Molecules, Morphology, Fossils, and the Relationship of Angiosperms and Gnetales

James A. Doyle

Section of Evolution and Ecology, University of California, Davis, California 95616

Morphological analyses of seed plant phylogeny agree that Gnetales are the closest living relatives of angiosperms, but some studies indicate that both groups are monophyletic, while others indicate that angiosperms are nested within Gnetales. Molecular analyses of several genes agree that both groups are monophyletic, but differ on whether they are related. Conflicts among morphological trees depend on the interpretation of certain characters; when these are analyzed critically, both groups are found to be monophyletic. Conflicts among molecular trees may reflect the rapid Paleozoic radiation of seed plant lines, aggravated by the long branches leading to extant taxa. Trees in which angiosperms are not related to Gnetales conflict more with the stratigraphic record. Even if molecular data resolve the relationships among living seed plant groups, understanding of the origin of angiosperm organs will require integration of fossil taxa, necessarily using morphology. © 1998 Academic Press

INTRODUCTION

One of the most intractable problems in plant systematics and evolution is the origin of angiosperms, or flowering plants, the dominant plant group on earth today. One aspect of this question is which of the other seed plants, or "gymnosperms," are most closely related to angiosperms—their closest outgroup (sister group) or outgroups. To a paleobotanist, this involves a consideration of the many extinct Paleozoic and Mesozoic seed plants, but to a molecular systematist, who can study only extant organisms, the question is which living seed plants are closest to the angiosperms. The only living group that has attracted significant attention is the Gnetales (Martens, 1971; Friedman, 1996), which include three divergent genera. *Gnetum* is a tropical rainforest tree or vine, with leaves virtually identical to those of a dicotyledonous angiosperm. *Ephedra* is a desert shrub with small, pointed leaves, but as in *Gnetum* these are opposite. *Welwitschia,* in the deserts of southwest Africa, produces only one pair of strap-shaped leaves. All three genera resemble most angiosperms in having special water-conducting vessels in the wood, and they all have reproductive structures organized into compound strobili and seeds with a micropylar tube. Recently, Friedman (1990, 1992) and Carmichael and Friedman (1996) confirmed earlier reports that Gnetales have part of the angiosperm process of double fertilization: both of the two sperm produced by the male gametophyte (in the pollen) fuse with nuclei in the female gametophyte (in the ovule), but the second fusion produces an additional embryo or embryos rather than triploid endosperm tissue, as in angiosperms. The other living seed plants (cycads, *Ginkgo,* and conifers) have rarely been associated with angiosperms.

Molecular data may also help clarify the "rooting" of the angiosperms, or determining which angiosperm groups are "basal" (i.e., derived from the first splits in angiosperm phylogeny). This may help in reconstructing states in the common ancestor of living angiosperms and thus in connecting them with other groups. However, in this article I will consider only relationships of angiosperms with other seed plants.

At first sight, morphological and molecular data on the relationship of angiosperms and Gnetales may appear to conflict, but closer examination suggests that they are actually complementary. On one question, whether angiosperms and Gnetales are both monophyletic, molecular data are in agreement, whereas morphological analyses have given inconsistent results. However, on the question of whether the two groups are related at all, it is the molecular data that are inconsistent and morphological data that are unanimous. I will argue that the disagreements among morphological analyses are due to questionable interpretations of characters, and when these are corrected one obtains results that agree with molecular data. The disagreements among molecular analyses may reflect inherent limitations of molecular data in resolving ancient radiations. I will also argue that consideration of the great diversity of fossil plants is necessary in order to understand the origin of angiosperm organs, since structures in living seed plants are too highly modified, and that fossils may provide independent tests of hypotheses

448

based on molecular data, both as additional taxa and as sources of stratigraphic evidence that bears on tree topology. Since fossils can be integrated into phylogenies only through their morphological characters, these are arguments for continued attention to morphology, even if molecules become the primary basis for reconstructing relationships among living groups.

HYPOTHESES ON RELATIONSHIPS

Precladistic Views

One of the earliest hypotheses, proposed by Wettstein (1907) but often called the "englerian" theory, postulated that angiosperms were derived from Gnetales, which were in turn related to conifers. This was based on comparisons between the compound strobili of Gnetales, which consist of bracts with simple, unisexual "flowers" in their axils, with one orthotropous (erect) ovule in the female flowers, and the inflorescences of the wind-pollinated "Amentiferae," which are also made up of simple, unisexual flowers, with one orthotropous ovule. This view fell out of favor with increasing evidence that Amentiferae are advanced within angiosperms, compared to "Magnoliidae," such as *Magnolia*, with showy, bisexual flowers, and this conclusion has been amply confirmed by both fossil and molecular data (Doyle, 1978; Chase *et al.*, 1993; Crane *et al.*, 1995). However, englerian ideas are far from dead, since one magnoliid group, Chloranthaceae, is also like Gnetales in having opposite leaves, inflorescences of simple flowers, and one orthotropous ovule.

A competing theory was proposed by Arber and Parkin (1907), inspired by the large, bisexual "flowers" of Mesozoic Bennettitales, which had cycad-like leaves. In Bennettitales, the ovules were borne on simple stalks, rather than anything like a carpel, the closed ovule-bearing organ of angiosperms. Arber and Parkin therefore proposed that the two groups were derived from a common ancestor with pinnate sporophylls, with the megasporophylls folded to enclose the ovules in angiosperms, but reduced in Bennettitales. Arber and Parkin (1908) thought that Gnetales were also related to angiosperms, but they interpreted their simple flowers as reduced rather than primitive.

Angiosperms have also been associated with the Mesozoic genus *Caytonia*, which had palmately compound leaves with simple reticulate venation and megasporophylls bearing two rows of anatropous (reflexed) "cupules," each containing several ovules (Gaussen, 1946; Stebbins, 1974; Doyle, 1978). This would explain the fact that the ovules of angiosperms are usually anatropous and bitegmic (with two integuments). If the number of ovules in a *Caytonia* cupule was reduced to 1, the cupule would resemble an angiosperm ovule, with the outer integument corresponding to the cupule wall. The carpel might be derived by widening and folding of the sporophyll rachis.

Others have linked angiosperms to the Permian glossopterids, which had simple leaves with simple reticulate venation and leaflike ovulate structures bearing one or several cupules or "sporophylls" (Stebbins, 1974; Retallack and Dilcher, 1981). The cupules could be transformed into bitegmic ovules as proposed for *Caytonia*, while the carpel could be derived from the leaflike portion.

Phylogenetic Analyses Based on Morphology

Of the several cladistic analyses of seed plants based on morphological data, all agree that Gnetales are the closest living relatives of angiosperms, but some analyses indicate that both groups are monophyletic, while others indicate that angiosperms are nested within Gnetales. They also differ in how fossils are interpolated among living taxa—some link angiosperms and Gnetales with *Caytonia* and glossopterids, while others link them with coniferopsids (conifers, ginkgos, Paleozoic cordaites). For this reason, they vary greatly in implications for character evolution, such as origin of the carpel. Some of these scenarios might never be envisioned if only living taxa were considered, as is the case with molecular analyses (cf. Doyle and Donoghue, 1987, 1992).

In the analysis of Crane (1985), angiosperms and Gnetales were sister groups. The two were in turn related to Bennettitales and the Mesozoic genus *Pentoxylon*, and the whole clade was associated with glossopterids, *Caytonia*, and Triassic corystosperms. Since Bennettitales, Gnetales, and angiosperms all have flowerlike structures, this would support an Arber and Parkin (1907) scenario for floral evolution, with angiosperm carpels and bennettitalian stalked ovules both derived from pinnate sporophylls like those of *Caytonia*.

