Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes

Yin-Long Qiu[1], Libo Li, Tory A. Hendry, Ruiqi Li, David W. Taylor, Michael J. Issa, Alexander J. Ronen, Mona L. Vekaria & Adam M. White

[1]*Department of Ecology & Evolutionary Biology, The University Herbarium, University of Michigan, Ann Arbor, Michigan 48109-1048, U.S.A. ylqiu@umich.edu (author for correspondence).*

Three mitochondrial (*atp1*, *matR*, *nad5*), four chloroplast (*atpB*, *matK*, *rbcL*, *rpoC2*), and one nuclear (18S) genes from 162 seed plants, representing all major lineages of gymnosperms and angiosperms, were analyzed together in a supermatrix or in various partitions using likelihood and parsimony methods. The results show that *Amborella* + Nymphaeales together constitute the first diverging lineage of angiosperms, and that the topology of *Amborella* alone being sister to all other angiosperms likely represents a local long branch attraction artifact. The monophyly of magnoliids, as well as sister relationships between Magnoliales and Laurales, and between Canellales and Piperales, are all strongly supported. The sister relationship to eudicots of *Ceratophyllum* is not strongly supported by this study; instead a placement of the genus with Chloranthaceae receives moderate support in the mitochondrial gene analyses. Relationships among magnoliids, monocots, and eudicots remain unresolved. Direct comparisons of analytic results from several data partitions with or without RNA editing sites show that in multigene analyses, RNA editing has no effect on well supported relationships, but minor effect on weakly supported ones. Finally, comparisons of results from separate analyses of mitochondrial and chloroplast genes demonstrate that mitochondrial genes, with overall slower rates of substitution than chloroplast genes, are informative phylogenetic markers, and are particularly suitable for resolving deep relationships.

KEYWORDS: *Amborella*, basal angiosperms, magnoliids, mitochondrial genes, Nymphaeales, RNA editing.

INTRODUCTION

The basal angiosperms have been subject to some of the most intensive phylogenetic analyses ever conducted on any group of organisms using numerous methods on a large number of molecular markers over the last fifteen years (Martin & Dowd, 1991; Hamby & Zimmer, 1992; Chase & al., 1993; Qiu & al., 1993, 1999, 2000, 2001, 2005; D. Soltis & al., 1997, 2000; Hoot & al., 1999; Mathews & Donoghue, 1999, 2000; Parkinson & al., 1999; Renner, 1999; P. Soltis & al., 1999; Barkman & al., 2000; Graham & Olmstead, 2000; Savolainen & al., 2000; Nickrent & al., 2002; Zanis & al., 2002, 2003; Borsch & al., 2003; Goremykin & al., 2003, 2005; Hilu & al., 2003; Sauquet & al., 2003; Aoki & al., 2004; Kim & al., 2004; Stefanovic & al., 2004; Leebens-Mack & al., 2005; Löhne & Borsch, 2005). As a result, this abominably mysterious group of plants are among the best known for their long-thought unknowable evolutionary history among many similarly difficult groups of organisms on the tree of life. At present, the following consensus is emerging. *Amborella* or *Amborella* + Nymphaeales represent the earliest-diverging lineage of extant angiosperms, followed by Austrobaileyales (Mathews & Donoghue, 1999, 2000; Parkinson & al., 1999; Qiu & al., 1999, 2000, 2001, 2005; P. Soltis & al., 1999; Barkman & al., 2000; Graham & Olmstead, 2000; D. Soltis & al., 2000; Zanis & al., 2002; Borsch & al., 2003, 2005; Hilu & al., 2003; Aoki & al., 2004; Kim & al., 2004; Stefanovic & al., 2004; Goremykin & al., 2005; Leebens-Mack & al., 2005; Löhne & Borsch, 2005). The remaining angiosperms (euangiosperms, cf. Qiu & al., 1999) are divided into five lineages: Chloranthaceae, *Ceratophyllum*, monocots, magnoliids, and eudicots (Qiu & al., 1999, 2000, 2005; Zanis & al., 2002; Hilu & al., 2003). There is no strongly supported, generally agreed resolution on relationships among these five lineages, except that *Ceratophyllum* has been accepted to be sister to eudicots by The Angiosperm Phylogeny Group (2003). The magnoliids are composed of four well-defined taxa, Magnoliales, Laurales, Canellales, and Piperales, which form two pairs of sister groups. The monophyly of magnoliids and relationships among the four member lineages have received steady increase of bootstrap support as more data are added in the analysis (Mathews & Donoghue, 1999; Qiu & al., 1999, 2005; Zanis & al., 2002).

Exploration in several dimensions have been pursued in this exemplar case of phylogenetic reconstruction. Increasing taxon sampling represents one of the earliest and main assaults on this long-standing problem (Chase & al., 1993; D. Soltis & al., 1997, 2000; Savolainen & al., 2000; Hilu & al., 2003). Selection of markers from different genomes, encoding different functions, and evolving at different rates was also attempted right from the beginning and received a major boost after automated sequencing became available (Martin & Dowd, 1991; Hamby & Zimmer, 1992; Qiu & al., 1993, 1999, 2005; D. Soltis & al., 1997; Mathews & Donoghue, 1999; Graham & Olmstead, 2000; Savolainen & al., 2000; Borsch & al., 2003, 2005; Hilu & al., 2003; Aoki & al., 2004; Kim & al., 2004; Löhne & Borsch, 2005). Relative merits of character versus taxon sampling on influencing tree topology stability have been investigated in several recent studies (Graham & Olmstead, 2000; Goremykin & al., 2003, 2005; Stefanovic & al., 2004, Leebens-Mack & al., 2005). Finally, experimentation with different methods of data analysis (e.g., constraint analysis/alternative topology testing, noise reduction, compartmentalization, and synapomorphic sites identification) and rooting techniques (e.g., duplicated gene rooting and random sequence outgroup rooting), besides regular application of parsimony, likelihood, and distance methods as well as bootstrapping, jackknifing, and decay analyses, have been performed to specifically address the rooting issue and robustness of parts of the internal topology (Qiu & al., 1993, 2000, 2001, 2005; Mathews & Donoghue, 1999; Parkinson & al., 1999; Barkman & al., 2000; Zanis & al., 2002, 2003). These efforts have undoubtedly contributed to the progress on our understanding of relationships among basal angiosperms. To solve remaining problems, more data sources and analytic methods need to be explored.

Among three genomes within the plant cell, the mitochondrial genome was the last to be utilized for phylogenetic reconstruction. This situation might have been caused by the perception that several features of mitochondrial genes were undesirable. These include low substitution rates, the dramatic lineage-specific rate heterogeneity, abundant occurrence of introns (many of which are *trans*-spliced, making PCR difficult), idiosyncratic RNA editing patterns across genes and lineages, sometimes frequent losses from the genome, and horizontal transfers (Palmer, 1992; Bowe & dePamphilis, 1996; Malek & Knoop, 1998; Steinhauser & al., 1999; Adams & al., 2002; Bergthorsson & al., 2003; Won & Renner, 2003; Cho & al., 2004, C. Davis & Wurdack, 2004; Dombrovska & Qiu, 2004; Knoop, 2004; Qiu & Palmer, 2004; Parkinson & al., 2005). However, some of these features are actually advantageous for certain kinds of phylogenetic reconstruction (e.g., genes of low substitution rates can yield less homoplasious characters and intron presence/absence can be used as informative characters), while others may not be as problematic as previously thought, such as the lineage-specific rate heterogeneity, horizontal gene transfer, and RNA editing, as long as the care is taken to deal with these problems. A large number of studies that have employed mitochondrial genes and introns for phylogenetic reconstruction across land plants support this evaluation (Hiesel & al., 1994; Malek & al., 1996; J. Davis & al., 1998, 2004; Qiu & al., 1998, 1999, 2005, in press; Beckert & al., 1999, 2001; Duff & Nickrent, 1999; Vangerow & al., 1999; Barkman & al., 2000, 2004; Bowe & al., 2000; Chaw & al., 2000; Wang & al., 2000; Gugerli & al., 2001; Anderberg & al., 2002; Nickrent & al., 2002; Sanjur & al., 2002; Cox & al., 2004; Dombrovska & Qiu, 2004; Qiu & Palmer, 2004; Groth-Malonek & al., 2005; Guo & Ge, 2005; Wikström & Pryer, 2005). At a time when many problems in the angiosperm and land plant phylogenies are still unresolved and suitable markers from the nuclear genome that can be used on wide taxonomic scales without the problem of paralogy are limited in number, the mitochondrial genome remains the only other viable choice and a fertile ground for exploration outside the chloroplast genome. Furthermore, for many deep relationships in the angiosperm and land plant phylogenies, because of their controversial nature and volatility, it is imperative to use information from a different genome to evaluate independently the results that are derived largely from chloroplast genes (Qiu & Palmer, 1999). Therefore, the potential for mining historical signals from the mitochondrial genome to understand the evolutionary history of plants deserves a serious and careful assessment.

In this study, we analyze eight mitochondrial, chloroplast, and nuclear genes from all major lineages of gymnosperms and angiosperms using both maximum likelihood and maximum parsimony methods, aiming to achieve two goals. We first want to use these data (1) to determine whether the earliest-diverging lineages of extant angiosperms consists of *Amborella* alone or *Amborella* + Nymphaeales, (2) to evaluate the monophyly of magnoliids and sister relationships between Magnoliales and Laurales, and between Canellales and Piperales, and (3) to resolve relationships among Chloranthaceae, *Ceratophyllum*, magnoliids, monocots, and eudicots. We then want to use this case of reconstructing the basal angiosperm phylogeny, which employs several genes from both mitochondrial and chloroplast genomes, to assess relative contributions of each group of genes in conceptualizing our current understanding of the basal angiosperm relationships, particularly on controversy surrounding the basalmost node of the angiosperm phylogeny, placement of *Ceratophyllum*, and relationships among the four magnoliid member lineages.

MATERIALS AND METHODS

Taxon sampling. — Several different taxon sampling schemes have been employed to investigate the basal angiosperm phylogeny, ranging from small (Graham & al., 2000; Goremykin & al., 2003, 2005; Stefanovic & al., 2004; Leebens-Mack & al., 2005), medium (Qiu & al., 1999, 2005), to large-scale (Chase & al., 1993; D. Soltis & al., 1997, 2000; Savolainen & al., 2000; Hilu & al., 2003). While the small and large-scale taxon sampling strategies maximize signal retrieval through either extensive character or taxon sampling, the medium-scale taxon sampling strategy we have adopted takes advantage of dense, but not so extensive, sampling of both taxa and characters. In this study, we continue this same taxon sampling strategy but with increased densities within monocots and eudicots. These two groups had not been sampled densely in our previous studies (Qiu & al., 1999, 2005), which might have affected our ability to determine their relationships with magnoliids, *Ceratophyllum*, and Chloranthaceae. In this study, the number of monocots has been almost tripled and that of eudicots more than doubled. All other groups, *Amborella*, Nymphaeales, Austrobaileyales, Chloranthaceae, *Ceratophyllum*, magnoliids, and gymnosperms remain virtually unchanged in the number of taxa. A detailed list of taxa for all groups investigated is shown in the Appendix. A total of 162 gymnosperms and angiosperms were included.

Gene choices. — Our medium-scale taxon sampling strategy allows us to increase the number of genes relatively easily. Since the rRNA genes in the mitochondrial genomes of gymnosperms and basal angiosperms have many insertions and deletions, and the nuclear large subunit rRNA gene often has pseudogenes, which all make analysis problematic (Qiu & al., 2005), we focused on protein-coding genes in this study. Recent studies analyzing *matK* across angiosperms and *atpB* across land plants demonstrate that fast-evolving genes can be useful for reconstructing deep phylogenies if a proper taxon sampling density is ensured (Hilu & al., 2003; Qiu & al., in press). Hence, for the chloroplast genes used in this study, besides the widely used *atpB*, *matK*, and *rbcL*, we added *rpoC2* (encoding the beta chain of RNA polymerase), which has an evolutionary rate comparable to that of *matK* (Y. Qiu, unpublished data). For the mitochondrial genes, we added *nad5* (encoding nicotinamide adenine dinucleotide dehydrogenase subunit 5) to the commonly sequenced *atp1* and *matR*. This gene is slightly slower than *atp1* and *matR* (Y. Qiu, unpublished data), and has four group II introns in angiosperms (Malek & Knoop, 1998). We sequenced only the second exon with approximately 1.2 kb. For the nuclear gene, we used 18S rRNA gene. For *rpoC2* and *nad5*, we obtained sequences from almost all 162 taxa, and for the other six genes (atpB, *matK*, *rbcL*, *atp1*, *matR*, and 18S), we added data for species that lacked data in the GenBank. Many taxa had only a short sequence for *matK* in the GenBank, and we re-sequenced the gene to obtain a length of 1.4–1.5 kb so as to retrieve signal across entire length of the gene. Detailed information on sequences of these eight genes is provided in the Appendix.

DNA extraction, gene amplification, sequencing, alignment, and phylogenetic analyses. — The protocols of DNA extraction, gene amplification, and sequencing are the same as described in Qiu & al. (2000, 2005). Since *atp1*, *matR*, atpB, *matK*, *rbcL*, and 18S have been used extensively by us and others, we will not publish their primer sequences here, which are available upon request. For the two new genes used in this study, mitochondrial *nad5* and chloroplast *rpoC2*, the primer sequences are provided in Table 1. The genes were aligned individually using ClustalX (Thompson & al., 1997) with manual adjustment. They were then combined to form: (1) the 8-gene matrix, (2) the 4-cp (chloroplast) gene matrix, (3) the 4-cp + nu18S gene matrix, (4) the 4-cp + 3-mt (mitochondrial) gene matrix, and (5) the 3-mt gene matrix. Because RNA editing has been reported in mitochondrial *atp1*, *matR*, and *nad5* and chloroplast *matK* and *rpoC2* (see Table 2), and in the case of *nad5*, the editing is quite frequent (19 in 1248 = 1.5% nucleotides that we sequenced), we removed all known RNA editing sites from these five genes and assembled another set of matrices in parallel to the above five matrices to systematically investigate the effect of RNA editing on phylogenetic reconstruction, especially on deep relationships.

It is becoming increasingly clear in phylogenetics that obtaining assessment of confidence levels using statistical resampling methods such as bootstrapping or

Table 1. Primer sequences used to amplifying and sequencing *nad5* and *rpoC2*.

BA-*rpoC2*-F0	A GCT GTA GGT ATT ATW GCN GG
BA-*rpoC2*-F1	AGG GRG ARA TGC AYT GGA G
BA-*rpoC2*-F2	CTT GCY TAT TTY GAY GAT CC
BA-*rpoC2*-F3	CGA CCC GTA GTS NCA TAT G
BA-*rpoC2*-F4	TG TCA TCA TCC AAY TGK TYT C
BA-*rpoC2*-F5	GAT ATA ACN CAR GGT CTT CC
BA-*rpoC2*-R1	TTT CAG GCC TTT YAR CCA RTC
BA-*rpoC2*-R2	C ACC AGA TCT YGA YYT TTC
BA-*rpoC2*-R3	GAA CAA GAT YTN ATT ATG
BA-*rpoC2*-R4	CG AAC TAA TTG AMT RYT TGT G
BA-*rpoC2*-R5	G GGG ATC GTC RAA AWA RGC
BA-*rpoC2*-F12	GT CAT TTA TGG RTA TTR TCR G
BA-*rpoC2*-R12	AGA TGC YYT YGT TAT TCC
nad5-F1	CCG TAG TNA TGT YAA TTG TGG
nad5-F2	TCC CAC TCC WGT ATC YGC
nad5-R1	TAT CCT ACA AAR AKA CTM CC
nad5-R2	TTC AAA TRA AGG RGA RCA CC

Table 2. Information on RNA editing sites deleted from the matrix.

