# Toward a resolution of Campanulid phylogeny, with special reference to the placement of Dipsacales

**Richard C. Winkworth[1,2], Johannes Lundberg[3] & Michael J. Donoghue[4]**

[1] *Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Caixa Postal 11461–CEP 05422-970, São Paulo, SP, Brazil. rwinkworth@gmail.com (author for correspondence)*
[2] *Current address: School of Biology, Chemistry, and Environmental Sciences, University of the South Pacific, Private Bag, Laucala Campus, Suva, Fiji*
[3] *Department of Phanerogamic Botany, The Swedish Museum of Natural History, Box 50007, 104 05 Stockholm, Sweden*
[4] *Department of Ecology & Evolutionary Biology and Peabody Museum of Natural History, Yale University, P.O. Box 208106, New Haven, Connecticut 06520-8106, U.S.A.*

Broad-scale phylogenetic analyses of the angiosperms and of the Asteridae have failed to confidently resolve relationships among the major lineages of the campanulid Asteridae (i.e., the euasterid II of APG II, 2003). To address this problem we assembled presently available sequences for a core set of 50 taxa, representing the diversity of the four largest lineages (Apiales, Aquifoliales, Asterales, Dipsacales) as well as the smaller "unplaced" groups (e.g., Bruniaceae, Paracryphiaceae, Columelliaceae). We constructed four data matrices for phylogenetic analysis: a chloroplast coding matrix (*atpB*, *matK*, *ndhF*, *rbcL*), a chloroplast non-coding matrix (*rps16* intron, *trnT-F* region, *trnV-atpE* IGS), a combined chloroplast dataset (all seven chloroplast regions), and a combined genome matrix (seven chloroplast regions plus 18S and 26S rDNA). Bayesian analyses of these datasets using mixed substitution models produced often well-resolved and supported trees. Consistent with more weakly supported results from previous studies, our analyses support the monophyly of the four major clades and the relationships among them. Most importantly, Asterales are inferred to be sister to a clade containing Apiales and Dipsacales. Paracryphiaceae is consistently placed sister to the Dipsacales. However, the exact relationships of Bruniaceae, Columelliaceae, and an *Escallonia* clade depended upon the dataset. Areas of poor resolution in combined analyses may be partly explained by conflict between the coding and non-coding data partitions. We discuss the implications of these results for our understanding of campanulid phylogeny and evolution, paying special attention to how our findings bear on character evolution and biogeography in Dipsacales.

KEYWORDS: Apiales, Asterales, campanulids, character evolution, Dipsacales, phylogeny

## INTRODUCTION

Over the last two decades our understanding of angiosperm phylogeny has improved dramatically. Progress was initially based on large-scale analyses of chloroplast *rbcL* sequences (e.g., Chase & al., 1993; Savolainen & al., 2000b) but data matrices have more recently included loci from all three genomes (e.g., Zanis & al., 2002; Qiu & al., 2005). In addition, studies based on more targeted samples and/or testing novel genetic markers have generally confirmed the same broad patterns of relationship (e.g., Mathews & Donoghue, 2000). Increasing confidence in major relationships is reflected in the emergence of new classification schemes (e.g., APG II, 2003), sometimes entailing new, or at least newly defined, taxonomic names (e.g., Cantino & al., 2007). However, despite remarkable progress several parts of the broad angiosperm tree have remained poorly resolved and this is hampering evolution-

ary studies that depend directly upon such phylogenetic knowledge. Here we focus attention on one of these unresolved regions, specifically relationships among the major lineages within what has become known as the euasterid II (e.g., APG, 1998; APG II, 2003) or campanulid clade (Bremer & al., 2002).

The "asterid" concept dates back more than 200 years (Wagenitz, 1992). Over time views on the composition and circumscription of the group have varied widely, as have ideas about the relationships of asterids to other angiosperms (cf. Cronquist, 1981; Dahlgren, 1989; Thorne, 1992; Tahktajan, 1997). Our contemporary concept of Asteridae has emerged largely from molecular phylogenetic analyses and is therefore relatively recent (e.g., Chase & al. 1993; APG, 1998). The updated APG II (2003) system recognized ten major asterid lineages—two early branching groups (Cornales, Ericales), and the remaining eight split equally between two large clades, the euasterid I (or

lamiid) clade and the euasterid II (or campanulid) clade. Although the campanulid clade is now widely recognized, support for this grouping began to accumulate only in the early 1990s. Initially, chloroplast restriction site mapping (Downie & Palmer, 1992) and *rbcL* sequences (Olmstead & al., 1992, 1993) suggested close links between representatives of Apiales, Asterales, and Dipsacales. Later studies (e.g., Chase & al., 1993; Bremer & al., 1994, 2001, 2002; Backlund & Bremer, 1997; Olmstead & al., 2000; Savolainen & al., 2000a, b; Soltis & al., 2000, 2003; Albach & al., 2001; Kårehed, 2001) confirmed these relationships and, in some cases, suggested additional members of the group, such as *Corokia*, *Escallonia*, *Griselinia*, *Ilex*, and a variety of former Saxifragaceae s.l. and Icacinaceae. Current evidence suggests that the campanulid clade consists of four larger lineages (Apiales, Asterales, Aquifoliales, Dipsacales) and several smaller groups (e.g., Bruniaceae, Columelliaceae, Escalloniaceae, Parachryphiaceae). While we now have a relatively clear picture of the component lineages, we are far less certain about relationships among these groups. In previous analyses, higher-level relationships within campanulids have remained unresolved or the suggested relationships have received only weak support. The six-locus analysis by Bremer & al. (2002) provides a useful example. Here, as in several other analyses (e.g., Olmstead & al., 2000; Soltis & al., 2000) Aquifoliales (newly circumscribed to include *Ilex*, *Helwingia*, *Phyllonoma*, Cardiopteridaceae, and Stemonuraceae) was identified as sister to the rest of the campanulids (with 100% jackknife support). The remaining lineages fell into four groups— Apiales, a weakly supported Asterales-Bruniaceae-Columelliaceae clade, a weakly supported Dipsacales-*Paracryphia*-*Quintinia* clade, and a moderately supported lineage that included *Escallonia*, *Eremosyne*, *Polysoma*, and *Tribeles*. Relationships among these groups were not well resolved or well supported.

Previous analyses have not focused specifically on relationships within the campanulids. Instead, the problem has been addressed within broader studies of angiosperms (e.g., Soltis & al., 2000) or of the Asteridae (e.g., Bremer & al., 2001, 2002). Although these analyses have provided significant insights into campanulid phylogeny, and have effectively settled several important issues (e.g., placement of the root of the campanulids between Aquifoliales and the remaining groups), relationships among Apiales, Asterales, and Dipsacales, as well as the placement of the smaller clades, remain uncertain. In the hope of better resolving these relationships we pursued a more focused approach. Specifically, we compiled four large molecular datasets from currently available sequences for a carefully selected set of 50 campanulid taxa. Analyses of these datasets provide additional insights into relationships among the major campanulid lineage and suggest a possible explanation for why these relationships have remained difficult to resolve.