To test these results, Doyle and Donoghue (1986, 1992) added several characters that might support alternative relationships, but this had little effect (Fig. 1). Most importantly, angiosperms were separated from Gnetales, as the sister group of Bennettitales, *Pentoxylon*, and Gnetales. These four groups were called "anthophytes" because their common ancestor would have had flowerlike structures. *Caytonia* was the sister group of the anthophytes, supporting the homology of the bitegmic ovule with the *Caytonia* cupule. The flowers of Gnetales would represent the culmination of a trend for reduction and simplification. Ovules of Gnetales have two integuments; the outer of these would correspond to the perianth of Bennettitales. These inferences illustrate the importance of fossils: without fossils interpolated among extant taxa, there would be no *Caytonia* prototype for the bitegmic angiosperm ovule, and the outer integument of angiosperms might be derived from the outer integument of Gnetales.
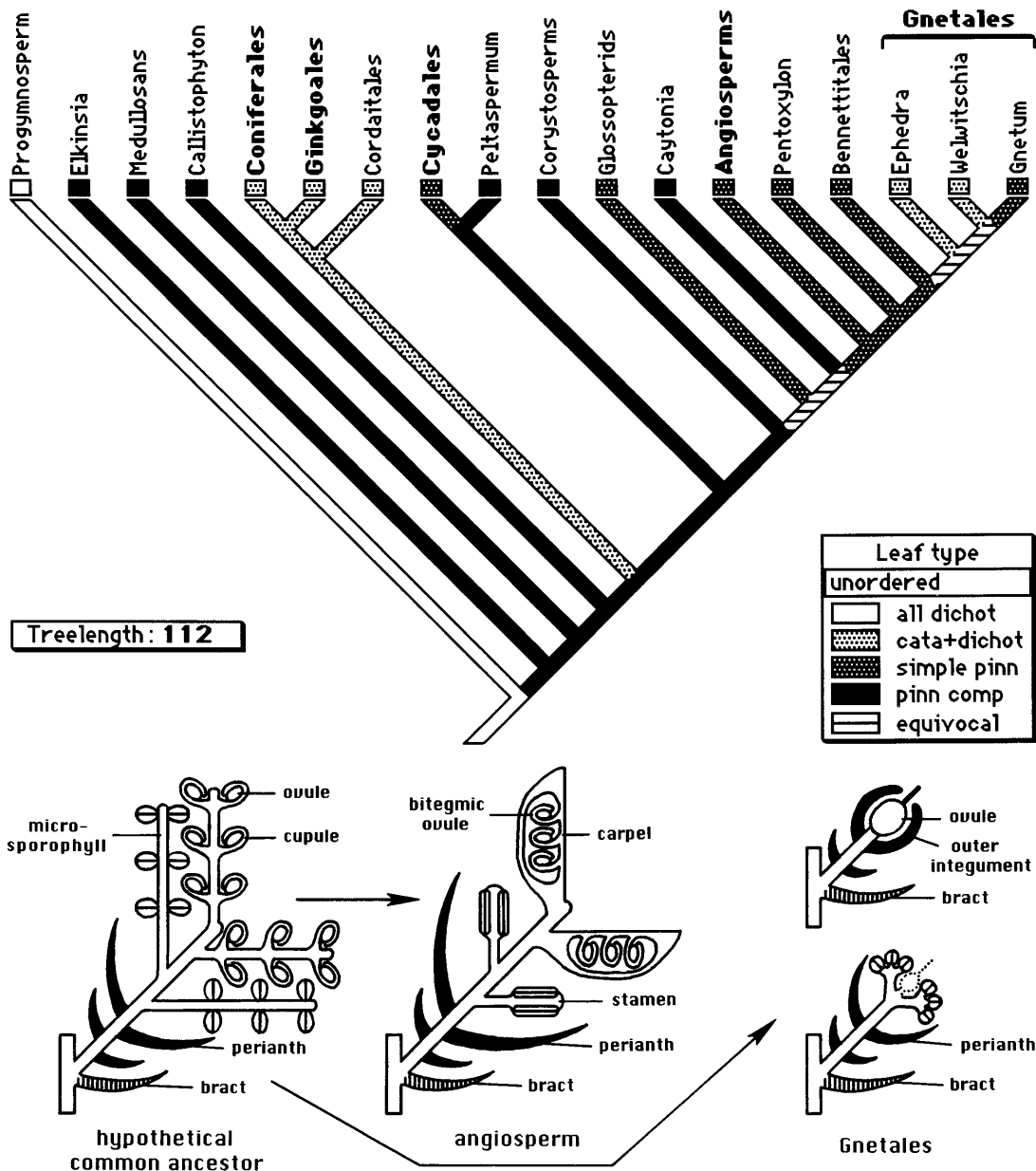
FIG. 1. Representative most parsimonious seed plant tree of Doyle and Donoghue (1992), showing evolution of leaf morphology and a scenario for floral evolution, with shading indicating homologous parts (modified from Doyle, 1994). Names of extant taxa are given in boldface type.

Unfortunately, these results were not robust: Doyle and Donoghue (1987, 1992) found very different trees that were only one or two steps longer. These include "neo-englerian" trees, with Gnetales basal in anthophytes and anthophytes linked with conifers (Fig. 2). This would imply that the simple flowers of Gnetales are primitive and homologous with the fertile short shoots of Paleozoic conifers. The angiosperm outer integument might be derived from the outer integument of Gnetales, itself derived from sterile appendages on the fertile short shoot; the carpel might corre-spond to the subtending bract (Doyle, 1994). Anthophytes were also linked with conifers in Rothwell and Serbet (1994), while angiosperms and Gnetales were sister groups within anthophytes. Angiosperms and Gnetales were linked with conifers in analyses of living plants only by Loconte and Stevenson (1990) and Doyle and Donoghue (1992). However, this result would not necessarily imply a neo-englerian scenario for floral evolution, since it would also be consistent with trees with glossopterids and *Caytonia* interpolated between conifers and anthophytes.
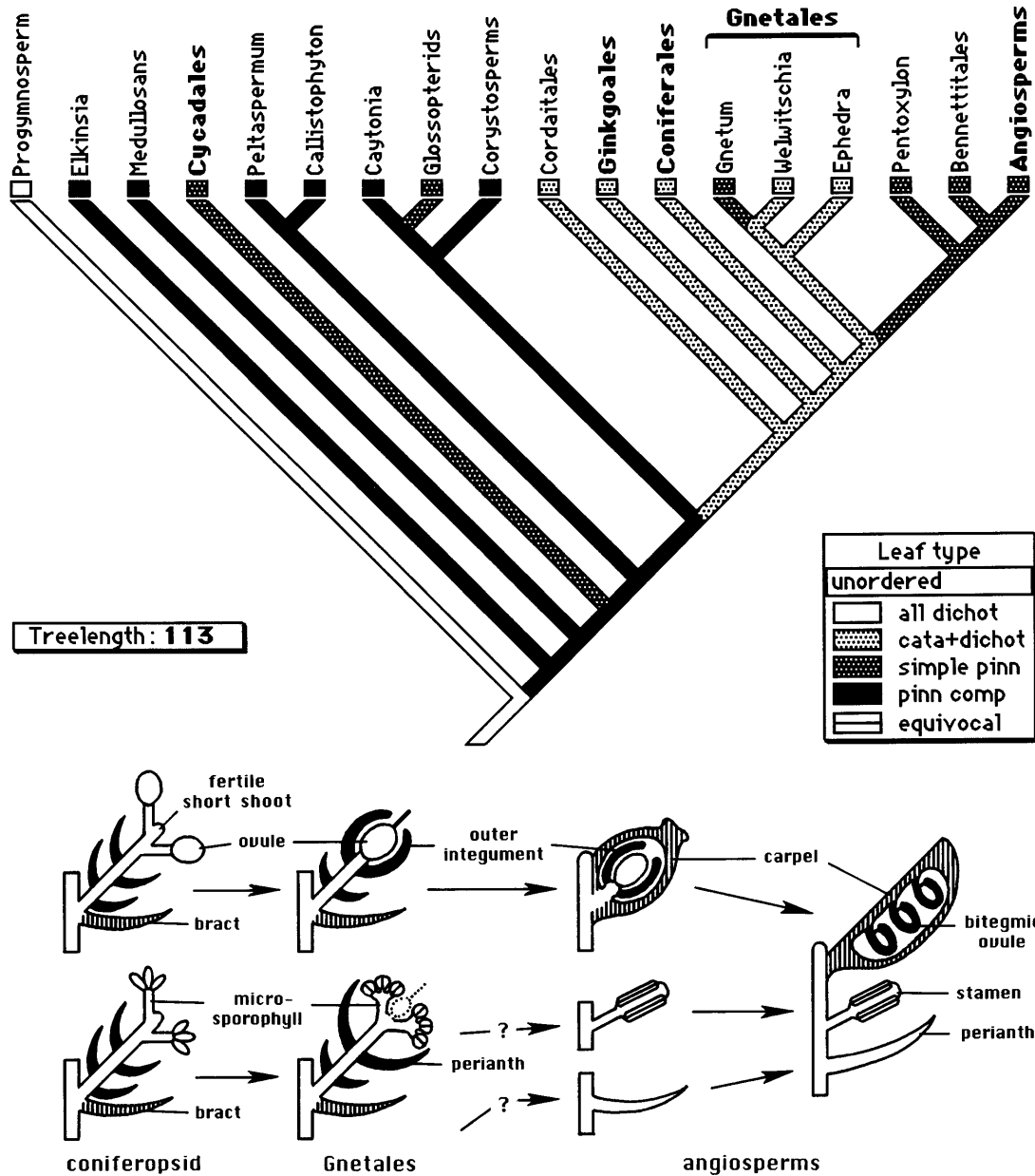
FIG. 2. Representative most parsimonious neo-englerian tree based on the data set of Doyle and Donoghue (1992), showing evolution of leaf morphology and a scenario for floral evolution (modified from Doyle, 1994).

Another weakness of these studies is that they treated angiosperms as a single taxon, which required debatable assumptions on basic states. For example, Doyle and Donoghue (1986, 1992) scored angiosperms as having spiral leaves and anatropous ovules. However, opposite leaves and orthotropous ovules occur in Chloranthaceae, which are among the oldest angiosperms in the Cretaceous fossil record (Crane *et al.*, 1995) and have been considered a link between angiosperms and Gnetales (Meeuse, 1972; Hickey and Taylor, 1996).

