

SPECIAL ISSUE INTRODUCTION

# Tree of death: The role of fossils in resolving the overall pattern of plant phylogeny

Gar W. Rothwell<sup>1,2,5</sup>, Ignacio H. Escapa<sup>3</sup>, and Alexandru M. F. Tomescu<sup>4</sup>

Manuscript received 9 July 2018; revision accepted 16 July 2018.

<sup>1</sup> Department of Environmental and Plant Biology, Ohio University, Athens, OH 45701, USA

<sup>2</sup> Department of Botany and Plant Pathology, Oregon State University, 2082 Cordley Hall, Corvallis, OR 97331, USA

<sup>3</sup> Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Museo Paleontológico Egidio Feruglio, Avenida Fontana 140, Trelew, Chubut 9100, Argentina

<sup>4</sup> Department of Biological Sciences, Humboldt State University, Arcata, CA 95521, USA

<sup>5</sup> Author for correspondence (e-mail: rothwell@ohio.edu)

**Citation:** Rothwell, G. W., I. H. Escapa, and A. M. F. Tomescu. 2018. Tree of death: The role of fossils in resolving the overall pattern of plant phylogeny. *American Journal of Botany* 105(8): 1–4.  
**doi:**10.1002/ajb2.1138

Systematics has a long history of conflicting results arising from analyses of different categories of biologically informative data and differing analytical methods. Until the advent of numerical methods in systematics in the 1960s, evolutionary relationships were most often inferred from a small subset of available characters (e.g., floral structure, fruit type, pollen characters, leaf form, cuticular anatomy), and hypotheses of relationships were not routinely tested against the results from other subsets of the data (see Nixon, 1996). In retrospect, we now realize that only partly accurate “phylogenies” became widely accepted, through either relatively universal popularity or by the force-of-will of influential authors (e.g., Haeckel, 1876). For example, while both the Takhtajan (1969) and Cronquist (1981) systems of classification for flowering plants have been extremely useful in a taxonomic context, they now are recognized to be collections of systematic hypotheses that were largely untested scientifically.

With the advent of cladistic methodologies that rely on shared derived characters (Hennig, 1950; Wagner, 1952), a standard method for formulating and testing systematic hypotheses (i.e., in the form of phylogenetic trees) was developed, thus producing much more systematically rigorous results. In the intervening period, remarkable progress has been made in resolving relationships among living plants, particularly with the continuing development of new sources of data (e.g., nucleotide sequences) and innovative analytical methods (e.g., maximum likelihood and Bayesian

modeling). Moreover, as the time necessary to develop nucleotide sequence characters has diminished dramatically, levels of systematic resolution are widely believed to have increased significantly.

In the current literature well over 95% of phylogenetic hypotheses are derived from studies of only living organisms and employ only nucleotide sequence data (see papers included in the recent special issue of the *American Journal of Botany*, “Using and navigating the plant tree of life”; e.g., Soltis et al., 2018). By contrast, nucleotide sequence characters have not become available for virtually any extinct species. At the same time, a small number of paleobotanists are discovering new data from the fossil record at the fastest rate ever, revealing that only a small percentage of potentially available data for extinct species has thus far been extracted. An additional current trend is for the usefulness of the most appropriate algorithm for systematic analysis of morphological characters (i.e., maximum parsimony) to be widely questioned, as maximum likelihood and Bayesian inference methodologies for modeling phylogenetic pattern have increased in popularity (e.g., O’Reilly et al., 2016, 2018; but see also Goloboff et al., 2017, 2018; Schrago et al., 2018).

Against this backdrop, many have come to accept the prediction of Hennig (1966) and hypothesis of Patterson (1981) that input from the fossil record does not alter patterns of relationships resolved from extant species alone; see Donoghue et al. (1989) for a more detailed history of this avenue of thought. However, studies that contradict those assertions extend back to the early 1980s, and

several examples of fossils altering the results of analyses conducted using only living species have since been published (e.g., Donoghue et al., 1989; Rothwell and Nixon, 2006; Hilton and Bateman, 2006; Atkinson, 2018). Even if adding fossil taxa to an analysis does not change relationships among living species, their inclusion is vital for understanding homologies, as well as the mode and tempo of evolution (Donoghue et al., 1989; Nixon, 1996). Thus, the question remains: is it valid to assume that relationships among living species represent the overall pattern of phylogeny for major clades? Or, does the exclusion of the extinct plants from phylogenetic analyses yield primarily relationships among living species that do not necessarily reflect overall phylogenetic pattern and that hold a tenuous relevance for the reconstruction of deep evolutionary history?

