

PHYLOGENETIC RELATIONSHIPS OF CORNACEAE AND CLOSE RELATIVES INFERRED FROM *MATK* AND *RBCL* SEQUENCES¹

QIU-YUN XIANG,² DOUGLAS E. SOLTIS, AND PAMELA S. SOLTIS

Department of Botany, Washington State University, Pullman, Washington 99164–4238

Phylogenetic relationships were inferred using nucleotide sequences of the chloroplast gene *matK* for members of Cornales, a well-supported monophyletic group comprising Cornaceae and close relatives. The shortest trees resulting from this analysis were highly concordant with those based on previous phylogenetic analysis of *rbcL* sequences. Analysis of a combined *matK* and *rbcL* sequence data set (a total of 2652 bp [base pairs]) provided greater resolution of relationships and higher internal support for clades compared to the individual data sets. Four major clades (most inclusive monophyletic groups) of Cornales are indicated by both sets of genes: (1) *Cornus-Alangium*, (2) nyssoids (*Nyssa-Davidia-Camptotheca*)-mastixioids (*Mastixia, Diplopanax*), (3) *Curtisia*, and (4) Hydrangeaceae-Loasaceae. The combined evidence indicates that clades 2 and 3 are sisters, with clade 4 sister to the remainder of Cornales. These relationships are also supported by other lines of evidence, including synapomorphies in fruit and pollen morphology and gynoecial vasculature. Comparisons of *matK* and *rbcL* sequences based on one of the most parsimonious *rbcL-matK* trees indicate that *matK* has a much higher A-T content (66.9% in *matK* vs. 55.8% in *rbcL*) and a lower transition:transversion ratio (1.23 in *matK* vs. 2.21 in *rbcL*). The total number of nucleotide substitutions per site for *matK* is 2.1 times that of *rbcL* in Cornales. These findings are similar to recent comparisons of *matK* and *rbcL* in other dicots. Variable sites of *matK* are almost evenly distributed among the three codon positions (1.0:1.0:1.3), whereas variable sites of *rbcL* are mostly at the third position (1.8:1.0:7.5). Among-lineages rates of nucleotide substitutions in *rbcL* are basically homogeneous throughout Cornales, but are more heterogeneous in *matK*.

Key words: Cornales; *matK*; molecular evolution; molecular phylogeny; *rbcL*.

Cornaceae sensu lato represent one of the taxonomically most problematic families of flowering plants. The composition and close allies of the family are extremely controversial, and relationships among genera ascribed to Cornaceae are poorly understood (Xiang et al., 1993; Xiang and Soltis, in press). For example, from one to as many as 17 diverse genera (*Cornus, Alangium, Aralidium, Aucuba, Camptotheca, Corokia, Curtisia, Davidia, Diplopanax, Garrya, Grisilinia, Helwingia, Kaliphora, Mastixia, Melanophylla, Nyssa, and Toricellia*) have been placed in Cornaceae by various authors at one time or another. The taxonomic history of Cornaceae was reviewed in detail by Xiang et al. (1993) and Xiang and Soltis (in press). Recent phylogenetic analyses of sequence data for the chloroplast gene *rbcL* (Xiang et al., 1993; Xiang and Soltis, in press) have greatly improved our understanding of the circumscription and affinities of Cornaceae. These studies revealed that many genera (*Aralidium, Aucuba, Corokia, Garrya, Grisilinia, Helwingia, Kaliphora, Melanophylla, and Toricellia*) previously placed in Cornaceae by some authors are only distantly related to *Cornus*. Analyses of *rbcL* sequence

data also suggested the existence of a cornaceous clade that consists of *Cornus, Alangium*, nyssoids (*Nyssa, Davidia, and Camptotheca*), mastixioids (*Diplopanax, Mastixia, Curtisia*, Hydrangeaceae, and Loasaceae (referred to as Cornales hereafter). This cornaceous clade, excluding Hydrangeaceae and Loasaceae, closely corresponds to Eyde's circumscription of Cornaceae (Eyde, 1988; Eyde and Xiang, 1990) followed by Thorne (1992), which consists of six of the above eight genera (*Cornus, Alangium, Nyssa, Davidia, Camptotheca, Diplopanax, and Mastixia*). Within this clade, four most inclusive monophyletic lineages were identified: (1) *Cornus-Alangium*, (2) nyssoids-mastixioids, (3) *Curtisia*, and (4) Hydrangeaceae-Loasaceae. However, relationships among these four major lineages remained unresolved by *rbcL* sequence data due to insufficient sequence variation in *rbcL*.

Chloroplast genes that evolve more rapidly than *rbcL* can be extremely useful for clarifying relationships at taxonomic levels below that of family. Recent studies (Johnson and Soltis, 1994, 1995; Plunkett, 1994; Steele and Vilgalys, 1994) have demonstrated the utility of *matK* for resolving generic and even species-level relationships. The chloroplast gene *matK* encodes a maturase and is part of the *trnK* intron (Neuhaus and Link, 1987; reviewed in Johnson and Soltis, 1995). Johnson and Soltis (1994, 1995) not only found that *matK* evolves approximately three times faster than *rbcL* in Saxifragaceae sensu stricto (s. s.), but also showed that analysis of *matK* sequence data provided fine-scale resolution of relationships within the family comparable to that achieved via restriction site analysis of cpDNA. Generic and species relationships have been resolved in Polemoniaceae using *matK* sequence data (Steele and Vilgalys, 1994; Johnson

¹ Manuscript received 22 March 1996; revision accepted 17 June 1997.

The authors thank L. Hufford, M. W. Chase, K. A. Kron, R. G. Olmstead, D. E. Boufford, I. Hey, B. Bartholomew, T. M. Hardig, Z.-H. Ji, H. Sun, S. Yonkowsky, S. J. Brunsfeld, T. Edwards, The Arnold Arboretum, Missouri Botanical Garden, The Royal Botanical Gardens, Kew, Strybing Arboretum, and the U. S. National Arboretum for providing DNA samples and leaf material, B. S. Gaut for providing comments and a computer program for conducting the relative rate tests. The project was supported in part by NSF grant DEB9307000.

² Author for correspondence, current address: Department of Biological Sciences, Idaho State University, Pocatello, ID 83209–8007 (FAX: 208–236-4570).

and Soltis, 1995; Johnson et al., in press). We therefore used comparative *matK* sequencing to elucidate further relationships within Cornales.

MATERIALS AND METHODS

DNA isolation and sequencing—Total DNAs were isolated from fresh or silica-dried leaf tissue via the miniprep procedure of Doyle and Doyle (1987) as modified by Soltis et al. (1991). Double-stranded DNAs of *matK* were amplified using the polymerase chain reaction (PCR) with *trnK*-3914F and *trnK*-2R as PCR primers. Although generally successful, this primer combination failed to generate double-stranded products for *Deutzia*, *Petalonyx*, and *Eucnide*; *matK*-2000R was therefore used in place of *trnK*-2R for these three genera. The double-stranded PCR products were then used as templates for asymmetric amplification of single-stranded DNA using only one PCR primer. Single-stranded DNAs were then cleaned and sequenced following the general methods used for *rbcL* (e.g., Morgan and Soltis, 1993; Xiang et al., 1993). In general, four sequencing primers, *matK*-1168R, *matK*-1470R, *matK*-1176F, and *matK*-1412F, were used to obtain sequence data for approximately two-thirds of the gene *matK*. Where these primers proved ineffective, presumably because of primer divergence, one or more of the following primers were also used: *matK*-710F, *matK*-1300F, *matK*-1235R, *matK*-1412R, *matK*-1560R, and *matK*-1848R. All primers were from Johnson and Soltis (1995), except *matK*-1300F (GAT GC[T/C] TCT TCT TTG GAT T), which was designed for this study; *trnK*-2R was designed by Kelly Steele, and *trnK*-3914F was originally designed by Jerry Learn and modified by Johnson and Soltis (1995).

We obtained *matK* sequences for 27 taxa (Table 1), 24 of which are members of the *rbcL*-based Cornales; three outgroup taxa also were included. The cornaleous genera sampled include: *Cornus*, *Alangium*, *Nyssa*, *Davidia*, *Camptotheca*, *Mastixia*, *Curtisia* (genera that have been previously placed in Cornaceae), *Hydrangea*, *Schizophragma*, *Deutzia*, *Philadelphus* (representatives of Hydrangeaceae), *Petalonyx*, and *Eucnide* (representatives of Loasaceae) (Table 1). This sampling represents all of the genera and families of the *rbcL*-based Cornales and all have been analyzed previously for *rbcL* (Xiang et al., 1993; Hempel and Jansen, 1994; Soltis, Xiang, and Hufford, 1995). Given the demonstrated monophyly of both Hydrangeaceae and Loasaceae (Soltis, Xiang, and Hufford, 1995; Hempel et al., 1995), only selected representatives of each family were included here. One species was sequenced for each genus for *matK* except for *Cornus* sensu lato (s. l.) for which 12 species were analyzed that represent well the diversity of clades present in the genus (see Murrell, 1993; Xiang et al., 1993; Xiang and Soltis, in press; Xiang et al., 1996). *Sarracenia*, *Fouquieria*, and *Roridula* were also sequenced for *matK* and used as outgroups because they are phylogenetically close to Cornales based on the global analysis of *rbcL* sequences of seed plants (Chase et al., 1993); they were also used as outgroups in the phylogenetic analysis of *rbcL* sequences for Cornaceae and Hydrangeaceae (Soltis, Xiang, and Hufford, 1995; Xiang and Soltis, in press).