## MATERIALS AND METHODS

**Taxon sampling and sequences. —** The analyses of Lundberg (2001) and Bremer & al. (2002) contained 36 and 41 campanulids, respectively. These studies broadly overlapped in taxon sampling (i.e., 26 taxa are common to both analyses), with the differences reflecting the inclusion of alternative exemplars (e.g., Lundberg used *Scaevola* to represent Goodeniaceae, whereas Bremer & al. used *Goodenia*) and contrasting sampling strategies (e.g., Lundberg included seven Escalloniaceae and four Dipsacales, whereas Bremer & al. included four Escalloniaceae and seven Dipsacales). Combining the sampling of these two studies resulted in a preliminary set of 51 taxa. Using recent phylogenetic studies of these lineages (e.g., Donoghue & al., 2003, for Dipsacales; Chandler & Plunkett, 2004 and Plunkett & al., 2004, for Apiales) we then refined the sampling within of each major group. Our final dataset contained 50 terminals: 11 representatives of Apiales, 14 of Asterales, 10 of Dipsacales, 5 of Aquifoliales, and 10 lineages not placed within one of the major groups in previous analyses (e.g., Bruniaceae, Escalloniaceae) (Appendix).

Although our taxon sampling broadly overlaps that of both Bremer & al. (2002; 39 terminals are shared) and Lundberg (2001; 29 terminals are shared) there are important differences. These reflect our focus on representing campanulid diversity and the selection of terminals that were well represented by currently available sequences. Most of the sampling differences fall within Apiales and Dipsacales. In Apiales we include representatives of five genera not considered by either Bremer & al. (2002) or Lundberg (2001): *Azorella*, *Hydrocotyle*, *Mackinlaya*, *Myodocarpus*, and *Panax*. In general these broaden the sampled diversity within Apiales, but *Panax* is a replacement for *Aralia* since all of the chloroplast sequences could be taken from a complete chloroplast genome sequence for *Panax ginseng* (GenBank accession AY582139). We included *Apium*, *Pennantia*, and *Torricelia* from the earlier studies, but not *Angelica*, *Hedera*, or *Melanophylla* since closely related exemplars were included. We also increased the sampled diversity within Dipsacales, adding *Adoxa*, *Patrinia*, and *Triplostegia* relative to both of the earlier studies, as well as *Diervilla* and *Morina* relative to the Lundberg (2001) matrix. In addition we substituted two taxa. Bremer & al. (2002) included *Linnaea* but the available *rbcL* sequence appears to be aberrant (see Donoghue & al., 2001) and so we used *Dipelta* as a representative of the Linnaeeae instead. Relative to the Lundberg (2001) dataset we replaced *Symphoricarpos* with *Lonicera* since this is better represented by available sequences. Within Asterales we added *Lobelia*, *Roussea* and *Stylidium* relative to one or both of the earlier studies; compared to Lundberg (2001) we used *Scaevola* instead

of *Goodenia*. We also included additional representatives of Aquifoliales—*Irvingbaileya* compared to Bremer & al. (2002) and both *Irvingbaileya* and *Cardiopterus* relative to Lundberg (2001). Finally, for the unplaced lineages we included *Berzelia* relative to Bremer & al. (2002) and excluded *Anopterus*, *Forgesia*, and *Valdivia* compared to Lundberg (2001) since sequences for only three of the coding markers (*atpB*, *ndhF*, *rbcL*) were available.

For each terminal we compiled the available DNA sequences for seven chloroplast (*atpB*, *matK*, *ndhF*, *rbcL*, *rps16* intron, *trnT-F* region, *trnV-atpE* IGS) and two nuclear loci (18S rDNA, 26S rDNA) from GenBank. Our sampling of chloroplast markers includes the six regions used by Bremer & al. (2002) plus *atpB* gene sequences. Whenever possible we used sequences from a single species to represent a terminal. However, in order to construct as complete a dataset as possible we also included composite terminals in which sequences from between two and four species were used to represent a lineage. For example *Griselinia* is represented by sequences from two species; the *matK* gene, *trnV-atpE* IGS, *rps16* intron, and *trnT-F* region sequences are from *Griselinia littoralis*, whereas the remaining five sequences are from *Griselinia lucida*. In some cases we included different accessions to those included by Bremer & al. (2002) or Lundberg (2001). Most often these replacement sequences are more complete or include less ambiguity than those available at the time of the earlier studies. However, we replaced the *Lonicera ndhF* sequence used by Bremer & al. (2002) since it appeared to be a chimera. A close inspection of preliminary alignments indicated that this sequence was quite unlike other available *Lonicera ndhF* sequences and suggested that it is a composite of sequences from *Lonicera* and *Viburnum*. GenBank accession numbers for all the sequences used are presented in the Appendix.

**Sequence alignment and datasets. —** Multiple sequence alignments were prepared for each of the nine marker loci using a two-stage procedure. First we constructed separate alignments for the representatives of each major lineage (Apiales, Aquifoliales, Asterales, Dipsacales) plus one for the unplaced taxa using ClustalX (Thompson & al., 1997). These initial alignments were visually inspected and adjusted manually for minor improvement. At the second stage we aligned the separate matrices to one another, again using ClustalX followed by visual inspection and manual adjustment.

Sequence alignments for the chloroplast protein coding regions (*atpB*, *ndhF*, *matK*, *rbcL*) were largely unambiguous. However, in several cases we inferred single nucleotide gaps that disrupted the reading frame in the affected sequences. These were assumed to be sequencing errors; insertions were excluded and deletions treated as missing data in subsequent analyses. We also excluded sections at the beginning and end of the coding matrices

either because of alignment ambiguity (which may also reflect sequencing errors) or because some sequences were incomplete and therefore these regions were represented in less than 50% of the taxa. A threshold of 50% representation was also used when considering within-frame length mutations. If a gap was inferred in more than half of the sequences then these positions were excluded, otherwise they were treated as missing data. We used these same general criteria when preparing the matrices for the chloroplast non-coding regions (i.e., *trnV-atpE* IGS, *rps16* intron, and *trnT-F* region) and nuclear rDNA loci (i.e., 18S and 26S), again excluding regions of ambiguous alignment or those represented in less than half of the taxa.

Recent analyses of angiosperm and asterid phylogeny have often used datasets containing loci representing different genomes (e.g., chloroplast or nuclear) or different functional classes (e.g., coding and non-coding). This approach allows the largest possible dataset to be analyzed. We constructed two datasets in this way—a combined chloroplast dataset was compiled by concatenating the seven chloroplast matrices and a combined genome matrix was prepared by adding the rDNA datasets to this. Although large combined matrices provide potentially substantial amounts of information there is the potential that different data partitions may provide conflicting signals (e.g., Winkworth & al., in press.). We constructed separate chloroplast coding (*atpB*, *matK*, *ndhF*, *rbcL*) and chloroplast non-coding (*trnV-atpE* IGS, *rps16* intron, *trnT-F* region) datasets to examine potential differences in signal; however, since there were rDNA sequences for only about half of the taxa in our sample we did not analyze these separately. If a taxon was not represented in one or more of the individual matrices it was treated as missing data in combined matrices. For example, *matK* sequences were not available for either *Lobelia* or *Irvingbaileya*, and therefore these were coded as missing data for the coding, combined chloroplast, and combined genome matrices.

Our trees and data matrices are available in Tree-BASE (study accession number S1799, matrix accession numbers M3284–M3287).

**Bayesian inference. —** We used heterogeneous models in a Bayesian framework to explore campanulid phylogeny. Specifically, we partitioned each of our data matrices by locus; the coding dataset contained four (*atpB*, *matK*, *ndhF*, *rbcL*), the non-coding three (*trnV-atpE* IGS, *rps16* intron, *trnT-F* region), the combined chloroplast seven, and the full combined matrix nine partitions. Based on preliminary analyses, individual partitions were assigned a GTR+I+Γ substitution model with parameter values (e.g., the rate matrix, proportion of invariable sites, gamma shape parameter) assigned independently.

