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 long-standing botanical challenge [1]. Numerous morphological and molecular systematic studies have attempted to reconstruct the early history of this group, including 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: 6 December 1999

Current Biology 1999, 9:1485–1488

0960-9822/99/$ – see front matter

Results and discussion

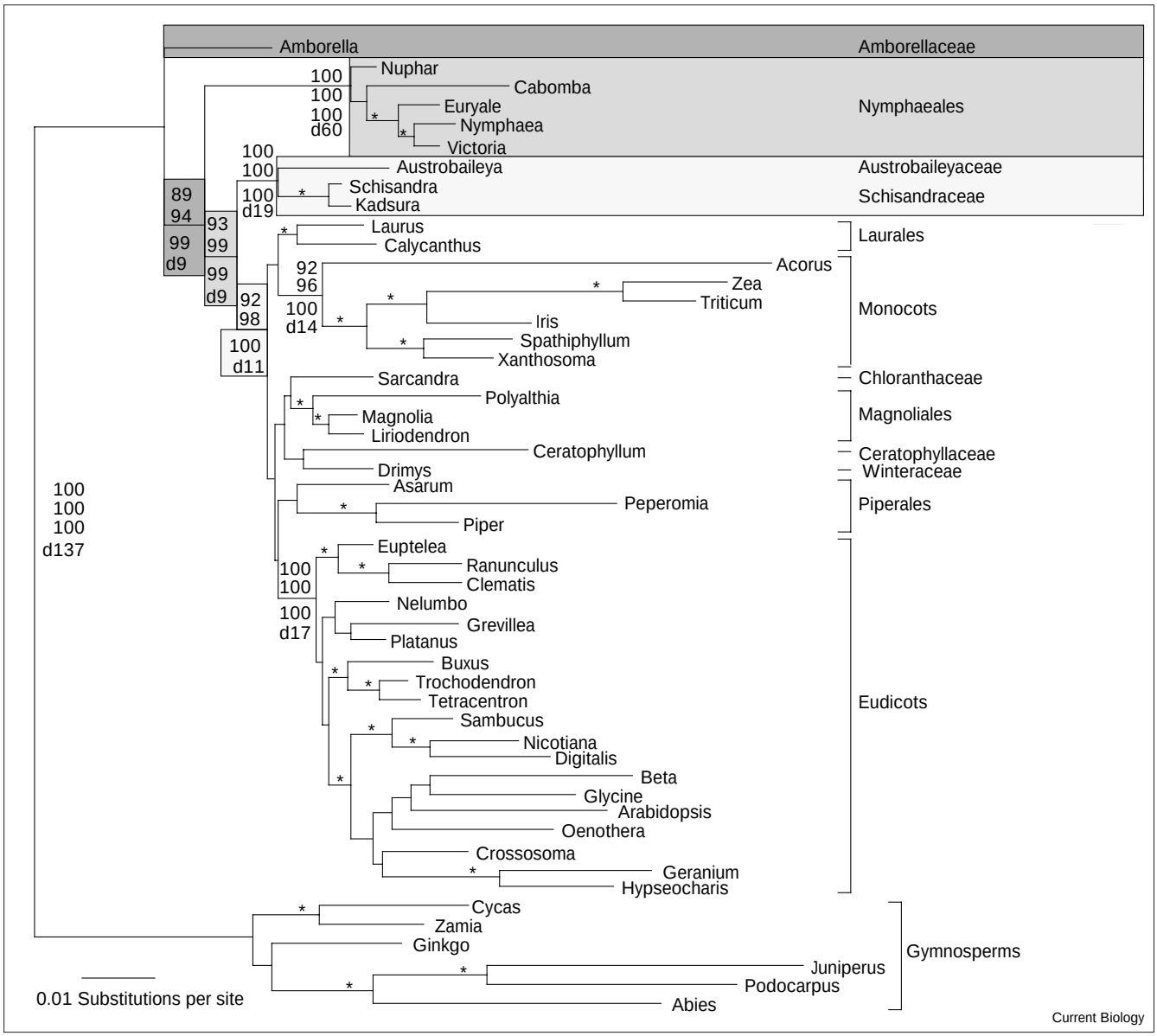
Our efforts to identify the earliest angiosperms emphasized mitochondrial genes, in order to capitalize on the low rate of nucleotide substitutions in plant mitochondrial genomes [8]. Most of the sequences for the three mitochondrial genes analyzed (mtSSU rDNA, *cox1* and *rps2*) were generated in this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