Gene	# of edited sites deleted	Reference taxa (GenBank accession #)*		
atp1	2	Os (BA000029)	Bn (AP006444)	—
matR	6	At (NC_001284)	Bn (AP006444)	Zm (AY506529)
nad5	19	At (NC_001284)	Bn (AP006444)	—
matK	3	At (NC_000932)	Ac (AY178864)	Zm (X86563)
rpoC2	3	Af (NC_004543)	Nt (Z00044)	Zm (X86563)

*Ac, *Adiantum capillus-veneris;* Af, *Anthoceros formosae;* At, *Arabidopsis thaliana;* Bn, *Brassica napus;* Nt, *Nicotiana tabacum;* Os, *Oryza sativa;* Zm, *Zea mays.*

jackknifing is more informative than obtaining the tree topology itself under the optimality criteria such as parsimony or likelihood (Nei & al., 1998). Thus, for phylogenetic analysis, we conducted only bootstrapping analysis using both maximum parsimony and maximum likelihood methods. The parsimony bootstrapping analyses (1000 replicates in all cases) were conducted in PAUP* (Swofford, 1998), using simple taxon addition, one tree held at each step during stepwise addition, tree-bisection-reconnection branch swapping, steepest descent option on, MulTree option on, and no upper limit of MaxTree set. For likelihood bootstrapping analyses (100 replicates in all cases), Modeltest 3.06 (Posada & Crandall, 1998) was used to select the best models of sequence evolution for each data set. The general time-reversible model (GTR+I+G) was suggested by the hierarchical Likelihood Ratio Test (hLRT) and the Akaike Information Criterion (AIC) as the best-fit model for all sequence data sets. Maximum likelihood bootstrapping analyses were conducted using PHYML version 2.4.4 (Guindon & Gascuel, 2003) under the optimal model of sequence evolution. For each data set, the GTR+I+G model was implemented with parameter values for the proportion of invariant sites and the gamma distribution as estimated by Modeltest 3.06 (Table 3).

For the 3-mt gene matrix with and without edited sites, only likelihood bootstrapping analyses were performed because parsimony bootstrapping analyses on this matrix could not be finished within a reasonable amount of time. Further, 11 taxa—two *Acorus* species, eight alismatids (*Triglochin, Zostera, Potamogeton, Butomus, Hydrocharis, Najas, Hydrocleys, Alisma*), and *Cyperus*—had to be removed from the matrix, as they formed a clade falling into eudicots with modest bootstrap support in two of the several nodes leading to their position. We interpreted this result as an artifact caused by long branch attraction because all three groups of monocots had highly divergent sequences that were likely due to rate acceleration in the mitochondrial genes, similar to the phenomenon reported in *Plantago* and *Pelargonium* (Cho & al., 2004; Parkinson & al., 2005). Our examination and phylogenetic analysis of the sequences did not show any indication that misplacement of these three groups were caused by horizontal gene transfer.

RESULTS

One identical topology was recovered by both likelihood and parsimony bootstrapping analyses of five of the ten matrices: 1) the 8-gene matrix without edited sites (Fig. 1); 2) and 3) the 4-cp gene matrix with and without edited sites, respectively (Fig. 2B); 4) and 5) the 4-cp + nu18S gene matrix with and without edited sites, respectively (Fig. 2C). This same topology was also recovered by the parsimony bootstrapping analyses of the 4-cp + 3-mt gene matrix with and without edited sites (Fig. 2D). In this topology, *Amborella* represents the first diverging lineage of all extant angiosperms, followed by Nymphaeales, Austrobaileyales, and Chloranthaceae sequentially. Among the remaining angiosperms, monocots are sister to a clade in which magnoliids are sister to eudicots plus *Ceratophyllum*.

In likelihood and parsimony bootstrapping analyses of the 8-gene matrix with edited sites, a different topology was obtained: Chloranthaceae and *Ceratophyllum* form a monophyletic group sister to the clade in which magnoliids are sister to eudicots plus monocots (Fig. 2A). The placements of *Amborella*, Nymphaeales, and Austrobaileyales are the same as in the first topology. A third topology, which differs from the first topology only in placing *Amborella* and Nymphaeales together as a

Table 3. The proportion of invariant sites and the gamma distribution shape parameter for the 8-gene matrix and its various partitions.

Matrix	Invariable site proportion (I)	Shape parameter (G)
8-gene	0.27	0.78
8-gene without edited sites	0.27	0.79
4-cp gene	0.20	1.07
4-cp gene without edited sites	0.20	1.07
4-cp + 18S rRNA gene	0.27	0.96
4-cp + 18S rRNA gene without edited sites	0.27	0.96
4-cp + 3-mt gene	0.23	0.82
4-cp + 3-mt gene without edited sites	0.23	0.82
3-mt gene	0.23	0.65
3-mt gene without edited sites	0.24	0.69

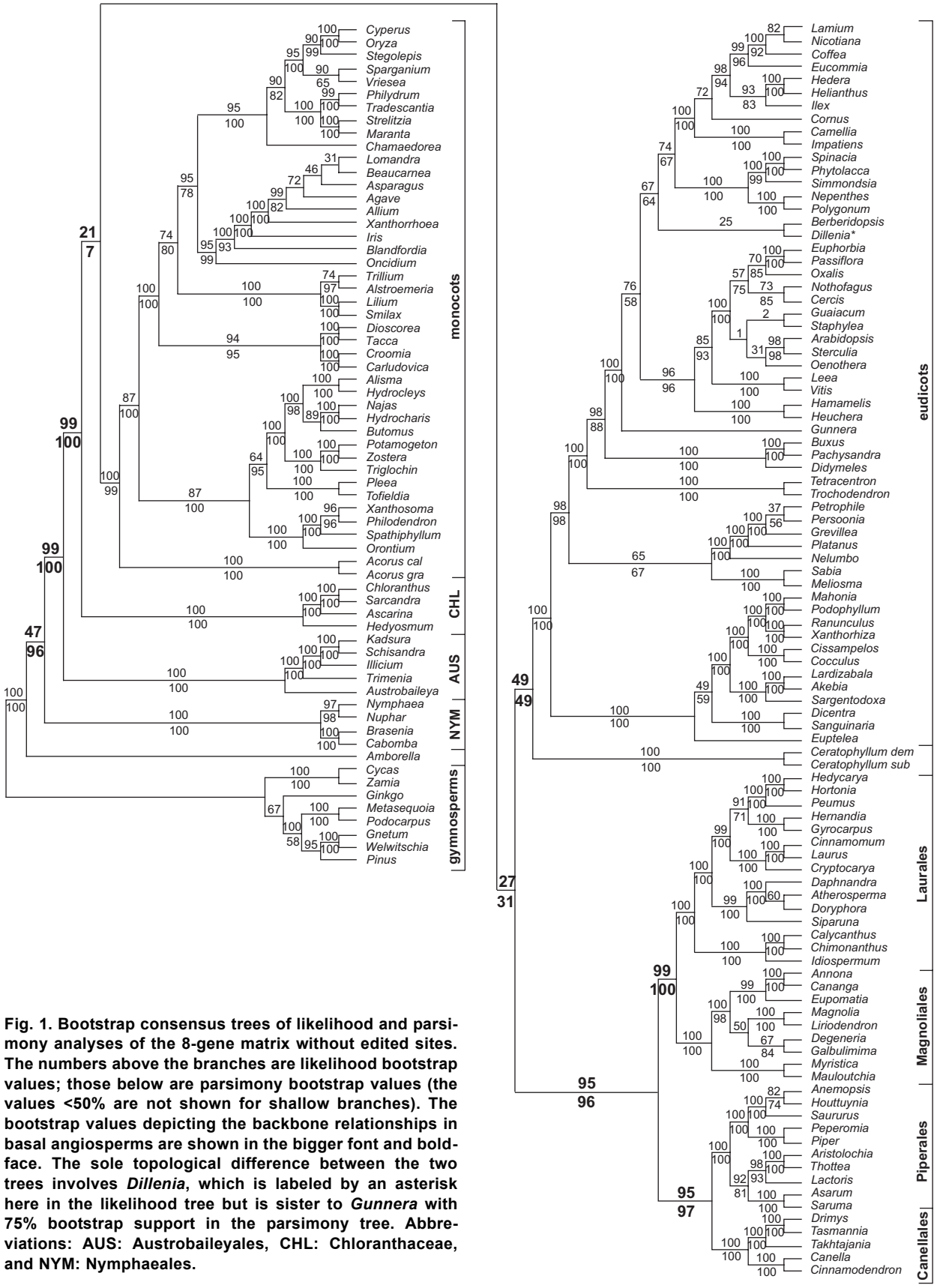

Fig. 1. Bootstrap consensus trees of likelihood and parsimony analyses of the 8-gene matrix without edited sites. The numbers above the branches are likelihood bootstrap values; those below are parsimony bootstrap values (the values <50% are not shown for shallow branches). The bootstrap values depicting the backbone relationships in basal angiosperms are shown in the bigger font and boldface. The sole topological difference between the two trees involves *Dillenia*, which is labeled by an asterisk here in the likelihood tree but is sister to *Gunnera* with 75% bootstrap support in the parsimony tree. Abbreviations: AUS: Austrobaileyales, CHL: Chloranthaceae, and NYM: Nymphaeales.

group sister to all other angiosperms, was recovered in the likelihood bootstrapping analyses of the 4-cp + 3-mt gene matrix with and without edited sites (Fig. 2D). Finally, in likelihood bootstrapping analyses of the 3-mt gene matrix with and without edited sites, a fourth topology was obtained (Fig. 3). *Amborella* and Nymphaeales together as a group are sister to all other angiosperms. Austrobaileyales diverge after this clade. Among euangiosperms, eudicots are sister to a clade in which Chloranthaceae and *Ceratophyllum* form a pair of sister groups and magnoliids and monocots form the other pair.

In terms of robustness of the relationships identified in these analyses, four categories can be recognized. First, in the four topologies recovered from all 18 analyses, the following relationships received strong bootstrap support, no matter how the data were partitioned and analyzed: the basalmost position of *Amborella* and Nymphaeales among all angiosperms, the intermediate position of Austrobaileyales between these two groups and euangiosperms, and monophyly of euangiosperms (Figs. 1–3). Second, the monophyly of magnoliids and sister relationships between Canellales and Piperales,

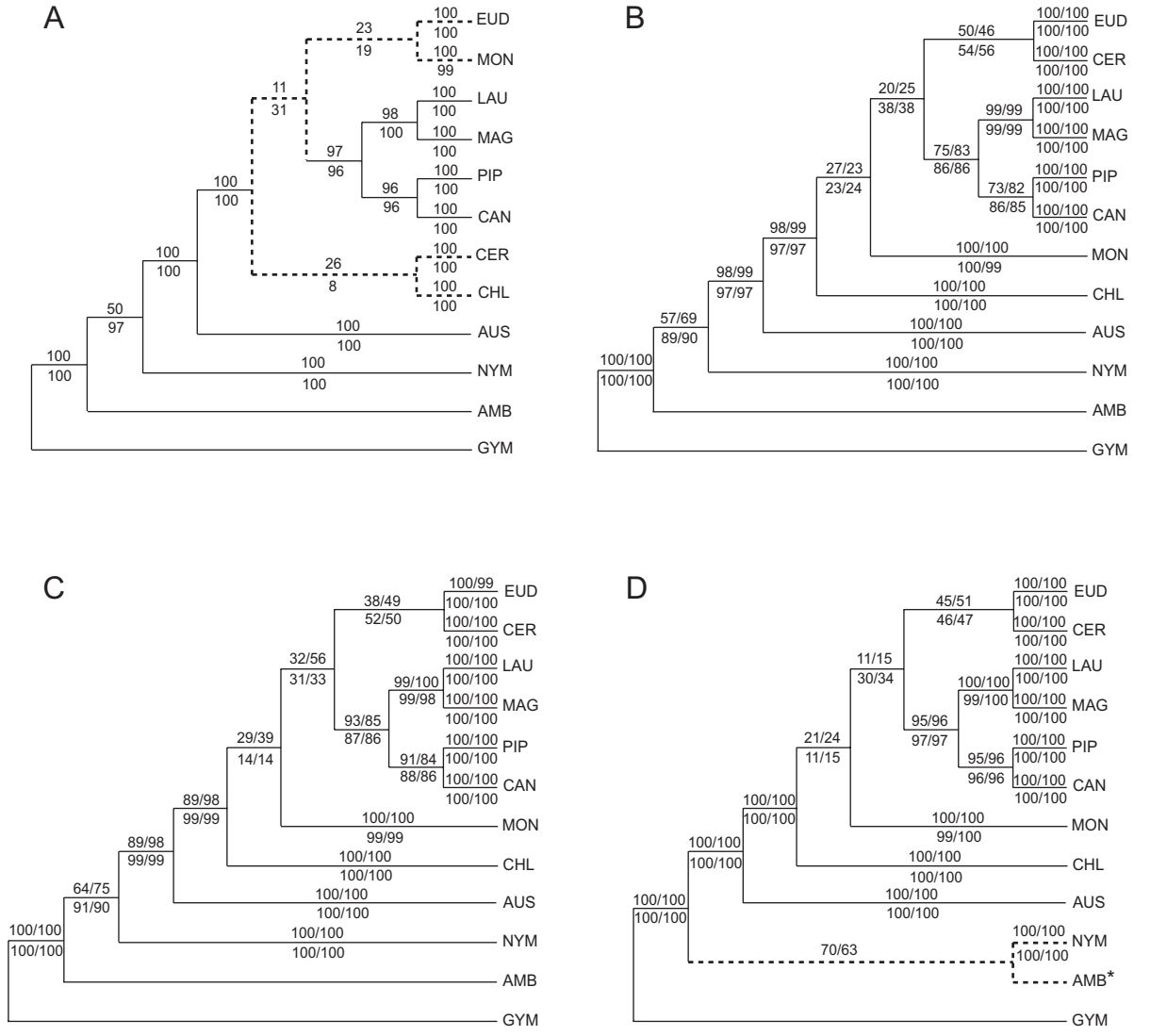

Fig. 2. Schematic presentation of bootstrap consensus trees of likelihood and parsimony analyses of the 8-gene matrix and its various partitions (details of the topology within the clades are virtually identical to those shown in Fig. 1). The numbers above the branches are likelihood bootstrap values; those below are parsimony bootstrap values (the bootstrap values before and after slashes in Figs. 2B, 2C, and 2D are from analyses of the matrices with and without edited sites, respectively). The thicker, dashed lines in Figs. 2A and 2D indicate relationships that are different from Figs. 2 B and 2C and Fig. 1. A, the 8-gene matrix with edited sites; B, the 4-cp gene matrix; C, the 4-cp + nu18S gene matrix; D, the 4-cp + 3-mt gene matrix (in the parsimony tree, *Amborella* is not sister to Nymphaeales, and instead is sister to Nymphaeales and all other angiosperms with 97/95% bootstrap support).

and between Magnoliales and Laurales, received steady increase of bootstrap support when nuclear 18S, three mitochondrial genes, and four chloroplast genes were combined in various ways (Figs. 1–3). Third, the placement of *Amborella* alone at the base of the angiosperm phylogeny received low bootstrap support in the likelihood analyses but high support in the parsimony analyses (Figs. 1, 2). In contrast, *Amborella* + Nymphaeales

