January 2011

Angiosperm Phylogeny: 17 Genes, 640 Taxa

Douglas E. Soltis  
*University of Florida*

Steven A. Smith  
*Brown University*

Nico Cellinese  
*Florida Museum of Natural History*

Kenneth J. Wurdack  
*National Museum of Natural History, Smithsonian Institution*

David C. Tank  
*University of Idaho*

See next page for additional authors

Follow this and additional works at: [http://thekeep.eiu.edu/bio_fac](http://thekeep.eiu.edu/bio_fac)

Part of the Botany Commons

Recommended Citation

Soltis, Douglas E.; Smith, Steven A.; Cellinese, Nico; Wurdack, Kenneth J.; Tank, David C.; Brockington, Samuel F.; Refulio-Rodriguez, Nancy F.; Walker, Jay B.; Moore, Michael J.; Carlsward, Barbara S.; Bell, Charles D.; Latvis, Maribeth; Crawley, Sunny; Black, Chelsea; Diouf, Diaga; Xi, Zhenxiang; Rushworth, Catherine A.; Gitzendanner, Matthew A.; Sytsma, Kenneth J.; Qiu, Yin-Long; Hilu, Khidir W.; Davis, Charles C.; Sanderson, Michael J.; Beaman, Reed S.; Olmstead, Richard G.; Judd, Walter S.; Donoghue, Michael S.; and Soltis, Pamela S., "Angiosperm Phylogeny: 17 Genes, 640 Taxa" (2011). *Faculty Research & Creative Activity*. 17.  
[http://thekeep.eiu.edu/bio_fac/17](http://thekeep.eiu.edu/bio_fac/17)

This Article is brought to you for free and open access by the Biological Sciences at The Keep. It has been accepted for inclusion in Faculty Research & Creative Activity by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.
**Angiosperm phylogeny: 17 genes, 640 taxa**

Douglas E. Soltis²,⁹, Stephen A. Smith³, Nico Cellinese⁴, Kenneth J. Wurdack⁵, David C. Tank⁶, Samuel F. Brockington⁷, Nancy F. Refulio-Rodriguez⁸, Jay B. Walker⁸, Michael J. Moore¹⁰, Barbara S. Carlsward¹¹, Charles D. Bell¹², Maribeth Latvis²,⁴, Sunny Crawley¹³, Chelsea Black¹³, Diaga Diof¹³,¹⁴, Zhenxiang Xi¹⁵, Catherine A. Rushworth¹⁵, Matthew A. Gitzendanner²,⁴, Kenneth J. Sytsma⁹, Yin-Long Qiu¹⁶, Khidir W. Hilu¹³, Charles C. Davis¹⁵, Michael J. Sanderson¹⁷, Reed S. Beaman⁴, Richard G. Olmstead⁸, Walter S. Judd², Michael J. Donoghue¹⁸, and Pamela S. Soltis⁴

²Department of Biology, University of Florida, Gainesville, Florida 32611-8525 USA; ³Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island 02912 USA; ⁴Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611-7800 USA; ⁵Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, D.C. 20013-7012 USA; ⁶Department of Forest Ecology and Biogeosciences & Stillinger Herbarium, University of Idaho, P. O. Box 441133, Moscow, Idaho 83844-1133 USA; ⁷Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK; ⁸Department of Biology, University of Washington, Seattle, Washington 98195-5325 USA; ⁹Department of Botany, University of Wisconsin, Madison, Wisconsin 53706 USA; ¹⁰Biography Department, Oberlin College, Oberlin, Ohio 44074-1097 USA; ¹¹Department of Biological Sciences, Eastern Illinois University, Charleston, Illinois 61920 USA; ¹²Department of Biological Sciences, University of New Orleans, New Orleans, Louisiana 70148 USA; ¹³Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061 USA; ¹⁴Département de Biologie Végétale, Université Cheikh Anta Diop, Dakar–Fann, BP 5005, Republic of Sénégal; ¹⁵Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138 USA; ¹⁶Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109-1048 USA; ¹⁷Department of Ecology and Evolutionary Biology, University of Arizona, 1041 East Lowell, Tucson, Arizona 85721-0088 USA; and ¹⁸Department of Ecology and Evolutionary Biology, Yale University, P. O. Box 208160, New Haven, Connecticut 06520-8106 USA.

**Premise of the study:** Recent analyses employing up to five genes have provided numerous insights into angiosperm phylogeny, but many relationships have remained unresolved or poorly supported. In the hope of improving our understanding of angiosperm phylogeny, we expanded sampling of taxa and genes beyond previous analyses.

**Methods:** We conducted two primary analyses based on 640 species representing 330 families. The first included 25,260 aligned base pairs (bp) from 17 genes (representing all three plant genomes, i.e., nucleus, plastid, and mitochondrion). The second included 19,846 aligned bp from 13 genes (representing only the nucleus and plastid).

**Key results:** Many important questions of deep-level relationships in the nonmonocot angiosperms have now been resolved with strong support. Amborellaceae, Nymphaeaceae, and Austrobaileyaceae are successive sisters to the remaining angiosperms (Mesangiospermae), which are resolved into Chloranthales + Magnoliidae as sister to Monocotyledoneae + [Ceratophyllaceae + Eu dicotyledoneae]. Eu dicotyledoneae contains a basal grade subtending Gunneridae. Within Gunneridae, Gunnerales are sister to the remainder (Pentapetaleae), which comprises (1) Superrosidae, consisting of Rosidae (including Vitaceae) and Saxifragales; and (2) Superasteridae, comprising Berberidopsidales, Santalales, Caryophyllales, Asteridae, and, based on this study, Dilleniaceae (although other recent analyses disagree with this placement). Within the major subclades of Pentapetaleae, most deep-level relationships are resolved with strong support.

**Conclusions:** Our analyses confirm that with large amounts of sequence data, most deep-level relationships within the angiosperms can be resolved. We anticipate that this well-resolved angiosperm tree will be of broad utility for many areas of biology, including physiology, ecology, paleobiology, and genomics.

**Key words:** angiosperms; bioinformatics; large data sets; molecular systematics; RAxML; Superasteridae; supermatrix; Superrosidae.

---

¹ Manuscript received 11 October 2010; revision accepted 9 February 2011. The authors thank the University of Florida High-Performance Computing Center, the UF Genetics Institute, and facilities supported in part by NSF Grant CNS-0821622 for providing computational resources that facilitated the completion of this research. The authors thank E. Mavrodiev for help in preparing some of the figures. This study was carried out as part of the Angiosperm Tree of Life Project (NSF EF-0431266 to D.E.S., P.S.S., W.S.J., and S. R. Manchester, EF-0431242 to C.C.D., DEB-9634239; EF-0431057 to K.H., EF-0431233 to K.J.S., DEB-0431239 to Y.L.Q.) with additional support from NSF grant DEB-0544039 to C.C.D. The authors thank the Royal Botanic Gardens, Kew, and the Missouri Botanical Garden for providing samples used in this study.

¹⁹ Author for correspondence (e-mail: dsoltis@botany.ufl.edu) doi:10.3732/ajb.1000404
It has been a decade since the 567-taxon, three-gene (\textit{rbcL}, \textit{atpB}, and 18S rDNA) parsimony-based phylogenetic analysis of angiosperms was published (Soltis et al., 1999, 2000). Since that time, other studies have expanded sampling of additional genes or used alternative methods to evaluate the results of that three-gene study. Hilu et al. (2003) conducted a broad analysis of angiosperms based on \textit{matK} sequences, with results that agreed closely with the three-gene topology. Additional progress was achieved by Davies et al. (2004) by constructing a supertree for angiosperms. Soltis et al. (2007) undertook a Bayesian analysis of the 567-taxon, three-gene data set and obtained a topology nearly identical to that obtained with parsimony. More recently, Bell et al. (2010) analyzed this same three-gene data set using a Bayesian relaxed clock model to simultaneously infer topology and divergence times within angiosperms and found results similar to Soltis et al. (2007).

In addition, Burleigh et al. (2009) inferred angiosperm phylogeny using five genes for the same 567 taxa analyzed in Soltis et al. (1999, 2000). Although the five-gene matrix had significantly more missing data (27.5\%) than the three-gene matrix (2.9\%), the five-gene analysis resulted in higher levels of bootstrap support across the tree.

Recent phylogenomic analyses have shown the value of constructing data sets of many genes to infer deep-level angiosperm phylogeny. Some of these analyses have employed nearly complete plastid genome sequence data (e.g., Leebens-Mack et al., 2005; Jansen et al., 2007; Moore et al., 2007, 2010), but all have been limited in sampling to fewer than 100 taxa. These and other studies based on many genes but focused only on major angiosperm clades (e.g., Schönenberger et al., 2005; Jian et al., 2008; H. Wang et al., 2009; Wurdack and Davis, 2009; Brockington et al., 2010; Tank and Donoghue, 2010) showed that with very large amounts of data (i.e., 13 to 83 genes), many, if not most, deep-level questions of angiosperm phylogeny can be resolved.

While the three- and five-gene analyses of 567 taxa have broad taxonomic coverage, support for many portions of the framework of angiosperm phylogeny is low in these studies. For example, relationships among members of \textit{Mesangiospermae} and of \textit{Pentapetalae} remain unclear. Conversely, studies employing complete plastid genome sequences have deep gene coverage and strong internal support, but taxonomic coverage is often sparse. Hence, it is important to assemble a data set having both broad taxonomic coverage as well as numerous genes. In the hope of further improving our understanding of angiosperm phylogeny, we have sequenced deeply over a broad representation of angiosperms. We constructed a 17-gene data set for 640 species representing 640 genera, 330 families, and 58 of 59 orders (sensu APG III, 2009) using genes that represent all three plant genomic compartments (nucleus, plastid, and mitochondrion).