1. Crane PR, Friis EM, Pedersen KR: The origin and early diversification of angiosperms. *Nature* 1995, 347:27-33.
2. Donoghue MJ, Mathews S: Duplicate genes and the root of angiosperms, with an example using phytochrome sequences. *Mol Phylogenet Evol* 1997, 9:489-500.
3. Doyle JA: Phylogeny of vascular plants. *Annu Rev Ecol Syst* 1998, 29:567-599.
4. Doyle JA: Molecules, morphology, fossils, and the relationship of angiosperms and Gnetales. *Mol Phylogenet Evol* 1998, 9:448-462.
5. Mathews S, Donoghue MJ: The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 1999, 286:947-950.
6. Soltis PS, Soltis DE, Chase MW: Angiosperm phylogeny inferred from multiple genes: a research tool for comparative biology. *Nature* 1999, 402:402-404.
7. 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, 402:404-407.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. Renner SS: Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. *Amer J Bot* 1999, 86:1301-1315.
14. Hasegawa M, Kishino H, Yano T: Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 1985, 22:160-174.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. Cronquist A: *An Integrated System of Classification of Flowering Plants.* New York: Columbia University Press; 1981.
23. Igershem A, Endress PK: Gynoecium diversity and systematics of the magnoliales and winteroids. *Bot J Linn Soc* 1997, 124:213-271.
24. Endress PK, Igershem A: Gynoecium diversity and systematics of the laurales. *Bot J Linn Soc* 1997, 125:93-168.
25. Igershem A, Endress PK: Gynoecium diversity and systematics of the paleoherbs. *Bot J Linn Soc* 1998, 127:289-370.
26. Bailey IW, Swamy BGL: ***Amborella trichopoda*** Bail., a new morphological type of vesselless dicotyledon. *J Arnold Arbor* 1948, 29:245-254.
27. Metcalfe CR: *Anatomy of the Dicotyledons.* Oxford: Clarendon Press; 1987.
28. 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.
29. Endress P: Floral structure and evolution of primitive angiosperms: recent advances. *Plant Syst Evol* 1994, 192:79-97.
30. 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.
31. 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.
32. Doyle JA, Donoghue MJ: Phylogenies and angiosperm diversification. *Paleobiology* 1993, 19:141-167.
33. Sanderson MJ, Donoghue MJ: Shifts in diversification rate with the origin of angiosperms. *Science* 1994, 264:1590-1593.

Supplementary material

Multigene analyses identify the three earliest lineages of extant flowering plants

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

Current Biology 6 December 1999, 9:1485–1488

Supplementary materials and methods

Sequence generation, alignment, and phylogenetic analyses

Total DNA was isolated as described in [S1]. Whenever possible, sequences from the same species were used; otherwise, generic placeholders were substituted, as indicated in Tables S1 and S2. Newly generated sequences were obtained from PCR products, either cloned (TOPO TA cloning kit, Invitrogen) or uncloned, using primers and conditions as described in [S2,S3] for four of the genes. For *rps2*, the following primers were used for amplification: rps2.F2, 5′AAGACACTRATTTGTTTACGAA-3′; and either of the following reverse primers, rps2.R3, 5′-AYGGGATAAGTKATTMKTTTAT-3′, or rps2.R4, 5′-TCMAGAATSMCTGTTTTSRT-3′. PCR was performed using 20 ng total cellular DNA, 0.8 mM $MgCl_2$, 1 mM of each dNTP, 2 μM of each primer, and Taq polymerase. Reaction conditions were 94°C for 10 sec, 50°C for 20 sec, and 72°C for 1 min, for 35 cycles in an Idaho Air thermal cycler.