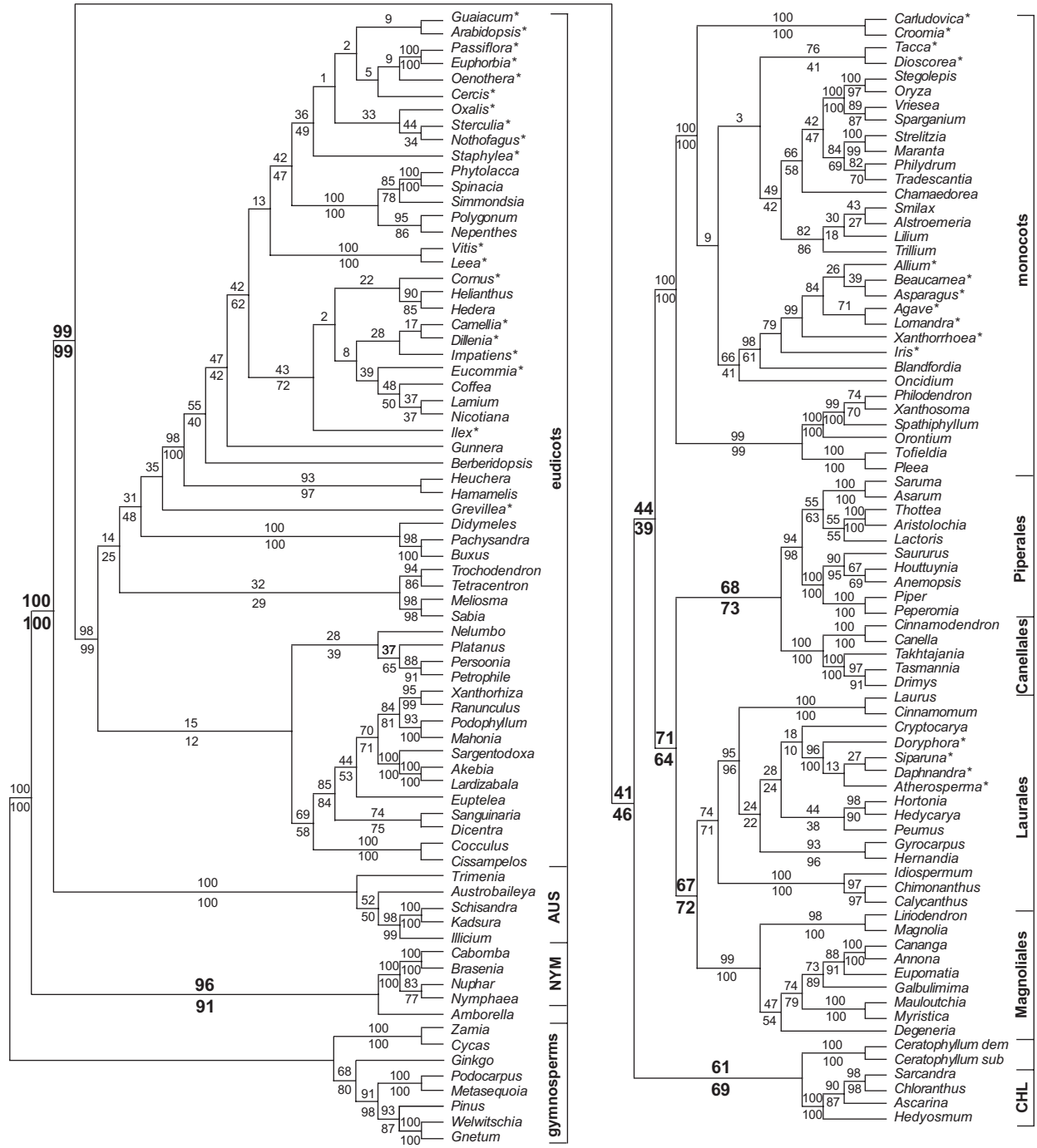

Fig. 3. Bootstrap consensus trees of likelihood analyses of the 3-mt gene matrix with and without edited sites. The numbers above and below the branches are bootstrap values from the analyses without and with edited sites, respectively. The bootstrap values depicting the backbone relationships in basal angiosperms are shown in the bigger font and boldface. The tree shown is from the analysis without edited sites, and the branches lacking bootstrap values and the taxa labeled by asterisks are arranged differently in the analysis with edited sites. Abbreviations: AUS: Austrobaileyales, CHL: Chloranthaceae, and NYM: Nymphaeales.

together as a sister to all other angiosperms received moderate to high bootstrap support in likelihood analyses of the 4-cp + 3-mt gene and the 3-mt gene matrices (Figs. 2D, 3). Finally, the relationships among Chloranthaceae, monocots, magnoliids, *Ceratophyllum*, and eudicots are resolved differently by different data partitions and are all weakly supported.

Analyses of the two series of matrices with and without edited sites in most cases produced identical topologies in major relationships with identical or slightly different (<10%) bootstrap support for most nodes (Figs. 2B, 2C, 2D, 3). Only in analyses of the 4-cp + nu18S gene and the 4-cp gene matrices, >10% bootstrap differences were seen on the nodes that are either controversial (the placement of *Amborella*) or weakly supported (the relationships among monocots, magnoliids, *Ceratophyllum*, and eudicots) (Figs. 2B, 2C). In analyses of the 8-gene matrix, the RNA editing sites had effect on topology, but again on parts of the tree that were only weakly supported (the placement of *Ceratophyllum* and the relationships among monocots, magnoliids, and eudicots) (Figs. 1, 2A).

DISCUSSION

Does *Amborella* or *Amborella* + Nymphaeales represent the first diverging lineage of angiosperms? — Since *Amborella*, Nymphaeales, and Austrobaileyales were identified as modern representatives of the earliest diverging lineages of angiosperms, several studies have raised the issue of exact taxonomic composition of the first diverging lineage of angiosperms. *Amborella* + Nymphaeales, and to a less extent Nymphaeales alone, have been found to be statistically undistinguishable alternatives to *Amborella* as the sister to all other angiosperms (Parkinson & al., 1999; Barkman & al., 2000; Graham & Olmstead, 2000; Qiu & al., 2000, 2005; Zanis & al., 2002; Hilu & al., 2003; Stefanovic & al., 2004; Leebens-Mack & al., 2005). The *Amborella* + Nymphaeales-basal topology is usually found with likelihood, Bayesian, or neighbor-joining analyses, and often with mitochondrial gene-dominant data sets.

In this study, we uncovered two pieces of evidence that provide significant support to the *Amborella* + Nymphaeales-basal hypothesis. One is that in analyses of all matrices except the 3-mt gene and the 4-cp + 3-mt gene matrices, the likelihood analyses uniformly obtained much lower bootstrap support to the *Amborella*-basal topology than the parsimony analyses (Figs. 1, 2A, 2B, 2C). The other piece of evidence comes from the likelihood analyses of the 3-mt gene and the 4-cp + 3-mt gene matrices, which showed that *Amborella* + Nymphaeales were sister to all other angiosperms with moderate to high bootstrap support (note that the high bootstrap support is from analyses of the 3-mt gene matrix when the chloroplast genes were excluded; Figs. 2D, 3). Because parsimony methods are more susceptible to long branch attraction than likelihood methods (Felsenstein, 1978) and the three mitochondrial genes used here have significantly lower substitution rates and thus much less homoplasy than the four chloroplast genes (data not shown, also see J. Davis & al., 1998), it is likely that the strongly supported topology of *Amborella* alone as the sister to all other angiosperms is a result of *local* long branch attraction. In this case, the divergent *Amborella* sequences might have been attracted away from Nymphaeales and instead to gymnosperms, even though placement of *Amborella*, Nymphaeales, and Austrobaileyales together at the base of angiosperm phylogeny has been shown to be free of long branch attraction effects (Qiu & al., 2001, 2005).

When the results from analyses of the 4-cp gene, the 4-cp + 3-mt gene, and the 3-mt gene matrices are compared (Figs. 2B, 2D, 3), it is clear that the mitochondrial genes support the *Amborella* + Nymphaeales-basal topology whereas the chloroplast genes favor the *Amborella*-basal topology. It is evident that the four chloroplast genes, particularly *matK* and *rpoC2*, contain enough noisy information that renders the likelihood method vulnerable to long branch attraction, since when the mitochondrial genes were not included, likelihood analyses of both 4-cp gene and 4-cp + nu18S gene matrices recovered the *Amborella*-basal topology with >50% bootstrap support (Figs. 2B, 2C). Because the mitochondrial genes have the lineage-specific rate heterogeneity problem, which caused misplacement of *Acorus*, alismatids, and *Cyperus* when these taxa were included in the 3-mt gene matrix (data not shown), it is reasonable to ask whether the *Amborella* + Nymphaeales-basal topology favored by the mitochondrial genes could be caused by this problem. The evidence against this possibility is that likelihood analyses of the 4-cp gene and the 4-cp + nu18S gene matrices recovered much lower bootstrap support than the parsimony analyses for the *Amborella*-basal topology (Figs. 2B, 2C). Our earlier parsimony bootstrapping analysis of a mitochondrial gene data set including *atp1*, *matR*, and SSU and LSU rRNA genes recovered 67% bootstrap support for the *Amborella*-basal topology (Qiu & al., 2005). We now think that both susceptibility of parsimony methods to long branch attraction and presence of poorly aligned sequences in the mitochondrial SSU and LSU rRNA genes might have been responsible for that result.

From the evidence presented here and reported in previous studies (Parkinson & al., 1999; Barkman & al., 2000; Graham & Olmstead, 2000; Qiu & al., 2000, 2005;

Zanis & al., 2002; Hilu & al., 2003; Stefanovic & al., 2004; Leebens-Mack & al., 2005), it is sufficiently clear that the first diverging lineage of extant angiosperms consists of *Amborella* + Nymphaeales. Future studies could target sequencing slow-evolving chloroplast genes and more mitochondrial genes to evaluate further the two competing hypotheses (the Nymphaeales-basal topology was not recovered in any of our analyses). Resolution of the issue of exact taxonomic composition of the first diverging lineage of angiosperms has profound implications on understanding the habitat in which angiosperms originated and diversified, and will influence future studies on origin and early evolution of angiosperms. An aquatic/wet habitat origin, or more likely early diversification in such an environment, of angiosperms deserves to be seriously considered given the following reasons. First, the taxonomic composition of the first diverging lineage of angiosperms now presents an equal probability of aquatic or terrestrial ancestry of this lineage, pushing the aquatic growth habit one node lower than the *Amborella*-basal topology. Second, aquatic/semi-aquatic plants have disproportionately high frequency of occurrence in Nymphaeales, basal monocots, Chloranthaceae, Piperales, *Ceratophyllum*, and basal eudicots (ranunculids and *Nelumbo*) (Cronquist, 1981; Xia & Brach, 1994; Xia & Jeremie, 1994). Even non-aquatic species in several basal lineages such as *Amborella*, Austrobaileyales, and Chloranthaceae occur in environments with high rainfall without appreciable dry season (Feild & al., 2003). Third, aquatic/streamside plants figure conspicuously among the best preserved, earliest documented angiosperm fossils, e.g., A*rchaefructus liaoningensis*, *A. sinensis*, and *A. eoflora*, and fossil water lilies (Friis & al., 2001; Sun & al., 2002; Ji & al., 2004). Finally, an aquatic origin or early diversification of angiosperms may better explain than some of the earlier hypotheses (Stebbins, 1965, 1974; Hickey & Doyle, 1977) how angiosperms during their earliest existence could have avoided competition with gymnosperms, most of which grew in and were well adapted to terrestrial environments.

Monophyly of the magnoliids and sister relationships between Magnoliales and Laurales, and between Canellales and Piperales. — Although magnoliids were assembled as a complex of taxa representing basal angiosperms nearly two centuries ago (see Qiu & al., 1993), the monophyly of magnoliids as currently circumscribed and relationships among their member lineages were not established until very recently (Qiu & al., 1999, 2000). Moreover, there is still lack of morphological or phytochemical characters to define the group and to determine relationships among its member lineages (Doyle & Endress, 2000). Hence, as in the case of identifying *Amborella*, Nymphaeales, and Austrobaileyales being the earliest diverging lineages of extant angiosperms, molecular phylogenetic studies are left alone to demonstrate that their results do not suffer from any analytical distortion and that sufficient support can be obtained to gain confidence on these results when little corroborative evidence is available from morphological studies.

In three earlier large-scale angiosperm phylogenetic studies, neither monophyly of magnoliids nor relationships among their four member lineages as currently defined were recognized (Chase & al., 1993; Savolainen & al., 2000; Soltis & al., 2000). These results were probably due to lack of sufficient signal in the one to three slowly evolving genes (chloroplast *atpB* and *rbcL*, and nuclear 18S) used, as a later, large-scale study of faster evolving *matK* did recover both results (Hilu & al., 2003). The relatively taxon-dense analyses of basal angiosperms with five chloroplast (*atpB*, *rbcL*), mitochondrial (*atp1*, *matR*), and nuclear (18S) genes by Qiu & al. (1999, 2000) first identified the sister relationships between Magnoliales and Laurales, and between Canellales and Piperales, as well as monophyly of the magnoliids consisted of these four lineages, but with only low to moderate bootstrap support. In this study, these relationships are all strongly supported in the 8-gene analyses (Figs. 1, 2A). Several pieces of evidence from this and other studies support that these relationships reflect accurately the evolutionary history of this group of plants. First, chloroplast and mitochondrial gene analyses recovered these relationships independently (Figs. 2B, 3). Second, likelihood and parsimony analyses of the 8-gene matrix and its various partitions obtained similar bootstrap support for these results (Figs. 1–3). Third, bootstrap values for these relationships increased steadily when chloroplast, mitochondrial, and nuclear genes were combined in this study and others that had similar taxon sampling schemes (Figs. 1, 2A, 2C, 2D; Nickrent & al., 2002; Zanis & al., 2002; Qiu & al., 2005). Fourth, several other studies employing many different chloroplast, mitochondrial, and nuclear genes and using relatively sparse taxon sampling in basal angiosperms (Mathews & Donoghue, 1999, 2000; Barkman & al., 2000; Graham & Olmstead, 2000; Borsch & al., 2003, 2005; Löhne & Borsch, 2005; Graham & al., 2006) and one large-scale analysis of angiosperms using *matK* (Hilu & al., 2003) all recovered these same results. Fifth, an insertion was found in a group II intron in the chloroplast gene *petD* that defines the magnoliids (Löhne & Borsch, 2005). Finally, although the cladistic analysis of largely morphological characters by Doyle & Endress (2000) did not identify any synapomorphy for the magnoliids, Kubitzki (1993) pointed out that there are several morphological and phytochemical features or a combination of them that suggest a link between Piperales (as defined here includ-

ing Aristolochiaceae and *Lactoris*) and Magnoliales, Laurales, and Canellales.

Firmly establishing monophyly of the magnoliids has important implications for understanding pattern and process of the initial radiation of angiosperms right after their origin. According to this new topology, nearly all living angiosperms can be grouped into three large clades, magnoliids, monocots, and eudicots, which contains 3%, 22%, and 75% of all angiosperm species, respectively (Herendeen & Crane, 1995). Only a few to dozens of species are found in each of the other four clades, *Amborella*-Nymphaeales, Austrobaileyales, Chloranthaceae, and *Ceratophyllum*. For a long time, magnoliids were regarded as a paraphyletic group at base of the angiosperm phylogeny (Takhtajan, 1969; Cronquist, 1981; Donoghue & Doyle, 1989; Qiu & al., 1993). With five groups (*Amborella*-Nymphaeales, Austrobaileyales, Chloranthaceae, *Ceratophyllum*, ranunculids) removed, the core of magnoliids stay together as a monophyletic group. Thus, the bushy appearance of the angiosperm tree seen in the earlier phylogenetic reconstructions (Donoghue & Doyle, 1989; Qiu & al., 1993) is greatly reduced. With only three major and at most four minor clades placed at base of angiosperm phylogeny, the initial radiation of angiosperms after their origin might not have been as explosive as originally thought. The dramatic increase of diversification rates in angiosperms did not seem to occur until well after emergence of magnoliids, monocots, and eudicots, and was probably not correlated with the so-called key innovations of angiosperms. Earlier studies that specifically examined causative relationship between key innovations and diversification rates in angiosperms had reached the same conclusion (Sanderson & Donoghue, 1994; Davies & al., 2004). The lack of clear morphological or phytochemical synapomorphy for magnoliids (Doyle & Endress, 2000) may be invoked to argue for a rapid radiation scenario, but an equally plausible hypothesis could be that the magnoliids originated in the ancestral niche that the earliest angiosperms had occupied and there was little selective pressure that resulted in fixation of many new adaptive changes. The fact that *Amborella*, Austrobaileyales, Chloranthaceae, many basal eudicots, and most magnoliids occur in similar environments can perhaps be taken as the evidence to support this hypothesis. More support for this idea comes from examination of monocots, which are the only one of three major angiosperm clades that possess numerous morphological synapomorphies, e.g., a single cotyledon, fibrous roots, parallel leaf venation, and scattered vascular bundles in the stem. These traits probably represent adaptations to the aquatic/swamp habitats in which monocots likely originated, as can be inferred from the habitats of many modern basal monocots (also see Cronquist, 1981).

