

# Multigene analyses identify the three earliest lineages of extant flowering plants

Christopher L. Parkinson, Keith L. Adams and Jeffrey D. Palmer

Flowering plants (angiosperms) are by far the largest, most diverse, and most important group of land plants, with over 250,000 species and a dominating presence in most terrestrial ecosystems. Understanding the origin and early diversification of angiosperms has posed a longstanding botanical challenge [1]. Numerous morphological and molecular systematic studies have attempted to reconstruct the early history of this group, identifying the root of the angiosperm tree. There is considerable disagreement among these studies, with various groups of putatively basal angiosperms from the subclass Magnoliidae having been placed at the root of the angiosperm tree (reviewed in [2–4]). We investigated the early evolution of angiosperms by conducting combined phylogenetic analyses of five genes that represent all three plant genomes from a broad sampling of angiosperms. *Amborella*, a monotypic, vesselless dioecious shrub from New Caledonia, was clearly identified as the first branch of angiosperm evolution, followed by the Nymphaeales (water lilies), and then a clade of woody vines comprising Schisandraceae and Austrobaileyaceae. These findings are remarkably congruent with those from several concurrent molecular studies [5–7] and have important implications for whether or not the first angiosperms were woody and contained vessels, for interpreting the evolution of other key characteristics of basal angiosperms, and for understanding the timing and pattern of angiosperm origin and diversification.

Address: Department of Biology, Jordan Hall, 1001 E. Third Street, Indiana University, Bloomington, Indiana 47405, USA.

Correspondence: Jeffrey D. Palmer  
E-mail: jpalmer@bio.indiana.edu

Received: 12 October 1999  
Revised: 8 November 1999  
Accepted: 8 November 1999

Published:

Current Biology 1999, 9:000–000

0960-9822/99/\$ – see front matter  
© 1999 Elsevier Science Ltd. All rights reserved.

## Results and discussion

Our efforts to identify the earliest angiosperms emphasized mitochondrial genes, in order to capitalize on the low rate of nucleotide substitutions in plant mitochondrial genomes [8]. Most of the sequences for the three mitochondrial genes analyzed (mtSSU rDNA, *cox1* and *rps2*) were generated in this study, whereas sequences for the