Individual gene alignments were generated as in [S2], with removal before analyses of all known and potential RNA editing sites, regions of problematical alignment, and the co-conversion tract downstream of the intron found in some *cox1* genes [S4]. The final combined alignment consisted of 51 taxa with 6564 nucleotides (mtSSU, 1794; nuSSU, 1678; *rbcL*, 1351; *cox1*, 1359; and *rps2*, 410). Alignments for all five genes used in the phylogenetic analyses are available from the authors.

Maximum parsimony analyses were conducted using PAUP* [S5] and used heuristic searches with random taxon addition (100 replicates), MULPARS on, and TBR branch-swapping. The combined analyses branch swapped to completion without removal of any taxa. Bootstrap support was assessed using heuristic searches with 1000 replications. All characters were weighted equally.

Maximum likelihood analyses were conducted using fastDNAml version 1.06 [S6]. We used the F84 model of Felsenstein [S7], with the initial transition/transversion (ti/tv) ratio estimated using PUZZLE (version 4.02) under the Tamura-Nei model of evolution with parameter estimation set to 'approximate' [S8]. Ten initial maximum-likelihood trees were inferred by randomizing 'input' order with jumble, and using 'global' swapping across all nodes (equivalent to subtree-pruning-regrafting). The optimal tree (best log-likelihood score) was then input into PAUP* [S5] to reoptimize the ti/tv ratio using a model which incorporates variability in rates of change. We used the F84 evolutionary model assuming a discrete gamma distribution with four categories of site-to-site rate variability. The resulting ti/tv ratio was used to infer a new tree as above, further optimizing branch lengths. This tree and the optimized ti/tv ratio were then used to estimate evolutionary rates of change for each sequence position by partitioning the sites into 35 'rate' categories using the program DNArates (S. Pract, R. Overbeek and G. Olsen, personal communication). A new maximum-likelihood tree, incorporating the rate categories and the re-optimized ti/tv ratio, was then inferred. This new optimal tree was then used for a second round of rates estimation and tree inference. This process was iterated until a stable topology was achieved.

For maximum-likelihood bootstrapping, the SEQBOOT program of PHYLIP [S7] was used to generate 100 pseudoreplicate data sets. These were then analyzed using fastDNAml version 1.06, with the resulting bootstrap numbers generated using CONSENSE (PHYLIP). Relative likelihood support scores were calculated using TREECONS [S9] after generating the 1000 best maximum-likelihood trees using the RESTART (from the best tree) and KEEP options in fastDNAml version 1.1. Decay analysis was performed using AUTODECAY [S10].

Analysis of single gene data sets

Only the *rbcL* and nuSSU rDNA data sets could be analyzed to completion – that is, branch swapping occurred to completion (using TBR) in both a heuristic search and in bootstrap analyses – in single-gene maximum-parsimony analyses that included all taxa. Individual analyses of mtSSU rRNA, *cox1*, and *rps2* were performed using maximum-parsimony heuristic searches, with 10 random additions, NNI branch swapping, and simple addition for 100 replications. Eleven eudicots were excluded from the *cox1* analysis and 14 angiosperms were excluded from the mtSSU analysis to enable completion of the analyses, while all eudicots were excluded from the *rps2* analysis (*rps2* appears to be missing from the mitochondrial genome of almost all eudicots and was probably transferred to the nucleus early in eudicot evolution: data not shown). The *rbcL*, *cox1*, and nuSSU analyses all gave a basal polychotomy of at least 10 clades in the 50% bootstrap consensus tree, the *rps2* analysis resolved *Amborella* as the deepest angiosperm with 88% bootstrap support (but followed by a massive polychotomy), and the mtSSU analysis placed *Amborella*, Nymphaeales and *Acorus* as the deepest angiosperms with 66% bootstrap support (again followed by a massive polychotomy; the clearly anomalously deep placement of the monocot *Acorus* in this analysis almost certainly reflects the extraordinary divergence of the mtSSU rRNA gene in *Acorus*).