Identification of the magnoliids as a major clade of angiosperms has further implications on interpreting fossil record and assessing the effect of extinction on shaping the extant angiosperm diversity. A large diversity of angiosperm pollen, leaves, and floral structures have been found in the Early Cretaceous (Doyle, 1969; Doyle & Hickey, 1976; Upchurch, 1984; Walker & Walker, 1984; Friis & al., 1994, 1999, 2000, 2001; Sun & al., 2002; Ji & al., 2004), but how they fit in a phylogeny reconstructed with extant angiosperms is subject to debate. Establishing monophyly of the magnoliids should help to search for phylogenetic relationships of many fossil taxa and to identify extinct lineages by reconstructing the archetype of magnoliids and ancestral character states at major internal nodes within this clade. Given that many of the fossils possess features seen on extant magnoliids, it is doubtful that any major clade of angiosperms at the level of magnoliids, monocots, and eudicots has been extinct. Instead, extinction probably occurred within these clades and might have also been responsible for loss of lineages at the level of the four minor clades, *Amborella*-Nymphaeales, Austrobaileyales, Chloranthaceae, and *Ceratophyllum*. The effect of extinction on shaping the extant angiosperm diversity may not have been so dramatic. On the other hand, if the current difficulty of resolving relationships among the three major clades and Chloranthaceae and *Ceratophyllum* is partly caused by extinction, morphological cladistic analyses such as those performed by Sun & al. (2002) and Eklund & al. (2004) should be conducted to incorporate information from fossil taxa. This kind of analyses will provide a desperately needed, independent evaluation of relationships among basal angiosperms that are currently inferred almost purely with molecular data.

Placement of Piperales securely as sister to Canellales within the magnoliids also narrows the search scope for relationship of monocots after Nymphaeales and ranunculids were excluded (Qiu & al., 2000). Previously, Burger (1977) had suggested that similar features seen in Arales and Piperales might indicate a piperalean origin of monocots. The morphological cladistic analysis by Doyle & Endress (2000) provided further support to that hypothesis. The topology obtained here indicates that the similar features in Piperales and basal monocots likely resulted from independent adaptations to the aquatic/swamp environment in which both groups of plants grow.

Finally, establishing monophyly of the magnoliids in this multigene analysis provides an excellent example to substantiate a long-held belief of systematists, namely, to pursue all kinds of evidence as they become available to reconstruct the evolutionary history of organisms. Among the three major clades of angiosperms, monocots were recognized as a natural group before modern tax-

onomy thanks to their numerous morphological and anatomical diagnostic features. That eudicots may represent a natural lineage was first noticed in surveys of angiosperm pollen diversity using light and electronic microscopic techniques (Wodehouse, 1935, 1936; Bailey & Nast, 1943; Hu, 1950; Walker & Doyle, 1975) and then confirmed in cladistic analyses of morphological data on extant and fossil taxa (Crane 1989; Donoghue & Doyle 1989; Doyle & Hotton, 1991). The large-scale molecular analysis of angiosperms by Chase & al. (1993) finally established monophyly of this large group of angiosperms beyond any doubt. With regard to monophyly of the magnoliids, it was not revealed in early molecular phylogenetic studies that had low information content (Chase & al., 1993; Qiu & al., 1993; Soltis & al., 1997; Savolainen & al., 2000). As data increased, the clade was often identified but with only low to moderate bootstrap support (Qiu & al., 1999, 2000; Zanis & al., 2002; Hilu & al., 2003). This clade is now firmly established with much large data sets (Figs. 1–3; Borsch & al., 2005). This continuous progress made toward unveiling the evolutionary history of angiosperms as an increasing amount of information becomes available for phylogenetic inference should encourage future studies to sample more data sources, from both molecules and morphology, to solve the remaining problems in plant phylogenetics.

Relationships among Chloranthaceae, *Ceratophyllum*, monocots, magnoliids, and eudicots. — The current difficulty of resolving relationships among the five euangiosperm lineages is likely due to rapid radiation and extinction experienced by angiosperms right after their origin (Doyle, 1969; Doyle & Hickey, 1976; Friis & al., 1999, 2000). However, as the number of genes is increased in the series of analyses we have performed, more relationships have been resolved with steady increase of bootstrap support (Qiu & al., 1993, 1999, 2005, this study). Thus, it can be expected that with more data these relationships will be resolved eventually. In fact, the analyses performed in this study recovered only a very limited set of relationships out of all possible arrangements among the five lineages.

Chloranthaceae, though a small family, exhibit a disproportionately large diversity in vegetative and reproductive morphology (Endress, 1987), and are one of the few extant angiosperm families that have a traceable history back to the Early Cretaceous (Upchurch, 1984; Walker & Walker, 1984; Friis & al., 1986, 1994, 1999; Pedersen & al., 1991). Extinction has clearly occurred within or along this lineage, as can be inferred from a comparison between extant and fossil chloranthoid taxa for their morphological diversity (Friis & al., 1986, 1994, 1999; Pedersen & al., 1991; Eklund & al., 2004). In the morphological cladistic analysis of basal angiosperms by Doyle & Endress (2000), Chloranthaceae were placed between *Amborella* and other angiosperms. In our analyses of all matrices except the 3-mt gene matrix, this family is shown to be sister to all other euangiosperms (Figs. 1, 2B, 2C, and 2D), and occasionally together with *Ceratophyllum* in such a position (Fig. 2A). Even in the analyses of the 3-mt gene matrix, Chloranthaceae are still very deep among all euangiosperms (Fig. 3). Nevertheless, all these positions have very low bootstrap support. If stratigraphic information is used to evaluate phylogenetic hypothesis derived from study of extant plants, occurrence of chloranthoid taxa together with other angiosperm fossils that cannot be easily assigned to individual extant lineages in the Early Cretaceous would certainly favor a topology in which Chloranthaceae branch out early in the euangiosperm phylogeny. The sister relationship between this family and monocots, with moderate jackknife support, shown in the *matK* analysis by Hilu & al. (2003) was not recovered in this study, even though *matK* was one of the eight genes we used.

Ceratophyllum exhibits highly divergent and reduced morphology, such as branched pollen tubes, lack of roots, and loss of stomata (Les, 1988). In the morphological cladistic analysis by Doyle & Endress (2000), its position was highly unstable. Molecular analyses have placed the genus with eudicots (Soltis & al., 2000; Hilu & al., 2003; Qiu & al., 2005; Graham & al., 2006; with 53–82% bootstrap or jackknife support), monocots (Nickrent & al., 2002; Zanis & al., 2002; Borsch & al., 2005; Qiu & al., 2005; with 57–88% bootstrap or jackknife support), or Chloranthaceae (Fig. 3; with 61–69% bootstrap support). Of these three placements, the one with monocots is mostly seen in the mitochondrial gene-dominant analyses. The divergent mitochondrial gene sequences of *Acorus* and alismatids used in those studies might have caused long branch attraction (Qiu & al., 2005). The placement with eudicots appears in the chloroplast gene-dominant analyses, and this relationship has recently been codified by The Angiosperm Phylogeny Group (2003) in an angiosperm classification system. Our earlier parsimony analysis of five chloroplast (*atpB*, *matK*, and *rbcL*) and mitochondrial (*atp1* and *matR*) protein-coding genes recovered this relationship with 74% bootstrap support (Qiu & al., 2005). It was partly based on this result that we decided to add fast-evolving *rpoC2* and to sample more eudicots and monocots in this study, anticipating higher bootstrap support for this relationship. On the contrary, the bootstrap support decreased sharply or disappeared entirely in our analyses of the 8-gene and the 4-cp gene matrices (Figs. 1, 2A, 2B). Moreover, our likelihood analyses of the 3-mt gene matrix, with divergent sequences of *Acorus*, alismatids, and *Cyperus* excluded, recovered the sister relationship between *Ceratophyllum* and Chloranthaceae with 61–69% bootstrap support (Fig. 3). In light of these

results and the general situation in which different placements of *Ceratophyllum* have been obtained in analyses that differ in gene choices, taxon sampling schemes, and analytic methods—in stark contrast to the robust consensus reached on the magnoliids in most studies discussed above, we consider it premature to conclude that this genus is sister to eudicots as done by The Angiosperm Phylogeny Group (2003).

How magnoliids, monocots, and eudicots are related to each other represents one of the most difficult problems in plant phylogenetics. In this study, all three possible arrangements among these three clades were recovered with varying but uniformly low bootstrap values, depending on the genes used, and inclusion or exclusion of the RNA editing sites (Figs. 1–3). Previously, magnoliids and monocots were placed in a clade together with Chloranthaceae to the exclusion of eudicots with 56% jackknife support by Soltis & al. (2000). Eudicots and magnoliids were found to be sister to each other with 56 and 78% bootstrap support by Zanis & al. (2002) and Qiu & al. (2005), respectively, and Nickrent & al. (2002) placed these two group together with Chloranthaceae with 55% bootstrap support. The third topology, monocots being sister to eudicots, was recovered by Graham & al. (2006), and it had bootstrap support of 67%. If stratigraphic information can be used to distinguish these three hypotheses, and if the current fossil evidence reflects the history of early angiosperm evolution without strong bias, the first topology should be excluded, as pollen, leaves, and floral structures of magnoliids and monocots predate those of eudicots in the fossil record (Doyle, 1969; Doyle & Hickey, 1976; Upchurch, 1984; Walker & Walker, 1984; Friis & al., 1994, 1999, 2000, 2004; von Balthazar & al., 2005). The late appearance of tricolpate pollen in comparison to monosulcate angiosperm pollen in the fossil record (Doyle, 1969; Friis & al., 1999) and strictly restricted occurrence of this type of pollen within eudicots among all living angiosperms (Wodehouse, 1935; Walker & Doyle, 1975) are in agreement with an interpretation that eudicots evolved after magnoliids and monocots. To resolve these relationships with high level of confidence, however, it seems necessary to design future studies to sample more genes from both chloroplast and mitochondrial genomes with different evolutionary rates (see below). Meanwhile, since our knowledge of morphology on extant and fossil angiosperms has increased significantly over the last two decades (Upchurch, 1984; Walker & Walker, 1984; Endress & Hufford, 1989; Friis & al., 1994, 1999, 2000; Endress & Igersheim, 2000; Floyd & Friedman, 2000; Mohr & Friis, 2000; Sampson, 2000; Thien & al., 2000, von Balthazar & al., 2005), comprehensive cladistic analyses incorporating information on both extant and well reconstructed extinct lineages may provide an independent perspective to complement the molecular studies. Only when congruence between molecular and morphological analyses is ultimately achieved, can the relationships among these three major clades of angiosperms be considered satisfactorily resolved.

The roles of mitochondrial genes in reconstructing the basal angiosperm phylogeny and the plant phylogeny in general. — Basal angiosperms are one of the plant groups where mitochondrial genes have been used prominently for phylogenetic reconstruction (Parkinson & al., 1999; Qiu & al., 1999, 2005; Barkman & al., 2000; this study), and represent one of the groups whose phylogeny was revealed largely through analyses of many chloroplast, mitochondrial, and nuclear genes using different methods over the last fifteen years. Hence, they are an excellent model to examine empirically merits and demerits of genes from different genomic compartments for resolving difficult phylogenetic problems.

One of the salient features of plant mitochondrial genes is their generally low rates of substitution, in comparison to those from the chloroplast and the nucleus (Palmer, 1992). This feature could potentially result in low information content when mitochondrial genes alone are used. However, if these genes are used in combination with faster evolving chloroplast genes, their demerit will be masked and the merit of yielding less homoplasious characters becomes manifested (J. Davis & al., 1998). In a study that identified *Amborella*, Nymphaeales, and Austrobaileyales as the earliest diverging lineages of angiosperms (Qiu & al., 1999), which sampled mitochondrial *atp1* and *matR*, together with two chloroplast genes, *atpB* and *rbcL*, and nuclear 18S rRNA gene, this point was made clear. Use of the three chloroplast and nuclear genes or two chloroplast genes in two analyses sampling a much larger number of species only obtained the same result with lower jackknife support (Soltis & al., 2000), or did not even obtain this result (Savolainen & al., 2000). In resolving the issue of whether *Amborella* or *Amborella* + Nymphaeales together constitute the first diverging-lineage of angiosperms as shown in this study, the merit of mitochondrial genes being slow-evolving and thus less susceptible to long branch attraction was again demonstrated.

The generally low substitution rates of mitochondrial genes can be a problem when they are used alone. In the analysis of three mitochondrial genes of basal angiosperms performed here, some parts of the tree exhibit erratic topologies in comparison to those from the 8-gene and the 4-cp gene analyses (Fig. 3). A remedy to this problem will be to sequence faster evolving mitochondrial genes. One such candidate is *rps3*, which shows a substitution rate significantly faster than *atp1*, the fastest of three mitochondrial genes used in this study.

The other problem of mitochondrial genes with regard to rates concerns their lineage-specific rate heterogeneity, which has been reported in two groups of angiosperms, *Plantago* (Cho & al., 2004) and *Pelargonium* (Parkinson & al., 2005), and is detected in three more groups here, *Acorus*, alismatids, and *Cyperus*. The extent of rate heterogeneity is so dramatic that the maximum likelihood method used here could not avoid long branch attraction. Thus, the only way to deal with the problem is to remove the sequences from the analysis, as was done in this study. However, based on our sampling of five mitochondrial genes, three used here plus SSU and LSU rRNA genes, from hundreds of land plants, the problem of rate heterogeneity occurs only sporadically and in most cases has narrow phylogenetic distribution, restricted to at most a few families such as alismatids, conifers, and Gnetales (Qiu & al., 2005, in press). For protein-coding genes in particular, this problem is not serious enough to conclude that mitochondrial genes should not be used for phylogenetic reconstruction.

RNA editing sometimes occurs frequently in plant mitochondrial (Steinhauser & al., 1999; Knoop, 2004) and chloroplast genomes (Kugita & al., 2003; Suzuki & al., 2005). It has gene- and lineage-specific distribution patterns, and its evolutionary mechanism is still poorly understood. Although a previous study contended that RNA editing should not have any effect on phylogenetic inference (Bowe & dePamphilis, 1996), few phylogenetic studies have directly compared the results with and without edited sites. Among the three mitochondrial genes used in this study, *nad5* has the most edited sites, 19 in 1248 (1.5%) nucleotides that we sequenced (Table 2). In the single gene analysis of *nad5*, *Buxus* and several monocots were grossly mis-placed when the edited sites were included, placed as a grade between gymnosperms and other angiosperms. When the edited sites were removed, *Buxus* was grouped with other two genera of Buxaceae, *Pachysandra* and *Didymeles*, and the *Amborella*-Nymphaeales were placed at the base of angiosperm phylogeny again (data not shown). In analyses of the 8-gene matrix, inclusion and exclusion of the edited sites resulted in different topologies, but only on weakly supported relationships (Figs. 1, 2A). In analyses of various partitions of the 8-gene matrix, however, inclusion and exclusion of the edited sites did not have any influence on topology but only on the level of bootstrap support (Figs. 2B, 2C, 2D, 3). It is especially noteworthy that in the analyses of the 3-mt gene matrix, inclusion and exclusion of the edited sites affected relationships within monocots and eudicots, presumably because these three genes had lower information content than the chloroplast genes, taxon sampling in these two groups was relatively sparse in comparison to other groups, and *nad5* and *matR* have significantly more edited sites than the chloroplast genes (Table 2). It seems that by combining different genes from the same or different genomes, independent origins of RNA editing events that produced lineage- and gene-specific editing distribution patterns can be properly sorted out on the tree. This phenomenon confirms a suggestion made by Källersjö & al. (1999) that homoplasy can increase tree structure in phylogenetic reconstruction. Therefore, we conclude that RNA editing does not pose a problem for use of mitochondrial (and chloroplast) genes in phylogenetic studies, but it is important to be aware of its existence, especially when weakly supported and conflicting relationships among different data partitions are investigated.