Bayesian analyses were performed using Metropolis-coupled Markov chain Monte Carlo as implemented in MrBayes ver. 3.1.1 (Huelsenbeck & Ronquist, 2001;

Ronquist & Huelsenbeck, 2003). Searches used default settings for an incremental heating scheme (i.e., three "heated" chains, and one "cold" chain) as well as defaults for the priors on the rate matrix (0–100), branch lengths (0–10), gamma shape parameter (0–10), and the proportion of invariable sites (0–1). A dirichlet distribution was used for base frequency parameters and an uninformative prior for the tree topology. Simultaneous, independent pairs of searches were initiated from random start trees and run for 10 million generations, sampling from the posterior distribution of trees every 100 generations (for a total of 100,000 samples). Several approaches were used to determine the appropriate burn-in (the number of generations before apparent stationarity) for each analysis: *(1)* we plotted overall –lnL versus generations, *(2)* we examined the standard deviation of split frequencies, *(3)* we examined the potential scale reduction factor (PSRF), and *(4)* we compared topology and clade support for majority rule consensus trees for the individual analyses in each pair.

**Comparing matrices and topologies. —** A visual inspection of the resulting trees indicated that the placement of several lineages differed between coding and non-coding analyses. We tested the significance of these differences using the Partition Homogeneity Test (ILD; Farris & al. 1994), significantly less parsimonious test (SLP test; Templeton, 1983), and the Shimodaira-Hasegawa Test (SH test; Shimodaira & Hasegawa, 1999) as implemented in PAUP* 4.0b10 (Swofford, 2002).

We conducted a series of ILD tests. We used a test on the full dataset to examine the overall level of incongruence and then a series of reduced matrices to evaluate the influence of specific differences. The full matrix for ILD tests contained 46 terminals—*Berzelia*, *Irvingbaileya*, *Lobelia*, and *Pennantia* were excluded from the test since we only had chloroplast coding sequences. Reduced matrices were constructed by selectively removing Bruniaceae, Columelliaceae, and the *Escallonia* clade either singly or in various combinations. All ILD tests used 1,000 replicates.

SLP and SH tests used 50% majority rule consensus topologies from the Bayesian searches as test trees. Since coding and non-coding analyses used different numbers of taxa we could not compare the topologies directly. Instead we constructed various rival trees by constraining the relationships of interest to reflect the other analysis. Specifically, we constructed *(1)* a rival tree that reflected both within and between clade differences, *(2)* a rival tree that reflected all of the within clade differences, and *(3)* a series of rival trees in which between clade differences where represented in various combinations (i.e., single differences, all possible pairs, and all three together). For SH tests we used a GTR model (with model parameters estimated on the topologies of interest) and estimated the test distribution using 1,000 RELL bootstrap replicates.

### ■ RESULTS

**Alignments of the individual loci. —** The number of available sequences differed for each marker. With the exception of *atpB*, matrices for the chloroplast coding regions were the most complete, with between 46 and 50 terminals represented. In contrast, matrices for the 26S and 18S ribosomal genes were less complete, containing 23 and 27 sequences, respectively. In general, alignments

**Table 1. Statistics for the aligned data matrices.**

|  | *atpB* | *matK* | *ndhF* | *rbcL* | *rps16* intron | *trnT-F* region | *trnV-atpE* IGS | **18S** rDNA | **26S** rDNA |
|---|---|---|---|---|---|---|---|---|---|
| No. of taxa | 34 | 48 | 46 | 50 | 38 | 45 | 37 | 27 | 23 |
| Aligned matrix length (nt) | 1,467 | 1,434 | 2,169 | 1,395 | 781 | 1,231 | 1,427 | 1,737 | 3,320 |
| No. constant sites | 1,008 | 457 | 996 | 915 | 289 | 478 | 756 | 1,529 | 2,690 |
| No. varied sites | 459 | 977 | 1,173 | 480 | 492 | 753 | 671 | 208 | 630 |
| No. parsimony informative sites | 250 | 607 | 731 | 276 | 285 | 447 | 364 | 77 | 251 |
| % GC content range (all sites) | 41.6–51.0 | 31.4–35.9 | 31.7–35.2 | 43.0–45.6 | 33.1–37.7 | 34.9–38.2 | 34.9–38.9 | 48.8–51.1 | 52.8–60.0 |
| % GC content mean (all sites) | 42.7 | 34.1 | 33.0 | 44.5 | 35.8 | 36.8 | 36.8 | 49.7 | 56.6 |
| % GC content range (varied sites only) | 36.2–49.7 | 37.3–43.9 | 31.4–37.9 | 42.2–48.5 | 36.0–41.4 | 39.4–45.8 | 41.0–46.0 | 54.8–72.8 | 57.3–71.7 |
| % GC content mean (varied sites only) | 45.2 | 41.4 | 35.0 | 45.4 | 39.1 | 43.0 | 43.5 | 63.9 | 63.7 |
| % missing data (in brackets, for 50 terminals) | 4.6 (35.1) | 4.0 (8.2) | 7.4 (12.2) | 2.5 (2.5) | 2.2 (25.7) | 5.6 (21.7) | 3.5 (28.6) | 2.0 (47.1) | 9.6 (58.4) |

for the chloroplast coding and nuclear rDNA loci required few gaps to maintain positional similarity. However, short regions flanking several gaps in the rDNA alignments were excluded due to alignment ambiguity. Gaps were more common in non-coding chloroplast alignments. In several cases non-identical but overlapping gaps led to alignment ambiguity; these regions were excluded from subsequent phylogenetic analyses. Statistics for the individual data matrices are presented in the Table 1.

For individual markers the data matrices contained less than 10% missing data. However, in datasets with all 50 terminals the number of cells coded as missing increased dramatically for the more poorly sampled loci. For example, while the amount of missing data remained less than 10% for the two most complete matrices (*matK*, *rbcL*), for the three most incomplete datasets (*atpB*, 18S rDNA, 26S rDNA) more than 30% of the cells were coded as missing.

**Data matrices and gene trees. —** Statistics for the four matrices and the gene trees resulting from Bayesian analyses are presented in Table 2. Likelihood plots, convergence diagnostics, and tree comparisons all suggested that searches had converged within the initial 50,000 generations. For example, after the initial 50,000 generations were excluded the PSRF was between 1.000 and 1.003 for all parameters in every analysis. Therefore, the final results from each analysis are based on 199,000 samples.