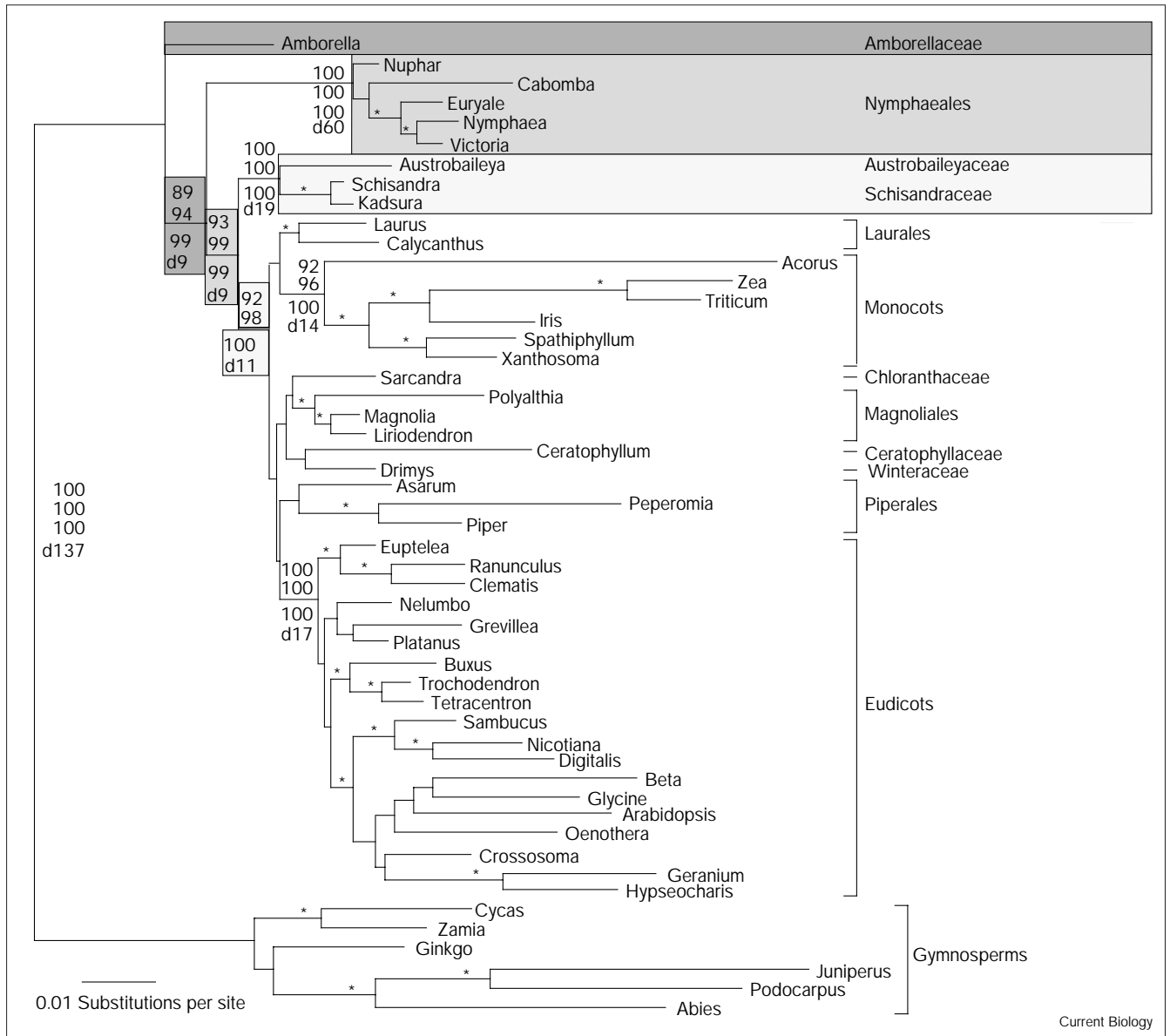
chloroplast *rbcL* and nuclear SSU rDNA genes are largely from GenBank. Forty-five diverse angiosperms, representing all major lineages of basal angiosperms, were included in the study, with six gymnosperms used as outgroups for comparison. Our sampling of angiosperms was based largely on the 1997 study by Soltis *et al.* [9] and the 1998 review by Doyle [3]. Gnetales, thought on morphological grounds to be the sister group of angiosperms [4], were not included as outgroups because recent molecular studies [7,10,11] indicate that they are instead gymnosperms with high rates of sequence evolution.

Individual analyses of the five genes yielded relatively poorly resolved trees; but importantly, the trees were not visibly incongruent with one another (see Supplementary material). Therefore, we deemed it both appropriate and necessary, in order to obtain better resolved and supported trees, to combine the five genes into a single, total molecular evidence data set. This yielded an alignment of 51 taxa with 6564 characters, of which 2393 were variable and 1391 were informative for parsimony analysis. The data were analyzed by maximum parsimony and maximum likelihood, using three methods to assess internal branch support (see Supplementary material).

The maximum-parsimony and maximum-likelihood analyses revealed that *Amborella trichopoda* (the sole member of the Amborellaceae) is the first branch of angiosperm evolution (Figure 1). This placement was strongly supported by both maximum-parsimony analyses (89% bootstrap support and a decay value of 9 steps) and maximum-likelihood analyses (94% bootstrap and 99% relative likelihood support). *Amborella* is an evergreen, dioecious shrub endemic to New Caledonia; it lacks vessels and contains many distinctive characteristics that are considered to be ancestral or erratic [12]. A monophyletic Nymphaeales (water lilies and related aquatic plants) was found to be the second branch of the angiosperm tree, while the third lineage was found to comprise Austrobaileyaceae and Schisandraceae (woody vines), with both placements being highly supported (Figure 1). Studies with more extensive taxonomic sampling have shown that Illiciaceae and Trimeniaceae also belong to the Austrobaileyaceae/Schisandraceae clade [6,7,9,13].