One goal of the angiosperm Assembling the Tree of Life project, that allows users to access and share data in real time, as well as among angiosperm phylogenists, is to assemble a data set having both broad taxonomic coverage as well as numerous genes (sensu APG III, 2009). We here also provide putative morphological synapomorphies of Saxifragales, although a more detailed morphological exploration of this clade is presented elsewhere (Carlsward et al., unpublished manuscript). This study of Saxifragales represents the first treatment in a series of analyses for large clades of angiosperms that will address morphological synapomorphies and the clades they support (W. S. Judd et al., unpublished data).

### MATERIALS AND METHODS

#### Data sets

Data samples for most of the species used here were extracted from either fresh or silica-dried material following the general method of Doyle and Doyle (1987) or modifications thereof that employ liquid nitrogen and higher CTAB concentrations (e.g., Soltis et al., 1991; Sytsma, 1994). We attempted to use the same species and DNA samples across all of the genes analyzed here, although multiple species were sometimes used as necessary placeholders to reduce missing data. Many of these DNA samples have been used in earlier analyses (e.g., Chase et al., 1993; Soltis et al., 2000).

We constructed a 17-gene data set for 640 species (for a complete list of taxa, voucher information, and GenBank numbers see Appendix 51). We supplemented Data online at http://www.amborg.org/cgi/content/full/ajl.1000434DC1 using genes from the nuclear, plastid, and mitochondrial genomes. Given the potential problems inherent in phylogeny reconstruction using mtDNA sequences (e.g., RNA editing, Bowe and dePamphilis, 1996; horizontal gene transfer or HGT, Berghorston et al., 2003), we also constructed a data set without the mtDNA data, resulting in a matrix of 13 genes. For both data sets, we used the following representatives of \textit{Acrogyynospermae} (extant angiosperms, sensu Cantino et al., 2007) as the outgroup: \textit{Cycas, Ginkgo, Gnetum, Metasequoia, Pinus, Podocarpus, Welwitschia}, and \textit{Zamia}.

Taxon sampling across major clades of angiosperms was not uniform, with poorly resolved clades (e.g., Malpighiales, Saxifragales) targeted for denser taxon sampling to seek improved resolution. We also largely avoided parasitic clades (except Orobancheae, Santalales, and \textit{Cuscuta}), which can create analytical problems due to gene loss, accelerated molecular evolution, and horizontal gene transfer (see Davis and Wurdack, 2004; Nickrent et al., 2004; Barkman et al., 2007).

The following 17 genes were sequenced: 18S and 26S rDNA from the nuclear genome; \textit{atpB}, \textit{matK}, \textit{ndhF}, \textit{psbBTNH} (four contiguous genes here treated as one region), \textit{rbcL}, \textit{rpoC2}, \textit{rps16}, and \textit{rps4} from the plastid genome; and \textit{atp1}, \textit{matR}, \textit{nad5}, and \textit{rps3} from the mitochondrial genome. The total length of the aligned 17-gene matrix was 25,260 bp and of the 13-gene matrix was 19,846 bp. The percentage of missing data for the full data set was 41\% and for the data set without mtDNA data was 42\%.

#### Alignment and phylogenetic analyses

All sequence data were stored and managed in TOLKIN (Beanan and Cellinese, 2010). TOLKIN is a web application, developed for distance collaboration as part of the Angiosperm Tree of Life project, that allows users to access and share data in real time, as well as among angiosperm phylogenists, is to assemble a data set having both broad taxonomic coverage as well as numerous genes. In the hope of further improving our understanding of angiosperm phylogeny, we have sequenced deeply over a broad representation of angiosperms. We constructed a 17-gene data set for 640 species representing 640 genera, 330 families, and 58 of 59 orders (sensu APG III, 2009) using genes that represent all three plant genomic compartments (nucleus, plastid, and mitochondrion).

One goal of the angiosperm Assembling the Tree of Life (AToL) project was to assess morphological synapomorphies for each of the major clades of angiosperms as well as across angiosperms as a whole. We here also provide putative morphological synapomorphies of Saxifragales, although a more detailed morphological exploration of this clade is presented elsewhere (Carlsward et al., unpublished manuscript). This study of Saxifragales represents the first treatment in a series of analyses for large clades of angiosperms that will address morphological synapomorphies and the clades they support (W. S. Judd et al., unpublished data).

All sequence data were stored and managed in TOLKIN (Beanan and Cellinese, 2010). TOLKIN is a web application, developed for distance collaboration as part of the Angiosperm Tree of Life project, that allows users to access and share data in real time, as well as among angiosperm phylogenists, is to assemble a data set having both broad taxonomic coverage as well as numerous genes. In the hope of further improving our understanding of angiosperm phylogeny, we have sequenced deeply over a broad representation of angiosperms. We constructed a 17-gene data set for 640 species representing 640 genera, 330 families, and 58 of 59 orders (sensu APG III, 2009) using genes that represent all three plant genomic compartments (nucleus, plastid, and mitochondrion).

One goal of the angiosperm Assembling the Tree of Life (AToL) project was to assess morphological synapomorphies for each of the major clades of angiosperms as well as across angiosperms as a whole. We here also provide putative morphological synapomorphies of Saxifragales, although a more detailed morphological exploration of this clade is presented elsewhere (Carlsward et al., unpublished manuscript). This study of Saxifragales represents the first treatment in a series of analyses for large clades of angiosperms that will address morphological synapomorphies and the clades they support (W. S. Judd et al., unpublished data).

All sequence data were stored and managed in TOLKIN (Beanan and Cellinese, 2010). TOLKIN is a web application, developed for distance collaboration as part of the Angiosperm Tree of Life project, that allows users to access and share data in real time, as well as among angiosperm phylogenists, is to assemble a data set having both broad taxonomic coverage as well as numerous genes. In the hope of further improving our understanding of angiosperm phylogeny, we have sequenced deeply over a broad representation of angiosperms. We constructed a 17-gene data set for 640 species representing 640 genera, 330 families, and 58 of 59 orders (sensu APG III, 2009) using genes that represent all three plant genomic compartments (nucleus, plastid, and mitochondrion).

One goal of the angiosperm Assembling the Tree of Life (AToL) project was to assess morphological synapomorphies for each of the major clades of angiosperms as well as across angiosperms as a whole. We here also provide putative morphological synapomorphies of Saxifragales, although a more detailed morphological exploration of this clade is presented elsewhere (Carlsward et al., unpublished manuscript). This study of Saxifragales represents the first treatment in a series of analyses for large clades of angiosperms that will address morphological synapomorphies and the clades they support (W. S. Judd et al., unpublished data).
were partitioned, and unlinked substitution models were applied to each gene. Bootstrapped (BS; Felsenstein, 1985) trees were summarized as majority-rule consensus trees with Phyutility (Smith and Dunn, 2008).

A series of maximum likelihood (ML) analyses was conducted on the combined data sets, ultimately requiring approximately over 1 year of analysis time and approximately 9 years of actual CPU time. Unexpected taxon placements required close examination of the alignment and subsequent realignment when problems were detected. For example, initial analyses indicated that the placement of Polyosma (Escalloniaceae) differed dramatically when mtDNA data were included vs. when they were omitted for this taxon (see Results). Partitioned analyses suggested that the mtDNA signal for Polyosma was providing a spurious placement. In our total evidence analysis (Kluge, 1989; Fig. 1, summary tree, and Fig. 2, complete cladogram; shown as a phylogram in Appendix S1), we therefore omitted the mtDNA data for Polyosma. We subsequently sequenced the mitochondrial genes surveyed here for an additional sample of Polyosma as well as for a second accession of Quintinia, the sister of Polyosma in the trees that included mtDNA sequence data.

Maximum parsimony (MP) has not performed well on some recent large angiosperm data sets; long-branch attraction has clearly played a role in previous analyses of deep-level relationships in angiosperms, especially in analyses with limited taxon sampling (reviewed in Leebens-Mack et al., 2005; Soltis et al., 2005, 2007; Burleigh et al., 2009). Given the current availability of ML programs (e.g., RAxML) that can readily handle large data sets of the size employed here, we focused much of our data analyses on this approach. Nonetheless, we also conducted MP searches, using the parsimony ratchet (Nixon, 1999) approach to thorough and rapid tree searching. These analyses were run in the program PAUP\(^*\) 4.0b10 (Swofford, 2002). Ratchet files were generated with PAUPRat (Sikes and Lewis, 2001) with 50 independent replicates of 500 iterations each. A majority-rule consensus of the best trees from each replicate was generated. Bootstrapping was conducted by generating 500 bootstrap data sets with the SeqBoot module of PHYLIP (Felsenstein, 2005) and running each of these with a PAUPRat-generated ratchet file for a single 500-iteration search. The parsimony searches took just over three CPU years of analysis time. It is noteworthy therefore that, for this data set, a thorough MP analysis took longer than a comparable ML analysis (each total evidence ML analysis required one CPU year of analysis time).