Phylogenetic analyses—The *matK* sequence data matrix contains 27 taxa each of 1212 base pairs (bp) (1098 bp from the 5' coding region of *matK* that covers approximately two-thirds of the gene and 114 bp from the *trnK* intron flanking the 5' end of *matK*) with eight small alignment gaps in the coding region (see Results; Table 2). Phylogenetic analyses of several data sets were conducted using PAUP version 3.1.1 (Swofford, 1993): (1) all of the aligned 1212 bp (base pairs) were included with gaps coded as missing data points but sequences across are gaps included; (2) only the coding region of the gene (1098 bp) was included; (3) sequences across the alignment gaps were excluded; (4) gap sequences excluded but each indel coded as a binary character. For all analyses, heuristic searches were performed with MULPARS, random taxon addition with 100 replicates, and tree-bisection-reconnection (TBR) branch-swapping, with character states specified as unordered

TABLE 1. Sources of plant materials sequenced for *matK*.

Species	Sources and location of vouchers	Genebank accession numbers
<i>Alangium platanifolium</i> (Sieb. & Zucc.) Harms	Soltis 2543, Japan	U96880
<i>Camptotheca acuminata</i> Decne.	Strybing Arboretum #74-180.	U96888
<i>Cornus alternifolia</i> L.	Xiang et al. 167, WS.	U96889
<i>Cornus canadensis</i> L.	Xiang et al. 198, WS.	U96890
<i>Cornus capitata</i> Wall.	Strybing Arboretum XY-2080.	U96891
<i>Cornus chinensis</i> Wanger.	Roy. Bot. Gard., Kew #604.88.08692.	U96892
<i>Cornus controversa</i> Hemsley	Arnold Arboretum #20458, WS.	U96893
<i>Cornus florida</i> L.	Xiang 250, WS.	U96894
<i>Cornus macrophylla</i> Wall.	Arnold Arboretum #577-51-A, WS.	U96895
<i>Cornus mas</i> L.	Arnold Arboretum #748-75, WS.	U96896
<i>Cornus nuttallii</i> Audubon	Arnold Arboretum #573-73-A, WS.	U96897
<i>Cornus obliqua</i> Raf.	Xiang et al. 173, WS.	U96898
<i>Cornus oblonga</i> Wall.	Sun, s.n., Bot. Gard. Kunming, China.	U96899
<i>Cornus sessilis</i> Torr. ex Durand	Soltis, California 5/3/1988.	U96900
<i>Curtisia dentata</i> (Burm.) G. A. Sm.	Edwards 918, NU.	U96901
<i>Duetzia rubens</i> Rehder	Arnold Arboretum #1003-86-Mass.	U96884
<i>Davidia involucreta</i> Baill.	U.S. National Arboretum, #12067.	U96885
<i>Eucnide uvens</i> (A. Gray) Parry	Hufford 552, WS.	U96902
<i>Fouquieria splendens</i> Engelm.	Matthaei Bot. Gard. #860162.	U96893
<i>Hydrangea quercifolia</i> Bartram	Missouri Bot. Gard. #821044.	U96882
<i>Mastixia caudatilimba</i> C. Y. Wu ex Soong	Zan-he Ji, s.n., WS.	U96887
<i>Nyssa ogeche</i> Marsh	U.S. National Arboretum, s.n.	U96886
<i>Petalonyx nitidus</i> S. Wats.	Hufford 554, DUL.	U96904
<i>Philadelphus hirsutus</i> Nutt.	Arnold Arboretum #320-79-4.	U96881
<i>Roridula</i> sp.	Soltis s.n.	U96905
<i>Sarracenia purpurea</i> L.	Xiang 252, WS.	U96906
<i>Schizophragma hydrangeoides</i> Sieb et Zucc.	Soltis 2516, Japan.	U96883

and equally weighted. To evaluate relative robustness of the clades found in the most parsimonious trees, bootstrap and decay analyses (Felsenstein, 1985; Bremer, 1988; Donoghue et al., 1992) were conducted. Because all four analyses resulted in highly similar trees (see Results), bootstrap and decay analyses were performed only for the first of the four data sets above. The bootstrap analysis (100 replicates) was conducted with random taxon addition (ten replicates per bootstrap replicate), TBR branch-swapping, and MULPARS. Decay values were obtained by examining trees up to five steps longer than the shortest following the general method of Johnson and Soltis (1994). This approach involves using a strict consensus of all most parsimonious trees as a constraint tree for heuristic searches that save only trees of one additional step that fail to satisfy the constraint topology.

Recent analyses of combined data sets (e.g., Soltis et al., 1993, 1996; Johnson and Soltis, 1994, 1995; Plunkett, 1994; Olmstead and Sweere, 1995) all provided more resolution and internal support for relationships than did the individual data sets. Thus, since initial analyses of *matK*

sequences for Cornales provided results highly concordant with those achieved via analysis of *rbcL* sequences, phylogenetic analyses of two combined *rbcL-matK* data sets were also conducted for Cornales to obtain a comprehensive view of relationships (e.g., Kluge, 1989; Donoghue and Sanderson, 1992). The first combined *rbcL-matK* sequence data set consists of 31 taxa that have been sequenced for at least one of the two genes to provide a broad sampling. In this data matrix three taxa were not sequenced for *matK* (*Cornus officinalis*, *Diplopanax*, and *Mentzelia*) and four taxa were not sequenced for *rbcL* (*Cornus sessilis*, *C. capitata*, *C. macrophylla*, and *Eucnide*). The missing sequences were filled as “?” in the data matrix. The second combined *rbcL-matK* sequence data set includes a subset of taxa sequenced for both *rbcL* and *matK*. This data matrix contains 23 taxa with 20 members of Cornales and three outgroups. A total of 2652 bp was analyzed (1440 bp of *rbcL*, 1212 bp of *matK*) for both data sets following the general methods described above.

The *matK* sequences of *Cornus* obtained in this study are the third cpDNA data set generated for the purpose of resolving relationships within this systematically controversial genus. Analyses of the *matK* sequences and the combined *rbcL-matK* sequences of Cornales provided results within *Cornus* similar to those suggested by the previous analyses of cpDNA restriction site data (Xiang et al., 1996) and *rbcL* sequence data (Xiang et al., 1993; Xiang and Soltis, in press). Thus, to obtain a comprehensive view of relationships in *Cornus*, a data matrix of *rbcL* and *matK* sequences, as well as cpDNA restriction sites, was constructed. This combined data set consists of 18 taxa (15 species of *Cornus* and three outgroups, *Alangium*, *Davidia*, and *Nyssa*) and a total of 2894 characters (1440 bp of *rbcL*, 1212 bp of *matK*, and 242 restriction sites). Species of *Cornus* that have at least two of the three possible character sets available were included in the data matrix. Phylogenetic analysis was conducted as above.

Comparing rate and pattern of variation between *matK* and *rbcL*—

The second combined *rbcL-matK* sequence data matrix that contains only taxa sequenced for both *matK* and *rbcL* was used for molecular evolutionary comparisons of the two genes. Nucleotide composition and distribution of variable sites at different codon positions were examined for both *matK* and *rbcL* using MEGA version 1.0 (Kumar, Tamura, and Nei, 1993). The transition:transversion ratio and the total number of nucleotide substitutions per site for both genes were estimated based on one of the most parsimonious trees (chosen at random) resulting from analysis of the data matrix.

Relative rate tests were performed to assess homogeneity of rates of both synonymous and nonsynonymous substitutions at *rbcL* and *matK* within Cornales using the method of Li and Bousquet (1992) and following Gaut et al. (1996). Numbers of synonymous (Ks) and nonsynonymous (Ka) substitutions per site between lineages were estimated using the method of Nei and Gojobori (1986). A computer program that performs relative rate tests of Li and Bousquet (1992) using the distance measures of Nei and Gojobori (1986) was used (kindly provided by B. S. Gaut). The *rbcL* and *matK* sequences of *Fouquieria*, one of the three outgroup taxa used above for the phylogenetic analyses, were chosen as the reference sequences for the relative rate tests among the major lineages of Cornales, and the *rbcL* and *matK* sequences of *Alangium*, the sister group of *Cornus*, were chosen as the reference sequences for tests between lineages within *Cornus*.

RESULTS

matK—The *matK* sequences obtained vary from 852 to 1191 bp in length among the 24 members of Cornales and the three outgroups before alignment. These sequences were aligned easily by hand using the sequence of *Cornus mas* as the reference, which has a total length of 1170 bp (109 bp from the intron flanking the 5' end of *matK* and 1062 bp from the 5' end of *matK* coding re-

gion). The aligned portion of sequence is 1212 bp with eight small alignment gaps, which were inferred as indels (Table 2). Six of the indels appear to be autapomorphies (indels A, B, C, F, G, and H; Table 2); two of these are potentially phylogenetically informative (indels D and E). Because sequencing primer *matK*-1412F was ineffective in *Cornus obliqua*, *C. oblonga*, *C. macrophylla*, *Campylothea*, and *Davidia* this primer was replaced by sequencing primer *matK*-1300F. However, sequences generated when *matK*-1300F was used did not provide a complete sequence for *matK*. As a result, the five taxa noted above were missing ~100 bp of data at the 3' end of the aligned portion of the sequence. *Deutzia* was missing ~60 bp of sequence data in the region between two of the sequencing primers, *matK*-1168R and *matK*-1470R.

Of the 1098 bp of the aligned portion from the protein-coding part of the gene, no stop codons were found; 388 sites (35.5%) are variable and 171 sites (15.6%) are potentially phylogenetically informative. These values could be slightly higher if sequence data were obtained for the missing areas in the several taxa noted above. Of the 114 bp obtained for the noncoding region flanking the 5'-end of *matK*, 40 sites (35.1%) are variable and 16 sites (14.0%) are potentially phylogenetically informative. These values are very similar to those for the *matK* coding region and suggest a similar rate of evolution of this noncoding region and *matK* in Cornales.