Our placement of *Amborella*, Nymphaeales, and Austrobaileyaceae/Schisandraceae as the three earliest groups of angiosperms does not appear to be an artifact of long branch attraction (the tendency of relatively divergent

Figure 1



Angiosperm relationships from a combined analysis of five genes. The topology shown is from maximum-likelihood analysis (log likelihood = -48543.97). Numbers on selected nodes are from top to bottom: maximum-parsimony bootstrap values, maximum-likelihood

bootstrap values, maximum-likelihood relative likelihood support values, and maximum-parsimony decay values. Asterisks indicate all maximum-parsimony bootstrap values that are > 90%.

branches in a phylogenetic tree to erroneously group together to the exclusion of intervening short branches due to excessive parallel and convergent changes on the long branches) to the very long branch separating the angiosperm ingroup from the gymnosperm outgroups. The branches leading to these three angiosperm groups are not notably long, and unrooted maximum-parsimony and maximum-likelihood analyses — that is, with gymnosperms excluded — of the combined data set yielded

unrooted networks that were topologically equivalent to the rooted trees of Figure 1 with respect to the placement of *Amborella*, Nymphaeales, and Austrobaileyaceae/Schisandraceae relative to each other and to other angiosperms. In addition, alternative topology testing using the maximum-likelihood KH test [14] was performed to investigate various hypotheses for the earliest branch of the angiosperm tree. Placement of *Amborella* as the basal-most member of Nymphaeales, or switching the position of

*Amborella* and Nymphaeales, was not statistically different at the 5% level from the topology presented in Figure 1. Significant differences were found, however, between the best maximum-likelihood tree and topologies in which the basal branch of angiosperms was designated as Austrobaileyaceae/Schisandraceae, the Magnoliales, *Ceratophyllum*, or the monocots. Thus, the maximum-likelihood analyses reject all angiosperms except for *Amborella* and/or Nymphaeales as the earliest angiosperms. It should be stressed that the KH test compares, for a particular data set, log likelihood scores for the entirety of the best tree with those of designated alternative topologies. Thus, a single nearest-neighbor interchange (as with *Amborella* and the Nymphaeales) might not cause a significant change in the overall tree likelihood score, even if it disrupts a node that is strongly supported by the bootstrap and other support indices.

Several concurrent multigene studies [5–7] (S. Graham and R. Olmstead, personal communication) have identified, with modest-to-high support, the same three basal branches of angiosperm evolution as recovered in our analyses (Figure 1). This remarkable confluence of congruent results was foreshadowed, in one part or another, in several earlier, mostly single-gene studies. *Amborella* was the most basal in a subset of nuSSU rDNA trees in the 1997 study by Soltis *et al.* [9], while two 1993 *rbcL* studies [15,16] first suggested that *Amborella* is closely related to the Nymphaeales (but did not place it as the first branching angiosperm). The Nymphaeales were placed at the base of the angiosperm tree in several early molecular studies [2,17–20], although *Amborella* was not included in any of them and support for the Nymphaeales placement was not high. An early origin of Austrobaileyaceae and relatives was first suggested by the 1997 nuSSU study of Soltis *et al.* [9].

The complete agreement between our study and concurrent multigene studies [5–7] (S. Graham and R. Olmstead, personal communication) as to the three basal lineages of angiosperms gives us great confidence that the evolutionary root of flowering plants has finally been resolved. Thus, other groups, such as Magnoliales, Ceratophyllaceae, and Chloranthaceae, which have previously been considered as candidates for the earliest angiosperms (reviewed in [2–4]), should no longer be regarded as such. Relationships are poorly resolved among these latter three groups and the five other, now clearly non-basal, groups in our study. Of the five multiply sampled groups, four (monocots, Laurales, Magnoliales, and eudicots) are well supported as being monophyletic (monophyly of Piperales is only weakly supported), but relationships among these groups and the Chloranthaceae, Ceratophyllaceae, and Winteraceae differ between maximum-parsimony and maximum-likelihood analyses and are poorly supported. Better sampling, of both taxa and genes, is evidently

needed to resolve these relationships (see for example [6,7]). Relationships within monocots are well resolved, with *Acorus* the most basal, as suggested in previous studies (for example [6,15]). Relationships within eudicots are generally consistent with other, more extensive studies (for example [6,7,15,21]); clade support is high for some groups but low for others.

