Plant Syst. Evol. 231: 77–89 (2002)

Plant Systematics and Evolution

Phylogenetic relationships in Rosaceae inferred from chloroplast *mat*K and *trn*L-*trn*F nucleotide sequence data

D. Potter[1], **F. Gao**[1], **P. Esteban Bortiri**[1], **S.-H. Oh**[1], and **S. Baggett**[2]

[1]Department of Pomology, University of California, Davis, USA
[2]Ph.D. Program Biology, Lehman College, City University of New York, New York, USA

Received February 27, 2001
Accepted October 11, 2001

Abstract. Phylogenetic relationships in Rosaceae were studied using parsimony analysis of nucleotide sequence data from two regions of the chloroplast genome, the *mat*K gene and the *trn*L-*trn*F region. As in a previously published phylogeny of Rosaceae based upon *rbc*L sequences, monophyletic groups were resolved that correspond, with some modifications, to subfamilies Maloideae and Rosoideae, but Spiraeoideae were polyphyletic. Three main lineages appear to have diverged early in the evolution of the family: 1) Rosoideae *sensu stricto*, including taxa with a base chromosome number of 7 (occasionally 8); 2) actinorhizal Rosaceae, a group of taxa that engage in symbiotic nitrogen fixation; and 3) the rest of the family. The spiraeoid genus *Gillenia*, not included in the *rbc*L study, was strongly supported as the sister taxon to Maloideae *sensu lato*. A New World origin of Maloideae is suggested. The position of the economically important genus *Prunus* and the status of subfamily Amygdaloideae remain unresolved.

Key words: Rosaceae, *Gillenia*, Maloideae, Amygdaloideae, Spiraeoideae, phylogeny, *mat*K, *trn*L-*trn*F.

The large and economically important angiosperm family Rosaceae has a worldwide distribution and includes over 3000 species in 122 genera (Heywood 1993). The vast majority of economically important fruits of temperate regions is produced by members of this family, including species of *Malus* (apples), *Pyrus* (pears), *Prunus* (plums, peaches, cherries, almonds, and apricots), *Rubus* (raspberries), and *Fragaria* (strawberries). The family also includes many ornamentals, e.g., species of *Rosa* (roses), *Potentilla* (cinquefoil), and *Sorbus* (mountain ash). A variety of growth habits, fruit types, and chromosome numbers is found within the family (Robertson 1974), which is traditionally divided into four subfamilies on the basis of fruit type (e.g., Schulze-Menz 1964). Spiraeoideae are characterized by follicles or capsules, Rosoideae by achenes, Amygdaloideae (Prunoideae) by drupes, and Maloideae by pomes. In the traditionally circumscribed Rosoideae, the base chromosome number is $x = 7$, 8, or 9; in Amygdaloideae, $x = 8$; in Spiraeoideae, $x = 9$ (15, 17); and in Maloideae, $x = 17$. The division of the family into four subfamilies has not been followed universally. For example, Hutchinson (1964) recognized 20 tribes and did not group these into subfamilies.

Because of their economic importance and diversity, the Rosaceae have been the subject of numerous taxonomic and evolutionary studies. The family is generally considered to

form a natural group united by floral characteristics (Robertson 1974, Dickson et al. 1992). Kalkman (1988) suggested that the presence of a hypanthium may be the only morphological synapomorphy for the group, but the uncertainty of the relationship of Rosaceae to other families casts some doubt on this conclusion. This, as well as the relationships among subfamilies, genera, and species, are subjects of considerable discussion and investigation (summarized in Morgan et al. 1994, Phipps et al. 1991). Recent phylogenetic analyses have employed a variety of data, including vegetative, floral, and fruit morphology (Kalkman 1988; Phipps et al. 1991; Rohrer et al. 1991, 1994), floral ontogeny (Evans and Dickinson 1999a, b), wood anatomy (Zhang 1992), pollen morphology (Hebda and Chinnappa 1994), chloroplast DNA sequences (Morgan et al. 1994), nuclear gene sequences (Potter 1997, Evans et al. 2000), and combined data from multiple sources (Evans and Dickinson 1997, 1999c), to address questions about the placement of problematic genera and the relationships of the subfamilies. Sequences of the internal transcribed spacer (ITS) regions of nuclear ribosomal DNA have been used in phylogenetic studies of two of the subfamilies (Maloideae, Campbell et al. 1995, Rosoideae, Eriksson et al. 1998). These studies have done much to improve our understanding of the affinities of particular taxa and the evolution of specific characters in Rosaceae, but a variety of questions remain, and all of the cited papers point out areas in which further resolution is needed.

In a phylogenetic study of *rbc*L gene sequence variation across the family (Morgan et al. 1994), monophyletic groups of genera were identified that corresponded, with some modifications, to all of the subfamilies except Spiraeoideae, which was shown to be grossly polyphyletic. The data strongly supported recognition of Rosoideae *sensu stricto*, excluding several taxa with $x = 9$, and of Maloideae *sensu lato*, including several spiraeoid taxa with $x = 15$ or 17. Weak support was found for Amygdaloideae *sensu lato*, including *Prunus*, *Prinsepia*, *Oemleria*, and *Exochorda*. The results suggested that, since some fruit types have evolved several times within the family, they are not as reliable as indicators of relationship as chromosome numbers.

Takhtajan (1997) incorporated some results from recent phylogenetic studies, such as that of Morgan et al. (1994), into his classification of the family, in which he recognized twelve subfamilies. *Exochorda*, previously classified in Spiraeoideae, was included in Amygdaloideae, Maloideae were expanded to include the traditionally spiraeoid genera *Kageneckia*, *Vauquelinia*, and *Lindleya*, and both Rosoideae and Spiraeoideae were subdivided.