For the data matrix used for comparison of *matK* and *rbcL*, 364 (33.2%) sites of the *matK* sequences are variable among the members of Cornales and 153 (13.9%) sites are potentially phylogenetically informative. These values are 2.3 and 1.9 times, respectively, those obtained for the *rbcL* sequences for the same suite of taxa (variable sites: 206/1425 = 14.5%; potentially informative sites: 106/1425 = 7.4%).

The first analysis, which used all of the aligned 1212 bp of *matK* sequence data, found ten most parsimonious trees each of 889 steps and all in one island (Maddison, 1991). The second analysis, which included only the coding region of *matK*, found 12 shortest trees each of 819 steps. The strict consensus of these 12 trees is identical to that resulting from the first analysis except for the portion within *Cornus*, which shows less resolution among the major clades within the genus. The third analysis, which excluded the alignment gaps in the data matrix, and the fourth analysis, which included alignment gaps that were coded as binary characters, both found ten shortest trees identical to those found in the first analysis.

Four major subclades (most inclusive monophyletic lineages) were found within Cornales in all most parsimonious trees resulting from all four analyses; they are identical to those revealed by the *rbcL* sequence analysis (Figs. 1, 2); they are: (1) *Cornus-Alangium*, (2) nyssoids-mastixioids, (3) *Curtisia*, and (4) Hydrangeaceae-Loasaceae. The *Cornus-Alangium* lineage (bootstrap value of 92%, decay value of 5) and Hydrangeaceae-Loasaceae lineage (bootstrap value of 82%, decay value of 3) are both strongly supported by bootstrap and decay analyses. Relationships among these four major lineages remain unresolved, however (Fig. 2). For example, although *Curtisia* appears as sister to the nyssoids and mastixioids in most of the shortest trees (80%), in other most parsi-

TABLE 2. Insertions and deletions (indels A–H) in *matK* sequences inferred in Cornales and outgroups. Dashes represent missing bases associated with indels. Dots in the sequence below the reference taxon (*Cornus mas*) indicate that the same nucleotide present in the reference taxon is also present in the species containing the indel. All species in which an indel occurs are listed and their sequences are given. P/A represents the number of species with nucleotides involved in indels present and absent, respectively. RN, the reference nucleotide, is the position of the bold sequential nucleotide preceding the indel in the sequences of the reference taxon. Starting site of *matK* of the reference sequence corresponds to that of the published sequences of mustard and tobacco (Neuhaus and Link, 1987).

Indel	P/A	Representative sp.	RN	Sequence region
A	26A	<i>Cornus mas</i>	115	T T AAATA-----GATCGATTTTGT
	1P	<i>Curtisia dentata</i>	TAAAAA.....
B	26A	<i>Cornus mas</i>	147	T T ATGAC-----AATAATCCAGCT
	1P	<i>Sarracenia purpurea</i>	TATGAC.....
C	26P	<i>Cornus mas</i>	188	T T TAATTACTCG-----AA
	1P	<i>Sarracenia purpurea</i>	ATTTAATTACTCA..
D	24A	<i>Cornus mas</i>	236	A T CAA--AATCCATTTTGGGGCAT
	2P	<i>Philadelphus hirsutus</i>		.C...CAA.....C
		<i>Deutzia rubens</i>		.C...CGA....C....C...C
E	18A	<i>Cornus mas</i>	565	T G GAATA-----ATCTTATTACTCC
	9P	<i>Cornus florida</i>		C.T...TGGAATA..T.G.....
		<i>Cornus nuttallii</i>		C.T...GGAATA..T.G.....
		<i>Cornus capitata</i>		C.T...GGAATAC.T.....
		<i>Cornus canadensis</i>		CAT...TGGAATC..T..T.....
		<i>Cornus obliqua</i>		C....TGGAATA.....G.....
		<i>Cornus macrophylla</i>		C.TAA.TGGAATA.....G.....
		<i>Cornus oblonga</i>		CATAA.TGGAATA.....
		<i>Cornus alternifolia</i>		CATAA.TGGAATA.....
		<i>Cornus controversa</i>		CATAA.TGGAATA.....
		<i>Cornus mas</i>		GAA A CCCGAGTTCCTCTTTTCAAAAA
1P	<i>Curtisia dentata</i>G.....		
G	26P	<i>Cornus mas</i>	782	A G GTTA---TCCTATGGTTGTTCAAG
	1P	<i>Cornus canadensis</i>		G....TTA.....
H	26A	<i>Cornus mas</i>	782	A G GTTATCCTATGGTTGTTCAAGGAT
	1P	<i>Roridula</i> sp.		..CA.....A.....

monious trees it appears as sister to a large clade containing the *Cornus-Alangium* and Hydrangeaceae-Loasaceae lineages. Hence, the strict consensus tree of all shortest trees shows a tetrachotomy among them.

Although relationships among the major subclades are not clear, relationships within each of the subclades of Cornales are well resolved by *matK* sequence data (Fig. 2). Within the *Cornus-Alangium* lineage, all shortest trees show that *Alangium* is sister to *Cornus* (11 nucleotide substitutions, bootstrap value of 92%, decay value of 5); *Cornus*, in turn, is also a well-supported clade (18, 99%, >5). Within *Cornus*, the cornelian cherries (11, 98%, >5) are sister to the remainder of *Cornus*. The remainder of the genus consists of two subclades: the blue-fruited dogwoods (5, 63%, 1) and the showy-bracted dogwoods (6, 58%, 1). Within the subclade of the blue-fruited dogwoods, all shortest trees placed *C. oblonga* as sister to the remainder of the blue-fruited dogwoods (Fig. 2). The showy-bracted dogwoods include both the dwarf dogwoods and the big-bracted dogwoods. Within this subclade, *C. canadensis*, the single species sampled for the dwarf dogwoods that consists of three species, is sister to a strongly supported (30, 99%, >5) big-bracted dogwood clade composed of *C. florida*, *C. nuttallii*, and *C. capitata* (Fig. 2).

Within the nyssoid-mastixioids subclade (Fig. 2), *Camptotheca* and *Nyssa* appear as sisters (18, 87%, 3), and are, in turn, sister to a clade composed of *Davidia* and *Mastixia* (11, 45%, 1). The Hydrangeaceae-Loasaceae lineage consists of two subclades, one containing representatives of Hydrangeaceae and the other containing members of Loasaceae. Among the four representa-

tives of Hydrangeaceae, two strongly supported clades were recognized: *Hydrangea* (*H. quercifolia*) - *Schizophragma* and *Deutzia* - *Philadelphus* (Fig. 2).

Combined *rbcL-matK*—The topologies generated by analyses of the two different combined *rbcL-matK* sequence data sets (see above) reveal identical relationships among taxa included in both data sets. Thus, only the results from the first data set, which includes a broader sampling of taxa, are described below.

The heuristic search found six most parsimonious trees, each of 1436 steps all in one island. All six shortest trees recognized the same major lineages within Cornales. Significantly, more resolution was obtained among the four major lineages of Cornales from analysis of the combined data set than with either *rbcL* or *matK* separately (Figs. 1–3). The strict consensus tree resulting from analysis of the combined *rbcL-matK* sequences shows that the Hydrangeaceae-Loasaceae lineage is the sister to the remainder of Cornales; *Curtisia* is the sister of the nyssoids-mastixioids lineage, which in combination are sister to the *Cornus-Alangium* lineage. The *Cornus-Alangium*-nyssoids-mastixioids-*Curtisia* clade is supported by six nucleotide substitutions, a bootstrap value of 44%, and decay value of one (Fig. 3). The sister relationship between nyssoids-mastixioids and *Curtisia* is found in all shortest trees in the combined *rbcL-matK* analysis and is supported by six nucleotide substitutions, a bootstrap value of 35%, and decay value of 1 (Fig. 3). In contrast, when these data sets were analyzed separately, other placements of *Curtisia* were sometimes seen. Thus, the strict consensus trees from analyses of *rbcL* or *matK* se-

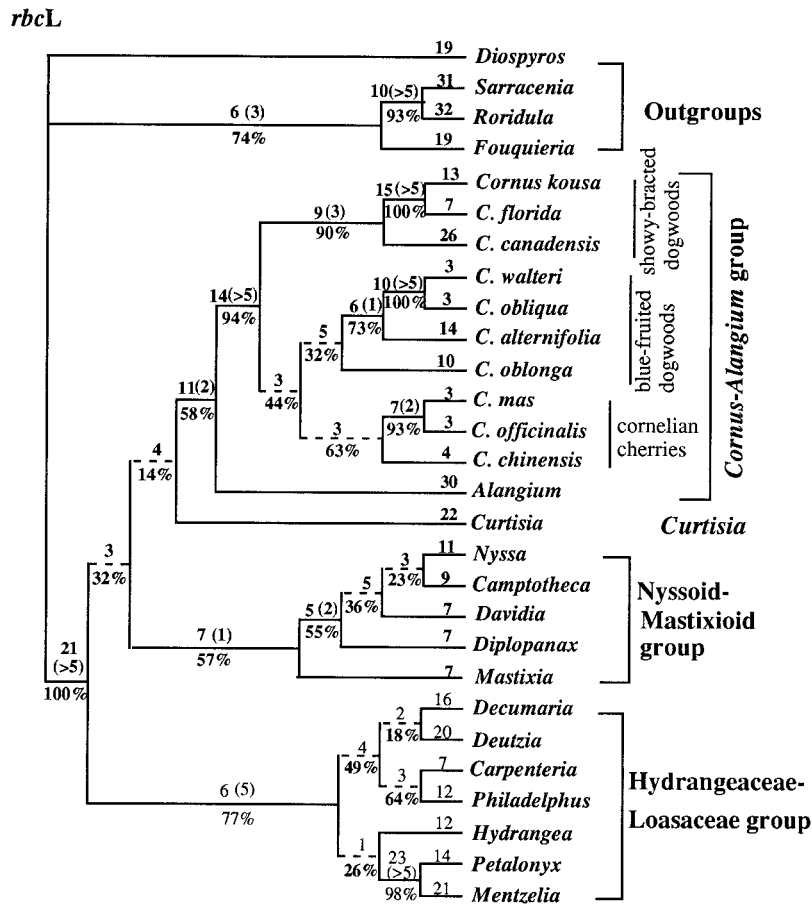


Fig. 1. *rbcL* cladogram of the cornalean clade from Xiang and Soltis (in press). This is one of the 97 most parsimonious trees resulting from phylogenetic analysis of *rbcL* sequences of Cornales and outgroup taxa (length = 568 steps, CI [Consistency Index] = 0.620 excluding uninformative characters, and RI [Retention Index] = 0.637). Dashed lines represent branches that are not present in all shortest trees. Numbers below branches are bootstrap values; numbers above branches indicate the inferred number of base substitutions; numbers in parentheses are decay values and indicate the number of steps longer a tree must be before a clade collapses.

