

# Seed plant phylogeny: Demise of the anthophyte hypothesis?

Michael J. Donoghue\* and James A. Doyle†

**Recent molecular phylogenetic studies indicate, surprisingly, that Gnetales are related to conifers, or even derived from them, and that no other extant seed plants are closely related to angiosperms. Are these results believable? Is this a clash between molecules and morphology?**

Addresses: \*Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138, USA. †Section of Evolution and Ecology, University of California, Davis, California 95616, USA.  
E-mail: mdonoghue@oeb.harvard.edu  
jadoyle@ucdavis.edu

**Current Biology** 2000, 10:R106–R109

0960-9822/00/\$ – see front matter

© 2000 Elsevier Science Ltd. All rights reserved.

These are exciting times for those interested in plant evolution. The ever-increasing supply of molecular, and particularly sequence, data is providing a new window on plant relationships, with some surprising results. One of the thorny issues in plant phylogeny that has been addressed recently concerns the enigmatic Gnetales, one of the five extant lineages of seed plants. In the early 1900s, Gnetales were considered the closest living relatives of the flowering plants, or angiosperms. Subsequently, Gnetales were dissociated from angiosperms and often linked with conifers, and then, for the last decade, they were again linked with angiosperms, in a group called the anthophytes. But the molecular data obtained recently by a number of groups indicate that Gnetales are associated with conifers after all, and not closely related to angiosperms.

The five extant groups of seed plants are angiosperms, which with 250,000 or so species dominate most terrestrial habitats; conifers, with 550 species; cycads, with 150 species; ginkgos, with just one species; and Gnetales, with 70 species belonging to three very distinct groups, *Ephedra*, *Welwitschia* and *Gnetum*. The problem is how these lines are related to one another, and which of them is most closely related to the angiosperms. Its solution is complicated by the fact that the extant lineages radiated rapidly very long ago: seed plants appeared in the Late Devonian, about 370 million years ago, and at least three of the five lines probably split within the Late Carboniferous, 290–320 million years ago. Despite such common features as the seed, the five living groups are separated by great morphological gaps. Some features are shared by two or more groups, but patterns are complicated by conflicting similarities, some of which must be the result of convergent evolution.

The first suggestions that Gnetales are related to angiosperms were based on several obvious morphological similarities — vessels in the wood, net-veined leaves in *Gnetum*, and reproductive organs made up of simple, unisexual, flower-like structures, which some considered evolutionary precursors of the flowers of wind-pollinated Amentiferae, but others viewed as being reduced from more complex flowers in the common ancestor of angiosperms, Gnetales and Mesozoic Bennettitales [1]. These ideas went into eclipse with evidence that simple flowers really are a derived, rather than primitive, feature of the Amentiferae, and that vessels arose independently in angiosperms and Gnetales. Vessels in angiosperms seem derived from tracheids with scalariform pits, whereas in Gnetales they resemble tracheids with circular bordered pits, as in conifers. Gnetales are also like conifers in lacking scalariform pitting in the primary xylem, and in the scale-like and strap-shaped leaves of *Ephedra* and *Welwitschia*. Some authors also drew homologies between the ‘flowers’ of Gnetales and the fertile short shoots of Paleozoic conifers, which were transformed into woody cone scales in later conifers.

Beginning in the mid-1980s [1,2], numerical phylogenetic analyses of morphological characters offered hope of a solution to this conundrum. All these studies, though based on different interpretations of many key characters, indicated that Gnetales are the closest living relatives of angiosperms. Some even nested angiosperms within Gnetales, closer to *Welwitschia* and *Gnetum* than to *Ephedra* [3]. Gnetales, angiosperms and Bennettitales were grouped in a clade named the ‘anthophytes’ to highlight their flower-like reproductive structures. In a few years, the anthophyte hypothesis went from heresy to near-dogma. It also fueled reinterpretation of character evolution in seed plants, especially the origin of such distinctive angiosperm features as the carpels that house the ovules (future seeds) and double fertilization, which gives rise to embryo and endosperm. If angiosperms are nested within Gnetales, then vessels, net-veined leaves and the simple flowers of Gnetales and angiosperms such as Chloranthaceae can be homologous. If Bennettitales and other fossils are interpolated between angiosperms and Gnetales, then these similarities are parallelisms; the features that unite the two groups are more obscure, such as lignin composition, a tunica layer in the apical meristem, granular pollen structure and a reduced megaspore wall.