Identification of the three earliest angiosperm groups provides the opportunity to infer features of the common ancestor of extant angiosperms, and to reevaluate the evolution of morphological, anatomical, and biochemical characteristics in basal angiosperms. *Amborella* and the Nymphaeales lack ethereal oil cells [22], and in all three first-branching groups, closure of carpel margins occurs by secretion [23–25]. Our phylogeny suggests that these traits are ancestral among angiosperms. A long-standing issue is whether the first angiosperms were woody or herbaceous. *Amborella* is a woody shrub, and the Austrobaileyaceae and Schisandraceae are both woody vines (the Illiciaceae and Trimeniaceae are lianas and small trees), whereas the Nymphaeales are herbaceous [22]. This suggests, although not persuasively, that the common ancestor of extant angiosperms was woody, with the Nymphaeales being derived from a woody ancestor. *Amborella* apparently lacks vessels [26,27], suggesting that the ancestral angiosperm condition was vesselless. The very recent discovery of vessels in some Nymphaeales ([28] and references therein), however, emphasizes the importance of reexamining *Amborella*. Our phylogeny suggests that the flowers of the first branching angiosperms were neither the small and very reduced flowers of the Piperales and Chloranthaceae, nor the large multiparted flowers of the Magnoliales (reviewed in [1,29]), but were more likely to be intermediate between these extremes. Although some Nymphaeales species have multiparted flowers, this has been proposed to represent a derived condition [30].

Results from this study also have implications for the timing and pattern of angiosperm origin and diversification. The earliest unambiguously angiosperm fossils are 120–130 million years old [1,31], and, where assignable, belong to groups that have been defined in our study as non-basal, such as Magnoliales, Winteraceae, Chloranthaceae, monocots, and eudicots [1,32]. This suggests an even earlier origin for *Amborella*, Nymphaeales, and the Austrobaileyaceae group. If fossils documenting this early period of angiosperm evolution are eventually recovered, it will be interesting to see how deeply they cut into what is now a very lengthy period (100–200 million years) of stem-group evolution that connects extant angiosperms to their sister group, either the extinct Bennettiales and *Caytonia* and/or extant gymnosperms [3,10,11]. That *Amborella*, the first branch of angiosperm evolution, is monotypic, and that the next two groups are relatively small (~160 species in total [22]), is consistent with the

suggestion of Sanderson and Donoghue [33] that early angiosperm evolution was not characterized by the high diversification rates found in many groups of latter-day angiosperms, although massive extinction within these early lineages cannot be ruled out either.

#### Supplementary material

Supplementary material, including a complete list of plant names, DNA voucher information, GenBank accession numbers for the sequences used in this study, and all molecular and phylogenetic methodology, is available at <http://current-biology.com/supmat/supmatin.htm>.

#### Acknowledgements

We thank B. Thomason, B. Hall, T. Vincent, Y. Cho, R. Price, and C. dePamphilis for providing some of the sequences for this study, Y.-L. Qiu for several DNAs, and M. Donoghue, S. Graham, S. Mathews, R. Olmstead, Y.-L. Qiu, D. Soltis, P. Soltis and M. Zanis for sharing unpublished data. Financial support was provided by NIH F32 GM-19225 to C.L.P., USDA training grant 95-38420-2214 to K.L.A., and NIH RO1-GM35087 to J.D.P.