quences alone show a polychotomy among *Curtisia*, Hydrangeaceae-Loasaceae, nyssoids-mastixioids, and *Cornus-Alangium*.

Relationships within *Cornus* suggested by the analysis of the combined *rbcL-matK* sequence data set were nearly identical to those suggested by the *matK* sequence data alone, but were less resolved among the major subclades of the genus (Fig. 3). A trichotomy was recognized among the cornelian cherries, showy-bracted dogwoods, and the blue-fruited dogwoods in the strict consensus tree resulting from the analysis of the combined data set, whereas in the *matK* sequence analysis, the cornelian cherries appeared as the sister of a clade consisting of the showy-bracted dogwoods and the blue-fruited dogwoods (including *C. oblonga*) (Fig. 2). Relationships within each of these three subgroups of *Cornus*, as well as relationships within each of the other major lineages of Cornales, are basically identical to those recognized by analysis of the *matK* sequence data alone (see Figs. 2, 3).

Combined *rbcL-matK* sequence and cpDNA restriction site data—The heuristic search for the combined *rbcL-matK* sequence and cpDNA restriction site data set

of *Cornus* found two shortest trees each of 845 steps [Fig. 4; for comparison with the shortest trees resulting from cpDNA restriction site data alone (taken from Xiang et al., 1996), see Fig. 5]. The two shortest trees differ only in the relationships among the three closely related, opposite-leaved, blue-fruited species, *C. obliqua*, *C. macrophylla*, and *C. walteri*. Both shortest trees recognize subgroups within *Cornus* identical to those suggested by the *matK* sequence data and the combined *rbcL-matK* sequence data. However, relationships among the subgroups suggested by the combined *rbcL-matK* sequence and cpDNA restriction site data set differ from those suggested by the *matK* sequence data. In the shortest trees resulting from analysis of the combined *rbcL-matK* sequence and cpDNA restriction site data set, the showy-bracted dogwoods are sister to the cornelian cherries, which together are the sister group of the blue-fruited dogwoods (including *C. oblonga*) (Fig. 4). These relationships are strongly supported by the bootstrap and decay analyses.

Rate and pattern of variation of *matK* vs. *rbcL*—The variable sites are fairly evenly distributed throughout *rbcL* and the 5' portion of *matK* that we have sequenced,

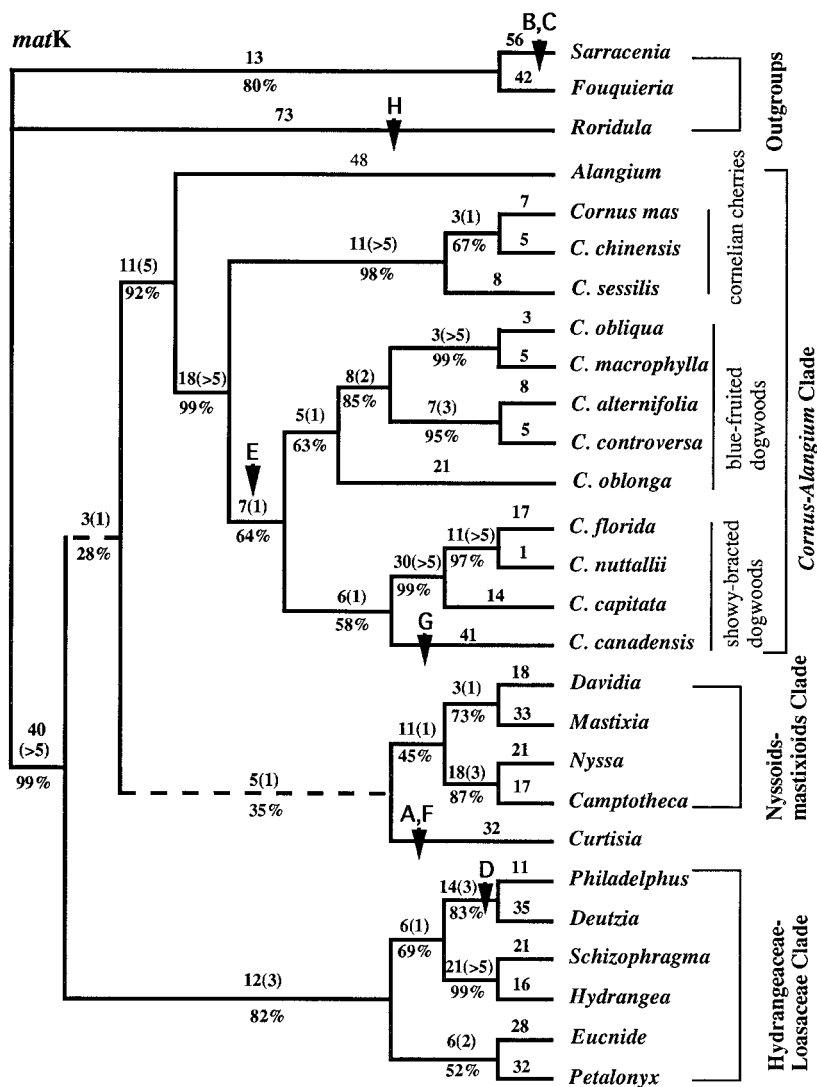


Fig. 2. One of the ten most parsimonious trees resulting from phylogenetic analysis of *matK* sequences of Cornales and outgroup taxa (length = 889 steps, CI = 0.588 excluding uninformative characters, and RI = 0.688). Base substitutions are indicated by numbers above branches; decay values are indicated by numbers in parentheses; bootstrap values are indicated by percentages below branches. Dashed lines represent branches that did not occur in all shortest trees; all other branches were present in all ten most parsimonious trees. A–H represent inferred sequence indels listed in Table 2.

with a greater level of variation observed in *matK*. Within Cornales, the total number of nucleotide substitutions per site estimated based on one of the shortest trees is 0.7031 for *matK*, a value 2.1 times higher than that determined for *rbcL* (0.3361). The ratio of variable sites at different codon positions differs greatly between *matK* and *rbcL*. In *matK*, variable sites appear to be nearly even in distribution among the three codon positions, 1.0:1.0:1.3 (109:112:143). In *rbcL*, in contrast, the ratio is 1.8:1.0:7.5 (36:20:150), with the third codon position much more variable. Nucleotide composition also differs between *matK* and *rbcL*. In *matK*, nucleotide composition is skewed with a much higher A-T than G-C content (an average of 66.9% vs. 33.1%), whereas the A-T and G-C nucleotide composition of *rbcL* is more evenly distributed (an average of 55.8% A-T vs. 44.2% G-C). The two genes also differ in the transition:transversion ratio. The transition:transversion ratio estimated from changes in-

ferred from one of the shortest trees is 1.21 in *matK* and 1.74 in *rbcL*. Five indels corresponding to the five alignment gaps were inferred in *matK*, whereas no indels occurred in *rbcL*.

Relative rate tests revealed that rates of synonymous (Ks) and nonsynonymous substitutions (Ka) in both *rbcL* and *matK* are largely homogeneous among the four major lineages of Cornales (*Cornus-Alangium*, nyssoids-mastixioids, Hydrangeaceae-Loasaceae, and *Curtisia*) (i.e., not significantly different from one another), with two exceptions (Table 3). That is, the nyssoids have a significantly slower Ks than the *Cornus-Alangium* lineage for both *rbcL* and *matK*; *Curtisia* has a slower Ka than the other three lineages and a faster Ks than the nyssoids for *matK*. Results of relative rate tests within *Cornus* indicated that for *rbcL*, Ks and Ka are also homogeneous across five major lineages (*C. oblonga*, the remaining blue-fruited dogwoods, the cornelian cherries, the big-