For higher clades, we consistently use PhyloCode names (see Cantino et al., 2007) whenever these are available; these names are always in italics (e.g., Pentapetalae, Mesangiospermae, Rosidae, Fabidae, Malvidae). Note that Rosidae (sensu Cantino et al., 2007) does include Vitaceae. Our use of family and ordinal names follows APG III (2009) as a formal point of reference; for Caryophyllales, we follow Cantino et al. (2007; hence, the use of italics), which matches the APG III circumscription. For additional recent discussion on families and their status, see the Angiosperm Phylogeny Website (Stevens, 2001 onward). We recognize that some broader family circumscriptions favored in these analyses suggested that the mtDNA signal for Polyosma was providing a spurious placement. In our total evidence analysis (Kluge, 1989; Fig. 1, summary tree, and Fig. 2, complete cladogram; shown as a phylogram in Appendix S1), we therefore omitted the mtDNA data for Polyosma. We subsequently sequenced the mitochondrial genes surveyed here for an additional sample of Polyosma as well as for a second accession of Quintinia, the sister of Polyosma in the trees that included mtDNA sequence data.

The 17-gene tree (with mtDNA for all taxa, including Polyosma) and the 13-gene tree (without mtDNA) are presented as online supplementary figures (Appendix S2 and Appendix S3, respectively). Additional online supplementary figures are provided for the trees resulting from the analysis of each genomic partition: nuclear rDNA tree (Appendix S4), plastid tree (Appendix S5), and mtDNA tree (Appendix S6), as well as the single-gene trees (Appendices S7–S20; note that psbBNTH, representing four genes, are combined in one tree). We also provide a tree resulting from analysis of a reduced, 17-gene, 86-taxon data set (Appendix S21) for comparison with a recent 86-taxon tree based on complete plastid genome sequences (Moore et al., 2010).

Finally, we also provide an MP total evidence topology, with mtDNA for all taxa, except Polyosma (Appendix S22). As reviewed below, the MP topology (Appendix S22) is very similar to the ML tree (Figs. 1, 2).

Taxon sampling is so dense, it is decisive for all possible topologies in the sense of Steel and Sanderson (2010; see also Sanderson et al., 2010). This means that there is sufficient taxon coverage to avoid problems of lack of resolution due solely to missing data (though lack of resolution stemming from insufficient sequence information may still be an issue).

**RESULTS**

Each ML analysis of the 17- and 13-gene data sets took 20–32 h on a 32-core (2.93 GHz Xeon ×7350) machine with 128 gb RAM, and analyses of individual genes took 1–11 h.

The best RAxML trees from analyses of the 17-gene and 13-gene matrices are very similar, but with a few noteworthy differences. The 17-gene tree has one major deviation in placement from expected relationships (Polyosma; see above). However, other differences are relatively minor, and in most cases the 17-gene tree gives higher BS support than the 13-gene tree. Hence, only the 17-gene tree (total evidence, but no mtDNA data for Polyosma; e.g., Lecointre and Deleporte, 2005) is discussed below. This tree (Fig. 1, summary tree) has been divided into separate, interconnected subtrees (Fig. 2a–I).

**DISCUSSION**

**Overview**—Broad phylogenetic analyses involving three to five genes have provided important insights into angiosperm phylogeny, but crucial portions of the backbone of the tree were either not resolved or not well supported (e.g., relationships among major lineages of Mesangiospermae and of Pentapetalae). Use of large numbers of genes from nearly compete plastid genome sequences established most deep branches of extant angiosperm diversity with strong support, but taxon density was low (Jansen et al., 2007; Moore et al., 2007, 2010). Here, we have employed numerous genes as well as broad taxonomic coverage. The general topology provided here is very similar to recently published trees that include broad taxonomic coverage of the angiosperms, albeit with fewer genes (e.g., Soltis et al., 2000, 2007; Burleigh et al., 2009) or far fewer taxa and nearly complete plastid genome sequences (e.g., Jansen et al., 2007; Moore et al., 2007, 2010). However, the 17-gene tree is a significant improvement in that it provides much higher levels of support for deep-level relationships than obtained with either three or five genes. Our analyses further support those topologies recovered using complete plastid genome analyses (but based on fewer than 100 exemplars).

Within Acrogynnospermae, the two Gnetales included here (Gnetum and Welwitschia) are embedded within the Coniferae (conifers) and are sister to Pinaceae (BS = 81%) in any potential rooting other than one in which Gnetales are sister to all other seed plants.

Within the angiosperms, Amborellaceae, Nymphaeales, and Austrobaileyales are subsequent sisters to all other extant flowering plants. These placements all receive BS support (BS > 80%; Fig. 1) and are a result obtained across 17-gene and 13-gene data sets. It is noteworthy, however, of the genome partitions, the mtDNA data have Amborella + Nymphaeales (BS = 88%), the cpDNA partition has Amborella + Nymphaeales (BS = 78%), and the 18S/26S rDNA partition has Nymphaeales followed by Amborella as sister to other angiosperms (BS < 50%; Appendices S4–S6). Our interpretation of these results is that the support
Fig. 1.  Summary tree of the maximum likelihood majority-rule consensus from the 17-gene analysis with mtDNA data removed for *Polyosma*. Names of the orders and families follow APG III (2009); other names follow Cantino et al. (2007). Numbers above branches are bootstrap percentages. The number and letter following clade names (i.e., 2a to 2l) refer to clade designations that are used to depict the separate portions of the complete tree in Fig. 2.
from this particular data set for the placement of *Amborella* is not strong. Nonetheless, the placement of *Amborella* as sister to all other extant angiosperms is well supported using complete plastid genome sequences (e.g., Leebens-Mack et al., 2005; Jansen et al., 2007; Moore et al., 2007), as well as in recent analyses of numerous nuclear gene trees (Jiao et al., in press).

Following Amborellaceae, Nymphaeales, and Austrobaileyales, the monophyly of the remaining angiosperms (*Mesangiospermae*) received maximum support. Chloranthales are sister to the magnoliid clade (*Magnoliidae*) with BS = 85%. This clade of *Magnoliidae* + Chloranthales is in turn sister to the remaining angiosperms. Monocotyledoneae (monocots), Ceratophyllaceae, and Eudicotyledoneae (eudicots) form a well-supported clade (BS = 86%) with Monocotyledoneae sister to a weakly supported clade (BS = 68%) of Ceratophyllaceae + Eudicotyledoneae. These deep-level relationships agree with those obtained using complete plastid genome sequences (Moore et al., 2007).

Within Eudicotyledoneae, there is a grade of basal taxa; in contrast to many previous studies, relationships among the members of this basal grade are well supported. Successively more distant sister groups from *Gunneridae* (= core eudicots) are Buxaceae, Trochodendraceae, Proteales plus Sabiaceae, and Ranunculales.

Within Gunneridae, Gunnerales are sister to the remainder of this clade (BS = 99%), which constitute Pentapetalae. *Pentapetalae* comprises (1) a well-supported (BS = 100%) “super-rosid” clade, here named *Superrosidae* (see below), consisting of *Rosidae* (including Vitaceae) and Saxifragales; and (2) a well-supported (BS = 87%) “super-asterid” clade, here named *Superasteridae* (see below), consisting of Berberidopsidales, Santalales, Caryophyllales, Asteridae, and Dilleniaceae, whose position here is at odds with other recent studies (see below).

The MP topology (Appendix S22) is similar to the ML tree (Figs. 1, 2), but there are several noteworthy differences that have also been reported in previous analyses involving MP and ML on smaller data sets (see Soltis et al., 2005, 2007). For example, in the MP tree, Ceratophyllaceae are sister to the Monocotyledoneae rather than Eudicotyledoneae, as found with ML. MP does recover the Superasteridae and Superrosidae clades, but there are differences in relationships among members of Superasteridae with MP and ML. For example, with MP, Dilleniaceae are sister to Caryophyllales (rather than sister to all other Superasteridae, as found with ML). We will not discuss the MP topology further.

**Basal angiosperms**—Our 17-gene analysis (Figs. 1, 2) places Amborellaceae, followed by Nymphaeales, and then Austrobaileyales as well-supported sisters to all other extant angiosperms, in agreement with a series of phylogenetic analyses that have been based on an ever-increasing number of gene sequences (reviewed in Leebens-Mack et al., 2005; Soltis et al., 2005).

Our results also agree with other analyses in providing strong support (BS = 100%) for the placement of Hydatellaceae in Nymphaeales as sister to other members of the clade (Saarela et al., 2007). *Trithuria* has ascidiate carpel development, consistent with placement in Nymphaeales, but as reviewed elsewhere, this placement is important in that it greatly expands the morphological diversity encompassed by Nymphaeales (Rudall et al., 2007).

Following *Amborella* and Nymphaeales, an Austrobaileyales clade (BS = 100%) of Schisandraceae, Austrobaileyaceae, and Trimeniaceae is sister to all remaining angiosperms. Within Austrobaileyales, Austrobaileyaceae are sister to Trimeniaceae + Schisandraceae (including Illiciaceae; see APG III, 2009).

**Magnoliidae + Chloranthales**—Chloranthales are the only member of Chloranthales, APG III, 2009) are well supported here as sister to *Magnoliidae* (BS = 85%). This same sister-group relationship also emerged (with BS = 72%) from complete plastid genome sequencing (Moore et al., 2010), but has not been apparent in previous analyses involving three or five genes (Soltis et al., 2000; Burleigh et al., 2009). Within Chloranthaceae, the relationships among the four genera (*Hedyosmum, (Ascarina + Sarcandra + Chloranthus)* agree with the results of focused analyses of the family (Qi et al., 1999; Doyle et al., 2003; Zhang and Renner, 2003; Eklund et al., 2004).

In agreement with other recent analyses (e.g., Zanis et al., 2002; Hilu et al., 2003; Qi et al., 1999, 2005; Soltis et al., 2007; Moore et al., 2010), *Magnoliidae* comprises two well-supported clades, each with BS = 100%: Magnoliales + Laurales and Piperales + Canellales. Relationships within these clades also agree, for the most part, with previous analyses and are summarized below.

