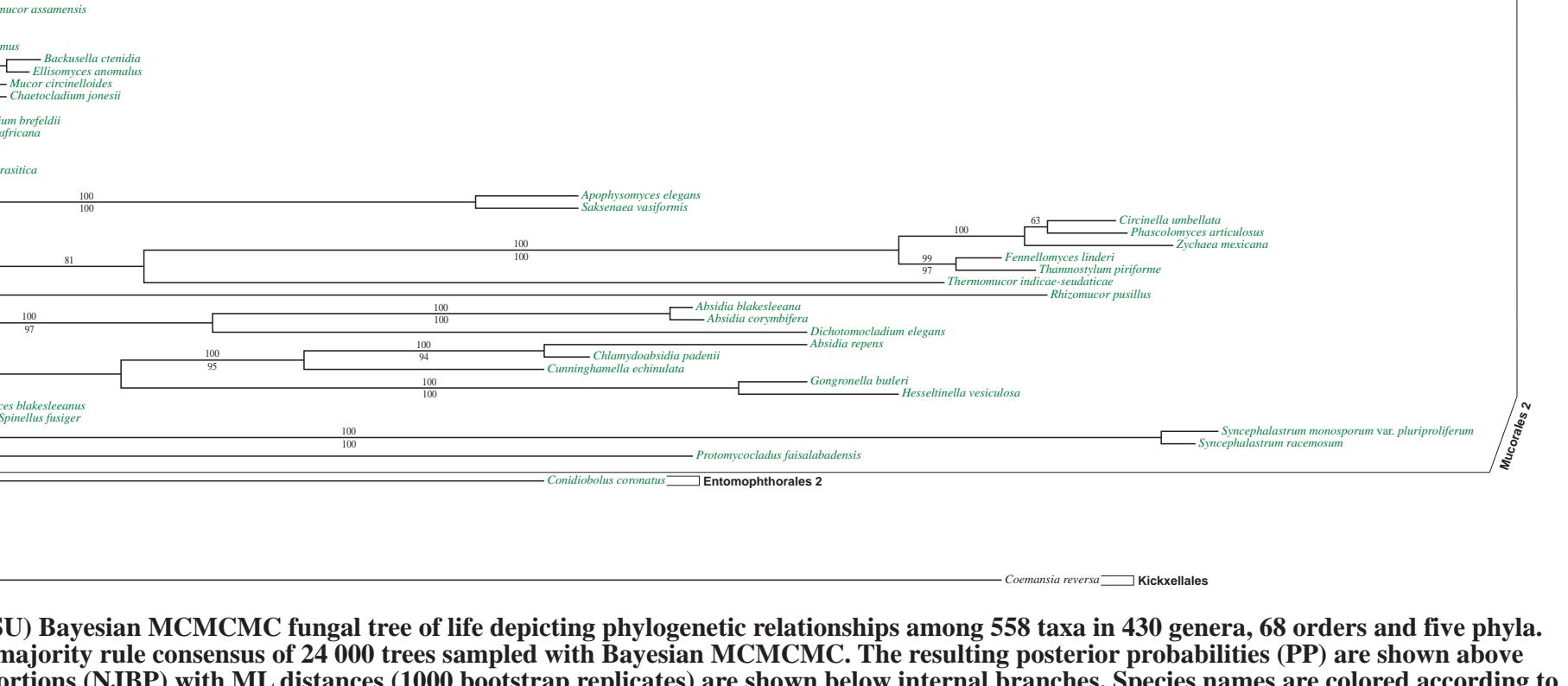


Figure 4. Phylogenetic relationships among 103 ascomycete and basidiomycete taxa based on combined evidence from nucSSU, nuLSU, mitSSU, and RPB2. This phylogeny resulted from a 50% majority rule consensus of 20,000 trees sampled with Bayesian MCMCMC.