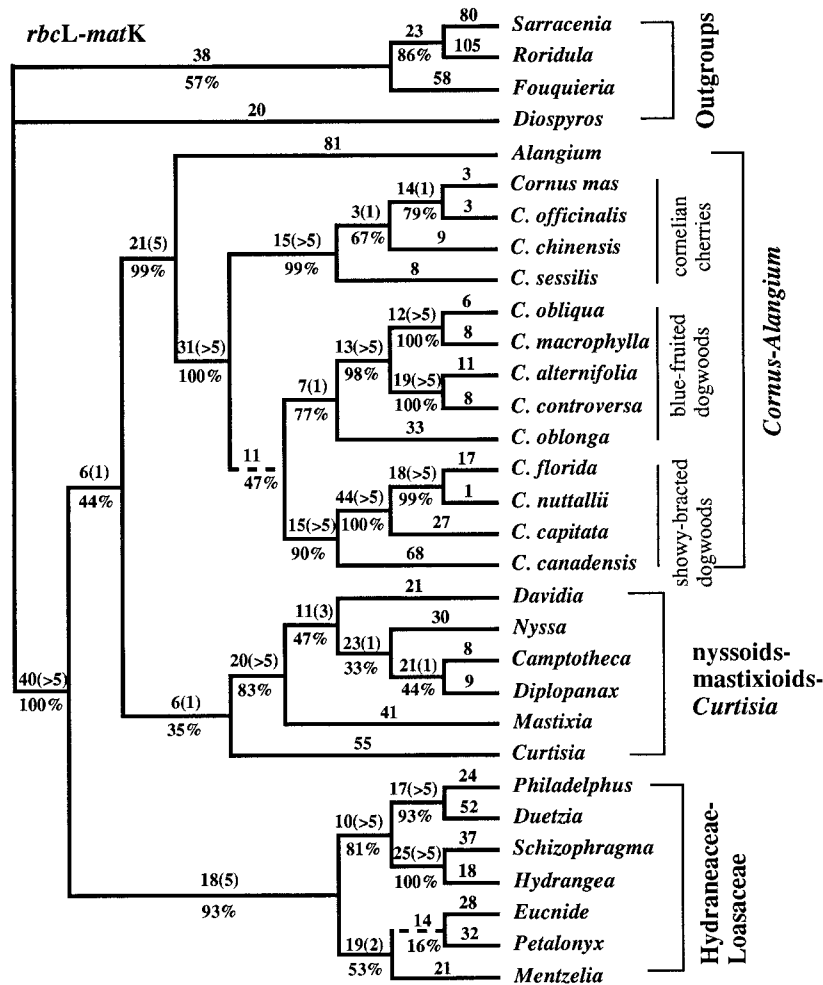


Fig. 3. One of the six most parsimonious trees resulting from phylogenetic analysis of a combined data set of *rbcL-matK* sequences for Cornales and outgroup taxa (length = 1436 steps, CI = 0.539 excluding uninformative characters, and RI = 0.671). Not all taxa are present in both data sets (see Materials and Methods). Base substitutions are indicated by numbers above branches; decay values are indicated by numbers in parentheses; bootstrap values are indicated by percentages below branches. Dashed lines represent branches that did not occur in all shortest trees; all other branches were present in all six most parsimonious trees.

bracted dogwoods, and the dwarf dogwood). A single exception was observed involving *C. oblonga* and the cornelian cherries; *K_a* in *C. oblonga* is significantly faster than that in the cornelian cherries. For *matK*, *K_s* and *K_a* are less homogeneous. The blue-fruited dogwoods have a significantly lower *K_s* than in *C. oblonga*, the cornelian cherries, and the dwarf dogwood; the big-bracted dogwoods have a lower *K_s* than observed for the dwarf dogwood. In addition, the dwarf dogwood has a faster *K_a* than the blue-fruited dogwoods and the cornelian cherries, and the big-bracted dogwoods also have a faster *K_a* than these two groups, as well as *C. oblonga*.

DISCUSSION

Comparison of *rbcL*, *matK*, and *rbcL-matK*—Separate analyses of the *rbcL* and *matK* data sets indicate that: (1) *Alangium* and *Cornus* are sister taxa, (2) four major lineages (or most inclusive monophyletic groups) are distinguishable by both genes in Cornales (*Cornus-Alangium*, nyssoids-mastixioids, *Curtisia*, and Hydrangeaceae-Loasaceae), and (3) three major lineages are pres-

ent within *Cornus* (blue-fruited dogwoods, cornelian cherries, and showy-bracted dogwoods; see Figs. 1, 2). Relationships among the four major lineages present in Cornales are poorly resolved in both analyses (Figs. 1, 2). However, *matK* sequence data provide greater resolution within subclades than do the *rbcL* sequences. For example, the shortest *rbcL* trees do not resolve relationships among the three major lineages within *Cornus* (Fig. 1). In contrast, the shortest *matK* trees do resolve relationships within *Cornus*, suggesting that the blue-fruited dogwoods are sister to the showy-bracted dogwoods, and the cornelian cherries are sister to all other *Cornus* species (Fig. 2). Relationships within the nyssoids-mastixioids lineage and Hydrangeaceae-Loasaceae lineage are also resolved poorly based on *rbcL* sequence data, but are well resolved with *matK* sequences (see Figs. 1, 2). In addition, the relationship of *Cornus oblonga* remains uncertain in the analysis of *rbcL* sequence data, but is resolved in the analysis of *matK* sequences. The shortest *matK* trees place *Cornus oblonga* as the sister to the remainder of the blue-fruited dogwoods (Fig. 2).

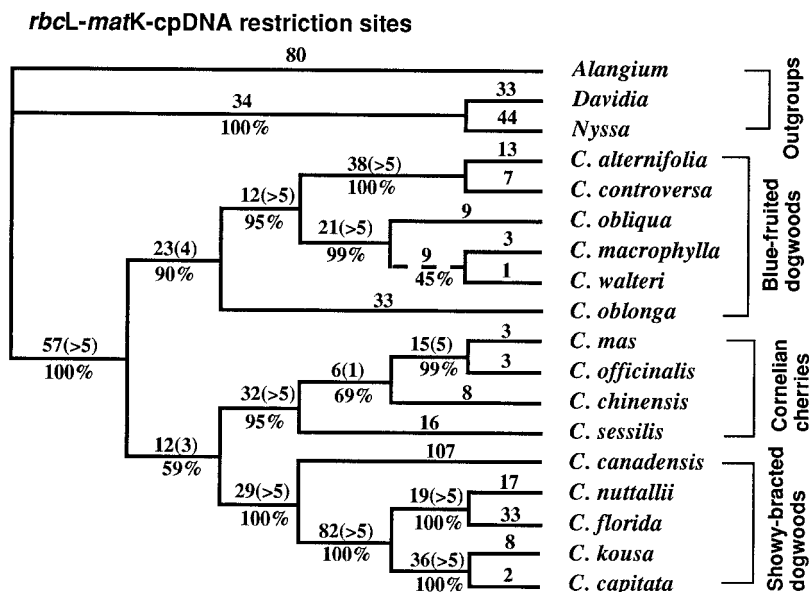


Fig. 4. One of the two most parsimonious trees resulting from phylogenetic analysis of a combined data set of *rbcl-matK* sequences and cpDNA restriction site data set for *Cornus* and outgroup taxa (length = 845, CI = 0.707 excluding uninformative characters, and RI = 0.823). Base substitutions are indicated by numbers above branches; decay values are indicated by numbers in parentheses; bootstrap values are indicated by percentages below branches. A dashed line represent a branch that did not occur in both shortest trees; all other branches were present in both most parsimonious trees.

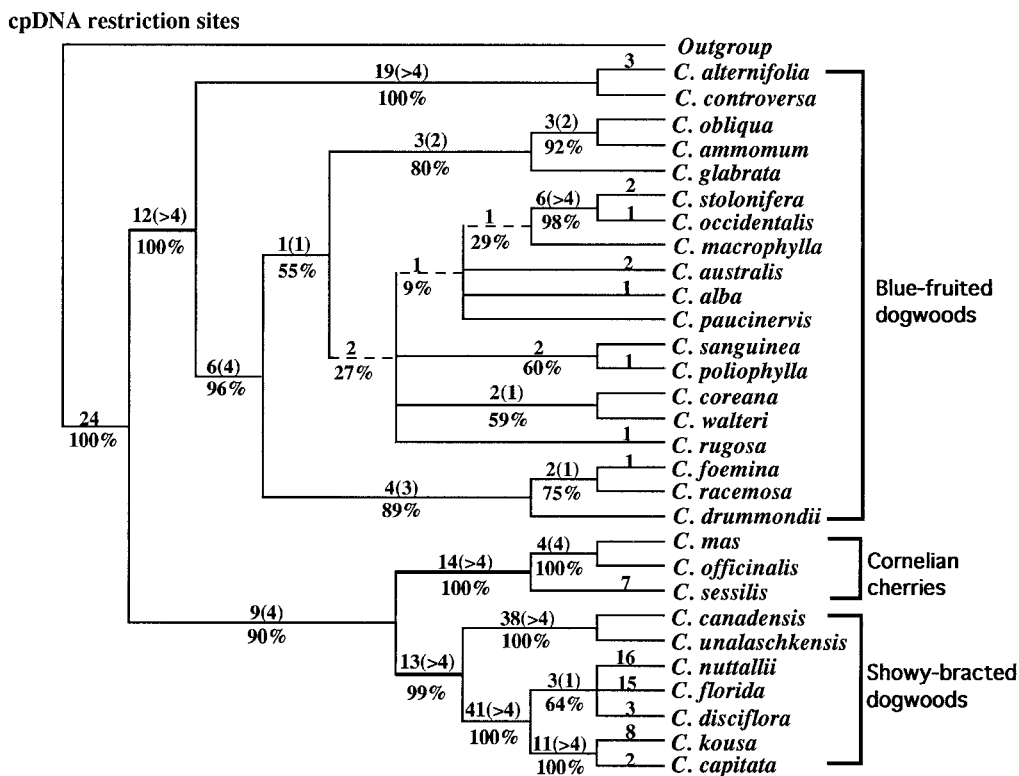


Fig. 5. One of the 27 most parsimonious trees (randomly selected) resulting from phylogenetic analysis of cpDNA restriction site data for *Cornus* (length = 284 steps, CI = 0.852 excluding uninformative characters, and RI = 0.940) (from Xiang et al., in press). Base substitutions are indicated by numbers above branches; decay values are indicated by numbers in parentheses; bootstrap values are indicated by percentages below branches. Dashed lines represent branches that did not occur in all shortest trees; all other branches were present in all 27 most parsimonious trees.