Generally our phylogenetic analyses resulted in well-resolved and supported topologies. For example, Fig. 1 shows the combined genome tree. In this analysis only eight nodes are supported by posterior probabilities (PP) of less than 0.95 (i.e., seven nodes have support values of 0.51–0.92 plus the unresolved basal node). Furthermore, trees from each of our analyses were broadly similar to one another. All analyses recovered the same set of eight major lineages—four large clades corresponding to Apiales, Aquifoliales, Asterales, and Dipsacales, plus four smaller ones: Bruniaceae, Columelliaceae, Paracryphi-
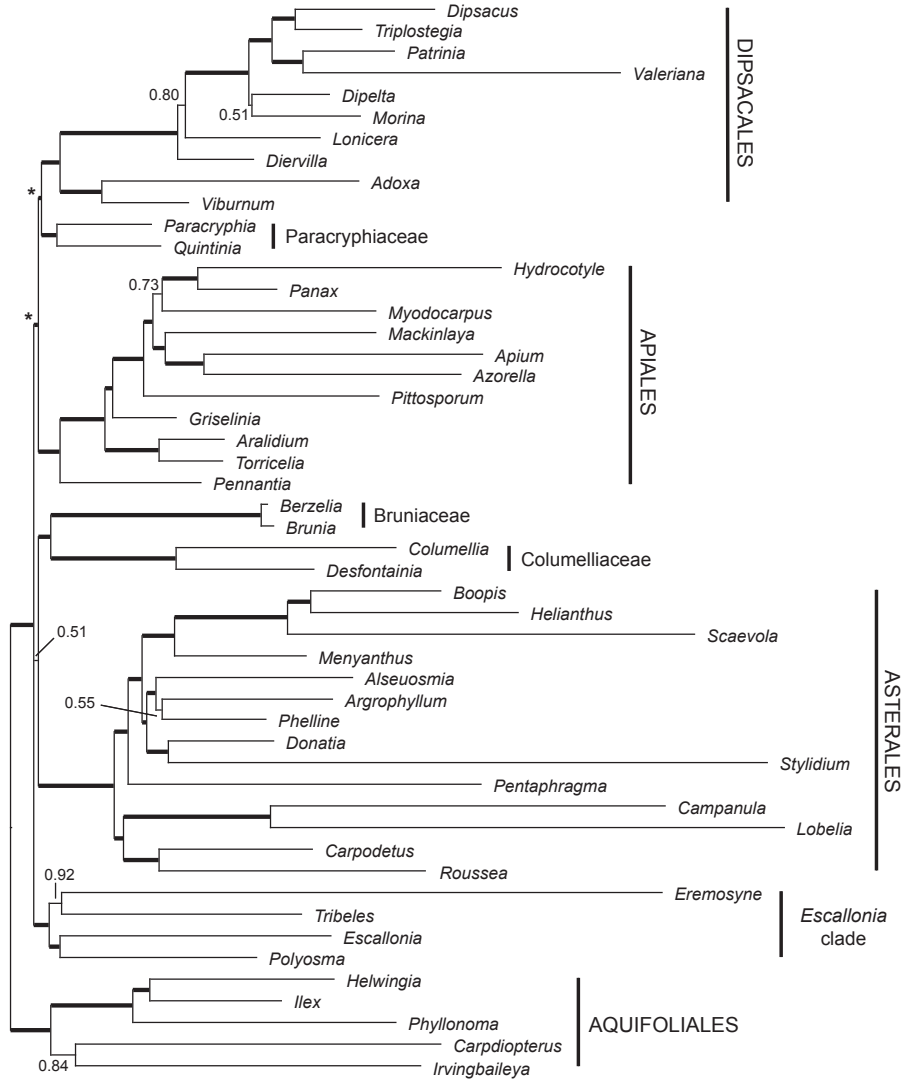
aceae, and a clade containing *Escallonia* and its relatives. These eight clades were well supported in all analyses, in most cases receiving PP of 1.0 (Fig. 2). There are also similarities in the relationships suggested among these lineages. In particular, all four analyses indicated that Apiales and Dipsacales are more closely related to one another than either is to Asterales. Furthermore, Paracryphiaceae is placed as the sister group of the Dipsacales in all of our analyses, although this relationship is only weakly supported by the chloroplast coding data alone (PP = 0.52). In contrast, the placement of the remaining three lineages (i.e., Bruniaceae, Columelliaceae, and the *Escallonia* clade) differed among our analyses. The non-coding data (i.e., Fig. 2A) placed Bruniaceae and Columelliaceae as successive sisters to Asterales; this entire clade is linked with the Apiales-Dipsacales-Paracryphiaceae clade. Rooting along the Aquifoliales branch, the *Escallonia* clade appeared as sister to all of the above in the non-coding analysis. On the other hand in the coding analysis (i.e., Fig. 2B) Bruniaceae, Columelliaceae, and the *Escallonia* clade are united sister to the Apiales-Dipsacales-Paracryphiaceae clade. The combined chloroplast and combined genome analyses favored topologies more consistent with the non-coding results, but with weaker support for the exact placements of the smaller clades (PP < 0.60).

Relationships within each of the major clades were also very similar among analyses, and often well supported (e.g., Fig. 1). In a few cases the suggested relationships differ, but these differences are not well supported in one or more of the analyses. Differences among the analyses include *(1)* the exact arrangement of *Mackinlaya* and *Myodocarpus* within Apiales; *(2)* the placement of *Helwingia* and *Phyllonoma* relative to *Ilex* within Aquifoliales; *(3)* the placements of *Campanula* and *Phelline* within Asterales (in the case of *Campanula* this may reflect only the absence of *Lobelia* in the non-coding analysis); *(4)* the relative placement of *Diervilla* and *Lonicera* within

**Table 2. Statistics for the combined matrices and analyses.**

|  | Chloroplast non-coding | Chloroplast coding | Combined chloroplast | Combined genome |
|---|---|---|---|---|
| No. of taxa | 46 | 50 | 50 | 50 |
| Aligned matrix length (nt) | 3,439 | 6,465 | 9,904 | 14,961 |
| No. varied sites | 1,916 | 3,089 | 5,005 | 5,843 |
| No. parsimony informative sites | 1,096 | 1,864 | 2,960 | 3,288 |
| % missing data | 19.0 | 15.5 | 19.0 | 31.2 |
| No. of model parameters | 30 | 40 | 70 | 90 |
| 95% credible tree set | 34,038 | 6,830 | 1,122 | 1,048 |
| Mean support | 0.906 | 0.922 | 0.936 | 0.941 |
| Mean and variance (in brackets) of tree length | 2.089 (0.003) | 2.246 (0.002) | 2.233 (0.001) | 2.063 (0.001) |

**Fig. 1. The 50% majority rule phylogram from Bayesian analysis of the combined genome data matrix. Both topology and support values are based on the combined post–burn-in generations from a simultaneous pair of analyses. Branches that received Bayesian PP of greater than 0.95 are thickened with values for the remaining edges indicated. Edges subtending the Dipsacales-Paracryphiaceae and the Dipsacales-Paracryphiaceae-Apiales clades (marked with asterisks) are short and the thickening is indistinct, in both cases these received support values of 1.00.**



## DISCUSSION

Dipsacales; and *(5)* the exact arrangement of terminals within the *Escallonia* clade.

**Comparing matrices and topologies. —** ILD tests of coding versus non-coding data suggested substantial differences in the signal provided by these two partitions ($P = 0.043$). Tests on datasets with Bruniaceae, Columelliaceae, and the *Escallonia* clade excluded in various combinations retained such differences—*P* values ranged from 0.022 when the *Escallonia* clade was deleted to 0.006 when Bruniaceae and Columelliaceae were removed. The topology tests are also consistent with substantial differences between coding and non-coding data partitions. *P* values ranged from 0.0005 to 0.1573 (N = 2–191) for SLP tests, and from 0.000 to 0.208 for SH tests (Table 3). These tests suggest that both within and between clade differences make substantial contributions to the overall level of incongruence.