In an attempt to correct this problem, Doyle *et al.* (1994) included nine potentially basal angiosperm taxa, to which Doyle (1996) added two more and a Jurassic relative of Gnetales, *Piroconites* (Kirchner, 1992; van Konijnenburg-van Cittert, 1992), which had glossopterid-like reproductive structures. The two analyses gave similar results. In Doyle (1996), Gnetales were the closest living group to angiosperms, but *Caytonia* was located on the angiosperm line, rather than below the anthophytes, and in many trees glossopterids were the sister group of anthophytes (Fig. 3). This implies that the common ancestor of angiosperms and Gnetales had glossopterid-like ovulate structures, retained up to
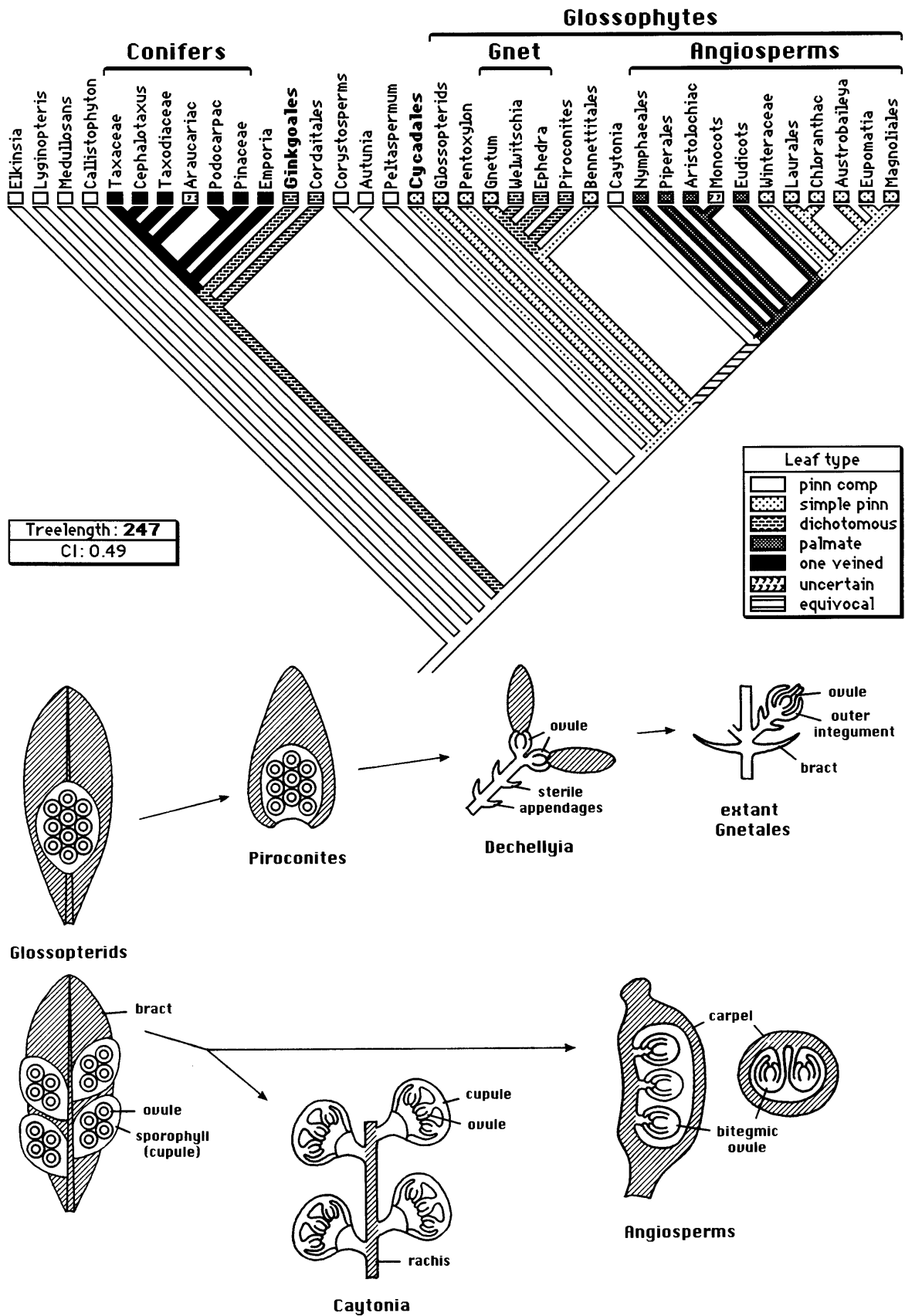
FIG. 3. Representative most parsimonious seed plant tree of Doyle (1996), showing evolution of leaf morphology and a scenario for evolution of the ovulate structures. GNET, Gnetales. *Dechellyia* is a Late Triassic fossil too incompletely known to be included in the analysis.

452

Piroconites on the line to Gnetales, and anatropous cupules arose on the *Caytonia*–angiosperm line, rather than being ancestral in anthophytes.

In contrast, in the analysis of Nixon *et al.* (1994), which included 18 angiosperms, angiosperms were nested within Gnetales, which were therefore paraphyletic rather than monophyletic (Fig. 4). Angiosperms were linked with *Gnetum* and *Welwitschia,* whereas *Ephedra* was one or two nodes lower. In most trees, anthophytes were linked with conifers, never with glossopterids and *Caytonia.* These results imply that angiosperms were derived from a gnetalian prototype—a neo-englerian scenario (cf. Fig. 2). Consistent with this, *Chloranthus* or *Casuarina* was basal in angiosperms, rather than Magnoliales or Nymphaeales (Donoghue and Doyle, 1989; Doyle *et al.,* 1994; Doyle, 1996). Angiosperms were also nested within Gnetales in a smaller analysis by Hickey and Taylor (1996), linked directly with *Gnetum.*

Phylogenetic Analyses of Molecular Data

In contrast, all major molecular analyses strongly support angiosperms and Gnetales as monophyletic groups, thus contradicting Nixon *et al.* (1994) and Hickey and Taylor (1996), but they disagree on whether the two groups are related at all. Simplified molecular trees are presented in Fig. 5, together with arrangements of extant taxa based on morphology (Figs. 5a–5d). I will pass over two early studies because of the short length of sequences and small number of taxa analyzed: Hori *et al.* (1985), based on 5S rRNA, which did not include Gnetales; and Troitsky *et al.* (1991), based on 5S, 5.8S, and short 18S sequences, which separated *Gnetum* and *Ephedra,* possibly because *Welwitschia* was not included.

In Hamby and Zimmer (1992), based on most of the 18S rRNA gene and smaller portions of 26S, angiosperms and Gnetales were both monophyletic, but their relationships were unstable. In a neighbor-joining analysis of 60 taxa, with seed plants rooted by *Psilotum* and *Equisetum,* Gnetales were linked with angiosperms, while cycads, *Ginkgo,* and conifers formed a clade at the base of seed plants (Fig. 5e). In a parsimony analysis the two angiosperm outgroups were reversed, so that Gnetales were basal in seed plants (Fig. 5f), but this was favored over the other arrangement by only one step. When 72 taxa were analyzed with parsimony (Doyle *et al.,* 1994), the relationship of angiosperms and Gnetales was favored. Both trees correspond to the same unrooted tree of seed plants; the difference is whether outgroups attach to the line to cycads, *Ginkgo,* and conifers or the line to Gnetales. There is reason to question any rooting based on living outgroups, which branched from seed plants before the Middle Devonian. In morphological analyses including fossils (Figs. 1–4), the crown-group (all derivatives of the most recent common ancestor of extant seed plants) was rooted
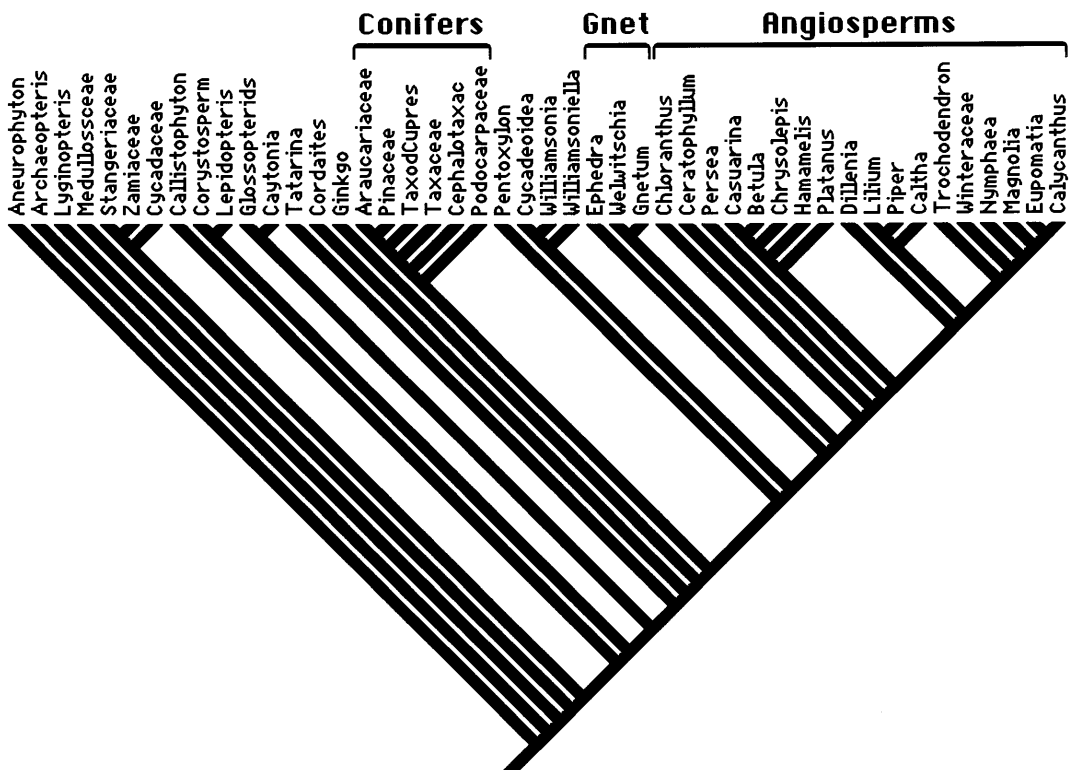

