Molecular systematics and character evolution in Caryophyllaceae

Anne K. Greenberg & Michael J. Donoghue

Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208105, New Haven, Connecticut 06520, U.S.A.
Author for correspondence: *Anne K. Greenberg, anne.greenberg@yale.edu*

Abstract The aim of the present study was to infer a substantially larger, more evenly sampled, phylogenetic tree for Caryophyllaceae in order to more confidently resolve relationships within this clade. This would allow us to evaluate previous classification schemes and to infer the evolution of a number of characters that have figured prominently in higher-level taxonomic treatments. We have inferred a 630-tip phylogeny (ca. 30% of the 2200 species) using maximum likelihood analyses of data from the nuclear ribosomal ITS region and five chloroplast genes and intergenic spacers: *matK*, *ndhF*, *trnL-trnF*, *trnQ-rps16*, and *trnS-trnfM*. Our results confirm that subfamily Paronychioideae is paraphyletic at the base of Caryophyllaceae. Alsinoideae and Caryophylloideae together form a clade, within which neither subfamily is monophyletic. With only a few exceptions, our results support the tribal classification presented by Harbaugh & al. (2010). In agreement with other recent studies, it appears that many of the larger genera are not strictly monophyletic. Our results imply that the first Caryophyllaceae had stipules, free sepals, small apetalous flowers with few stamens, and single-seeded indehiscent or irregularly dehiscing utricles. Stipules were lost along the branch to the Alsinoideae-Caryophylloideae clade, and the evolution of a tubular calyx marks Caryophylloideae. The evolution of petals, 10 stamens, and capsule fruits is inferred to have taken place along the branch subtending a clade that includes Sperguleae (mostly containing former members of Paronychioideae) and the remainder of Caryophyllaceae. As this previously unnamed major group is both well-supported in molecular phylogenetic studies and marked by clear-cut apomorphies, we propose the name Plurcaryophyllaceae for this clade and provide a phylogenetic definition.

Keywords classification; molecular phylogeny; phylogenetic nomenclature; Plurcaryophyllaceae

Supplementary Material Figures S1–S8 and the Appendix are available in the free Electronic Supplement to the online version of this article (http://ingentaconnect.com/content/iapt/tax).

■ INTRODUCTION

Caryophyllaceae (Caryophyllales) form a clade of approximately 2200 species of herbs and subshrubs traditionally placed in some 70 to 86 genera (Bittrich, 1993b). Although it has a primarily Holarctic distribution with a center of diversity in the Mediterranean and Irano-Turanean regions, there are genera endemic to North America, South America, Africa, and Asia. Many species occupy dry, exposed habitats, but some are found in moist habitats such as temperate forests and meadows. At least seven genera have circumpolar distributions, several species grow at the very highest elevations occupied by plants (ca. 7000 meters), and *Colobanthus quitensis* is one of the very few angiosperms in Antarctica (Bittrich, 1993b). Despite their popularity in horticulture (e.g., carnations, *Dianthus*), and the importance of weedy species (e.g., the chickweed, *Stellaria media*), our knowledge of phylogenetic relationships in Caryophyllaceae is still limited.

The monophyly of Caryophyllaceae has been supported by molecular evidence (Downie & al., 1997; Cuenod & al., 2002; Brockington & al., 2009) and by two morphological synapomorphies: incomplete or complete reduction of the septa in the ovary and P III c'f sieve tube element plastids (Bittrich, 1993b). Traditionally, Caryophyllaceae are subdivided into three subfamilies: Alsinoideae, Caryophylloideae, and Paronychioideae. Alsinoideae are distinguished by nectar glands located at the abaxial base of the episepalous stamens (Bittrich, 1993b) and Caryophylloideae by a tubular calyx tube and jointed/clawed petals (Chrtek & Slavikova, 1987). Bittrich (1993b) suggested that Alsinoideae and Caryophylloideae together form a monophyletic group based on caryophyllad-type embryology, as compared with solanad embryology in Paronychioideae. Solanad, caryophyllad, and chenopodiad (which is the dominant type in the closely related Amaranthaceae; Townsend, 1993) embryologies are very similar, all beginning their development with a transverse division of the terminal cell in the two-celled embryo. In solanad and chenopodiad embryo development, the basal cell also undergoes a transverse division, while the basal cell in caryophyllad embryos does not divide (Raghavan, 2006). Chromosome number, which tends to be higher in Alsinoideae (base numbers $x = 6–19$) and Caryophylloideae (base numbers $x = 10–18$) than in Paronychioideae (base numbers $x = 8–9$; Bittrich, 1993b), may also support the monophyly of Alsinoideae plus Caryophylloideae. Chromosome numbers have also suggested to some authors that Caryophylloideae are derived from within Alsinoideae (Fernandes & Leitao, 1971; Bittrich, 1993b), as Caryophylloideae species tend to have higher base numbers. In addition to the probably plesiomorphic solanad embryo development, Paronychiodeae are characterized by the presence of stipules (Bittrich, 1993b). It has been noted that homoplasy in morphological characters within Caryophyllaceae has been problematical in circumscribing subfamilies (Bittrich, 1993b; Fior & al., 2006; Harbaugh & al., 2010), emphasizing the need for molecular data.

Molecular phylogenetic studies to date (Smissen & al., 2002; Fior & al., 2006; Frajman & al., 2009; Harbaugh & al., 2010) have suggested that the traditionally recognized subfamilies are non-monophyletic. In the most recent study, Harbaugh & al. (2010) found that the genera historically placed in Paronychioideae form a basal grade, while those traditionally classified as Alsinoideae and Caryophylloideae form a monophyletic group within which there are two major clades that do not correspond completely with the two traditional subfamilies. The monophyly of subfamily Caryophylloideae is not fully contradicted in any of these molecular studies, but support has not been uniformly strong (Fior & al., 2006; Harbaugh & al., 2010). Doubts about the monophyly of the traditional subfamilies led Harbaugh & al. (2010) to propose a new classification for the group. In their treatment, 11 major hypothesized clades were named as tribes.

The previous molecular phylogenetic studies of Caryophyllaceae have been limited in terms of the sample of species included. For example, Fior & al. (2006) included 40 species (~2% of the estimated total), while Harbaugh & al. (2010) sampled 126 species (~6%). More thorough sampling might increase confidence in relationships among the major lineages, and would certainly clarify the placement of many smaller genera that have not yet been sampled. The primary aim of the present study was to infer a substantially larger, more evenly sampled, phylogenetic tree for Caryophyllaceae. Although this sets the stage for studies of diversification, biogeography, and niche evolution, our intention in this paper is to resolve phylogenetic relationships with greater confidence within Caryophyllaceae, and to understand better the phylogenetic distribution of the characters that have figured prominently in taxonomic treatments and previous classification systems. Specifically, we have inferred a 630-tip phylogeny using data from the nuclear ribosomal ITS region and five chloroplast genes: *matK*, *ndhF*, *trnL-trnF*, *trnQ-rps16*, and *trnS-trnfM*. In the context of the resulting phylogenetic trees, we evaluate the classification of the group, focusing especially on the traditional subfamilies and the tribal classification system of Harbaugh & al. (2010). We also consider the evolution of several conspicuous floral characters, fruit type, and stipules.

■ MATERIALS AND METHODS

Samples. — Sequence data from 630 accessions were obtained from herbarium specimens and from GenBank (Table 1; July 2009). Sequences for 196 species previously unavailable in GenBank were added by this study, and a total of 1199 new sequences were generated. See the Appendix (Electronic Supplement) for voucher information and GenBank GI numbers of all specimens used in this study. The most widely used classification of Caryophyllaceae (Bittrich, 1993b) was taken into consideration when selecting taxa, which were chosen to proportionally cover all major groups within Caryophyllaceae. The final dataset included 615 Caryophyllaceae species, 14 Amaranthaceae species, and *Mollugo verticillata* (Molluginaceae). While Molluginaceae are traditionally used as an outgroup for Caryophyllaceae (Smissen & al., 2002), recent studies (Brockington & al., 2009; Schaferhoff & al., 2009) find that Amaranthaceae are more closely related to Caryophyllaceae. To infer character states at the base of Caryophyllaceae we rooted our trees along the *Mollugo verticillata* branch.

DNA extraction, PCR protocols, and sequencing. — Total DNAs were extracted using a modified Qiagen DNeasy Plant Mini Kit (Qiagen, La Jolla, California, U.S.A.) protocol (Wurdack & al., 2004). Double-stranded copies of all regions were amplified using standard Polymerase Chain Reaction (PCR) in 25-µl reactions. All reactions included an initial heating for 2 min at 80°C and 2 min at 94°C; then 5 cycles of 1 min denaturation at 94°C, 40 s annealing at 50°C, 3 min extension at 68°C; 35 cycles of 30 s denaturation at 94°C, 20 s annealing at 48°C, 1.5 min extension at 86°C; and a final extension for

Table 1. Taxa sampled in this study and two previous studies. Tribal classification from Bittrich (1993b).

Subfamily	Tribe	Total genera	Genera (species) sampled in Fior & al., 2006	Genera (species) sampled in Harbaugh & al., 2010	Genera (species) sampled in this study
	Polycarpeae	16	6 (7)	6 (8)	8 (16)
Paronychioideae	Paronychieae	15	7 (13)	4 (6)	11 (44)
	Corrigioleae	2	2 (2)	2 (3)	2 (3)
	Alsineae	28	9 (32)	18 (74)	18 (244)
	Pycnophylleae	1	1 (1)	1 (2)	1 (3)
Alsinoideae	Sclerantheae	2	1 (3)	1 (3)	1 (3)
	Habrosieae	1	0 (0)	0 (0)	1 (1)
	Geocarpeae	1	0 (0)	1 (1)	1 (1)
	Caryophylleae	17	7 (11)	8 (15)	10 (74)
Caryophylloideae	Drypideae	1	1 (1)	1 (1)	1 (1)
	Sileneae	8	4 (9)	4 (13)	8 (225)
TOTAL		92	38 (79)	46 (126)	62 (615)

10 min at 86°C. Amplifications were purified using polyethylene glycol (PEG) DNA precipitation. Dye-terminator cycle sequencing analyses were carried out using the ABI PRISM "BigDye" v.3.0 primer cycle sequencing ready reaction kit (ABI, Foster City, California, U.S.A.) and were analyzed using an ABI 3730 Genetic Analyzer (ABI). New sequences were combined with those downloaded from GenBank.

A total of six gene regions were amplified for this study: nrITS, *matK*, *ndhF*, *trnL-trnF*, *trnQ-rps16*, and *trnS-trnfM*. Primers for ITS (Schmidt & Schilling, 2000), *trnL-trnF* (Taberlet & al., 1991), *trnQ-rps16* (Shaw & al., 2007), and *trnS-trnfM* (Shaw & al., 2005) are universal primers. Primers for *matK* and *ndhF* were those used in previous studies of Caryophyllaceae (Smissen & al., 2002; Fior & al., 2006).

Sequence alignment and phylogenetic analysis. — Contiguous sequences were assembled using Sequencher v.4.7 (Gene Codes Corp., Madison, Wisconsin, U.S.A.). Groups of closely related sequences were aligned with PRANK, a phylogeny-aware gap placement alignment program (Loytynoja & Goldman, 2008). Multiple sequence alignments were then aligned using MUSCLE profile-profile alignment (Edgar, 2004). This was especially useful in improving ITS alignments. Alignment errors were identified and corrected with the pairwise alignment tool in the Align package in Mesquite (Maddison & Maddison, 2010). The resulting data matrices (and the trees presented here) are available in TreeBase (www.treebase.org) or upon request from the first author.

