Int. J. Plant Sci. 172(4):541–558. 2011.

1058-5893/2011/17204-0007$15.00 DOI: 10.1086/658923

PHYLOGENETIC ANALYSIS OF THE PLASTID INVERTED REPEAT FOR 244 SPECIES: INSIGHTS INTO DEEPER-LEVEL ANGIOSPERM RELATIONSHIPS FROM A LONG, SLOWLY EVOLVING SEQUENCE REGION

Michael J. Moore,[1,2,]* Nasr Hassan,[1,]† Matthew A. Gitzendanner,‡ Riva A. Bruenn,* Matthew Croley,* Alexia Vandeventer,‡ James W. Horn,§ Amit Dhingra,‖ Samuel F. Brockington,# Maribeth Latvis,‡ Jeremy Ramdial,‡ Roolse Alexandre,‡ Ana Piedrahita,‡ Zhenxiang Xi,** Charles C. Davis,** Pamela S. Soltis,†† and Douglas E. Soltis‡

*Biology Department, Oberlin College, Oberlin, Ohio 44074, U.S.A.; †Botany Department, Faculty of Science, Assiut University, Assiut 71516, Egypt; ‡Department of Biology, University of Florida, Gainesville, Florida 32611, U.S.A.; §Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, U.S.A.; ‖Department of Horticulture and Landscape Architecture, Washington State University, Pullman, Washington 99164, U.S.A.; #Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom; **Harvard University Herbaria, Cambridge, Massachusetts 02138, U.S.A.; and ††Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, U.S.A.

Recent plastid phylogenomic studies have helped clarify the backbone phylogeny of angiosperms. However, the relatively limited taxon sampling in these studies has precluded strongly supported resolution of some regions of angiosperm phylogeny. Other recent work has suggested that the 25,000-bp plastid inverted repeat (IR) region may be a valuable source of characters for resolving these remaining problematic nodes. Consequently, we aligned all available angiosperm IR sequences to produce a matrix of 24,702 aligned bases for 246 accessions, including 36 new accessions. Maximum likelihood analyses of the complete data set yielded a generally well-supported topology that is highly congruent with those of recent plastid phylogenomic analyses. However, reducing taxon sampling to match a recent 83-gene plastid analysis resulted in significant changes in bootstrap support at some nodes. Notably, IR analyses resolved *Pentapetalae* into three well-supported clades: (1) superasterids (comprising Santalales, *Caryophyllales*, Berberidopsidales, and *Asteridae*), (2) superrosids (comprising Vitaceae, Saxifragales, and *Rosidae*), and (3) Dilleniaceae. These results provide important new evidence for a stable, well-supported phylogenetic framework for angiosperms and demonstrate the utility of IR data for resolving the deeper levels of angiosperm phylogeny. They also reiterate the importance of carefully considering taxon sampling in phylogenomic studies.

Keywords: Angiosperm Tree of Life, plastid inverted repeat, phylogenetics, large data sets.

Online enhancement: appendix.

Introduction

A well-supported phylogeny of *Angiospermae* (extant angiosperms) is crucial for understanding the developmental, morphological, and ecological history of flowering plants. Over the past 20 years, studies employing targeted sequencing of one to several genes have brought much of the backbone phylogeny of *Angiospermae* into focus (reviewed in Judd and Olmstead 2004 and Soltis et al. 2005). For example, such studies have revealed that the species-poor clades *Amborella*, Nymphaeales, and Austrobaileyales form a grade sister to all other *Angiospermae* and that the bulk of angiosperm diversity is represented by the large clades *Eudicotyledoneae* and *Gunneridae* (=core eudicots; order and family names throughout follow APG III 2009, whereas the names for more inclusive clades follow Cantino et al. 2007). Nevertheless, many deep-level *Angiospermae* relationships, such as those among the major clades of *Mesangiospermae* (Chloranthaceae, *Magnoliidae*, *Ceratophyllum*, *Monocotyledoneae*, and *Eudicotyledoneae*) and *Pentapetalae* (=core eudicots excluding Gunnerales), have eluded resolution when only a few genes were used (Soltis et al. 2000, 2005, 2007; Hilu et al. 2003; Worberg et al. 2007; Burleigh et al. 2009). Recently, several researchers have turned to analyses of 61–83 plastid genes to improve resolution among major clades of *Angiospermae* (Leebens-Mack et al. 2005; Cai et al. 2006; Jansen et al. 2007; Moore et al. 2007, 2010). For example, analyses of large numbers of plastid genes have provided the strongest support yet obtained for relationships among the five major clades of *Mesangiospermae*, including moderate to strong support for a clade consisting of *Monocotyledoneae*, *Ceratophyllum*, and *Eudicotyledoneae* (Jansen et al. 2007; Moore et al. 2007). Likewise, plastid phylogenomics has further refined *Eudicotyledoneae* relationships, revealing that *Pentapetalae* comprise two major clades: a superrosid clade consisting of *Rosidae*, Saxifragales, and Vitaceae and a superasterid clade consisting of Berberidopsidales, Santalales, *Caryophyllales*, and *Asteridae* (Moore et al. 2010).

[1] These two authors contributed equally to this research.

[2] Author for correspondence; e-mail: michael.moore@oberlin.edu.

Manuscript received May 2010; revised manuscript received December 2010.

However, even with complete plastid genome sequence data, several deep-level *Angiospermae* relationships have yet to be resolved. For example, the placement of Dilleniaceae within *Pentapetalae* remains problematic. In studies involving one to several genes, the family has frequently appeared as sister to *Caryophyllales* (in the superasterids), albeit with weak support (Soltis et al. 2000, 2007; Burleigh et al. 2009). Maximum likelihood (ML) analyses of 83 plastid genes reject this placement of Dilleniaceae, instead resolving Dilleniaceae as sister to the superrosid clade, although again with weak support (Moore et al. 2010). Likewise, relationships among some deep-level *Eudicotyledoneae* clades (e.g., Buxales, Trochodendrales, and *Gunneridae*), as well as those among the three major superrosid clades (Vitaceae, *Rosidae*, and Saxifragales), remain unsettled. Support for relationships among these clades has been low or moderate even with gene-rich plastid genome data sets (Soltis et al. 2005, 2007; Jansen et al. 2007; Moore et al. 2007, 2010; Worberg et al. 2007; Burleigh et al. 2009; Wang et al. 2009).

A serious current limitation of plastid phylogenomic studies is relatively sparse taxon sampling. For example, complete plastid genome sequences are available for just four species of *Caryophyllales*, two species each of Saxifragales and Santalales, and only one species each of Vitaceae, Dilleniaceae, and Berberidopsidales. The potential problems and pitfalls of low taxon density for topology estimation and support values are well known (Soltis and Soltis 2004; Leebens-Mack et al. 2005; Heath et al. 2008). However, improving taxon sampling through sequencing additional complete plastid genomes can be limited by expense and requirements for fresh tissue (although cost is declining rapidly and alternative methods of isolating plastid DNA are being explored; Cronn et al. 2008). To increase taxon density efficiently and economically in crucial portions of angiosperm phylogeny, we have sequenced the complete or nearly complete inverted repeat (IR) region of the plastid genome for more than 134 phylogenetically pivotal angiosperm species (Jian et al. 2008; Brockington et al. 2009; Soltis et al. 2009; Wang et al. 2009), including 36 accessions reported here for the first time. Combining these sequences with existing IR data has enabled us to assemble a large, character-rich data set of nearly 250 species of *Angiospermae* that more fully represents flowering plant diversity, compared to recent studies employing all angiosperm plastid protein-coding and rRNA genes.

A feature of nearly all land plant plastid genomes, the IR divides the plastid genome into large and small single-copy regions, which in *Angiospermae* are ~85 kilobase pairs (kb) and 20 kb in length, respectively (Downie and Palmer 1992; Raubeson and Jansen 2005). In most *Angiospermae*, the IR is between 22 and 26 kb in length, with much of this size variation due to expansion and contraction of the ends of the IR as well as pseudogenization of some genes (e.g., *rpl23*, *ndhB*, and *ycf2*) and occasional loss of introns (e.g., *ndhB* and *rpl2* introns) in various taxa (Palmer 1991; Downie and Palmer 1992; Plunkett and Downie 2000; Raubeson and Jansen 2005; Wang et al. 2008).

The IR is also characterized by a highly conserved structure and an evolutionary rate that is low compared to that of the single-copy regions (Wolfe et al. 1987; Downie and Palmer 1992; Graham et al. 2000; Raubeson and Jansen 2005; Moore et al. 2007; Jian et al. 2008), facilitating its amplification (Dhingra and Folta 2005) and potentially reducing the effects of mutational saturation on phylogenetic estimation (Felsenstein 1978; Bull et al. 1993). The length of the IR may also enhance its phylogenetic utility, as simulation studies have suggested that increasing sequence length by using slowly or moderately evolving regions may improve phylogenetic accuracy more than utilizing sequences with a relatively fast rate (Wortley et al. 2005). Recent phylogenetic studies in Saxifragales, *Rosidae*, and *Caryophyllales* have indeed demonstrated that the IR is an appropriate marker for obtaining high resolution of deep-level angiosperm divergences (Jian et al. 2008; Brockington et al. 2009; Wang et al. 2009). Consequently, the IR region may hold great promise in helping to resolve remaining problem areas within *Angiospermae* phylogeny. Combining all existing flowering plant IR sequences with a number of new IR sequences may mitigate the problems with low taxon density seen in complete plastid genome data sets and may therefore provide increased resolution and support for deeper-level *Angiospermae* relationships. In this study, we report analyses that test this hypothesis, with a particular focus on relationships within *Eudicotyledoneae*. Despite the low rate of evolution within the IR, we find that ML and maximum parsimony (MP) analyses of the IR region are highly congruent with recent plastid phylogenomic analyses, although at several deeper nodes MP trees display evidence of systematic error associated with long branches.

Material and Methods

Taxon Sampling

The complete data matrix contained 246 accessions (including 244 species), including 235 accessions of *Angiospermae* and 11 *Acrogymnospermae* (=extant gymnosperms) as outgroups. All major deeper-level clades and all but five orders of *Angiospermae* (sensu APG III 2009) were represented, as were the major *Acrogymnospermae* lineages: *Cycadophyta*, *Ginkgo*, *Pinaceae*, *Cupressophyta*, and all three genera of *Gnetophyta* (app. A). All available complete *Angiospermae* IR sequences were downloaded from GenBank (app. A). To help resolve the phylogenetic positions of several angiosperm groups, especially among the basalmost lineages of *Eudicotyledoneae* and *Pentapetalae*, we increased sampling by obtaining complete or partial IR sequence data from several additional taxa, including one species of Nymphaeales (*Trithuria*), one species of Chloranthaceae (*Hedyosmum*), one species of *Magnoliidae* (*Aristolochia*), four species of *Monocotyledoneae* (*Lilium*, *Spathiphyllum*, *Tradescantia*, and *Vallisneria*), and 29 species of *Eudicotyledoneae*, distributed among numerous orders: Berberidopsidales (*Aextoxicon*), Boraginales (*Ehretia*), Dilleniaceae (*Acrotrema*, *Curatella*, *Davilla*, *Dillenia*, *Hibbertia*, *Pinzona*, and *Tetracera*), Dipsacales (*Lonicera*), Ericales (*Rhododendron*), Garryales (*Aucuba*), Gentianales (*Nerium*), Gunnerales (*Myrothamnus*), Lamiales (*Antirrhinum*), Malpighiales (*Caryocar*, *Erythrospermum*, and *Malesherbia*), Proteales (*Nelumbo*), Sabiaceae (*Meliosma*), Santalales (*Heisteria*, *Phoradendron*, and *Ximenia*), Saxifragales (*Heuchera* and *Liquidambar*), Trochodendrales (*Trochodendron*), unplaced

(*Cynomorium*), and Vitaceae (*Cissus* and *Tetrastigma*); see appendix A.

DNA Amplification and Sequencing

DNA was extracted from fresh leaf material and from herbarium specimens with the method of Doyle and Doyle (1987), but including the modifications of Soltis et al. (1991). New IR sequences were generated in one of two ways. The IR regions of *Acrotrema*, *Aextoxicon*, *Caryocar*, *Cissus*, *Cynomorium*, *Curatella*, *Davilla*, *Erythrospermum*, *Hibbertia*, *Malesherbia*, *Myrothamnus*, *Pinzona*, *Tetracera*, and *Tetrastigma* were sequenced using the ASAP primers of Dhingra and Folta (2005), following the protocols of Dhingra and Folta (2005), Jian et al. (2008), and Wang et al. (2009). Final sequences were annotated using default parameters in the program DOGMA (Wyman et al. 2004), with "Ns" placed in the alignment for missing sequence. The IR regions of *Antirrhinum*, *Aristolochia*, *Aucuba*, *Ehretia*, *Dillenia*, *Hedyosmum*, *Heisteria*, *Heuchera*, *Lilium*, *Liquidambar*, *Lonicera*, *Meliosma*, *Nelumbo*, *Nerium*, *Phoradendron*, *Rhododendron*, *Spathiphyllum*, *Tradescantia*, *Trithuria*, *Trochodendron*, *Vallisneria*, and *Ximenia* were generated as part of complete plastid genome sequencing using the chloroplast isolation and rolling-circle amplification protocols of Moore et al. (2006, 2007). For these taxa, sequencing was performed on a 454 GS 20 or GS FLX sequencer (454 Life Sciences, Branford, CT), following the protocols of Moore et al. (2006, 2007).