A long-standing question concerns the origin of Maloideae. Because of a base chromosome number of 17, it is generally accepted that the subfamily originated either via polyploidization of a spiraeoid ancestor with $x = 9$ (e.g. Gladkova 1972) or by hybridization between two lineages, most likely a spiraeoid with $x = 9$ and an amygdaloid with $x = 8$, followed by polyploidization (e.g. Sax 1933). The exact identity of the putative parental lineage(s) remains controversial, however. The *rbc*L analysis by Morgan et al. (1994) indicated that members of the spiraeoid tribe Sorbarieae (plus the traditionally rosoid genus *Adenostoma*) are sister to Maloideae *sensu lato*, a relationship consistent with data from carpel anatomy (Sterling 1966), but this relationship was weakly supported. Morgan et al. (1994) pointed out that, since *rbc*L is a chloroplast-encoded gene, phylogenetic hypotheses based upon it represent phylogenies of maternal lineages only. Thus, the results are not inconsistent with an ancestral member of Amygdaloideae with $x = 8$ having been one of the parents involved in an ancient hybridization that led to the origin of Maloideae (Sax 1933, Phipps et al. 1991). Recent phylogenetic studies of the nuclear gene *waxy* by Evans et al. (2000), however, have not provided any evidence for an amygdaloid ancestor having been involved in the origin of Maloideae; nor have phylogenetic studies of several nuclear genes in our laboratory (e.g. Potter et al. 1999).

Recent molecular phylogenetic analyses (Chase et al. 1993, Morgan et al. 1994, Soltis et al. 1995, Källersjö et al. 1998, Soltis et al. 2000) place Rosaceae in a clade with Moraceae, Rhamnaceae, Urticaceae, and several others, which together have been designated as order Rosales (Angiosperm Phylogeny Group 1998). Data from several genes further support a sister relationship of Rosaceae to the rest of the families in the order (Evans and Campbell 2000, Soltis et al. 2000).

The goal of this study was to assess phylogenetic relationships across Rosaceae using new chloroplast DNA data, with a primary emphasis on the mostly woody taxa not included in Rosoideae *sensu stricto*, i.e., genera traditionally classified in Spiraeoideae, Maloideae, and Amygdaloideae, and taxa of Rosoideae with $x = 9$. We wanted to test the relationships identified by Morgan et al. (1994) and we were especially interested in seeing whether or not we could obtain better resolution than was provided by the *rbc*L data in the following areas: 1) the deep branches within the Rosaceae phylogenetic tree; 2) the affinities of *Prunus* and other taxa traditionally classified in Amygdaoideae; and 3) the relationships of Maloideae to spiraeoid and/or amygdaloid taxa. In addition, we included *Gillenia* of tribe Gillenieae (Spiraeoideae; Schulze-Menz 1964), a group not represented in Morgan et al.'s (1994) *rbc*L study. Several members of other families of Rosales *sensu* Angiosperm Phylogeny Group (1998) were included as outgroups.

We examined nucleotide sequences from two regions of the chloroplast genome. The first of these, the *mat*K gene, has provided useful information in phylogenetic studies at the intrafamilial level in a number of angiosperm groups (e.g., Xiang et al. 1998, Hu et al. 2000). The maturase-encoding gene occurs as a 1.5 kb region embedded within a 2.5 kb intron that interrupts the two *trn*K exons (Sugita et al. 1985). The second region of the chloroplast genome we have examined is a fragment of about 1 kb, including the *trn*L intron and the *trn*L-*trn*F spacer. This is a noncoding region of cpDNA for which universal primers exist (Taberlet et al. 1991) and which has been useful in phylogenetic studies of angiosperm groups at various taxonomic levels (e.g. Gielly and Taberlet 1996, Karol et al. 2000, Potter et al. 2000, Scott et al. 2000, Bortiri et al. in press).

Materials and methods

The species sampled for this study are listed in Table 1. Voucher specimens are deposited in the U. C. Davis Herbarium (DAV). Total DNA was extracted from fresh leaf material using the CTAB method (Doyle and Doyle 1987) or modifications thereof. In the case of two species (*Fragaria vesca* and *Oemleria cerasiformis*), the *mat*K and *trn*L-*trn*F sequences were determined from different accessions.

Primers for PCR and sequencing were purchased from Genosys Biotechnologies, Inc. PCR amplifications were carried out using the Perkin-Elmer GeneAmp II kit. For *mat*K, a 1.5–1.9 kb fragment was amplified using trnK685F and trnK2R as primers as described by Hu et al. (2000). The forward PCR primer, trnK685F, is located at 685 bp upstream of the *trn*K 5′ exon relative to the *Pisum* sequence (Boyer and Mullet 1988). The reverse PCR primer, trnK2R, occurs at site 2475 on the *trn*K 3′ exon. The *trn*L-*trn*F region was amplified using primers c and f (Taberlet et al. 1991). For both regions, PCR conditions were as follows: 1 minute at 95 °C; 40 cycles of 30 seconds at 95 °C, 1 minute at 55 °C, and 2 minutes at 72 °C; 7 minutes at 72 °C.

PCR products were purified from agarose gels with the QIAquick Gel Extraction Kit (Qiagen Inc.). PCR products and sequencing primers were submitted to one of two sequencing facilities on the U. C. Davis campus, each of which uses an ABI/Prism 377 automated sequencer. For *mat*K, the two PCR primers, plus two internal primers, matK3F (5′-TCCCTCTTCTTTGCATTTATTACG-3′) and matK4R (5′-GCGTTACAAAATTTCACTTTAGCC-3′), were used for sequencing. For *trn*L-*trn*F, primers c, d, e, and/or f (Taberlet et al. 1991) were used, in various combinations, because some primers worked better than others for sequencing some templates.

Sequences were edited with SEQUENCHER™ 3.1.1 (Gene Codes Corporation). Boundaries of the *mat*K coding regions were

Table 1. Accessions included in this study, arranged by subfamilies as treated by Schulze-Menz (1964)