The data were analyzed using maximum likelihood (ML) methods. Tree searches for ML analyses were executed under the GTRMIX nucleotide substitution model in RAxML v.7.2.2 (Lemmon & Moriarty, 2003; Alfaro & Huelsenbeck, 2006; Stamatakis, 2006a, b; Stamatakis & al., 2008; Ripplinger & Sullivan, 2008). This model first uses the GTRCAT model to perform tree inferences. GTRCAT is a nucleotide substitution model that uses a general time reversible approximation and optimizes individual per-site substitution rates in a fixed number of rate categories (CAT). The CAT model is empirically equivalent to the GAMMA model, but has faster inference times and requires significantly less memory. However, alternative tree topologies cannot be compared using CAT-based likelihood values (Stamatakis, 2006a). Therefore, GTRMIX switches to a GAMMA (general time reversible substitution model with a gamma model of rate heterogeneity) model to evaluate final tree topologies and obtain stable likelihood values. The data were partitioned by gene for the ML analyses (a total of six partitions), and nonparametric bootstrapping under ML was carried out on the dataset using 100 bootstrap replicates. Trees were inferred from the ITS dataset (25.7% missing data), the chloroplast dataset (68.2% missing data), the full combined dataset (73.5% missing data), and a reduced combined dataset that included only species with both ITS and chloroplast data (58.5% missing data).

Morphological characters and ancestral state reconstructions. — Data were collected for seven morphological characters: (1) fruit type (capsule or utricle), (2) petal presence/absence, (3) tubular calyx presence/absence, (4) stipule presence/absence, (5) fruit dehiscence, (6) stamen number, and (7) maximum seed number. Morphological data were obtained from floras, species keys, and herbarium specimens (Bittrich 1993a, b; Townsend, 1993; Lu & al., 2001; Rabeler & Hartman, 2005).

Characters 1–4 are discrete, binary characters and ancestral states for all nodes excluding the root of the entire tree (*Mollugo verticillata*) were reconstructed with maximum likelihood methods using LASRDisc v.1.0 (Likelihood Ancestral State Reconstruction for Discrete Characters; Jackson, 2004). Character 5 is a multi-state character with three discrete states: dehiscent fruit, indehiscent fruit, and fruit that ruptures irregularly. Character 6 has six discrete states: 1, 3, 4, 5, 8, or 10 stamens. However, because 1, 3, 4, or 8 stamens are rare, Fig. 5D is simplified to show only two character states: 5 (rarely 1, 3, or 4) stamens, and 10 (rarely 8) stamens. Species with 1, 3, 4, or 8 stamens are listed in the figure caption. In addition, we note that several shifts between the 5- and 10-stamen states were initially inferred to have occurred via a single-stamen intermediate. These results were further investigated in Mesquite and, as expected, the single stamen state was found to be unlikely at these nodes. Figure 5D reflects the Mesquite analyses. Character 7, seed number, is also a multi-state discrete character, but potentially has a very large number of states. However, based on an initial evaluation of the distribution of seed numbers, we recognized four categories: 1, 2–50, 51–99, and 100+ seeds. Maximum values were generally used for characters 6 and 7. However, if a maximum stated value was designated as "rare", the highest common value was used. Ancestral states for characters 5–7 were reconstructed for all nodes excluding the root of the entire tree using maximum likelihood methods in BayesTraits (Pagel, 1999).

Ancestral states at nodes along the backbone of the tree were further tested by comparing likelihood scores to examine whether a particular character state was significantly (>2 lnL; Edwards, 1992) more likely. The "fossil" command in BayesTraits (Pagel, 1999) was used to fix the state at the node of interest. This was repeated for each character state, which allowed us to obtain the likelihood associated with each alternative state. Character states with the highest significantly different likelihood were considered the most likely ancestral state.

■ RESULTS

The final combined dataset included 630 species, 196 of which are newly sequenced. A single Molluginaceae species (*Mollugo verticillata*) was included in the dataset for rooting purposes, along with 14 species of Amaranthaceae, and 615 species of Caryophyllaceae. The combined data matrix consisted of 6547 aligned nucleotides, 5165 of which were variable (~79%). The most variable marker was *trnQ-rps16*, followed by ITS, *matK*, *trnL-trnF*, *trnS-trnfM*, and *ndhF* (Table 2). The ITS data matrix consisted of 706 aligned nucleotides, 638 (~90%) of which were variable. The combined chloroplast data matrix consisted of 5841 aligned nucleotides, 4527 (~78%) of which were variable.

Table 2. Statistics from data matrices used for maximum likelihood analyses.

	ITS	*matK*	*ndhF*	*trnL-trnF*	*trnQ-rps16*	*trnS-trnfM*	Chloroplast	Combined
No. of sequences	472	374	86	223	146	135	454	630
Aligned length [bp]	706	1690	1222	1055	841	1033	5841	6547
No. of variable sites	638	1486	574	921	774	772	4527	5165
% variable	90.37	87.93	46.97	87.30	92.03	74.73	77.50	78.89

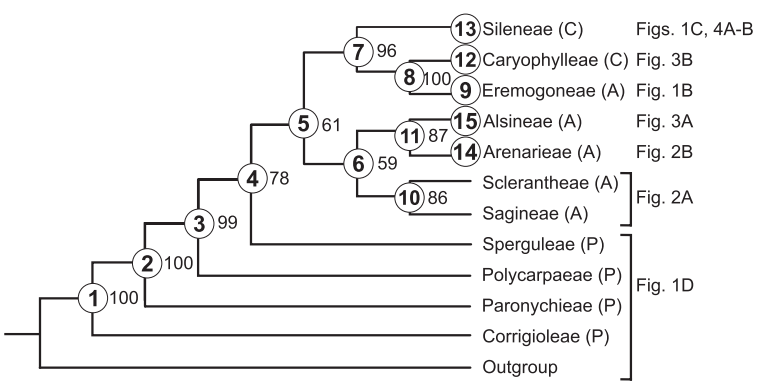
Major relationships within Caryophyllaceae. — Our phylogenetic results are largely consistent with other recent analyses (Figs. 1–4; Fior & al., 2006; Harbaugh & al., 2010), although the sample of species was greatly increased (Table 1). Caryophyllaceae are supported as a clade (100% BS; node 1, Fig. 1A, D). The first divergence within Caryophyllaceae separates a clade comprised of *Telephium* and *Corrigiola* (the Corrigioleae of Harbaugh & al., 2010) from the rest of Caryophyllaceae (100% BS; node 2, Fig. 1A, D). The first divergence within the larger clade (node 2, Fig. 1A, D) divides a clade of Paronychioideae species (the Paronychieae of Harbaugh & al., 2010; including all *Paronychia*) from the rest of Caryophyllaceae (99% BS; node 3, Fig. 1A, D; Harbaugh & al., 2010). The first divergence within node 3 separates another clade of Paronychioideae (designated as tribe Polycarpaeae by Harbaugh & al., 2010) from the rest of Caryophyllaceae (78% BS; node 4, Fig. 1A, D), which diverges into the final clade of Paronychiodeae (designated as tribe Sperguleae by Harbaugh & al., 2010) and the rest of Caryophyllaceae (61% BS; node 5, Fig. 1A). The genera present in tribes Paronychieae, Polycarpaeae, and Sperguleae are largely consistent with those in Harbaugh & al. (2010), although additional small genera have been added to each tribe in our analyses.

The large remaining Caryophyllaceae clade (61% BS; node 5, Fig. 1A) comprises all members of subfamilies Alsinoideae and Caryophylloideae, and is split into two large clades (59% BS and 96% BS, respectively; nodes 6 and 7, Fig. 1A). The majority of Alsinoideae form a clade at node 6, while node 7 is comprised of two large clades (100% BS and 96% BS, respectively; nodes 8 and 13, Fig. 1A, C) containing all Caryophylloideae species (other than *Drypis spinosa*) as well as a small clade (100% BS; node 9, Fig. 1B) traditionally assigned to Alsinoideae, which contains species of *Arenaria* and *Minuartia* (tribe Eremogone of Harbaugh & al., 2010). The clade at node 6 is comprised of two clades (nodes 10 and 11, Fig. 1A). The clade at node 10 (86% BS, Figs. 1A, 2A) comprises two named tribes: Scleranthea and Sagineae, both of which are also clades in our analysis. As in Harbaugh & al. (2010), Scleranthea (100% BS) includes the Hawaiian genus *Schiedea* as well as *Geocarpon*, *Honckenya*, *Scleranthus*, and *Wilhelmsia*. Our analyses also include a single *Spergularia* species (*Spergularia heldreichii*) and four *Stellaria* species in this clade. Sagineae (84% BS) includes *Bufonia*, *Colobanthus*, *Sagina*, and *Drypis spinosa*, which is also consistent with the findings of Harbaugh & al. (2010). *Drypis* is a monotypic genus usually included in subfamily Caryophylloideae (originally placed in the monotypic tribe Drypideae). Our analyses also include *Habrosia* and *Arenaria fontinalis* in Sagineae. As in Harbaugh & al. (2010), Scleranthea and Sagineae also include the majority of *Minuartia* species. The clade at node 11 (87% BS, Fig. 1A) is comprised of two clades that have been recognized as tribes. Alsineae (95% BS, Fig. 3A) contains *Stellaria*, *Cerastium*, and some *Arenaria* species, and many smaller genera including *Holosteum*, *Lepyrodiclis*, *Moenchia*, *Myosoton*, *Plettkea*, and *Pseudostellaria*. Arenarieae (99% BS, Fig. 2B) contains *Moehringia* and the majority of *Arenaria* species.

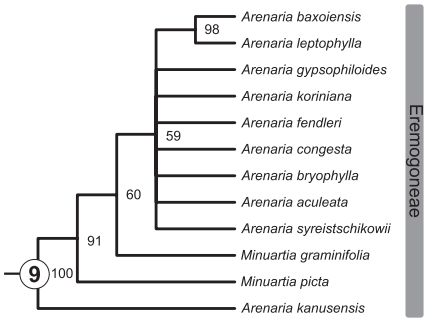
The clade at node 8 (Fig. 1A) is also split into two clades (nodes 9 and 12, Fig. 1A). The clade at node 9 (100% BS, Fig. 1B) corresponds to tribe Eremogone, and is traditionally included within subfamily Alsinoideae. The clade at node 12 (100% BS, Fig, 3B) corresponds to tribe Caryophylleae, which includes *Gypsophila* and *Dianthus*, along with smaller genera such as *Petrorhagia*, *Saponaria*, *Allochrusa* and *Vaccaria*. Our analyses also include *Psammosilene* and *Bolbosaponaria* in this clade. The clade at node 13 (96% BS, Figs. 1C, 4A–B), corresponding to tribe Sileneae (Harbaugh & al., 2010), is comprised of all *Silene* species as well as smaller genera such as *Agrostemma*, *Lychnis*, and *Petrocoptis*.

Although comparison is difficult owing to lack of resolution and differences among the datasets, tree topologies for the five chloroplast markers are generally congruent with one another, with ITS, and with the combined dataset (Figs. S1–S6 in the Electronic Supplement). In all six gene trees, Sileneae and Caryophylleae each form clades. In gene trees where the backbone is sufficiently resolved (*trnS-trnfM*, *trnQ-rps16*, *ndhF*, *matK*), these clades are closely related to one another and to Eremogoneae. While the placement of Eremogoneae as sister to Caryophylleae is consistent in the combined dataset, the combined chloroplast tree, ITS, and *trnS-trnfM*, Eremogoneae are sister to Sileneae in the *matK* dataset. Other chloroplast markers do not directly address the placement of Eremogoneae, either due to lack of resolution (*trnL-trnF*) or the absence of relevant species. Alsineae, Arenarieae, Scleranthea, and Sagineae form clades in all six gene trees, but the placement of

Fig. 1. Maximum likelihood (ML) bootstrap consensus phylogeny of Caryophyllaceae based on analyses of nrITS, *matK*, *ndhF*, *trnL-trnF*, *trnQ-rps16*, and *trnS-trnfM* sequences. Un-circled numbers are ML bootstrap (BS) values. Circled numbers mark major nodes discussed in the text. Collapsed branches signify BS values <50. **A,** Backbone phylogeny showing the position of each tribe of Harbaugh & al. (2010). Letters in parentheses show the assignment of each tribe to a traditionally recognized subfamily (A, Alsinoideae; C, Caryophylloideae; P, Paronychioideae). This tree points to figures that show detailed relationships within each clade.