The abundant occurrence of group II introns in the land plant mitochondrial genome should not pose any problem for their use in phylogenetic studies, as sources of either presence/absence type of information or sequence characters. Their stable inheritance and infrequent structural changes (*cis*- to *trans*-splicing) warrant the high quality of information they can generate in terms of gains and losses and evolution of *trans*-splicing (Qiu & al., 1998; Malek & Knoop, 1998; Dombrovska & Qiu, 2004; Qiu & Palmer, 2004). Sequencing of intron regions also provides sufficient variation for resolving lower level phylogenetic relationships (Sanjur & al., 2002; Guo & Ge, 2005), just as appropriate chloroplast and nuclear markers chosen for solving similar types of problems. Hence, no particular prejudice should be held against mitochondrial introns in this regard.

Sometimes frequent losses and occasional horizontal transfers of mitochondrial genes from the genome could seriously undermine their phylogenetic utility (Adams & al., 2002; Bergthorsson & al., 2003; Won & Renner, 2003; C. Davis & Wurdack, 2004). A remedy to this problem would be to choose genes that have low tendency of loss or horizontal transfer, and published surveys such as that of Adams & al. (2002) can be used for this purpose. Our sampling of *atp1*, *matR*, *nad5*, and SSU and LSU rRNA genes from hundreds of land plants have uncovered only one potential case of inter-genome transfer and no inter-organism horizontal transfer at all (Qiu & al., 2005, in press, and this study). Hence, the problem is clearly not as serious as originally thought.

The above discussion on features of mitochondrial genes clearly indicates that they can be useful phylogenetic markers. A more important, somewhat philosophical reason why mitochondrial genes should be explored for phylogenetic use is related to the goal of organismal phylogenetics, which is to estimate evolutionary history of organisms by retrieving information stored in their genomes. In the case of plants, there are three of them, and the chloroplast genome seems to be the most suitable for most studies. However, as shown in this and previous studies, use of mitochondrial genes together with chloro-

plast and nuclear genes has significantly advanced our understanding on several key issues of the basal angiosperm phylogeny. First, identification of the earliest-diverging lineage of angiosperms benefited significantly from examining information from all three, not just one, plant genomes (Chase & al., 1993; Qiu & al., 1993, 1999, 2005; Mathews & Donoghue, 1999; Parkinson & al., 1999; Barkman & al., 2000; Graham & Olmstead, 2000; Soltis & al., 2000). In this particular case, the analyses of chloroplast genes with parsimony methods produced misleading results more than once, suggesting first *Ceratophyllum* and then *Amborella* alone as the earliest angiosperm. It was the inclusion of mitochondrial genes and the combined use of parsimony and likelihood methods that identified both *Amborella* and Nymphaeales as members of the first diverging lineage of extant angiosperms. Second, placement of *Ceratophyllum* as sister to eudicots, which is already codified by The Angiosperm Phylogeny Group (2003) based on chloroplast gene-dominant analyses (Soltis & al., 2000; Hilu & al., 2003; Graham & al., 2006), is now challenged in light of the mitochondrial gene analysis performed in this study (Fig. 3). Finally, recognition of magnoliids as a major clade of angiosperms and resolution of relationships among their four member lineages receive a significant increase of support, when independent analyses of mitochondrial and chloroplast genes recovered the same topology with similar levels of support and a combined analyses of both groups of genes obtained even higher bootstrap support (Figs. 2B, 2D, 3). As shown in Qiu & al. (2005), the number of informative sites supporting resolution of a particular deep node like the one between *Amborella*-Nymphaeales-Austrobaileyales and other angiosperms can be very small (71 out of 26,990 nucleotide positions). It is dangerous to draw conclusion based on data from one genome. This point is especially relevant in the current state of molecular systematics as most deep relationships in the angiosperm and land plant phylogenies are not resolved or have only weak support. Further, many relationships resolved in molecular phylogenies lack independent morphological corroboration, e.g., monophylies of eudicots and magnoliids, and placement of *Amborella*, Nymphaeales, and Austrobaileyales at the base of angiosperm phylogeny. It is a fact that chloroplast genes are excellent, information-rich molecules for tracing plant evolutionary history, but they are by no means perfect, problem-free phylogenetic markers. There is definitely a niche for mitochondrial genes, and in most cases for nuclear genes as well, to occupy in molecular systematics, even though they have their own limitations. Thus, we wish to emphasize that a sound molecular phylogenetic study should include many genes from, if possible, all three plant genomes. Only when congruent, robust results are obtained from all single genome analyses, can the organismal phylogeny be considered correctly reconstructed. Independent studies to obtain reproducible results have always been a pillar of scientific methods.

ACKNOWLEDGEMENTS

We thank C. C. Davis for stimulating discussion and N. Al-Bassam, J. C. Deering, Y. Duanmu, P. L. Haynes, and L. M. Minnema for technical assistance. YQ was supported by an Early Career award (DEB 0332298) and an ATOL grant (DEB 0431239) from the NSF.

LITERATURE CITED

Adams, K. L., Qiu, Y.-L., Stoutemyer, M. & Palmer, J. D. 2002. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer during angiosperm evolution. *Proc. Natl. Acad. Sci. U.S.A.* 99: 9905–9912.

Anderberg, A. A., Rydin, C. & Källersjö, M. 2002. Phylogenetic relationships in the order Ericales s.l.: Analyses of molecular data from five genes from the plastid and mitochondrial genomes. *Amer. J. Bot.* 89: 677–687.

Aoki, S., Uehara, K., Imafuku, M., Hasebe, M. & Ito, M. 2004. Phylogeny and divergence of basal angiosperms inferred from *APETALA3*- and *PISTILLATA*-like MADS-box genes. *J. Plant Res.* 117: 229–244.

Bailey, I. W. & Nast, C. G. 1943. The comparative morphology of the Winteraceae I. Pollen and stamens. *J. Arnold Arbor.* 24: 340–346.

Barkman, T. J., Chenery, G., McNeal, J. R., Lyons-Weiler, J., Ellisens, W. J., Moore, G., Wolfe, A. D. & dePamphilis, C. W. 2000. Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc. Natl. Acad. Sci. U.S.A.* 97: 13166–13171.

Barkman, T. J., Lim, S. H., Salleh, K. M. & Nais, J. 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. *Proc. Natl. Acad. Sci. U.S.A.* 101: 787–792.

Beckert, S., Muhle, H., Pruchner, D. & Knoop, V. 2001. The mitochondrial *nad2* gene as a novel marker locus for phylogenetic analysis of early land plants: a comparative analysis in mosses. *Mol. Phylog. Evol.* 18: 117–126.

Beckert, S., Steinhauser, S., Muhle, H. & Knoop, V. 1999. A molecular phylogeny of bryophytes based on nucleotide sequences of the mitochondrial *nad5* gene. *Pl. Syst. Evol.* 218: 179–192.

Bergthorsson, U., Adams, K. L., Thomason, B. & Palmer, J. D. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424: 197–201.

Borsch, T., Hilu, K. W., Quandt, D., Wilde, V., Neinhuis, C. & Barthlott, W. 2003. Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol.* 16: 558–576.

Borsch, T., Löhne, C., Müller, K., Hilu, K. W., Wanke, S.,

Worberg, A., Barthlott, W., Neinhuis, C. & Quandt, D. 2005. Towards understanding basal angiosperm diversification: recent insights using rapidly evolving genome regions. *Nova Acta Leopoldina* 92: 85–110.

Bowe, L. M. & dePamphilis, C. W. 1996. Effects of RNA editing and gene processing on phylogenetic reconstruction. *Mol. Biol. Evol.* 13: 1159–1166.

Bowe, L. M., Coat, G. & dePamphilis, C. W. 2000. Phylogeny of seed plants based on all three plant genomic compartments: extant gymnosperms are monophyletic and Gnetales are derived conifers. *Proc. Natl. Acad. Sci. U.S.A.* 97: 4092–4097.

Burger, W. C. 1977. The Piperales and the monocots. Alternative hypotheses for the origin of monocotyledonous flowers. *Bot. Rev.* 43: 345–393.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y.-L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, W. J., Karol, K. G., Clark, W. D., Hedren, M., Gaut, B. S., Jansen, R. K., Kim, K.-J., Wimpee, C. F., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q.-Y., Plunkett, G. M., Soltis, P. S., Swensen, S, Williams, S. E., Gadek, P. A., Quinn, C. J., Eguiarte, L. E., Golenberg, E., Learn, G. H., Jr., Graham, S. W., Barrett, S. C. H., Dayanandan, S. & Albert, V. A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Missouri Bot. Gard.* 80: 528–580.

Chaw, S.-M., Parkinson, C. L., Cheng, Y., Vincent, T. M. & Palmer, J. D. 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proc. Natl. Acad. Sci. U.S.A.* 97: 4086–4091.

Cho, Y., Mower, J. P., Qiu, Y.-L. & Palmer, J. D. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* 101: 17741–17746.

Cox, C. J., Goffinet, B., Shaw, A. J. & Boles, S. B. 2004. Phylogenetic relationships among the mosses based on heterogeneous Bayesian analysis of multiple genes from multiple genomic compartments. *Syst. Bot.* 29: 234–250.

Crane, P. R. 1989. Paleobotanical evidence on the early radiation of nonmagnoliid dicotyledons. *Pl. Syst. Evol.* 162: 165–191.

Cronquist, A. 1981. *An Integrated System of Classification of Flowering Plants.* Columbia University Press, New York.

Davies, T. J., Barraclough, T. G., Chase, M. W., Soltis, P. S., Soltis, D. E. & Savolainen, V. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 101: 1904–1909.

Davis, C. C. & Wurdack, K. J. 2004. Host-to-parasite gene transfer in flowering plants: Phylogenetic evidence from Malpighiales. *Science* 305: 676–678.

Davis, J. I., Simmons, M. P., Stevenson, D. W. & Wendel, J. F. 1998. Data decisiveness, data quality, and incongruence in phylogenetic analysis: An example from the monocotyledons using mitochondrial *atpA* sequences. *Syst. Biol.* 47: 282–310.

Davis, J. I., Stevenson, D. W., Petersen, G., Seberg, O., Campbell, L. M., Freudenstein, J. V., Goldman, D. H., Hardy, C. R., Michelangeli, F. A., Simmons, M. P., Specht, C. D., Vergara-Silva, F. & Gandolfo, M. A. 2004. Phylogeny of the monocots, as inferred from *rbcL* and *atpA* sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. *Syst. Bot.* 29: 467–510.

Dombrovska, O. & Qiu, Y.-L. 2004. Distribution of introns in the mitochondrial gene *nad1* in land plants: phylogenetic and molecular evolutionary implications. *Mol. Phylog. Evol.* 32: 246–263.

Donoghue, M. J. & Doyle, J. A. 1989. Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. Pp. 17–45 in: Crane, P. R. & Blackmore, S. (eds.), *Evolution, Systematics, and Fossil History of the Hamamelidae,* vol. 1. Clarendon Press, Oxford.

Doyle, J. A. 1969. Cretaceous angiosperm pollen of the Atlantic Coastal Plain and its evolutionary significance. *J. Arnold Arbor.* 50: 1–35.

Doyle, J. A. & Endress, P. K. 2000. Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. *Int. J. Plant Sci.* 161: S121–S153.

Doyle, J. A. & Hickey, L. J. 1976. Pollen and leaves from the mid-Cretaceous Potomac group and their bearing on early angiosperm evolution. Pp. 139–206 in: Beck, C. B. (ed.), *Origin and Early Evolution of Angiosperms.* Columbia University Press, New York.

Doyle, J. A. & Hotton, C. L. 1991. Diversification of early angiosperm pollen in a cladistic context. Pp. 169–195 in: Blackmore, S. & Barnes, S. H. (eds.), *Pollen and Spores.* Clarendon Press, Oxford.

Duff, R. J. & Nickrent, D. L. 1999. Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. *Amer. J. Bot.* 86: 372–386.

Eklund, H., Doyle, J. A. & Herendeen, P. S. 2004. Morphological phylogenetic analysis of living and fossil Chloranthaceae. *Int. J. Plant Sci.* 165: 107–151.

Endress, P. K. 1987. The Chloranthaceae: reproductive structures and phylogenetic position. *Bot. Jahrb. Syst.* 109: 153–226.

Endress, P. K. & Hufford, L. D. 1989. The diversity of stamen structures and dehiscence patterns among Magnoliidae. *Bot. J. Linn. Soc.* 100: 45–85.

Endress, P. K. & Igersheim, A. 2000. Gynoecium structure and evolution of basal angiosperms. *Int. J. Plant Sci.* 161: S211–S213.

Feild, T. S., Arens, N. C. & Dawson, T. E. 2003. The ancestral ecology of angiosperms: emerging perspectives from extant basal lineages. *Int. J. Plant Sci.* 164: S129–S142.

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

Floyd, S. K. & Friedman, W. E. 2000. Evolution of endosperm developmental patterns among basal flowering plants. *Int. J. Plant Sci.* 161: S57–S81.

Friis, E. M., Crane, P. R. & Pedersen, K. R. 1986. Floral evidence for Cretaceous chloranthoid angiosperms. *Nature* 320: 163–164.

Friis, E. M., Pedersen, K. R. & Crane, P. R. 1994. Angiosperm floral structures from the Early Cretaceous of Portugal. *Pl. Syst. Evol.* [Suppl.] 8: 31–49.

Friis, E. M., Pedersen, K. R. & Crane, P. R. 1999. Early

angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. *Ann. Missouri Bot. Gard.* 86: 259–296.

Friis, E. M., Pedersen, K. R. & Crane, P. R. 2000. Reproductive structure and organization of basal angiosperms from the Early Cretaceous (Barremian or Aptian) of western Portugal. *Int. J. Plant Sci.* 161: S169–S182.

Friis, E. M., Pedersen, K. R. & Crane, P. R. 2001. Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. *Nature* 410: 357–360.

Friis, E. M., Pedersen, K. R. & Crane, P. R. 2004. Araceae from the Early Cretaceous of Portugal: Evidence on the emergence of monocotyledons. *Proc. Natl. Acad. Sci. U.S.A.* 101: 16565–16570.

Goremykin, V. V., Hirsch-Ernst, K. I., Wölfl, S. & Hellwig, F. H. 2003. Analysis of the *Amborella trichopoda* plastid genome sequence suggests that *Amborella* is not a basal angiosperm. *Mol. Biol. Evol.* 20: 1499–1505.

Goremykin, V. V., Holland, B., Hirsch-Ernst, K. I. & Hellwig, F. H. 2005. Analysis of *Acorus calamus* chloroplast genome and its phylogenetic implications. *Molec. Biol. Evol.* 22: 1813–1822.

Graham, S. W. & Olmstead, R. G. 2000. Utility of 17 plastid genes for inferring the phylogeny of the basal angiosperms. *Amer. J. Bot.* 87: 1712–1730.

Graham, S. W., Zgurski, J. M., McPherson, M. A., Cherniawsky, D. M., Saarela, J. M., Horne, E. S. C., Smith, S. Y., Wong, W. A., O'Brien, H. E., Biron, V. L., Pires, J. C., Olmstead, R. G., Chase, M. W. & Rai, H. S. 2006. Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. Pp. 3–20 in: Columbus, J. T., Friar, E. A., Hamilton, C. W., Porter, J. M., Prince, L. M. & Simpson, M. G. (eds.), *Monocots: Comparative Biology and Evolution.* 2 vols. Rancho Santa Ana Botanic Garden, Claremont, California.