The first molecular phylogenetic studies of seed plants (reviewed in [4]) strongly supported the monophyly of Gnetales, refuting the view that they gave rise to

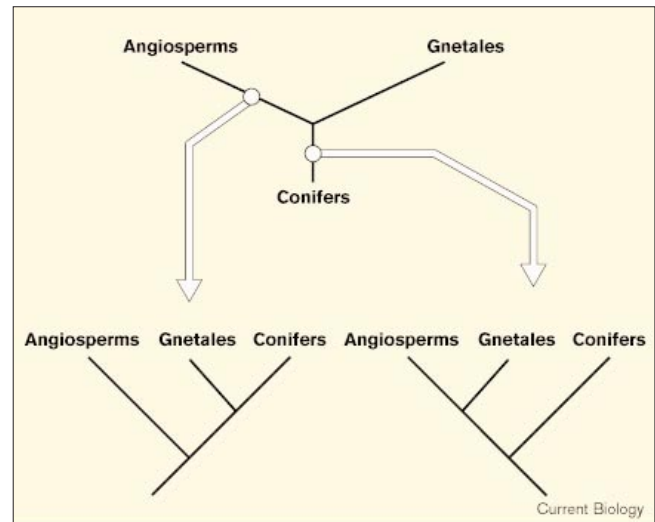
angiosperms. Only a few analyses of ribosomal DNA linked angiosperms and Gnetales, and this with low statistical support. Analyses of the chloroplast gene *rbcL* placed Gnetales at the base of the seed plants, followed by angiosperms (an arrangement supported by more recent studies of photosystem genes, M. Sanderson, personal communication). Studies of chloroplast internal transcribed spacer (ITS) sequences [5] and 18S rDNA [6] placed angiosperms at the base of seed plants and linked Gnetales with conifers. These trees more seriously contradict the anthophyte hypothesis, as there is no way they can be rerooted so that angiosperms and Gnetales form a clade.

The results of these early molecular studies were thus rather inconclusive — given the apparent disagreements among genes, low levels of statistical support and inconsistencies with the fossil record, it still seemed that the morphological inference that angiosperms and Gnetales form a clade could be correct [7]. This picture has changed with a flood of new multigene studies [8–12], which provide much stronger evidence that Gnetales are related not to angiosperms, but rather to conifers. Rejection of the anthophyte hypothesis is news in itself, as a possible example of serious conflict between molecules and morphology. Some studies go further, however, in indicating that Gnetales originated within conifers, and in addition that the four living non-angiosperm seed plant groups — the extant ‘gymnosperms’ — form a clade, so that none of them is more closely related than any other to the angiosperms.

Hansen *et al.* [8] studied a very long sequence of chloroplast DNA, but in only one gnetalian (*Gnetum*), one conifer (*Pinus*) and three angiosperms; as cycads and *Ginkgo* were not included, they were not able to address the status of extant gymnosperms. As support for angiosperm monophyly is strong, this becomes a three-taxon problem, with angiosperms, Gnetales and conifers radiating from a central point (Figure 1). The relationships among these groups depend entirely on placement of the root of the tree: if it is rooted on the branch leading to angiosperms, as inferred by Hansen *et al.* [8], Gnetales are linked with conifers; if it is rooted on the conifer branch, Gnetales are linked with angiosperms. Hansen *et al.* rooted the tree by including the liverwort *Marchantia* as an outgroup. Liverworts have, however, been diverging from seed plants for about 450 million years; too much sequence divergence could cause *Marchantia* to attach spuriously to one of the longer seed plant branches, which happen to be the lines to angiosperms and *Gnetum*. Such ‘long-branch’ artifacts are difficult to prove, but the possibility should be investigated, for instance by including in the analysis sequences from a closer outgroup, such as a fern, or another gnetalian, preferably *Ephedra*, to subdivide the long branch to *Gnetum*.