13 Sileneae (C) Figs. 1C, 4A-B
12 Caryophylleae (C) Fig. 3B
9 Eremogoneae (A) Fig. 1B
15 Alsineae (A) Fig. 3A
14 Arenarieae (A) Fig. 2B
Scleranthaeae (A)
Sagineae (A) Fig. 2A
Sperguleae (P)
Polycarpaeae (P)
Paronychieae (P) Fig. 1D
Corrigioleae (P)
Outgroup

A.


Arenaria baxoiensis
Arenaria leptophylla
Arenaria gypsophiloides
Arenaria koriniana
Arenaria fendleri
Arenaria congesta
Arenaria bryophylla
Arenaria aculeata
Arenaria syreistschikowii
Minuartia graminifolia
Minuartia picta
Arenaria kanusensis
Eremogoneae

B.

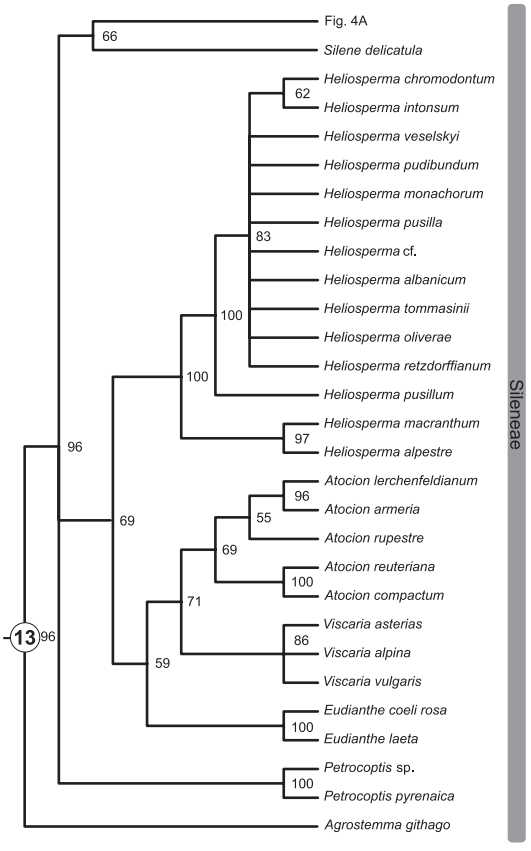
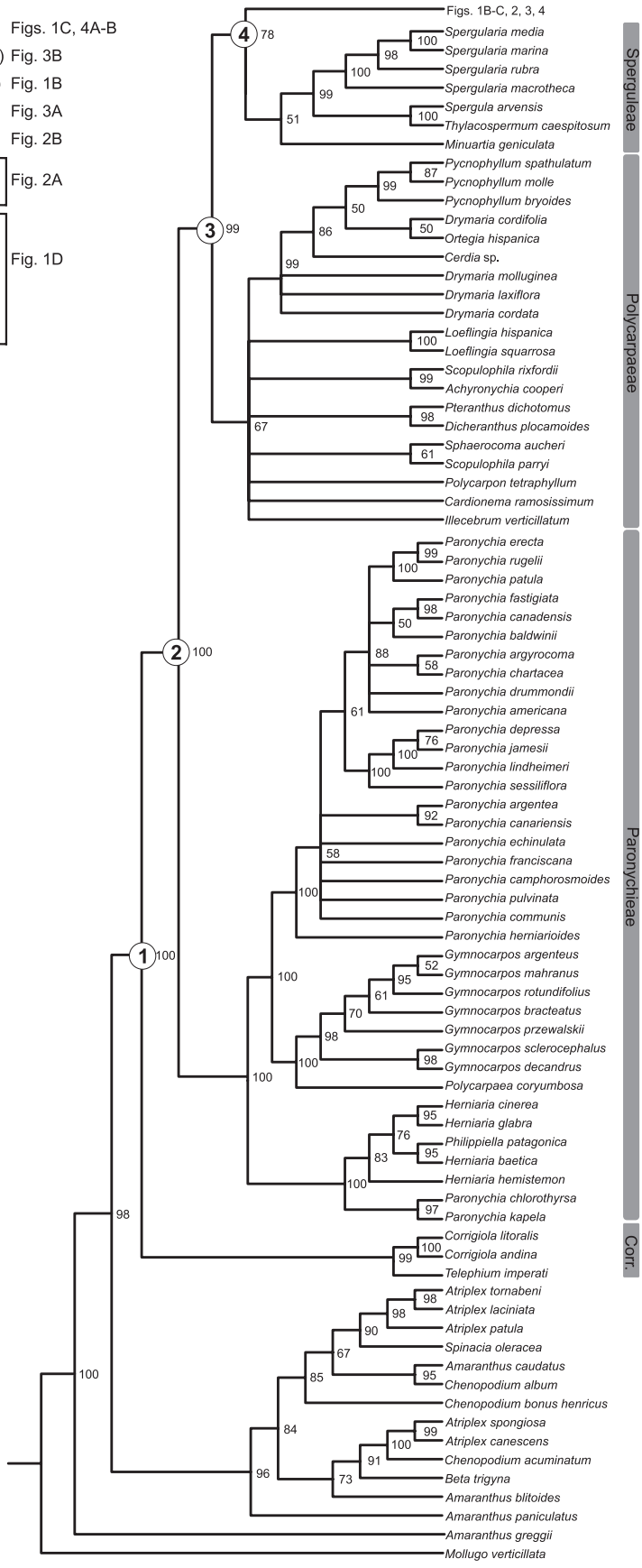


Fig. 4A
Silene delicatula
Heliosperma chromodontum
Heliosperma intonsum
Heliosperma veselskyi
Heliosperma pudibundum
Heliosperma monachorum
Heliosperma pusilla
Heliosperma cf.
Heliosperma albanicum
Heliosperma tommasinii
Heliosperma oliverae
Heliosperma retzdorffianum
Heliosperma pusillum
Heliosperma macranthum
Heliosperma alpestre
Atocion lerchenfeldianum
Atocion armeria
Atocion rupestre
Atocion reuteriana
Atocion compactum
Viscaria asterias
Viscaria alpina
Viscaria vulgaris
Eudianthe coeli rosa
Eudianthe laeta
Petrocoptis sp.
Petrocoptis pyrenaica
Agrostemma githago
Sileneae

C.


Figs. 1B-C, 2, 3, 4
Spergularia media
Spergularia marina
Spergularia rubra
Spergularia macrotheca
Spergula arvensis
Thylacospermum caespitosum
Minuartia geniculata
Sperguleae
Pycnophyllum spathulatum
Pycnophyllum molle
Pycnophyllum bryoides
Drymaria cordifolia
Ortegia hispanica
Cerdia sp.
Drymaria molluginea
Drymaria laxiflora
Drymaria cordata
Loeflingia hispanica
Loeflingia squarrosa
Scopulophila rixfordii
Achyronychia cooperi
Pteranthus dichotomus
Dicheranthus plocamoides
Sphaerocoma aucheri
Scopulophila parryi
Polycarpon tetraphyllum
Cardionema ramosissimum
Illecebrum verticillatum
Polycarpaeae
Paronychia erecta
Paronychia rugelii
Paronychia patula
Paronychia fastigiata
Paronychia canadensis
Paronychia baldwinii
Paronychia argyrocoma
Paronychia chartacea
Paronychia drummondii
Paronychia americana
Paronychia depressa
Paronychia jamesii
Paronychia lindheimeri
Paronychia sessiliflora
Paronychia argentea
Paronychia canariensis
Paronychia echinulata
Paronychia franciscana
Paronychia camphorosmoides
Paronychia pulvinata
Paronychia communis
Paronychia herniarioides
Gymnocarpos argenteus
Gymnocarpos mahranus
Gymnocarpos rotundifolius
Gymnocarpos bracteatus
Gymnocarpos przewalskii
Gymnocarpos sclerocephalus
Gymnocarpos decandrus
Polycarpaea coryumbosa
Herniaria cinerea
Herniaria glabra
Philippiella patagonica
Herniaria baetica
Herniaria hemistemon
Paronychia chlorothyrsa
Paronychia kapela
Paronychieae
Corrigiola litoralis
Corrigiola andina
Telephium imperati
Corr.
Atriplex tornabeni
Atriplex laciniata
Atriplex patula
Spinacia oleracea
Amaranthus caudatus
Chenopodium album
Chenopodium bonus henricus
Atriplex spongiosa
Atriplex canescens
Chenopodium acuminatum
Beta trigyna
Amaranthus blitoides
Amaranthus paniculatus
Amaranthus greggii
Mollugo verticillata

D.

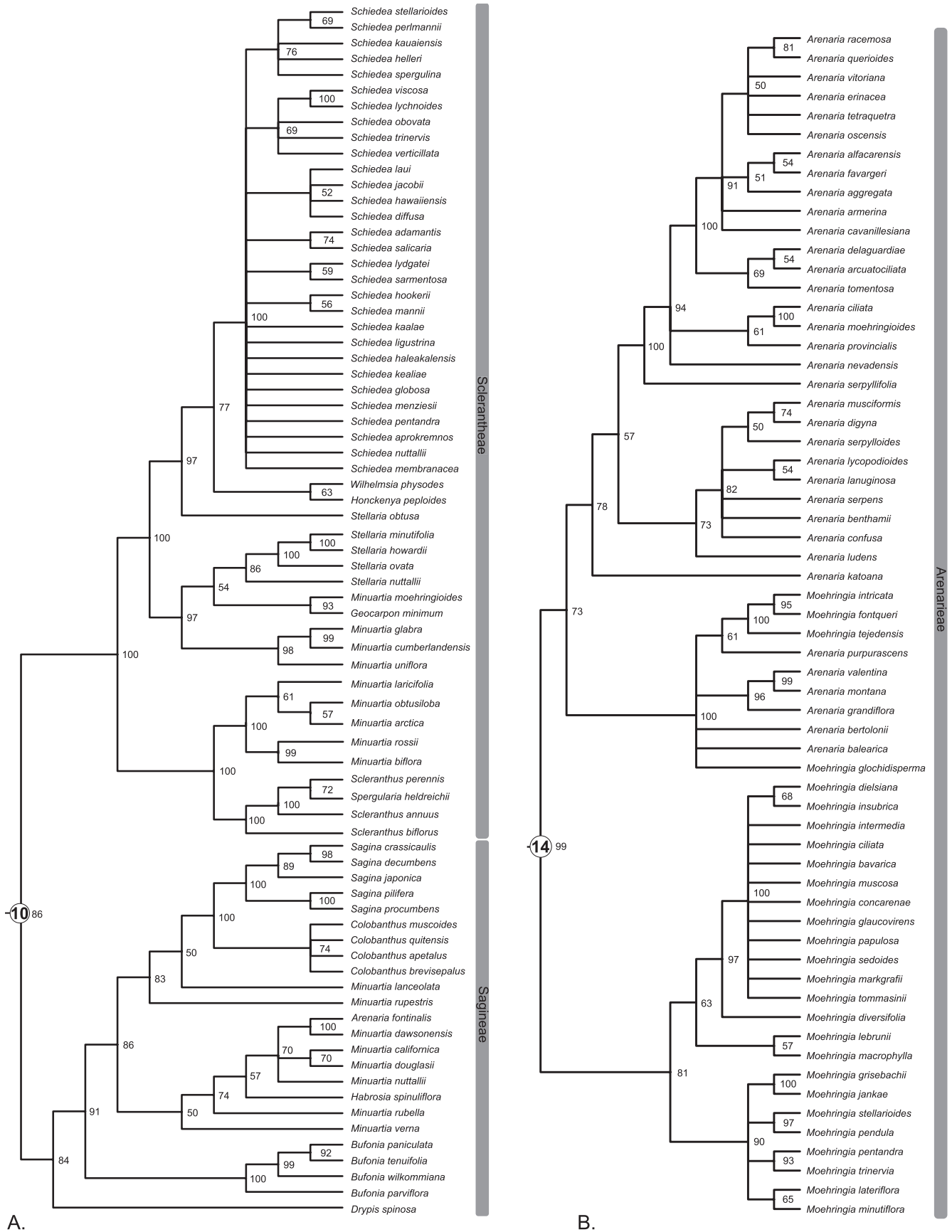


Fig. 2. Maximum likelihood (ML) bootstrap consensus phylogeny of Caryophyllaceae continued from Fig. 1.