Groth-Malonek, M., Pruchner, D., Grewe, F. & Knoop, V. 2005. Ancestors of *trans*-splicing mitochondrial introns support serial sister group relationships of hornworts and mosses with vascular plants. *Mol. Biol. Evol.* 22: 117–125.

Gugerli, F., Sperisen, C., Büchler, U., Brunner, I., Brodbeck, S., Palmer, J. D. & Qiu, Y.-L. 2001. The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. *Mol. Phylog. Evol.* 21: 167–175.

Guindon, S. & Gascuel, O. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* 52: 696–704.

Guo, Y. L. & Ge, S. 2005. Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. *Amer. J. Bot.* 92: 1548–1558.

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

Herendeen, P. S. & Crane, P. R. 1995. The fossil history of the monocotyledons. Pp. 1–21 in: Rudall, P. J., Cribb, P. J., Cutler, D. F. & Humphries, C. J. (eds.), *Monocotyledons: Systematics and Evolution.* Royal Botanic Gardens, Kew.

Hickey, L. J. & Doyle, J. A. 1977. Early Cretaceous fossil evidence for angiosperm evolution. *Bot. Rev.* 43: 3–104.

Hiesel, R., Haseler, A. & Brennicke, A. 1994. Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 91: 634–638.

Hilu, K. W., Borsch, T., Müller, K., Soltis, D. E., Soltis, P. S., Savolainen, V., Chase, M. W., Powell, M. P., Alice, L. A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T. A. B., Rohwer, J. G., Campbell, C. S., Chatrou, L. W. 2003. Angiosperm phylogeny based on *matK* sequence information. *Amer. J. Bot.* 90: 1758–1776.

Hoot, S. B., Magallon, S. & Crane, P. R. 1999. Phylogeny of basal tricolpates based on three molecular data sets: *atpB*, *rbcL*, and 18S nuclear ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* 86: 1–32.

Hu, H.-H. 1950. A polyphyletic system of classification of angiosperms. *Science Record* 3: 221–230.

Ji, Q., Li, H., Bowe, L. M., Liu, Y. & Taylor, D. W. 2004. Early Cretaceous *Archaefructus eoflora* sp. nov. with bisexual flowers from Beipiao, western Liaoning, China. *Acta Geol. Sin.* 78: 883–896.

Källersjö, M., Albert, V. A. & Farris, J. S. 1999. Homoplasy increases phylogenetic structure. *Cladistics* 15: 91–93.

Kim, S. T., Yoo, M. J., Albert, V. A., Farris, J. S., Soltis, P. S. & Soltis, D. E. 2004. Phylogeny and diversification of B-function MADS-box genes in angiosperms: Evolutionary and functional implications of a 260-million-year-old duplication. *Amer. J. Bot.* 91: 2102–2118.

Knoop, V. 2004. The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* 46: 123–139.

Kubitzki, K. 1993. Introduction. Pp. 1–12 in: Kubitzki, K. (ed.), *The Families and Genera of Vascular Plants,* Vol. II (Kubitzki, K., Rohwer, J. G. & Bittrich, V., vol. eds.), *Flowering Plants, Dicotyledons: Magnoliid, Hamamelid and Caryophyllid Families.* Springer-Verlag, Berlin.

Kugita, M., Yamamoto, Y., Fujikawa, T., Matsumoto, T. & Yoshinaga K. 2003. RNA editing in hornwort plastids makes more than half the genes functional. *Nucleic Acids Res.* 31: 2417–2423.

Leebens-Mack, J., Raubeson, L. A., Cui, L. Y., Kuehl, J. V., Fourcade, M. H., Chumley, T. W., Boore, J. L., Jansen, R. K. & dePamphilis, C. W. 2005. Identifying the basal angiosperm node in chloroplast genome phylogenies: Sampling one's way out of the felsenstein zone. *Molec. Biol. Evol.* 22: 1948–1963.

Les, D. H. 1988. The origin and affinities of the Ceratophyllaceae. *Taxon* 37: 326–435.

Löhne, C. & Borsch, T. 2005. Molecular evolution and phylogenetic utility of the *petD* group II intron: a case study in basal angiosperms. *Molec. Biol. Evol.* 22: 317–332.

Malek, O., Lättig, K., Hiesel, R., Brennicke, A. & Knoop, V. 1996. RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO J.* 15: 1403–1411.

Malek, O. & Knoop, V. 1998. *Trans*-splicing group II introns in plant mitochondria: the complete set of *cis*-arranged homologs in ferns, fern allies, and a hornwort. *RNA* 4: 1599–1609.

Martin, P. G. & Dowd, J. M. 1991. Studies of angiosperm phylogeny using protein sequences. *Ann. Missouri Bot. Gard.* 78: 296–337.

Mathews, S. & Donoghue, M. J. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286: 947–950.

Mathews, S. & Donoghue, M. J. 2000. Basal angiosperm phy-

logeny inferred from duplicated phytochromes A and C. *Int. J. Plant Sci.* 161: S41–S55.

Mohr, B. A. R. & Friis, E. M. 2000. Early angiosperms from the Lower Cretaceous Crato Formation (Brazil), a preliminary report. *Int. J. Plant Sci.* 161: S155–S167.

Nei, M., Kumar, S. & Takahashi, K. 1998. The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. *Proc. Natl. Acad. Sci. U.S.A.* 95: 12390–12397.

Nickrent, D. L., Blarer, A., Qiu, Y.-L., Soltis, D. E., Soltis, P. S. & Zanis, M. 2002. Molecular data place Hydnoraceae with Aristolochiaceae. *Amer. J. Bot.* 89: 1809–1817.

Palmer, J. D. 1992. Mitochondrial DNA in plant systematics: applications and limitations. Pp. 36–49 in: Soltis, P. S., Soltis, D. E. & Doyle J. J. (eds.), *Molecular Systematics of Plants.* Chapman and Hall, New York.

Parkinson, C. L., Adams, K. L. & Palmer, J. D. 1999. Multigene analyses identify the three earliest lineages of extant flowering plants. *Curr. Biol.* 9: 1485–1488.

Parkinson, C. L., Mower, J. P., Qiu, Y.-L., Shirk, A. J., Song, K., Young, N. D., dePamphilis, C. W. & Palmer, J. D. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evol. Biol.* 5: 73.

Pedersen, K. R., Crane, P. R. & Friis, E. M. 1991. Fruits from the mid-Cretaceous of North America with pollen grains of the *Clavatipollenites* types. *Grana* 30: 577–590.

Posada, D. & Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Qiu, Y.-L., Chase, M. W., Les, D. H. & Parks, C. R. 1993. Molecular phylogenetics of the Magnoliidae: cladistic analyses of nucleotide sequences of the plastid gene *rbcL*. *Ann. Missouri Bot. Gard.* 80: 587–606.

Qiu, Y.-L., Cho, Y., Cox, J. C. & Palmer, J. D. 1998. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* 394: 671–674.

Qiu, Y.-L., Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V. & Chase, M. W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402: 404–407.

Qiu, Y.-L., Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V. & Chase, M. W. 2000. Phylogeny of basal angiosperms: analyses of five genes from three genomes. *Int. J. Plant Sci.* 161: S3–S27.

Qiu, Y.-L., Lee, J., Whitlock, B. A., Bernasconi-Quadroni, F. & Dombrovska, O. 2001. Was the ANITA rooting of the angiosperm phylogeny affected by long branch attraction? *Molec. Biol. Evol.* 18: 1745–1753.

Qiu, Y.-L. & Palmer, J. D. 1999. Phylogeny of basal land plants: insights from genes and genomes. *Trends Plant Sci.* 4: 26–30.

Qiu, Y.-L. & Palmer, J. D. 2004. Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J. Molec. Evol.* 59: 80–89.

Qiu, Y.-L., Dombrovska, O., Lee, J., Li, L., Whitlock, B. A., Bernasconi-Quadroni, F., Rest, J. S., Davis, C. C., Borsch, T., Hilu, K. W., Renner, S. S., Soltis, D. E., Soltis, P. S., Zanis, M. J., Cannone, J. J., Gutell, R. R., Powell, M., Savolainen, V., Chatrou, L. W. & Chase, M. W. 2005. Phylogenetic analysis of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* 166: 815–842.

Qiu, Y.-L., Li, L., Wang, B., Chen, Z., Dombrovska, O., Lee, J., Kent, L., Li, R., Jobson, R. W., Hendry, T. A., Taylor, D. W., Testa, C. M. & Ambros, M. A. In press. Non-flowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial and nuclear genes. *Int. J. Plant Sci.*

Renner, S. S. 1999. Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. *Amer. J. Bot.* 86: 1301–1315.

Sampson, F. B. 2000. Pollen diversity in some modern magnoliids. *Int. J. Plant Sci.* 161: S193–S210.

Sanderson, M. J. & Donoghue, M. J. 1994. Shifts in diversification rate with the origin of angiosperms. *Science* 264: 1590–1593.

Sanjur, O. I., Piperno, D. R., Andres, T. C. & Wessel-Beaver, L. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proc. Natl. Acad. Sci. U.S.A.* 99: 535–540.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W. & Le Thomas, A. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Bot. J. Linn. Soc.* 142: 125–186.

Savolainen, V., Chase, M. W., Hoot, S. B., Morton, C. M., Soltis, D. E., Bayer, C., Fay, M. F., de Bruijn, A. Y., Sullivan, S. & Qiu, Y.-L. 2000 Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcL* gene sequences. *Syst. Biol.* 49: 306–362.

Soltis, D. E., Soltis, P. S., Chase, M. W., Mort, M. E., Albach, D. C., Zanis, M., Savolainen, V., Hahn, W. H., Hoot, S. B., Fay, M. F., Axtell, M., Swensen, S. M., Prince, L. M., Kress, W. J., Nixon, K. C. & Farris, J. S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL* and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.

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., Swensen, S. M., Zimmer, E. A., Chaw, S.-M., Gillespie, L. J., Kress, W. J. & Sytsma, K. J. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* 84: 1–49.

Soltis, P. S., Soltis, D. E. & Chase, M. W. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.

Stebbins, G. L. 1965. The probable growth habit of the earliest flowering plants. *Ann. Missouri Bot. Gard.* 52: 457–468.

Stebbins, G. L. 1974. *Flowering Plants: Evolution above the Species Level.* Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Stefanovic, S., Rice, D. W. & Palmer, J. D. 2004. Long branch attraction, taxon sampling, and the earliest angiosperms: *Amborella* or monocots? *BMC Evol. Biol.* 4: Art. No. 35.

Steinhauser, S., Beckert, S., Capesius, I., Malek, O. & Knoop, V. 1999. Plant mitochondrial RNA editing. *J. Molec. Evol.* 48: 303–312.

Sun, G., Ji, Q., Dilcher, D. L., Zheng, S., Nixon, K. C. &

Wang, X. 2002. Archaefructaceae, a new basal angiosperm family. *Science* 296: 899–904.

Suzuki, K., Tuzi, S., Hasebe, M., Ueda, K., Nishiuchi, T., Nishiyama, T. & Yamaguchi, K. 2005 The plastid genome and RNA editing with a high frequency of *Selaginella uncinata*, a microphyllophyte. *J. Plant Res.* 118: S59 (abstract).

Swofford, D. L. 1998. *PAUP*4.0b10: Phylogenetic Analysis Using Parsimony.* Sinauer, Sunderland, Massachusetts.

Takhtajan, A. 1969. *Flowering Plants: Origin and Dispersal.* Oliver and Boyd, Edinburgh.

The Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Bot. J. Linn. Soc.* 141: 399–436.

Thien, L. B., Azuma, H. & Kawano, S. 2000. New perspectives on the pollination biology of basal angiosperms. *Int. J. Plant Sci.* 161: S225–S235.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24: 4876–4882.

Upchurch, G. R. 1984. Cuticle evolution in early Cretaceous angiosperms from the Potomac Group of Virginia and Maryland. *Ann. Missouri Bot. Gard.* 71: 522–550.

Vangerow, S., Teekorn, T. & Knoop, V. 1999. Phylogenetic information in the mitochondrial *nad5* gene of pteridophytes: RNA editing and intron sequences. *Plant Biol.* 1: 235–243.

von Balthazar, M., Pedersen, K. & Friis, E. M. 2005. *Teixeiria lusitanica*, a new fossil flower from the Early Cretaceous of Portugal with affinities to Ranunculales. *Pl. Syst. Evol.* 255: 55–75.

Walker, J. W. & Doyle, J. A. 1975. The bases of angiosperm phylogeny: palynology. *Ann. Missouri Bot. Gard.* 62: 664–723.

Walker, J. W. & Walker, A. G. 1984. Ultrastructure of Lower Cretaceous angiosperm pollen and the origin and early evolution of flowering plants. *Ann. Missouri Bot. Gard.* 71: 464–521.

Wang, X.-Q., Tank, D. C. & Sang, T. 2000. Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Molec. Biol. Evol.* 17: 773–781.

Wikström, N. & Pryer, K. M. 2005. Incongruence between primary sequence data and the distribution of a mitochondrial atp1 group II intron among ferns and horsetails. *Molec. Phylog. Evol.* 36: 484–493.

Won, H. & Renner, S. S. 2003. Horizontal gene transfer from flowering plants to *Gnetum*. *Proc. Natl. Acad. Sci. U.S.A.* 100: 10824–10829.

Wodehouse, R. P. 1935. *Pollen Grains: their Structure, Identification and Significance in Science and Medicine.* McGraw-Hill Book Co., Inc. New York.

Wodehouse, R. P. 1936. Evolution of pollen grains. *Bot. Rev.* 2: 67–84.

Xia, N. & Brach, A. R. 1994. Saururaceae. Pp. 108–110 in: Wu, Z.-Y. & Raven, P. H. (eds.), *Flora of China*. Science Press, Beijing and Missouri Botanical Garden, St. Louis, Missouri.

Xia, N. & Jeremie, J. 1994. Chloranthaceae. Pp. 132–138 in: Wu, Z.-Y. & Raven, P. H. (eds.), *Flora of China*. Science Press, Beijing and Missouri Botanical Garden, St. Louis, Missouri.

Zanis, M. J., Soltis, D. E., Soltis, P. S., Mathews, S. & Donoghue, M. J. 2002 The root of the angiosperms revisited. *Proc. Natl. Acad. Sci. U.S.A.* 99: 6848–6853.

Zanis, M. J., Soltis, P. S., Qiu, Y.-L., Zimmer, E. A. & Soltis, D. E. 2003. Phylogenetic analyses and perianth evolution in basal angiosperms. *Ann. Missouri Bot. Gard.* 90:129–150.

Appendix. Taxa and GenBank accession numbers of the sequences used in this study*. Species (when a different species is used for a particular gene, the name is shown in the parenthesis after the gene), and GenBank accession numbers of *atp1, matR, nad5, atpB, matK, rbcL, rpoC2,* 18S. New sequences generated in this study are labeled with asterisks, with voucher information provided in the GenBank accessions. Dashes indicate missing data.