Relationships among the major campanulid lineages have remained uncertain despite considerable effort to resolve the broad patterns of angiosperm phylogeny (e.g., Chase & al. 1993; Savolainen & al., 2000; Qiu & al., 2005). However, since previous analyses have often been very broad in phylogenetic scope it is perhaps not surprising that some areas of the angiosperm tree have remained poorly resolved. Here we have focused exclusively on higher-level relationships within the campanulid clade, using the currently available sequence data for a carefully selected set of 50 terminal taxa to evaluate the phylogeny. Below we describe the general structure and implications of the topologies we recovered, as well as highlight areas of particular importance for future analyses.
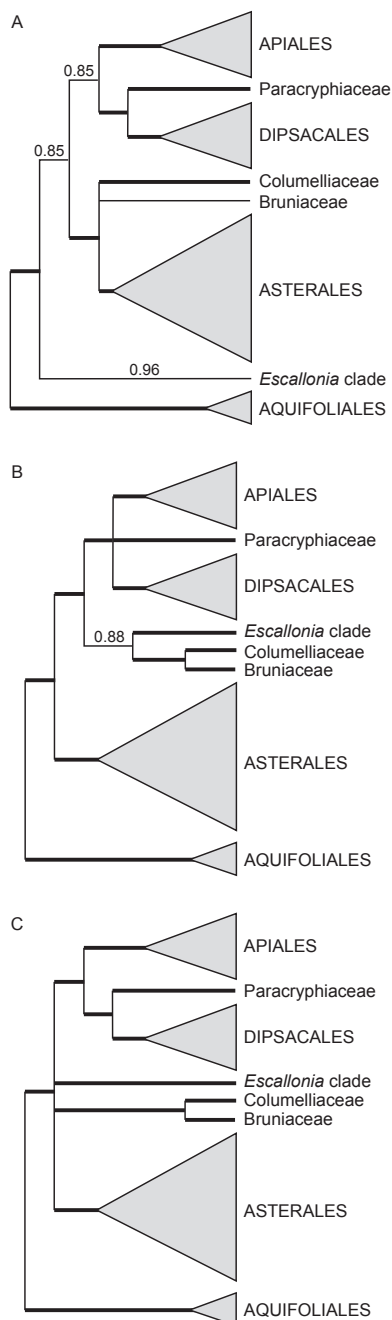
**Fig. 2. Schematic diagrams summarizing relationships among and support for the main campanulid clades in the four analyses. A, chloroplast non-coding; B, chloroplast coding; C, chloroplast combined and combined genome. Diagrams are based on the 80% majority rule consensus of samples remaining after the burn-in was discarded. The four largest clades are represented by triangles, the smaller lineages as single terminals. Based on previous studies trees are rooted using Aquifoliales. Branches that received Bayesian PP of 1.00 are thickened, values for less well-supported edges are given. Note that there is no support value associated with the Bruniaceae in the chloroplast non-coding analysis since the lineage was represented by a single terminal in this matrix.**

**Gene trees. —** Phylogenetic analyses of coding, non-coding, and combined datasets all resulted in broadly similar topologies. These consistently resolve and strongly support relationships among the four major lineages. Specifically, assuming that the root of the campanulids falls along the Aquifoliales branch (based on previous studies, e.g., Bremer & al. 2002), our analyses indicate that Apiales and Dipsacales are more closely related to one another than either is to Asterales. This basic relationship received PP of 1.00 in all but our non-coding analysis. Beyond this, our analyses strongly support the recognition of four smaller clades—specifically, Bruniaceae (including *Brunia* and *Berzelia*), Columelliaceae (including *Columellia* and *Desfontainia*), Paracryphiaceae (including *Paracryphia* and *Quintinia*), and an *Escallonia* clade (including *Escallonia*, *Polyosma*, *Tribeles*, and *Eremosyne*). Moreover, Paracryphiaceae is strongly supported as sister to Dipsacales in all but the coding analysis and there is also strong support for a clade including Bruniaceae and Columelliaceae in all but the non-coding analysis. This effectively reduces the problem to just two clades with uncertain relationship—Bruniaceae-Columelliaceae and the *Escallonia* clade. The non-coding data suggested that the Bruniaceae-Columelliaceae clade is most closely related to Asterales, with the *Escallonia* lineage falling outside of the entire Apiales-Asterales-Dipsacales clade (Fig. 2A). In contrast, coding sequences united the Bruniaceae and Columelliaceae clade with the *Escallonia* clade, and placed this clade as sister to the Apiales-Dipsacales-Paracryphiaceae clade (Fig. 2B). Analyses of combined chloroplast and combined genome datasets recover relationships consistent with those suggested by the non-coding dataset. However, these relationships were more poorly resolved and supported in the combined chloroplast and combined genome analyses (Figs. 1, 2C).

Many of the relationships within major clades were also consistently resolved and well supported in our analyses (Fig. 1). The few exceptions are weakly supported by one of the data partitions. For example, the contrasting placements of *Diervilla* and *Lonicera* within Dipsacales received support values of 0.63 and 0.94 in coding and non-coding analyses, respectively. In this case the combined analysis recovered the better-supported non-coding arrangement, but with a lower PP (0.80). The placement of Diervilleae sister to the remaining Caprifoliaceae (sensu Donoghue & al., 2001) coincides with the results of analyses focused on Dipsacales (e.g., Bell & al., 2001; Donoghue & al., 2003). However, conflict between the data partitions does not always result in reduced support for relationships in the combined analyses. Despite incongruence between coding and non-coding data the placements of *Campanula*, *Helwingia*, and *Mackinlaya* all remain strongly supported. Indeed, in combined analyses support

for relationships within the *Escallonia* clade increases relative to that in separate tests.

**Implications for resolving broad-scale campanulid relationships. —** Our analyses go some way toward resolving broad-scale relationships in the campanulid clade. Specifically, we have found support for Asterales being sister to an Apiales-Dipsacales clade, as well as for Dipsacales plus Paracryphiaceae, and Bruniaceae plus Columelliaceae. These results provide welcome confirmation for relationships that have been suggested in previous analyses, but that were only weakly supported. For example, in their combined analysis Bremer & al. (2002) recovered these same patterns of relationship, but with less than 55% jackknife support. In contrast, the placements of the Bruniaceae-Columelliaceae clade and of the *Escallonia* clade remain uncertain. The relationships of these lineages differed markedly in our coding and non-coding analyses, and support was weak in combined analyses (Fig. 2C).

Although we have not been able to fully resolve broad-scale campanulid phylogeny our analyses do provide some insights into why these relationships may have remained uncertain in previous analyses. Conflict between data partitions may have been an important factor. Specifically, competing signals from the coding and non-coding datasets tend to cancel each other out in our combined analyses, leaving relationships poorly resolved and supported (e.g., Winkworth & al., in press). Whether conflict between data partitions also influenced the results of Bremer & al. (2002) is difficult to ascertain since similar compari-

sons are not reported in quantitative terms. Differences in analytical approach may also help explain problems confidently resolving campanulid phylogeny. Specifically, comparisons with preliminary parsimony and maximum likelihood analyses indicated that although these methods often recover the same among-clade relationships, parsimony tests provide lower support values than model-based approaches. Differences in analytical approach may also help to explain differences between Bremer & al. (2002) and our study with regard to within clade relationships. Most noticeably, in Asterales the Bremer & al. (2002) topology united *Stylidium* with *Campanula*, and we recovered this relationship in our own preliminary parsimony analyses. In contrast, consistent with other studies (e.g., Kårehed, 2002; Lundberg & Bremer, 2003) model-based analyses suggest a *Stylidium-Donatia* clade and also a direct link between this clade and the *Alseuosmia* clade. Since *Stylidium* and *Campanula* are among the longest edges in the tree we suspect that long-branch attraction may have influenced the parsimony result in this case (Fig. 1). Sequence alignment may also have played a role in limiting the ability of previous studies to recover relationships within the campanulid clade. Specifically, these studies focused on resolving broad patterns across angiosperm (e.g., Qiu & al., 2005) or Asteridae phylogeny (e.g., Bremer & al., 2002) rather than specifically on relationships within the campanulid clade. The much wider taxon sampling required for such studies may have introduced additional conflicting signal or led to key lineages being omitted from the analyses.