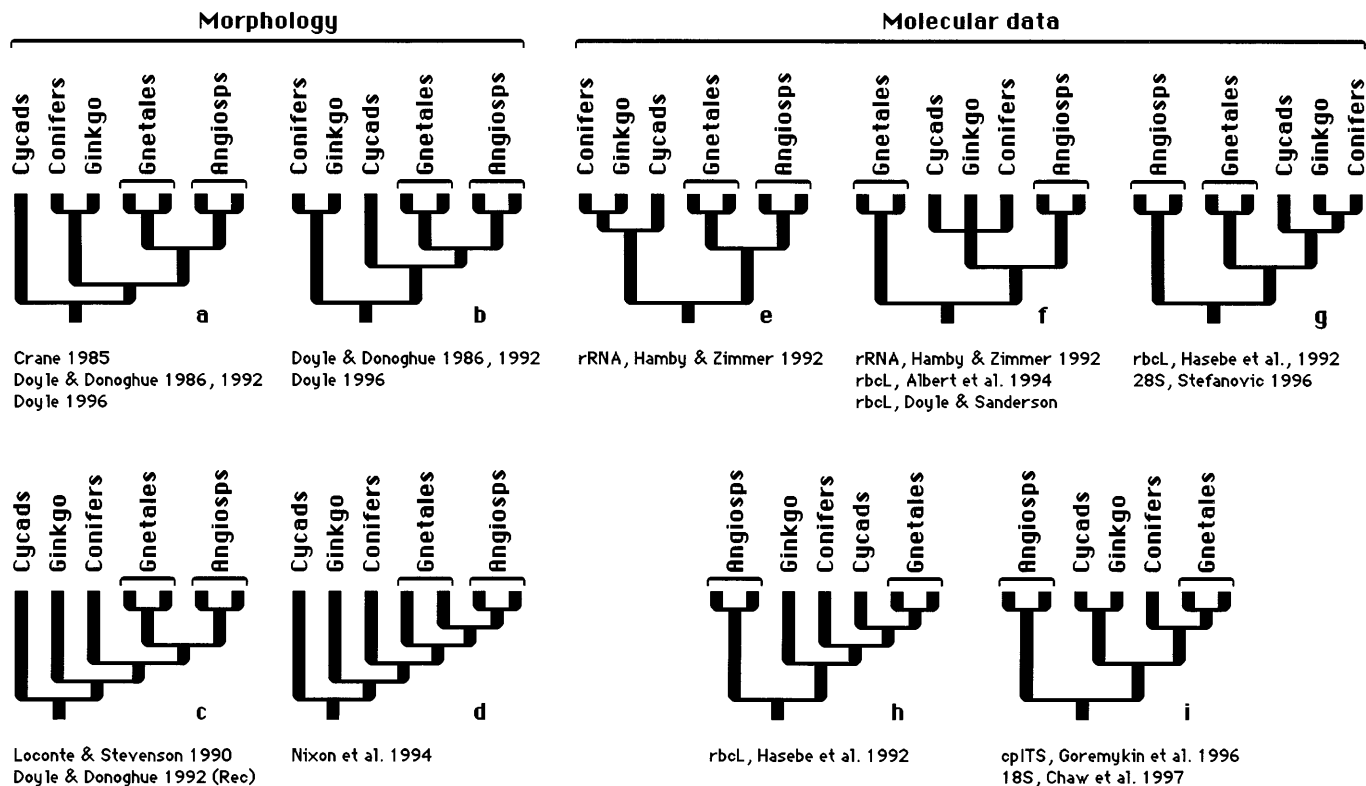**FIG. 4.** Representative seed plant tree of Nixon *et al.* (1994).

FIG. 5. Relationships of living groups found in morphological and molecular analyses of seed plants.

with Late Devonian–Carboniferous progymnosperms and seed ferns, which are closer than living outgroups, and the root was attached among cycads, *Ginkgo,* and conifers.

Doyle *et al.* (1994) performed bootstrap and decay analyses to evaluate the strength of the rRNA results. Because they questioned the rooting based on *Psilotum* and *Equisetum,* this was an unrooted analysis. Angiosperms and Gnetales were strongly supported as clades, at bootstrap percentages of 100 and 99%, respectively, and the two were linked at the 88% level. This does not directly evaluate whether angiosperms and Gnetales form a clade, but it does measure support for an unrooted tree in which the two taxa could be a clade under some rootings.

Trees with Gnetales basal in seed plants and angiosperms linked with cycads, *Ginkgo,* and conifers (Fig. 5f) were also found in analyses of the chloroplast gene *rbc*L by Albert *et al.* (1994) and Doyle and Sanderson (1997). In the latter analysis, bootstrap support for this rooting (i.e., for seed plants other than Gnetales) was 83%, but Gnetales and angiosperms were monophyletic at the 96 and 99% levels, respectively. However, Hasebe *et al.* (1992) found a different tree when they converted *rbc*L data into amino acid sequences and analyzed them by neighbor joining, with angiosperms the sister group of other seed plants (Fig. 5g). Gymnosperms therefore formed a clade, but its bootstrap support was

only 63%. Gnetales were basal in gymnosperms, so angiosperms and Gnetales could be linked if seed plants were rooted differently. Stefanovic (1996) found a similar tree based on 28S rDNA; again, bootstrap support for the gymnosperms was low, 66%.

In contrast, when Hasebe *et al.* (1992) analyzed their *rbc*L data with maximum likelihood, angiosperms were the sister group of gymnosperms but Gnetales were nested within the latter, linked with cycads (Fig. 5h). This tree cannot be converted into one with angiosperms and Gnetales linked by rerooting seed plants. Similar trees, but with Gnetales linked with conifers (Fig. 5i), were found by Goremykin *et al.* (1996), based on chloroplast rDNA ITS sequences, by Chaw *et al.* (1997), based on 18S rDNA, and by Bowe and DePamphilis (1997), based on the mitochondrial gene *coxI.* However, in Goremykin *et al.,* bootstrap support was only 53–58% for the gymnosperms and 50–56% for Gnetales plus conifers, in contrast to 90–92% for angiosperms and 99–100% for Gnetales. In Chaw *et al.* (1997), bootstrap support was 100% for both angiosperms and Gnetales, but only 56–75% for gymnosperms. Support for Gnetales plus conifers was higher, 84%, but it should be noted that this is less than the 88% value for angiosperms plus Gnetales in the rRNA analysis of Doyle *et al.* (1994). Chaw *et al.* (1997) cited absences of angiosperm indels as evidence for the monophyly of gymnosperms, but such absences are

presumably symplesiomorphies that occurred in the common ancestor of all seed plants and are therefore irrelevant to the status of gymnosperms.

DISCUSSION

As summarized in Fig. 5, all morphological analyses (Figs. 5a–5d) agree that angiosperms and Gnetales are related, but some indicate that both groups are monophyletic (Figs. 5a–5c), while others indicate that Gnetales are paraphyletic and angiosperms are nested within them (Fig. 5d). The relationships of other seed plants are poorly resolved, but they always form a paraphyletic basal series. Molecular studies (Figs. 5e–5i) all strongly support the monophyly of both angiosperms and Gnetales. Some indicate that the two groups are related (Fig. 5e), but most imply that they are not. In some analyses (Figs. 5f and 5g), angiosperms and Gnetales could be related if seed plants were rooted differently, but in others (Figs. 5h and 5i), they cannot be related under any rooting. However, support for these arrangements is weak.

In the following sections, I will adduce evidence that the molecular results are more credible in cases where their internal support is strong—that angiosperms and Gnetales are both monophyletic—and the morphological results where these are strongly supported—that the two groups are related. It should be noted that the disagreements among analyses of the two kinds are somewhat different. With the molecular studies, many of the conflicts are between analyses of different genes (although some involve method of analysis or taxon sampling), whereas with the morphological analyses the conflicts are between data sets made up of overlapping characters, which, however, have been interpreted differently. The fact that molecular systematics involves fewer problems of character analysis (except sometimes in sequence alignment) is one of its great strengths. The conflicts among morphological analyses might thus seem to confirm the view that morphological data are hopelessly subjective and can be reinterpreted to support any hypothesis one likes. Part of the following discussion is therefore a defense of morphology against this charge: it is often possible to decide objectively between interpretations of characters. In addition, I will argue that fossils may provide independent criteria for choosing among molecular hypotheses.

Conflicts among Morphological Analyses

With morphological data, the main disagreement on relations of living taxa concerns whether Gnetales are monophyletic or paraphyletic, as in Nixon et al. (1994) and Hickey and Taylor (1996). One could simply claim that the molecular data refute these hypotheses, and this might be a valid position. However, the same conclusion can be reached by closer examination of the relevant morphological characters. I will concentrate

on the analysis of Nixon et al.; that of Hickey and Taylor was based on too few taxa and character definitions too heavily linked with one evolutionary scenario.

Even without considering individual characters, there is evidence that support for the paraphyly of Gnetales in the Nixon et al. (1994) data set is weak. In an analysis of this data set, Donoghue, Doyle, and Friedman (cited in Doyle, 1996) found that bootstrap support for the angiosperm–Gnetum–Welwitschia clade was only 54%. Furthermore, when Albert et al. (1994) combined the Nixon et al. data set with rbcL data for extant taxa, they obtained trees in which Gnetales were monophyletic. An even more intriguing result was that anthophytes were linked with Caytonia and glossopterids rather than with conifers. This may seem paradoxical, since there are no molecular data on Caytonia. However, it can be explained as a result of strong molecular support for Gnetales, weak morphological support for the angiosperm–Gnetum–Welwitschia clade, and the fact that angiosperms share morphological features with Caytonia, even in the Nixon et al. data set. When angiosperms are nested within Gnetales, the latter features have no effect on the position of anthophytes as a whole. But when the rbcL data pull Gnetales together as a clade, angiosperms go to the base of anthophytes, and their features draw the whole group together with Caytonia.

The key characters responsible for the result of Nixon et al. (1994) are the five synapomorphies that link angiosperms with Gnetum and Welwitschia, to the exclusion of Ephedra. Consideration of these characters (documented in Doyle, 1996) may illustrate some general principles about pitfalls in the use of morphological characters and ways to avoid them. These examples emphasize the necessity of considering structures in their positional and developmental contexts.

One character is cellularization of the female gametophyte (in the ovule). The basic condition in seed plants is alveolar: gametophyte ontogeny begins with a phase of free-nuclear divisions and then cell walls form in a honeycomb-like pattern around individual nuclei. However, Gnetum and Welwitschia show a nonalveolar pattern, in which cell walls form irregularly around several nuclei. In angiosperms, the female gametophyte is reduced to an egg cell and two synergids at the micropylar end, three antipodal cells at the other end, and two free polar nuclei in the middle. Nixon et al. (1994) scored angiosperms like Gnetum and Welwitschia, presumably because the middle of the gametophyte remains free-nuclear. However, the egg, synergids, and antipodals of angiosperms are all uninucleate cells, and in this sense they are more like other seed plants. One solution would be to score angiosperms as unknown. However, it is also possible to redefine the character in terms of a related but less ambiguous distinction: cellularization resulting in uninucleate vs multinucleate cells. Under this definition, angiosperms

can be unambiguously scored as having the basic state and *Gnetum* and *Welwitschia* the derived state. This suggests a general principle: when choosing among alternative distinctions between character states, one should define states such that critical taxa can be scored unambiguously.

