American Journal of Botany 88(6): 1131–1138. 2001.

PHYLOGENETIC RELATIONSHIPS WITHIN *CORNUS* (CORNACEAE) BASED ON 26S rDNA SEQUENCES[1]

CHUANZHU FAN[2,3] AND (JENNY) QIU-YUN XIANG[3]

Department of Biological Sciences, Idaho State University, Pocatello, Idaho 83209-8007 USA

Phylogenetic relationships within the dogwood genus *Cornus* have been highly controversial due to the great morphological heterogeneity. Earlier phylogenetic analyses of *Cornus* using chloroplast DNA (cpDNA) data (including *rbcL* and *matK* sequences, as well as restriction sites) and morphological characters suggested incongruent relationships within the genus. The present study generated sequence data from the nuclear gene 26S rDNA for *Cornus* to test the phylogenetic hypotheses based on cpDNA and morphological data. The 26S rDNA sequence data obtained represent 16 species, 13 from *Cornus* and three from outgroups, having an aligned length of 3380 bp. Both parsimony and maximum likelihood analyses of these sequences were conducted. Trees resulting from these analyses suggest relationships among subgroups of *Cornus* consistent with those inferred from cpDNA data. That is, the dwarf dogwood (subg. *Arctocrania*) and the big-bracted dogwood (subg. *Cynoxylon* and subg. *Syncarpea*) clades are sisters, which are, in turn, sister to the cornelian cherries (subg. *Cornus* and subg. *Afrocrania*). This red-fruited clade is sister to the blue- or white-fruited dogwoods (subg. *Mesomora,* subg. *Kraniopsis,* and subg. *Yinquania*). Within the blue- or white-fruited clade, *C. oblonga* (subg. *Yinquania*) is sister to the remainder, and subg. *Mesomora* is sister to subg. *Kraniopsis.* These relationships were also suggested by the combined 26S rDNA and cpDNA data, but with higher bootstrap and Bremer support in the combined analysis. The 26S rDNA sequence data of *Cornus* consist of 12 expansion segments spanning 1034 bp. These expansion segments evolve approximately four times as fast as the conserved core regions. The study provides an example of phylogenetic utility of 26S rDNA sequences below the genus level.

Key words: 26S rDNA; combining data; Cornaceae; *Cornus*; cpDNA; molecular evolution; molecular phylogeny.

Cornus L. sensu lato consists of ~55 species that are mostly trees and shrubs and rarely perennial herbs with woody rhizomes. The genus is widely distributed in the northern hemisphere, with centers of diversity in eastern Asia, eastern North America, and western North America. Two species of the genus are endemic to South America and one species is endemic to tropical Africa (see Table 1). Species of *Cornus* are morphologically diverse. Various types of inflorescences are found within the genus, including open compound cymes with minute, nonmodified bracts; umbellate cymes with four basal, scale-like bracts; capitate cymes with four large, basal, and showy bracts; and minute compound cymes with four basal showy bracts. In addition, the color of fruits also varies among species (see Table 1). Due to the great morphological diversity encompassed by the Linnaean circumscription of *Cornus,* the taxonomy and phylogenetic relationships of subgroups within the genus are highly controversial (see Eyde, 1988; Murrell, 1993; Xiang et al., 1996). The genus has been divided into several distinct genera (e.g., Rafinesque, 1838; von Berchtold and Opiz, 1838; Spach, 1839; Nakai, 1909; Hutchinson, 1942; Pojarkova, 1950; also see Murrell, 1993), or into various numbers of subgenera (e.g., Harms, 1898; Wangerin, 1910; Ferguson, 1966; Eyde, 1987, 1988; Xiang, 1987). Following the broad view of *Cornus,* a total of ten subgenera have been recognized by different authors at one time or another (see Table 1; Ferguson, 1966; Xiang, 1987; Eyde, 1988; Murrell, 1993).

In order to understand relationships within *Cornus,* several sets of data have been recently collected for the genus for phylogenetic analyses. These include molecular data from the chloroplast genome (Xiang et al., 1996; Xiang, Soltis, and Soltis, 1998) and morphological characters (Murrell, 1993). Phylogenetic analyses of these data identified five major lineages within *Cornus*: (1) *C. oblonga,* an enigmatic blue-fruited dogwood, (2) other blue- or white-fruited dogwoods, (3) the cornelian cherries, (4) the big-bracted dogwoods, and (5) the dwarf dogwoods (the herbaceous species). However, relationships among these groups suggested by cpDNA data are different from those suggested by morphological data. The cpDNA data (combined *rbcL* and *matK* sequences and restriction sites) suggested that the dwarf dogwoods are sisters of the big-bracted dogwoods with the cornelian cherries sister to them, and *C. oblonga* is a member of the blue- or white-fruited dogwoods and sister to the remainder of the group (see Fig. 1; Xiang et al., 1996; Xiang, Soltis, and Soltis, 1998). These results are consistent with the hypothesis of Eyde (1988) based on synthesis of information available without performing phylogenetic analyses. In contrast, cladistic analysis of 28 morphological, anatomical, chemical, and cytological characters of *Cornus* by Murrell (1993) suggested that the cornelian cherries and the big-bracted dogwoods are sister groups, and the dwarf dogwoods are sister to them, with *C. oblonga* being the first-branching lineage in *Cornus* (see Fig. 2; Murrell, 1993; Hardin and Murrell, 1997). To explore possible sources of incongruence between the cpDNA and morphological data, additional phylogenetic analyses, especially analyses of molecular data from the nuclear genome, are necessary.