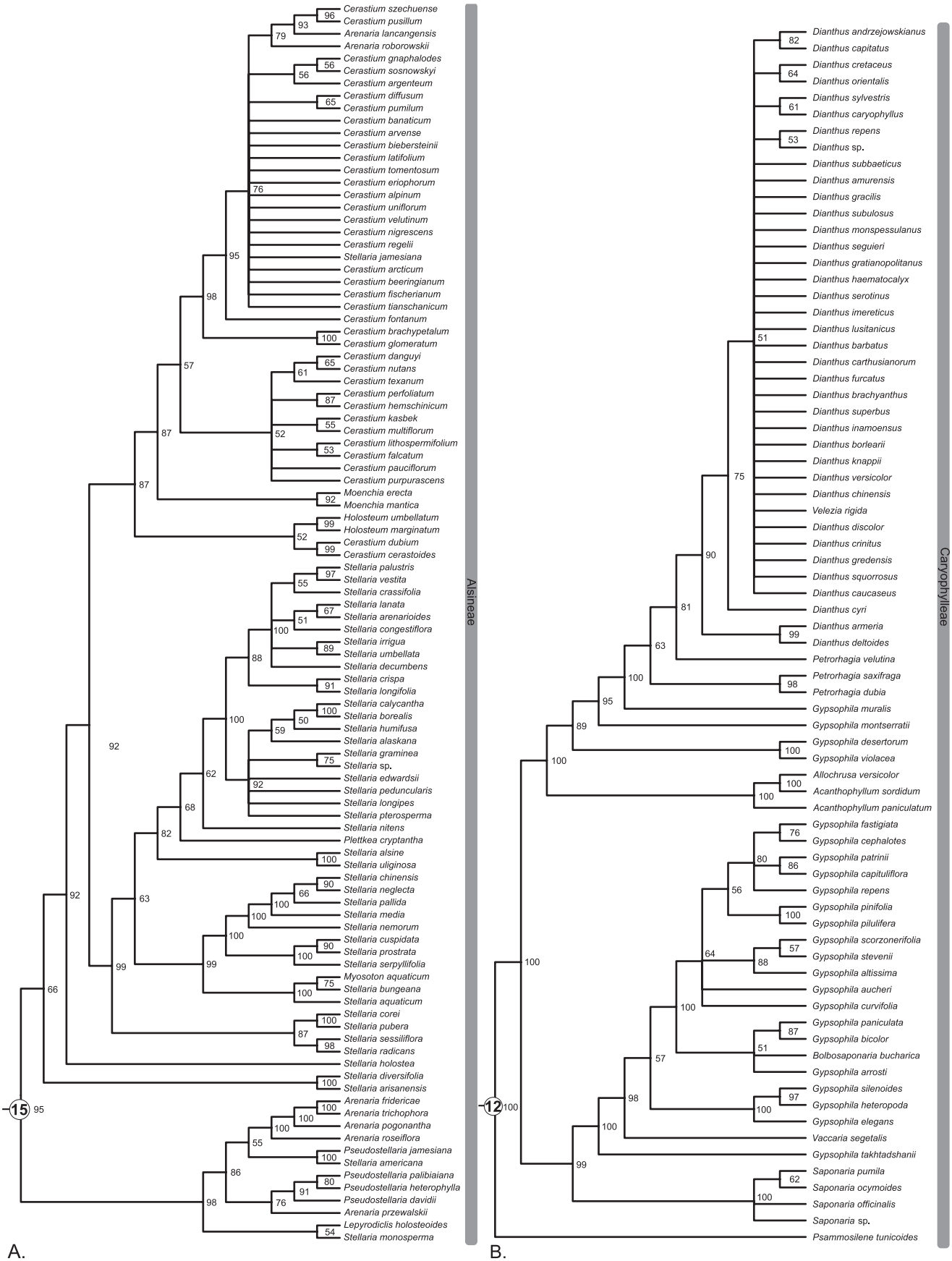


Fig. 3. Maximum likelihood (ML) bootstrap consensus phylogeny of Caryophyllaceae continued from Fig. 1.

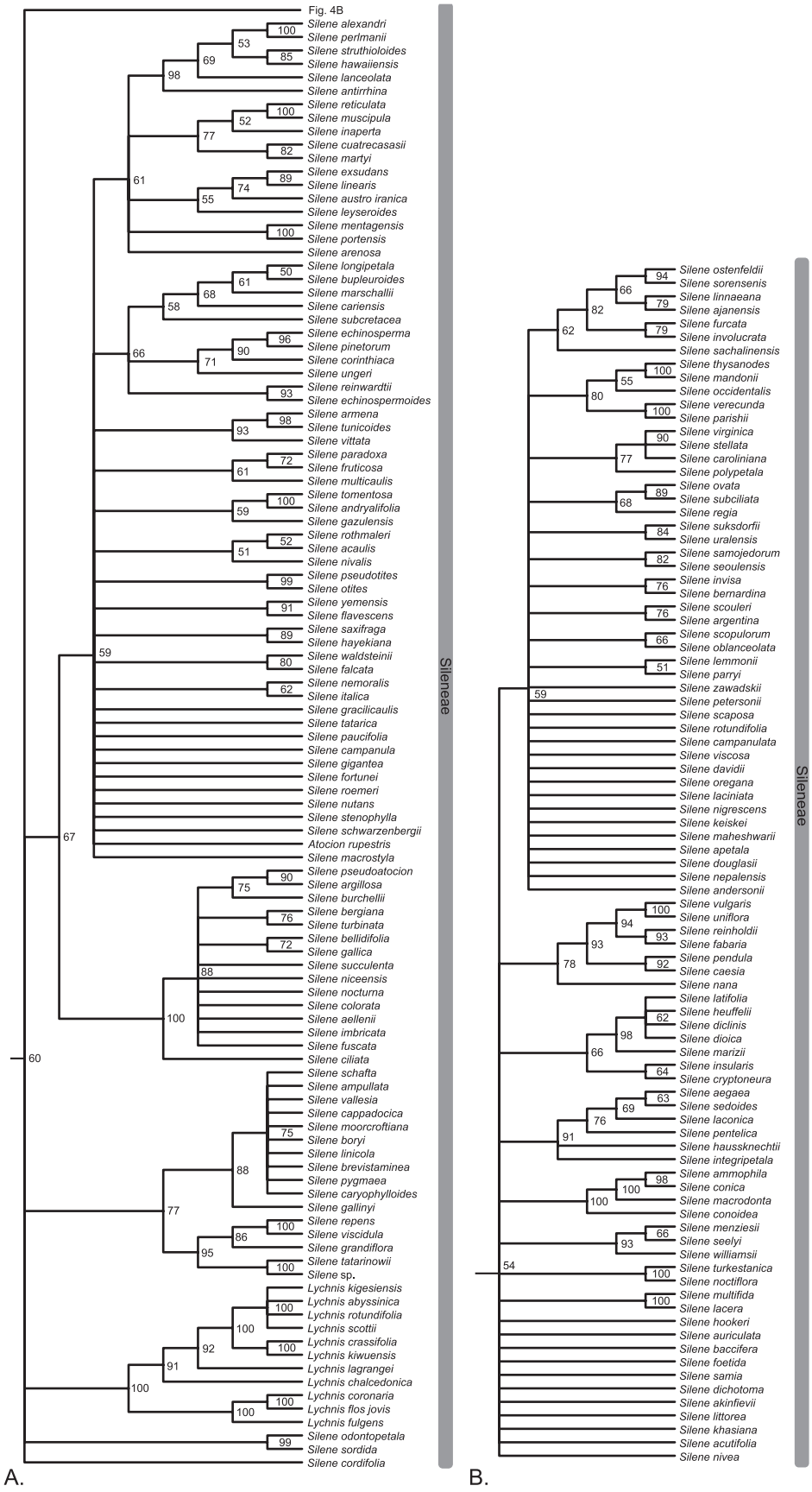


Fig. 4. Maximum likelihood (ML) bootstrap consensus phylogeny of Caryophyllaceae continued from Fig. 1.

these clades tends to be unresolved. However, in *trnS-trnfM* and *matK*, Alsineae and Arenarieae are sister to a clade consisting of Sileneae and Caryophylleae, and Scleranthеae and Sagineae are sister to the clade containing all four of the proceeding clades. In ITS, the clade containing Scleranthеae and Sagineae is sister to Alsineae, and Arenarieae is a separate clade. Paronychieae are also resolved as a clade in all six gene trees. Resolution of backbone nodes is greatly reduced in the ITS tree (Fig. S6), and somewhat reduced in the chloroplast tree (Fig. S8), as compared to the full combined dataset. The reduced combined dataset was largely congruent with the full combined dataset (Fig. S7). Beyond support for small groups of species, relationships within the major numbered lineages are generally poorly resolved (Figs. 1–4). Additional molecular data, especially faster evolving markers, would likely increase within-lineage resolution. However, it is possible that hybridization is also contributing to poor resolution, as it has been detected in multiple Caryophyllaceae lineages (Chinnappa & al., 2005; Popp & Oxelman, 2007; Erixon & Oxelman, 2008; Rautenberg & al., 2008; Frajman & al., 2009; Balao & al., 2010).

Ancestral state reconstructions. — While ancestral states were reconstructed for all nodes in the tree (Figs. 5–6), we paid special attention to the numbered nodes in Fig. 1A. The earliest diverging nodes within Caryophyllaceae are inferred to have had stipulate leaves, apetalous flowers, few stamens, free sepals (no tubular calyx; Fig. 5), and single-seeded, indehiscent or irregularly rupturing utricles (Fig. 6). These character states persist through the backbone nodes with branches leading to the several clades containing species traditionally assigned to Paronychioideae. Node 5, which includes Alsinoideae and Caryophylloideae species is reconstructed as having exstipulate leaves, petals, many stamens, and multi-seeded, dehiscent capsules. The node at which the shift from free sepals to a tubular calyx occurs is less clear, as Eremogoneae (sister to Caryophylleae) have free sepals while Caryophylleae and Sileneae have a tubular calyx. It is possible that this shift occurred twice, once on the branch leading to Sileneae (node 13, Fig. 1A), and once on the branch leading to Caryophylleae (node 12, Fig. 1A). However, it is equally likely that this character was gained only once, on the branch leading to the Eremogoneae/Caryophylleae/Sileneae clade (node 7, Fig. 1A), and then lost in Eremogoneae. Likelihood scores at these nodes are virtually equal for both character states. Shifts in fruit type, dehiscence, petal presence/absence, and the number of stamens are inferred to have occurred between nodes 3 and 4 (Figs. 5–7), while the shift from stipulate to exstipulate leaves is inferred between nodes 4 and 5, immediately subtending the large Alsinoideae/Caryophylloideae clade (Figs. 5, 7).

Comparison of likelihood scores at each of the numbered nodes generally supports these results. The highest log likelihood value for seed number at the root of Caryophyllaceae (node 1) indicates that the earliest members of this clade were single-seeded (>3 lnL). The log likelihood value for the fruit type at the root indicates that a utricle is over 6 lnL higher than a capsule, and the log likelihood value for apetalous flowers is 5 lnL higher than that of having petals. Log likelihood values at the root for the presence of stipules and the absence of the tubular calyx are over 12 lnL higher than the alternate states. The log likelihood values for indehiscent and irregularly rupturing fruits at the root are not significantly different from one another, but they are both almost 5 lnL higher than dehiscence. Log likelihoods for stamen number at the root are more ambiguous. Log likelihood values for 3, 4, or 5 stamens are not significantly different from one another, but these values do differ significantly from higher stamen numbers.