A related character is the presence or absence of archegonia, the original egg-containing organ in land plants. *Ephedra* has archegonia, with an egg and neck cells, but they are absent in angiosperms, *Gnetum,* and *Welwitschia.* This may seem like a clear, objective distinction, but closer examination shows that the conditions in angiosperms and the two gnetalian genera are very different. In *Gnetum,* the micropylar end of the gametophyte remains free-nuclear, and free nuclei function as eggs (Martens, 1971; Carmichael and Friedman, 1996); in *Welwitschia,* the whole gametophyte becomes cellularized, but individual nuclei of multinucleate cells act as eggs (Carmichael and Friedman, personal communication). In contrast, although angiosperms have no neck cells and thus no archegonia, the egg is a normal, uninucleate cell. I see no basis for equating these two states or for considering one a modification of the other; in fact, one might argue that they resulted from different modifications of the basic gametophyte ontogeny. In the face of such divergent scenarios for character evolution, bias in favor of one scenario over another can be avoided by treating the alternative conditions (three in this case) as states of an unordered multistate character.

Similar problems concern embryogeny, which passes through an initial free-nuclear phase in cycads, *Ginkgo,* and conifers but proceeds entirely by cellular divisions in angiosperms, *Gnetum,* and *Welwitschia.* The critical taxon is *Ephedra,* which Nixon *et al.* (1994) and previous workers (Crane, 1985; Doyle and Donoghue, 1986, 1992) interpreted as free-nuclear, because the two zygotes divide to produce eight free nuclei. However, as emphasized by Friedman (1994), the following stages are more like those in angiosperms, *Gnetum,* and *Welwitschia,* in that each of the eight free nuclei (now surrounded by cell walls) produces a separate embryo by cellular divisions, whereas in cycads, *Ginkgo,* and conifers many free nuclei contribute to one embryo. This suggests that the occurrence of two free-nuclear divisions in *Ephedra* is a secondary advance. Whether or not this hypothesis is correct, it casts doubt on the scoring of *Ephedra.* Again, this character can be replaced with a related one, whether each embryo is derived from several free nuclei or from a single cell by cellular divisions, for which *Ephedra* can be unambiguously scored as having the latter state.

A final case involves redundancy of two characters used by Nixon *et al.* (1994): number of nuclei in the male gametophyte (in the pollen) and presence or absence of a stalk cell. In the basic male gametophyte ontogeny in seed plants, successive divisions produce three or more sterile nuclei, the last called the stalk cell, and two sperm, but in *Gnetum* and *Welwitschia* the number of nuclei is reduced to four and in angiosperms to three. However, reduction from five to four nuclei in *Gnetum* and *Welwitschia* occurs by loss of the stalk cell, so the changes in the two characters are equivalent. The stalk cell would also be lost in angiosperms during reduction to three nuclei, whether this occurred by the same or another route. Treating the stalk cell as a separate character therefore counts the same evolutionary event twice; on the general principle that such weighting should be avoided, this character should be eliminated.

When these characters are redefined in a more neutral manner (Doyle, 1996) and the rest of the Nixon *et al.* (1994) data set is left unchanged, one obtains trees in which Gnetales and angiosperms are both monophyletic. The bootstrap support for Gnetales is only moderate (75%), but this is as high as for many groups whose monophyly has rarely been questioned (e.g., cycads).

Conflicts among Molecular Analyses

In discussing the results of molecular analyses, I noted reasons to doubt the rootings obtained for seed plants as a whole. Seed plants are a good example of a radiation that occurred in a relatively short period a long time ago, a situation in which molecular data are least likely to be reliable, because of the small number of molecular synapomorphies between nodes, the possibility that these were erased by later changes (multiple hits), and the related problem of long branch attraction (Felsenstein, 1978; Donoghue *et al.,* 1989; Donoghue and Sanderson, 1992).

This point is illustrated by Fig. 6, the tree in Fig. 3 (Doyle, 1996) with conifers and angiosperms reduced to single taxa, plotted against geologic time. Black lines show the known ranges of each group, white lines "ghost lineages" (Norell, 1992), where the cladogram predicts that the line existed because its sister group is already known in the fossil record, but the line itself is not. Ghost lineages have been extended downward the least possible distance (minimum implied gap: Benton and Storrs, 1994), resulting in several apparent polychotomies. Four living seed plant lines separated in the Carboniferous; plants on their stem-lineages occur in the Late Carboniferous or Permian, and most if not all of the many seed plants known from the Early Carboniferous are more primitive than their reconstructed common ancestor (e.g., lacking such crown-group apomorphies as endarch primary xylem, pollen with honeycomb-type exine structure and air sacs, and bilateral ovule symmetry). Although the first records of the two anthophyte lines are Late Triassic, it is possible that they diverged in the Permian. If glossopterids are paraphyletic (they have only two questionable autapomorphies, saccate and striate pollen), the ghost lin-
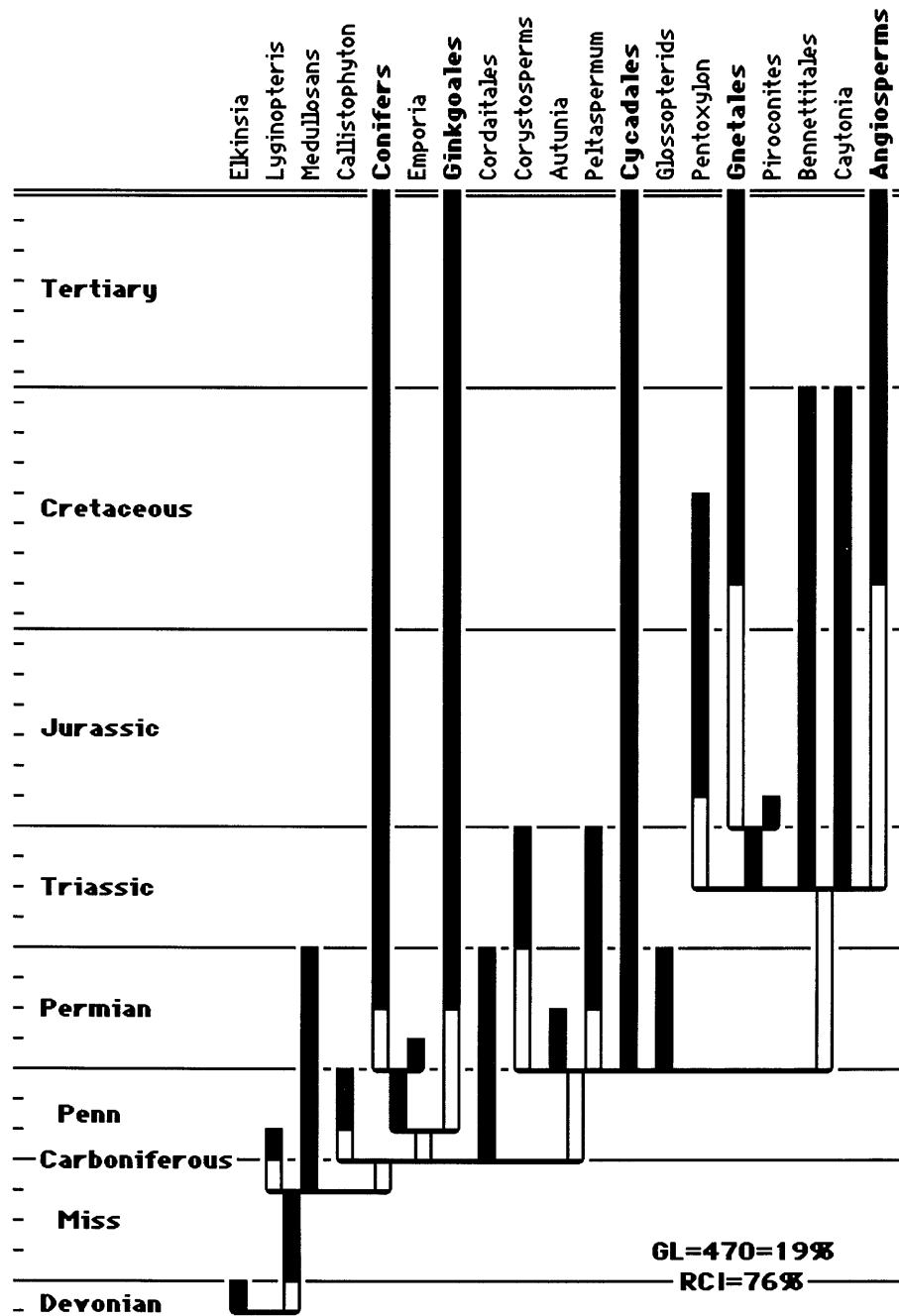
FIG. 6. Most parsimonious seed plant tree of Doyle (1996) (Fig. 3) plotted against geologic time (Palmer, 1983), approximated to 10-million-year intervals, with known ranges, ghost lineages implied by the tree, and measures of consistency with the stratigraphic record (see text for discussion). Ranges based on Stewart and Rothwell (1993) and other general references.

eages leading to both lines may be represented in the Permian by plants known as glossopterids (Doyle, 1996).