To address these fundamental questions of phylogenetic accuracy, we have assembled a representative body of studies that span the systematic gamut of land plants to explore the role of paleontological data in phylogenetic studies. The goal of this enterprise is to clarify the level of phylogenetic resolution that has been achieved to date and to help focus future studies on the most productive approaches for ultimately resolving the overall pattern of land plant phylogeny. This special issue of the *American Journal of Botany* includes 12 studies that focus primarily on paleontological data to illustrate the role of fossils in the reconstruction of phylogenetic pattern, to reveal otherwise unavailable aspects of character evolution, to establish minimum clade ages, and to help clarify relationships among land plants. Although the included studies encompass an incomplete sampling of terrestrial embryophytes and test phylogenetic hypotheses to differing degrees, they all contribute to increasing the sampling density from the paleontological record and clearly illustrate the vital role played by fossils in phylogenetic resolution.

The collection begins with the analysis of polytrichacean moss phylogeny by Bippus et al. (2018), which reveals that resolved topologies are substantially altered by the use of alternate outgroups. This study includes discrete and continuous morphological data, analyzed together and showing the increasing relevance of continuous characters in plant phylogenetic analyses (see also Elgorriaga et al., 2018; Escapa et al., 2018). The problem of outgroup selection in tree searches is also illustrated by the Rothwell et al. (2018) analysis of marattialean fern phylogeny, which cannot be adequately addressed without dramatically increasing the sampling density of extinct species.

Several of the contributions to this volume emphasize that living species comprise only minor segments of many clades with long fossil histories. In addition to dissecting long branches on phylogenetic trees, all of these analyses contribute valuable information about homologies of organs and patterns of morphological evolution. Toledo et al. (2018) explore relationships at the base of the lignophyte clade, as revealed by the most ancient stem-group seed plants and several other related groups, none of which is represented in the modern flora, but all of which are vital for our understanding of the evolution and homologies of leaves, stelar architecture, branching, gymnospermous reproduction, and seeds.

Elgorriaga et al. (2018) have developed a phylogenetic morphological analysis including the best-represented equisetalean plants in the fossil record, thus producing a first step toward understanding the early evolution of this ancient lineage. This study emphasizes that equisetalean sphenopsids have been vital components of terrestrial biotas since the Upper Devonian and that the sphenopsid clade cannot be assumed to represent the basal node of a fern clade (e.g., as resolved in Rothfels et al., 2015, for living species) simply because it

has only about 15 living species. Inclusion of extinct species in the sphenopsid analysis produces no substantial change in relationships among living *Equisetum* species (Elgorriaga et al., 2018), but when coupled with information from genetic regulatory mechanisms (Tomescu et al., 2017), relationships among the fossil and living taxa provide otherwise inaccessible data about evolution and homologies of organs, of body plans, and of reproductive structures.

Rothwell et al. (2018) and Choo and Escapa (2018) address the overall patterns of phylogeny for marattialean ferns and dipteridacean ferns, respectively. Both studies highlight the fact that living species comprise only minor components of diverse and species-rich clades with long fossil records and high levels of extinction. For the Marattiales, alternative outgroup selections for the clade reveal that relationships among genera with living species are not as clearly resolved as commonly believed (Rothwell et al., 2018). For the Dipteridaceae, the Choo and Escapa (2018) analysis reveals that morphological diversity found in living genera is only a fraction of the morphological diversity present in the Mesozoic. This study represents the first morphological phylogenetic analysis for Dipteridaceae, and taxonomic treatments for extinct species of the family are proposed.

Escapa et al. (2018) contrast the results for analyses of araucarian conifers that either exclude or include extinct species to emphasize that the two main lineages, *Araucaria* and the agathoid clade, co-occurred in the Mesozoic, refuting previous hypotheses that proposed a Cenozoic age for the Araucariaceae crown group. This study also provides a stratigraphic context for understanding the pattern of seed cone evolution within Araucariaceae. The position of particular fossil species within the agathoid clade suggests morphological trends for the evolution of important features within the family (e.g., seed wings).

Among families of conifers with living representatives, Pinaceae has the most densely sampled fossil record, particularly for seed cones (Smith et al., 2016). Gernandt et al. (2018) employ this rich fossil record in conjunction with both nucleotide sequence characters and morphological characters of living species to provide the most comprehensive phylogenetic analysis for the family and to refine our emerging understanding of the overall pattern of phylogeny for the Pinaceae.