Piperales consist of two well-supported clades: Pipaceae + Saururaceae (BS = 100%) as sister to Aristolochiaceae + Lactidicales (BS = 99%). The parasitic Hydnoraceae, also of Piperales (APG III, 2009), were not included here. Canellales comprise Canellaceae + Winteraceae.

In Magnoliales, Magnoliaceae are sister to the remaining Magnoliales: Degeneriaceae + Myristicaceae are sister to Himantandraceae + *[Eupomatiaceae + Annonaceae]*, albeit without BS > 50%. These results differ from the three-gene (Soltis et al., 2000, 2007) and five-gene (Burleigh et al., 2009) analyses, as well as focused studies on the clade (Sauquet et al., 2003), all of which have placed Myristicaceae as sister to all other Magnoliales. The problem with the placement of Myristicaceae here appears to be the result of the rDNA data. Both plastid and mtDNA place Myristicaceae as sister to all remaining Magnoliales, following earlier analyses. However, the combined 18S + 26S rDNA tree places a strongly supported clade (BS = 94%) of Myristicaceae plus Degeneriaceae (all with very long branches) as embedded well within Magnoliales, sister to Annonaceae (Appendix S4). This result was not apparent in earlier studies that included 18S rDNA (Soltis et al., 2000). However, in this study, 18S rDNA placed Myristicaceae with Chloranthaceae (albeit without BS support > 50%), and the 26S rDNA data placed the family as indicated for the combined data set with strong BS support.

Within Laurales, Calycanthaceae are well supported as sister to the remainder of the clade, as in previous analyses (e.g., Qi et al., 1999; Soltis et al., 1999, 2000). The current analyses also provide more resolution and support of relationships within the remainder of Laurales than the three- and five-gene studies (Soltis et al., 2000; Burleigh et al., 2009), although reasonable resolution and support were obtained with six genes and morphology (e.g., Renner, 1999). Our trees then divide the remaining Laurales into two subclades: (1) Lauraceae sister to (Hernandiaceae + Monimiaceae), and (2) Siparunaceae sister to (Atherospermataceae + Gomortegaceae). These same two clades were recovered by Renner (1999), although the relationships found here within the two clades are not identical to those recovered in that study.

**Monocotyledoneae**—The monocots, or Monocotyledoneae, are not extensively sampled here, in deference to the ongoing
Fig. 2. The maximum likelihood majority-rule consensus from the 17-gene analysis shown as a cladogram with mtDNA data removed for *Polyosma*. This single large tree has been divided into a series of interconnected trees in which subclades are labeled 2a through 2l; these designations match those given in Fig. 1. Names of the orders and families follow APG III (2009); other names follow Cantino et al. (2007). Numbers above branches are bootstrap percentages.
work of the Monocot AToL research group (see Givnish et al., 2010). Nonetheless, the relationships depicted here within the Monocotyledoneae mirror those produced in multigene phylogenetic analyses focused on the clade (e.g., Chase et al., 2006; Graham et al., 2006). On a broad scale, Monocotyledoneae is well supported (BS = 100%), with Acoraceae (Acorales), followed by Alismatales, each well supported (BS = 100%) as subsequent sisters to all other monocots (Petrosaviales were not sampled). A clade of Dioscoreales + Pandanales (BS = 95%) is sister to the remaining Monocotyledoneae. Within this remainder, Liliales followed by Asparagales are sisters to Commelinidae (i.e., commelinid clade; BS = 100%; Arecales, Zingiberales,
Arecales, Poales, and Dasypogonaceae, a family not included here). Relationships within Commelinidae also agree with focused analyses of Monocotyledoneae, with Arecales sister to Commelinidae + Zingiberales; this clade is in turn sister to Poales. Relationships obtained here within the larger Monocotyledoneae clades also largely agree with more focused analyses that have included more taxa (see Chase et al., 2006; Graham et al., 2006; Givnish et al., 2010).

**Eudicotyledoneae**—**Basal eudicots**—Basal eudicot relationships parallel those in other broad analyses (e.g., Soltis et al., 1999, 2000; Burleigh et al., 2009), albeit with higher BS support here. Ranunculales (BS = 100%) are sister to remaining Eudicotyledoneae, which form a clade with strong support (BS = 100%); Ranunculales are then followed by a clade of Sabiaceae + Proteales (BS = 59% in total evidence tree; 65% without mtDNA data); a similar sister-group relationship was recovered with strong support (BS = 80%) with complete plastid genome data (Moore et al., 2010). Following Proteales, both analyses provided strong support for the placements of Trochodendraceae (BS = 100%) followed by Buxaceae (BS = 98%), as subsequent sisters to the core eudicots (Gunneridae). The same relationships were recovered by Moore et al. (2010).

Relationships within the basal eudicot clades are generally well resolved and supported. Within Ranunculales, Eupteleaceae are sister to the remaining taxa (BS = 76%), followed by Papaveraceae, which are well supported (BS = 100%) as sister to the rest of the clade. The remaining Ranunculales form two clades; (1) Circiaeasteraceae + Lardizabalaceae (BS = 78%), and (2) Menispermaeae sister to (Berberidaceae + Ranunculaceae) (BS = 100%). These results are comparable to most recent analyses (e.g., Kim et al., 2004; W. Wang et al., 2009), although there has been disagreement in terms of the sister group to all other Ranunculales. In some analyses, Papaveraceae have appeared in this position (Soltis et al., 2000), whereas in other studies it has been Eupteleaceae (Kim et al., 2004; W. Wang et al., 2009).

Within Proteales, Nelumboaceae are sister to Platanaceae + Proteaceae (with BS = 100%) as in other molecular analyses; recent investigations of floral morphology also indicate similarities between Proteaceae and Platanaceae (von Balthazar and Schönberner, 2009).

**Gunneridae**—Gunnerales are sister to all other core eudicots (i.e., Pentapetalae) with BS = 100%. Pentapetalae, in turn, comprises: (1) the Superrosidae, consisting of Rosidae (including Vitaceae) and Saxifragales (BS = 100%); and (2) the Superasteridae, consisting of Berberidopsidales, Santalales, Caryophyllales, Asteridae, and possibly Dilleniidae (BS = 97%).

**Superrosidae**—Superrosidae consists of Saxifragales as sister to a well-supported (BS = 85%) Rosidae; Rosidae comprises Vitaceae as sister to Malvidae + Fabidae (see
Fig. 2. Continued.
Fig. 2. Continued.
Fig. 2. Continued.
Fig. 2. Continued.
Complete plastid genome data (Moore et al., 2010) as well as inverted repeat (IR) sequence data (Moore et al., in press) also recover Superrosidae with strong support, but relationships differ among the three constituent clades (Vitaceae, all remaining Rosidae, Saxifragales). Complete plastid genome data suggest that Rosidae (excluding Vitaceae) are sister to Saxifragales + Vitaceae, whereas IR sequence data recover the topology found here (Vitaceae sister to all remaining Rosidae + Saxifragales). Focused studies of Superrosidae suggest Saxifragales are sister to Vitaceae + all remaining Rosidae + Saxifragales), in agreement with the results here (see H. Wang et al., 2009). The divergent results produced by analyses of complete plastid genomes may be due to more limited taxon sampling in that analysis compared to the IR analysis (244 taxa; Moore et al., in press) and the present study.

We provide a formal definition of Superrosidae in Appendix 1.

**Saxifragales**—In the 17-gene tree, Peridiscaceae are sister to all remaining Saxifragales, which are well supported (BS = 98%) and form two major clades: (1) a weakly supported clade (BS = 59%) of Paeoniaceae + the “woody clade” (BS = 100%), comprising Cercidiphyllaceae, Daphniphyllaceae, Altingiaceae, and Hamamelidaceae, and (2) core Saxifragales (BS = 100%), comprising two clades, each with BS = 100%: Crassulaceae + Haloragaceae s.l. (i.e., Aphanopetalum, Tetracarpaeaceae, Penthoraceae, and Haloragaceae) and the Saxifragaceae alliance. Within the latter, Saxifragaceae + Grossulariaceae are sister to Iteaceae + Pterostemonaceae. These results agree with the focused analyses of Jian et al. (2008).

Morphological synapomorphies for Saxifragales include basified anthers, usually with the filament attached at a basal pit (but reversals to the dorsifixed condition occur in Ribes, Pterostemon, and Iteaceae), and latrorse dehiscence. All but Peridiscaceae also may be united by follicle fruits, a homoplasious character that is correlated with the ovaries being at least distally distinct. Additionally, the presence of violoid to theoid teeth is possibly synapomorphic, but salicoid teeth occur in Aphanopetalum, and nonglandular teeth occur in Hamamelis,
Itea, and Haloragis, while the condition cannot be assessed in taxa with entire leaves.

Rosidae—Despite extensive progress in elucidating relationships within the angiosperms, Rosidae has long stood out as the largest poorly resolved major clade; deep relationships within Rosidae have been particularly problematic. Our analyses agree with recent analyses of this clade based on parsimony and ML analyses of 12-gene (>20000 bp) and IR (>24000 bp) data sets for over 100 rosid species (H. Wang et al., 2009). Rosidae forms
a well-supported clade (BS = 85%), in which Vitaceae are sister to the remainder (BS = 100%), which in turn consists of two large clades: (1) Fabidae, which includes the nitrogen-fixing clade, Celastrales, Huaceae, Zygophyllales, Malpighiales, and Oxalidales; and (2) Malvidae, which includes Tapiaceae, Brassicales, Malvales, Sapindales, Geraniales, Myrtales, Crossopternales, and Picramniaceae.