Species	Origin and Collection Number	GenBank Accession Numbers: *mat*K/ *trn*L-*trn*F
Maloideae:		
Cotoneaster pannosa Franchet	Santa Cruz County, CA DXP 033	AF288098/ AF348540
Crataegus monogyna Jacq.	Guemes Island, WA DP 970517-08	AF288099/ AF348541
Photinia serrulata Lindl.	U. C. Davis campus DP 970911-01	AF288111/ AF348552
Pyrus caucasica Fed.	UCDA[a] A69.0879	AF288120/ AF348564
Sorbus californica E. Greene	Placer County, CA DXP 073	AF288126/ AF348570
Amygdaloideae (Prunoideae):		
Oemleria cerasiformis (Hook. & Arn.) J. W. Landon	Point Reyes, CA DXP 059	AF288110/ —
Oemleria cerasiformis (Hook. & Arn.) J. W. Landon	Guemes Island, WA DXP 069	—/ AF348551
Prinsepia sinensis (Oliv.) Oliv. ex Bean	Harvard University DXP 191	AF288114/ AF348558
Prunus laurocerasus L.	UCDA[a] EB 88	AF288116/ AF348559
Prunus persica (L.) Sieb. & Zucc. "54P 455"	Cultivar SA[b]	AF288117/ AF348560
Prunus virginiana L.	NCGR[c] DPRU 393	AF288118/ AF348561
Rosoideae:		
Adenostoma fasciculatum Hook. & Arn.	Winters, CA SO 970424-01	AF288093/ AF348535
Cercocarpus betuloides Torrey & A. Gray	UCDA[a] DXP 037	AF288095/ AF348537
Chamaebatia foliolosa Benth.	Yosemite, CA DP 970427-02	AF288096/ AF348538
Fallugia paradoxa (D. Don) Endl.	UCDA[a] DXP 610	AF288101/ AF348543
Fragaria vesca L.	San Mateo, CA DXP 026	AF288102/ —
Fragaria vesca L.	Yosemite, CA DP 970427-01	—/ AF348545
Neviusia alabamensis A. Gray	BBG[d] 93.0973	AF288109/ AF348550
Potentilla anserina L.	Montara, CA DP 970309-02	AF288113/ AF348556
Purshia tridentata (Pursh) DC.	Luther Pass, CA DP 970831-02	AF288119/ AF348562
Rhodotypos scandens (Thunb.) Mak.	BBG[d] 86.0616	AF288122/ AF348566
Rosa californica Cham. & Schldl.	UCDA[a] T00292	AF288123/ AF348567
Rubus ursinus Cham. & Schldl.	San Mateo, CA DXP 027	AF288124/ AF348568
Spiraeoideae:		
Aruncus dioicus (Walter) Fern.	BBG[d] 83.0466	AF288094/ AF348536
Chamaebatiaria millefolium (Torr.) Maxim.	UCDA[a] A74.0245	AF288097/ AF348539
Exochorda racemosa (Lindl.) Rehder	Henan, China DXP 613	AF288100/ AF348542
Gillenia stipulata (Muhl. ex Willd.) Baillon	BBG[d] 92.0438	AF288103/ AF348554
Gillenia trifoliata (L.) Moench	Cornell University DXP 192	AF288104/ AF348555

Table 1 (continued)

Species	Origin and Collection Number	GenBank Accession Numbers: matK/ trnL-trnF
Holodiscus discolor (Pursh) Maxim.	Guemes Island, WA DXP 070	AF288105/ AF348546
Kageneckia oblonga Ruiz & Pav.	BBG[d] 88.0176	AF288106/ AF348547
Lyonothamnus floribundus A. Gray	UCDA[a] A84.0082	AF288107/ AF348548
Neillia thyrsiflora D. Don	RBGE[e] 19841790	AF288108/ AF348549
Physocarpus capitatus (Pursh) Kuntze	UCDA[a] DP 970702-01	AF288112/ AF348553
Sorbaria sorbifolia (L.) A. Braun	BBG[d] 83.0529	AF288125/ AF348569
Spiraea densiflora Torr. & A. Gray	Tahoe N.F., CA DP 970619-02	AF288127/ AF348571
Stephanandra chinensis Hance	USNA[f] 59954	AF288128/ AF348572
Vauquelinia californica (Torr.) Sarg.	UCDA[a] A77.0200	AF288129/ AF348573
Outgroups:		
Rhamnus californica Eschsch.	UCDA[a] A93.0177	AF288121/ AF348565
Morus alba L.	U. C. Davis campus DXP 100	AF400590/ AF400592
Ulmus procera Salisb.	U. C. Davis campus DXP 347	AF400591/ AF400593

[a] UCDA = U. C. Davis Arboretum
[b] SA = DNA kindly provided by S. Arulsekar, U. C. Davis
[c] NCGR = National Clonal Germplasm Repository, Davis
[d] BBG = Berkeley Botanic Garden
[e] RBGE = Royal Botanic Garden, Edinburgh
[f] USNA = U. S. National Arboretum

determined by comparison to that of *Pisum* (Boyer and Mullet 1988). Only the coding regions were included in the alignment. The *mat*K sequences were aligned by eye; the *trn*L-*trn*F sequences were aligned using ClustalX (Thompson et al. 1997) and refined by eye. Sequences and alignments were submitted to GenBank (Table 1); these are also available, upon request, from the first author.

All sequence comparisons and phylogenetic analyses were carried out using PAUP* (Swofford 2000). Pairwise divergences (Jukes-Cantor distances) among sequences were calculated. Phylogenetic analysis of the data employing maximum parsimony was implemented using heuristic searches and 100 replicates of random taxon addition with TBR branch-swapping and MULPARS in effect. All positions were weighted equally; gaps were treated as missing values. The *mat*K and *trn*L-*trn*F data were analyzed separately and in combination. In order to determine whether or not there was significant conflict between the data from the two regions, the incongruence length difference (ILD) test (Farris et al. 1994) was used, with heuristic searches, as described above, and 100 replicates. Relative support for clades was assessed using phylogenetic bootstrapping (Felsenstein 1985), with 1000 replicates, and by decay indices (Mishler et al. 1991). The latter were determined as follows: After the initial heuristic search was completed and the most parsimonious trees and their strict consensus tree had been examined, the process was repeated eight times, each time increasing the maximum length of trees to be saved by one step over the previous search, beginning with one step longer than the most parsimonious trees identified in the initial analysis. The strict consensus tree from each of these searches was then compared to the initial consensus tree to see which branches had collapsed.

Results

The *mat*K sequences (coding region) ranged from 1503 to 1521 bp long, and the final alignment included 1551 sites. The *trn*L-*trn*F region sequences ranged from 930 to 1083 bp long, and the final alignment included 1295 sites. The final combined data matrix therefore contained 2846 characters, of which 1592 were constant, 540 were phylogenetically uninformative, and 714 (385 from *mat*K and 329 from *trn*L-*trn*F) were phylogenetically informative. The ILD test indicated that there was no significant incongruence ($p = 0.30$) between these two regions of the chloroplast genome.