Supplementary references

S1. Qiu YL, Cho Y, Cox JC, Palmer JD: The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* 1998, 394:671-674.

S2. 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.

S3. 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.

S4. Cho Y, Qiu YL, Kuhlman P, Palmer JD: Explosive invasion of plant mitochondria by a group I intron. *Proc Natl Acad Sci USA* 1998, 95:14244-14249.

S5. Swofford DL: *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, 4.0-B2 edition. Sunderland, Massachussetts: Sinauer; 1999.

S6. Olsen GJ, Matsuda H, Hagstrom R, Overbeek R: fastDNAml: A tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput Appl Biosci* 1994, 10:41-48.

S7. Felsenstein J: *PHYLIP: Phylogeny Inference Package.*, 3.5c edition. Seattle, Washington: University of Washington; 1993.

S8. Strimmer K, von Haeseler A: Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol Biol Evol* 1996, 13:964-969.

S9. Jermiin LS, Olsen GJ, Mengersen KL, Easteal S: Majority-rule consensus of phylogenetic trees obtained by maximum-likelihood analysis. *Mol Biol Evol* 1997, 14:1296-1302.

S10. Eriksson T: *Autodecay (Hypercard Stack Program),* 4.0 edition. Stockholm: Botaniska Institutionen, Stockholm University, 1998.

Table S1

Species used, DNA voucher information and GenBank accession numbers for each taxon for the three mitochondrial genes used in the analyses.