Malvidae—Within a well-supported Malvidae (BS = 97%), a clade (BS = 79%) of Myrtales (BS = 100%) + Geraniales (BS = 68%) is sister to the remaining Malvidae (BS = 99%). Within the latter, a well-supported (BS = 100%) Crossopterales are sister to the rest (BS = 100%), followed by Picramniaceae as sister to a clade (BS = 100%) of Sapindales (BS = 100%) + Huertaeales + (Malvales + Brassicaceae; each with BS = 100%). These results are identical to recent topologies resulting from analyses focused on Rosidae (H. Wang et al., 2009).

Over half of the families of Brassicales (Akaniaceae, Emblicaeeae, Gyrostemonaceae, Koeberliniaceae, Moringaceae, Pentadipandraceae, Salvadoraceae, Setchellanthaceae, and Tovariaceae) were not included in our study. Of those Brassicales sampled, relationships agree closely with other analyses (see Rodman et al., 1996, 1998; Hall et al., 2004). Tropaeolaceae and Caricaceae are sisters to a well-supported Limnanthaceae followed by Bataceae, as subsequent sisters to Resedaceae + (Capparidaceae + Brassicaceae). Malvales are also poorly sampled, with only 5 of 10 families included, but the topology matches other studies (e.g., Bayer et al., 1999; Alverson et al., 1999; H. Wang et al., 2009).

Within Sapindales, Nitrariaceae are sister to two clades (although the sister-group relationship is not very well supported):
Fig. 2. Continued.
Fig. 2. Continued.
(1) Burseraceae + Anacardiaceae, and (2) a clade (also without BS >50%) of Sapindaceae sister to Meliaceae + (Simaroubaceae + Rutaceae). These relationships largely agree with other analyses (e.g., Gadek et al., 1996; Mueller et al., 2007); different relationships were reported by H. Wang et al. (2009), albeit with poor sampling and low support.

In Crossoxomatales, there are two well-supported (BS = 100%) clades: Staphyleaceae + [Stachyuraceae + Crossoxomataceae] and Aphelocaceae + Strasburgeriaceae (including Ixeraceae). In Myrtales, two major clades were recovered: one comprising Onagraceae + Lythraceae and the other comprising the large families Myrtaceae and Melastomataceae and associated smaller families. These results are consistent with all previous studies of the order (e.g., Conti et al., 1997; Sytsma et al., 2004). The placement of Combretaceae, however, is not strongly supported, a result that is in agreement with previous studies.

**Fabidae—** Fabidae is well supported (BS = 99%), with Zygophyllales sister (albeit with low support, BS = 57%) to two major clades (each with BS = 100%): the nitrogen-fixing clade (Cucurbitales, Fabales, Fagales, and Rosales) and a clade of Celastrales, Oxalidales (including Huaceae; see APG III), and Malpighiales (the COM group; Endress and Matthews, 2006). These results agree with H. Wang et al. (2009).

In the COM clade, Celastrales (BS = 100%) are sister to a clade (BS = 59%) of Malpighiales (BS =100%) + Oxalidales (BS = 100%). This topology agrees with recent analyses (e.g., Soltis et al., 2007; H. Wang et al., 2009; Wurdack and Davis, 2009). Alternative topologies in which Malpighiales and Oxalidales did not form a clade have been strongly rejected (Wurdack and Davis, 2009).

**Nitrogen-fixing clade—** Within the nitrogen-fixing clade (BS = 100%), Fabales, Fagales, Rosales, and Cucurbitales all have BS = 100%. Fabales appear as sister with strong support (BS = 84%) to the remainder of the clade; Rosales are then sister to a clade (BS = 78%) of Fagales + Cucurbitales. These results are in agreement with recent studies (e.g., Soltis et al., 2005, 2007; H. Wang et al., 2009).

In Rosales (Barbeyaceae and Dirachmaceae were not sampled here), Rosaceae are sister to the remainder of the clade (BS = 100%), which forms two subclades: (1) a well-supported clade (BS = 100%) of Ulmaceae sister to (Cannabaceae + (Urticaceae + Moraceae)) and (2) a clade (BS = 73%) of Elaeagnaceae + Rhamnaceae. These results agree with other recent studies (e.g., Sytsma et al., 2002; H. Wang et al., 2009).

Relationships in Fagales also agree with broad (Soltis et al., 2000, 2007) as well as focused analyses (Manos and Steele, 1997; Li et al., 2002). Nothofagaceae and Fagaceae are successive sisters to the remaining Fagales, which form two clades: (1) Casuarinaceae + Betulaceae and (2) Myricaceae + Juglandaceae (Ticodendraceae and Rhoipteleaceae were not sampled here).

Our results for Cucurbitales are similar to recent analyses (with Corynocarpaceae and Tetrameleaceae not sampled here) (e.g., Zhang et al., 2006; Soltis et al., 2007). Anisophyllaceae and Coriariaceae are subsequent sisters to a clade (BS = 78%) of (Datiscaceae + Begoniaceae) + Cucurbitaceae.

In Fabales, Polygaliaceae + Quillajaceae (BS support < 50%) are sister to a clade (BS = 62%) of Surianaceae + Fabaceae. The topologies for Fabales in recent analyses have varied greatly, but most of these studies have involved only a few DNA regions (see Wojciechowski et al., 2004; Banks et al., 2008; Bruneau et al., 2008; Bello et al., 2009).

**COM clade—** In Celastrales, Lepidobotryaceae are sister to a broadly defined Celastraceae, following other recent focused analyses (Simmons et al., 2001a, b; Zhang and Simmons, 2006). Within Oxalidales, Huaceae are sister to the remainder of the clade, which then forms two well-supported (BS = 100%) sub-clades: (1) Connaraeaceae + Oxalidaceae, and (2) Brunelliaceae, Cephalotaceae, Cunoniaceae, and Elaeocarpaceae. Within the second clade, Brunelliaceae are sister to Elaeocarpaceae + (Cephalotaceae + Cunoniaceae). These relationships differ only slightly from previous analyses. In Zhang and Simmons (2006), results were well supported and identical to those here, but Cephalotaceae were not sampled. In Wurdack and Davis (2009) a weakly supported (BS = 59%) Cephalotaceae + Cunoniaceae were unresolved with Brunelliaceae and Elaeocarpaceae. In H. Wang et al. (2009), Cunoniaceae were sister to (Cephalotaceae + Elaeocarpaceae), albeit without BS support > 50%; the current study also has increased representation of families within Oxalidales over the H. Wang et al. (2009) analysis of Rosidae.

Malpighiales are well represented in the current study, and relationships largely agree with recent analyses (e.g., Davis and Wurdack, 2004; Davis et al., 2001, 2005; Wurdack and Davis, 2009; Ruhfel et al., 2011). The strong agreement between our results and those of Wurdack and Davis (2009) is not surprising given that both studies share considerable data. We overview the major features of this clade below.

We recovered a well-supported (BS = 99%) parietal-placentation clade (sensu Wurdack and Davis, 2009). Within this clade, Lacistemataceae are the well-supported sister to Salicaceae, and Goupiaceae are weakly supported (BS = 68%) as sister to Violaceae. The placement of Goupiaceae was unresolved in Wurdack and Davis (2009). The clusoid clade sensu Wurdack and Davis (2009) was well supported (BS = 100%), and within this clade Calophyllaceae were sister (BS = 74%) to Podostemaceae + Hypericaceae. These relationships are also in agreement with Ruhfel et al. (2011). Balanopaceae are sister to a clade of Trigoniacese + Dicapotetaceae, all of which is sister to Chrysobalanaceae + Euphorbiaceae. Other well-supported (BS = 100%) sister-group relationships recovered here include: Erythroxylaceae + Rhizophoraceae, Elatiaceae + Malpighiaceae, Lophopyxidaceae + Putranjivaceae, and Phyllanthaceae + Picrodendraceae. The recently recognized (Wurdack and Davis, 2009) Euphorbiaceae segregate Peraceae are recovered here as sister to Euphorbiaceae. APG III (2009) has deferred recognition of Peraceae pending additional support from other non-mtDNA gene regions for placement of Rafflesiaceae within Euphorbiaceae + Peraceae. Similar to recent findings (i.e., Davis et al., 2007; Wurdack and Davis, 2009), our preliminary 17-gene analyses that sampled Rafflesiaceae (later excluded with most other parasites in our final analyses, but available upon request) placed that holoparasitic lineage with this clade. Despite some advances in our understanding of Malpighiales here, relationships of many constituent clades still remain unclear (see Wurdack and Davis, 2009).

**Superasteridae—** Superasteridae (BS = 87%) consists of Berberidopsisidae, Santalales, Caryophyllales, Asteridae, and possibly Dilleniaceae (see below). The 17-gene analysis placed Santalales as sister to a weakly supported (BS < 50%) clade of (Berberidopsisidae + Caryophyllales) + Asteridae. The sister grouping of Berberidopsisidae + Caryophyllales received BS = 75%.

Complete plastid genome data (Moore et al., 2010) and IR sequence data (Moore et al., in press) similarly recovered a
well-supported Superasteridae clade; hence, we formally name this clade here (see Appendix 2). However, these three analyses differ in the relationships suggested among members of Superasteridae. Complete plastid genome sequence data place Santalales as sister to the rest, but the precise placement of the enigmatic Berberidopsidae is unclear.