For comparisons between outgroup and ingroup taxa, pairwise sequence divergence values ranged from 0.1173 (*Rhamnus californica* and *Lyonothamnus floribundus*) to 0.1846 (*Ulmus procera* and *Potentilla anserina*). Among the ingroup taxa, these values ranged from 0.0024 (between *Stephanandra chinensis* and *Neillia thyrsiflora*) to 0.1461 (between *Potentilla anserina* and *Aruncus dioicus*).

Phylogenetic analysis of the combined data matrix produced four most parsimonious trees (Length = 2472, CI excluding autapomorphies = 0.58, RI = 0.71), one of which is shown in Fig. 1. Separate analyses of the *mat*K and *trn*L-*trn*F data produced trees (not shown) that were generally similar, in terms of levels of homoplasy and clades recovered, to those from the combined analysis. An exception is discussed below.

In terms of intergeneric relationships in Rosaceae, our results were generally consistent with those from *rbc*L data (Morgan et al. 1994). As in that study, a number of clades were strongly supported, but support for relationships among those groups was generally weak.

The taxa of Rosoideae *sensu stricto* ($x = 7$) formed a well-supported clade, as did most of the remaining taxa in the family, with $x = 8$ or above. *Adenostoma*, *Neviusia*, and *Rhodotypos*, traditionally classified in Rosoideae but with $x = 9$, all appeared within the latter clade. The position of the clade including three actinorhizal taxa, *Cercocarpus*, *Chamaebatia*, and *Purshia*, also traditionally placed in Rosoideae but with $x = 9$, varied among the most parsimonious trees. In two trees, this clade was sister to Rosoideae *sensu stricto* (hereafter, topology "A"); in the other two (e.g., Fig. 1), it was sister to the clade including all other members of the family *except* Rosoideae *sensu stricto* (hereafter, topology "B"). These results were sensitive to both the outgroups and the characters included in the analysis. When the combined data were analyzed with just *Rhamnus* or just *Morus* and *Ulmus* as outgroups, only trees of topology "B" were obtained; any other combination of outgroups produced at least some trees of topology "A". When the *trn*L-*trn*F data were analyzed alone, only trees of topology "B" were obtained, regardless of the outgroup(s) included. In contrast, when the *mat*K data were analyzed alone, only trees of topology "A" were obtained with all combinations of outgroups except *Rhamnus* alone, which produced only trees of topology "B".

The traditional Spiraeoideae and Amygdaloideae were shown to be polyphyletic. Monophyly of the traditional Maloideae, all of which produce pomes, was strongly supported, as was that of Maloideae *sensu lato*, including *Vauquelinia californica* (Spiraeoideae, $x = 15$) and *Kageneckia oblonga* (Spiraeoideae, $x = 17$) plus the pome-producing taxa. *Gillenia* was strongly supported as the sister group to Maloideae *sensu lato*. Specific similarities and differences between our analysis and that of Morgan et al. (1994) are discussed below.

Discussion

Phylogenetic resolution. Clades within Rosaceae that were strongly supported in our analysis (80% or greater bootstrap support and decay index values of 5 or greater in Fig. 1) and that of Morgan et al. (1994; decay index values of 4 or greater in their Fig. 2) include: 1) Rosoideae *sensu stricto*, including members of the subfamily with $x = 7$; our sampling within this group was admittedly limited; 2) the actinorhizal Rosaceae, including *Cercocarpus*,

Purshia, and *Chamaebatia* (the last not included in the *rbc*L study), a group of taxa distributed in western North America that form symbiotic relationships with nitrogen-fixing actinomycetes of the genus *Frankia*; 3) *Rhodotypos* and *Neviusia*; 4) *Exochorda*, *Oemleria*, and *Prinsepia*; 5) tribe Neillieae, comprising *Physocarpus*, *Neillia*, and *Stephanandra* (the last not included in the *rbc*L study); 6) Maloideae *sensu lato*, including the pome-bearing Maloideae plus *Kageneckia* and *Vauquelinia*; 7) *Spiraea*, *Aruncus*, and *Holodiscus*; 8) *Adenostoma*, *Chamaebatiaria*, and *Sorbaria*.

Several differences between our results and those of Morgan et al. (1994) merit discussion. The first concerns the deep branches