Among all of the papers in this volume, Atkinson's treatment of Cornales (Atkinson, 2018) provides the most definitive evidence for how inclusion of extinct species resolves deep node polytomies that occur in the results of analyses that include only living species. Atkinson's analysis of cornalean fruits through time both increases the resolution of deep internal nodes for phylogenetic analyses of early asterids and documents the pattern of morphospace exploration for cornalean fruits during the time interval from the Late Cretaceous to the present.

Four additional studies of extinct species explore flowering plant evolution and phylogeny from fagalean inflorescences (Gandolfo et al., 2018), ericalean flowers (Crepet et al., 2018), zingiberalean monocot seeds (Smith et al., 2018), and Early Cretaceous seeds that are similar to those of basal living angiosperms (Friis et al., 2018). By comparing key morphological characters of fossil representatives to the pattern of structural features on the fagalean tree, Gandolfo et al. (2018) help dissect branches on the tree of living species and provide crucial data for employing fossils in future tests of systematic hypotheses that currently are restricted to living species.

Results of the analysis of Ericales by Crepet et al. (2018) are comparable to those from the analysis of molecular characters for

living species of the order, except that the position of fossils on the tree is resolved. Until the density of sampling for fossil species is high enough to provide a substantial body of data for meaningful hypothesis tests, dissection of branches on trees derived from the analysis of living species, such as in this analysis of Ericales, will continue to be the principal contribution of fossils to the pattern of phylogeny for many clades.

The inclusion of fossil seeds in morphological and combined morphological/nucleotide sequence analyses by Smith et al. (2018) helps explain how different topologies from morphological and molecular data sets result from plesiomorphic characters shared by Musaceae, Zingiberaceae, Costaceae, and many zingiberalean fossils. This study emphasizes the pressing need for fossils to be incorporated into phylogenetic analyses, not only to place those fossils on the tree, but also to better understand morphological evolution, to establish character polarities, recognize homoplasy, and identify apparent data conflicts.

In the final contribution to this issue, Friis et al. (2018) examine Early Cretaceous seeds similar to those of living Austrobaileyales and Nymphaeales and highlight the extent to which the morphology of extant angiosperms is not representative of the diversity that once existed among early-diverging members of the clade. These authors clearly illustrate that relictual living species often do not reflect the majority of character combinations that actually characterize basal angiosperms.

#### **OF WHAT USE IS THE FOSSIL RECORD IN RESOLVING THE OVERALL PATTERN OF PLANT PHYLOGENY?**

Living biodiversity represents only a small fraction of the diversity of life that populated the Earth along its geologic history (Niklas, 1997). Consideration of the fossil record is the only way to access the extensive diversity of extinct plant structures, leading to a more accurate and complete picture of evolution. Well-understood extinct species frequently reveal patterns of past species richness and novel combinations of characters that are crucial for reconstructing evolution and phylogeny and that cannot be predicted based on the morphology of living species. Because biological evolution is a chaotic process in many respects, it is important to remember that “the origins of complex morphologies are unique events, and thus do not conform well to probabilistic models” (K. Nixon, Cornell University, personal communication, 2018). Therefore, neither probability estimates nor theoretical models can replace real evidence from the fossil record in reconstructing the overall pattern of phylogeny.

In studies of relationships at shallow phylogenetic nodes (i.e., most assemblages that represent species of a single genus or genera of a single family), the effect of fossils is often minimal, and the extinct species tend to nest within clades that are determined primarily by living species. However, when successively deeper phylogenetic nodes are addressed, the clades are increasingly likely to have suffered large percentages of extinction. Where 90% or more of species richness has been pruned and is absent from the living flora, very deep nodes of the tree may not be represented by any living species at all (e.g., Toledo et al., 2018). As a result, the inclusion of extinct species has much greater potential to alter the topology of the resulting trees when deeper internal nodes are addressed. In these instances, topologies of clades may be obscured by extinction to a level where the overall patterns of relationships either remain equivocal or are not accurately represented by analyses of only

living species (e.g., sphenopsids [Elgorriaga et al., 2018] and flowering plants [Atkinson, 2018]).