In an analysis by Winter *et al.* [9] of genes encoding ‘MADS box’ proteins — believed to be transcription

Figure 1



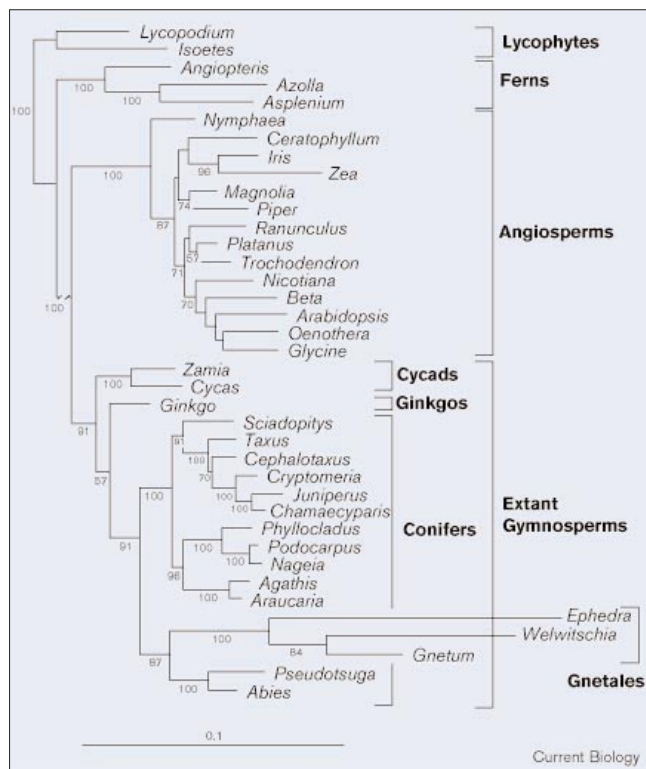
The effect of alternative rootings on inferred relationships of Gnetales, angiosperms and conifers.

factors with roles in regulating development — the only gnetalian group sampled was again *Gnetum*, and no cycad or *Ginkgo* sequences were included. Thirteen MADS box genes from *Gnetum* were analyzed, along with genes from angiosperms, conifers and ferns. In the four cases where *Gnetum* genes were linked with conifer and angiosperm genes, they were closer to the conifer genes. Two cases might be dismissed because of weak ‘bootstrap’ values — a measure of statistical support for a particular phylogenetic inference — but bootstrap values were very high in the other cases.

In three even more recent studies [10–12], the potential for long-branch artifacts was reduced by including cycads and *Ginkgo*, more conifers and angiosperms, and two or all three genera of Gnetales. These analyses add new mitochondrial sequences — *cox1*, *atp1*, *atpA*, *matR* and small subunit rDNA — to data on *rbcL*, nuclear 18S rDNA and in one case [12] the chloroplast gene *atpB*. Those analyses that included outgroups [10,11] indicate that extant gymnosperms form a clade, while they all place the Gnetales within the conifers, directly linked with Pinaceae (Figure 2) — the ‘gnepines’ hypothesis [11]. Although seed plants can be rerooted so that extant gymnosperms are paraphyletic, there is no way to reroot them so that Gnetales and angiosperms form a clade.