Comparisons of likelihood scores at all other numbered nodes were also performed and support the ancestral state reconstructions described above. We paid special attention to the base of the large Alsinoideae/Caryophylloideae clade (at node 5), the base of the majority of Alsinoideae (at node 6), and the base of the Eremogoneae/Caryophylleae/Sileneae clade (at node 7). Our results indicate that the first Alsinoideae/Caryophylloideae had petals, 10 stamens, and dehiscent capsules with between 2 and 50 seeds. The ancestral species at this node lacked stipules and a tubular calyx. Ancestral Alsinoideae were similar, but our comparison of likelihood scores at this node cannot reject the hypothesis that ancestral Alsinoideae had fewer stamens. As discussed above, the tubular calyx is less clear. It may have evolved on the branch subtending the Eremogoneae/Caryophylleae/Sileneae clade, or independently on branches leading to Sileneae and Caryophylleae.

■ DISCUSSION

Taxon sampling. — Algorithmic developments and computational advances have encouraged the development of phylogenies with increasingly large numbers of tips (e.g., Sanderson & al., 2008; Stamatakis & al., 2008; Goloboff & al., 2009; Smith & al., 2009). In turn, larger phylogenies have increased our ability to elucidate broad-scale evolutionary patterns (e.g., Moles & al., 2005; Smith & Donoghue, 2008; Smith & Beaulieu, 2009; Edwards & Smith, 2010; Goldberg & al., 2010). However, large size often does not equate with dense or proportional taxon sampling (but see Pelser & al., 2007; Thomson & Shaffer, 2010). Indeed, the percentage of species covered by phylogenetic studies remains rather small for most angiosperm groups of substantial size. At the present time, for example, less than 20% of the species belonging to many major angiosperm clades are represented by any sequences in GenBank (Smith & al., 2011). Here, we have increased the density of sampling in Caryophyllaceae from ca. 6% in the latest molecular phylogenetic study (Harbaugh & al., 2010) to approximately 30% of the estimated 2200 species of Caryophyllaceae, which has allowed us to include not only representatives of the major genera, but also most of the smaller ones that have been recognized. While we have largely confirmed recent results, our data increase support for deep nodes within Caryophyllaceae and suggest a number of new relationships within and among the major lineages (see below). We hope that our attempt to proportionally sample Caryophyllaceae taxonomy will provide a better basis for classification, but also for downstream comparative analyses of all kinds (e.g., of character evolution, diversification rates, biogeography).

Phylogenetic relationships. — While Caryophyllaceae have traditionally been divided into three large subfamilies—Alsinoideae, Caryophylloideae, and Paronychioideae (Chrtek & Slavikova 1987; Bittrich 1993b)—on the basis of morphological characters, it has been unclear to what extent these are supported or contradicted by molecular evidence (Smissen & al., 2002; Fior & al., 2006; Frajman & al., 2009; Harbaugh & al., 2010). However, based on evidence of the non-monophyly of at least the Paronychioideae, Harbaugh & al. (2010) put forward an alternative tribal classification for the group.

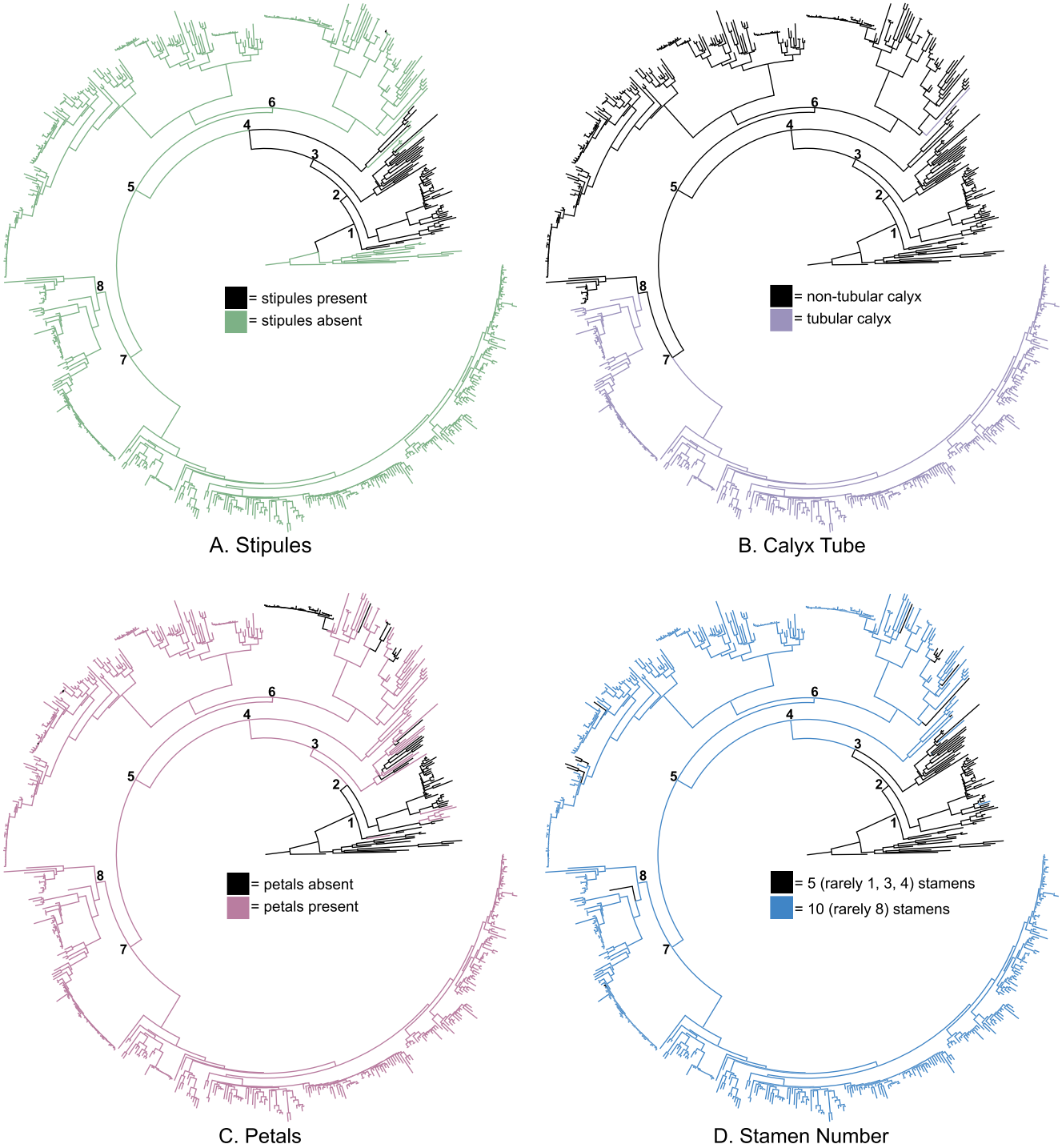


Fig. 5. Ancestral state reconstructions of morphological characters discussed in the text. Numbers represent the major nodes discussed in the text (see Fig. 1A). **A,** Stipules; **B,** calyx tube; **C,** petals; **D,** stamen number. While the analysis for stamen number included additional rare character states (1, 3, 4, and 8 stamens), the figure was simplified to include only 5 and 10 stamens. The following species/genera have one stamen: *Cerdia*; three stamens: *Dicheranthus plocamoides*; four stamens: *Pteranthus dichotomus*, *Philippiella patagonica*, *Moenchia erecta*, and *Moenchia mantica*; and eight stamens: *Bufonia*.

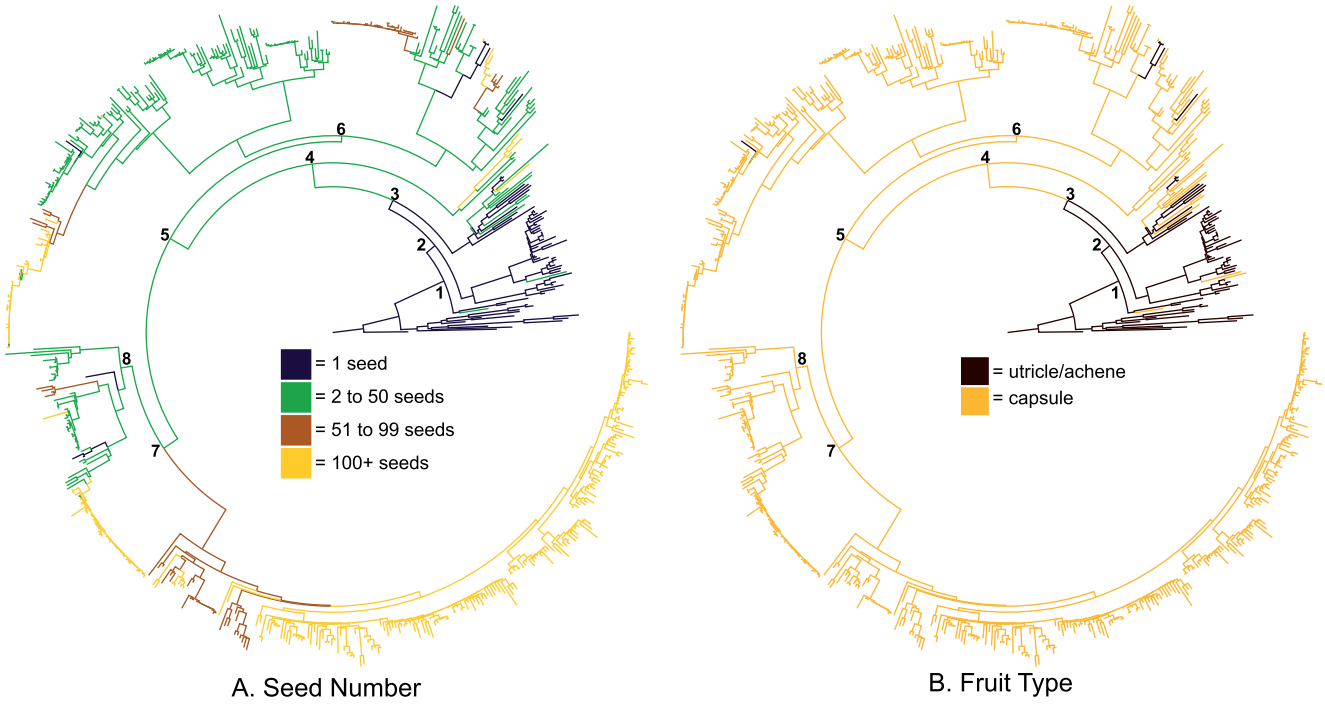


Fig. 6. Ancestral state reconstructions of morphological characters discussed in the text. Numbers represent major nodes discussed in the text (see Fig. 1A). **A,** Seed number; **B,** fruit type.

Our results do not support the monophyly of any of the three traditional subfamilies within Caryophyllaceae. However, our results are strongly consistent with those of Harbaugh & al. (2010).

Our phylogenies (Figs. 1–4) suggest that: (1) subfamily Paronychioideae is a non-monophyletic grade of early diverging lineages; (2) subfamilies Alsinoideae and Caryophylloideae together form a clade; (3) Caryophylloideae is non-monophyletic as it includes a clade traditionally placed within Alsinoideae (Eremogoneae; Harbaugh & al., 2010); (4) most of the members of subfamily Alsinoideae form a clade, but Eremogoneae is included within Caryophylloideae, and *Thylacospermum caespotosum*, *Minuartia geniculata*, and a small clade of *Pycnophyllum*, are placed within the early diverging paronychiod lineages; and (5) none of the eight largest genera (*Arenaria*, *Cerastium*, *Dianthus*, *Gypsophila*, *Minuartia*, *Paronychia*, *Silene*, *Stellaria*) appear to be strictly monophyletic.