Dilleniaceae were recovered as sister to the remaining Superasteridae (BS = 97%). This position is much better supported than other recent placements of this enigmatic family (e.g., Moore et al., 2010, in press), but despite this high support, we remain cautious about its true placement because of differing placements in recent studies. Using complete plastid genome sequence data, Moore et al. (2010) placed Dilleniaceae as sister to Superrosidae (BS = 64%), but topology tests did not reject alternative positions of Dilleniaceae as sister to Asteridae or to all remaining Pentapetalae. Analysis of IR sequences for 244 taxa placed Dilleniaceae sister to Superrosidae + Superasteridae (BS = 79%) (Moore et al., in press). Because of these varied placements, Dilleniaceae are only tentatively placed here in Superasteridae based on their position in Fig. 1. Additional research is required to place Dilleniaceae definitively.

Santalales—Within Santalales (BS = 100%), the two sampled genera of “Olacaceae”, a family now regarded as nonmonophyletic (see Malécot and Nickrent, 2008; Nickrent et al., 2010), are sister to the well-supported (BS = 100%) remainder of the clade. Santalaceae and Opiliaceae are successive sisters to a clade (BS = 79%) of Schaepeaceae + Lorantheae. Represen-
tation of Santalales is sparse, notably missing Erythropalaceae, which may be sister to the rest of Santalales, and the holoparasitic Balanophoraceae, which have recently been recognized as part of this clade. Relationships among major clades of Santalales have not been resolved with strong BS support in recent analyses, and many segregate families have been newly recognized (e.g., Malécot, 2002; Nickrent et al., 2010).

Caryophyllales—The topology resolved here is similar to that of Brockington et al. (2009). Two major subclades were recovered: core Caryophyllales (Caryophyllineae) and noncore Caryophyllales (Polygonineae) (clade names of Judd et al., 2008). Polygonineae in turn comprise two clades: (1) Plumbaginaceae + Polygonaceae and Tamaricaceae + Frankeniaceae, and (2) Droseraceae + Nepenthaceae sister to Drosophyllaceae + Asteridae + Superasteridae (BS = 98%) of Actinidiaceae + Sarraceniaceae sister to a clade (BS = 92%), (2) a clade (BS = 100%) of Clethraceae + (Cyrillaceae + Ericaceae) (BS = 100%), and (3) a Primulaceae s.l. clade (BS = 100%) in which Maesa is sister to Clavija + Primulaceae s.s. (BS = 100%). As in the 11-gene analysis (Schönengerber et al., 2005), the relationships of Diapensiaceae, Ebenaceae, Saxifragaceae, and Theaceae are not well supported here. A recent 23-gene analysis with approximately 90 taxa of Ericales resolves many of these latter placements (K. Sysmsa et al., unpublished manuscript).

Lamiidae—Lamiidae are well supported (BS = 100%). Within this clade, Icacinaceae + Garryales, and Vahliales are sister to the remaining lamiids, which comprise four major clades: Boraginaeae, Lamiales, Solanales, and Gentianales, among which there is no well-supported resolution. The position of Vahliales as immediate sister to the rest of Lamiidae is consistent with the parsimony analysis of N. F. Refullo-Rodriguez and R. G. Olmstead (unpublished manuscript); however, in their ML and Bayesian analyses, Vahlia is sister to Solanales.

Asteridae—Within Asteridae, Cornales are sister to the remaining taxa (BS = 97%) within which Ericales are sister to a clade (BS = 100%) consisting of two subclades: Lamiales (BS = 100%) + Campanulidae (BS = 100%). Recognition of Asteridae comprising these four groups was first based on a molecular phylogenetic analysis (Olmstead et al., 1992), and the relationships among these groups are congruent with most prior molecular studies (Olmstead et al., 2000; Albach et al., 2001; Bremer et al., 2002), although with stronger support.

Ericales—The monophyly of Ericales is well supported (BS = 100%), but as in other recent studies, resolution within the clade is more problematic. Our study retrieves a well-supported clade (BS = 100%) of Tetrameristaceae + (Maregraviaceae + Balsaminaceae) (see Schönengerber et al., 2010) as sister to the remaining Ericales, which are also well supported as monophyletic (BS = 95%), but within which relationships are even more uncertain, mirroring earlier analyses by Schönengerber et al. (2005) based on 11 genes. Lecythidaceae are strongly supported (BS = 95%) as sister to the remainder, although in the 11-gene study with much greater taxon sampling (Schönengerber et al., 2005), Lecythidaceae were placed in a polytomy at this node. Three other well-supported clades were recovered: (1) Polemoniaceae + Fouquieriaceae (BS = 92%), (2) a clade (BS = 98%) of Actinidiaceae + Sarraceniaceae sister to a clade (BS = 100%) of Clethraceae + (Cyrillaceae + Ericaceae) (BS = 100%), and (3) a Primulaceae s.l. clade (BS = 100%) in which Maesa is sister to Clavija + Primulaceae s.s. (BS = 100%). As in the 11-gene analysis (Schönengerber et al., 2005), the relationships of Diapensiaceae, Ebenaceae, Saxifragaceae, and Theaceae are not well supported here. A recent 23-gene analysis with approximately 90 taxa of Ericales resolves many of these latter placements (K. Sysmsa et al., unpublished manuscript).
A more thorough analysis of the clade is provided by N. F. Refulio-Rodriguez and R. G. Olmstead (unpublished manuscript) with a total of 129 species of Lamidaceae, including all recognized families (APG III, 2009). Their results differ mostly in places where resolution is weak in the results presented here, such as the branching order among the early-diverging groups, Oncothecaceae, Icacinaceae, and Garryales, and among the four major clades of Lamidaceae, Lamiales, Solanales, Gentianales, and Boraginaceae.

Lamiales—Lamiales (BS = 100%) comprise Plocospermataceae, Oleaceae, and Tetrachondraceae as well-supported subsequent sisters (each with BS = 100%) to the rest of Lamiales. Pelikanthera (unassigned to family) plus (Gesneriaceae + Calceolariaceae) form a clade (BS = 97%) sister to the remaining Lamiales, which form a strongly supported core (BS = 100%). Within this core, few relationships have BS values >50% in our analyses, indicating the difficulties in resolving relationships in this clade. Relationships with support >70% include Verbenaceae, Oleaceae, and Tetrachondraceae as well-supported sister groups, followed by Phrymaceae in their results. However, most of these families are represented here by only a single species.

Solanales—The topology for Solanales agrees with other recent analyses, although most have not had complete family coverage of the order (e.g., Soltis et al., 2000). Complete family coverage was obtained in Bremer et al. (2002), although the current study provides much stronger internal support because of the large number of genes included and added sampling within Solanaceae and Convulvulaceae. We found Montiniae are sister to Spencelocaceae + Hydroloaceae; this well-supported clade (BS = 98%) is in turn sister to Convulvulaceae + Solanaceae (BS = 100%).

Gentianales—Rubiaceae are sister to the remainder of the clade, in agreement with other studies (e.g., Backlund et al., 2000; Frasier, 2009). Relationships among the remaining families have been unclear in most previous studies (e.g., Olmstead et al., 2000; Soltis et al., 2000; Bremer et al., 2002), but here Loganiaceae are weakly supported as sister to Gelsemiaceae + (Apoxyaceae + Gentianaceae), which is congruent with a study of four plastid DNA regions (Frasier, 2009). Although the current study does reveal areas of strong support, sampling in the families of this order is low.

Campanulidae—A Campanulidae clade was recovered with strong support (BS = 100%), and relationships within largely agree with recent analyses, especially Tank and Donoghue (2010; also see Lundberg, 2001; Winkworth et al., 2008). This is not surprising given the very large overlap in data between the current study and that of Tank and Donoghue (2010), the major difference being the addition of mtDNA and nuclear rDNA data in the current study. Tank and Donoghue (2010) used their phylogenetic results to discuss floral evolution in Campanulidae, and analyses of biogeographic patterns, especially in relation to past continental connections, are underway.

Each of the four major lineages of Campanulidae—Apiaceae, Aquifoliales, Asterales, and Dipsacales—was recovered here with strong support (BS = 100%). The several smaller Campanulidae clades, including Escalloniaceae sensu Tank and Donoghue, 2010) and Paracyphiaceae (= Paracyphiaceae sensu Tank and Donoghue, 2010), also receive strong support in the current study (BS = 91% and 99%, respectively). Although recovered in the best ML tree (Figs. 1 and 2), bootstrap support for Bruniiales (the Columelliaceae-Bruniaceae clade sensu Tank and Donoghue, 2010) is <50% in the total evidence analysis reported here.

Relationships among the primary Campanulidae clades also agree with previous analyses. We recovered the sister-group relationship of Aquifoliales and Apiidae (i.e., the remaining Campanulidae taxa; Cantino et al., 2007). Likewise, relationships among the six major lineages within Apiidae—Apiaceae, Asterales, Bruniiales, Dipsacales, Escalloniaceae, and Paracyphiaceae—were identical to earlier results (e.g., Zhang et al., 2003; Winkworth et al., 2008; Tank and Donoghue, 2010). To facilitate the study of evolutionary patterns within Apiidae, Tank and Donoghue (2010) named several clades that they recovered with strong statistical support, including Dipsapidae (the Apiales-Dipsacales-Paracyphiaceae clade) and Dipsidae (the Dipsacales-Paracyphiaceae clade). Both of these clades were recovered in the total evidence analysis, albeit with <50% bootstrap support.