Although the most widely used nuclear phylogenetic markers below the generic level have been the ITS (internal transcribed spacer) sequences of ribosomal genes (see reviews by Baldwin et al., 1995; Soltis and Soltis, 1998), our initial anal-

[1] Manuscript received 29 June 2000; revision accepted 9 November 2000.

The study was supported by the Research Office and the Molecular Research Core Facility at Idaho State University. Our special thanks go to Erin O'Leary-Jepsen for running the sequencing gels, to Zack Murrell for providing DNAs of *C. volkensii,* to Chris Brochmann for collecting plant material of *C. suesica,* to David E. Boufford for collecting *C. officinalis,* to Gregory M. Plunkett and Robert K. Kuzoff for their insightful review comments on the manuscript.

[2] Author for reprint requests (Tel: 919 515-3345; FAX: 919 515-3436; e-mail: cfan3@unity.ncsu.edu).

[3] Current address: Dept. of Botany, North Carolina State University, Campus Box 7612, Raleigh, NC 27695-7612 USA.

TABLE 1. Morphological characteristics of the subgenera of *Cornus.* All subgenera are woody and hermaphroditic with terminal inflorescence and opposite leaves except those indicated. Taxonomic treatment was synthesized from Ferguson (1996), Xiang (1987), and Murrell (1993). Common names of subgenera are provided below the scientific names.

Subgroup	Fruits and inflorescence	Size and distribution
Subg. *Yinquania* (Zhu) Murrell (oblong-blue-fruited dogwood)	Blue, oblong fruits; open compound, cymes; bracts minute.	1 spp.; East Asia
Subg. *Kraniopsis* Raf. (blue- or white-fruited dogwoods)	Blue or white, and round fruits; open compound cymes; bracts minute.	~30 spp.; mostly East Asia, North America; 2 or 3 Europe; 1 or 2 South America
Subg. *Mesomora* Raf. (alternate-leaved, blue-fruited dogwoods)	Blue fruits; alternate leaves; open compound cymes; bracts minute.	2 spp.; East Asia, eastern North America
Subg. *Afrocrania* Harms (African cornelian cherry)	Red fruits; umbellate cymes subtended by four nonshowy bracts; dioecious.	1 spp.; tropical Africa
Subg. *Cornus* (Cornelian cherries)	Red fruits; umbellate cymes terminal subtended by four nonshowy bracts.	3 spp.; East Asia, western North America, Europe
Subg. *Sinocornus* Q. Y. Xiang (Chinese cornelian cherry)	Red fruits; umbellate cymes axillary subtended by four nonshowy bracts.	1 spp.; China
Subg. *Discocrania* (Mexican disciflorous dogwood)	Red fruits; capitulate cymes subtended by four early deciduous bracts.	1 or 2 spp.; Central America
Subg. *Cynoxylon* Raf. (American big-bracted dogwoods)	Red fruits; capitulate cymes subtended by four large, showy bracts; fruit separate.	2 or 3 spp.; eastern and western North America extended to Mexico
Subg. *Syncarpea* (Nakai) Xiang. (Asian big-bracted dogwoods)	Red fruits; capitulate cymes subtended by four large, showy bracts; fruits fused.	4–12 spp.; East Asia
Subg. *Arctocrania* Endl. Ex Reichenbach (dwarf dogwoods)	Red fruit; minute cymes subtended by four small, showy bracts; herbaceous.	3 spp.; circumboreal

yses within *Cornus* revealed a high level of ITS sequence divergence among species from the five major lineages, which made unambiguous alignment of sequences unfeasible (Xiang and Fan, unpublished data). Therefore, we turned to a more slowly evolving nuclear gene 26S rDNA. Recent studies have demonstrated the great potential of 26S rDNA sequences in elucidating phylogenetic relationship at various taxonomic levels of seed plants (at and above the generic level) (e.g., Mishler et al., 1994, used rRNA sequences only; Ro, Keener, and McPheron, 1997; Kuzoff et al., 1998; Soltis and Soltis, 1998; Stefanovic et al., 1998; Ashworth, 2000). In our ongoing study of 26S rDNA sequencing for Cornales, several species representing the five different subgroups of *Cornus* were included. Our preliminary results revealed that sufficient sequence variation exists among *Cornus* species. Thus, we employed comparative 26S rDNA sequencing to reconstruct a nuclear phylogeny of *Cornus.* The goals of this study were: (1) to determine phylogenetic relationships among subgroups of *Cornus* using nuclear 26S rDNA sequences; (2) to test the phylogenetic hypotheses based on morphology and cpDNA data; (3)