Alignment

To facilitate alignment, all IR sequences were divided into 40 gene, intron, and intergenic spacer regions, and each region was aligned individually. Not all regions were present for all taxa; table B1, in the online edition of the *International Journal of Plant Sciences*, indicates the regions missing for each taxon, as well as the percentage of data present per taxon and per region. For taxa lacking the IR (e.g., *Pinaceae* and the IR-lacking clade of Fabaceae), gene regions corresponding to the IR were extracted and aligned. Intron and spacer regions were aligned initially with Clustal X (Jeanmougin et al. 1998) with default settings; these alignments were then adjusted manually with MacClade 4.08 (Maddison and Maddison 2003). In regions of unambiguous alignment, misaligned indel boundaries in noncoding sequence were shifted according to the criteria of Kelchner (2000). All IR coding regions were aligned manually, with the 86-taxon alignments of Moore et al. (2010) as a starting point. Inversions are frequently encountered in noncoding regions of plastid DNA (Graham et al. 2000; Borsch and Quandt 2009), and consequently all 40 individual region alignments were thoroughly searched at least twice by eye for inversions, which were reversed and complemented before analysis. Regions that were globally difficult to align were excluded, as were all insertions present in only one or a few taxa. It was difficult to align some stretches of the IR for a few taxa with elevated rates of evolution (e.g., portions of intergenic spacers and *ycf2* in conifers and *Gnetophyta*). To preserve the phylogenetic signal present within remaining taxa at these positions, the difficult-to-align sequence regions in taxa with elevated evolutionary rates were moved into manually created alignment gaps that were subsequently excluded from analyses. The final aligned length (after character exclusion) for each region is provided in table B2 in the online edition of the *International Journal of Plant Sciences*.

Phylogenetic Analyses

MP and ML analyses were performed. MP analyses were conducted in PAUP* 4.0b10 (Swofford 2000). Shortest trees were obtained by means of heuristic searches with 1000 random taxon addition replicates and tree-bisection-reconnection (TBR) branch swapping and saving all shortest trees per replicate. Bootstrap support (BS) for relationships (Felsenstein 1985) was estimated from 1000 bootstrap replicates using 10 random taxon additions per replicate, with TBR branch swapping (saving all trees).

ML analyses were implemented in RAxML, version 7.2.6 (Stamatakis 2006). All ML analyses used the general time-reversible with GAMMA distributed rates (GTRGAMMA) model of nucleotide sequence evolution. We conducted both unpartitioned and partitioned analyses. Several partitioning schemes were investigated, involving between three and eight data partitions and generally separating protein-coding, RNA-coding, and noncoding characters. Likelihood ratio tests determined that the eight-partition analysis produced significantly better likelihood scores than other partitioning schemes (including a single partition, hereafter termed "unpartitioned"). For all analyses, we used a single set of branch lengths for all partitions, because missing data created unrealistic branch lengths when branch lengths were estimated on a per-partition basis. The eight partitions were (1) first and second codon positions of all coding genes except *ycf1* and *ycf2*, (2) third codon positions of all coding genes except *ycf1* and *ycf2*, (3) first and second codon positions of *ycf1* and *ycf2*, (4) third codon positions of *ycf1* and *ycf2*, (5) rRNA genes, (6) tRNA genes (excluding tRNA introns), (7) intergenic spacers, and (8) introns. For both unpartitioned and eight-partition analyses, 50 independent ML searches were conducted. Bootstrap support was estimated with 800 bootstrap replicates. Bootstrap proportions were drawn on the tree with highest likelihood score from the 50 independent searches.

Finally, to evaluate how the taxon sampling in our 244-taxon analyses influenced topology and phylogenetic support values, compared to taxon sampling in the most taxon-rich plastid phylogenomic analyses currently available (Moore et al. 2010), we analyzed a smaller IR data set of 87 species that corresponds in taxon sampling to the 86-taxon, 83-gene analyses of Moore et al. (2010), with *Cynomorium* also included. ML bootstrap analyses were conducted as above, using eight partitions.

Results

Data Set Characteristics

The total aligned length of the IR data set, after character exclusion, was 24,702 characters. The total amount of missing data (including gaps) after character exclusion was

16.1%. General data set characteristics, including aligned length and number of parsimony-informative characters for each of the 40 regions of the IR, are given in table B2. The regions with the highest numbers of parsimony-informative characters per aligned sequence length were spacer regions and the protein-coding genes *ycf1* and *ycf2*, whereas RNA genes had the lowest numbers. All 40 regions (including spacer regions) were at least partially alignable across all *Angiospermae*, and most regions (including all coding regions and introns) were alignable across *Acrogymnospermae* and *Angiospermae*. The number of included characters per region for each taxon is given in table B2. The full data set has been deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S10503) and is also available from the authors upon request. The newly generated IR sequences are available in GenBank (app. A).

ML Results for the 244-Taxon Data Matrix

ML analyses of the 244-taxon matrix under the eight-partition and unpartitioned strategies yielded nearly identical topologies, with very similar BS values throughout (figs. 1, 2; fig. B1, available in the online edition). Consequently, we focus below primarily on results from the eight-partition ML analysis; unless otherwise noted, the BS proportions given below refer to this analysis (fig. 2).

Topology Overview

Within *Acrogymnospermae*, the branch separating *Cycas* and *Ginkgo* from the remaining *Spermatophyta* was supported at BS = 72% (fig. 2*A*). Conifers + Gnetophyta formed a clade with 100% ML BS. Within this clade, *Cryptomeria* (*Cupressophyta*) + *Gnetophyta* formed a clade with BS = 100%.

The root of *Angiospermae* was poorly supported in both partitioned and unpartitoned analyses. Both found *Amborella* to be sister to the remaining *Angiospermae*, though with BS < 50% (figs. 1, 2, B1). All 50 partitioned replicate ML searches found this result and were identical throughout except for the placement of *Rivina*, *Sarcobatus*, and *Phytolacca*. However, in the unpartitioned searches, only 10 of the 50 replicates found *Amborella* to be sister to the remaining *Angiospermae*, while the other 40 found *Amborella* + Nymphaeales to be sister to the remaining *Angiospermae*. However, the tree with the highest ln *L* score from the unpartitioned analyses did place *Amborella* as sister to the remaining *Angiospermae*. Following these two lineages, Austrobaileyales + *Mesangiospermae* formed a strongly supported clade (BS = 100%; fig. 2*A*).

Within *Mesangiospermae*, five major strongly supported (BS = 100%) basal clades were recovered: *Monocotyledoneae*, *Magnoliidae*, *Eudicotyledoneae*, *Ceratophyllum*, and Chloranthaceae. However, relationships among them were not strongly supported (fig. 2*A*). In the ML tree, *Monocotyledoneae*, *Magnoliidae*, and a *Ceratophyllum* + Chloranthaceae clade were successively sister to *Eudicotyledoneae*, but only the branch uniting *Eudicotyledoneae*, *Ceratophyllum*, and Chloranthaceae (BS = 76%) had BS > 55% in the 244-taxon tree (fig. 2*A*). Within *Monocotyledoneae*, *Acorus* (Acoraceae) was sister to a clade consisting of Alismatales + all remaining *Monocotyledoneae* (BS = 100%; fig. 2*A*). Relationships among Dioscoreales (represented by *Dioscorea*), Liliales (represented by *Lilium*), Asparagales, and *Commelinidae* were weakly supported, although the lattermost clade was well supported (BS = 98%; fig. 2*A*). Within *Magnoliidae*, Canellales + Piperales (BS = 63%) were sister to Magnoliales + Laurales (BS = 98%; fig. 2*A*).

Eudicotyledoneae

Within *Eudicotyledoneae*, Ranunculales (BS = 100%) were sister to the remaining *Eudicotyledoneae* with strong support (BS = 99%), followed by a *Meliosma* (Sabiaceae) + Proteales clade (fig. 2*A*), which received BS < 52% in all IR ML analyses. The IR data strongly supported (BS = 100%) a clade consisting of Buxales, Trochodendrales, and *Gunneridae* (=core eudicots), with the latter two clades strongly supported as sisters (BS = 93% in partitioned and unpartitioned 244-taxon analyses; BS = 98% in the 87-taxon tree; figs. 2, 3, B1).

Within *Gunneridae* (BS = 100%), the well-supported Gunnerales (100%) were sister to the strongly supported *Pentapetalae* (BS = 98%; fig. 2*A*). *Pentapetalae* were resolved into three major clades: (1) Dilleniaceae; (2) a "superasterid" clade consisting of Santalales, *Caryophyllales*, Berberidopsidales, and *Asteridae*; and (3) a "superrosid" clade consisting of Vitaceae, Saxifragales, and *Rosidae* (fig. 2*A*). Support for the individual monophyly of both superasterids and superrosids was moderately high (BS ≈ 85% for both clades in both partitioned and unpartitioned 244-taxon trees; figs. 2*A*, B1). Support for the sister relationship of Dilleniaceae to superasterids + superrosids was moderate in both partitioned (BS = 70%) and unpartitioned (BS = 79%) 244-taxon analyses (figs. 2*A*, B1).

Superrosids

In all trees, Saxifragales were sister to *Rosidae* (figs. 2, 3, B1), but this clade was only weakly supported (BS = 59%) in the 244-taxon trees (fig. 2*B*). The monophyly of Saxifragales was well supported (BS = 92%; fig. 2*B*). In the 244-taxon tree, Peridiscaceae were sister to a well-supported clade (BS = 100%) containing all remaining Saxifragales, which were divided into two strongly supported sister clades (fig. 2*B*): (1) Paeoniaceae plus the woody clade (BS = 98%), comprising Cercidiphyllaceae, Daphniphyllaceae, Altingiaceae, and Hamamelidaceae; and (2) core Saxifragales (BS = 100%), comprising Crassulaceae, Haloragaceae s.l., and the Saxifragaceae alliance (Saxifragaceae, Grossulariaceae, Iteaceae, and Pterostemonaceae).

Rosidae were strongly supported as monophyletic (BS = 94%) and were resolved into two major subclades, each of which was also strongly supported (fig. 2*B*): (1) *Fabidae* (fabids/eurosids I; BS = 97%), which included the nitrogen-fixing clade (Fabales, Cucurbitales, Fagales, and Rosales), the COM clade (Celastrales, Oxalidales, and Malpighiales), and Zygophyllales; and (2) *Malvidae* (BS = 95%), which included Brassicales, Malvales, Huerteales, Sapindales, and Picramniaceae as well as Crossosomatales, Myrtales, and Geraniales (we follow Judd et al. [forthcoming] in recognizing

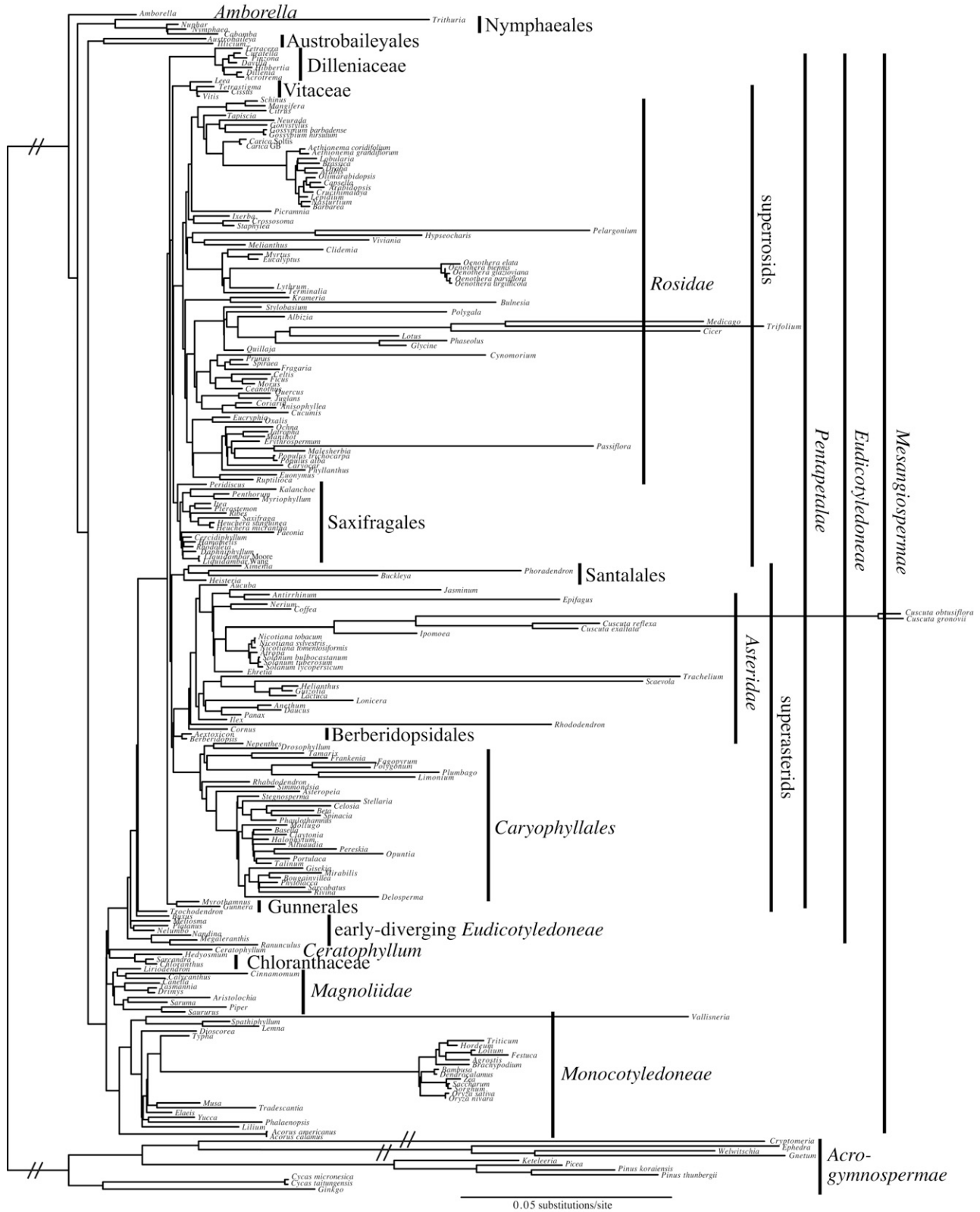

Fig. 1 Phylogram depicting the best maximum likelihood tree ($\ln L = -424{,}640.15$) for the 244-taxon data set using eight partitions. Hatch marks indicate branches that have been shortened to allow the topology to fit on one page.