Aquifoliales—Aquifoliales were resolved as the sister group to Apiidae. Within Aquifoliales, there are two major lineages: Cardiopetalideae + Stemonuraceae and a clade of Aquifoliaceae + (Phyllonomaclaceae + Helwingiaceae). As did Tank and Donoghue (2010), we found strong support for Phyllonoma-Helwingia as sister to Ilex (Fig. 2). Although Phyllonoma and Helwingia share several conspicuous morphological apomorphies, including ephiphyllous inflorescences, most earlier studies (based on far less data) placed Helwingia directly with Ilex to the exclusion of Phyllonoma (e.g., Morgan and Soltis, 1993; Soltis and Soltis, 1997; Olmstead et al., 2000).

Asterales—Each of the major lineages of Asterales was recovered with strong support, and, for the most part, relationships among them are also well supported (Fig. 2). One exception is at the base of the Asterales, where the successive sister-group relationship of Pentaphragmataceae and the Rousaceae + Campanulaceae clade to the remainder of Asterales received <50% bootstrap support (Fig. 2). In earlier studies, these relationships were reversed, with the Rousaceae-Campanulaceae
clade sister to the rest of Asterales (Lundberg and Bremer, 2003; Winkworth et al., 2008; Tank and Donoghue, 2010). In our analyses of the plastid partition alone, we recovered the same relationships as the earlier studies with strong support (Appendix S5). By contrast, the mtDNA partition provides strong support for the alternative resolution of these relationships (Appendix S6). When the mtDNA gene trees are examined separately, *atp1* and *matR* recovered the mtDNA topology (BS < 50%), while *rps3* resolved the plastid topology (BS values < 50%) (Appendices S17–S20).

**Bruniaceae**—As noted above, Bruniaceae were recovered with strong support by Tank and Donoghue (2010), but were only weakly supported (<50% bootstrap) in our combined analysis. This case is discussed further below (see *Comparison of mtDNA, rDNA, and plastid topologies*). This clade is of biogeographic interest in that Columelliaceae are restricted to South America, while Bruniaceae are a primarily South African lineage; this appears to be an ancient vicariance event in *Campanulidaceae*, and the detailed historical biogeography of this clade is the subject of current research.

**Dipsapiidae**—Tank and Donoghue (2010) named this clade to emphasize the close relationship between Apiaceae and Dipsacaceae (along with Paracyphiaceae, = Paracyphiales sensu APG III)—a result that appeared in several earlier studies (e.g., Bremer et al., 2002; Winkworth et al., 2008) and is strongly supported in their plastid analyses (Tank and Donoghue, 2010). Surprisingly, this clade is not recovered with >50% bootstrap support in our total evidence analyses, although it does appear in the ML BS tree (Figs. 1 and 2). Within *Dipsapiidae*, Paracyphiaceae are resolved as sister to Dipsacales (= Dipsidae sensu Tank and Donoghue, 2010), albeit with <50% bootstrap support. As with *Dipsapididae*, *Dipsidae* were strongly supported in the analyses of both Winkworth et al. (2008) and Tank and Donoghue (2010).

Relationships within Dipsacaceae largely agree with earlier studies (e.g., Bell et al., 2001; Donoghue et al., 2001, 2003; Bell, 2004; Winkworth et al., 2008; Tank and Donoghue, 2010), as do relationships within Apiaceae (e.g., Kärhehd, 2001, 2003; Chandler and Plunkett, 2004; Winkworth et al., 2008; Tank and Donoghue, 2010), with only minor differences in areas with little to no bootstrap support (e.g., within Araliaceae).

**Possible taxon sampling issues**—Several relationships do not agree with more detailed phylogenetic studies focused on particular clades (see also Tank and Donoghue, 2010). For example, within Dipsacaceae (Dipsacaceae), *Pterocephalodes* is sister to *Dipsacus*, and *Scabiosa* is their subsequent sister. The sister group of *Pterocephalodes* and *Dipsacus* was very strongly supported in the analyses of Tank and Donoghue (2010), but is less strongly supported here (BS = 54%). This result is incongruent with the much more densely sampled analyses of Carlson et al. (2009), where *Pterocephalodes* emerged as sister to all other Dipsacaceae. In contrast, while Tank and Donoghue (2010) found a well-supported sister-group relationship between *Echinops* and *Gerbera* within Asteraceae, here we recovered the successive sister-group relationship of *Gerbera* and *Echinops* to the remainder of Asteraceae that appears in more detailed studies of Asteraceae (e.g., Panero and Funk, 2008). In general, these examples highlight that although statistical support may be high for particular relationships, taxon and gene sampling can both have major effects on the outcome, especially when the diversity of the group in question has been grossly undersampled.

**Comparison of mtDNA, rDNA, and plastid topologies**—In the present study, there is some conflict among trees derived from mtDNA, rDNA, and plastid data. However, resolution and support of the rDNA topology is so low that there are no instances of mutual incongruence with BS greater than 75% involving it and either of the other partitions.

One major case of incongruence involves relationships among major lineages of Rosidae. The analysis of mitochondrial genes shows that the weakly supported COM clade is sister to the strongly supported core members of Malvidae (Sapindales, Malvaceae, Brassicales, Huerteales, Picramniaceae) with 92% BS support. This result parallels results obtained in a recent mtDNA analysis of angiosperms (Qiu et al., 2010). An earlier analysis of *matK* also recovered the same relationship, albeit with lower bootstrap support (Zhu et al., 2007). In all previous large-scale analyses of angiosperms or eudicots, the COM clade is sister to the nitrogen-fixing clade (Chase et al., 1993; Savolainen et al., 2000a, b; Soltis et al., 2000; Hilu et al., 2003; Burleigh et al., 2009), occasionally with high support (e.g., BS = 89% in Burleigh et al., 2009). These analyses were based either entirely on plastid genes alone or in combination with nuclear rDNA (18S and 26S rDNA). Thus far, only one large-scale analysis of angiosperms has been performed on nuclear gene data, 18S rDNA (Soltis et al., 1997). In that study, clades corresponding to Fabidae and Malvidae were not recovered, but this has been a general problem with 18S and 26S rDNA data—limited resolution and low support due to low phylogenetic signal (see Soltis et al., 1997).

One line of evidence that lends some support to the possible sister relationship between the COM clade and Malvidae suggested by mtDNA data comes from a broad survey of floral structural characters (Endress and Matthews, 2006). Twenty-two COM families and 18 families of Malvidae share a type of ovule with a thicker inner integument than the outer one, a situation that is otherwise very rare in eudicots (with only one other occurrence, in Trochodendraces). That survey did not find any morphological synapomorphy supporting the monophyly of Fabidae, but this is not surprising in that morphological synapomorphies remain elusive for many major clades of angiosperms.

Likewise, within *Campanulidaceae* there are several instances where mtDNA and plastid analyses are in conflict. The most obvious example of conflict among data sets involved the placement of *Polyosma*. With mtDNA sequence data included, *Polyosma* (Escalloniaceae) is placed sister to Quintinia in the Paracyphiaceae (Appendix S2), but when mtDNA data for *Polyosma* are removed, it is placed in Escalloniaceae. Both alternative placements received high BS support, and nuclear rDNA data do not provide sufficient resolution of the placement of this taxon. These results suggest either a biological phenomenon (e.g., HGT of mtDNA from some member of Paracyphiaceae to the *Polyosma* lineage) or human error (e.g., incorrect labeling of material). All of the plastid sequences for *Polyosma* came from two accessions—one used in Bremer et al. (2002) and one a silica-gel collection obtained from MO (Tank and Donoghue, 2010). The individual plastid gene trees all agree with the placement of *Polyosma* in Escalloniaceae. The *atp1* sequence for *Polyosma* was obtained from our MO accession, while the other three mtDNA genes came from a separate accession (Kew M285-TL265). Because sequences from the MO accession used by Tank and Donoghue (2010) in their plastid analyses agree with previous studies based on a different accession (e.g., Bremer et al., 2002), it is possible that the M285-TL265 accession was misidentified, or tissue or tubes were mislabeled. Significantly, sequencing of the mitochondrial genes surveyed
here in a second accession of *Polysoma*, as well as a second sample of *Quintinia* (the sister of *Polysoma* when mtDNA are included), yielded nearly identical sequences and the same topologies. Thus, available data favor HGT as the more likely cause of incongruence involving *Polysoma*.

Further evidence of conflict between the mtDNA and plastid partitions in *Campanulidae* comes from the lack of bootstrap support in the combined analyses for several clades that were well supported in our analyses of the plastid data alone. In analyses of the plastid partition (Appendix S5), we recovered a Bruniales clade with 79% BS support, in agreement with Tank and Donoghue (2010). However, as noted above, in the combined analysis this relationship received <50% bootstrap support. These results suggest that there is enough conflict between the mtDNA partition and the plastid partition to collapse this relationship. We observed several other such cases, including the *Dipsapinidae* and *Dipsidae* clades and relationships at the base of Asterales, as noted above.

**Impact of taxon density**—One concern with the complete plastid genome analysis of angiosperms by Moore et al. (2010) is that only 86 taxa were employed, raising questions regarding the impact of taxon density on the topology. To test this, we trimmed the 17-gene data set from 640 taxa to the 86 taxa in the analysis by Moore et al. (2010).