Other insights may be gained by considering how unparsimonious trees found with molecular data are in terms of a morphological data set (Doyle, 1996). Figure 7 shows one of two trees found after forcing angiosperms to be the sister group of other living seed plants

(by adding 10 characters, with angiosperms and *Elkinsia* scored as 0, other living seed plants as 1, and fossils as unknown, then subtracting 10 steps from the resulting trees). Contrary to Troitsky *et al.* (1991), this arrangement does not mean that angiosperms and gymnosperms were independently derived from Devonian progymnosperms. Because of the stem, pollen, and ovule advances shared by living seed plants, the angio-
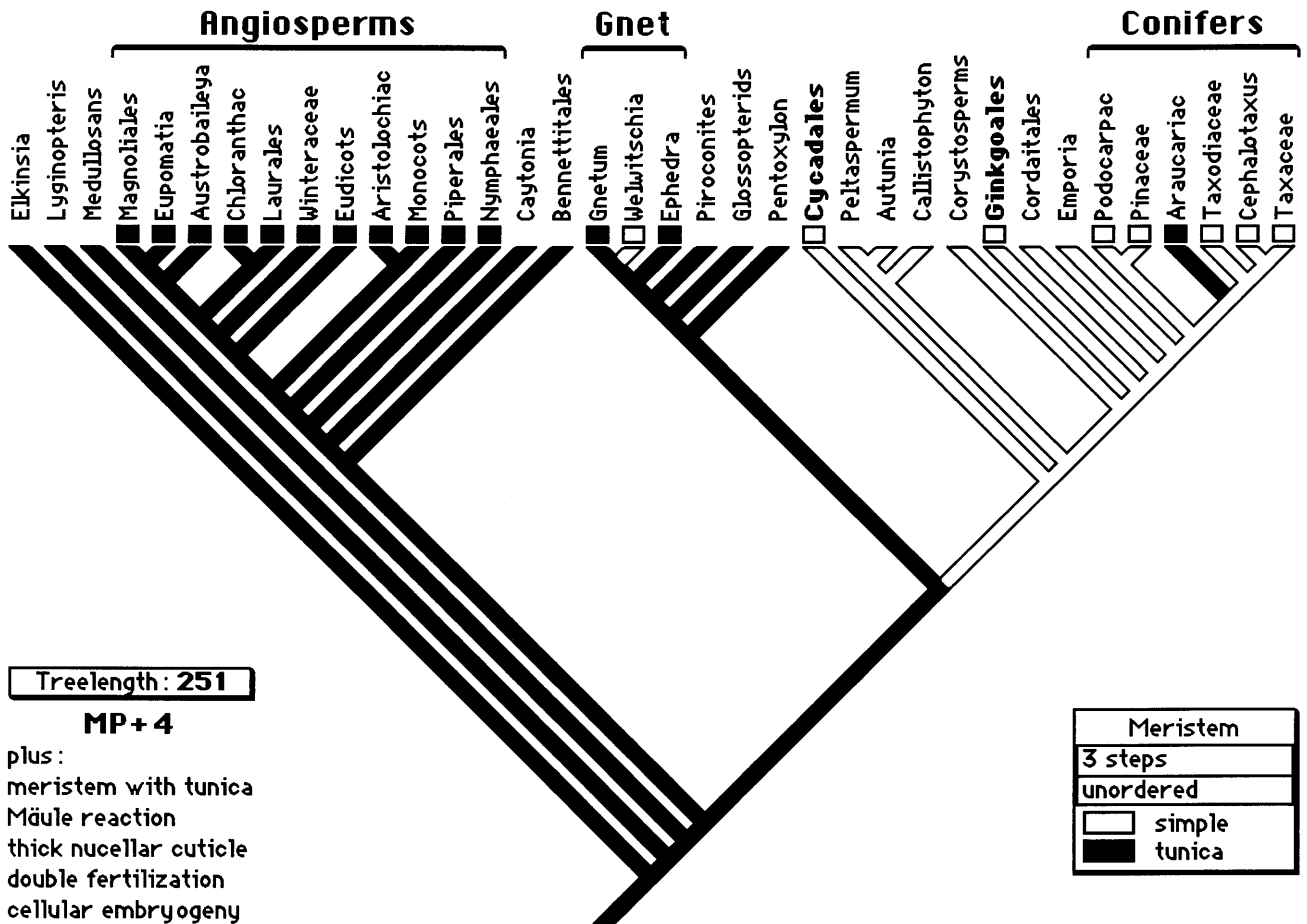
FIG. 7. Seed plant tree with angiosperms forced below other living seed plants, as in some molecular analyses, based on the data set of Doyle (1996), showing evolution of the apical meristem. MP, most parsimonious; below are listed five characters not documented in fossils that the tree implies are ancestral in seed plants but are more likely to be derived based on more distant outgroups or other evidence, implying additional steps (see text for discussion).

sperm line splits off above medullosans, still linked with Bennettitales and *Caytonia.* The next branch leads to Gnetales, *Pentoxylon,* glossopterids, and *Piroconites;* the last to cycads, *Ginkgo,* conifers, and various fossils. In effect, this tree reroots the seed plant crown-group with angiosperms and their fossil relatives at the base.

Based on the present data set, this tree is 4 steps longer than the most parsimonious trees, but it may actually be even less parsimonious, because angiosperms and Gnetales have five other character states, listed in Fig. 7, that are reconstructed as ancestral but are more likely derived. Because the basal taxa are fossils in which these characters are unknown, it is most parsimonious to assume that they had the state found in angiosperms and Gnetales. However, more distant outgroups or other data suggest that the earliest seed plants actually had the alternative state. Each character may therefore involve a hidden extra step.

One of these characters is plotted in Fig. 7, presence of a tunica in the apical meristem. This is an outer cell layer that produces only epidermal tissue, whereas in other groups the surface cells produce both epidermal and internal tissues. The latter condition is more like that in lower vascular plants, which have one apical cell that produces both outer and inner tissues. Similarly, the Mäule reaction (or the related abundance of syringaldehyde as a lignin oxidation product) is absent in most lower vascular plants (except *Selaginella, Equisetum,* and *Dennstaedtia;* Gibbs, 1957; Logan and Thomas, 1985). Logan and Thomas (1985) linked abundance of syringaldehyde to presence of fibers, which are lacking in the wood of early seed plants. Thickness of the nucellar cuticle is poorly documented in Paleozoic seeds, but the thickened condition of *Caytonia* and most anthophytes was first noted by Harris (1954) in Mesozoic seeds. It is unlikely that the first seed plants had double fertilization, which is associated with persistence of a nucleus next to the egg (Donoghue and Scheiner, 1992); in lower vascular plants, the adjacent neck and ventral canal cells degenerate before fertilization, as in cycads, *Ginkgo,* and conifers. Lower vascular
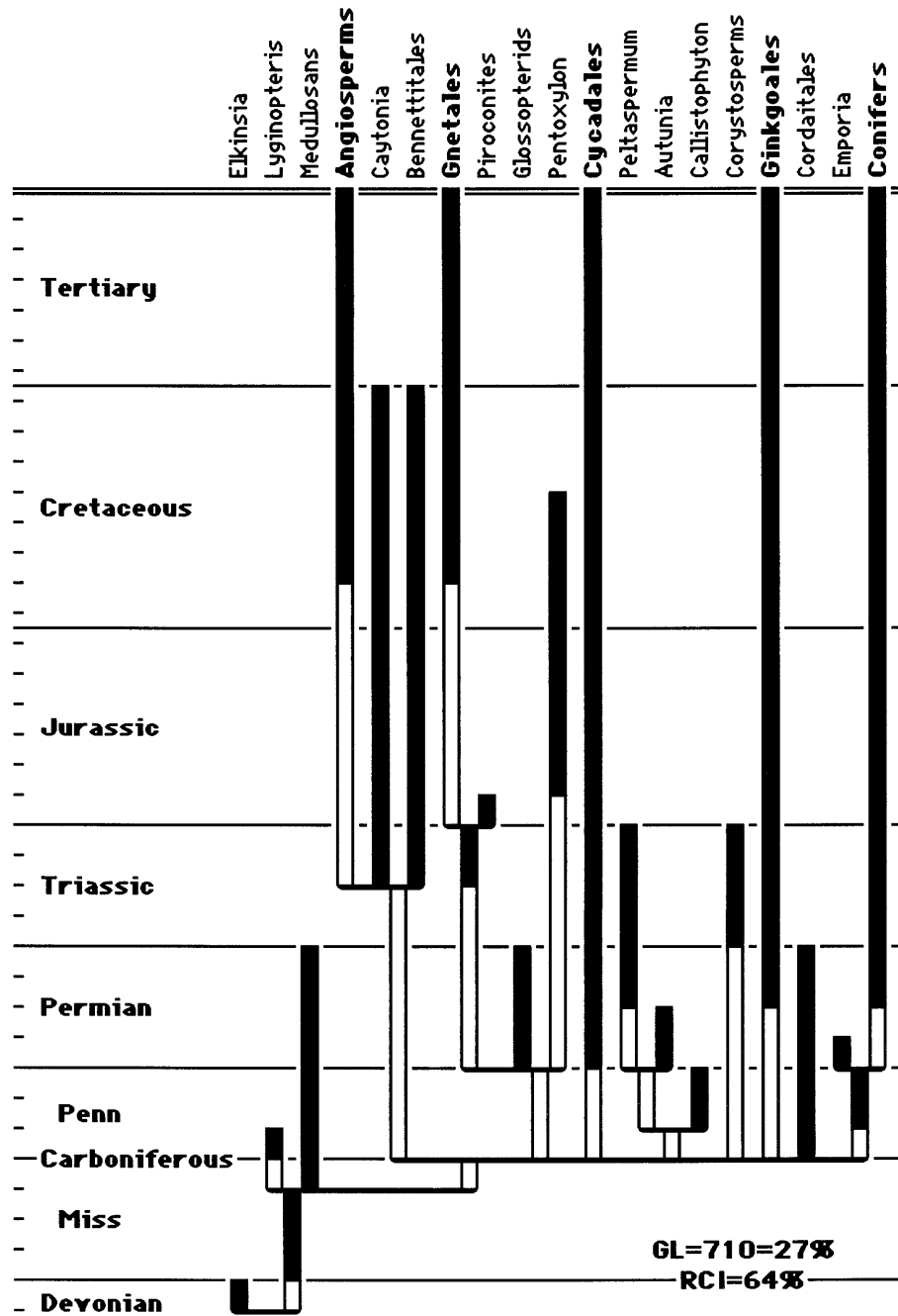
FIG. 8. Seed plant tree with angiosperms forced below other living groups of seed plants (Fig. 7), plotted against geologic time, with known ranges, ghost lineages implied by the tree, and measures of consistency with the stratigraphic record (see text for discussion).

plants are like angiosperms and Gnetales in having cellular embryogeny, but early seed plants had large egg cells (Stewart and Rothwell, 1993), which are correlated with free-nuclear embryogeny, presumably because it is difficult to partition a large cell.