TABLE 3. Results of relative rate tests at *rbcL* and *matK* in Cornales. *rbcL* and *matK* sequences of *Fouquieria* were used as the outgroup in tests among major lineages of Cornales (Taxa A–D); *rbcL* and *matK* sequences of *Alangium* were used as the outgroup in tests within *Cornus* (Taxa E–I). A test statistic with absolute value >1.96 (indicated in boldface) is significant at the 0.05 level. A = *Cornus-Alangium* lineage, B = nyssoids-mastixioids lineage, C = Hydrangeaceae-Loasaceae lineage, D = *Curtisia*, E = *Cornus oblonga*, F = blue-fruited dogwoods, G = cornelian cherries, H = big-bracted dogwoods, and I = *C. canadensis*. For species sampled for each group, see Materials and Methods. K1 – K2 = Difference in the weighted number of substitutions per site between lineage 1 and lineage 2, Ks = Synonymous substitutions, and Ka = Nonsynonymous substitutions.

Taxa compared		<i>rbcL</i>		<i>matK</i>		
1	2	K1 – K2	Test statistic	K1 – K2	Test statistic	
Between major lineages of Cornales						
A	B	Ks:	0.054	3.772	0.038	1.965
		Ka:	0.003	1.126	0.010	1.408
A	C	Ks:	0.019	1.120	0.006	0.331
		Ka:	0.002	1.083	0.008	1.159
A	D	Ks:	–0.000	–0.001	–0.022	–0.811
		Ka:	0.002	0.699	0.027	3.554
B	C	Ks:	–0.009	–0.634	–0.031	–1.683
		Ka:	0.000	0.177	–0.001	–0.192
B	D	Ks:	–0.028	–1.723	–0.060	– 2.342
		Ka:	0.000	0.089	0.017	2.193
C	D	Ks:	–0.019	–1.059	–0.029	–1.148
		Ka:	–0.000	–0.030	0.019	2.639
Between major lineages within <i>Cornus</i>						
E	F	Ks:	–0.012	–1.024	0.062	3.037
		Ka:	0.002	0.675	0.006	0.904
E	G	Ks:	0.002	0.216	0.019	0.799
		Ka:	0.006	1.977	0.011	1.574
E	H	Ks:	–0.005	–0.327	0.029	1.221
		Ka:	0.001	0.160	–0.022	– 2.469
E	I	Ks:	–0.032	–1.694	–0.021	–0.791
		Ka:	0.002	0.501	–0.014	–1.607
F	G	Ks:	0.014	1.241	–0.043	– 2.140
		Ka:	0.004	1.957	0.005	0.779
F	H	Ks:	0.007	0.401	–0.033	–1.594
		Ka:	–0.001	–0.478	–0.028	– 3.302
F	I	Ks:	–0.020	–0.988	–0.083	– 3.274
		Ka:	0.000	0.012	–0.020	– 2.628
G	H	Ks:	–0.008	–0.504	0.010	0.454
		Ka:	–0.005	–1.566	–0.033	– 4.104
G	I	Ks:	–0.034	–1.805	–0.040	–1.634
		Ka:	–0.004	–1.130	–0.025	– 3.016
H	I	Ks:	–0.026	–1.237	0.050	–1.939
		Ka:	0.001	0.560	0.008	0.789

Because the separate analyses of *matK* and *rbcL* sequences produced topologies that are largely congruent (as discussed above), the two data sets were combined and phylogenetic analysis of this combined data set was conducted. The results of this analysis again reveal the *Cornus-Alangium*, Hydrangeaceae-Loasaceae, and nyssoids-mastixioids clades. In the combined analysis, *Curtisia* is always sister to nyssoids-mastixioids. This relationship is also revealed in 26% of the *rbcL* shortest trees and 80% of the *matK* shortest trees. Relationships among the major subclades of Cornales are also better resolved in the analysis of the combined *rbcL-matK* data set, although not strongly supported by the bootstrap and decay analyses (compare Figs. 1–3). All most parsimonious trees resulting from the analysis of the combined data set

indicate that the Hydrangeaceae-Loasaceae subclade is the sister of the remainder of Cornales; *Curtisia* and nyssoids-mastixioids form a monophyletic group sister to *Cornus-Alangium* (Fig. 3). These relationships are also shown in 40% of shortest *matK* trees and 21% of shortest *rbcL* trees. The lack of strong support for relationships among these major lineages may indicate a relatively rapid radiation of Cornales following its origin or may be a sampling artifact. The combined *rbcL-matK* analysis again reveals the same three major subclades within *Cornus* retrieved in the analysis of *matK* sequences alone (Figs. 2, 3).

Relationships within each of the large subclades of Cornales are well resolved in the analysis of the combined *rbcL-matK* data set and are concordant with those suggested by the analyses of *matK* and *rbcL* sequences alone. Differences between the strict consensus trees (not shown) resulting from analyses of the combined data set and the *matK* and *rbcL* sequences alone involve only the relationships among the three major subgroups within *Cornus* and the placement of *Mastixia*. Unlike the shortest *matK* trees, those resulting from the analyses of the combined data set and *rbcL* sequences alone did not clearly resolve relationships among the subclades within *Cornus*. That is, the cornelian cherries, showy-bracted dogwoods, and the blue-fruited dogwoods (including *C. oblonga*) form a trichotomy in the strict consensus tree of all shortest trees. Similarly, whereas the most parsimonious *matK* trees place *Mastixia* as sister to *Davidia*, the shortest trees resulting from analyses of the combined data set and *rbcL* sequences alone place *Mastixia* as sister to a clade consisting of all nyssoids and *Diplopanax* (Figs. 1–3).

Phylogenetic relationships among major lineages of Cornales and circumscription of Cornaceae—The analysis of a combined *rbcL* and *matK* sequence data set suggests that Cornales may have split early into two major groups. One group comprises Hydrangeaceae and Loasaceae, and the other group subsequently diverged into two clades, the *Cornus-Alangium* clade and the nyssoids-mastixioids-*Curtisia* clade. The latter clade further diverged into two subclades: (1) *Curtisia*, and (2) the nyssoids and mastixioids. These relationships were suggested by all most parsimonious trees resulting from analysis of the combined *rbcL-matK* data set, as well as by 40% of *matK* and 21% of *rbcL* shortest trees. Evidence from other sources also lends support to these relationships. For example, *Curtisia* and the nyssoids-mastixioids share similar pollen morphology and the same chromosome number ($N = 13$ in *Curtisia* and some species of *Mastixia*; see Ferguson, 1977; Goldblatt, 1978; Ferguson and Hideux, 1980). Several features also seem to unite the *Cornus-Alangium* clade with the nyssoids-mastixioids-*Curtisia* clade. All of these taxa share the presence of germination valves on the fruit stone, H-shaped thinning of the pollen aperture, one ovule per locule, and drupaceous fruits, and all but *Curtisia* lack central bundles in their gynoeceal vasculature; these features are not found in Hydrangeaceae, Loasaceae, and the outgroup taxa (Hutchinson, 1967; Eyde, 1988; Xiang et al., 1993). The *Cornus-Alangium*-nyssoids-mastixioids-*Curtisia* clade corresponds well to the circumscription of Cornaceae of

Eyde (1988), Eyde and Xiang (1990), and Thorne (1992), which included all these taxa, except *Alangium* and *Curtisia*. According to the results of the phylogenetic analyses of the combined *rbcL-matK* sequences, *Alangium* should be considered the closest relative of *Cornus*, and *Curtisia* should also be placed in Cornaceae sensu Eyde (1988).

Phylogenetic relationships within major lineages of Cornales—Phylogenetic analyses of *matK* sequences and the combined *rbcL-matK* data set resolve relationships within each of the major lineages that make up Cornales. For the most part, *matK* and the combined *rbcL-matK* trees depict lower level relationships within *Cornus* identical to those revealed by cpDNA restriction sites (Xiang et al., 1996; compare Figs. 2, 3, 5). Differences among trees resulting from analyses of *matK* sequences, the combined *rbcL-matK* sequence data set, and the cpDNA restriction sites involve only the placement of the cornelian cherry group. The *matK* sequence data suggest that: (1) the cornelian cherries are sister to all other *Cornus* species and (2) the blue-fruited dogwoods and showy-bracted dogwoods are sisters (Fig. 2). The sister-group relationship between the blue-fruited dogwoods and the showy-bracted dogwoods is supported by seven nucleotide substitutions, a bootstrap value of 64%, and a decay value of 1. It is supported further by a six-bp insertion in *matK* (Table 2), which is present only in the blue-fruited dogwoods and showy-bracted dogwoods. In contrast, analysis of cpDNA restriction site variation (Fig. 5), which involved more extensive sampling of species within *Cornus* (Xiang et al., 1996) indicates that the cornelian cherries are the sister group of the showy-bracted dogwoods and that the blue-fruited dogwoods are in turn sister to this large cornelian cherry-showy bracted lineage. The sister-group relationship between the cornelian cherries and showy-bracted dogwoods is strongly supported (nine restriction site mutations, bootstrap value of 90%, and decay value of four) (Fig. 5). Analysis of *rbcL* sequences alone and also analysis of the combined *rbcL-matK* data set did not resolve the relationships among the cornelian cherries, the blue-fruited dogwoods, and the showy-bracted dogwoods (Figs. 1, 3). Analysis of the combined *rbcL-matK* sequence-cpDNA restriction site data set revealed relationships among these groups identical to those suggested by the cpDNA restriction site data alone (Fig. 4).