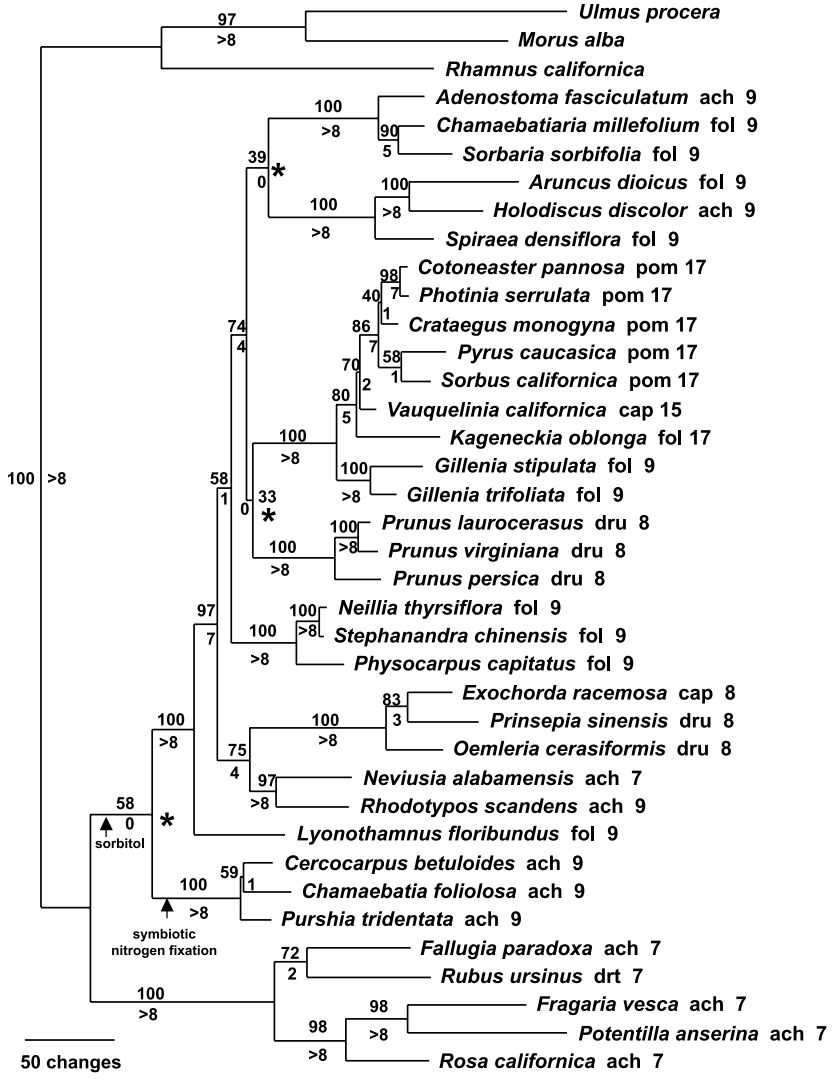

Fig. 1. One of four most parsimonious trees (Length = 2472, CI excluding autapomorphies = 0.58, RI = 0.71) from phylogenetic analysis of *mat*K and *trn*L-*trn*F sequences of selected Rosaceae and outgroups. Bootstrap values appear above or to the left of, and decay index values below or to the right of, the branches. Branches with decay index values of 0 are also marked with asterisks; they were not present in the strict consensus tree. Fruit types (ach = achene; cap = capsule; dru = drupe; drt = drupelet; fol = follicle; pom = pome) and base chromosome numbers are indicated after species names. The branches along which it is inferred that the ability to accumulate sorbitol and to form symbiotic relationships with nitrogen-fixing bacteria arose are indicated

within Rosaceae. The *rbc*L data provided weak support (decay index 2) for monophyly of a group comprised of all taxa except Rosoideae *sensu stricto*, i.e., including taxa traditionally classified in Spiraeoideae, Amygdaloideae, and Maloideae, plus several traditionally classified in Rosoideae but with $x = 9$. Most of the deep branches in that group were weakly supported. Nested within this large group was a weakly supported (decay index 1) clade comprised of *Lyonothamnus* and the actinorhizal group. In contrast, in our study, the position of the actinorhizal clade varied among the most parsimonious trees, appearing as either sister to Rosoideae *sensu stricto* or as sister to all members of the family *except* Rosoideae *sensu stricto*. The former topology was recovered in some or all trees resulting from analyses that included the *mat*K data (alone or in combination with the *trn*L-*trn*F data) and certain outgroup combinations. Although this relationship is consistent with the classification of these taxa in Rosoideae (Schulze-Menz 1964) based on fruit type and floral characters, a variety of data, including *trn*L-*trn*F and *rbc*L sequences and several non-molecular characters (Morgan et al. 1994; see below) support the inclusion of the actinorhizal clade in the spiraeoid/amygdaloid/maloid group. We therefore favor the topology shown in Fig. 1, in which the actinorhizal clade appears as the first branch in the spiraeoid/amygdaloid/maloid clade, but we acknowledge that this requires further study. In any case, our data strongly suggest that the actinorhizal group diverged from the other lineages early in the evolution of the family.

We did find strong support for monophyly of a group including all taxa in the family except Rosoideae *sensu stricto* and the actinorhizal clade, and for the sister position of *Lyonothamnus* to all other taxa within that group. Neither of these relationships was recovered in the *rbc*L analysis. We concur with Morgan et al. (1994) that *Lyonothamnus*, a monotypic genus found only on the Channel Islands of California, appears to be an isolated taxon within the family and its divergence near the base of the spiraeoid/amygdaoid/maloid clade makes sense in this regard. This view is consistent with the classification system of Takhtajan (1997), who placed the genus in its own subfamily Lyonothamnoideae.

The sister relationship of *Sorbaria* and *Chamaebatiaria*, and that of *Adenostoma* to those two genera, were strongly supported in our analysis, whereas relationships among those three genera were unresolved in the *rbc*L study. Our result is consistent with tribal classifications (Hutchinson 1964, Takhtajan 1997) that unite *Sorbaria* and *Chamaebatiaria*, whose members produce follicles, in tribe Sorbarieae, and place *Adenostoma*, which produces achenes, in its own tribe, Adenostomateae.

Finally, the relative positions of *Prunus* and the other drupe-producing genera in the family, and of *Exochorda*, differ between our analysis and that of Morgan et al. (1994). Our results, like those of Morgan et al.'s (1994) *rbc*L study and Lee and Wen's (2001) phylogenetic analysis of ITS sequences of Amygdaloideae, strongly supported a clade including *Oemleria*, *Prinsepia*, and *Exochorda*. The *rbc*L data rather weakly supported a sister relationship between that clade and *Prunus*, a relationship supported also by their shared base chromosome number ($x = 8$; Goldblatt 1976) and wood anatomy (Zhang 1992). Because Lee and Wen's (2001) ITS analysis included only taxa traditionally classified in Amygdaloideae plus *Exochorda* and one ougroup (*Lyonothamnus*), it did not provide a test of the monophyly of the group including *Prunus* plus *Oemleria*/*Prinsepia*/*Exochorda*. The results of our analyses place the latter clade sister to *Rhodotypos* and *Neviusia*, with moderate support. The position of *Prunus* varied among the four most parsimonious trees in our study, appearing either as sister to the clade including Maloideae *sensu lato* (Fig. 1) or as sister to *Adenostoma*/*Chamaebatiaria*/*Sorbaria*, but we found no support at all for a close relationship between *Prunus* and *Oemleria*/*Prinsepia*/*Exochorda*.