From a practical perspective, inclusion/exclusion experiments with fossils have identified five principal ways in which results of phylogenetic analyses can be affected by the inclusion of extinct species. (1) Fossils may attach to the tree in such a way as to dissect branches and increase clade completeness, but otherwise do not alter the overall topology of the tree. This case is most typical of studies including fossils that are part of the crown of living genera or sometimes families. The analysis by Crepet et al. (2018) is an example of this outcome. (2) Fossils may substantially alter the overall topology of the tree, such that relationships among living species are changed (i.e., the Rothwell et al. [2018] analysis of Marattiales). (3) Based either on conflict of characters or lack of information, fossils may cause nodes to collapse. (4) Fossils may allow internal polytomies to be resolved, such that deep internal nodes are more highly resolved (i.e., the Atkinson [2018] analysis of asterids). (5) Fossils may enrich the topology and overall species richness of clades that have suffered extremely high levels of extinction. Understanding phylogenetic relationships within a clade is also important for clarification of the basal node of the clade. In turn, this information is crucial for understanding the relationships of that clade with other lineages, such that phylogenetic patterns are more finely resolved or revealed for the first time (e.g., several contributions in this volume, particularly the Elgorriaga et al. [2018] analysis of equisetalean sphenopsids).

Data developed in the studies included in this special issue repeatedly contradict the assumptions of Hennig (1966) and falsify the hypothesis of Patterson (1981), which together form the foundation upon which all of the arguments for building phylogenies from living species alone (and then adding fossils in a post-hoc fashion) are built. Therefore, all hypotheses of overall patterns of phylogeny that derive from systematic relationships resolved for living species alone require testing by inclusion/exclusion experiments with extinct species. To assume otherwise invites a major epistemic flaw because phylogenetic hypotheses based exclusively on living species cannot be tested empirically if fossils, the unique independent source of evidence that can support or reject those hypotheses, are not considered in the analyses. These problems are compounded further in the generation of hypotheses for clade ages, if fossils are used only to calculate the clade ages after an untested “phylogeny” based exclusively on living species has been generated (e.g., Wilf and Escapa, 2015, 2016).

#### **OUTLOOK AND FUTURE DIRECTIONS**

This special issue is about the vital importance of fossils for accurately reconstructing the overall pattern of phylogeny for land plants, which also highlights the importance of continuing to work on the morphology and anatomy of both living and fossil species. We cannot emphasize too strongly that morphological and anatomical features of both living and extinct plants have been only superficially examined to date. The idea that such features are well known for living plants is highly inaccurate, and the widely held belief that the fossil record is both well sampled and woefully incomplete is not justified. Organismal botany, and particularly structural botany, simply went out of fashion before the work was done. If this compendium of studies is to make a meaningful contribution to the future development of plant biology and plant phylogeny, that

contribution will be to reverse the practice of excluding fossils from systematic analyses and thereby to transform the results from simply clarifying relationships among living species to resolving overall patterns of plant phylogeny.