The conflict between the *matK* and cpDNA restriction site trees is not as great as it might seem because reciprocal tree topology tests (i.e., analyzing one data set using the tree topology found with the other data set as a constraint tree and saving trees compatible with the constraint tree) indicate that to have the *matK* trees showing the same relationships within *Cornus* as suggested by the cpDNA restriction site data, only one additional step is required compared to the most parsimonious trees. To obtain trees showing the cpDNA restriction site topologies with the *rbcL-matK* data set requires two additional steps; to obtain the *matK* topologies requires one additional step. However, the cpDNA restriction site topologies may be the more correct topology. First, the cpDNA restriction site topologies are supported by much higher bootstrap and decay values (Fig. 5) compared to the *matK* topology (Fig. 2). To obtain cpDNA restriction site trees showing the *matK* topologies

requires four extra steps compared to the shortest. Other lines of evidence also lend support to the relationships suggested by the cpDNA restriction site data. For example, the cornelian cherries and the showy-bracted dogwoods share a number of features that are different from those found in the blue-fruited dogwoods, such as red fruits, inflorescences normally subtended by four decussate bracts, tenuinucellate ovules, similar pollen traits, and the production of iridoids. All of these features are synapomorphies of the cornelian cherries and the showy-bracted dogwoods except for the production of iridoids which is a symplesiomorphy. The blue-fruited dogwoods, in contrast, produce blue fruits, inflorescence with early-deciduous bracts, crassinucellate ovules, and do not produce iridoids (see Eyde, 1988; Xiang et al., 1996). These lines of evidence all point to a closer relationship between the cornelian cherries and the showy-bracted dogwoods than between either of these and the blue-fruited dogwoods.

The taxonomy and relationships of *Cornus oblonga*, a blue-fruited species, have long been controversial. *Cornus oblonga* is morphologically very similar to other blue-fruited dogwoods in having blue fruits and open cymes with small bracts. However, this species also exhibits a number of characters that are distinct from other species. For example, *C. oblonga* possesses morphologically distinctive pollen grains that are larger than those produced by other *Cornus* species (Zhu, 1984), and vessel elements that are shorter and wider than those of other *Cornus* species (Adams, 1949). *Cornus oblonga* also produces one to four embryo sacs with only one reaching maturity, whereas other *Cornus* species have only one embryo sac (Chopra and Kaur, 1965). In addition, *C. oblonga* has flowers that bloom in the fall rather than the spring, larger inflorescence bracts compared to other blue-fruited *Cornus* species, bracts that often persist to the mature stage of the inflorescence (in other blue-fruited species the bracts are early deciduous), and subopposite leaf arrangement (in other *Cornus* species the leaves are either typically opposite or alternate).

Cornus oblonga is typically considered a member of the blue-fruited group (Wangerin, 1910; Ferguson, 1966; Xiang, 1987; Eyde, 1988); however, as a result of the distinctive morphology of this species, it has been treated variously as a monotypic subsection (subsect. *Oblongifoliae*) within the section of the opposite-leaved, blue-fruited dogwoods (Wangerin, 1910), a distinct monotypic genus (*Yinquania*; Zhu, 1984), and a monotypic subgenus (subgen. *Yinquania*) of *Cornus* (Murrell, 1993). A recent morphological analysis (Murrell, 1993) placed *C. oblonga* as the sister to the remainder of the genus *Cornus*. Analysis of *matK* sequences, as well as analyses of the combined *rbcL-matK* sequence data matrix and the combined *rbcL-matK*-cpDNA restriction site data matrix, in contrast, all suggest that *C. oblonga* is sister to the remainder of the blue-fruited dogwoods (Figs. 2–4). This relationship is strongly supported in the analysis of the combined *rbcL-matK* sequence-cpDNA restriction site data set (23 mutations, a bootstrap value of 90%, and a decay value of 4) (Fig. 4). A 25-bp insertion in the 3' end of *rbcL* that extends to the flanking 3' noncoding region found in *C. oblonga* and the Asian opposite-leaved blue-fruited species (Xiang and Soltis, in press) may add additional support to a close relationship of *C. oblonga*

to other blue-fruited dogwoods. This insertion may have occurred early in the evolutionary history of the blue-fruited group and was later lost in the alternative-leaved and North American opposite-leaved, blue-fruited species.

Relationships among the nyssoids remain incompletely resolved. Analysis of *matK* sequences suggests that *Nyssa* is more closely related to *Camptotheca* than it is to *Davidia*, supporting the hypothesis of Titman (1949) and in agreement with evidence from wood anatomy, fatty acids, morphology, palynology, and serology (Fig. 2; Xiang et al., 1993). The *matK* sequence data also suggest a close relationship between *Mastixia* and *Davidia* (Fig. 2), but this relationship is only moderately supported (bootstrap value 73%, decay value of 1). This *matK* result differs from that suggested by analyses of the *rbcL* sequences and the combined *rbcL-matK* data set, which place *Mastixia* as sister to the remainder of the nyssoids-mastixioids group (Figs. 1, 3), a relationship having lower bootstrap, but higher decay support (bootstrap value of 55%, decay value of 2 in the *rbcL* trees, and bootstrap value of 47%, decay value of 3 in the combined *rbcL-matK* trees). Other lines of evidence (e.g., leaf, floral, and fruit morphology) suggest, however, that *Mastixia* is most closely related to *Diplopanax* (Eyde and Xiang, 1990). The latter appears as the sister of *Camptotheca* in all the combined *rbcL-matK* trees (Fig. 3) and in 63% of the *rbcL* trees. *Diplopanax* was not sequenced for *matK*, due to inadequate material.

Analyses of both *matK* sequences and the combined *rbcL-matK* data set are congruent in suggesting that Hydrangeaceae and Loasaceae are sister families, as well as in suggesting a close relationship between *Philadelphus* and *Deutzia*, and between *Schizophragma* and *Hydrangea* (Figs. 2, 3). These results are consistent with the findings of more extensive *rbcL* studies of Hydrangeaceae and Loasaceae (Hempel et al., 1995; Soltis, Xiang, and Hufford, 1995). The close relationship between *Philadelphus* and *Deutzia* is supported further by an insertion of 3 bp in *matK* at position 236 in these two genera. A recent analysis of *rbcL* sequence data for Loasaceae and putative relatives suggested that the enigmatic aquatic genus *Hydrostachys* (Hydrostachyaceae) might be a member of the Hydrangeaceae-Loasaceae clade more closely allied with Hydrangeaceae. Unfortunately, a sequence of *matK* is not available for the genus.

***matK* indels**—Three of the five *matK* indels inferred in Cornales are autapomorphies (two in *Curtisia* and one in *Cornus canadensis*), with the remaining (one in Hydrangeaceae and one in *Cornus*) potentially phylogenetically informative and adding support to relationships suggested by base substitutions. These findings are similar to those for Saxifragaceae s. s., Polemoniaceae, and Apiales where several indels were shown to be phylogenetically informative (Plunkett, 1994; Johnson and Soltis, 1995). The insertion (indel E) inferred in *Cornus* (in the blue-fruited dogwoods and the showy-bracted dogwoods) may be interpreted in two different ways. According to the *matK* phylogeny, the insertion can be considered most parsimoniously as a single event occurring in the common ancestor of the blue-fruited dogwoods and the showy-bracted dogwoods. However, the insertion also

can be considered to have evolved once early in the common ancestor of *Cornus* with a subsequent loss in the cornelian cherries. The latter hypothesis is consistent with the cpDNA restriction site phylogeny, as well as the phylogeny inferred from analysis of the combined *rbcL-matK* sequence and cpDNA restriction site data set of *Cornus* (Figs. 4, 5).

Rate and pattern of variation in *matK* and *rbcL*—

Our analyses indicate that *matK* has a different rate of evolution and pattern of variation from those observed for *rbcL* in Cornales (see also Johnson and Soltis, 1995). The *matK* sequences are more variable than those of *rbcL*. Furthermore, in *matK* the variable sites are more evenly distributed among codon positions (ratio of variable sites by codon position: 1.0:1.0:1.3). In contrast, the variable sites in *rbcL* sequences are mostly at the third position (1.8:1.0:7.5). The transition:transversion ratio in *matK* is ~70% lower than that in *rbcL* (1.21 vs. 1.74). These results are similar to the findings of Steele and Vilgalys (1994) and Johnson and Soltis (1995) for Saxifragaceae and Polemoniaceae, and suggest that *matK* may be less functionally constrained than *rbcL*. The *matK* gene is found to evolve 2.1 times faster than *rbcL* in Cornales as estimated from examining substitutions on one of the shortest trees. This rate difference between *matK* and *rbcL* is very close to those observed in Apiales (2.3 times faster, see Plunkett, 1994), and lower than that in Saxifragaceae s. s. (approximately three times faster, see Johnson and Soltis, 1994, 1995).

The *rbcL* gene has been found to evolve at different rates among different monocot lineages (Wilson, Gaut, and Clegg, 1990; Gaut et al., 1992; Gaut, Muse, and Clegg, 1993) with a long life span associated with a lower rate. A slower rate of *rbcL* evolution was also found in the gymnosperm family Taxodiaceae (Brunsfield et al., 1994; see also Soltis and Soltis, 1995). Relative rate tests and estimated differences in rates between groups reveal that rates of synonymous (Ks) and nonsynonymous substitutions (Ka) in *rbcL* are basically homogeneous within Cornales except in the cornelian cherries, in which Ka is significantly slower than that in *C. oblonga* (Table 3), and in nyssoids and mastixioids, Ks is significantly slower than that in *Cornus* and *Alangium*. In contrast, the *matK* gene evolves at more heterogeneous rates in Cornales (see Table 3). No apparent correlation between rates and life span are observed. However, the only herbaceous group of *Cornus*, containing three rhizomatous perennial dwarf dogwoods represented by *C. canadensis* in this study, does have the highest Ks in the genus at both *rbcL* and *matK*.