Evolution of particular characters. Soltis et al. (1995) suggested that the predisposition

for nitrogen fixation may have been present in the common ancestor of all members of the Eurosid I (Angiosperm Phylogeny Group 1998) clade, although members of only ten extant families actually form these relationships. Four of these families (Rosaceae, Rhamnaceae, Elaeagnaceae, Celtidaceae) fall within the clade now designated as Rosales (Angiosperm Phylogeny Group 1998); in all of these except Celtidaceae (in which *Parasponia* spp. are nodulated by *Rhizobium* spp.), nitrogen fixation occurs in root nodules as a result of a symbiosis involving the host plant and members of the actinomycete genus *Frankia*. Our results strongly support the monophyly of the actinorhizal Rosaceae. Although we included only three of the five genera involved, this conclusion was also strongly supported in phylogenetic analyses of *rbc*L sequences by Swensen and Mullin (1997), who included the other two (*Cowania* and *Dryas*). Thus, our phylogenetic trees suggest that symbiotic nitrogen fixation evolved only once in Rosaceae (Fig. 1); however, given the lack of resolution, discussed above, among the actinorhizal clade, Rosoideae *sensu stricto*, and the rest of the family, we cannot dismiss the possibility that the ability to form this symbiosis was present in the common ancestor of Rosaceae and lost early in the evolution of the family. More thorough studies of relationships among actinorhizal and non-actinorhizal members of the entire order Rosales should help resolve this question.

As discussed above, Morgan et al. (1994) found weak support (decay index 2), for monophyly of a group including all taxa except Rosoideae *sensu stricto*; this group was supported in two of the four most parsimonious trees in our analysis and received 58% bootstrap support. This group is distinguished not only by the higher base chromosome numbers, but also by a number of ecological and biochemical characters (Morgan et al. 1994), such as the ability to accumulate and transport the sugar alcohol sorbitol (Zimmermann and Ziegler 1975, Wallaart 1980). Accumulation of sugar alcohols has not been reported from any other family in Rosales (Zimmerman and Ziegler 1975). As explained above, we favor the position of the actinorhizal clade illustrated in Fig. 1, which is consistent with the hypothesis that sorbitol synthesis evolved once within the family (Fig. 1).

Morgan et al. (1994) stated that their *rbc*L-based phylogeny was consistent with achenes, follicles, or drupes having been the ancestral fruit type for the family, with each of these other types having evolved at least twice. In Fig. 1, it is most parsimonious to consider either achenes or follicles as ancestral; an additional step would be required if any other fruit type were ancestral. This conclusion is based on the assumption that transitions among all of these fruit types are equally likely. Regardless of which fruit type was ancestral, our trees suggest that the achenes of some taxa traditionally classified in Rosoideae (e.g., *Adenostoma*) evolved independently (perhaps as a reversal) of those in Rosoideae *sensu stricto*. Similarly, capsules and follicles each may have evolved two or more times, and our results suggest that the drupes of *Oemleria* and *Prinsepia* evolved independently of those in *Prunus*. In contrast, the pome, the unique fruit type that characterizes Maloideae *sensu stricto*, appears to have evolved just once within the family. Our data support the hypothesis (Morgan et al. 1994) that the pome was derived from a spiraeoid follicle (as in *Gillenia*). Because there is considerable diversity, among pomes, in the degree of carpel connation and of ovary adnation to the hypanthium (Rohrer et al. 1991, 1994), determination of the evolutionary sequence in which carpel fusion and development of the fleshy hypanthium occurred, and of whether or not one or both of those occurred independently in different lineages, require more thorough sampling of Maloideae than is presented here.

Origin of Maloideae. Our data are consistent with *rbc*L sequences in suggesting that taxa with $x = 9$ traditionally classified in Spiraeoideae are sister to Maloideae plus

Spiraeoideae with $x = 15$ and 17. The particular taxa involved are different, however. Whereas Morgan et al. (1994) found that the sister group to the Maloideae *sensu lato* was the clade including *Sorbaria, Chamaebatiaria*, and *Adenostoma*, our analysis strongly supported a sister relationship between *Gillenia* and Maloideae *sensu lato*, with *Adenostoma*, *Sorbaria*, and *Chamaebatiaria* more distantly related. *Gillenia* was not included in the *rbc*L study, but the similarity in carpel anatomy of members of that genus to Maloideae (and, interestingly, also to *Sorbaria*) has been noted (Saunders 1927, Sterling 1966). A sister relationship between *Gillenia* and Maloideae *sensu lato* was also found by Evans and Dickinson (1999c) in an analysis of combined molecular and non-molecular data. Since the results presented here, like those of Morgan et al. (1994), are based on maternally inherited markers, they are not inconsistent with the hypothesis that the polyploidization event that gave rise to Maloideae *sensu lato* involved hybridization between two evolutionarily distinct lineages. We can, however, conclude that *Gillenia* is the closest extant descendant of at least one of the lineages involved in that event.

Figure 1 suggests a New World origin of Maloideae. *Gillenia* and *Vauquelinia* are both distributed within North America, while *Kageneckia* comprises three Chilean species. The two species of *Lindleya*, another genus traditionally placed in Spiraeoideae with a base chromosome number of 17, which was not included in this study but appeared sister to *Vauquelinia* plus Maloideae *sensu stricto* in the *rbc*L analysis (Morgan et al. 1994), are found in Mexico.

Taxonomic Implications. Some of the clades resolved in this study and that of Morgan et al. (1994) are equivalent to taxonomic groups recognized in the classification of the family proposed by Takhtajan (1997). For example, he included *Kageneckia*, *Vauquelinia*, and *Lindleya* in subfamily Pyroideae (Maloideae), *Kerria*, *Neviusia*, and *Rhodotypos* in Kerrioideae, and *Lyonothamnus* alone in Lyonothamnoideae. On the other hand, he maintained all of the following in Spiraeoideae: *Physocarpus*, *Neillia*, *Stephanandra* (together forming tribe Neillieae), *Spiraea*, *Aruncus* (both in tribe Spiraeeae), *Holodiscus* (alone in tribe Holodisceae), *Sorbaria*, *Chamaebatiaria* (tribe Sorbarieae), *Adenostoma* (alone in tribe Adenostomateae) and *Gillenia* (in tribe Gillenieae). Both our results and those of Morgan et al. (1994) suggest that this circumscription of the subfamily, while less diverse than the traditional Spiraeoideae, is still a polyphyletic assemblage. Our results also indicate that Maloideae could be further expanded to include *Gillenia*, a circumscription that we support since it would maintain monophyletic taxa without requiring classification of *Gillenia* alone in its own subfamily.