Table 3. SLP and SH test scores from comparisons of coding and non-coding datasets with various rival constraints.

|  | Test data and tree | Rival constraint | SLP test | | SH test |
|---|---|---|---|---|---|
|  |  |  | N | *P* value | *P* value |
| Full test | Coding | Non-coding | 191 | 0.0095 | 0.000 |
|  | Non-coding | Coding | 92 | 0.0047 | 0.000 |
| Among three clades | Coding | Non-coding | 42 | 0.0005 | 0.020 |
|  | Non-coding | Coding | 15 | 0.0707 | 0.066 |
| Bruniaceae and Columelliaceae | Coding | Non-coding | 43 | 0.0055 | 0.024 |
|  | Non-coding | Coding | 10 | 0.0578 | 0.068 |
| Bruniaceae and *Escallonia* clade | Coding | Non-coding | 38 | 0.0022 | 0.043 |
|  | Non-coding | Coding | 14 | 0.0325 | 0.044 |
| Columelliaceae and *Escallonia* clade | Coding | Non-coding | 41 | 0.0046 | 0.031 |
|  | Non-coding | Coding | 6 | 0.0143 | 0.065 |
| Bruniaceae alone | Coding | Non-coding | 46 | 0.0390 | 0.054 |
|  | Non-coding | Coding | 10 | 0.0114 | 0.061 |
| Columelliaceae alone | Coding | Non-coding | 46 | 0.0768 | 0.073 |
|  | Non-coding | Coding | 4 | 0.0455 | 0.097 |
| *Escallonia* alone | Coding | Non-coding | 16 | 0.1336 | 0.208 |
|  | Non-coding | Coding | 2 | 0.1573 | 0.201 |
| Only within-clade differences | Coding | Non-coding | 166 | 0.0512 | 0.001 |
|  | Non-coding | Coding | 80 | 0.0116 | 0.002 |

in ovary position along the edge leading to Dipsacales. However, the current analyses suggest a different interpretation of ovary evolution. Based on the apparent similarity of the semi-inferior ovary in *Quintinia* and Adoxaceae it seems possible that an inferior ovary is not a synapomorphy of Dipsacales but instead arose earlier in campanulid history. Instead the phylogenetic relationships suggested by our trees imply that opposite leaves and sympetalous corollas might be apomorphies for Dipsacales; members of Paracryphiaceae have alternate leaves and free perianth parts. Obviously such reconstructions remain speculative and further studies are needed before we can confidently identify synapomorphies for Dipsacales. Importantly we need to confirm the present phylogenetic results with data from additional markers, especially those from the nuclear and mitochondrial genomes. It will also be critical to further test Lundberg's (2001) finding that *Sphenostemon* (formerly allied with Aquifoliaceae) is linked with Paracryphiaceae. In addition detailed morphological and anatomical analyses of Paracryphiaceae, as well as direct comparisons with the corresponding characters in Dipsacales, are needed. Such studies will be crucial for establishing synapomorphies for Dipsacales, but they would also help to clarify the morphological links between Dipsacales and Paracryphiaceae.

A direct link between Dipsacales and Paracryphiaceae also has important biogeographic implications. Dipsacales has a predominantly north-temperate distribution, occurring in habitats ranging from temperate and boreal forests to the seasonally arid Mediterranean, as well as the high mountains of Asia, Europe and North America. A few lineages have distributions that extend into the Southern Hemisphere—notably *Valerianaceae* and *Viburnum* in the mountains of Latin America. Previous studies have led to the view that the Dipsacales originated in Asia, where each of the major Dipsacales lineages began to diversify before moving out around the Northern Hemisphere. The few clades that reached South America appear to have done so only very recently (Donoghue & al., 2003; Bell, 2004; Bell & Donoghue, 2005a, b; Winkworth & Donoghue, 2005). Given this scenario for the historical biogeography of Dipsacales a direct link with the primarily South American Columelliaceae is problematical since it suggests an ancient presence in South America. In contrast, the direct link with Paracryphiaceae suggested here is at least consistent with an origin of Dipsacales in the Old World. The current distribution of Paracryphiaceae from the Philippines through New Guinea and northeastern Australia, to New Caledonia and New Zealand suggests a split—presumably by the late Cretaceous (Bremer & al., 2004; Bell & Donoghue, 2005a)—between an Australasian Paracryphiaceae clade and a more northern, Asian, Dipsacales clade adapted to existence in colder climates.

## LITERATURE CITED

**Albach, D.C., Soltis, P.S., Soltis, D.E. & Olmstead, R.G.** 2001. Phylogenetic analysis of asterids based on sequences of four genes. *Ann. Missouri Bot. Gard.* 88: 163–212.

**APG (Angiosperm Phylogeny Group).** 1998. An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Gard.* 85: 531–553.

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

**Backlund, A. & Bremer, B.** 1997. Phylogeny of the Asteridae s. str. based on *rbcL* sequences, with particular reference to the Dipsacales. *Pl. Syst. Evol.* 207: 225–254.

**Backlund, A. & Donoghue, M.J.** 1996. Morphology and phylogeny of the order Dipsacales. In: Backlund, A., *Phylogeny of the Dipsacales*. Doctoral Dissertation. Uppsala University, Uppsala.

**Bell, C.D.** 2004. Preliminary phylogeny of Valerianaceae (Dipsacales) inferred from nuclear and chloroplast DNA sequence data. *Molec. Phylog. Evol.* 31: 340–350.

**Bell, C.D. & Donoghue, M.J.** 2005a. Dating the Dipsacales: comparing models, genes, and evolutionary implications. *Amer. J. Bot.* 92: 284–296.

**Bell, C.D. & Donoghue, M.J.** 2005b. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms Divers. Evol.* 5: 147–159.

**Bell, C.D., Edwards, E.J., Kim, S.-T. & Donoghue, M.J.** 2001. Dipsacales phylogeny based on chloroplast DNA sequences. *Harvard Pap. Bot.* 6: 481–499.

**Bremer, K., Backlund, A., Sennblad, B., Swenson, U., Andreasen, K., Hjertson, M., Lundberg, J., Backlund, M. & Bremer, B.** 2001. A phylogenetic analysis of 100+ genera and 50+ families of euasterids based on morphological and molecular data with notes on possible higher level morphological synapomorphies. *Pl. Syst. Evol.* 229: 137–169.

**Bremer, B., Bremer, K., Heidari, N., Erixon, P., Olmstead, R.G., Anderberg, A.A., Källersjö, M. & Barkhoddarian, E.** 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. *Molec. Phylog. Evol.* 24: 274–301.

**Bremer, K., Friis, E.M. & Bremer, B.** 2004. Molecular phylogenetic dating of asterid flowering plants shows Early Cretaceous diversification. *Syst. Biol.* 53: 496–505.