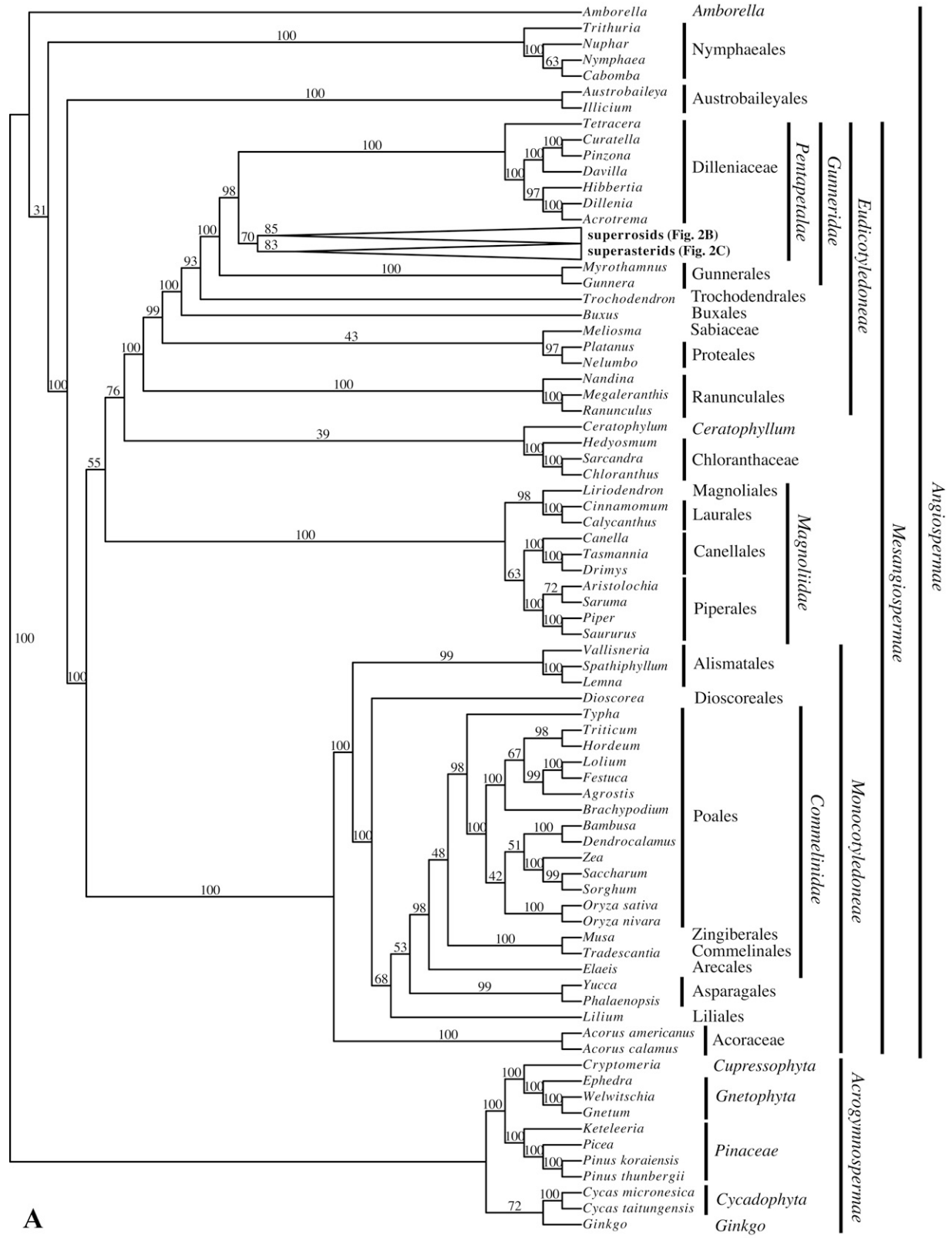

Fig. 2 Cladogram depicting the best maximum likelihood (ML) tree for the 244-taxon data set using eight partitions. Numbers above branches are ML bootstrap proportions. *A*, Deeper-level relationships, with superrosids and superasterids collapsed; *B*, superrosid relationships; *C*, superasterid relationships.

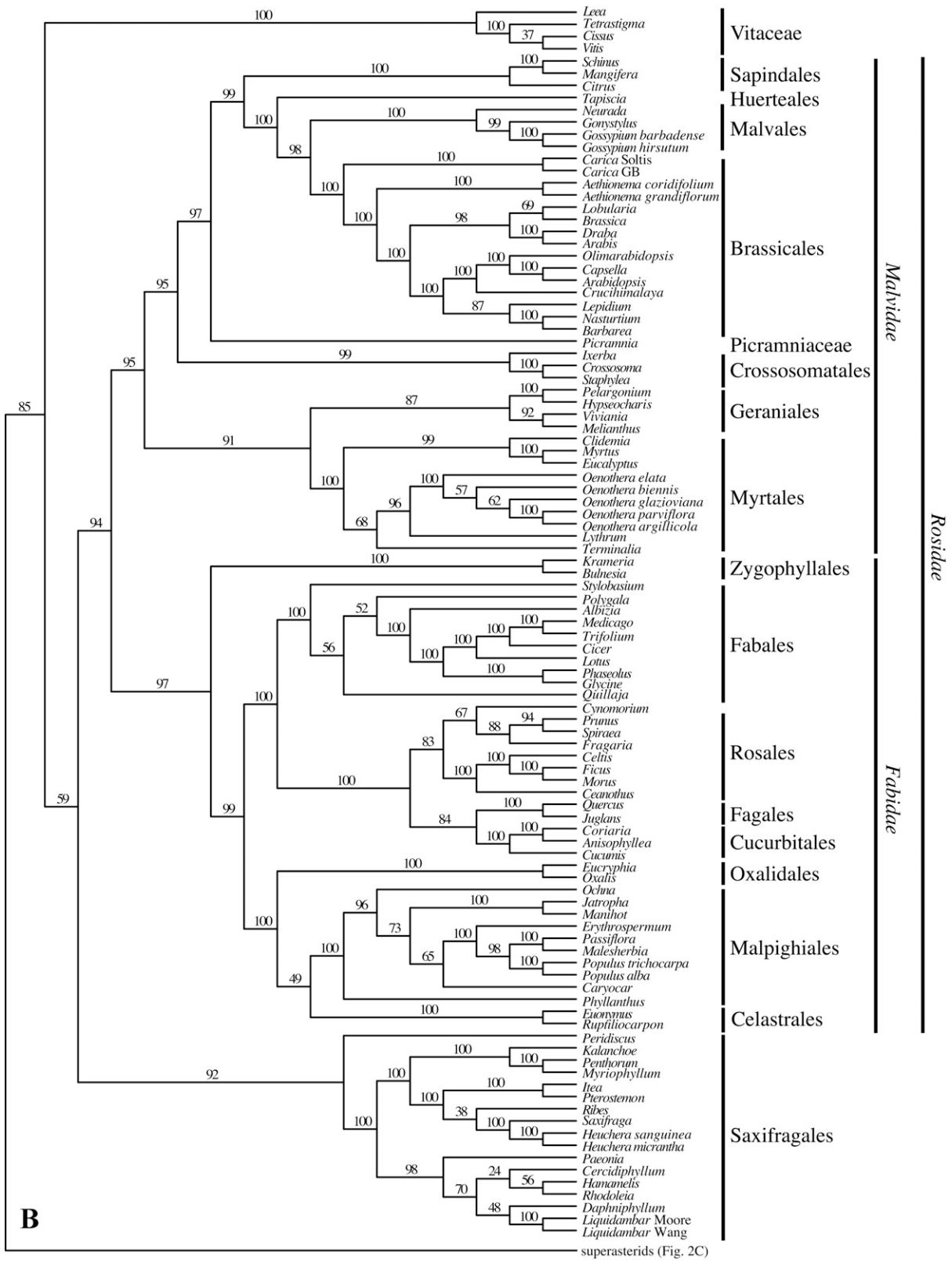

Fig. 2 (*Continued*)

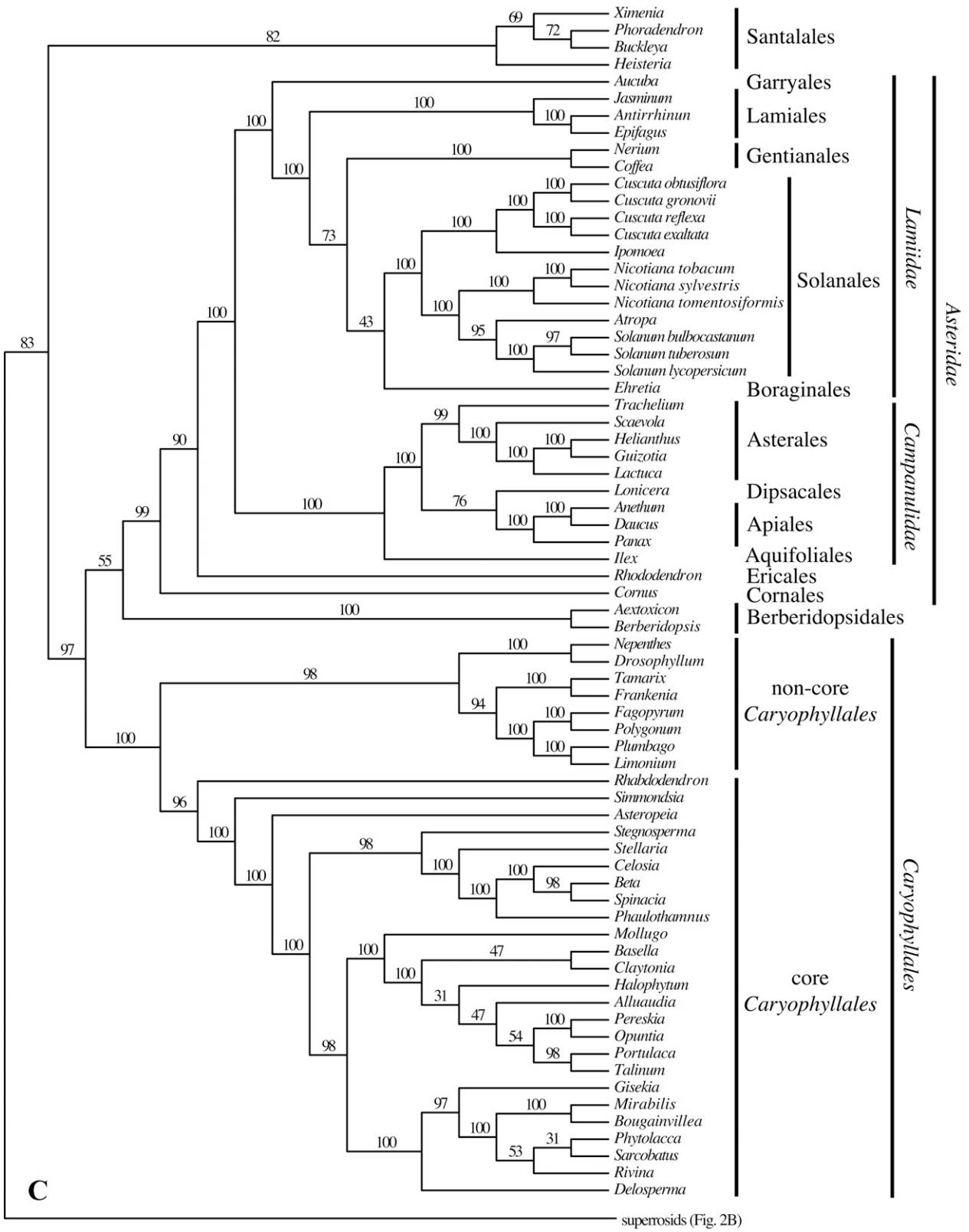

Fig. 2 (*Continued*)