	Species name	DNA number	mtSSU rRNA	*cox1*	*rps2*
1	*Ginkgo biloba*	gb*/gb*/Qiu94015	AB029355	AF020565	AF193904
2	*Cycas revoluta/C.r./C.r.*	Qiu94051/gb*/Qiu94051	AB029356	AF020562	AF193905
3	*Zamia floridana/Z. bracteata/Z.f.*	gb*/gb*/Qiu95035	AB029357	AF020583	AF193906
4	*Abies homolepis/A. macrophyllus/ Pinus sp.*	Qiu96224/Qiu96224/Qiu94013	AB029360	AF020556	AF193907
5	*Podocarpus costalis/P. macrophyllus*	gb*/gb*/Qiu96148	AB029369	AF020575	no data
6	*Juniperus chinensis/J. virginiana/*	gb*/gb*/	AB029368	AF020567	no data
7	*Zea mays*	gb*/gb*/	X00794	X02660	AF202318
8	*Triticum aestivum*	gb*	K01229	Y00417	Y13920
9	*Iris sp./I.s./Trillium sp.*	Qiu95091/Qiu95091/Qiu95016	AF161087	unpub	AF193909
10	*Xanthosoma mafatta/X.m./ Philodendron oxycardium*	Qiu95063/Qiu95063/Qiu94065	AF193974	AJ223807	AF193910
11	*Spathiphyllum clevelandii*	Qiu94140	AF193975	AJ007554	AF193911
12	*Acorus calamus*	Qiu94052	AF193976	AF193944	AF195652
13	*Ceratophyllum demersum*	Qiu95003	AF193977	AF193945	AF193912
14	*Asarum canadense*	Qiu96018/ /Qiu96018	AF193978	unpub†	AF193913
15	*Nymphaea sp.*	gb*/gb*/Qui91029	AF161091	AF020570	AF193914
16	*Victoria sp.*	palmer788	AF193979	AF193946	AF193915
17	*Euryale sp.*	palmer790	AF193980	AF193947	AF193916
18	*Nuphar sp.*	palmer689	AF193981	AF193948	AF193917
19	*Cabomba sp.*	palmer688	AF193982	AF193949	AF193918
20	*Nelumbo nucifera*	palmer686	AF193983	AF193950	AF193919
21	*Schisandra spenanthera*	Qiu94165	AF193984	AF193951	AF193920
22	*Kadsura japonica*	Qiu94159	AF193985	AF193952	AF193921
23	*Drimys winteri*	palmer572/ /palmer572	AF197162	unpub†	AF193922
24	*Peperomia fosterii/–/P. argyreia*	Qiu96002/–/Qiu96001	AF193986	unpub‡	AF193923
25	*Piper betle/–/P. nigrum*	Qiu91048/ /Qiu97028	AF161088	unpub†	AF193924
26	*Amborella trichopoda*	Qiu97123	AF193987	AF193953	AF193925
27	*Austrobaileya scandens*	Qiu90030	AF193988	AF193954	AF193926
28	*Calycanthus floridus*	Qiu94155	AF193989	AF193955	AF193927
29	*Laurus nobilis*	Qiu94209	AF193990	AF193956	AF193928
30	*Polyalthia suberosa*	Qiu94008	AF193991	AF193957	AF193929
31	*Sarcandra grandifolia*	Qiu92002	AF193992	AF193958	AF193930
32	*Magnolia grandiflora*	gb*/gb*/palmer612	AF161089	AF020568	AF193931
33	*Liriodendron tulipifera*	Qiu94126	AF193993	AF193959	AF193932
34	*Clematis sp.*	Qiu95085	AF193994	AF193960	AF193933
35	*Ranunculus sp.*	gb*/ /Qui95024	AF161093	unpub†	AF193934
36	*Grevillea robusta*	Qiu94087	AF193995	AF193961	AF193935
37	*Buxus sp.*	Qiu94069	AF193996	AF193962	T§
38	*Platanus occidentalis*	gb*/ /Qiu94152	AF161090	unpub†	AF193936
39	*Euptelea polyandra*	Qiu95098	AF193997	AF193963	T§
40	*Tetracentron sinense*	Qiu94166	AF193998	AF193964	T§
41	*Trochodendron araliodes*	gb*	AF161092	AF020581	T§
42	*Nicotiana tabacum*	gb/Qiu94122	AF161095	unpub†	T§
43	*Glycine max*	gb*	M16859	M16884	T§
44	*Oenothera berteriana*	gb*	X61277	X05465	T§
45	*Beta vulgaris*	gb*	AF161094	M57645	T§
46	*Digitalis feruginca/purpurea*	Cho1/gb*	unpub‡	AJ223415	T§
47	*Arabidopsis thaliana*	gb*	Y08502	Y08502	T§
48	*Sambucus canadensis*	Qiu94098	AF194000	AF193965	T§
49	*Crossosoma bigelovii*	palmer1103	AF194001	AF193966	T§
50	*Hypseocharis pimpinellifolium*	palmer1102	AF194002	AF193967	T§
51	*Geranium himalayense*	CLP1	AF194003	AF193968	T§

Where placeholder taxa were used, the species name and DNA voucher information are separated by a slash (/): the order listed follows the gene order. *Sequence from GenBank. †Unpublished sequence of C. dePamphilis. ‡Unpublished sequence of Y. Cho. §*rps2* is thought to have been transferred to the nucleus (see Supplementary materials and methods).

Table S2

Species used, DNA voucher information and GenBank accession numbers for each taxon for the chloroplast rbcL and nuSSU rDNA.