**Bremer, B., Olmstead, R.G., Struwe, L. & Sweere, J.A.** 1994. *rbcL* sequences support exclusion of *Retzia*, *Desfontainia*,

and *Nicodemia* (Buddlejaceae) from Gentianales. *Pl. Syst. Evol.* 190: 213–230.

Cantino, P.D., Doyle, J.A., Graham, S.W., Judd, W.S., Olmstead, R.G., Soltis, D.E., Soltis, P.S. & Donoghue, M.J. 2007. Towards a phylogenetic nomenclature of *Tracheophyta*. *Taxon* 56: 822–846.

Chandler, G.T. & Plunkett, G.M. 2004. Evolution in Apiales: nuclear and chloroplast markers together in (almost) perfect harmony. *Bot. J. Linn. Soc.* 144: 123–147.

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

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

Dahlgren, G. 1989. An updated angiosperm classification. *Bot. J. Linn. Soc.* 100: 197–203.

Donoghue, M.J., Bell, C.D. & Winkworth, R.C. 2003. The evolution of reproductive characters in Dipsacales. *Int. J. Pl. Sci.* 164: S453–S464.

Donoghue, M.J., Eriksson, T., Reeves, P.A. & Olmstead, R.G. 2001. Phylogeny and phylogenetic taxonomy of Dipsacales, with special reference to *Sinadoxa* and *Tetradoxa* (Adoxaceae). *Harvard Pap. Bot.* 6: 459–479.

Donoghue, M.J., Ree, R.H. & Baum, D.A. 1998. Phylogeny and the evolution of flower symmetry in the Asteridae. *Trends Pl. Sci.* 3: 311–317.

Downie, S.R. & Palmer, J.D. 1992. Restriction site mapping of the chloroplast DNA inverted repeat: a molecular phylogeny of the Asteridae. *Ann. Missouri Bot. Gard.* 79: 266–283.

Farris, J.S., Källersjö, M., Kluge, A.G. & Bult, C. 1994. Testing significance of incongruence. *Cladistics* 10: 315–319.

Howarth, D.G. & Donoghue, M.J. 2005. Duplications in the CYC-like genes of Dipsacales correlate with floral form. *Int. J. Pl. Sci.* 166: 357–370.

Huelsenbeck, J.P. & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.

Kårehed, J. 2001. Multiple origin of the tropical forest tree family Icacinaceae. *Amer. J. Bot.* 88: 2259–2274.

Kårehed, J. 2002. *Evolutionary Studies in Asterids Emphasising Euasterids II*. Doctoral Dissertation. Uppsala University, Uppsala.

Lundberg, J. 2001. *Phylogenetic Studies in the Euasterids II With Particular Reference to Asterales and Escalloniaceae*. Doctoral Dissertation. Department of Systematic Botany, EBC, Uppsala University, Uppsala.

Lundberg, J. & Bremer, K. 2003. A phylogenetic study of the order Asterales using one morphological and three molecular data sets. *Int. J. Pl. Sci.* 164: 553–578.

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

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

Olmstead, R.G., Bremer, B., Scott, K.M. & Palmer, J.D. 1993. A parsimony analysis of the Asteridae sensu lato based on *rbcL* sequences. *Ann. Missouri Bot. Gard.* 80: 700–722.

Olmstead, R.G., Kim, K.-J., Jansen, R.K. & Wagstaff, S.J. 2000. The phylogeny of the Asteridae sensu lato based on chloroplast *ndhF* gene sequences. *Molec. Phylog. Evol.* 16: 96–112.

Olmstead, R.G., Michaels, H.J., Scott, K.M. & Palmer, J.D. 1992. Monophyly of the Asteridae and identification of their major lineages from DNA sequences of *rbcL*. *Ann. Missouri Bot. Gard.* 79: 249–265.

Plunkett, G.M., Chandler, G.T., Lowry, P.P., II, Pinney, S.M. & Sprenkle, T.S. 2004. Recent advances in understanding Apiales and a revised classification. *S. African J. Bot.* 70: 371–381.

Qiu, Y.-L., Dombrovska, O., Lee, J., Whitlock, B.A., Bernasconi-Quadroni, F., Rest, J.S., Davis, C.C., Borsch, T., Hilu, K.W., Renner, S.S., Soltis, D.E., Soltis, P.S., Zanis, M.J., Cannone, J.J., Gutell, R.R., Powell, M., Savolainen, V. & Chase, M.W. 2005. Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Pl. Sci.* 16: 815–842.

Ree, R.H. & Donoghue, M.J. 1999. Inferring rates of change in flower symmetry in asterid angiosperms. *Syst. Biol.* 48: 633–641.

Ronquist, F. & Huelsenbeck, J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

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

Savolainen V., Fay, M.F., Albach, D.C., Backlund, A., van der Bank, M., Cameron, K.M., Johnson, S.A., Lledó, M.D., Pintuad, J.-C., Powell, M., Sheahan, M.C., Soltis, D.E., Soltis, P.S., Weston, P., Whitten, W.M., Wurdack, K.J. & Chase, M.W. 2000b. Phylogeny of the eudicots: a nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bull.* 55: 257–309.

Shimodaira, H. & Hasegawa, M. 1999. Multiple comparisons of loglikelihoods with applications to phylogenetic inference. *Molec. Biol. Evol.* 16: 1114–1116.

Soltis, D.E., Senters, A.E., Kim, S., Thompson, J.D., Soltis, P.S., Zanis, M.J., de Craene, L.S., Endress P.K. & Farris, J.S. 2003. Gunnerales are sister to other core eudicots and exhibit floral features of early-diverging eudicots. *Amer. J. Bot.* 90: 461–470.

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

Swofford, D.L. 2002. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4. Sinauer Associates, Sunderland, Massachusetts.

Takhtajan, A.L. 1997. *Diversity and Classification of Flowering Plants*. Columbia Univ. Press, New York.

Templeton, A.R. 1983. Phylogenetic inference from restriction

endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37: 221–244.

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

Thorne, R.F. 1992. An updated phylogenetic classification of the flowering plants. *Aliso* 13: 365–389.

Wagenitz, G. 1992. The Asteridae: evolution of a concept and its present status. *Ann. Missouri Bot. Gard.* 79: 209–217.

Wiens, J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* 52: 528–538.

Winkworth, R.C. & Donoghue, M.J. 2005. *Viburnum* phylogeny based on combined molecular data: implications for taxonomy and biogeography. *Amer. J. Bot.* 92: 653–666.

Winkworth, R.C., Bell, C.D. & Donoghue, M.J. In press. Mitochondrial sequence data and Dipsacales phylogeny: mixed models, partitioned Bayesian analyses, and model selection. *Molec. Phylogen. Evol.*

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

Zhang, W.-H., Chen, Z.-D., Li, J.H., Chen, H.-B. & Tang, Y.-C. 2003. Phylogeny of the Dipsacales s.l. based on chloroplast *trnL-F* and *ndhF* sequences. *Molec. Phylog. Evol.* 26: 176–189.