In summary, analyses of both *matK* sequences and the combined *rbcL-matK* data set have helped to elucidate phylogenetic relationships within Cornales. Results of these analyses suggested that *Cornus*, *Alangium*, *Nyssa*, *Camptotheca*, *Davidia*, *Diplopanax*, *Mastixia*, and *Curtisia* comprise a well-supported clade that should be considered to represent Cornaceae. This circumscription differs from all previously proposed Cornaceae (see Xiang et al., 1993; Xiang and Soltis, in press), but is very close to those of Eyde (1988), Eyde and Xiang (1990), and Thorne (1992). The rate and pattern of evolution of *matK* and *rbcL* are different in Cornales, with *rbcL* having a

slower and more homogeneous rate of nucleotide substitutions among different lineages, a higher transition:transversion ratio, and a lower A-T content compared to *matK*. Variable sites in *rbcL* are mostly at third position, whereas variable sites in *matK* are almost evenly distributed among the three codon positions.

LITERATURE CITED

- ADAMS, J. E. 1949. Studies in the comparative anatomy of the Cornaceae. *Journal of the Elisha Mitchell Scientific Society* 65: 218–244.
- BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.
- BRUNSFELD, S. J., P. S. SOLTIS, D. E. SOLTIS, P. A. GADEK, C. J. QUINN, D. D. STRENGE, AND T. A. RANKER. 1994. Phylogenetic relationships among genera of the conifer families Taxodiaceae and Cupressaceae: evidence from *rbcL* sequences. *Systematic Botany* 19: 253–262.
- CHASE, M. C., et al. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–580.
- CHOPRA, R. N., AND H. KAUR. 1965. Some aspects of the embryology of *Cornus*. *Phytomorphology* 15: 353–359.
- DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, AND J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 79: 333–345.
- , AND M. J. SANDERSON. 1992. The suitability of molecular and morphological evidence in reconstructing plant phylogeny. In P. S. SOLTIS, D. E. SOLTIS, AND J. J. DOYLE [eds.], *Molecular systematics of plants*, 340–368, Chapman and Hall, New York, NY.
- DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.
- EYDE, R. H. 1988. Comprehending *Cornus*: puzzles and progress in the systematics of the dogwoods. *Botanical Review* 54: 233–351.
- , AND Q.-Y. XIANG. 1990. Fossil mastixioid (Cornaceae) alive in eastern Asia. *American Journal of Botany* 77: 689–692.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- FERGUSON, L. K. 1966. Notes on the nomenclature of *Cornus*. *Journal of the Arnold Arboretum* 47: 100–105.
- . 1977. World pollen and spore flora 6: Cornaceae Dum. Almqvist & Wiksell, Stockholm.
- , AND M. J. HIDEUX. 1980. Some aspects of the pollen morphology and its taxonomic significance in Cornaceae sens. lat. *Proceedings of IV International Palynology Conference, Lucknow* 1: 240–249.
- GAUT, B. S., B. MORTON, B. C. MCCAIG, AND M. T. CLEGG. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. *Proceedings of the National Academy of Sciences, USA* 93: 10274–10279.
- , S. V. MUSE, W. D. CLARK, AND M. T. CLEGG. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. *Journal of Molecular Evolution* 35: 292–303.
- , ———, AND M. T. CLEGG. 1993. Relative rates of nucleotide substitution in the chloroplast genome. *Molecular Phylogenetics and Evolution* 2: 89–96.
- GOLDBLATT, P. 1978. A contribution of cytology in Cornales. *Annals of the Missouri Botanical Garden* 65: 650–655.
- HEMPEL, A. L., AND R. K. JANSEN. 1994. Phylogenetic relationships of the Mentzelioideae (Loasaceae). *American Journal of Botany* 81: 160.
- , P. A. REEVES, R. G. OLMSTEAD, AND R. K. JANSEN. 1995. Implications of *rbcL* sequence data for higher order relationships of the Loasaceae and the anomalous aquatic plant *Hydrostachys* (Hydrostachyaceae). *Plant Systematics and Evolution* 194: 25–37.
- HUTCHINSON, J. 1967. The genera of flowering plants, vol. 2. Oxford University Press, London.
- JOHNSON, L. A., AND D. E. SOLTIS. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae sensu stricto. *Systematic Botany* 19: 143–156.
- , AND ———. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Annals of the Missouri Botanical Garden* 82: 149–175.
- , J. L. SCHULTZ, D. E. SOLTIS, AND P. S. SOLTIS. In press. Monophyly and generic relationships of Polemoniaceae based on *matK* sequences. *American Journal of Botany*.
- KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* 38: 7–25.
- KUMAR, S., K. TAMURA, AND M. NEI. 1993. MEGA: molecular evolutionary genetic analysis, version 1.0. Pennsylvania State University, University Park, PA.
- LI, P., AND J. BOUSQUET. 1992. Relative-rate test for nucleotide substitutions between two lineages. *Molecular Biology and Evolution* 9: 1185–1189.
- MADDISON, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Systematic Zoology* 40: 315–328.
- MORGAN, D. R., AND D. E. SOLTIS. 1993. Phylogenetic relationships among members of Saxifragaceae sensu lato based on *rbcL* sequence data. *Annals of the Missouri Botanical Garden* 80: 631–660.
- MURRELL, Z. E. 1993. Phylogenetic relationships in *Cornus* (Cornaceae). *Systematic Botany* 18: 469–495.
- NEI, M., AND T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3: 418–426.
- NEUHAUS, H., AND G. LINK. 1987. The chloroplast tRNA^{Lys} (UUU) gene from mustard. *Current Genetics* 11: 251–257.
- OLMSTEAD, G. R., AND J. A. SWEERE. 1995. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.
- PLUNKETT, G. M. 1994. A molecular-phylogenetic approach to the “family-pair dilemma” in Apiales and Cyperales. Ph. D. dissertation. Washington State University, Pullman, WA.
- SOLTIS, D. E., R. K. KUZOFF, E. CONTI, R. GORNALL, AND K. FERGUSON. 1996. *matK* and *rbcL* gene sequence data indicate that *Saxifraga* (Saxifragaceae) is polyphyletic. *American Journal of Botany* 83: 371–382.
- , D. R. MORGAN, A. GRABLE, P. S. SOLTIS, AND R. KUZOFF. 1993. Molecular systematics of Saxifragaceae sensu stricto. *American Journal of Botany* 80: 1056–1081.
- , P. S. SOLTIS, T. G. COLLIER, AND M. L. EDGERTON. 1991. Chloroplast DNA variation in *Lithophragma* (Saxifragaceae): evidence for chloroplast transfer and paraphyly. *American Journal of Botany* 78: 1091–1112.
- , Q.-Y. XIANG, AND L. HUFFORD. 1995. Relationships and evolution of Hydrangeaceae based on *rbcL* sequence data. *American Journal of Botany* 82: 504–514.
- SOLTIS, P. S., AND D. E. SOLTIS. 1995. Plant molecular systematics: inferences of phylogeny and evolutionary processes. *Evolutionary Biology* 28: 139–194.
- STEELE, K. P., AND R. VILGALYS. 1994. Phylogenetic analysis of Polemoniaceae using nucleotide sequences of the plastid gene *matK*. *Systematic Botany* 19: 126–142.
- SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by Center of Biodiversity, Illinois Natural History Survey, Champaign, IL.
- THORNE, R. F. 1992. An updated phylogenetic classification of the flowering plants. *Aliso* 13: 365–389.
- TITMAN, P. W. 1949. Studies in the woody anatomy of the family Nysaceae. *Journal of the Elisha Mitchell Scientific Society* 65: 245–261.
- WANGERIN, W. 1910. Cornaceae. In W. Engelmann [ed.], *Das Pflanzenreich*, Series IV, family 229 (Helf 41). A. Engler, Leipzig.
- WILSON, M. A., B. GAUT, AND M. T. CLEGG. 1990. Chloroplast DNA evolves slowly in the palm family. *Molecular Biology and Evolution* 7: 303–314.
- XIANG, Q.-Y. 1987. A neglected character of *Cornus* L. s. l. with special reference to a new subgenus—*Sinocornus* Q. Y. Xiang. *Acta Phytotaxonomica Sinica* 25: 125–131.
- , S. J. BRUNSFELD, D. E. SOLTIS, AND P. S. SOLTIS. 1996. Phy-

logenetic relationships in *Cornus* based on chloroplast DNA restriction sites: implications for biogeography and character evolution. *Systematic Botany* 21: 515–534.

———, AND D. E. SOLTIS. In press. *rbcL* sequence divergence and phylogenetic relationships of Cornaceae sensu lato. In D. E. Boufford and H. Ohba [eds.], *Sino-Japanese flora—its characteristics and diversification*. University of Tokyo Press, Tokyo.

———, ———, D. R. MORGAN, AND P. S. SOLTIS. 1993. Phylogenetic relationships of *Cornus* L. sensu lato and putative relatives inferred from *rbcL* sequence data. *Annals of the Missouri Botanical Garden* 80: 723–734.

ZHU, Z.-Y. 1984. *Yinquania* Z. Y. Zhu—a new genus of Cornaceae from China. *Bulletin of Botanical Research (Harbin)* 4(4): 121–128.