Our sampling within Rosoideae *sensu stricto* was not thorough enough to test Takhtajan's (1997) taxonomic treatment of most of the genera in that group. As discussed above, some of our trees supported a sister relationship of a clade comprised of the actinorhizal Rosaceae to all members of Rosoideae *sensu stricto* examined here (one species each of *Fallugia*, *Fragaria*, *Potentilla*, *Rosa*, and *Rubus*). Our data, however, provide no support for Takhtajan's (1997) classification of the actinorhizal genera (*Cercocarpus*, *Chamaebatia*, *Cowania*, *Dryas*, and *Purshia*) within Potentilloideae, since his circumscription of that subfamily includes *Potentilla*, *Fragaria*, and *Fallugia*, but not *Rubus* and *Rosa*.

As discussed above, the positions of *Prunus* and other genera with $x = 8$ have varied among analyses to date. Recognition of a taxonomic group including *Oemleria*, *Exochorda*, and *Prinsepia*, each of which is placed in its own tribe by Takhtajan (1997), seems warranted. Such a group has not been recognized in any existing treatment but is strongly supported by molecular data. Whether or not that group should be united with *Prunus* (as in Takhtajan's Amygdaloideae) requires further investigation, but it appears that there is little support from molecular data for such a taxon.

We reached a similar conclusion using ITS and *trn*L-*trn*F data from a subset of the taxa included here (Bortiri et al. 2001).

The weak support for some relationships indicates that additional evidence is required before final decisions can be made about higher-level classification of the family. Some, but not all, of the subfamilies and/or tribes recognized by Takhtajan (1997) appear to represent monophyletic groups, based on both the current study and results from analysis of *rbc*L sequences (Morgan et al. 1994). The issue of how many subfamilial ranks should be recognized will be influenced by the degree to which resolution among strongly supported clades can be improved in future analyses. The branching order near the base of our trees was, in general, more strongly supported than in the *rbc*L study, but the relationships among several of the well-supported clades remain poorly resolved. The low support values on the internal branches are due to a combination of homoplasy in the data, which might be attributable to a complex history perhaps involving reticulation among lineages, and relatively short branch lengths, which may reflect rapid radiation of major lineages within the family. If resolution remains poor, it may be most appropriate to recognize only one supergeneric rank (e.g., tribes, as in Hutchinson (1964)), which would best reflect the fact that there are a number of well-supported groups but the relationships among them are not well understood. On the other hand, if strong resolution among clades is eventually obtained by integrating multiple data sets, this resolution should be incorporated into the hierarchical structure of the taxonomy of the family.

Technical assistance from Patrick Cheung, Dean Lavelle, Millie Lee, Teri Nguyen, and Mindy Sin is gratefully acknowledged. The following individuals assisted us in obtaining plant material: S. Arulsekar, Phillip Ashby, Wei Cao, Holly Forbes, Mary Hirshfeld, Jim Luby, and Warren Roberts. Kenneth Sytsma, Kenneth Robertson, and one anonymous reviewer provided many helpful comments on the manuscript. Financial support was provided by the Agricultural Experiment Station of the University of California.

References

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

Bortiri E., Oh S., Jiang J., Baggett S., Granger A., Weeks C., Buckingham M., Potter D., Parfitt D. (2001) Phylogeny and systematics of the genus *Prunus* L. (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trn*L-*trn*F spacer DNA. Syst. Bot. 26: 797–807.

Boyer S. K., Mullet J. E. (1988) Pea chloroplast tRNA-Lys (UUU) gene: transcription and analysis of an intron-containing gene. Photosyn. Res. 17: 7–22.

Campbell C. S., Donoghue M. J., Baldwin B. G., Wojciechowski M. F. (1995) Phylogenetic relationships in Maloideae (Rosaceae): evidence from sequences of the internal transcribed spacers of nuclear ribosomal DNA and its congruence with morphology. Amer. J. Bot. 82: 903–918.

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

Dickson E. E., Arumuganathan K., Kresovich S., Doyle J. J. (1992) Nuclear DNA content variation within the Rosaceae. Amer. J. Bot. 79: 1081–1086.

Doyle J. J., Doyle J. L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

Eriksson T., Donoghue M. J., Vretblad M. (1998) Phylogenetic analysis of *Potentilla* using DNA sequences of nuclear ribosomal internal tran-

scribed spacers (ITS), implications for the classificaiton of Rosoideae (Rosaceae). Plant. Syst. Evol. 211: 155–179.

Evans R. C., Alice L. A., Campbell C. S., Kellogg E. A., Dickinson T. A. (2000) The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. Mol. Phylogenet. Evol. 17: 388–400.

Evans R. C., Campbell C. S. (2000) [abstract] A multigene tree of the Rosales. Amer. J. Bot. 87 (6 Suppl.): 125–126.

Evans R. C., Dickinson T. A. (1997) [abstract] Phylogenetic relationships in the Rosaceae: a total evidence approach. Amer. J. Bot. 84 (6 suppl.): 190.

Evans R. C., Dickinson T. A. (1999a) Floral ontogeny and morphology in subfamily Amygdaloideae T & G (Rosaceae). Int. J. Pl. Sci. 160: 955–979.

Evans R. C., Dickinson T. A. (1999b) Floral ontogeny and morphology in subfamily Spiraeoideae Endl. (Rosaceae). Int. J. Pl. Sci. 160: 981–1012.

Evans R. C., Dickinson T. A. (1999c) [abstract] Rosaceae floral evolution: hypothesis from analysis of combined molecular and non-molecular data. Abstracts from the XVI International Botanical Congress, August 1–7, 1999, St. Louis, MO, p. 39.

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

Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Gielly L., Taberlet P. (1996) A phylogeny of the European gentians inferred from chloroplast *trn*L (UAA) intron sequences. Bot. J. Linn. Soc. 120: 57–75.

Gladkova V. N. (1972) On the origin of subfamily Maloideae. Bot. Zhurn. 57: 42–49.

Goldblatt P. (1976) Cytotaxonomic studies in the tribe Quillajeae (Rosaceae). Ann. Missouri Bot. Gard. 63: 200–206.

Hebda R. J., Chinnappa C. C. (1994) Studies on pollen morphology of Rosaceae. Acta Bot. Gallica 141: 183–193.

Heywood V. H. (1993) Flowering plants of the world. Oxford University Press, New York.