#### References

- Crane PR, Friis EM, Pedersen KR: **The origin and early diversification of angiosperms.** *Nature* 1995, **347**:27-33.
- Donoghue MJ, Mathews S: **Duplicate genes and the root of angiosperms, with an example using phytochrome sequences.** *Mol Phylogenet Evol* 1997, **9**:489-500.
- Doyle JA: **Phylogeny of vascular plants.** *Annu Rev Ecol Syst* 1998, **29**:567-599.
- Doyle JA: **Molecules, morphology, fossils, and the relationship of angiosperms and Gnetales.** *Mol Phylogenet Evol* 1998, **9**:448-462.
- Mathews S, Donoghue MJ: **The root of angiosperm phylogeny inferred from duplicate phytochrome genes.** *Science* 1999, **286**:947-950.
- Soltis PS, Soltis DE, Chase MW: **Angiosperm phylogeny inferred from multiple genes: a research tool for comparative biology.** *Nature* 1999, in press.
- Qiu Y-L, Lee J, Bernasconi-Quadroni B, Soltis DE, Soltis PS, Zanis M, et al.: **The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes.** *Nature* 1999, in press.
- Wolfe KH, Li W-H, Sharp PM: **Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs.** *Proc Natl Acad Sci USA* 1987, **84**:9054-9058.
- Soltis DE, Soltis PS, Nickrent DL, Johnson LA, Hahn W, Hoot S, et al.: **Angiosperm phylogeny inferred from 18S ribosomal DNA sequences.** *Ann Missouri Bot Gard* 1997, **84**:1-49.
- Chaw S-M, 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* 1999, in press.
- 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* 1999, in press.
- Nandi OI, Chase MW, Endress PK: **A combined cladistic analysis of angiosperms using *rbcl* and non-molecular data sets.** *Ann Missouri Bot Gard* 1998, **85**:137-212.
- Renner SS: **Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data.** *Amer J Bot* 1999, **86**:1301-1315.
- Hasegawa M, Kishino H, Yano T: **Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.** *J Mol Evol* 1985, **22**:160-74.
- Chase MW, Soltis DE, Olmstead RG, Morgan D, Duvall MR, Price RA, et al.: **Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcl*.** *Ann Missouri Bot Gard* 1993, **80**:528-580.
- Qiu Y-L, Chase MW, Les DH, Parks CR: **Molecular phylogenetics of the Magnoliidae: cladistic analysis of nucleotide sequences of the plastid gene *rbcl*.** *Ann Missouri Bot Gard* 1993, **80**:587-606.
- Hamby RK, Zimmer EA: **Ribosomal RNA as a phylogenetic tool in plant systematics.** In *Molecular Systematics of Plants*. Edited by Soltis PS, Soltis DE, Doyle JJ. New York: Chapman and Hall, 1992.
- Doyle JA, Donoghue MJ, Zimmer EA: **Integration of morphological and ribosomal RNA data on the origin of angiosperms.** *Ann Missouri Bot Gard* 1994, **81**:419-450.
- 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.
- Hoot S, Magallon S, Crane PR: **Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcl*, and 18S nuclear ribosomal sequences.** *Ann Missouri Bot Gard* 1999, **86**:1-32.
- Cronquist A: *An Integrated System of Classification of Flowering Plants*. New York: Columbia University Press; 1981.
- Igershem A, Endress PK: **Gynoecium diversity and systematics of the magnoliales and winteroids.** *Bot J Linn Soc* 1997, **124**:213-271.
- Endress PK, Igershem A: **Gynoecium diversity and systematics of the laurales.** *Bot J Linn Soc* 1997, **125**:93-168.
- Igershem A, Endress PK: **Gynoecium diversity and systematics of the paleoherbs.** *Bot J Linn Soc* 1998, **127**:289-370.
- Bailey IW, Swamy BGL: ***Amborella trichopoda* Bail., a new morphological type of vesselless dicotyledon.** *J Arnold Arbor* 1948, **29**:245-254.
- Metcalfe CR: *Anatomy of the Dicotyledons*. Oxford: Clarendon Press; 1987.
- Schneider EL, Carlquist S: **Vessels in *Brasenia* (Cabombaceae): new perspectives on vessel origin in primary xylem of angiosperms.** *Amer J Bot* 1996, **83**:1236-1240.
- Endress P: **Floral structure and evolution of primitive angiosperms: recent advances.** *Plant Syst Evol* 1994, **192**:79-97.
- Les DH, Schneider EL, Padgett DL, Soltis PS, Soltis DE, Zanis M: **Phylogeny, classification and floral evolution of water lilies (Nymphaeaceae; Nymphaeales): A synthesis of non-molecular, *rbcl*, *matK*, and 18S rDNA data.** *Syst Bot* 1999, **24**:28-46.
- Swisher III CC, Wang Y-Q, Wang X-L, Xu X, Wang Y: **Cretaceous age for the feathered dinosaurs of Liaoning, China.** *Nature* 1999, **400**:58-61.
- Doyle JA, Donoghue MJ: **Phylogenies and angiosperm diversification.** *Paleobiology* 1993, **19**:141-167.
- Sanderson MJ, Donoghue MJ: **Shifts in diversification rate with the origin of angiosperms.** *Science* 1994, **264**:1590-1593.