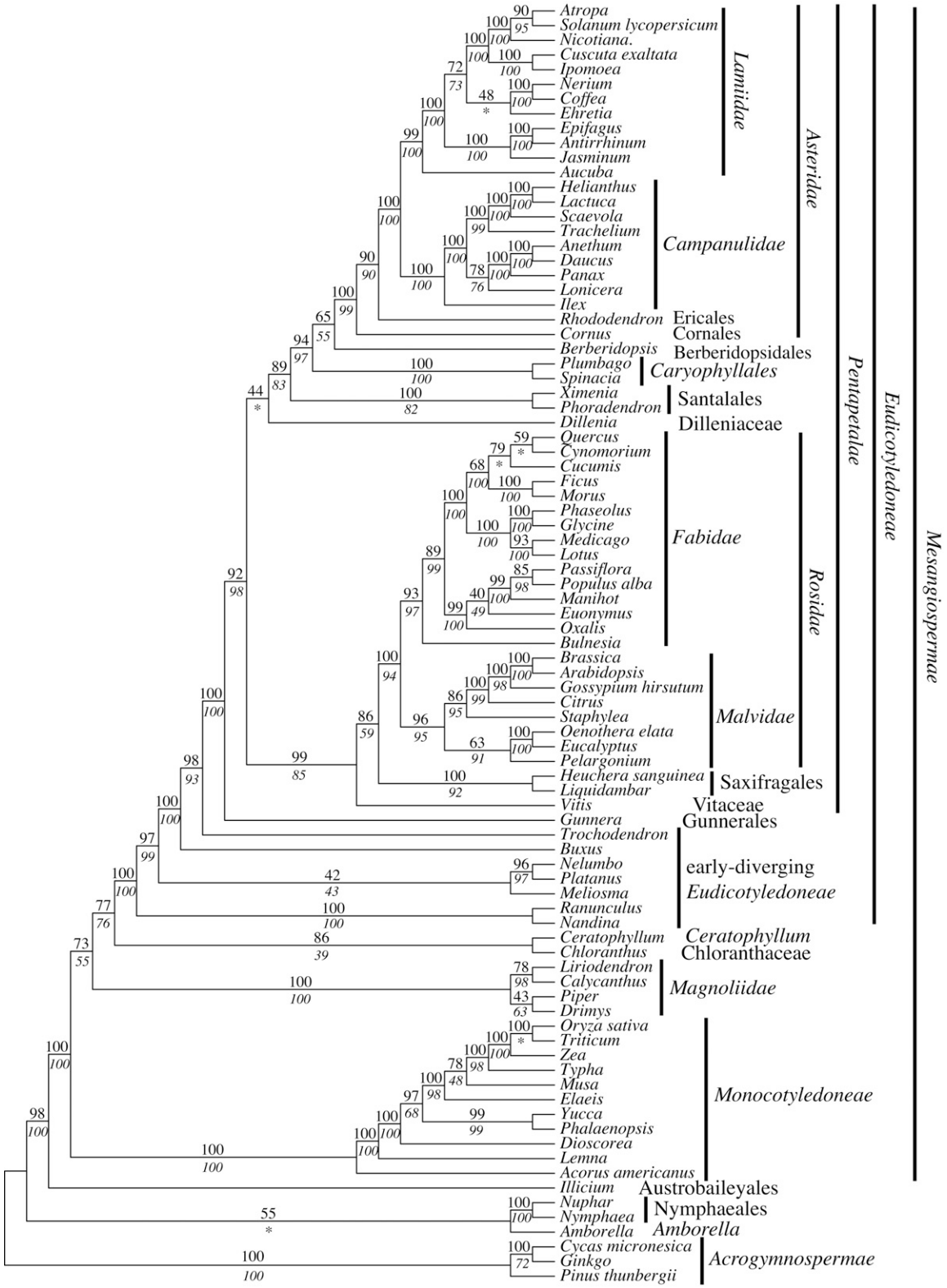

Fig. 3 Cladogram depicting the best maximum likelihood (ML) tree for the 87-taxon data set (ln$L$ = −238,578.95). Numbers above branches are ML bootstrap support (BS) proportions based on the 87-taxon data set; italicized numbers below the branches indicate ML BS proportions for the 244-taxon, eight-partition analysis for the corresponding clades (i.e., from fig. 2). For example, the italicized ML BS proportion for the branch uniting *Ximenia* and *Phoradendron* (Santalales) indicates the BS proportion for the branch uniting all species of Santalales in the 244-taxon analysis (fig. 2). Asterisks under branches indicate branches that are not shared between the 87- and 244-taxon analyses.

an expanded *Malvidae*). Relationships among major subclades within *Rosidae* were generally well supported, with the exception of relationships among the orders comprising the COM clade. Within this clade, Celastrales + Malpighiales formed a clade with BS < 50%. The holoparasite *Cynomorium* was resolved within Rosales as sister to Rosaceae with weak support (BS = 67% for Rosaceae + *Cynomorium*; BS = 83% for the entire clade of Rosales including *Cynomorium*; fig. 2*B*).

Superasterids

Four major basal clades of superasterids were recovered in all ML analyses: Santalales, Berberidopsidales, *Caryophyllales*, and *Asteridae*. Santalales (BS = 82%) were sister to a strongly supported clade containing all remaining superasterids (BS = 97%). Within the latter clade, *Caryophyllales* were sister to a weakly supported Berberidopsidales + *Asteridae* clade (BS = 55% in the partitioned 244-taxon analysis; BS = 68% in the unpartitioned analysis; figs. 2*C*, B1). *Caryophyllales* were strongly supported as monophyletic in the 244-taxon tree (BS = 100%) and were resolved into two well-supported sister clades (fig. 2*C*): noncore *Caryophyllales* (BS = 98%), consisting of Polygonaceae, Plumbaginaceae, Tamaricaceae, Frankeniaceae, and the carnivorous clade (represented by Nepenthaceae and Drosophyllaceae in the IR tree); and core *Caryophyllales* (BS = 96%), consisting of all remaining families. Within the strongly supported but more sparsely sampled *Asteridae* (BS = 99%), Cornales were sister to the remaining *Asteridae* (BS = 90%; fig. 2*C*). Both of the major core *Asteridae* clades, *Lamiidae* and *Campanulidae*, were strongly supported (BS = 100%). Within *Campanulidae*, Aquifoliales (represented by *Ilex*) were sister to the remaining members of the clade (BS = 100%; fig. 2*C*).

Influence of Taxon Sampling on Topology and Branch Support

Although ML topologies based on analyses of the 87- and 244-taxon matrices were highly congruent, three important topological differences were observed between them (figs. 2, 3). In the 87-taxon tree, *Amborella* was sister to Nymphaeales, although with weak support (BS = 55%; fig. 3). Likewise, Dilleniaceae were sister to superasterids in the 87-taxon tree rather than to superrosids + superasterids, but again with poor support (BS = 44%; fig. 3). Finally, *Cynomorium* formed a moderately supported (BS = 79%) clade with *Quercus* and *Cucumis* in the 87-taxon tree, whereas it was weakly supported as sister to Rosaceae in the 244-taxon tree.

BS values occasionally varied substantially between the 87- and 244-taxon ML trees. BS values were sometimes lower in the 244-taxon partitioned-analysis tree in regions where taxon sampling was increased; examples include the branch separating Vitaceae from Saxifragales and *Rosidae* (BS = 86% in the 87-taxon tree and 59% in the 244-taxon tree) and the branches separating the five major lineages of *Mesangiospermae* (which declined from BS values of 73%–86% in the 87-taxon tree to 39%–76% in the 244-taxon tree; figs. 2, 3). However, in some instances BS improved markedly with additional taxon sampling. ML BS for the sister relationship of Dilleniaceae to the remaining *Pentapetalae* improved from <50% in the 87-taxon tree to 70%–79% (depending on partitioning scheme) in the 244-taxon trees (figs. 2, 3). Likewise, support for a clade consisting of Cucurbitales, Fagales, Rosales, and *Cynomorium* improved significantly with additional taxon sampling (from BS = 68% to BS = 100%), as did support for a clade consisting of Myrtales and Geraniales (from BS = 63% to BS = 91%; figs. 2, 3). BS for relationships among the four major clades of *Magnoliidae* also improved markedly: support for Magnoliales + Laurales improved from 78% to 98%, while support for Piperales + Canellales improved from 43% to 63% (figs. 2, 3).

MP Topology

Relationships within the MP tree were largely congruent with those within the ML tree, but there were several key topological differences (fig. B2). For example, in the MP tree Nymphaeales were not monophyletic; instead, *Trithuria* (Hydatellaceae) and the remaining Nymphaeales were recovered as successive sisters to all other *Angiospermae*, with BS = 79% (fig. B3). Several taxa characterized by long branches also formed clades in the MP tree that were absent in the ML trees. These included clades of *Vallisneria* (Hydrocharitaceae, Alismatales) and grasses, with BS = 69% (the remaining Alismatales were sister to this clade, but with BS < 50%); *Rhododendron* (Ericaceae), *Scaevola* (Goodeniaceae), and *Trachelium* (Campanulaceae), with BS = 53%; and Zygophyllales within Geraniales, with BS = 65% (fig. B3). *Monocotyledoneae* displayed other differences between the MP and ML topologies as well; most notably, in the MP tree, *Typha* (Typhaceae) was sister to *Lilium* (Liliaceae) rather than to Poaceae (fig. B3). Other parts with weaker BS also differed between MP and ML trees. For example, in the MP tree, *Paeonia* (Paeoniaceae) was sister to *Cercidiphyllum* (Cercidiphyllaceae) rather than to the woody Saxifragales clade, and Oxalidales and Celastrales were sisters in the MP tree (fig. B3) whereas Celastrales was sister to Malpighiales in the ML trees (figs. 2, 3).

Discussion

Overall, the IR ML results are highly congruent with those of recent gene-rich plastid studies (Jansen et al. 2007; Moore et al. 2007, 2010), suggesting that we have converged on a stable and largely well-supported view of relationships at deeper levels of the angiosperm tree of life. Relationships among *Amborella*, Nymphaeales, Austrobaileyales, and *Mesangiospermae* are congruent topologically with those found in most previous studies (Leebens-Mack et al. 2005; Jansen et al. 2007; Moore et al. 2007; Saarela et al. 2007; Soltis et al. 2007; Moore et al. 2010), although the positions of *Amborella* and Nymphaeales are inconsistent, depending on optimality criterion and taxon sampling (figs. 2, 3, B1, B3). The root of *Angiospermae* has been similarly weakly or moderately supported in many other recent phylogenetic analyses based on smaller numbers of genes (Barkman et al. 2000; Zanis et al. 2002; Saarela et al. 2007; Worberg et al. 2007;

Graham and Iles 2009), although recent 61- and 83-gene plastid analyses have provided 100% ML BS for an *Amborella* + remaining *Angiospermae* topology (Jansen et al. 2007; Moore et al. 2007, 2010). Relationships within (but not among; see below) the five major clades of *Mesangiospermae* are also highly congruent between the IR and other recent plastid genome analyses. Deeper-level relationships within *Magnoliidae* and *Monocotyledoneae* are identical between our IR tree and other recent analyses (Cai et al. 2006; Chase et al. 2006; Givnish et al. 2006; Graham et al. 2006; Jansen et al. 2007; Li and Zhou 2007; Moore et al. 2007, 2010; Saarela et al. 2007), while *Eudicotyledoneae* relationships are nearly so. For example, *Pentapetalae* are resolved into three major clades (Dilleniaceae, superasterids, and superrosids) in both our IR ML tree (fig. 2*A*) and the ML trees of Burleigh et al. (2009) and Moore et al. (2010). Likewise, almost all deeper relationships within *Caryophyllales*, *Rosidae*, and *Asteridae* are congruent with those of recent large-scale analyses (Bremer et al. 2002; Worberg et al. 2007, 2009; Brockington et al. 2009; Schäferhoff et al. 2009; Wang et al. 2009; Moore et al. 2010).

Differences in topology between the IR tree and previous trees based on 61–83 plastid genes (Jansen et al. 2007; Moore et al. 2007, 2010) are confined to nodes that were only weakly or moderately supported in the IR and/or the 61–83-gene analyses. For example, Jansen et al. (2007) and Moore et al. (2007, 2010) recovered Chloranthaceae + *Magnoliidae* clade as sister to a *Monocotyledoneae* + (*Ceratophyllum* + *Eudicotyledoneae*) clade, using 61–83 plastid genes, although relationships were generally only moderately supported in these analyses. Moreover, topology tests could not reject a number of alternative relationships among these five clades in the 61-gene analyses of Moore et al. (2007). Likewise, the well-supported position of Buxales as sister to *Gunneridae* in the IR tree (fig. 2*A*) also differs from the findings of many previous analyses, which have recovered Buxales as sister to *Gunneridae*, although generally with only weak or moderate support (Qiu et al. 2006; Worberg et al. 2007; Burleigh et al. 2009; Moore et al. 2010). The position of Dilleniaceae has never been consistently strongly supported in ML analyses (Qiu et al. 2006; Worberg et al. 2007; Moore et al. 2010), and the position of Dilleniaceae as sister to the remaining *Pentapetalae* is similarly only moderate in the 244-taxon IR trees (fig. 2*A*). Moreover, topology tests using the 83-gene data matrix of Moore et al. (2010) could not reject the sister relationships of Dilleniaceae to all other *Pentapetalae* or to superasterids, which was the position recovered for Dilleniaceae in both the 87-taxon IR analyses and the five-gene, 567-taxon analyses of Burleigh et al. (2009).