Hu J.-M., Lavin M., Wojciechowski M. F., Sanderson M. J. (2000) Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trn*K/*mat*K sequences and its implications for evolutionary patterns in Papilionoideae. Amer. J. Bot. 87: 418–430.

Hutchinson J. (1964) The genera of flowering plants, vol. 1. Clarendon Press, Oxford, U.K.

Kalkman C. (1988) The phylogeny of the Rosaceae. Bot. J. Linn. Soc. 98: 37–59.

Källersjö M., Farris J. S., Chase M. W., Bremer B., Fay M. F., Humphries C. J., Petersen G., Sebert O., Bremer K. (1998) Simultaneous parsimony jackknife analysis of 2358 *rbc*L DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. Plant. Syst. Evol. 213: 259–287.

Karol K. G., Suh Y., Schatz G. E., Zimmer E. A. (2000) Molecular evidence for the phylogenetic position of *Takhtajania* in the Winteraceae: inference from nuclear ribosomal and chloroplast gene spacer sequences. Ann. Missouri Bot. Gard. 87: 414–432.

Lee, S., Wen. J. (2001) A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. Amer. J. Bot. 88: 150–160.

Mishler B. D., Donoghue M. J., Albert V. A. (1991) [abstract] The decay index as a measure of relative robustness within a cladogram. Tenth Meeting of the Willi Hennig Society, Toronto.

Morgan D. R., Soltis D. E., Robertson K. R. (1994) Systematic and evolutionary implications of *rbc*L sequence variation in Rosaceae. Amer. J. Bot. 81: 890–903.

Phipps J. B., Robertson K. R., Rohrer J. R., Smith P. G. (1991) Origins and evolution of subfam. Maloideae (Rosaceae). Syst. Bot. 16: 303–332.

Potter D. (1997) [abstract] Variation and phylogenetic relationships of nucleotide sequences for putative genes encoding polygalacturonase inhibitor protein (PGIPs) among taxa of Rosaceae. Amer. J. Bot. 84 (6 Suppl.): 224.

Potter D., Gao F., Oh S., Baggett S. (1999) [abstract] Molecular phylogenetic studies in Rosaceae. Abstracts from the XVI International Botanical Congress, August 1–7, 1999, St. Louis, MO, p. 39.

Potter D., Luby J. J., Harrison R. E. (2000) Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Syst. Bot. 25: 337–348.

Robertson K. R. (1974) The genera of the Rosaceae in the southeastern United States. J. Arnold Arbor. 55: 303–332, 344–401, 611–662.

Rohrer J. R., Robertson K. R., Phipps J. B. (1991) Variations in structure among fruits of Maloideae (Rosaceae). Amer. J. Bot. 78: 1617–1635.

Rohrer J. R., Robertson K. R., Phipps J. B. (1994) Floral morphology of Maloideae (Rosaceae) and its systematic relevance. Amer. J. Bot. 81: 574–581.

Saunders E. R. (1927) On carpel polymorphism. II. Ann. Bot. 41: 569–627.

Sax K. (1933) The origin of the Pomoideae. Proc. Amer. Soc. Hort. Sci. 30: 147–150.

Schulze-Menz G. K. (1964) Rosaceae. In: Melchior H. (ed.) Engler's Syllabus der Pflanzenfamilien II. 12th edn. Gebrüder Borntraeger, Berlin, pp. 209–218.

Scott K. D., McIntyre C. L., Playford J. (2000) Molecular analyses suggest a need for a significant rearrangement of Rutaceae subfamilies and a minor reassessment of species relationships within *Flindersia*. Plant. Syst. Evol. 223: 15–27.

Soltis D. E., Soltis P. S., Morgan D. R., Swensen S. M., Mullin B. C., Dowd J. M., Martin P. G. (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl. Acad. Sci. USA 92: 2647–2651.

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, *rbc*L, and *atp*B sequences. Bot. J. Linn. Soc. 133: 381–461.

Sterling C. (1966) Comparative morphology of the carpel in the Rosaceae. VIII. Spiraeoideae: Quillajeae, Sorbarieae. Amer. J. Bot. 53: 951–960.

Sugita M., Shinozaki K., Sugura M. (1985) Tobacco chloroplast $tRNA^{Lys}$ (UUU) gene contains a 2.5-kilobase-pair intron: an open reading frame and a conserved boundary sequence in the intron. Proc. Natl. Acad. Sci. U.S.A. 82: 3557–3561.

Swensen S. M., Mullin B. C. (1997) The impact of molecular systematics on hypotheses for the evolution of root nodule symbioses and implications for expanding symbioses to new host plant genera. Pl. & Soil 194: 185–192.

Swofford D. L. (2000) "PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods)," Version 4. Sinauer Associates, Sunderland, MA.

Taberlet P., Gielly L., Pautou G., Bouvet J. (1991) Universal primers for amplification of three noncoding regions of chloroplast DNA. Pl. Mol. Biol. 17: 1105–1109.

Takhtajan A. (1997) Diversity and classification of flowering plants. Columbia Univ. Press, New York.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., Higgins D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25: 4876–4882.

Wallaart R. A. M. (1980) Distribution of sorbitol in Rosaceae. Phytochemistry 19: 2603–2610.

Xiang Q.-Y., Soltis D. E., Soltis P. S. (1998) Phylogenetic relationships of Cornaceae and close relatives inferred from *mat*K and *rbc*L sequences. Amer. J. Bot. 85: 285–297.

Zhang S.-Y. (1992) Systematic wood anatomy of the Rosaceae. Blumea 37: 81–158.

Zimmermann M. H., Ziegler H. (1975) List of sugars and sugar alcohols in sieve-tube exudates. In: Zimmermann, M. H., Milburn, J. A. (eds.) Transport in plants. I. phloem transport. Springer, Heidelberg, pp. 480–502.

Addresses of the authors: Daniel Potter (e-mail: dpotter@ucdavis.edu), Fangyou Gao, Pedro Esteban Bortiri, Sang-Hun Oh, Department of Pomology, 1045 Wickson Hall, University of California, One Shields Avenue, Davis, California 95616, USA. Scott Baggett, Ph.D. Program Biology, Lehman College, City University of New York, 250 Bedford Park Boulevard West, Bronx, New York 10458, USA.