## LITERATURE CITED

- Atkinson, B. A. 2018. The critical role of fossils in inferring deep-node phylogenetic relationships and macroevolutionary patterns in Cornales. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1064>.
- Bippus, A. C., I. H. Escapa, and A. M. F. Tomescu. 2018. Wanted dead or alive (probably dead): Stem group Polytrichaceae. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1096>.
- Choo, T. Y. S., and I. H. Escapa. 2018. Assessing the evolutionary history of the fern family Dipteridaceae (Gleicheniales) by incorporating both extant and extinct members in a combined phylogenetic study. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1121>.
- Crepet, W. L., K. C. Nixon, and A. Weeks. 2018. Mid-Cretaceous angiosperm radiation and an asterid origin of bilaterality: Diverse and extinct "Ericales" from New Jersey. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1131>.
- Cronquist, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, NY, NY, USA.
- Donoghue, M. J., J. A. Doyle, J. Gauthier, A. G. Kluge, and T. Rowe. 1989. The importance of fossils in phylogeny reconstruction. *Annual Review of Ecology and Systematics* 20: 431–460.
- Elgorriaga, A., I. H. Escapa, G. W. Rothwell, A. M. F. Tomescu, and N. R. Cúneo. 2018. Origin of *Equisetum*: Evolution of horsetails (Equisetales) within the major euphyllophyte clade Sphenopsida. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1125>.
- Escapa, I. H., A. Iglesias, P. Wilf, S. A. Catalano, M. A. Caraballo-Ortiz, and N. R. Cúneo. 2018. *Agathis* trees of Patagonia's Cretaceous–Paleogene death landscapes and their evolutionary significance. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1127>.
- Friis, E. M., P. R. Crane, and K. R. Pedersen. 2018. *Tanispermum*, a new genus of hemi-orthotropous to hemi-anatropous angiosperm seeds from the Early Cretaceous of eastern North America. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1124>.
- Gandolfo, M. A., K. C. Nixon, W. L. Crepet, and D. A. Grimaldi. 2018. A Late Cretaceous fagalean inflorescence preserved in amber from New Jersey. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1103>.
- Germant, D. S., C. R. Arias, T. Terrazas, X. A. Duga, and A. Willyard. 2018. Incorporating fossils into the Pinaceae tree of life. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1139>.
- Goloboff, P. A., A. Torres, and J. S. Arias. 2017. Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics* 34: 407–437.
- Goloboff, P. A., A. Torres Galvis, and J. S. Arias. 2018. Parsimony and model-based phylogenetic methods for morphological data: comments on O'Reilly et al. *Palaeontology* 61: 625–630.
- Haeckel, F. M. V. 1876. The history of creation, vol. 2. Appleton and Co., NY, NY, U.S.A.
- Hennig, W. 1950. Grundzüge einer Theorie der phylogenetischen Systematik. Deutscher Zentralverlag, Berlin, Germany.
- Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana, IL, USA.
- Hilton, J., and R. M. Bateman. 2006. Pteridosperms are the backbone of seed-plant phylogeny. *Journal of the Torrey Botanical Society* 131: 119–168.
- Niklas, K. 1997. The evolutionary biology of plants. University of Chicago Press, Chicago, IL, USA.
- Nixon, K. C. 1996. Paleobotany in cladistics and cladistics in paleobotany: enlightenment and uncertainty. *Review of Palaeobotany and Palynology* 90: 361–373.
- O'Reilly, J. E., M. N. Puttick, L. Parry, A. R. Tanner, J. E. Tarver, J. Fleming, D. Pisani, and P. C. J. Donoghue. 2016. Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biology Letters* 12: 20160081.
- O'Reilly, J. E., M. N. Puttick, D. Pisani, and P. C. J. Donoghue. 2018. Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. *Palaeontology* 61: 105–118.
- Patterson, C. 1981. Significance of fossils in determining evolutionary relationships. *Annual Review of Ecology and Systematics* 12: 195–223.
- Rothfels, C. J., F.-W. Li, E. M. Sigel, et al. 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *American Journal of Botany* 102: 1089–1107.
- Rothwell, G. W., and K. C. Nixon. 2006. How does the inclusion of fossil data change our conclusions about the phylogenetic history of euphyllophytes? *International Journal of Plant Sciences* 167: 737–749.
- Rothwell, G. W., M. A. Millay, and R. A. Stockey. 2018. Resolving the overall pattern of marattialean fern phylogeny. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1115>.
- Schrago, C. G., B. O. Aguiar, and B. Mello. 2018. Comparative evaluation of maximum parsimony and Bayesian phylogenetic reconstruction using empirical morphological data. *Journal of Evolutionary Biology*. <https://doi.org/10.1111/jeb.13344>.
- Smith, S. Y., W. J. D. Iles, J. C. Benedict, and C. D. Specht. 2018. Building the monocot tree of death: Progress and challenges emerging from the macrofossil-rich Zingiberales. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1123>.
- Smith, S. Y., R. A. Stockey, G. W. Rothwell, and S. A. Little. 2016. A new species of *Pityostrobus* (Pinaceae) from the Cretaceous of California: moving towards understanding the Cretaceous radiation of Pinaceae. *Journal of Systematic Palaeontology*. <https://doi.org/10.1080/14772019.2016.1143885>.
- Soltis, D. E., M. J. Moore, E. B. Sessa, S. A. Smith, and P. S. Soltis. 2018. Using and navigating the plant tree of life. *American Journal of Botany* 105: 287–290.
- Takhtajan, A. L. 1969. Flowering plants: origin and dispersal. Oliver and Boyd, Edinburgh, UK.
- Toledo, S., A. C. Bippus, and A. M. F. Tomescu. 2018. Buried deep beyond the veil of extinction: Euphyllophyte relationships at the base of the spermatophyte clade. *American Journal of Botany*. <https://doi.org/10.1002/ajb2.1102>.
- Tomescu, A. M. F., I. H. Escapa, G. W. Rothwell, A. Elgorriaga, and N. R. Cúneo. 2017. Developmental programmes in the evolution of *Equisetum* reproductive morphology: a hierarchical modularity hypothesis. *Annals of Botany* 119: 489–505.
- Wagner, W. H. 1952. The fern genus *Diellia*: structure, affinities, and taxonomy. *University of California Publications in Botany* 26: 1–212.
- Wilf, P., and I. H. Escapa. 2015. Green Web or megabiased clock? Plant fossils from Gondwanan Patagonia speak on evolutionary radiations. *New Phytologist* 207: 283–290.
- Wilf, P., and I. H. Escapa. 2016. Molecular dates require geologic testing. Reply to Wang & Mao (2016; in this issue of *New Phytologist*, pp. 1353–1358). *New Phytologist* 209: 1359–1362.