	Species name	DNA number	*rbcL*	*nuSSU rRNA*
1	*Ginkgo biloba*	gb*	D10733	D16448
2	*Cycas circinalis/C. taitungensis*	gb*	L12674	D85297
3	*Zamia floridana/Z. pumila*	gb*	X58391	M20017
4	*Abies homolepis/A. lasiocarpa*	gb*	X58131	X79497
5	*Podocarpus costalis*	gb*	L12537	D38473
6	*Juniperis chinesis/Callitris rhomboidea*	gb*	289851	D38443
7	*Zea mays*	gb*	X86563	K02202
8	*Triticum aestivum/Oryza sativa*	gb*	344052	AF069218
9	*Iris germanica/Gladiolus buckerveldii*	gb*	L05307	L54602
10	*Xanthosoma mafatta/Gymnostachys anceps*	gb*	349165	AF069200
11	*Spathiphyllum clevelandii*	gb*/	4138464	unpub†
12	*Acorus calamus*	gb*	336205	L24078
13	*Ceratophyllum demersum*	gb*	1817557	D85300
14	*Asarum canadense/A. hayatanum*	gb*	348025	D29774
15	*Nymphaea odorata/A. tuberosa*	gb*	M77034	L24404
16	*Victoria cruziana*	gb*	343646	AF096698
17	*Euryale ferox*	gb*	336945	AF096694
18	*Nuphar variegata*	gb*	342742	AF096695
19	*Cabomba caroliniana/C. sp.* (Zanis 1998)	gb*	336459	AF096691
20	*Nelumbo nucifera*	gb*	342683	L75835
21	*Schisandra spenanthera*	gb*	294862	75842
22	*Kadsura japonica*	Qiu 94159	AF192969	AF192937
23	*Drimys winteri*	gb*	290206	U42823
24	*Peperomia sp./P. serpens*	gb*	294248	L24411
25	*Piper betle/P. kadsura*	gb*	L12660	D29778
26	*Amborella trichopoda*	gb*	289057	U42497
27	*Austrobaileya scandens*	gb*	289219	U42503
28	*Calycanthus floridus*	gb*	348081	U38318
29	*Laurus nobilis/Sassafras albidum*	Qiu 94209/gb*	AF192970	U52031
30	*Polyalthia suberosa*	Qiu 94008	AF192971	AF192938
31	*Sarcandra grandifolia*	gb*	294842	unpub†
32	*Magnolia macrophylla/M. acuminata*	gb*	X54345	D29776
33	*Liriodendron tulipifera*	gb*	Golenberg, 1990	unpub†
34	*Clematis sp./Xanthorhiza simplicissima*	Qiu 95085/gb*	AF192972	L75839
35	*Ranunculus trichophyllus/R. taisanensis*	gb*	L08766	D29780
36	*Grevillea robusta*	Qiu 94087	AF192973	AF192939
37	*Buxus sempervirens*	gb*	AF093717	L54065
38	*Platanus occidentalis*	gb*	L01943	U42794
39	*Euptelea polyandra*	gb*	290666	L75831
40	*Tetracentron sinense*	gb*	295282	U42814
41	*Trochodendron araliodes*	gb*	L01958	U42816
42	*Nicotiana tabacum/Brunsfelsia pauciflora*	gb*	Z00044	L49274
43	*Glycine max*	gb*	Z95552	X02623
44	*Oenothera/Clarkia xantiana*	gb*	Werman *et al.*	U67930
45	*Atriplex patula/Beta vulgaris*	gb*	X15925	AF161095
46	*Digitalis purpurea/D. feruginca*	gb*/Cho1	1490237	AF192940
47	*Arabidopsis thalina*	/gb*	unpub‡	X16077
48	*Sambucus racemosa/S. canadensis*	gb*/Qiu94098	294834	AF192941
49	*Crossosoma sp./C. bigelovii*	/palmer1103	unpub‡	AF192942
50	*Hypseocharis sp./sp.*		unpub‡	unpub†
51	*Geranium himalayense*	/CLP1	unpub‡	AF192943

Where placeholder taxa were used, the species name and DNA voucher information are separated by a slash (/): the order listed follows the gene order. *Sequences are from GenBank. †Unpublished sequence of D. Soltis, P. Soltis and/or M. Zanis. ‡Unpublished sequence of B. Price.