Acorus calamus L., AF197621, — , DQ406951*, AJ235381, AB040154, M91625, DQ407038*, L24078; *Acorus gramineus* Soland., AF197622, — , DQ406994*, AF197616, AB040155, D28866, DQ407037*, AF197584; *A. attenuata* Salm-Dyck, AY299703 (*A. ghiesbreghtii* K. Koch), DQ401408*, DQ407009*, AF209521 (*A. ghiesbreghtii* K. Koch), DQ401364*, AF206729 (*A. ghiesbreghtii* K. Koch), DQ407146*, AF206841 (*A. ghiesbreghtii* K. Koch); *Akebia quinata* Decne., AF197642, AF197810, DQ406891*, AF209523, AB069851, L12627, DQ407030*, L31795; *Alisma plantago-aquatica* L., AF197717, AF197815, DQ406947*, DQ007417*, AB040179 (*A. canaliculatum* A. Br. & Bouche), L08759, DQ407035*, AF197585; *Allium cepa* L., DQ401321*, DQ401400*, DQ407007*, AF209525 (*A. altaicum* Pall.), AB017307 (*A. grayi* Regel), D38294, DQ407092*, AF168825 (*A. thunbergii* G. Don.); *Alstroemeria* sp., AF039254 (*A. caryophyllaea* Jacq.), AY453076, DQ407001*, DQ401335*, AY624481, AY120366 (*A. andina* Phil.), DQ407111*, — ; *Amborella trichopoda* Baill., DQ007412, AF197813, AY832180, AJ235389, NC_005086, L12628, DQ407043*, U42497; *Anemopsis californica* (Nutt.) Hook & Arn., AF197631, AF197747, DQ406979*, AF197608, AY437810, AF197597, DQ407110*, AF197576; *Annona muricata* L., AF197695, AF197766, DQ406917*, AJ235393, AF543722, L12629, DQ407047*, AF206850; *Arabidopsis thaliana* (L.) Heynh., NC_001284, NC_001284, NC_001284, NC_000932, NC_000932, NC_000932, NC_000932, NC_003071; *Aristolochia macrophylla* Lam., AF197669, AF197732, DQ406911*, AJ235399, AF543724 (*A. pistolochia* L.), L12630, DQ407095*, AF206855; *Asarum canadense* L., AF197671, AF197751, DQ406915*, U86383, AB060737 (*A. yakusimense* Masam.), L14290, DQ407096*, L24043; *Ascarina* sp., AF197667, AF197755, DQ406865*, AF239775 (*A. lucida* Hook. f.), DQ401338*, AF197592, DQ407136*, see Zanis et al., 2002 18S of this sp.; *Asparagus officinalis* L., AF197713, AF197736, DQ407000*, AJ235400, AB029805 (*A. filicinus* Buch.-Ham. ex. D. Don), L05028, DQ407080*, AF069205 (*A. falcatus* L.); *Atherosperma moschatum* Labill, AF197683, AF197799, DQ406929*, AF197604, DQ401343*, AF121362, DQ407100*, AF197578; *Austrobaileya scandens* C.T. White, AF197664, AF197742, DQ406986*, AJ235403, DQ401347*, L12632, DQ407130*, AF206858; *Beaucarnea (~Nolina) recurvata* Lem., DQ401301*, DQ401405*, DQ407008*, AF168888, AB089626 (*B. gracilis* Lem.), AB089639 (*B. gracilis* Lem.), DQ407104*, AF206971; *Berberidopsis beckleri* (F. Muell.) Veldkamp, DQ401303*, DQ401394*, DQ406898*, AJ235409 (*B. corallina* Hook. f.), DQ401358*, AJ235773 (*B. corallina* Hook. f.), DQ407158*, AF206866 (*B. corallina* Hook. f.); *Blandfordia grandiflora* R. Br., AY299727, DQ401412*, DQ406966*, AJ235412 (*B. punicea* Sweet), AB017315 (*B. punicea* Sweet), Z73694 (*B. punicea* Sweet), DQ407090*, AF206869 (*B. punicea* Sweet); *Brasenia schreberi* J. Gmelin, AF197640, AF197728, DQ406956*, AJ235418, AF092973, M77031, DQ407039*, AF206874; *Butomus umbellatus* L., AY299733, — , DQ406968*, AY147593, DQ401367*, AY149345, DQ407032*, L39846; *Buxus sempervirens* L., AF197636, AF197786, DQ406879*, AF092110, AF543728, AF093717, DQ407018*, L54065; *Cabomba caroliniana* A. Gray, AF197641 (C. sp.), AF197729 (C. sp.), DQ406957* (C. sp.), AF187058, AF543729, M77027, DQ407122* (C. sp.), AF206878; *Camellia japonica* L., AF420952 (*C. sinensis* Kuntze), AF421034 (*C. sinensis* Kuntze), DQ406870*, AF209553, AJ429305 (*C. sinensis* Kuntze), L12602, DQ407056*, U42815; *Cananga odorata* (Lam.) Hook. F. & Thomson, AF197700, AF197763, DQ406918*, DQ401322*, AY220438, L12636, DQ407046*, AF469770;

Appendix (continued)

Canella winterana (L.) Gaertn., AF197676, AF197757, DQ406920*, AJ235424, AF543731, AJ131928, DQ407108*, AF206879; *Carludovica palmata* Ruiz & Pavon, AF197707, AF197734, DQ406948*, AF293861, AB088793, AF197596, DQ407079*, AF293756; *Calycanthus floridus* L., AF197678, AF197777, DQ406922*, AJ235422, NC_004993, AJ235422, DQ407074*, U38318; *Ceratophyllum demersum* L., AF197627, AF197730, DQ406988*, AJ235430, AF543732, M77030, DQ407133*, U42517; *Ceratophyllum submersum* L., AF197628, AF197731, DQ406989*, AF293860, DQ401361*, AF197599, DQ407134*, AF197582; *Cercis canadensis* L., DQ401293*, DQ401409*, DQ406868*, DQ401328*, AY386948 (*C. gigantea* F.C. Cheng & Keng f.), U74188, DQ407054*, — ; *Chamaedorea tenella* H. Wendl., DQ401295*, DQ401392*, DQ407003*, AF233083 (*C. seifrizii* Burret), DQ401368*, AJ404787 (*C. microspadex* Burret), DQ407083*, AF069210 (*C. seifrizii* Burret); *Chimonanthus praecox* (L.) Link, AF197679, AF197778, DQ406928*, AF197605, AY525340, L12639, DQ407075*, AF503352; *Chloranthus multistachys* Pei, AF197665, AF197753, DQ406864*, AJ235431 (*C. japonicus* Siebold), AF543733 (*C. brachystachys*), L12640 (*C. japonicus* Siebold), DQ407151*, AF206885; *Cinnamodendron ekmanii* Sleum., AF197677, AF197758, DQ406921*, AJ235435, DQ401342*, AJ235776, DQ407156*, AF206887; *Cinnamomum camphora* (L.) T. Nees & Eberm., AF197681, AF197797, DQ406933*, AJ235436, AJ247154, L12641, DQ407116*, AF206888; *Cissampelos pareira* L., AF197645, AF197775, DQ406892*, AF197613, DQ401349*, AF197590, DQ407153*, AF293758; *Cocculus trilobus* (Thunb.) DC, AF197646, AF197776, DQ406893*, AF197614, AB069852 (*C. orbiculatus* DC.), L12642, DQ407029*, AF197581; *Coffea arabica* L., DQ401313*, DQ401399*, DQ406878*, AJ235441, DQ401346*, X83631, DQ407115*, — ; *Cornus suecica* L., AF420915, AF420990, DQ407012* (C. sp), AJ235444 (*C. mas* L.), AJ429275 (*C. mas* L.), L11215 (*C. florida*), DQ407055* (C. sp), U52033 (*C. officinalis* Siebold & Zucc.); *Croomia pauciflora* Miq., AF197708, AF197735, DQ406939*, AF308039 (*C. japonica* Miq.), AY437815, D28154 (*C. heterosepala* Torr ex Torr. & Gray), DQ407078*, AF168835; *Cryptocarya meissneriana* Frodin, AF197702, AF197804, DQ406932*, AF197602, AJ627923 (*C. subtriplinervia*), U06841 (*C. obovata* R. Br.), DQ407113*, AF293757; *Cycas revoluta* Thunb., AF197623, AF197720, AJ130743, AF469657, AF410157 (*C. fairylakea* D. Yue. Wang), AF531199, DQ407139*, D85297 (*C. taitungensis* C.F. Shen, K.D. Hill, C.H. Tsou & C.J. Chen); *Cyperus alternifolius* L., DQ401297*, DQ401381*, DQ406944*, AF168906 (*C. albostriatus* Schrad.), DQ401373*, AF449514 (*C. tenellus*), — , AF168838 (*C. albostriatus* Schrad.); *Daphnandra micrantha* (Tul.) Benth., AF197684, AF197800, DQ406977*, AF293857, DQ401344*, AF052195 (*D. repandula* F. Muell.), DQ407119*, AF197579; *Degeneria vitiensis* I.W. Bailey & A.C. Sm., AF293752, AF197771, DQ406991*, AJ235451, AY220440 (*D. roseiflora* John M. Mill.), L12643, DQ407154*, AF206898; *Dicentra* sp., AF197649, AF197796, DQ406890*, AJ235454 (*D. chrysantha* Walp.), AF543734 (*D. eximia* Torrey), L08761, DQ407066*, L37908 (*D. eximia* Torrey); *Didymeles perrieri* Olivier, AF197637, AF197811, DQ406993*, AF092119, DQ401354*, AF061994, DQ407149*, AF094541; *Dillenia indica* L., DQ401306*, AY163747, DQ406882*, AF095732 (*D. retusa* Thunb.), DQ401359*, L01903, DQ407070*, U38314 (*D. alata* (DC.) Martelli); *Dioscorea* sp., AF197709, AF197737, DQ406959*, AF308014 (*D. pentaphylla* L.), AB040208 (*D. alata* L.), AJ235803 (*D. polygonoides* Humb & Bonpl), DQ407077*, AF206903 (*D. polygonoides* Humb & Bonpl); *Doryphora sassafras* Endl., AF197688, AF197801, DQ406978*, AF293858, AF542568, L77211 (*D. aromatica* (F.M. Bailey) L.S. Smith), DQ407112*, AF293754; *Drimys winteri* J.R. & G. Forster, AF197673, AF197781, DQ406919*, AF093425, AY437816, L01905, DQ407045*, U42823; *Eucommia ulmoides* Oliver, DQ401311*, DQ401387*, DQ406872*, AJ235469, AF345323, L01917, DQ407058*, L54066; *Euphorbia milii* var. *splendens* Desmoul., DQ401317*, AY674512 (*E. polychroma* Kern.), DQ406908*, AJ235472 (*E. polychroma* Kern.), AY491656 (*E. polyacantha* Boiss.), AF530850 (*E. humifusa*), DQ407118*, AF530762 (*E. humifusa*); *Eupomatia bennettii* F. Muell., AF197692, AF197772, DQ406927*, AJ235473, DQ401341*, L12644, DQ407048*, AF469771; *Euptelea polyandra* Sieb. & Zucc., AF197650, AF197787, DQ406873*, U86384, DQ401348*, L12645, DQ407031*, L75831; *Galbulimima belgraveana* (F. Muell.) Sprague., AF197693, AF197773, DQ406992*, AJ235478, AY220441, L12646, DQ407125*, AF206916; *Ginkgo biloba* L., AF197625, AF197722, AJ409109, AJ235481, AF279806, AJ235804, DQ407120*, D16448; *Gnetum gnemon* L., AF197617, AF197718, AJ409110, AF187060, AF280994 (*G. montanum* Markgr.), L12680, DQ407159*, U42416; *Grevillea robusta* Cunn. & R. Br., AF197712, AF197808, DQ406886*, AF060434 (*G. baileyana* R. Br.), AF542583 (*G. banksii* R. Br.), AF197589, DQ407023*, AF197577; *Guaiacum officinale* L., DQ401291*, AY674517, DQ406954*, AJ235486 (*G. sanctum* L.), DQ401366*, AJ131770 (*G. sanctum* L.), DQ407117*, AY674599; *Gunnera monoica* Raoul, DQ401302*, DQ401383*, DQ406897*, AF093374 (*G. hamiltonii* Kirk ex W. Ham), AY042596 (*G. perpensa* L.), AF307918, DQ407016*, U43787 (*G. manicata* Linden); *Gyrocarpus americanus* Jacq., AF197701, AF197805 (G. sp.), DQ406931*, AJ235487, DQ401370*, L12647 (G. sp), DQ407164*, AF206923; *Hamamelis mollis* Oliver ex Forb. & Hemsl., DQ401289*, AY453082, DQ407011*, AF093380 (*H. virginiana* L.), AF248618 (*H. mexicana* Standl.), L01922, DQ407106*, AF094551 (*H. virginiana* L.); Hedera helix L., DQ401310*, DQ401390*, DQ406955*, AJ235489, AJ319067, L01924, DQ407062*, U42500; *Hedycarya arborea* J.R. & G. Forst., AF197689, AF197806, DQ406909*, AJ235490, AJ627927, L12648, DQ407161*, AF206924; *Hedyosmum arborescens* Sw., AF197668, AF197756, DQ406863*, AJ235491, DQ401339*, L12649, DQ407042*, AF206925; *Helianthus annuus* L., X55963, AY453114, AF258785, AJ236205, AJ429380, L13929, DQ407063*, AF107577; *Hernandia ovigera* L., DQ007413, DQ007424, DQ406930*, DQ401325*, AJ247165 (*H. nymphaeifolia* (Presl) Kub.), L12650, DQ407109*, DQ007407; *Heuchera* sp., DQ401290*, DQ401398*, DQ406953*, AF093399 (*H. sanguinea* Engelm.), L34127 (*H. rubescens* Torr.), L01925 (*H. micrantha* Dougl.), DQ407053*, L28139 (*H. micrantha* Dougl.); *Hortonia floribunda* Wight ex Arn., DQ007414, DQ401404*, DQ406975*, DQ401323*, AY437811, AF040663, DQ407102*, DQ007408; *Houttuynia cordata* Thunb., AF197632, AF197749, DQ406980*, AF093397, AF543737, AF332100, DQ407093*, AF206929; *Hydrocharis morsus-ranae* L., — , — , DQ406963*, DQ401334*, AB002572 (H. dubia), U80701, DQ407067*, AY952398 (*H. dubia*); *Hydrocleys nymphoides*, — , — , DQ406999*, DQ401332*, AB002580, AB004900, DQ407143*, AY952397; *Ilex* sp., AY741812 (*I. verticillata* A. Gray), AY453090 (*I. aquifolium* L.), DQ406884*, AJ235502 (*I. crenata* Thunb.), AJ429376, X98735 (*I. brasiliensis*), DQ407061*, AF206938 (*I. opaca*); *Idiospermum australiense* (Diels) S. T. Blake, AF197680, AF197779, DQ406974*, AJ235500, AY525342, L12651, DQ407076*, AF206937; *Illicium floridanum* Ellis, AF197663, AF197740, DQ406985*, U86385 (*I. parviflorum* Michx. ex Vent), AF543738, AF543713, DQ407132*, L75832 (*I. parviflorum* Michx. ex Vent); *Impatiens pallida* Nutt., AF420933 (*I. parviflora* DC.), AF421011 (*I. parviflora* DC.), DQ406952*, AJ235503, AJ429280 (*I. capensis* Meerb.), AB043533 (*I. discolor*), DQ407057*, L24148; *Iris* sp., DQ401300*, DQ401386*, DQ407006*, AY147620 (*I. missouriensis*), AB088786 (*I. japonica* Thunb.), D28332 (*I. ensata* Thunb.), DQ407091*, AY952396 (*I. japonica* Thunb.); *Kadsura japonica* (L.) Dunal, AF197661, AF197738, DQ406971*, AF197607, AF542565, AF197595, DQ407128*, AF293763; *Lactoris fernandeziana* Phil., AF197710, AF197812, DQ406910*, AJ235515, AF543739, L08763, DQ407126*, U42783; *Lamium* sp., DQ401312*, DQ401385*, DQ406871*, DQ401329*, AJ429332 (*L. album* L.), Z37403 (*L. purpureum* L.), DQ407060*, L49287 (*L. amplexicaule* L.); *Lardizabala biternata* Ruiz & Pavon, AF197643, AF197789, DQ406867*, DQ401326*, AY437809, L37919, DQ407073*, L37910; *Laurus nobilis* L., AF197682, AF197798, DQ406923*, AJ235518, AJ247167 (*L. azorica* (Seub.) Franco), AF197593, DQ407165*, AF197580; *Leea guineensis* G. Don., DQ401304*, AY674530, DQ406899*, AJ235520, AF274621, AJ235783, DQ407015*, AF206951; *Lilium* sp., AY394729 (*L. tigrinum* Ker Gawl.), DQ401403*, DQ407002*, AF209618 (*L. superbum* L.), AB030847 (*L. columbianii* Hanson), L12682 (*L. superbum* L.), DQ407081*, D29775 (*L. formosanum* Wallace); *Liriodendron chinense* (Hemsl.) Sarg, AF197690, AF197774, DQ406926*, AJ235522 (*L. tulipifera* L.), AF123481, L12654, DQ407049, AJ23F5981; *Lomandra obliqua*, DQ401296, DQ401380, DQ406942, AJ417591 (*L. ordii*), DQ401356, L05039 (*L. longifolia* Labill.), DQ407099, — ; *Magnolia tripetala* L., AF197691, AF197770, DQ406916, AJ235526, AF548649, AJ131927, DQ407050*, AF206956; *Mahonia bealei* (Fortune) Carr., AF197659, AF197761, DQ406888*, AF197611, AB038184 (*M. japonica* Thunb. or DC.), L12657, DQ407069*, AF293755; *Maranta leuconeura* E. Morr., AY299801, DQ401410*, DQ406943*, AF168927 (*M. bicolor* Vell.), AY140303, L05040, DQ407085*, AF069225 (*M. bicolor* Vell.); *Mauloutchia chapelieri* Warb., AF197699, AF197769, DQ406960*, AF197606, AY220451, AF197594, DQ407160*, DQ007409; *Meliosma squamulata* Hance., AF197656, DQ007426, DQ406896*, AF209626 (*M. veitchiorum* Hemsl.), DQ401353*, AF197587 (*M. simplicifolia* (Roxb.) Walp.), DQ407021*, AF206961 (*M. veitchiorum* Hemsl.); *Metasequoia glyptostroboides* Hu & Cheng, AF197619, — , DQ406973*, AF469660, AB030122, AJ235805, — , L00970; *Myristica fragrans* Houtt., AF197698, AF197768, DQ406967* (*M. maingayi* Hook. f.), AJ235539, DQ401374* (*M. maingayi* Hook. f.), AF206798, DQ407142*, AF206968; *Najas*