These studies provide the best evidence yet against the anthophyte hypothesis and in favor of a connection between Gnetales and conifers. The strongest results are those based on mitochondrial genes, where the conifer–Gnetales clade is supported by high bootstrap values. Analyses of *rbcL* and nuclear rDNA show weaker support for this clade — for

Figure 2



Seed plant phylogeny obtained in a maximum likelihood analysis of *rbcL*, nuclear 18S rDNA and mitochondrial small subunit rDNA sequences. The numbers associated with branches are bootstrap values from a parsimony analysis. (Modified from [11].)

example, it is obtained only after exclusion of third codon positions in *rbcL*, as is also the case in studies of photosystem genes (M. Sanderson, personal communication). And a few analyses of individual genes yield trees compatible with the anthophytes (if seed plants are rerooted). Nevertheless, we are impressed by the fact that virtually no signal for the anthophytes seems to exist in these data sets, as it certainly should if such a clade exists. The most worrisome point is that Gnetales are an unusually long branch (Figure 2), especially in the mitochondrial trees.

With regard to the conclusion that extant gymnosperms are monophyletic, which also seems most strongly supported by mitochondrial data, more work is needed — for example, using Kishino-Hasegawa tests — to determine just how strongly the data reject rootings in which extant gymnosperms are paraphyletic. A position of the root between the cycad and *Ginkgo* nodes might be very difficult to detect, because this branch is so short compared to the long branches to angiosperms and Gnetales.

The novel conclusion that Gnetales are nested within conifers is troublesome for other reasons. The nuclear

rDNA data show fairly strong support for the monophyly of the conifers, with Gnetales their sister group. Furthermore, all conifers — but not Gnetales — have lost one copy of the inverted repeat in the chloroplast genome [13]. It is worth considering whether this incongruence among genomes might mean that the gene trees are genuinely different, most likely due to lineage sorting. For example, if the nuclear rDNA tree — with conifers forming a monophyletic group — reflects the true species tree, the mitochondrial gene trees might reflect the retention of one ancestral mitochondrial type in Gnetales and Pinaceae, and a second type in other conifers. This may seem far-fetched, but it could be addressed by measures of data set heterogeneity and analyses of additional nuclear genes.

What can we conclude from these analyses? The most important and believable result is that the anthophyte hypothesis is incorrect — as first emphasized in [5] — and that Gnetales are more closely related to conifers than to angiosperms. Gnetales are also linked with conifers in analyses of *rpoC1* sequences [14] and in some unpublished analyses by Graham and Olmstead, Frohlich, and Sanderson and colleagues (personal communications). We are less convinced that extant gymnosperms form a clade, or that Gnetales are derived from within conifers.

Is this a severe clash between morphology and molecules? To some extent, the conflict may have been exaggerated. Much has been made of the agreements among morphological trees; less attention has been paid to levels of support for the anthophytes. In the most recent analysis [15], bootstrap values in this part of the tree were low. Nevertheless, forcing angiosperms to the base of extant seed plants and linking Gnetales and conifers added ten evolutionary steps [7], suggesting that the conflict is not easily dismissed. The morphological data conflict still more with trees that place Gnetales within conifers. This arrangement would entail the evolution and then loss in Gnetales of several conifer features, such as resin canals and tiered proembryos. Ironically, it would also eliminate a prime morphological argument for a relationship between the two groups — the similarity between the fertile short shoots of Paleozoic conifers and the ‘flowers’ of Gnetales, as in modern conifers the short shoots are represented by woody cone scales that look nothing like flowers.

Contrary to what has been implied in some articles, few morphological phylogenetic analyses have used flowers and double fertilization as characters, and most of them indicate that vessels and dicot-like leaves arose independently in angiosperms and Gnetales. Focus on these features distracts attention from the more cryptic characters that *have* united the two groups. It is these that need scrutiny, especially with an eye toward understanding how they might have led us astray. For example, some of these characters — such as megaspore wall or air sacs on the

pollen — involve reduction, which may produce similar results, even when it occurs in independent lines.