The topologies of the 17-gene 640-taxon and reduced 86-taxon (Appendix S21) data sets are very similar, but with a few differences. Topological similarities include strong support for: Chloranthaceae + Magnoliidae (BS = 94%); Monocotyledoneae + (Ceratophyllaceae + Eudicotyledoneae) (BS = 97%); and Trochodendron followed by Buxaceae as subsequent sisters to Gunneridae (BS = 100%; 95%). However, there are noteworthy differences between the trees. First, in the reduced tree, *Amborella* is sister to Nymphaeales (BS = 97%), rather than *Amborella* followed by Nymphaeales as sister to all other extant angiosperms. Second, Nelumbonaceae and Meliosma are subsequent sisters to other angiosperms in the reduced tree, rather than a clade, as in the larger tree. Third, *Superrosidae* are not recovered in the reduced tree; instead, a modified *Superasteridae* (BS = 50%) includes Vitaceae (a member of *Superrosidae* in the complete analysis). Fourth, Dilleniaceae are sister to Caryophyllales (BS < 50%), rather than sister to all *Superasteridae*. This placement of Dilleniaceae as sister to Caryophyllales was also seen in the three-gene tree (*e.g.*, Soltis et al., 2000), but not in the larger tree here. All of these placements not only differ from the 17-gene, 640-taxon analysis, but also from the complete plastid genome topology and from an analysis of the IR for 244 taxa (Moore et al., in press). These results further confirm that reduced taxon density can have a major impact on topology, even when numerous genes are employed. Furthermore, some of these differences could also reflect conflict between plastid and mtDNA data.

**LITERATURE CITED**


Evans, Doyle, Fan, Davis, 726


Fraser, C. L. 2009. Evolution and systematics of the angiosperm order Gentianales with an in-depth focus on Loganiaceae and its species-rich and toxic genus Strachys. PhD dissertation, Rutgers University, New Brunswick, New Jersey, USA.


Manos Muellner

Olmstead

Soltis et al. — Angiosperm phylogeny

Moore

Li

Moore, R.-Q.

Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and sequences.

International Journal of Plant Sciences 164: 553–578.


Stevens, P. F. 2001 onward. Angiosperm Phylogeny Website, version 9, June 2008 [and more or less continuously updated since]. Website http://www.mobot.org/MOBOT/research/APweb/.


Superrosidae W. S. Judd, D. E. Soltis, and P. S. Soltis

Definition
The most inclusive crown clade that contains Rosa cinnamomea L. 1753 (Rosidae) but not Aster amellus L. 1753 (Asteridae). This is a branch-modified node-based definition. Abbreviated definition: >\( V \) Rosa cinnamomea L. 1753 - Aster amellus L. 1753.

Etymology
From the Latin super (above, over, or on top) and Rosidae, a converted clade name based on Rosa, the Latin name for rose (and probably originally from the Greek, rhodon), in reference to the fact that the Superrosidae is intended to apply to a crown clade “above Rosidae.”

Reference phylogeny
The primary reference phylogeny is this paper, see Figs. 1 and 2; see also Burleigh et al. (2009, see fig. 3 and the “full tree” included with their online supplemental information). Therefore, although the existence of the Superrosidae was suspected, and this clade was informally named in Moore et al. (2010), we have not previously provided a formal name for this clade. Only with the present results in hand are we sufficiently confident that Superrosidae represents a well-supported clade and thus is in need of a formal scientific name.

There is some disagreement, however, among recent phylogenetic analyses regarding the position of Dilleniaceae. In the 17-gene analysis (reported here), Dilleniaceae (represented by Tetracera, Hibbertia, and Dillenia) are strongly placed as sister to a clade comprising Santalales, Caryophyllales, Berberidopsidales, and Asteridae, while in an analysis of complete plastid genome sequence data (Moore et al., 2010) they are placed as sister to a Saxifragales + Rosidae clade. Finally, in an analysis of inverted repeat (IR) sequences (see Moore et al., 2010), Dilleniaceae are placed as the sister group to a large clade comprising Rosidae, Saxifragales, Asteridae, Berberidopsidales, Caryophyllales, and Santalales. The use of branch-modified, node-based definitions for both Superrosidae and Superasteridae accommodates placement of Dilleniaceae in either Superrosidae or Superasteridae, or positioned as the sister taxon to a Superrosidae + Superasteridae clade, within the Pentapetalae (which in turn is nested within the Gunneridae).

The essential feature of our concept of Superrosidae is its inclusion of everything that is more closely related to Rosidae (i.e., Fabidae, Malvidae, Vitaceae) than to Asteridae (i.e., Lamidaceae, Campanulidaceae, Ericales, Cornales). This is the feature that we have tried to capture in our definition. Furthermore, when this definition is used in conjunction with our reciprocal definition of Superasteridae (see Appendix 2), it ensures that Superrosidae and Superasteridae are always mutually exclusive, regardless of the placement of Dilleniaceae.

Appendix 2. PhyloCode description of Superasteridae

Superasteridae W. S. Judd, D. E. Soltis, & P. S. Soltis

Definition
The most inclusive crown clade that contains Aster amellus L. 1753 (Asteridae) but not Rosa cinnamomea L. 1753 (Rosidae). This is a branch-modified node-based definition. Abbreviated definition: >\( V \) Aster amellus L. 1753 - Rosa cinnamomea L. 1753.

Etymology
From the Latin super (above, over, or on top) and Asteridae, a converted clade name based on Aster (derived from the Latin, aster, meaning star, so called because of the form of the radiate floral heads of these plants), in reference to the fact that Superasteridae is intended to refer to a crown clade “above Asteridae.”

Reference phylogeny
The reference phylogeny is this paper; see Figs. 1 and 2. See also Burleigh et al. (2009, including online supplemental files 6 and 7). Aster amellus is used as a specifier because it is the type species of Asteraceae (and thus Asteridae, under ICBN, which forms part of this clade name); its close relationship to the species of Rosaceae included in the primary reference phylogeny (i.e., Barnadesia arborea, Gerbera jamesonii, Echinops bannaticus, Tragopogon dubius, T. porrifolius, Cichorium intybus, Lactuca sativa, Tagetes erecta, Guizotia abyssinica, and Helianthus annuus) is supported by the series of phylogenetic studies of Asteraceae presented in Funk et al. (2009).

Composition
Rosidae (incl. Vitaceae), Saxifragales, and possibly Dilleniaceae.

Diagnostic apomorphies
No non-DNA synapomorphies are known.

Synonyms
There are no synonymous scientific names but the informal name “superasterids” was used for this clade in Moore et al. (2010), while the name “superasters” is used in version 9 of Stevens (2001, onwards).

Comments
This is a recently discovered clade that is strongly supported by the extensive molecular analyses reported in this study, although we note that this clade was also well supported in both the 12-gene ML analysis of H. Wang et al. (2009; fig. 1), a study that focused only on relationships within the Rosidae and thus included many fewer sampled taxa, especially among nonrosid taxa, and the 83-gene ML and MP analyses of Moore et al. (2010; fig. 1), which also were based on a much smaller array of sampled taxa. The clade Superasteridae was recovered (but without strong support) in all most parsimonious trees in the three-gene analyses of Soltis et al. (2000; see figs 1(B), 5, and 6) and in some of the most parsimonious trees resulting from the analysis of atpB and rbcL sequences in Savolainen et al. (2000a). The clade received weak support in the ML analyses of angiosperms based on five genes (Burleigh et al., 2009, see fig. 3 and the “full tree” included with their online supplemental information).
Diagnostic apomorphies
No non-DNA synapomorphies are known.

Synonyms
There are no synonymous scientific names, but the informal names “superasterids” (Moore et al., 2010) and “superasterids” Stevens (2001, onward, version 9) have been used for this clade.

Comments
This is a recently discovered clade that is strongly supported by the extensive molecular analyses reported in this study, although we note that this clade was also strongly supported in the 83-gene ML and MP analyses of Moore et al. (2010; see fig. 1), which were based on a much smaller array of sampled taxa. The Superasteridae received weak support in the ML analyses of angiosperms based on five genes (Burleigh et al., 2009, see fig. 3 and their “full tree” included with the online supplemental information). Therefore, although the existence of the Superasteridae was suspected, and this clade was informally named by Moore et al. (2010), we have not previously provided a formal name for this clade. Only with the present results in hand are we sufficiently confident that Superasteridae represents a well-supported clade, and thus is in need of a formal scientific name.

There is some disagreement among recent phylogenetic analyses regarding the position of Dilleniaceae. In the 17-gene analysis (reported here) Dilleniaceae (represented by Tetracera, Hibbertia, and Dillenia) are strongly placed as sister to a clade comprising Santalales, Caryophyllales, Berberidopsidales, and Asteridae, while in an analysis of complete plastid genome sequence data (Moore et al., 2010 they are placed as sister to a Saxifragales + Rosidae clade. Finally, in an analysis of IR sequences (see Moore et al., unpublished) Dilleniaceae are placed as the sister group to a large clade comprising Rosidae, Saxifragales, Asteridae, Berberidopsidales, Caryophyllales, and Santalales. The use of branch-modified node-based definitions for both Superasteridae and Superrosidae accommodates placement of Dilleniaceae in either Superasteridae or Superrosidae, or positioned as the sister taxon to a Superasteridae + Superrosidae clade, within the Pentapetalae (which is in turn nested within the Gunneridae).

The essential feature of our concept of Superasteridae that we have tried to capture in our definition is its inclusion of everything that is closer to Asteridae (i.e., Lamiales, Campanulales, Ericales, Cornales) than to Rosidae (i.e., Fabidae, Malvidae, Vitaceae). Furthermore, when this definition is used in conjunction with our reciprocal definition of Superrosidae (see Appendix 1), it ensures that Superasteridae and Superrosidae are always mutually exclusive, regardless of the placement of Dilleniaceae.