Other examples of weakly supported incongruence between the IR tree and other recent analyses include the position of Berberidopsidales, the relationships among the three basalmost lineages of superrosids (*Rosidae*, Saxifragales, and Vitaceae), and relationships among the COM clade. In contrast to the position of Berberidopsidales as sister to *Asteridae* (with Caryophyalles sister to this clade) in the IR tree (fig. 2*C*), Moore et al. (2010) found stronger support for a Berberidopsidales + (*Caryophyllales* + *Asteridae*) clade. The positions of Saxifragales and Vitaceae with respect to *Rosidae* are also weakly supported in the IR tree (fig. 2*B*). While these three lineages have consistently formed a strongly supported clade in recent large-scale analyses (Qiu et al. 2006; Worberg et al. 2007; Burleigh et al. 2009; Moore et al. 2010), relationships among them have varied without strong support and consequently remain unsettled. Finally, in the IR ML tree, Celastrales + Malpighiales formed a clade with BS < 50%, in contrast to the findings of other recent large-scale studies, all of which have recovered differing topologies among the major subclades within the COM clade: Celastrales were sister to Oxalidales in Moore et al. (2010), Celastrales + Huaceae were sister to Malpighiales + Oxalidales in Wang et al. (2009), and Celastrales were strongly supported as sister to the remaining members of the COM clade in the 13-gene, taxon-rich analyses of the COM clade in Wurdack and Davis (2009).

The generally lower support observed in the IR trees for branches that have been difficult to resolve in previous analyses likely reflects the relatively low evolutionary rate within the IR (Graham and Olmstead 2000; Jian et al. 2008), which results in fewer characters supporting individual branches, despite the presence of nearly 25,000 characters in the analysis. Nevertheless, this low rate has proved advantageous in the reconstruction of relationships within some groups of *Eudicotyledoneae*. In IR-based analyses of Saxifragales (Wang et al. 2009) and *Rosidae* (Moore et al. 2010), MP and ML topologies were identical, even in difficult areas of the tree (e.g., the position of Paeoniaceae within Saxifragales and the position of Zygophyllales within *Rosidae*). This suggests that the lower evolutionary rate of the IR may in general reduce problems in topology and support due to systematic error, particularly for MP (Felsenstein 1978; Bull et al. 1993), although Yang (1998) found that in simulated data sets of varying evolutionary rates, ML also recovers a higher percentage of correct partitions when more slowly evolving sequences are analyzed. Recent simulations of phylogenetic informativeness as a function of evolutionary rate support our empirical observation: Klopfstein et al. (2010) found that sequences with slower rates of evolution become increasingly advantageous for phylogenetic reconstruction as the number of taxa in an analysis increases because the amount of homoplasy is much less dependent on taxon number in slowly evolving sequences than in more quickly evolving sequences.

However, the lack of congruence between the MP and ML trees reported here in some deeper angiosperm relationships indicates that even within the IR, systematic error may sometimes be a problem. This is not very surprising, given that most of the incongruence between the MP and ML trees is confined to lineages that are characterized by long branches, such as *Trithuria*, *Vallisneria*, Poaceae, *Scaevola*, and *Trachelium* (figs. 1, B2). Most of these lineages either have a reduced plastid genome as a result of being parasitic (e.g., *Cuscuta* [McNeal et al. 2007], *Epifagus* [dePamphilis and Palmer 1990], and *Phoradendron* [M. J. Moore, P. S. Soltis, and D. E. Soltis, unpublished data]) or possess a highly rearranged plastid genome (e.g., *Scaevola* [Downie and Palmer 1992], *Trachelium* [Haberle et al. 2008], *Pelargonium* [Chumley et al. 2006], *Vallisneria* [M. J. Moore, P. S. Soltis, and D. E. Soltis, unpublished data], and *Trifolium* [Cai et al. 2008]). Reduction and rearrangement of the plastid genome have

both been shown to be correlated with an elevated rate of nucleotide substitution (Jansen et al. 2007; Guisinger et al. 2008). Contraction or loss of the IR might also be expected to cause a substantial increase in evolutionary rate in genes normally found in the IR, by releasing these genes from recombination. Two groups of seed plants—conifers and the IR-lacking clade of papilionoid legumes (represented by *Cicer*, *Medicago*, and *Trifolium* in the 244-taxon tree)—are known to have completely or essentially completely lost the IR (Raubeson and Jansen 1992; Wojciechowski et al. 2004; Jansen et al. 2008), and these two groups are indeed characterized by extremely long branches in the IR ML tree (fig. 1). A similar phenomenon might be expected in regions that have migrated outside of the IR because of IR contraction, which has occurred in a handful of taxa included in our analyses: *rpl2* and *rpl23* lie outside the IR in *Anethum* (Plunkett and Downie 2000); *ycf1*, *rpl2*, and the 3′ end of *rpl23* lie outside the IR in *Lonicera* (M. J. Moore, P. S. Soltis, and D. E. Soltis, unpublished data); and *trnR-ACG*, *trnN-GUU*, and *ycf1* lie outside the IR in *Phoradendron* (M. J. Moore, P. S. Soltis, and D. E. Soltis, unpublished data). No other examples of IR loss or contraction in our data set are evident, although we would not be able to detect such a phenomenon in IR sequences that we sequenced conventionally using ASAP primers. An elevated rate of evolution due to IR contraction or loss might lead to problems in phylogenetic estimation because of rate heterogeneity among lineages (i.e., heterotachy; Philippe et al. 2005), although none of the taxa with a reduced or absent IR included in the 244-taxon ML analyses occurs in a phylogenetic position that is unusual compared to that in trees from previous analyses based largely on normally single-copy plastid genes (Burleigh et al. 2009; Moore et al. 2010).

In other regions where the 244-taxon IR tree and the 83-gene plastid genome ML tree of Moore et al. (2010) differ, such as the position of Buxales with respect to Trochodendrales and *Gunneridae* and the position of Dilleniaceae within *Pentapetalae*, it is difficult to say whether the more slowly evolving IR provides the "true" phylogenetic signal (i.e., species tree) for these groups, particularly given the lower BS values that characterize these nodes in the IR and/or the 83-gene analyses. Future combined IR + 83-gene analyses with additional taxon sampling may resolve this issue.

The strong congruence in overall topology and support values observed between the trees based on analyses of the complete 244-taxon and reduced 87-taxon matrices suggests that most relationships and support values obtained in the similar 86-taxon, 83-gene ML tree of Moore et al. (2010) may not be significantly influenced by the addition of new taxa. Nevertheless, the few significant changes in support values observed at certain key nodes between the 87-taxon and 244-taxon IR data sets (e.g., increased support for Dilleniaceae + remaining *Pentapetalae* and reduced support for basal Mesangiospermae; figs. 2, 3) indicate that increased taxon sampling in future plastid phylogenomic analyses may dramatically influence support values in some currently poorly sampled regions of the 83-gene tree. Collectively, these shifts in support values reiterate the importance of carefully considering taxon sampling in the design of phylogenomic sequencing projects. Future plant phylogenomic analyses should be guided by prior phylogenetic studies with far more taxa but far fewer genes. In particular, investigators should include samples representing all major clades identified from these taxon-rich analyses, with a special effort to span the basal node of each of these clades, where known.

Although not explicit foci of the current work, the position of *Gnetophyta* as sister to *Cupressophyta* and the position of the holoparasite *Cynomorium* as sister to Rosaceae in our IR analyses are also worth noting (figs. 1–3). In all recent analyses, *Gnetophyta* have been closely associated with *Coniferae*, albeit in differing positions (reviewed in Mathews 2009). *Gnetophyta* have most frequently been recovered as sister to *Pinaceae* (plastid structural data seem to support this as well; Braukmann et al. 2009), but they have also been resolved as sister to *Cupressophyta* and to *Cupressophyta* + *Pinaceae*. IR branch lengths in the ML tree were quite long for both *Cryptomeria* and *Gnetophyta* (fig. 1), reflecting a highly elevated rate of evolution, and it is therefore impossible with the available taxon sampling to rule out the effects of systematic error causing *Cryptomeria* and *Gnetophyta* to be sister, even under ML. Finally, our results confirm the IR placement of *Cynomorium* as sister to Rosaceae found in the study of Zhang et al. (2009), which contained limited taxon sampling. In contrast to plastid data, nuclear and mitochondrial data place *Cynomorium* within either Saxifragales or Sapindales (Nickrent et al. 2005; Barkman et al. 2007).

Conclusions

The overall congruence between our ML trees and those of recent plastid genome-scale studies improves confidence in the emerging view of deeper-level *Angiospermae* relationships, particularly given the improved taxon sampling in IR trees and the fact that much of the phylogenetic signal in the IR data set arises from regions (e.g., spacers and introns) that were not sampled across *Angiospermae* in previous genome-scale analyses. In the future, it will be valuable to combine IR sequences with as much alignable sequence as possible from the single-copy regions to examine phylogenetic signal rigorously throughout the plastid genome. Furthermore, our results emphasize that the IR, while not necessarily immune to problems due to rate heterogeneity, is nevertheless a valuable source of data at deeper levels of angiosperm phylogeny, particularly at the level of major clades such as *Rosidae* and Saxifragales. Our results further suggest that the IR may be useful for phylogenetic inference outside of *Angiospermae* as well. Moreover, because it is easily amplified with conserved primers, the IR can be a useful source of data for deeper-level phylogenetic studies when complete plastid genome sequence data are implausible or impractical. Efforts in the near future should aim at expanding taxonomic coverage of the IR in poorly sampled lineages such as *Monocotyledoneae* and *Asteridae*.

Acknowledgments

We thank three anonymous reviewers for helpful comments that improved the manuscript. We also thank Bob Jansen and Jim Leebens-Mack for access to unpublished ge-

nome sequences and Tom Givnish for leaf material of *Lilium superbum*. We acknowledge the University of Florida (UF) High-Performance Computing Center and the UF Genetics Institute for providing computational resources and support that have contributed to the research results reported in this article. We also thank Rob Ferl and Beth Laughner for lab space and general lab assistance and Gordon Burleigh for helpful correspondence. This study was carried out as part of the Angiosperm Tree of Life Project (NSF grant EF-0431266 to D. E. Soltis and P. S. Soltis, among others).

Appendix A

Taxa Included in Phylogenetic Analyses

The following information is given for each species in the list: APG III (2009) ordinal classification (for *Angiospermae*) or higher clade (*Acrogymnospermae*), family, species, voucher specimen (for all taxa newly deposited in GenBank), and GenBank accession number(s). An asterisk indicates species whose IR was sequenced for this study.

Acrogymnospermae

Cupressophyta, Cupressaceae, *Cryptomeria japonica* D. Don, . . ., NC_010548; *Cycadophyta*, Cycadaceae, *Cycas micronesica* K.D.Hill, . . ., EU016807, EU016810, EU016813, EU016817, EU016820, EU016823, EU016827, EU016845, EU016851, EU016861; *Cycadophyta*, Cycadaceae, *Cycas taitungensis* C. F. Shen, K. D. Hill, C. H. Tsou, & C. J. Chen, . . ., NC_009618; *Gnetophyta*, Ephedraceae, *Ephedra equisetina* Bunge, . . ., NC_011954; *Gnetophyta*, Gnetaceae, *Gnetum parvifolium* (Warb.) Cheng, . . ., NC_011942; *Gnetophyta*, Welwitschiaceae, *Welwitschia mirabilis* Hook.f., . . ., NC_010654; *Pinaceae*, Pinaceae, *Keteleeria davidiana* Beissn., . . ., NC_011930; *Pinaceae*, Pinaceae, *Picea sitchensis* Bong., . . ., NC_011152; *Pinaceae*, Pinaceae, *Pinus koraiensis* Siebold & Zucc., . . ., NC_004677; *Pinaceae*, Pinaceae, *Pinus thunbergii* Parl., . . ., NC_001631; *Ginkgo*, Ginkgoaceae, *Ginkgo biloba* L., . . ., DQ069410, DQ069446, DQ069494, EU016964, EU016966, EU016967, EU016969, EU016973, EU016976, EU016980.

Early-Diverging *Angiospermae*

Amborellales, Amborellaceae, *Amborella trichopoda* Baill., . . ., NC_005086; Austrobaileyales, Austrobaileyaceae, *Austrobaileya scandens* C. T. White, Y.-L. Qiu 90030 (IND), HQ664628; Austrobaileyales, Illiciaceae, *Illicium oligandrum* Merr. & Chun, . . ., NC_009600; Nymphaeales, Cabombaceae, *Cabomba caroliniana* A. Gray, M.-J. Yoo s.n., HQ664629; Nymphaeales, Hydatellaceae, *Trithuria filamentosa* Rodway,* B.G. Briggs 9859 (NSW), HQ664617; Nymphaeales, Nymphaeaceae, *Nuphar advena* Ait., . . ., NC_008788; Nymphaeales, Nymphaeaceae, *Nymphaea alba* L., . . ., NC_006050.