Gnetales are not associated with conifers in Fig. 7, as they were in cpITS, 18S, and *coxI* trees (Goremykin *et al.,* 1996; Chaw *et al.,* 1997; Bowe and DePamphilis, 1977; Fig. 5i). Apparently, even when angiosperms are

assumed to be basal in seed plants, there is more morphological support for an arrangement in which angiosperms and Gnetales form adjacent lines than there is for a link between Gnetales and conifers. Trees with angiosperms basal and Gnetales (plus *Piroconites* and Bennettitales) linked with conifers are 10 steps longer than the shortest trees. Trees with Gnetales basal in living seed plants, as in *rbc*L analyses (Albert *et al.,* 1994; Doyle and Sanderson, 1977; Fig. 5f), are 5

steps longer than the shortest trees, not counting the same five characters that probably involve additional cryptic steps.

Molecular trees that disagree with morphology can also be questioned because they conflict with the stratigraphic record. This is illustrated by two trees plotted against geologic time, with known ranges and ghost lineages: Fig. 6, a morphological tree; and Fig. 8, the tree in Fig. 7 with angiosperms basal, as in molecular trees of the type in Fig. 5g. The percentage of the total tree length consisting of ghost lineages is an inverse measure of consistency with the stratigraphic record (Smith and Littlewood, 1994). A direct measure is the relative completeness index (RCI), or 1 minus the ratio of ghost lineages to observed ranges, expressed as a percentage (Benton and Storrs, 1994). In the morphological tree (Fig. 6), the taxa that split off first are the oldest Late Devonian and Carboniferous seed ferns, followed by Permian and then Mesozoic groups. This implies 470 My of ghost lineages and 1960 My of observed ranges, or 19% ghost lineages and RCI = 76%. However, when angiosperms are basal (Fig. 8), the sequence is from the oldest to the youngest groups of seed plants and then back to taxa of intermediate age.

As a result, this tree implies 710 My of ghost lineages, or 27% ghost lineages and RCI = 64%. The numbers are identical when Gnetales are basal among living seed plants (Fig. 5f). These contrasts are all the more significant because most of the increase in ghost lineages involves extension of the lines leading to angiosperms and Gnetales down into the Carboniferous, which probably has the best plant record of any period, whereas the Late Permian and Early Triassic gaps seen in both trees fall in the most poorly known interval.

The reasons that molecular data may give incorrect relationships of angiosperms and Gnetales are suggested by a study on *rbc*L and the age of angiosperms by Doyle and Sanderson (1997), which indicates that evolution of this gene is far from clocklike. Assuming that hepatics split from other land plants in the Late Ordovician, the shortest tree, with Gnetales basal in seed plants (Fig. 5f), implies that conifers and cycads diverged in the Middle Jurassic, 130 million years after the first records of both lines (Middle Pennsylvanian and Early Permian, respectively). Figure 9 shows an *rbc*L tree found after imposing five constraints: *Lycopodium* basal in vascular plants, as inferred from fossil and molecular data (Raubeson and Jansen, 1992; Ken-
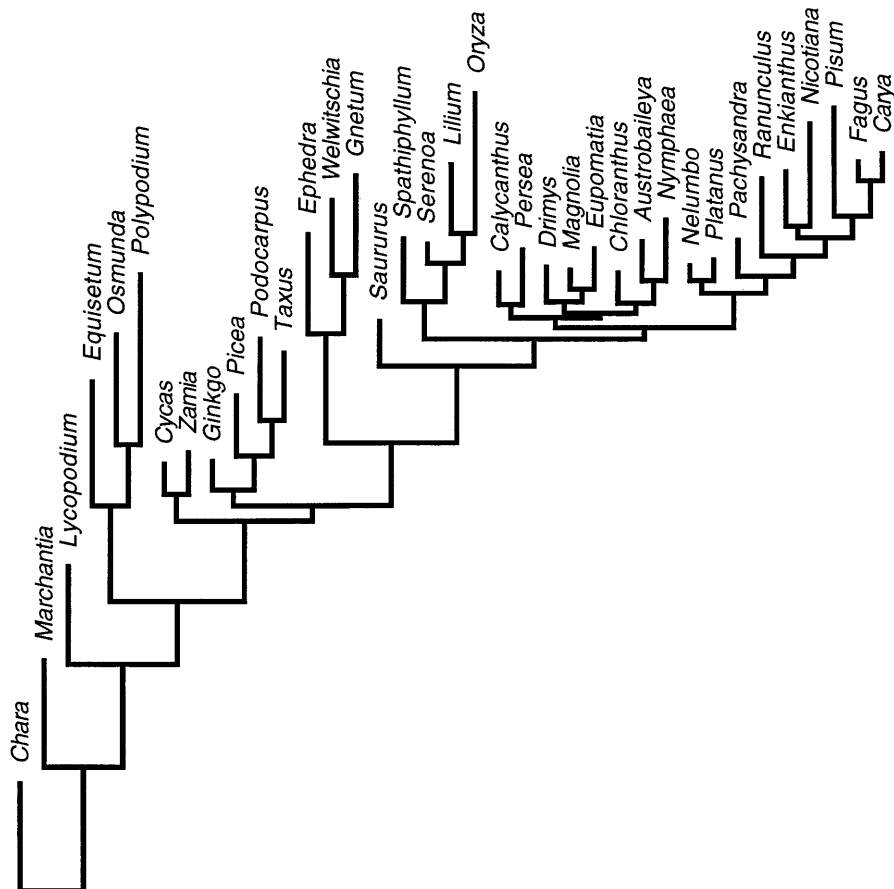

FIG. 9. Phylogram of seed plants based on *rbc*L (Doyle and Sanderson, 1997), with five topological constraints for greater congruence with other evidence (see text). Cycads, *Ginkgo,* and conifers are short branches relative to angiosperms and Gnetales.

rick and Crane, 1997); cycads, *Ginkgo,* and conifers as a paraphyletic grade and Gnetales linked with angiosperms, as in morphological analyses; and monocots and eudicots as clades, as indicated by larger *rbc*L analyses (Chase *et al.,* 1993; otherwise angiosperms are rooted next to *Oryza* or *Pisum*). Figure 9 is a phylogram, so branch lengths correspond to the amount of molecular evolution. Conifers and especially cycads and *Ginkgo* are very short branches relative to both lower groups and angiosperms and Gnetales, while within angiosperms some herbaceous lines (*Oryza, Pisum, Nicotiana*) are especially long (as noted by Bousquet *et al.,* 1992). These results can be explained if *rbc*L evolution slowed going from early vascular plants to early seed plants, but then accelerated again in angiosperms and Gnetales and still more in herbaceous angiosperms. As a result, angiosperms or Gnetales may be drawn to the base of seed plants by long branch effects, specifically reversals to outgroup states on the long lines leading to these taxa.

CONCLUSIONS

These exercises support the view that morphological and molecular data may complement each other in different parts of a phylogeny: morphology resolves conflicts among molecular analyses by indicating that angiosperms and Gnetales are related, whereas molecular data resolve morphological conflicts by implying that both groups are monophyletic. It is possible that the relationship of angiosperms and Gnetales will eventually be resolved using molecular data alone: combining data from many genes, discovery of structural changes, or improvements in methods of analysis that overcome the problem of long branch attraction. However, morphological analysis will still be needed in order to integrate fossils into the tree of living taxa, which is essential for understanding how the living groups and their structures originated.

ACKNOWLEDGMENTS

REFERENCES

Albert, V. A., Backlund, A., Bremer, K., Chase, M. W., Manhart, J. R., Mishler, B. D., and Nixon, K. C. (1994). Functional constraints and *rbc*L evidence for land plant phylogeny. *Ann. Mo. Bot. Gard.* **81:** 534–567.

Arber, E. A. N., and Parkin, J. (1907). On the origin of angiosperms. *J. Linn. Soc. London Bot.* **38:** 29–80.

Arber, E. A. N., and Parkin, J. (1908). Studies on the evolution of the angiosperms: The relationship of the angiosperms to the Gnetales. *Ann. Bot. London* **22:** 489–515.

Benton, M. J., and Storrs, G. W. (1994). Testing the quality of the fossil record: Paleontological knowledge is improving. *Geology* **22:** 111–114.

Bousquet, J., Strauss, S. H., Doerksen, A. H., and Price, R. A. (1992). Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. *Proc. Natl. Acad. Sci. USA* **89:** 7844–7848.

Bowe, L. M., and DePamphilis, C. W. (1997). Conflict and congruence among three plant genetic compartments: Phylogenetic analysis of seed plants from gene sequences of *coxI, rbc*L and 18S rDNA. *Am. J. Bot.* **84**(6, Suppl.): 178–179.

Carmichael, J. S., and Friedman, W. E. (1996). Double fertilization in *Gnetum gnemon* (Gnetaceae): Its bearing on the evolution of sexual reproduction within the Gnetales and the anthophyte clade. *Am. J. Bot.* **83:** 767–780.