**Appendix. Taxa, authorities, and GenBank accession numbers for the sequences used in this study. Accession numbers are given in the following order: *atpB*, *matK*, *ndhF*, *rbcL*, *rps16* intron, *trnT-F* region, *trnV-atpE* IGS, 18S rDNA, 26S rDNA. An dash indicates the sequence was not available for the species.**

**APIALES:** *Apium graveolens* L., AJ235396, AJ429370, AJ429124, L01885, AJ431081, AJ430956, AJ429716, AF206852, AF479195; *Aralidium pinnatifidum* Miq., –, U58627, AJ429126, AF299087, AJ431083, AJ430083, AJ429718, –, AY189036; *Azorella selago* Hook.f., –, AF271762, –, AY188417, –, –, –, –, AY189046; *Azorella trifurcata* Pers., –, –, –, –, AF110599, –, –, –, –; *Griselinia littoralis* (Raoul) Raoul, –, AJ429372, –, –, AJ431084, AJ430958, AJ429719, –, –; *Griselinia lucida* (J.R. Forst. & G. Forst) G. Forst, AF209595, –, AF130205, L11225, –, –, –, AF206922, AF479197; *Hydrocotyle bowlesioides* Mathias & Constance, –, U58600, –, U50235, –, AY393699, –, –, AY189076; *Hydrocotyle rotundifolia* Roxburgh ex DC., –, –, –, –, AF110607, –, –, –, –; *Hydrocotyle sinensis* J.F. Gmel., –, –, –, –, –, –, X16605, –; *Mackinlaya macrosciadia* (F. Muell.) F. Muell., –, U58613, –, U50247, –, AY393701, –, –, AY189083; *Myodocarpus fraxinifolius* Brongn. & Gris, –, –, –, –, –, AY393702, –, –, –; *Myodocarpus involucratus* Dubard & Viguier, –, AF271749, –, AY188430, –, –, –, –, AY189095; *Panax ginseng* C.A. Mey, AY582139, AY582139, AY582139, AY582139, AY582139, AY582139, AY582139, D83275, –; *Panax quinquefolium* L., –, –, –, –, –, –, –, –, AF479197; *Pennantia cunninghamii* Miers, AJ494841, AJ494845, AJ312970, AJ494843, –, –, –, –, –; *Pennantia corymbosa* J.R. Forst. & G. Forst, –, –, –, –, –, –, –, –, AY189099; *Pittosporum fairchildi* (Cheesem.), AJ235561, –, –, –, –, –, –, –, –; *Pittosporum japonicum* Hort. ex Presl, –, –, –, –, –, –, –, L28142, –; *Pittosporum tobira* Dryand. ex Ait., –, –, AF130201, U50261, –, –, –, –, AY189151; *Pittosporum undulatum* Vent., –, AJ429374, –, –, AJ431086, AJ430960, AJ429720, –, –; *Torricelia tiliaefolia* DC., –, AJ429375, AJ429127, AF299089, AJ431087, AJ430961, AJ429721, –, AY189113.

**AQUIFOLIALES:** *Cardiopterus quinqueloba* Hassk., –, AJ429310, AJ312963, AJ402936, AJ431021, AJ430898, AJ429665, –, –; *Helwingia japonica* (Thunb. ex Murray) F. Dietrich, AF209596, AJ430195, AF130207, L11226, AJ431089, AJ430963, AJ429723, U42524, –; *Ilex crenata* Thunb., AJ235502, –, –, –, –, –, –, –, –; *Ilex mitis* (L.) Radlk., –, –, –, X98730, –, –, –, –, –; *Ilex opaca* Soland., –, –, –, –, –, –, –, AF206938, AF479203; *Ilex* sp., –, AJ429376, AJ429128, –, AJ431088, AJ430962, AJ429722, –, –; *Irvingbaileya* sp., AJ236219, –, AJ236255, AF156733, –, –, AJ235999, AF479202; *Phyllonoma laticuspis* Engl., AJ236216, –, AF130208, L11201, –, –, U42546, –; *Phyllonoma ruscifolia* Willd. ex Schult., –, AJ429377, –, –, AJ431090, AJ430964, AJ429724, –, –.

**ASTERALES:** *Alseuosmia macrophylla* A. Cunn., AJ236198, AJ429378, AF060157, X87377, AJ431091, AJ430965, AJ429725, AF206844, –; *Argophyllum* sp., AJ318965, AJ429379, AJ238335, X87379, AJ431092, AJ430966, AJ429726, –, –; *Boopis anthemoides* Juss., –, –, L39384, L13860, –, AJ430969, –, –, –; *Boopis graminea* Phil., AJ236199, AJ429382, –, –, AJ431095, –, AJ429729, AF107583, AF479184; *Campanula elatines* L., –, AJ430387, –, –, –, AJ430970, AJ430391, –, –; *Campanula ramosissima* Sibth. & Sm., –, –, L39387, L13861, –, –, –, U42510, –; *Campanula trachelium* Brot., AJ235423, –, –, –, –, –, –, –, AF479191; *Carpodetus serratus* Forst., AJ318967, AJ429383, AJ238336, Y08461, –, AJ430971, AJ429535, –, –; *Donatia fascicularis* Forst., –, AJ429384, AJ225074, X87385, AJ431096, AJ430972, –, –, –; *Donatia* sp., AJ236203, –, –, –, –, –, –, U42510, AF479189; *Helianthus annuus* L., AJ236205, AJ429380, L39383, AF097517, AJ431093, AJ430967, AJ429727, AF107577, AF479183; *Lobelia angulata* Forst., AJ235524, –, –, –, –, –, –, U42785, –; *Lobelia cardinalis* L., –, –, AF130187, AF042659, –, –, –, –, –; *Lobelia puberula* Michx., –, –, –, –, –, –, –, –, AF148276; *Menyanthes trifoliata* L., AJ235533, AJ429386, L39388, L14006, AJ431098, AJ430974, AJ429731, AJ236009, AF479185; *Pentaphragma ellipticum* Poulsen, AJ318980, AJ429387, AF130183, L18794, AJ431099, AJ430975, AJ429732, –, –; *Phelline billardierei* Pancher ex Loes., –, –, –, –, –, –, –, AF206989, –; *Phelline comosa* Labill., AJ235557, –, –, –, –, –, –, –, AF479188; *Phelline lucida* Viell. ex Baill., –, AJ429388, AJ238343, AJ238347, AJ431100, AJ430976, AJ429733, –, –; *Roussea simplex* Sm., AJ235586, AJ429389, AJ277384, AF084477, AJ431101, AJ430977, AJ429734, U42548, AF479243; *Scaevola aemula* R. Br., AF209670, –, –, –, –, –, –, AJ236008, –; *Scaevola frutescens* K. Krause, –, –, L39385, L13932, –, –, –, –, –; *Scaevola* sp., –, AJ429385, –, –, AJ431097, AJ430973, AJ429730, –, –; *Stylidium bulbiferum* Benth., –, AJ429390, AJ225077, AJ225054, AJ431102, AJ430978, AJ429735, –, –; *Stylidium graminifolium* Sw., AJ236201, –, –, –, –, –, –, AJ236011, –.

**DIPSACALES:** *Adoxa moschatellina* L., –, AF446900, AF447020, L01883, –, AF366927, –, –, –; *Diervilla rivularis* Gatt., –, AJ429392, –, –, AJ431105, AJ430981, AJ429738, –, –; *Diervilla sessilifolia* Buckl., –, –, AF060164, Z29672, –, –, –, –, –; *Dipelta yunnanensis* Franch., –, AF446910, AF447030, AF446940, –, AF446970, –, –, –; *Dipsacus sativus* Garsault, AF209577, AJ429393,

AF130190, L13864, AJ431106, AJ430982, AJ429739, –, AF479231; *Dipsacus* sp.,