Our results are generally consistent with the tribal classification of Harbaugh & al. (2010), although direct comparisons are difficult owing to our addition of many taxa. Also, because phylogenetic definitions (see Cantino & al., 2007) have not yet been provided for these taxa, we are unable to precisely specify their boundaries. Nevertheless, with only a few exceptions, our tree supports the monophyly of all of the tribes recognized by Harbaugh & al. (2010).

Eremogoneae, a small clade (100% BS) made up of *Arenaria* subg. *Eremogone* and subg. *Eremogoneastrum*, and *Minuartia* subg. *Spergella*, is found in our analyses within Caryophylloideae, as sister to Caryophylleae (containing *Dianthus* and *Gyposphila*). This renders both the traditional Alsinoideae and Caryophyllodieae non-monophyletic. Consistent results were obtained by Fior & al. (2006) and by Harbaugh & al. (2010). Indeed, Harbaugh & al. (2010) suggested based on their phylogenetic results, as well as on morphological features such as grass-like leaves, that the *Arenaria* species in this clade be included in a new tribe, Eremogoneae, and that *Phlebanthia*

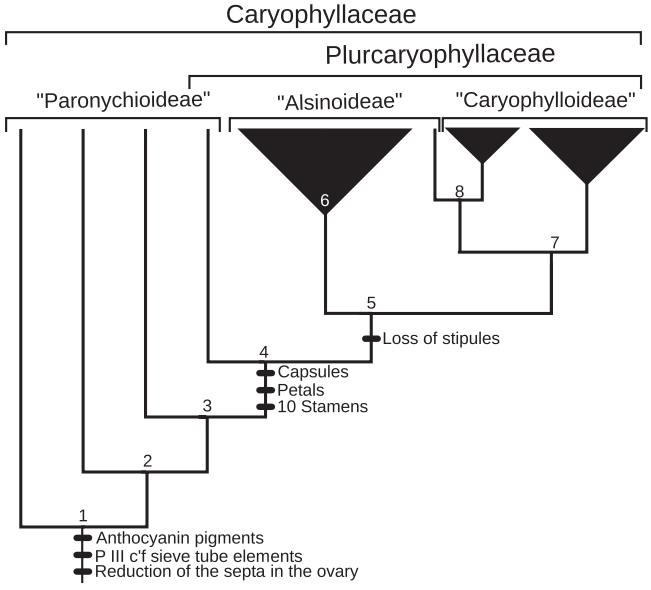


Fig. 7. A synthesis of major phylogenetic results and ancestral state reconstructions, showing the traditionally recognized subfamilies and the newly named Plurcaryophyllaceae.

be re-established as a genus to include the *Minuartia* species in this clade. Species in this clade have Northern Hemisphere distributions, with disjunctions between eastern Asia, western North America, and Europe and the Middle East.

For the Alsineae, Arenarieae, and Corrigioleae of Harbaugh & al. (2010), we recover clades that correspond directly in terms of the genera that are represented. As a result of the increased sampling in our analysis, additional species and genera are seen to be added within clades that otherwise correspond to Caryophylleae, Paronychieae, Sagineae, Sclerantheae, Polycarpaeae, and Sileneae. *Bolbosaponaria* and *Psammosilene* are included in Caryophylleae, while *Polycarpaea* and *Philippiella* are included in Paronychieae. *Habrosia* and *Drypis* are included in Sagineae along with a single species of *Arenaria* (*A. fontinalis*). *Drypis*, found in the eastern Mediterranean, has traditionally been placed in Caryophylloideae based on the presence of a tubular calyx. The base chromosome number in *Drypis* ($x = 15$) is also consistent with its placement within Caryophylloideae. However, it is noteworthy that it has usually been placed in a monogeneric tribe (Drypideae) based on its indehiscent fruits, spiny leaves, and zygomorphic flowers. Our analysis, along with previous analyses (Oxelman & Liden, 1995; Fior & al., 2006; Harbaugh & al., 2010), indicates that *Drypis* should be removed from Caryophylloideae and included in Sagineae. *Habrosia*, a monotypic genus present in Iran, Iraq, and Syria, was placed sister to *Drypis* by Smissen & al. (2003), and Harbaugh & al. (2010) suggested that if they were indeed closely related, then they might be placed in their own tribe. However, although our results support the placement of the two genera in the same clade, the species are not sister to each other and therefore should simply remain in Sagineae. *Arenaria fontinalis*, a rare species endemic to Kentucky and Tennessee in the United States, has alternatively been placed in S*agina*, *Stellaria*, and S*pergula*, but has characters such as 4-merous flowers that suggest that it belongs in either *Sagina* or *Minuartia* (Rabeler & Hartman, 2005). Our results place this species within a clade of *Minuartia* and close to *Sagina*.

A group of four *Stellaria* species is included in Sclerantheae along with a single species of *Spergularia* (*S. heldreichii*; 75% BS). This represents a slight divergence from the results of Harbaugh & al. (2010), in which all sampled *Spergularia* species were included in Sperguleae. The four *Stellaria* species found in Sclerantheae are found in western North America, Mexico, South America, or the Dominican Republic. *Stellaria howardii* and *S. minutifolia*, both found in the Dominican Republic, also share 4-merous flowers and shallowly cleft petals with *S. ovata*, which is present in Mexico and mountainous South America (Maguire, 1958). While *Stellaria obtusa*, found in western North America, is less morphologically similar to these species, it does occupy moist wooded areas at relatively high elevations. These *Stellaria* species appear to be most closely related to species of *Geocarpon*, *Minuartia*, and *Wilhelmsia*, which are also present mainly in such habitats in the Western Hemisphere.

A number of small genera including *Cerdia* (4 spp.), *Pteranthus* (1 sp.), *Cardionema* (6 spp.), *Scopulophila* (1 sp.), *Achyronychia* (2 spp.), *Illecebrum* (1 sp.), and *Sphaerocoma* (2 spp.) are found within Polycarpaeae, along with the other small genera (*Dicheranthus*, 1 sp.; *Loeflingia*, 7 spp.; *Ortegia*, 1 sp.) included in this tribe by Harbaugh & al. (2010). Several larger genera, including *Drymaria*, *Polycarpon*, and *Pycnophyllum*, are also included in Polycarpaeae, but relationships between the smaller genera and the larger groups remain unclear. Increased sampling of the larger genera might help to clarify relationships within this clade. The placement of *Pycnophyllum*, traditionally included in Alsinoideae, within Polycarpaeae (traditionally a Paronychioideae lineage) is supported by Harbaugh & al. (2010), and by morphological evidence. While all other Alsinoideae species have nectar glands at the abaxial base of the episepalous stamens, *Pycnophyllum* species do not. These species do have connate styles, a character not present within other Alsinoideae species but commonly found in traditional Paronychioideae. In addition, *Pycnophyllum* has tricolpate pollen, which is common in Paronychioideae, while Alsinoideae species have pantoporate pollen (Bittrich, 1993b). A number of smaller genera are also placed within Sileneae, including *Oberna*, *Atocion*, *Uebelinia*, *Viscaria*, *Eudianthe*, and *Heliosperma*.

Thylacospermum caespitosum, which was placed within Eremogoneae by Harbaugh & al. (2010), is found to be nested within Sperguleae (100% BS) in our tree. This is surprising given that *T. caespitosum* shares a high-elevation, rocky habitat, a perennial life history, white flowers, and a woody base with many Eremogoneae species (in contrast to the lower-elevation, coastal habitats, annual life history, and pink flowers of other Sperguleae species; Lu & al., 2001; Conti, 2003; Rabeler & Hartman, 2005). We note, however, that this species is represented by a single gene region (*matK*) in both studies, and we therefore consider its placement to be uncertain. *Minuartia geniculata*, also found nested in Sperguleae, is the only *Minuartia* species found within one of the four early diverging (Paronychioideae) lineages in our tree. This species is strikingly similar to other species within Sperguleae. It, like many *Spergularia* species, has pink flowers, and is found in coastal environments in the Mediterranean region. It also has tricolpate pollen instead of the pantoporate pollen found in the rest of *Minuartia* (Jackson, 1933; Candau, 1978; Bittrich, 1993b; Rabeler & Hartman, 2005). In fact, the only clades without pantoporate pollen that have traditionally been placed within Alsinoideae or Caryophylloideae are *Minuartia* subg. *Rhodalsine* (which includes *M. geniculata*) and *Pycnophyllum* (see above; Bittrich, 1993b), both of which we place within early diverging lineages.

The phylogenetic placement of the tribes in our analysis is also generally consistent with that of Harbaugh & al. (2010). The four earliest diverging clades (Corrigioleae, 99% BS; Paronychieae, 100% BS; Polycarpaeae, 67% BS; Sperguleae, 51% BS) diverge in the same order as they do in Harbaugh & al. (2010), and the remaining seven tribes also comprise a large clade (61% BS). In both of our analyses, and in previous studies, Sagineae and Sclerantheae are directly linked (86% BS), as are Alsineae and Arenarieae (87% BS), and Sileneae, Caryophylleae, and Eremogone together form a clade (96% BS). However, our results suggest some alternative placements

for these clades. Harbaugh & al. (2010) found that the clade containing Sagineae and Sclerantheae is sister to a clade containing Alsineae, Arenarieae, Sileneae, Eremogoneae, and Caryophylleae, while our results suggest that Sileneae, Eremogoneae, and Caryophylleae (96% BS) are sister (61% BS) to a clade containing Sagineae plus Sclerantheae and Alsineae plus Arenarieae (59% BS). While our support values are generally lower than those of Harbaugh & al. (2010), this is to be expected, as bootstrap values tend to decline with larger datasets and increased amounts of missing data. However, despite lower bootstrap values, large datasets have been shown to reflect the previously well-supported clades of smaller datasets (Sanderson, 2007; Sanderson & al., 2010), as is generally the case here.

While our results do generally support the tribal delimitations of Harbaugh & al. (2010), the monophyly of the major genera is not well supported. Similar results have been obtained in previous analyses (Fior & al., 2006; Frajman & al., 2009; Harbaugh & al., 2010). Multiple genera including *Lychnis*, *Atocion*, *Viscaria*, *Eudianthe*, and *Heliosperma* are sometimes included in *Silene*. Oxelman & al. (2000) suggested that *Eudianthe* be restored as a genus and that *Viscaria*, *Heliosperma*, and *Atocion* should also be recognized as genera. While these genera tend to form well-supported clades within Sileneae and have been previously recognized as clades (Greuter, 1995), the relationships between these clades and *Silene* remain unclear, and it is still possible that they are nested within *Silene*.

Dianthus would be non-monophyletic if *Velezia* were recognized as a separate genus. A recent study of the European radiation of *Dianthus* also supported the inclusion of *Velezia* within this genus (Valente & al., 2010), as did Harbaugh & al. (2010). While the majority of the *Gypsophila* species form a clade, *Gypsophila* in its entirety appears to be non-monophyletic; *Vaccaria*, a small genus of one to four species that is thought to be closely related to *Saponaria* (Bittrich, 1993b), is found nested within *Gypsophila*, and a small clade of *Saponaria* is sister to the *Gypsophila*/*Vaccaria* clade. *Gypsophila*, *Vaccaria*, and *Saponaria* have similar distributions that include temperate Eurasia, the Mediterranean, and the Irano-Turanean region. There are also four *Gypsophila* species present outside of the main *Gypsophila* clade. These species are interspersed with *Petrorhagia* in a grade subtending *Dianthus*.