Chase, M. W., and 41 others (1993). Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. *Ann. Mo. Bot. Gard.* **80:** 526–580.

Chaw, S.-M., Zharkikh, A., Sung, H.-M., Lau, T.-C., and Li, W.-H. (1997). Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nuclear 18S rRNA sequences. *Mol. Biol. Evol.* **14:** 56–68.

Crane, P. R. (1985). Phylogenetic analysis of seed plants and the origin of angiosperms. *Ann. Mo. Bot. Gard.* **72:** 716–793.

Crane, P. R., Friis, E. M., and Pedersen, K. R. (1995). The origin and early diversification of angiosperms. *Nature* **374:** 27–33.

Donoghue, M. J., and Doyle, J. A. (1989). Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. *In* "Evolution, Systematics, and Fossil History of the Hamamelidae" (P. R. Crane and S. Blackmore, Eds.), Vol. 1, pp. 17–45. Clarendon, Oxford.

Donoghue, M. J., Doyle, J. A., Gauthier, J., Kluge, A. G., and Rowe, T. (1989). The importance of fossils in phylogeny reconstruction. *Annu. Rev. Ecol. Syst.* **20:** 431–460.

Donoghue, M. J., and Sanderson, M. J. (1992). The suitability of molecular and morphological evidence in reconstructing plant phylogeny. *In* "Molecular Systematics of Plants" (P. S. Soltis, D. E. Soltis, and J. J. Doyle, Eds.), pp. 340–368. Chapman and Hall, New York.

Donoghue, M. J., and Scheiner, S. M. (1992). The evolution of endosperm: A phylogenetic account. *In* "Ecology and Evolution of Plant Reproduction" (R. E. Wyatt, Ed.), pp. 356–389. Chapman and Hall, New York.

Doyle, J. A. (1978). Origin of angiosperms. *Annu. Rev. Ecol. Syst.* **9:** 365–392.

Doyle, J. A. (1994). Origin of the angiosperm flower: A phylogenetic perspective. *Plant Syst. Evol. Suppl.* **8:** 7–29.

Doyle, J. A. (1996). Seed plant phylogeny and the relationships of Gnetales. *Int. J. Plant Sci.* **157**(6, Suppl.): S3–S39.

Doyle, J. A., and Donoghue, M. J. (1986). Seed plant phylogeny and the origin of angiosperms: An experimental cladistic approach. *Bot. Rev.* **52:** 321–431.

Doyle, J. A., and Donoghue, M. J. (1987). The importance of fossils in elucidating seed plant phylogeny and macroevolution. *Rev. Palaeobot. Palynol.* **50:** 63–95.

Doyle, J. A., and Donoghue, M. J. (1992). Fossils and seed plant phylogeny reanalyzed. *Brittonia* **44:** 89–106.

Doyle, J. A., Donoghue, M. J., and Zimmer, E. A. (1994). Integration of morphological and ribosomal RNA data on the origin of angiosperms. *Ann. Mo. Bot. Gard.* **81:** 419–450.

Doyle, J. A., and Sanderson, M. J. (1997). Fossils, molecular clocks, and the age of angiosperms. *Am. J. Bot.* **84**(6, Suppl.): 132.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27:** 401–410.

Friedman, W. E. (1990). Sexual reproduction in *Ephedra nevadensis* (Ephedraceae): Further evidence of double fertilization in a nonflowering seed plant. *Am. J. Bot.* **77:** 1582–1598.

Friedman, W. E. (1992). Evidence of a pre-angiosperm origin of endosperm: Implications for the evolution of flowering plants. *Science* **255:** 336–339.

Friedman, W. E. (1994). The evolution of embryogeny in seed plants and the developmental origin and early history of endosperm. *Am. J. Bot.* **81:** 1468–1486.

Friedman, W. E. (Ed.) (1996). Biology and evolution of the Gnetales. *Int. J. Plant Sci.* **157**(6, Suppl.): S1–S125.

Gaussen, H. (1946). Les Gymnospermes, actuelles et fossiles. *Trav. Lab. Forest. Toulouse,* Tome II Etud. Dendrol., Sect. 1, Vol. 1, Fasc. 3, Chap. 5, pp. 1–26.

Gibbs, R. D. (1957). The Mäule reaction, lignin, and the relationships between woody plants. *In* "The Physiology of Forest Trees" (K. V. Thimann, Ed.), pp. 269–312. Ronald Press, New York.

Goremykin, V., Bobrova, V., Pahnke, J., Troitsky, A., Antonov, A., and Martin, W. (1996). Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbc*L data do not support gnetalean affinities of angiosperms. *Mol. Biol. Evol.* **13:** 383–396.

Hamby, R. K., and Zimmer, E. A. (1992). Ribosomal RNA as a phylogenetic tool in plant systematics. *In* "Molecular Systematics of Plants" (P. S. Soltis, D. E. Soltis, and J. J. Doyle, Eds.), pp. 50–91. Chapman and Hall, New York.

Harris, T. M. (1954). Mesozoic seed cuticles. *Sven. Bot. Tidskr.* **48:** 281–291.

Hasebe, M., Kofuji, R., Ito, M., Kato, M., Iwatsuki, K., and Ueda, K. (1992). Phylogeny of gymnosperms inferred from *rbc*L gene sequences. *Bot. Mag.* **105:** 673–679.

Hickey, L. J., and Taylor, D. W. (1996). Origin of the angiosperm flower. *In* "Flowering Plant Origin, Evolution and Phylogeny" (D. W. Taylor and L. J. Hickey, Eds.), pp. 176–231. Chapman and Hall, New York.

Hori, H., Lim, B.-L., and Osawa, S. (1985). Evolution of green plants as deduced from 5S rRNA sequences. *Proc. Natl. Acad. Sci. USA* **82:** 820–823.

Kenrick, P., and Crane, P. R. (1997). "The Origin and Early Diversification of Land Plants: A Cladistic Study," Smithsonian Institution, Washington, DC.

Kirchner, M. (1992). Untersuchungen an einigen Gymnospermen der fränkischen Rhät-Lias-Grenzschichten. *Palaeontographica Abt. B* **224:** 17–61.

Loconte, H., and Stevenson, D. W. (1990). Cladistics of the Spermatophyta. *Brittonia* **42:** 197–211.

Logan, K. J., and Thomas, B. A. (1985). Distribution of lignin derivatives in plants. *New Phytol.* **99:** 571–585.

Martens, P. (1971). "Les Gnétophytes," Encyclopedia of Plant Anatomy, Vol. 12, No. 2, Borntraeger, Stuttgart.

Meeuse, A. D. J. (1972). Taxonomic affinities between Piperales and Polycarpicae and their implications in interpretative floral morphology. *In* "Advances in Plant Morphology" (Y. S. Murty, B. M. Johri, H. Y. Mohan Ram, and T. M. Varghese, Eds.), pp. 3–27. Sarita Prakashan, Meerut, India.

Nixon, K. C., Crepet, W. L., Stevenson, D., and Friis, E. M. (1994). A reevaluation of seed plant phylogeny. *Ann. Mo. Bot. Gard.* **81:** 484–533.

Norell, M. A. (1992). Taxic origin and temporal diversity: The effect of phylogeny. *In* "Extinction and Phylogeny" (M. J. Novacek and Q. D. Wheeler, Eds.), pp. 88–118. Columbia Univ. Press, New York.

Palmer, A. R. (1983). The Decade of North American Geology 1983 geologic time scale. *Geology* **11:** 503–504.

Raubeson, L. A., and Jansen, R. K. (1992). Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* **255:** 1697–1699.

Retallack, G., and Dilcher, D. L. (1981). Arguments for a glossopterid ancestry of angiosperms. *Paleobiology* **7:** 54–67.

Rothwell, G. R., and Serbet, R. (1994). Lignophyte phylogeny and the evolution of spermatophytes: A numerical cladistic analysis. *Syst. Bot.* **19:** 443–482.

Smith, A. B., and Littlewood, D. T. J. (1994). Paleontological data and molecular phylogenetic analysis. *Paleobiology* **20:** 259–273.

Stebbins, G. L. (1974). "Flowering Plants: Evolution above the Species Level," Harvard Univ. Press, Cambridge, MA.

Soltis, D. E., Soltis, P. S., Nickrent, D. L., Johnson, L. A., Hahn, W. J., Hoot, S. B., Sweere, J. A., Kuzoff, R. K., Kron, K. A., Chase, M. W., Swenson, S. M., Zimmer, E. A., Chaw, S.-M., Gillespie, L. J., Kress, W. J., and Sytsma, K. J. (1997). Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Mo. Bot. Gard.* **84:** 1–49.

Stefanovic, S. (1996). "Phylogénie Moléculaire et Relations des Familles de Coniférales Basée sur le Séquençage de l'ADN Ribosomique 28S," Diplôme d'Études Approfondies, Université de Paris.

Stewart, W. N., and Rothwell, G. W. (1993). "Paleobotany and the Evolution of Plants," 2nd ed., Cambridge Univ. Press, Cambridge, U.K.

Taylor, E. L., and Taylor, T. N. (1992). Reproductive biology of the Permian Glossopteridales and their suggested relationship to flowering plants. *Proc. Natl. Acad. Sci. USA* **89:** 11495–11497.

Troitsky, A. V., Melekhovets, Yu. F., Rakhimova, G. M., Bobrova, V. K., Valiejo-Roman, K. M., and Antonov, A. S. (1991). Angiosperm origin and early stages of seed plant evolution deduced from rRNA sequence comparisons. *J. Mol. Evol.* **32:** 253–261.

Van Konijnenburg-van Cittert, J. H. A. (1992). An enigmatic Liassic microsporophyll, yielding *Ephedripites* pollen. *Rev. Palaeobot. Palynol.* **71:** 239–254.

Wettstein, R. R. von (1907). "Handbuch der Systematischen Botanik," Vol. 2, Franz Deuticke, Leipzig/Vienna.