Appendix (continued)

gracillima, — , — , DQ406969*, DQ401333*, AB002579 (*N. minor* All.), AB004899, DQ407127*, — ; *Nelumbo nucifera* Gaertner, AF197654, AF197795, DQ406894*, AF093387 (*N. lutea* (Willd.) Pers.), AF543740, M77033, DQ407025*, L75835 (*N. lutea* (Willd.) Pers.); *Nepenthes* × *kosobe*, DQ401307*, DQ401379*, DQ406900*, AJ235542 (*N. alata* Blanco), AF315864 (*N. gymnamphora*), L01936 (*N. alata* Blanco), DQ407068*, U42787 (*N.* sp.); *Nicotiana tabacum* L., NC_006581, AY453113 (*N. sylvestris* Speg. & Comes), NC_006581, AF035909, NC_001879, NC_001879, DQ407059*, AJ236016; *Nothofagus moorei* (F. Muell.) Maiden, DQ401292*, DQ401401*, DQ406905*, AY605515, AB015464 (*N. solandri* (Hook. f.) Oerst.), L13356, DQ407147*, AY147111; *Nuphar* sp., AF197638, AF197726, DQ406982*, AF209640 (*N. variegata* Durand), AF543741 (*N. lutea* L.), M77029 (*N. variegata* Durand), DQ407040*, AF206972 (*N. variegata* Durand); *Nymphaea* sp., AF197639, AF197727, DQ406981*, AJ235544 (*N. odorata* Aiton), NC_006050 (*N. alba*), M77034 (*N. odorata* Aiton), DQ407123*, AF206973 (*N. odorata* Aiton); *Oenothera biennis*, X04023, AY453083, X07566 (*O.* sp.), NC_002693 (*O. elata* ssp. *hookeri*), NC_002693 (*O. elata* ssp. *hookeri*), AF495770 (*O. brachycarpa* A. Gray), DQ407114* (*O. organensis* S. Emerson), — ; *Oncidium sphacelatum* Lindl., DQ401299*, DQ401393*, DQ407005*, AJ235548 (*O. excavatum* Lindl.), AF350630 (*O. dasytyle* Rchb. f.), AF074201 (*O. excavatum* Lindl.), DQ407103*, AF168864 (*O. ornithoglossum*); *Orontium aquaticum* L., AF197705, AF197745, DQ406996*, AF197610, AF543744, AJ005632, DQ407064*, AF293753; *Oryza sativa* L., BA000029, DQ401382*, BA000029, NM_195988, AF148650, NM_195989, X15901, AF069218; *Oxalis corniculata* L., DQ401314* (*O.* sp.), AY453111, DQ406907* (*O.* sp.), AF530730, AF542605 (*O. stricta*), L01938 (*O. dillenii* Jacq.), DQ407138* (*O.* sp.), AF206978 (*O. dillenii* Jacq.); *Pachysandra procumbens* Michx., AF197634, AF197784, DQ406887*, AF092111, AF542581 (*P. terminalis* Sieb. & Zucc.), AF093718, DQ407017*, AF094533; *Passiflora suberosa* L., DQ401315*, AY453071 (*P. edulis* Sims), DQ406902*, AJ235553 (*P. coccinea* Aubl.), DQ401363*, L01940 (*P. quadrangularis* L.), DQ407157*, AF206981 (*P. standleyi* Killip.); *Peperomia obtusifolia* A. Dietr., AF197629, AF197814, DQ406924*, AJ235556, AF542574 (*P. graveolens* Rauh & Barthlott), L12661 (*P.* sp.), DQ407166* (*P. argyreia* E. Morr.), AF206985; *Persoonia katerae* P. Weston & L. Johnson, AF197652, AF197794, DQ406984*, AF197615, AY437813, U79178 (*P. lanceolata*), DQ407145*, AF293762; *Petrophile canescens* Cunn. Ex R. Br., AF197653, AF197807, DQ406983*, AF060401 (*P. circinata* Kippist ex Meisn.), DQ401365*, U79181 (*P. biloba* R. Br.), DQ407148*, AF293761; *Peumus boldus* Molina, AF197686, AF197803, DQ406990*, AF209650, AJ247183, AF206807, DQ407162*, AF206988; *Philodendron oxycardium* Schott, DQ401319*, DQ401395*, DQ406937*, DQ401331*, DQ401355*, AJ005623, DQ407121*, — ; *Philydrum lanuginosum* Banks ex Gaertn., AY299824, DQ401406*, — , AY147607, DQ401369*, U41596, DQ407124*, AY952390; *Phytolacca americana* L., DQ401288*, DQ401389*, DQ406904*, AF093391, DQ401362*, M62567, DQ407052*, U42793; *Pinus thunbergii* Parl., AF197626 (*P.* sp.), AF197723 (*P.* sp.), AY832181, NC_001631, AB161019 (*P. hartwegii rudis* (Endl.) Silba), AY115755 (*P. maximartinezii* Rzed.), AY228468 (*P. koraiensis* Siebold & Zucc.), D16446 (*P. luchuensis* Mayr); *Piper betle* L., AF197630, AF197750, DQ406925*, AJ235560, AB040153 (*P. nigrum*), L12660, DQ407167*, AF206992; *Platanus occidentalis* L., AF197655, AF197793, AY832177, U86386, AF543747, AF081073, DQ407024*, U42794; *Pleea tenuifolia* Michaux, AF197703, AF197743, DQ406995*, AJ235564, AB183407, AJ131774, DQ407131*, AF206995; *Podocarpus macrophyllus* (Thunb.) Sweet, AF197620, — , DQ406962*, AF469661 (*P. chinensis* Sweet), AF228111, AF249616, DQ407150*, U87301 (*P. totara* D. Don); *Podophyllum peltatum* L., AF197660, AF197762, DQ406889*, AF197612, AB069843, AF197591, — , L24413; *Polygonum* sp., DQ401308*, DQ401388*, DQ406901*, AJ235569 (*P. sachalinense* F. Schdt ex Maxim), AF204859 (*P. bistorta*), AF297128 (*P. erectum* L.), DQ407051*, AF206996; *Potamogeton berchtoldii* Fieber, AF197715, AF197724, DQ406938*, AF197600, AB002581 (*P. distinctus* Arth. Benn), AB088809 (*P. distinctus* Arth. Benn), DQ407034*, DQ007410; *Ranunculus* sp., AF197714, AF197759, DQ406876*, DQ401327*, AB069847 (*R. silerifolius* H. Lev.), L08766, DQ407028*, D29780 (*R. taisanensis* Hayata); *Sabia* sp., AF197657, AF197780, DQ406895*, AF093395 (*S. swinhoei* Hemsl.), DQ401352*, L12662, DQ407022*, L75840 (*S. swinhoei* Hemsl.); *Sanguinaria canadensis* L., AF197651, AF197788, DQ406877*, U86387, DQ401350*, L01951, DQ407152*, AF293760; *Sarcandra chloranthoides* Gardner, AF197666, AF197754, DQ406866*, AJ235593 (*S. grandiflora* Subr. & Henry), DQ401340*, AY236833, DQ407041*, AF207012; *Sargentodoxa cuneata* (Oliv.) Rehder & Wilson, AF197644, AF197790, DQ406875*, AF093396, DQ401351*, AF093731, DQ407026*, L75841; *Saruma henryi* Oliv., AF197672, AF197752, DQ406912*, AJ235595, AF543748, L12664, DQ407097*, AF207013; *Saururus cernuus* L., AF197633, AF197748, DQ406934*, AF093398, AF543749, L14294, DQ407094*, U42805; *Schisandra sphenanthera* Rehd. & Wils., AF197662, AF197739, DQ406972*, AJ235599, AF543750 (*S. rubriflora*), L12665, DQ407129*, L75842 (*S. chinensis* (Turcz.) Baill.); *Simmondsia californica* Nutt., DQ401309*, DQ401397*, DQ406903*, AF093401 (*S. chinensis* C.K. Schneid.), AY514854 (*S. chinensis* C.K. Schneid.), AF093732 (*S. chinensis* C.K. Schneid.), DQ407105*, AF094562 (*S. chinensis* C.K. Schneid.); *Siparuna brasiliensis* A. DC., AF197687 (*S. decipiens* A. DC), AF197809 (*S. decipiens* A. DC), DQ406976*, DQ401324*, DQ401375*, AF129016 (*S. glycycarpa*), DQ407163*, DQ007411 (*S. decipiens* A. DC); *Smilax glauca* Walter, AF039251 (*S. rotundifolia*), DQ401391* (*Smilax* sp.), DQ406940* (*Smilax* sp.), AF209677, AB040204 (*S. china*), AF206822, DQ407082* (*Smilax* sp.), AF207022; *Sparganium americanum* Nutt., AY124509 (*S. eurycarpum* Engelm. in A. Gray), DQ401396*, DQ407010*, AF209678, AB088802 (*S. stoloniferum* Buch.-Ham.), M91633, DQ407086*, AF069220 (*S. eurycarpum* Engelm. in A. Gray); *Spathiphyllum clevelandii*, AF197706, AF197746, DQ406997*, AJ235606 (*S. wallisii* Hort), AF542575 (*S. floribundum* (Lind & Andre) N.E. Br.), AJ005626, DQ407072*, AF207023 (*S. wallisii* Hort); *Spinacia oleracea* L., DQ401287*, AY453110, DQ406883*, AF528861, NC_002202, NC_002202, DQ407013*, L24420; *Staphylea trifolia* L., DQ401294*, AY453105, DQ406906*, AJ235611, AF542582 (*S. pinnata* L.), AJ238406, DQ407101*, AJ235978; *Stegolepis* sp., AY124535 (*S. parvipetala* Steyerm.), DQ401411, DQ407004*, DQ401336*, AY614013 (*S. ligulata* Maguire), AY123242 (*S. parvipetala* Steyerm.), DQ407089*, — ; *Sterculia balanghas* L., DQ401316*, DQ401402*, DQ406869*, AJ233089 (*S. apetala* Karst.), AY321178 (*S. tragacantha* Lindl.), AF022126 (*S. tragacantha* Lindl.), DQ407107*, AF207029 (*S. recordiana* Standl.); *Strelitzia nicolai* Regel & K. Koch, AY299843, AY453112 (*S. reginae*), DQ406965* (*S. reginae*), AF168948, AF434874 (*S. alba* Skeels), AF243846, DQ407084* (*S. reginae*), AF069229; *Tacca chantrieri* Andre, AF039252 (*T. pinnatifida* J.R. Forst. & G. Forst.), DQ401377*, DQ406941*, AJ235618, AB088792 (*T.* sp.), AJ235810, DQ407098*, U42063 (*T. plantaginea* (Hance) Drenth); *Takhtajania perrieri* M. Baranova & J. Leroy, DQ007416, DQ007427, DQ406913*, AF209683, DQ401371*, AF206824, DQ407137*, AF207032; *Tasmannia insipida* DC, AF197674, AF197782, DQ406970*, AF093424, AF543735 (*T. lanceolata* Smith), L01957, DQ407044*, AF207035; *Tetracentron sinense* Oliv., AF197647, AF197791, DQ406874*, AF093422, AF274633, L12668, DQ407020*, U42814; *Thottea tomentosa* Ding Hou, AF197670, AF197733, DQ406914*, AF197609, AB060738, AF197598, DQ407155* (*T. borneensis* Valeton), DQ007406; *Tofieldia calyculata* (L.) Wahlenb., AF197704, AF197744, DQ406935*, AJ235627, AB040160 (*T. racemosa*), AJ235798, DQ407036*, AF207043; *Tradescantia* sp., DQ401320*, AY453108, DQ406950*, AF168950 (*T. ohiensis* Raf.), DQ401372*, L05041, DQ407088*, AF069213 (*T. ohiensis* Raf.); *Triglochin maritimum* L., AF197716, AF197725, DQ406998*, AF197601, AB088782, AB088811, DQ407071*, AF197586; *Trillium* sp., AF039253 (*T. grandiflorum* (Michx) Salisb.), DQ401407*, DQ406949*, AF209692 (*T. erectum* var. *album* (Michx) Pursh), AB017416 (*T. viridescens* Nutt.), AB018848 (*T. undulatum* Willd.), DQ407065, AF207048 (*T. erectum* var. *album* (Michx) Pursh); *Trimenia moorei* W.R. Philipson, DQ007415, AF197741, DQ406987*, AY116653, DQ401360*, AY116658, DQ407141*, see Zanis et al., 2002 for 18S; *Trochodendron aralioides* Sieb. & Zucc., AF197648, AF197792, DQ406880*, AF093423, AF543751, L01958, DQ407019*, U42816; *Vitis* sp., DQ401305*, AY453123 (*V. riparia*), DQ406881*, AJ235643 (*V. aestivalis* Michx.), AJ429274 (*V. vinifera* L.), AJ635355 (*V. vinifera* L.), DQ407014*, AF207053; *Vriesea splendens* Lem., DQ401298*, DQ401378*, DQ406945*, DQ401337*, AY614045, AY614411, DQ407087*, — ; *Welwitschia mirabilis* Hook. f., AF197618, AF197719, DQ406958*, AF239795, AF280996, AJ235814, DQ407140*, AF207059; *Xanthorhiza simplicissima* Marshall, AF197658, AF197760, DQ406885*, AF093394, AB069848, L12669, DQ407027*, L75839; *Xanthorrhoea quadrangulata* F. Muell., AF039250 (*X. australis* R. Br.), DQ401384*, DQ406946*, AF168952, DQ401345*, Y17339 (*X. minor* R. Br.), DQ407135*, U42064; *Xanthosoma mafaffa* Schott, DQ401318*, DQ401376*, DQ406936*, DQ401330*, DQ401357*, AJ007543, DQ407033*, — ; *Zamia floridana* A. DC, AF197624, AF197721, DQ406961*, AF188845 (*Z. furfuracea* Aiton), AF279804, AF531218, — , M20017 (*Z. inermis* A.P. Vovides, J.D. Rees, & M. Vazquez-Torres); *Zostera capricorni* Aschers., — , — , DQ406964*, AF209700 (*Z. noltii* Hornem.), AB096167, AY077963, DQ407144*, AF207058 (*Z. noltii* Hornem.).