Some of the morphological features claimed to link angiosperms and Gnetales do differ in detail: the tunica consists of two cell layers in angiosperms, one in Gnetales; the megaspore wall is thin in Gnetales, but absent in angiosperms. It seemed reasonable to code these states as potentially homologous, but they might not meet stricter *a priori* standards of comparability. This sort of reassessment is not just cathartic — we need to understand morphological characters much better than we apparently do if we are to understand the evolution of seed plants. If no living seed plants are closely related to angiosperms, the only way to reconstruct the origin of angiosperms is by fitting fossils into the picture, and this can only be done by analysis of morphological characters.

Although the studies reviewed here provide extremely important new data on seed plant phylogeny, additional tests are needed before a new dogma sets in. Most generally, we need detailed and creative analyses of factors underlying conflicts in the molecular data, along the lines of work by Graham and Olmstead and by Sanderson and colleagues (personal communications). Studies of more nuclear genes are needed to explore conflicts regarding the monophyly of conifers. More attention must also be paid to morphological characters, both to make sense of their evolution and to enhance the credibility of future phylogenetic analyses using fossils.

#### Acknowledgements

We thank C. dePamphilis, M. Frohlich, S. Graham, J. Palmer, R. Olmstead, and M. Sanderson for sharing unpublished data and manuscripts.

#### References

- Doyle JA, Donoghue MJ: **Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach.** *Bot Rev* 1986, **52**:321-431.
- Crane PR: **Phylogenetic analysis of seed plants and the origin of angiosperms.** *Ann Missouri Bot Gard* 1985, **72**:716-793.
- Nixon KC, Crepet WL, Stevenson D, Friis EM: **A reevaluation of seed plant phylogeny.** *Ann Missouri Bot Gard* 1994, **81**:484-533.
- Doyle JA: **Phylogeny of vascular plants.** *Annu Rev Ecol Syst* 1998, **29**:567-599.
- Goremykin V, Bobrova V, Pahnke J, Troitsky A, Antonov A, Martin W: **Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcl* data do not support gnetalean affinities of angiosperms.** *Mol Biol Evol* 1996, **13**:383-396.
- Chaw SM, Zharkikh A, Sung HM, Lau TC, Li WH: **Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences.** *Mol Biol Evol* 1997, **14**:56-68.
- Doyle JA: **Molecules, morphology, fossils, and the relationship of angiosperms and Gnetales.** *Mol Phylog Evol* 1998, **9**:448-462.
- Hansen A, Hansmann S, Samigullin T, Antonov A, Martin W: **Gnetum and the angiosperms: molecular evidence that their shared morphological characters are convergent, rather than homologous.** *Mol Biol Evol* 1999, **16**:1006-1009.
- Winter KU, Becker A, Münster T, Kim JT, Saedler H, Thiessen G: **MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants.** *Proc Natl Acad Sci USA* 1999, **96**:7342-7347.
- Bowe LM, Coat G, dePamphilis CW: **Phylogeny of seed plants based on all three plant genomic compartments: extant gymnosperms are monophyletic and Gnetales are derived conifers.** *Proc Natl Acad Sci USA*, in press.
- Chaw, SM, Parkinson CL, Cheng Y, Vincent TM, Palmer JD: **Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers.** *Proc Natl Acad Sci USA*, in press.
- Qiu YL, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen Z, Savolainen V, Chase MW: **The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes.** *Nature* 1999, **402**:404-407.
- Raubeson LA, Jansen RK: **A rare chloroplast-DNA structural mutation is shared by all conifers.** *Biochem Syst Ecol* 1992, **20**:17-24.
- Samigullin TKh, Martin WF, Troitsky AV, Antonov AS: **Molecular data from the chloroplast *rpoC1* gene suggest a deep and distinct dichotomy of contemporary spermatophytes into two monophyla: gymnosperms (including Gnetales) and angiosperms.** *J Mol Evol* 1999, **49**:310-315.
- Doyle JA: **Seed plant phylogeny and the relationships of Gnetales.** *Int J Plant Sci* 1996, **157**:S3-S39.