While a few species of *Stellaria* appear in Sclerantheae, the majority form a clade within Alsineae; however, when all of the species are considered, *Stellaria* is non-monophyletic, and forms a grade subtending *Cerastium*, a few species of *Arenaria*, and the smaller genera *Pseudostellaria*, *Holosteum*, *Moenchia*, *Myosoton*, *Plettkea*, and *Lepyrodiclis*. *Cerastium* forms a clade nested within *Stellaria*. While the majority of the *Arenaria* species form a clade within Arenarieae, it is noteworthy that species of *Arenaria* appear scattered throughout the tree (as described above), making this group non-monophyletic. *Minuartia* species appear in Sclerantheae, Sagineae, Sperguleae, and Eremogoneae, and are non-monophyletic. While all of the *Paronychia* species appear within one of the early diverging lineages, two species (*P. kapela*, *P. chlorothyrsa*) are outside of the main *Paronychia* clade, being closer to *Hernaria* and *Phillppiella*. The Hawaiian lineage, *Schiedea*, is monophyletic within Sclerantheae, nested within a clade that contains a number of small genera and most of the species of *Minuartia*.

It is not entirely surprising that the large genera do not appear to be strictly monophyletic. Circumscription of these groups has been inconsistent (Oxelman & Liden, 1995; Oxelman & al., 2000, 2002), with the recognition in some treatments of many small, segregate genera. Likewise, there has been limited attention to whether the character states that mark these genera are apomorphic, or to the possibility of morphological homoplasy. However, we also cannot entirely rule out the possibility in some cases that misidentified species and/or missing data have contributed to the apparent incongruence with traditional generic delimitations. Misidentifications, if they exist, often persist in GenBank and can be incorporated into multiple studies, giving a false impression of independent support. Indeed, it is worrisome that sequences of several of the species that seem out of place (e.g., *Velezia regida* in *Dianthus*; *Allochrusa versicolor* in *Gypsophila*; *Plettkea cryptantha* and *Myosoton aquaticum* in *Stellaria*) in our analyses were obtained from GenBank. We hope that our analyses will stimulate experts in these groups to evaluate the identity of the voucher specimens, and that special efforts will be made to generate new sequences for these controversial species. It is also possible that biases are introduced by the fact that some of the species in the dataset are represented by just one of the gene regions, and these might tend to cluster with one another. Again, we hope that our analyses will stimulate additional sequencing efforts.

Another possibility is that hybridization, which has been extensively detected within major lineages of Caryophyllaceae (Chinnappa & al., 2005; Popp & Oxelman, 2007; Erixon & Oxelman, 2008; Rautenberg & al., 2008; Frajman & al., 2009; Balao & al., 2010), could be contributing to the finding of non-monophyly in some cases. Although our ITS and chloroplast trees were largely congruent with one another and with the combined dataset, the position of Eremogoneae in the *matK* phylogeny was incongruent with that in the ITS, combined, and chloroplast trees. As noted above, tribe Eremogoneae is sister to Caryophylleae in all analyses except *matK*, where it is sister to Sileneae. In addition, the clade descending from node 10 (including Sagineae and Sclerantheae), which is sister to the clade descending from node 11 (Alsineae and Arenarieae) in the combined and chloroplast analyses, is sister to Alsineae in the ITS tree (Figs. S1–S8). These results could be attributed to ancient hybridization events, which would be consistent with the extensive polyploidy found within major lineages such as *Arenaria*, *Dianthus*, *Cerastium*, *Silene*, and *Stellaria* (Bittrich, 1993b). However, we note that much of the documented hybridization within Caryophyllaceae occurs within genera, and this within-lineage hybridization may help to explain the lack of resolution within clades such as *Silene*, *Cerastium*, and *Dianthus*. Other factors, such as lineage sorting or horizontal gene transfer, cannot yet be dismissed (Rautenberg & al., 2008; Kim & Donoghue, 2008).

Missing data may also play a role in the observed incongruencies. Although 302 species in our dataset are represented by both ITS and at least one chloroplast marker, the other 352 species are represented by only ITS or by only chloroplast genes. While the combined dataset proportionally represents all major groups within Caryophyllaceae, for historical reasons the ITS and chloroplast datasets emphasize different sets of species. For example, many *Silene* species are represented only by ITS, while many *Cerastium* species are represented only by chloroplast markers. To explore the effects of such missing data, we also inferred a tree using only the 302 species represented by both ITS and chloroplast data. This dataset generated generally higher support values, especially along the backbone of the tree. We note that the topology of this "reduced" tree is very similar to both the full combined tree and the chloroplast tree.

Despite the evident non-monophyly of the major genera, it is important to appreciate that our results do support the following relationships among the clades formed by the *majority* of the species in each genus. *Paronychia* is found, along with smaller genera such as *Hernaria* and *Philippiella*, in Paronychieae. Within Caryophylleae, the majority of *Dianthus*, along with several smaller groups, form a clade that is sister to a clade including the majority of *Gypsophila*. Caryophylleae are, in turn, sister to Sileneae, which contain all of the *Silene* species and several small segregate genera. Within Alsineae, *Cerastium* forms a clade nested within several lineages that include the majority of *Stellaria* species. Tribe Alsineae is linked with Arenarieae, which include a clade containing most of the species of *Moehringia*, a clade including most of *Arenaria*, and a smaller clade including some species of both genera. Scleranthеae are linked with Sagineae, with species of *Minuartia* appearing in both of these clades.

Character evolution. — We inferred ancestral states for several taxonomically important morphological characters, paying special attention to the numbered nodes that are deeply nested within the tree (Figs. 5–6). Paronychioideae are traditionally characterized by the presence of stipules, and prominent stipules are present in all species of Corrigioleae, Paronychieae, and Polycarpaeae, and in the majority of species in Sperguleae. Our analyses strongly confirm that Paronychioideae are paraphyletic and that stipules are ancestral within Caryophyllaceae. The earliest nodes in the tree (nodes 1–4 in Figs. 1A, 7) are all reconstructed as having stipules, and this result is supported by comparisons of likelihood scores, with the presence of stipules being at least 8 lnL higher than the alternative character state at nodes 1, 2, and 3. The situation at node 4 is less clear, where there is not a significant difference in likelihood scores. Stipules were either lost at this node and then regained within some Sperguleae, or they were lost along the branch leading to node 5 and also in some Sperguleae. We note that stipules are absent in Amaranthaceae, Achatocarpaceae, and other outgroups, so it is likely that they arose along the branch leading to Carophyllaceae (node 1; Figs. 1A, 7).

Subfamily Caryophylloideae has traditionally been delimited by the presence of a tubular calyx and jointed/clawed petals. As Eremogoneae (containing species with free sepals) are nested within the traditional Caryophylloideae in our analyses, it appears that the tubular calyx either evolved twice (once along the branch subtending Sileaneae, and once along the branch subtending Caryophylloideae) or that it evolved only once and was lost in Eremogoneae. Clearly, this needs to be investigated carefully, including from a developmental standpoint. In the meantime, we note that in examining the individual gene trees (Figs. S1–S6) we find little support for Caryophyllodieae monophyly; that is, a direct link between Caryophylleae and Sileneae, exclusive of Eremogoneae. However, neither do we see strong support for a Caryophylleae-Eremogoneae clade. In fact, *matK* instead provides support for a direct link between Eremogoneae and Sileneae. In any case, it is likely that the tubular calyx evolved independently in *Drypis* within Sagineae.

Although they have not been used to differentiate the traditional subfamilies, we also reconstructed ancestral states for fruit type, petal presence/absence, fruit dehiscence, stamen number, and seed number (Figs. 5–6). Overall, our results suggest a general trend along the backbone of Caryophyllaceae (Fig. 7). Our results imply that the first species of Caryophyllaceae had small apetalous flowers with few stamens and single-seeded indehiscent or irregularly dehiscing utricles. These characters are similar to those found in both Amaranthaceae and Achatocarpaceae, which suggests that they arose before the branch leading to Caryophyllaceae. More nested lineages (nodes 4–8 in Fig. 1A) have larger flowers with frequently notched or bifid petals, 10 stamens, and multi-seeded, regularly dehiscing capsules. The shift from the ancestral to the derived condition of each of these characters is inferred to have occurred along the branch between nodes 3 and 4 (Fig. 7). These results are generally strongly supported by the comparison of likelihood scores. As noted above, our ancestral state reconstruction indicates that ancestral Caryophyllaceae had five stamens, but the lnL values for 3, 4, or 5 stamens are not significantly different. However, we can conclude that there was an increase in stamen number within Caryophyllaceae. It is particularly interesting that stamen number increases from 5 to 10 in nested lineages, evidently correlated with the acquisition of petals. As petals within Caryophyllales generally appear to be derived from stamens (Ronse de Craene, 2007), we might have expected the opposite; that is, the loss of stamens with the gain of petals. This clearly warrants further study. Stamen number in the outgroups ranges from 5 to 10 to 20, although Amaranthaceae generally have 5 stamens, consistent with early Caryophyllaceae (Townsend, 1993).

Classification. — The results presented here begin to provide the basis for a revised higher-level classification of Caryophyllaceae. Although the need for several major changes is now clear, including the abandonment of the paraphyletic Paronychioideae, we believe that a comprehensive new treatment must await clarification on several fronts. Specifically, it will be important to further test the placement of the many small genera among the larger genera, and to resolve the instances of conflict between nuclear and chloroplast genes noted above.

In the meantime, we believe that work on Caryophyllaceae phylogeny and evolution will be well served by the naming of

one major new clade that has been identified in our study and in several previous analyses (Fior & al., 2006; Harbaugh & al., 2010). Specifically, we are referring to the clade descending from node 4 (Fig. 7) that includes Sperguleae (mostly containing former members of Paronychiodeae) and the remainder of Caryophyllaceae (including Alsinoideae and Caryophylloideae). As shown in Fig. 7, and as discussed above, this well supported clade cuts directly across the traditional subfamily boundaries, and therefore seems especially important to flag. As we have emphasized, this clade is well-marked by clear-cut apomorphies, including the presence of petals, 10 stamens, and capsule fruits (Fig. 7). We here provide the name Plurcaryophyllaceae for this clade, and, for clarity, also provide a formal phylogenetic definition following the conventions in Cantino & al. (2007).

Plurcaryophyllaceae A.K. Greenberg & M.J. Donoghue, new clade name.

Comments on name. – There is no preexisting scientific name for this clade. The new name *Plurcaryophyllaceae* is proposed because this clade contains the majority of Caryophyllaceae species. "*Plur-*" means majority or most.

Definition (node-based). – The least inclusive clade containing *Spergularia rubra* (L.) J. Presl & C. Presl 1819 and *Silene nutans* L. 1753.

Reference phylogeny. – This paper. See also Harbaugh & al. (2010) and Fior & al. (2006).

Composition. – Traditional Alisnoideae and Caryophylloideae of the Caryophyllaceae (tribes Alsineae, Arenarieae, Caryophylleae, Eremogoneae, Sagineae, Sclerantheae, and Sileneae of Harbaugh & al., 2010) plus Sperguleae (sensu Harbough & al., 2010).

Synapomorphies.– Presence of petals, 10 stamens, and capsule fruits (see above, Fig. 7).

Synonymy. – None.

We are certain that it will eventually be useful to formally name several other major clades in Fig. 7, including Caryophyllaceae itself. Specifically, it may soon become helpful to name the clade eminating from node 5, which includes Alsinoideae and Caryophylloideae. This clade is marked by the loss of stipules and, it appears, by a shift in pollen morphology from tricolpate to pantoporate (Candau, 1978; Bittrich 1993b). However, we note that at the present time this clade is not as strongly supported as *Plurcaryophyllaceae*, nor as clearly marked by derived morphological character states.

■ ACKNOWLEDGEMENTS

We thank S. Smith, J. Beaulieu, R. Rabeler, and E. Edwards and her lab for valuable feedback. S. Smith also provided helpful scripts for Figs. 5 and 6 along with advice about running analyses for large datasets. We also thank B. Oxelman for valuable feedback and nomenclatural advice. We thank the Yale University Herbarium, NYBG Herbarium, Harvard Herbaria, RBG Kew, and RBG Edinburgh, as well as R. Ree for leaf material used in this study. This work was supported by Yale University funds provided to MJD.