Ceratophyllum

Ceratophyllales, Ceratophyllaceae, *Ceratophyllum demersum* L., . . ., NC_009962.

Chloranthaceae

Chloranthales, Chloranthaceae, *Chloranthus spicatus* Makino, . . ., NC_009598; Chloranthales, Chloranthaceae, *Hedyosmum mexicanum* Cordem. ex Baill.,* M. J. Moore s.n., HQ664615; Chloranthales, Chloranthaceae, *Sarcandra chloranthoides* Gardn., Y.-L. Qiu 92002 (IND), HQ664633.

Eudicotyledoneae

Apiales, Apiaceae, *Anethum graveolens* L., . . ., EU016726, EU016729, EU016732, EU016736, EU016739, EU016742, EU016746, EU016764, EU016770, EU016780; Apiales, Apiaceae, *Daucus carota* L., . . ., NC_008325; Apiales, Araliaceae, *Panax ginseng* C.A.Mey., . . ., NC_006290; Aquifoliales, Aquifoliaceae, *Ilex cornuta* Lindl. & Paxton, M. J. Moore 308 (FLAS), HQ664579; Asterales, Asteraceae, *Guizotia abyssinica* (L.f.) Cass., . . ., NC_010601; Asterales, Asteraceae, *Helianthus annuus* L., . . ., NC_007977; Asterales, Asteraceae, *Lactuca sativa* L., . . ., NC_007578; Asterales, Campanulaceae, *Trachelium caeruleum* L., . . ., NC_010442; Asterales, Goodeniaceae, *Scaevola aemula* R.Br., . . ., EU017144, EU017147, EU017150, EU017154, EU017157, EU017159, EU017163, EU017180, EU017186, EU017196; Berberidopsidales, Aextoxicaceae, *Aextoxicon punctatum* Ruiz & Pav.,* M. W. Chase 959 (K), HQ664619; Berberidopsidales, Berberidopsidaceae, *Berberidopsis corallina* Hook.f., M. J. Moore 326 (FLAS), HQ664598; Boraginaceae, Boraginaceae, *Ehretia acuminata* R.Br., M. J. Moore 317 (FLAS), HQ664577; Brassicales, Brassicaceae, *Aethionema coridifolium* DC., . . ., NC_009265; Brassicales, Brassicaceae, *Aethionema grandiflorum* Boiss. & Hohen., . . ., NC_009266; Brassicales, Brassicaceae, *Arabidopsis thaliana* (L.) Heynh., . . ., NC_000932; Brassicales, Brassicaceae, *Arabis hirsuta* (L.) Scop., . . ., NC_009268; Brassicales, Brassicaceae, *Barbarea verna*

(Mill.) Asch., . . ., NC_009269; Brassicales, Brassicaceae, *Brassica rapa* L., . . ., DQ231548; Brassicales, Brassicaceae, *Capsella bursa-pastoris* (L.) Medik., . . ., NC_009270; Brassicales, Brassicaceae, *Crucihimalaya wallichii* (Hook.f. & Thompson) Al-Shehbaz, O'Kane & R. A. Price, . . ., NC_009271; Brassicales, Brassicaceae, *Draba nemorosa* L., . . ., NC_009272; Brassicales, Brassicaceae, *Lepidium virginicum* L., . . ., NC_009273; Brassicales, Brassicaceae, *Lobularia maritima* (L.) Desv., . . ., NC_009274; Brassicales, Brassicaceae, *Nasturtium officinale* W. T. Aiton, . . ., NC_009275; Brassicales, Brassicaceae, *Olimarabidopsis pumila* (Stephan) Al-Shehbaz, O'Kane & R.A.Price, . . ., NC_009267; Brassicales, Caricaceae, *Carica papaya* L. [GB], . . ., NC_010323; Brassicales, Caricaceae, *Carica papaya* L. [Soltis], D. Soltis 2707 (FLAS), HQ664551; Buxales, Buxaceae, *Buxus microphylla* Siebold & Zucc., . . ., NC_009599; Caryophyllales, Achatocarpaceae, *Phaulothamnus spinescens* A. Gray, J. Manhart s.n., HQ664653; Caryophyllales, Aizoaceae, *Delosperma napiforme* Schwantes, S. Brockington S700, HQ664643; Caryophyllales, Amaranthaceae, *Beta vulgaris* L., . . ., EF534108; Caryophyllales, Amaranthaceae, *Celosia cristata* L., Y.-L. Qiu 94153 (IND), HQ664641; Caryophyllales, Amaranthaceae, *Spinacia oleracea* L., . . ., NC_002202; Caryophyllales, Asteropeiaceae, *Asteropeia micraster* Hallier f., Civeyrel s.n. (K), HQ664638; Caryophyllales, Basellaceae, *Basella alba* L., Y.-L. Qiu 02055, HQ664639; Caryophyllales, Cactaceae, *Opuntia microdasys* (Lehm.) Pfeiff., D. Soltis s.n., HQ664651; Caryophyllales, Cactaceae, *Pereskia aculeata* Mill., D. Soltis 2645 (FLAS), HQ664652; Caryophyllales, Caryophyllaceae, *Stellaria media* (L.) Vill., D. Soltis s.n., HQ664662; Caryophyllales, Didiereaceae, *Alluaudia ascendens* Drake, Y.-L. Qiu 97030 (IND), HQ664637; Caryophyllales, Drosophyllaceae, *Drosophyllum lusitanicum* (L.) Link, J. Cortez s.n. (GBG), HQ664644; Caryophyllales, Frankeniaceae, *Frankenia pulverulenta* L., Collenette 6/93 (K), HQ664645; Caryophyllales, Halophytaceae, *Halophytum ameghinoi* Speg., Tortosa, Bartoli, and Chubut s.n., HQ664646; Caryophyllales, Molluginaceae, *Gisekia africana* Kuntze, . . ., HQ664636; Caryophyllales, Molluginaceae, *Mollugo verticillata* L., M. J. Moore 321 (FLAS), HQ664649; Caryophyllales, Montiaceae, *Claytonia virginica* L., . . ., HQ664642; Caryophyllales, Nepenthaceae, *Nepenthes alata* Blanco, cultivated in Botany Department greenhouse, University of Florida, HQ664650; Caryophyllales, Nyctaginaceae, *Bougainvillea glabra* Choisy, M. J. Moore 323 (FLAS), HQ664640; Caryophyllales, Nyctaginaceae, *Mirabilis jalapa* L., D. Soltis 2638 (FLAS), HQ664648; Caryophyllales, Phytolaccaceae, *Phytolacca americana* L., S. Brockington s.n., HQ664654; Caryophyllales, Phytolaccaceae, *Rivina humilis* L., D. Soltis 2643 (FLAS), HQ664658; Caryophyllales, Plumbaginaceae, *Limonium arborescens* Kuntze, M. J. Moore 318 (FLAS), HQ664647; Caryophyllales, Plumbaginaceae, *Plumbago auriculata* Lam., M. J. Moore 306 (FLAS), HQ664581; Caryophyllales, Polygonaceae, *Fagopyrum esculentum* Moench ssp. *ancestrale* Ohnishi, . . ., NC_010776; Caryophyllales, Polygonaceae, *Polygonum virginianum* L., D. Soltis 2656 (FLAS), HQ664655; Caryophyllales, Portulacaceae, *Portulaca oleracea* L., M. J. Moore 322 (FLAS), HQ664656; Caryophyllales, Rhabdodendraceae, *Rhabdodendron amazonicum* (Spruce ex Benth.) Huber, E. Ribéiro 1187 (K), HQ664657; Caryophyllales, Sarcobataceae, *Sarcobatus vermiculatus* (Hook.) Torr. in Emory, King & Garvey 13892 (MO), HQ664659; Caryophyllales, Simmondsiaceae, *Simmondsia chinensis* C. K. Schneid., Y.-L. Qiu 96120 (IND), HQ664660; Caryophyllales, Stegnospermataceae, *Stegnosperma halimifolium* Benth., Martin et al. s.n. (MO), HQ664661; Caryophyllales, Talinaceae, *Talinum paniculatum* (Jacq.) Gaertn., D. Soltis 2646 (FLAS), HQ664663; Caryophyllales, Tamaricaceae, *Tamarix pentandra* Hampe. ex Bunge, Chase 252 (NCU), HQ664664; Celastrales, Celastraceae, *Euonymus americanus* L., W. S. Judd 8071 (FLAS), HQ664608; Celastrales, Lepidobotryaceae, *Ruptiliocarpon caracolito* Hammel & N. Zamora, M. W. Chase 2311 (K), HQ664571; Cornales, Cornaceae, *Cornus florida* L., M. J. Moore 328 (FLAS), HQ664596; Crossosomatales, Crossosomataceae, *Crossosoma californicum* Nutt., cultivated at Rancho Santa Ana Botanical Garden, HQ664556; Crossosomatales, Staphyleaceae, *Staphylea colchica* Steven, D. E. and P. S. Soltis 2711 (FLAS), HQ664600; Crossosomatales, Strasburgeriaceae, *Ixerba brexioides* A. Cunn., Small s.n. (WS), HQ664559; Cucurbitales, Anisophylleaceae, *Anisophyllea fallax* Scott-Elliot, L. Gautier 3256 (MO), HQ664550; Cucurbitales, Coriariaceae, *Coriaria nepalensis* Wall., M. W. Chase 6416 (K), HQ664555; Cucurbitales, Cucurbitaceae, *Cucumis sativus* L., . . ., NC_007144; Dilleniaceae, Dilleniaceae, *Acrotrema costatum* Jack,* K. J. Williams 00–201 (DUKE), HQ664618; Dilleniaceae, Dilleniaceae, *Curatella americana* L.,* T. W. Henkel s.n. (DUKE), HQ664621; Dilleniaceae, Dilleniaceae, *Davilla nitida* (Vahl) Kubitzki,* J. W. Horn 3290 (DUKE), HQ664622; Dilleniaceae, Dilleniaceae, *Dillenia indica* L., M. J. Moore 340 (FLAS), HQ664593; Dilleniaceae, Dilleniaceae, *Hibbertia banksii* (DC.) Benth.,* J. W. Horn 4322 (DUKE), HQ664623; Dilleniaceae, Dilleniaceae, *Pinzona coriacea* Mart. & Zucc.,* J. W. Horn 3371 (DUKE), HQ664625; Dilleniaceae, Dilleniaceae, *Tetracera scandens* L. (Merrill),* P. Manos s.n. (DUKE), HQ664665; Dipsacales, Caprifoliaceae, *Lonicera japonica* Thunb., M. J. Moore 312 (FLAS), GQ997460, HQ664582; Ericales, Ericaceae, *Rhododendron simsii* Planch., M. J. Moore 324 (FLAS), HQ664585, HQ664586; Fabales, Fabaceae, *Albizia julibrissin* Durazz., D. Soltis 2633 (FLAS), HQ664549; Fabales, Fabaceae, *Cicer arietinum* L., . . ., NC_011163; Fabales, Fabaceae, *Glycine max* Merr., . . ., NC_007942; Fabales, Fabaceae, *Lotus corniculatus* L., . . ., NC_002694; Fabales, Fabaceae, *Medicago truncatula* Gaertn., . . ., NC_003119; Fabales, Fabaceae, *Phaseolus vulgaris* L., . . ., EU196765; Fabales, Fabaceae, *Trifolium subterraneum* L., . . ., NC_011828; Fabales, Polygalaceae, *Polygala cruciata* L., M. W. Chase 155 (K), HQ664569; Fabales, Quillajaceae, *Quillaja saponaria* Molina, M. W. Chase 10931 (K), HQ664570; Fabales, Surianaceae, *Stylobasium spathulatum* Desf., Brummitt, George & Oliver 21242 (K), HQ664573; Fagales, Fagaceae, *Quercus nigra* L., M. J. Moore 311 (FLAS), HQ664601; Fagales, Juglandaceae, *Juglans nigra* L., D. Soltis 2520 (WS), HQ664560; Garryales, Garryaceae, *Aucuba japonica* Thunb., M. J. Moore 327 (FLAS), HQ664599; Gentianales, Apocynaceae, *Nerium oleander* L., W. S. Judd 8076 (FLAS), HQ664607; Gentianales, Rubiaceae, *Coffea arabica* L., . . ., NC_008535; Geraniales, Geraniaceae, *Hypseocharis bilobata* Killip, M. W. Chase 2785 (K), HQ664558; Geraniales, Geraniaceae, *Pelargonium* × *hortorum* L.H. Bailey, . . ., NC_008454; Geraniales, Melianthaceae, *Melianthus comosus* Vahl., R. G. Olmstead 2005–4 (WTU), HQ664564; Geraniales, Vivianiaceae, *Viviania marifolia* Cav., Penalillo 91000 (IND), HQ664576; Gunnerales, Gunneraceae, *Gunnera manicata* Linden ex André, M. J. Moore 325 (FLAS), HQ664604; Gunnerales, Myrothamnaceae,