■ LITERATURE CITED

Alfaro, M.E. & Huelsenbeck, J.P. 2006. Comparative performance of Bayesian and AIC-based measures of phylogenetic model uncertainty. *Syst. Biol.* 55: 89–96.

Balao, F., Valente, L.M., Vargas, P., Herrera, J. & Talavera, S. 2010. Radiative evolution of polyploid races of the Iberian carnation *Dianthus broteri* (Caryophyllaceae). *New Phytol.* 187: 542–551.

Bittrich, V. 1993a. Achatocarpaceae. Pp. 35–36 in: Kubitzki, J. (ed.), *The families and genera of vascular plants*, vol. 2. Berlin: Springer.

Bittrich, V. 1993b. Caryophyllaceae. Pp. 206–236 in: Kubitzki, J. (ed.), *The families and genera of vascular plants*, vol. 2. Berlin: Springer.

Brockington, S.F., Alexandre, R., Ramdial, J., Moore, M.J., Crawley, S., Dhingra, A., Hilu, K., Soltis, D.E. & Soltis, P.E. 2009. Phylogeny of the Caryophyllales sensu lato: Revisiting hypotheses on pollination biology and perianth differentiation in the core Caryophyllales. *Int. J. Pl. Sci.* 170: 627–643.

Candau, P. 1978. Palinologia de Caryophyllaceae del sur de Espana: 1. Subfamilia Paronychioideae. *Lagascalia* 7: 143–157.

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.

Chinnappa, C.C., Donald, G.M., Sasidharan, R. & Neil Emery, R.J. 2005. The biology of *Stellaria longipes* (Caryophyllaceae). *Canad. J. Bot.* 83: 1367–1383.

Chrtek, J. & Slavikova, Z. 1987. Leitbündelanordnung in den Kronblättern von ausgewählten Arten der Familie Stellariaceae. *Preslia* 60: 11–21.

Conti, F. 2003. *Minuartia graminifolia* (Caryophyllaceae), a south-east European species. *Bot. J. Linn. Soc.* 143: 419–432.

Cuenod, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J. & Chase, M.W. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *Amer. J. Bot.* 89: 132–144.

Downie, S.R., Katz-Downie, D.S. & Cho, K.-J. 1997. Relationships in the Caryophyllales as suggested by phylogenetic analyses of partial chloroplast DNA ORF2280 homolog sequences. *Amer. J. Bot.* 84: 253–273.

Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* 32: 1792–1797.

Edwards, A.W.F. 1992. *Likelihood*, 2nd ed.. Baltimore: The Johns Hopkins University Press.

Edwards, E.J. & Smith, S.A. 2010. Phylogenetic analyses reveal the shady history of C_4 grasses. *Proc. Natl. Acad. Sci. U.S.A.* 107: 2532–2537.

Erixon, P. & Oxelman, B. 2008. Reticulate or tree-like chloroplast DNA evolution in *Sileneae* (Caryophyllaceae)? *Molec. Phylogenet. Evol.* 48: 313–325.

Fernandes, A. & Leitao, M.T. 1971. Contribution a la connaissance cytotaxonomique des Spermatophyta do Portugal. III. Caryophyllaceae. *Bol. Soc. Brot.* 45: 167–176.

Fior, S., Karis, P.O., Casazza, G., Minuto, L. & Sala, F. 2006. Molecular phylogeny of the Caryophyllaceae (Caryophyllales) inferred from chloroplast *matK* and nuclear rDNA ITS sequences. *Amer. J. Bot.* 93: 399–411.

Frajman, B., Eggens, F. & Oxelman, B. 2009. Hybrid origins and homoploid reticulate evolution within *Heliosperma* (Sileneae, Caryophyllaceae): A multigene phylogenetic approach with relative dating. *Syst. Biol.* 58: 328–345.

Goldberg, E.E., Kohn, J.R., Lande, R., Robertson, K.A., Smith, S.A. & Igic, B. 2010. Species selection maintains self-incompatibility. *Science* 328: 587–591.

Goloboff, P.A., Catalano, S.A., Mirande, J.M., Szumik, C.A., Arias, J.S., Källersjö, M. & Farris, J.S. 2009. Phylogenetic analysis of 73,060 taxa corroborates major eukaryotic groups. *Cladistics* 25: 211–230.

Greuter, W. 1995. Studies in Greek Caryophylloideae: *Agrostemma*, *Silene*, and *Vaccaria*. *Willdenowia* 25: 105–142.

Harbaugh, D.T., Nepokroeff, M., Rabeler, R.K., McNeill, J., Zimmer, E.A. & Wagner, W.L. 2010. A new lineage-based tribal classification of the family Caryophyllaceae. *Int. J. Pl. Sci*. 171: 185–198.

Jackson, A.K. 1933. Two plants new to the flora of Cyprus. *Kew Bull*. 8: 407–408.

Jackson, V.K. 2004. LASRDisc: Likelihood Ancestral State Reconstruction for Discrete Characters, version 1.0. http://ceb.csit.fsu.edu/lasrdisc/.

Kim, S.-T. & Donoghue, M.J. 2008. Incongruence between cpDNA and nrITS trees indicates extensive hybridization within *Eupersicaria* (Polygonaceae). *Amer. J. Bot.* 95: 1122–1135.

Lemmon, A.R. & Moriarty, E.C. 2003. The importance of proper model assumption in Bayesian phylogenetics. *Syst. Biol*. 53: 265–277.

Loytynoja, A. & Goldman, N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320: 1632–1635.

Lu, D., Wu, Z., Wu, C.Y., Zhou, L., Chen, S., Gilbert, M.G., Liden, M., McNeill, J., Morton, J.K., Oxelman, B., Rabeler, R.K., Thulin, M., Turland, N.J. & Wagner, W.L. 2001. *Caryophyllaceae through Lardizabalaceae*. Vol. 6 of: Wu, Z.Y., Raven, P.H. & Hong, D.Y. (eds.), *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.

Maddison, W.P. & Maddison, D.R. 2010. Mesquite: A modular system for evolutionary analysis, version 2.73. http://mesquiteproject.org.

Maguire, B. 1958. Two Dominican Stellarias. *Brittonia* 10: 32.

Moles, A.T., Ackerly, D.D., Webb, C.O., Tweddle, J.C., Dickie, J.B. & Westoby, M. 2005. A brief history of seed size. *Science* 307: 576–580.

Oxelman, B. & Liden, M. 1995. Generic boundaries in the tribe Sileneae (Caryophyllaceae) as inferred from nuclear rDNA sequences. *Taxon* 44: 525–542.

Oxelman, B., Ahlgren, B. & Thulin, M. 2002. Circumscription and phylogenetic relationships of *Gymnocarpos* (Caryophyllaceae-Paronychioideae). *Edinburgh J. Bot.* 59: 221–237.

Oxelman, B., Liden, M., Rabeler, R.K. & Popp, M. 2000. A revised generic classification of the tribe Sileneae (Caryophyllaceae). *Nord. J. Bot.* 2: 743–748.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol*. 48: 612–622.

Pelser, P.B., Nordenstam, B., Kadereit, J.W. & Watson, L.E. 2007. An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of *Senecio* L. *Taxon* 56: 1077–1104.

Popp, M. & Oxelman, B. 2007. Origin and evolution of North American polyploid *Silene* (Caryophyllaceae). *Amer. J. Bot.* 94: 330–349.

Rabeler, R.K. & Hartman, R.L. 2005. *Caryophyllaceae*. Vol. 5 of: Flora of North America Editorial Committee (eds.), *Flora of North America north of Mexico*. New York and Oxford: Oxford University Press.

Raghavan, V. 2006. *Double fertilization: Embryo and endosperm development in flowering plants*. Berlin, Heidelberg: Springer.

Rautenberg, A., Filatov, D., Svennblad, B., Heidari, N. & Oxelman, B. 2008. Conflicting phylogenetic signals in the SIX1/Y1 gene in *Silene*. *B.M.C. Evol. Biol.* 8: 299. DOI: 10.1186/1471-2148-8-299.

Ripplinger, J. & Sullivan, J. 2008. Does choice in model selection affect maximum likelihood analysis? *Syst. Biol.* 57: 76–85.

Ronse de Craene, L. 2007. Are petals sterile stamens or bracts? The origin and evolution of petals in the Core Eudicots. *Ann. Bot.* 100: 621–630.

Sanderson, M.J. 2007. Construction and annotation of large phylogenetic trees. *Austral. Syst. Bot.* 20: 287–301.

Sanderson, M.J., Boss, D., Chen, D., Cranston, K.A. & Wehe, A. 2008. The PhyLoTA Browser: Processing GenBank for molecular phylogenetics research. *Syst. Biol.* 57: 335–346.

Sanderson, M.J., McMahon, M.M. & Steel, M. 2010. Phylogenomics with incomplete taxon coverage: The limits to inference. *B.M.C. Evol. Biol.* 10: 155. DOI: 10.1186/1471-2148-10-155.

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

Schmidt, G.J. & Schilling, E.E. 2000. Phylogeny and biogeography of *Eupatorium* (Asteraceae: Eupatorieae) based on nuclear ITS sequence data. *Amer. J. Bot.* 87: 716–726.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, S.B., Winder, C.T., Schilling, E.E. & Small, R.L. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Amer. J. Bot.* 92: 142–166.

Shaw, J., Lickey, E.B., Schilling, E.E. & Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Amer. J. Bot.* 94: 275–288.

Smissen, R.D., Clement, J.C., Garnock-Jones, P.J. & Chambers, G.K. 2002. Subfamilial relationships within Caryophyllaceae as inferred from 5′ *ndhF* sequences. *Amer. J. Bot.* 89: 1336–1341.

Smissen, R.D., Garnock-Jones, P.J. & Chambers, G.K. 2003. Phylogenetic analysis of ITS sequences suggests a Pliocene origin for the bipolar distribution of Scleranthus (Caryophyllaceae). *Austral. Syst. Bot.* 16: 301–315.

Smith, S.A. & Beaulieu, J.M. 2009. Life history influences rates of climatic niche evolution in flowering plants. *Proc. Roy. Soc. B, Biol. Sci.* 276: 4345–4352.

Smith, S.A. & Donoghue, M.J. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322: 86–88.

Smith, S.A., Beaulieu, J.M. & Donoghue, M.J. 2009. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *B.M.C. Evol. Biol.* 9: 37. DOI: 10.1186/1471-2148-9-37.

Smith, S.A., Beaulieu, J.M., Stamatakis, A. & Donoghue, M.J. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *Amer. J. Bot.* 98: 404–414.

Stamatakis, A. 2006a. Phylogenetic models of rate heterogeneity: A high performance computing perspective. *Proceedings of 20th International Parallel and Distributed Processing Symposium*. April 25–29, 2006; Rhodes Island, Greece. IEEE/ACM.

Stamatakis, A. 2006b. RaxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stamatakis, A., Hoover, P. & Rougemont, J. 2008. A rapid bootstrap algorithm for the RaxML Web servers. *Syst. Biol.* 57: 758–775.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Pl. Molec. Biol.* 17: 1105–1109.

Thomson, R.C. & Shaffer, H.B. 2010. Sparse supermatrices for phylogenetic inference: Taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. *Syst. Biol.* 59: 42–58.

Townsend, C.C. 1993. Amaranthaceae. Pp. 70–91 in: Kubitzki, J. (ed.), *The families and genera of vascular plants*, vol 2. Berlin: Springer.

Valente, L.M., Savolainen, V. & Vargas, P. 2010. Unparalleled rates of species diversification in Europe. *Proc. Roy. Soc. B, Biol. Sci.* 277: 1489–1496.

Wurdack, K.J., Hoffmann, P., Samuel, R., de Bruijn, A., Van der Bank, M. & Chase, M.W. 2004. Molecular phylogenetic analysis of Phyllanthaceae (Phyllanthoideae pro parte, Euphorbiaceae sensu lato) using plastid *rbcL* DNA sequences. *Amer. J. Bot.* 91: 1882–1900.