Myrothamnus flabellifolia Welw.,* Winter 72 (RAV), HQ664624; Huerteales, Tapisciaceae, *Tapiscia sinensis* Oliv., M. W. Chase 1201 (K), HQ664574; Lamiales, Oleaceae, *Jasminum nudiflorum* Lindl., . . ., NC_008407; Lamiales, Orobanchaceae, *Epifagus virginiana* (L.) W.P.C.Barton, . . ., NC_001568; Lamiales, Plantaginaceae, *Antirrhinum majus* L., M. J. Moore 314 (FLAS), HQ664592; Malpighiales, Achariaceae, *Erythrospermum phytolaccoides* Gardn.,* M. W. Chase 1277 (K), HQ664612; Malpighiales, Caryocaraceae, *Caryocar glabrum* Pers.,* Mori 22997 (NY), HQ664613; Malpighiales, Euphorbiaceae, *Jatropha curcas* L., . . ., NC_012224; Malpighiales, Euphorbiaceae, *Manihot esculenta* Crantz, . . ., NC_010433; Malpighiales, Ochnaceae, *Ochna mossambicensis* Klotzsch, D. Soltis 2715 (FLAS), HQ664566; Malpighiales, Passifloraceae, *Malesherbia linearifolia* Poir.,* M. W. Chase 609 (K), HQ664611; Malpighiales, Passifloraceae, *Passiflora biflora* Lam., . . ., EU017068, EU017071, EU017074, EU017078, EU017081, EU017083, EU017087, EU017103, EU017109, EU017119; Malpighiales, Phyllanthaceae, *Phyllanthus calycinus* Labill., M. W. Chase 2163 (K), HQ664567; Malpighiales, Salicaceae, *Populus alba* L., . . ., NC_008235; Malpighiales, Salicaceae, *Populus trichocarpa* Torr. & A. Gray, . . ., NC_009143; Malvales, Malvaceae, *Gossypium barbadense* L., . . ., NC_008641; Malvales, Malvaceae, *Gossypium hirsutum* L., . . ., NC_007944; Malvales, Neuradaceae, *Neurada procumbens* L., Collenette 8193 (K), HQ664565; Malvales, Thymelaeaceae, *Gonystylus bancanus* Gilg, . . ., EU849490; Myrtales, Combretaceae, *Terminalia cattapa* L., E. Conti 103 (WIS), HQ664575; Myrtales, Lythraceae, *Lythrum flagellare* Shuttlew. ex Chapm., D. Soltis 2704 (FLAS), HQ664563; Myrtales, Melastomataceae, *Clidemia dentata* D. Don, D. Soltis 2708 (FLAS), HQ664554; Myrtales, Myrtaceae, *Eucalyptus globulus* Labill., . . ., NC_008115; Myrtales, Myrtaceae, *Myrtus communis* L., W. Judd 8066 (FLAS), HQ664632; Myrtales, Onagraceae, *Oenothera argillicola* Mack., . . ., NC_010358; Myrtales, Onagraceae, *Oenothera biennis* L., . . ., NC_010361; Myrtales, Onagraceae, *Oenothera elata* Kunth ssp. *hookeri* (Torr. & A.Gray) W.Dietr. & W.L.Wagner, . . ., NC_002693; Myrtales, Onagraceae, *Oenothera glazioviana* Micheli, . . ., NC_010360; Myrtales, Onagraceae, *Oenothera parviflora* L., . . ., NC_010362; Oxalidales, Cunoniaceae, *Eucryphia lucida* (Labill.) Baill., Strybing Arboretum accession 74.0135, HQ664557; Oxalidales, Oxalidaceae, *Oxalis latifolia* Kunth, M. J. Moore 316 (FLAS), HQ664602; Picramniales, Picramniaceae, *Picramnia pentandra* Sw., M. W. Chase 2571 (K), HQ664568; Proteales, Nelumbonaceae, *Nelumbo nucifera* Gaertn., S. B. Davis 1076 (FLAS), HQ664584; Proteales, Platanaceae, *Platanus occidentalis* L., . . ., NC_008335; Ranunculales, Berberidaceae, *Nandina domestica* Thunb., . . ., NC_008336; Ranunculales, Ranunculaceae, *Megaleranthis saniculifolia* Ohwi, . . ., NC_012615; Ranunculales, Ranunculaceae, *Ranunculus macranthus* Scheele, . . ., NC_008796; Rosales, Cannabaceae, *Celtis occidentalis* L., D. Soltis 2701 (FLAS), HQ664553; Rosales, Moraceae, *Ficus* sp., M. J. Moore 315 (FLAS), HQ664605; Rosales, Moraceae, *Morus indica* L., . . ., NC_008359; Rosales, Rhamnaceae, *Ceanothus prostratus* Benth., D. Soltis 2712 (FLAS), HQ664552; Rosales, Rosaceae, *Fragaria* × *ananassa* (Weston) Duchesne ex Rozier, . . ., DQ768221; Rosales, Rosaceae, *Prunus persica* (L.) Batsch, . . ., DQ768222; Rosales, Rosaceae, *Spiraea tomentosa* L., D. Soltis 2691 (FLAS), HQ664572; Sabiaceae, Sabiaceae, *Meliosma* aff. *cuneifolia* Franch., M. J. Moore 333 (FLAS), HQ664583; Santalales, Erythropalaceae, *Heisteria concinna* Standl.,* M. J. Moore 336 (FLAS), HQ664616; Santalales, Olacaceae, *Ximenia americana* L., W. S. Judd 8070 (FLAS), HQ664594; Santalales, Santalaceae, *Buckleya henryi* Diels, . . ., FJ895899–FJ895903; Santalales, Santalaceae, *Phoradendron leucarpum* (Raf.) Reveal & M.C.Johnst., M. J. Moore 332 (FLAS), GQ997778, HQ664580; Sapindales, Anacardiaceae, *Mangifera indica* L., . . ., EF205595; Sapindales, Anacardiaceae, *Schinus terebinthifolius* Raddi, D. Soltis 2705 (FLAS), HQ664548; Sapindales, Rutaceae, *Citrus sinensis* Osbeck, . . ., NC_008334; Saxifragales, Altingiaceae, *Liquidambar styraciflua* L. [Moore], M. J. Moore 307 (FLAS), HQ664606; Saxifragales, Altingiaceae, *Liquidambar styraciflua* L. [Wang], . . ., EF207449; Saxifragales, Cercidiphyllaceae, *Cercidiphyllum japonicum* Siebold & Zucc., . . ., EF207443; Saxifragales, Crassulaceae, *Kalanchoë daigremontiana* Raym.-Hamet & H. Perrier, . . ., EF207448; Saxifragales, Daphniphyllaceae, *Daphniphyllum* sp., . . ., EF207444; Saxifragales, Grossulariaceae, *Ribes americanum* Mill., . . ., EF207456; Saxifragales, Haloragaceae, *Myriophyllum spicatum* L., . . ., EF207450; Saxifragales, Hamamelidaceae, *Hamamelis japonica* Siebold & Zucc., . . ., EF207445; Saxifragales, Hamamelidaceae, *Rhodoleia championii* Hook.f., . . ., EF207455; Saxifragales, Iteaceae, *Itea virginica* L., . . ., EF207447; Saxifragales, Paeoniaceae, *Paeonia brownii* Douglas ex Hook., . . ., EF207451; Saxifragales, Penthoraceae (Haloragaceae s.l.), *Penthorum chinense* Pursh, . . ., EF207452; Saxifragales, Peridiscaceae, *Peridiscus lucidus* Benth., . . ., EF207453; Saxifragales, Pterostemonaceae, *Pterostemon rotundifolius* Ramirez, . . ., EF207454; Saxifragales, Saxifragaceae, *Heuchera micrantha* Douglas ex Lindl., . . ., EF207446; Saxifragales, Saxifragaceae, *Heuchera sanguinea* Engelm., M. J. Moore 329 (FLAS), HQ664603; Saxifragales, Saxifragaceae, *Saxifraga stolonifera* Curt., . . ., EF207457; Solanales, Convolvulaceae, *Cuscuta exaltata* Engelm., . . ., NC_009963; Solanales, Convolvulaceae, *Cuscuta gronovii* Willd. Ex Roem. & Schult., . . ., NC_009765; Solanales, Convolvulaceae, *Cuscuta obtusiflora* Kunth, . . ., NC_009949; Solanales, Convolvulaceae, *Cuscuta reflexa* Roxb., . . ., NC_009766; Solanales, Convolvulaceae, *Ipomoea purpurea* (L.) Roth, . . ., NC_009808; Solanales, Solanaceae, *Atropa belladonna* L., . . ., NC_004561; Solanales, Solanaceae, *Nicotiana sylvestris* Speg., . . ., NC_007500; Solanales, Solanaceae, *Nicotiana tabacum* L., . . ., NC_001879; Solanales, Solanaceae, *Nicotiana tomentosiformis* Goodsp., . . ., NC_007602; Solanales, Solanaceae, *Solanum bulbocastanum* Dunal, . . ., NC_007943; Solanales, Solanaceae, *Solanum lycopersicum* L., . . ., NC_007898; Solanales, Solanaceae, *Solanum tuberosum* L., . . ., NC_008096; Trochodendrales, Trochodendraceae, *Trochodendron aralioides* Siebold & Zucc., cultivated at University of Washington Arboretum: accession 581–4C, HQ664595; Vitales, Leeaceae, *Leea guineensis* G. Don, M. W. Chase 17676 (K), HQ664562; Vitales, Vitaceae, *Cissus rotundifolia* (Forssk.) Vahl,* D. Soltis s.n. (FLAS), HQ664620; Vitales, Vitaceae, *Tetrastigma yunnanense* Gagnep.,* J. Wen 8479 (F), HQ664626; Vitales, Vitaceae, *Vitis vinifera* L., . . ., NC_007957; Zygophyllales, Krameriaceae, *Krameria lanceolata* Torr., M. W. Chase 103 (K), HQ664561; Zygophyllales, Zygophyllaceae, *Bulnesia arborea* (Jacq.) Engl., M. J. Moore 334 (FLAS), HQ664597; unplaced, Cynomoriaceae, *Cynomorium songaricum* Rupr., . . ., FJ895885–FJ895893.

Magnoliidae

Canellales, Canellaceae, *Canella winterana* (L.) Gaertn., Y.-L. Qiu 90017 (IND), HQ664630; Canellales, Winteraceae, *Drimys granadensis* L.f., . . ., NC_008456; Canellales, Winteraceae, *Tasmannia lanceolata* (Poir.) A. C. Sm., S. Kim s.n., HQ664627; Laurales, Calycanthaceae, *Calycanthus floridus* L., . . ., NC_004993; Laurales, Lauraceae, *Cinnamomum camphora* (L.) J. Presl, D. Soltis 2718 (FLAS), HQ664631; Magnoliales, Magnoliaceae, *Liriodendron tulipifera* L., . . ., NC_008326; Piperales, Aristolochiaceae, *Aristolochia littoralis* D. Parodi 'elegans',* M. J. Moore 330 (FLAS), HQ664614; Piperales, Aristolochiaceae, *Saruma henryi* Oliv., D. Soltis 2663b, HQ664634; Piperales, Piperaceae, *Piper cenocladum* C.DC., . . ., NC_008457; Piperales, Saururaceae, *Saururus cernuus* L., M. Buzgo 1155 (FLAS), HQ664635.

Monocotyledoneae

Acorales, Acoraceae, *Acorus americanus* Raf., . . ., NC_010093; Acorales, Acoraceae, *Acorus calamus* L., . . ., NC_007407; Alismatales, Araceae, *Lemna minor* L., . . ., NC_010109; Alismatales, Araceae, *Spathiphyllum* sp.,* M. J. Moore 338 (FLAS), HQ664609; Alismatales, Hydrocharitaceae, *Vallisneria americana* Michx.,* M. J. Moore 339 (FLAS), HQ664587–HQ664591; Arecales, Arecaceae, *Elaeis oleifera* (Kunth) Cortés ex Prain, . . ., EU016888, EU016891, EU016894, EU016898, EU016901, EU016904, EU016908, EU016925, EU016931, EU016941; Asparagales, Agavaceae, *Yucca schidigera* Ortgies, . . ., DQ069414, DQ069450, DQ069498, EU016682, EU016684, EU016685, EU016687, EU016691, EU016694, EU016698; Asparagales, Orchidaceae, *Phalaenopsis aphrodite* Rchb.f., . . ., NC_007499; Commelinales, Commelinaceae, *Tradescantia ohiensis* Raf.,* M. J. Moore 337 (FLAS), HQ664610; Dioscoreales, Dioscoreaceae, *Dioscorea elephantipes* Engl., . . ., NC_009601; Liliales, Liliaceae, *Lilium superbum* L.,* cultivated in University of Wisconsin greenhouse, HQ664578; Poales, Poaceae, *Agrostis stolonifera* L., . . ., NC_008591; Poales, Poaceae, *Bambusa oldhamii* Munro, . . ., NC_012927; Poales, Poaceae, *Brachypodium distachyon* (L.) P. Beauv., . . ., NC_011032; Poales, Poaceae, *Dendrocalamus latiflorus* Munro, . . ., NC_013088; Poales, Poaceae, *Festuca arundinacea* Schreb., . . ., NC_011713; Poales, Poaceae, *Hordeum vulgare* L. ssp. *vulgare*, . . ., NC_008590; Poales, Poaceae, *Lolium perenne* L., . . ., NC_009950; Poales, Poaceae, *Oryza nivara* S. D. Sharma & Shastry, . . ., NC_005973; Poales, Poaceae, *Oryza sativa* L., . . ., NC_001320; Poales, Poaceae, *Saccharum officinarum* L., . . ., NC_006084; Poales, Poaceae, *Sorghum bicolor* (L.) Moench, . . ., NC_008602; Poales, Poaceae, *Triticum aestivum* L., . . ., NC_002762; Poales, Poaceae, *Zea mays* L., . . ., NC_001666; Poales, Typhaceae, *Typha latifolia* L., . . ., NC_013823; Zingiberales, Musaceae, *Musa acuminata* Colla, . . ., EU016988, EU016991, EU016994, EU016998, EU017001, EU017004, EU017008, EU017026, EU017032, EU017042.

Literature Cited

APG (Angiosperm Phylogeny Group) III 2009 An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc 161:105–121.

Barkman TJ, G Chenery, JR McNeal, J Lyons-Weiler, WJ Ellisens, G Moore, AD Wolfe, CW dePamphilis 2000 Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. Proc Natl Acad Sci USA 97:13166–13171.

Barkman TJ, JR McNeal, SH Lim, G Coat, HB Croom, ND Young, CW DePamphilis 2007 Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. BMC Evol Biol 7:248.

Borsch T, D Quandt 2009 Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. Plant Syst Evol 282:169–199.

Braukmann TW, M Kuzmina, S Stefanović 2009 Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. Curr Genet 55:323–337.

Bremer B, K Bremer, N Heidari, P Erixon, RG Olmstead, AA Anderberg, M Källersjö, E Barkhordarian 2002 Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Mol Phylogenet Evol 24:274–301.

Brockington SF, R Alexandre, J Ramdial, MJ Moore, S Crawley, A Dhingra, K Hilu, DE Soltis, PS Soltis 2009 Phylogeny of the Caryophyllales sensu lato: revisiting hypotheses on pollination biology and perianth differentiation in the core Caryophyllales. Int J Plant Sci 170:627–643.

Bull JJ, JP Huelsenbeck, CW Cunningham, DL Swofford, PJ Waddell 1993 Partitioning and combining data in phylogenetic analysis. Syst Biol 42:384–397.

Burleigh JG, KW Hilu, DE Soltis 2009 Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. BMC Evol Biol 9:61.

Cai Z, M Guisinger, HG Kim, E Ruck, JC Blazier, V McMurtry, JV Kuehl, J Boore, RK Jansen 2008 Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. J Mol Evol 67:696–704.

Cai Z, C Penaflor, JV Kuehl, J Leebens-Mack, JE Carlson, CW dePamphilis, JL Boore, RK Jansen 2006 Complete plastid genome sequences of *Drimys*, *Liriodendron*, and *Piper*: implications for the phylogenetic relationships of magnoliids. BMC Evol Biol 6:77.

Cantino PD, JA Doyle, SW Graham, WS Judd, RG Olmstead, DE Soltis, PS Soltis, MJ Donoghue 2007 Towards a phylogenetic nomenclature of *Tracheophyta*. Taxon 56:822–846.

Chase MW, MF Fay, DS Devey, O Maurin, N Ronsted, TJ Davies, Y Pillon, et al 2006 Multigene analyses of monocot relationships: a summary. Aliso 22:63–75.

Chumley TW, JD Palmer, JP Mower, HM Fourcade, PJ Calie, JL Boore, RK Jansen 2006 The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol 23:2175–2190.

Cronn R, A Liston, M Parks, DS Gernandt, R Shen, T Mockler 2008 Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. Nucleic Acids Res 36: e122, doi:10.1093/nar/gkn502.

dePamphilis CW, JD Palmer 1990 Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348:337–339.

Dhingra A, KM Folta 2005 ASAP: amplification, sequencing and annotation of plastomes. BMC Genomics 6:176.

Downie SR, JD Palmer 1992 Restriction site mapping of the chloroplast DNA inverted repeat: a molecular phylogeny of the Asteridae. Ann Mo Bot Gard 79:266–283.

Doyle JJ, JL Doyle 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15.

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

——— 1985 Confidence limits on phylogeny: an approach using the bootstrap. Evolution 39:783–791.

Givnish TJ, JC Pires, SW Graham, MA McPherson, LM Prince, TB Patterson, HS Rai, et al 2006 Phylogenetic relationships of monocots based on the highly informative plastid gene *ndh*F: evidence for widespread concerted convergence. Aliso 22:28–51.

Graham SW, WJD Iles 2009 Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny. Am J Bot 96:216–227.

Graham SW, RG Olmstead 2000 Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am J Bot 87: 1712–1730.

Graham SW, PA Reeves, ACE Burns, RG Olmstead 2000 Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int J Plant Sci 161(suppl):S83–S96.

Graham SW, JM Zgurski, MA McPherson, DM Cherniawsky, JM Saarela, ESC Horne, SY Smith, et al 2006 Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. Aliso 22:3–20.

Guisinger MM, JV Kuehl, JL Boore, RK Jansen 2008 Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. Proc Natl Acad Sci USA 105: 18424–18429.

Haberle RC, HM Fourcade, JL Boore, RK Jansen 2008 Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J Mol Evol 66:350–361.

Heath TA, SM Hedtke, DM Hillis 2008 Taxon sampling and the accuracy of phylogenetic analyses. J Syst Evol 46:239–257.

Hilu KW, T Borsch, K Müller, DE Soltis, PS Soltis, V Savolainen, MW Chase, et al 2003 Inference of angiosperm phylogeny based on *matK* sequence information. Am J Bot 90:1758–1776.

Jansen RK, Z Cai, LA Raubeson, H Daniell, CW DePamphilis, J Leebens-Mack, KF Müller, et al 2007 Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104: 19369–19374.

Jansen RK, MF Wojciechowski, E Sanniyasi, S-B Lee, H Daniell 2008 Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among legumes (Leguminosae). Mol Phylogenet Evol 48:1204–1217.

Jeanmougin F, JD Thompson, M Gouy, DG Higgins, TJ Gibson 1998 Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405.

Jian S, PS Soltis, MA Gitzendanner, MJ Moore, R Li, TA Hendry, YL Qiu, A Dhingra, CD Bell, DE Soltis 2008 Resolving an ancient, rapid radiation in Saxifragales. Syst Biol 57:38–57.

Judd WS, RG Olmstead 2004 A survey of tricolpate (eudicot) phylogenetic relationships. Am J Bot 91:1627–1644.

Judd WS, DE Soltis, PS Soltis, PD Cantino Forthcoming Malvidae. *In* K de Queiroz, J Gauthier, PD Cantino, eds. Phylonyms: a companion to the PhyloCode. University of California Press, Berkeley.

Kelchner SA 2000 The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann Mo Bot Gard 87:482–498.

Klopfstein S, C Kropf, DL Quicke 2010 An evaluation of phylogenetic informativeness profiles and the molecular phylogeny of Diplazontinae (Hymenoptera, Ichneumonidae). Syst Biol 59:226–241.

Leebens-Mack J, LA Raubeson, L Cui, JV Kuehl, MH Fourcade, TW Chumley, JL Boore, RK Jansen, CW dePamphilis 2005 Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone. Mol Biol Evol 22: 1948–1963.

Li X-X, Z-K Zhou 2007 The higher-level phylogeny of monocots based on *mat*K, *rbc*L, and 18S rDNA sequences. Acta Phytotaxon Sin 45:113–133.

Maddison WP, DR Maddison 2003 MacClade, version 4.0. Sinauer, Sunderland, MA.

Mathews S 2009 Phylogenetic relationships among seed plants: persistent questions and the limits of molecular data. Am J Bot 96:228–236.

McNeal JR, JV Kuehl, JL Boore, CW dePamphilis 2007 Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. BMC Plant Biol 7:57.

Moore MJ, CD Bell, PS Soltis, DE Soltis 2007 Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci USA 104:19363–19368.

Moore MJ, A Dhingra, PS Soltis, R Shaw, WG Farmerie, KM Folta, DE Soltis 2006 Rapid and accurate pyrosequencing of angiosperm plastid genomes. BMC Plant Biol 6:17.

Moore MJ, PS Soltis, CD Bell, JG Burleigh, DE Soltis 2010 Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc Natl Acad Sci USA 107:4623–4628.

Nickrent DL, JP Der, FE Anderson 2005 Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. BMC Evol Biol 5:38.

Palmer JD 1991 Plastid chromosomes: structure and evolution. Pages 5–53 *in* RG Hermann, ed. Cell culture and somatic cell genetics of plants. Vol 7A. The molecular biology of plastids. Academic Press, Vienna.

Philippe H, Y Zhou, H Brinkmann, N Rodrigue, F Delsuc 2005 Heterotachy and long-branch attraction in phylogenetics. BMC Evol Biol 5:50.

Plunkett GM, SR Downie 2000 Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. Syst Bot 25:648–667.

Qiu Y-L, L Li, TA Hendry, R Li, DW Taylor, MJ Issa, AJ Ronen, ML Vekaria, AM White 2006 Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. Taxon 55:837–856.

Raubeson LA, RK Jansen 1992 A rare chloroplast-DNA structural mutation is shared by all conifers. Biochem Syst Ecol 20:17–24.

——— 2005 Chloroplast genomes of plants. Pages 45–68 *in* RJ Henry, ed. Plant diversity and evolution: genotypic and phenotypic variation in higher plants. CABI, Cambridge, MA.

Saarela JM, HS Rai, JA Doyle, PK Endress, S Mathews, AD Marchant, BG Briggs, SW Graham 2007 Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. Nature 446:312–315.

Schäferhoff B, KF Müller, T Borsch 2009 *Caryophyllales* phylogenetics: disentangling *Phytolaccaceae* and *Molluginaceae* and description of *Microteaceae* as a new isolated family. Willdenowia 39: 209–228.

Soltis DE, MA Gitzendanner, PS Soltis 2007 A 567-taxon data set for angiosperms: the challenges posed by Bayesian analyses of large data sets. Int J Plant Sci 168:137–157.

Soltis DE, PS Soltis 2004 *Amborella* not a "basal angiosperm"? Not so fast. Am J Bot 91:997–1001.

Soltis DE, PS Soltis, MW Chase, ME Mort, DC Albach, M Zanis, V Savolainen, et al 2000 Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Bot J Linn Soc 133:381–461.

Soltis DE, PS Soltis, TG Collier, ML Edgerton 1991 Chloroplast DNA variation within and among genera of the *Heuchera* group (Saxifragaceae): evidence for chloroplast transfer and paraphyly. Am J Bot 78:1091–1112.

Soltis DE, PS Soltis, PK Endress, MW Chase 2005 Phylogeny and evolution of angiosperms. Sinauer, Sunderland, MA.

Soltis PS, SF Brockington, M-J Yoo, A Piedrahita, M Latvis, MJ Moore, AS Chanderbali, DE Soltis 2009 Floral variation and floral genetics in basal angiosperms. Am J Bot 96:110–128.

Stamatakis A 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Swofford DL 2000 PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, MA.

Wang H, MJ Moore, PS Soltis, CD Bell, SF Brockington, R Alexandre, CC Davis, M Latvis, SR Manchester, DE Soltis 2009 Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc Natl Acad Sci USA 106:3853–3858.

Wang RJ, CL Cheng, CC Chang, CL Wu, TM Su, SM Chaw 2008 Dynamics and evolution of the inverted repeat–large single copy junctions in the chloroplast genomes of monocots. BMC Evol Biol 8:36.

Wojciechowski MF, M Lavin, MJ Sanderson 2004 A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. Am J Bot 91:1846–1862.

Wolfe KH, WH Li, PM Sharp 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058.

Worberg A, MH Alford, D Quandt, T Borsch 2009 Huerteales sister to Brassicales plus Malvales, and newly circumscribed to include *Dipentodon*, *Gerrardina*, *Huertea*, *Perrottetia*, and *Tapiscia*. Taxon 58:468–478.

Worberg A, D Quandt, A-M Barniske, C Löhne, KW Hilu, T Borsch 2007 Phylogeny of basal eudicots: insights from non-coding and rapidly evolving DNA. Org Divers Evol 7:55–77.

Wortley AH, PJ Rudall, DJ Harris, RW Scotland 2005 How much data are needed to resolve a difficult phylogeny? case study in Lamiales. Syst Biol 54:697–709.

Wurdack KJ, CC Davis 2009 Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. Am J Bot 96:1551–1570.

Wyman SK, RK Jansen, JL Boore 2004 Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3255.

Yang Z 1998 On the best evolutionary rate for phylogenetic analysis. Syst Biol 47:125–133.

Zanis MJ, DE Soltis, PS Soltis, S Mathews, MJ Donoghue 2002 The root of the angiosperms revisited. Proc Natl Acad Sci USA 99:6848–6853.

Zhang Z-H, C-Q Li, J Li 2009 Phylogenetic placement of *Cynomorium* in Rosales inferrred from sequences of the inverted repeat region of the chloroplast genome. J Syst Evol 47:297–304.