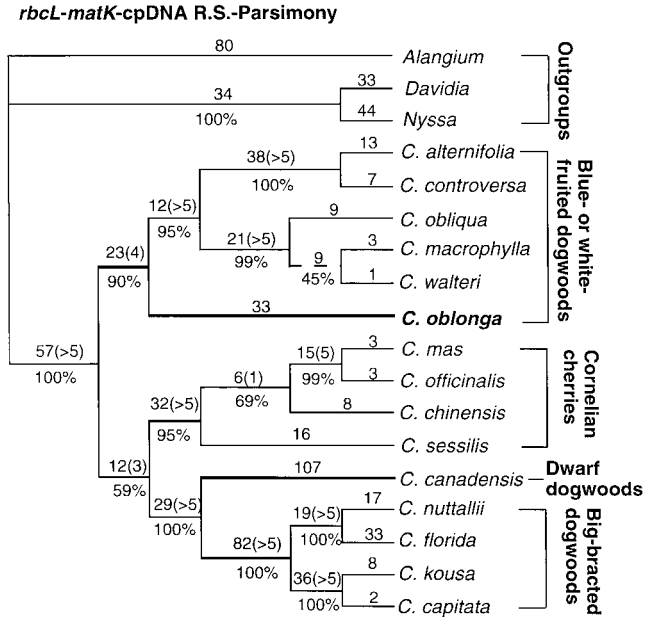

Fig. 1. One of the most parsimonious trees resulting from phylogenetic analyses of combined data set of *rbcL* and *matK* sequences and cpDNA restriction site data for *Cornus* (length = 845, consistency index = 0.707, excluding uninformative characters, and retention index = 0.823) (modified from Xiang, Soltis, and Soltis, 1998). Base substitutions are indicated above branches; bootstrap values are indicated below branches; decay values are indicated by numbers in parentheses.

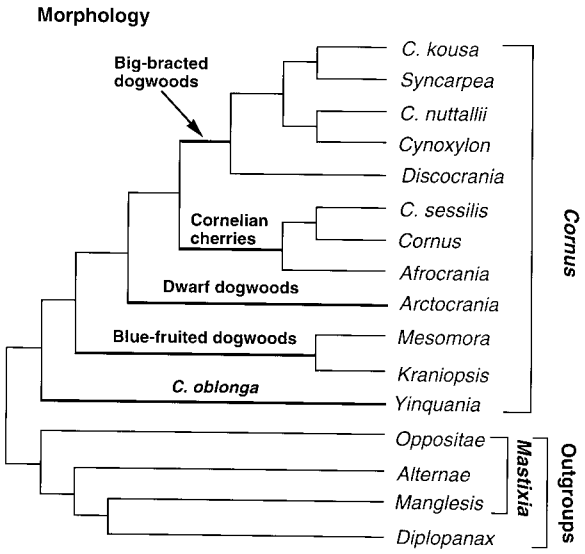

Fig. 2. The phylogenetic tree derived from cladistic analysis of 28 morphological, anatomical, chemical, and cytological characters modified from Murrell (1993).

TABLE 2. Species sampled in the study of 26S rDNA sequencing of *Cornus*. Species of *Cornus* are listed according to classifications of Ferguson (1966), Murrell (1993), and Xiang (1987).

Species	Sources and location of vouchers	GenBank accession no.[a]
Cornus L.		
Subgen. *Yinquania*		
C. oblonga Wall.	Sun, s.n., Bot. Gard. Kunming, China.	GBAN-AF297539
Subgen. *Mesomora*		
C. controversa Hemsl.	Arnold Arboretum 20458, WS.	GBAN-AF297541
Subgen. *Kraniopsis*		
C. racemosa Lam.	Xiang et al. 157, WS.	GBAN-AF297538
C. walteri Wangerin	Arnold Arboretum 414-67-1, WS.	GBAN-AF297540
Subgen. *Afrocrania*		
C. volkensii Harms	Knox 2528, Africa.	GBAN-AF297542
Subgen. *Cornus*		
C. mas L.	Arnold Arboretum 577-51-A, WS.	GBAN-AF297535
C. officinalis Seib. et Zucc.	Boufford et al. 26065, GH.	GBAN-AF297536
C. sessilis Torr. Ex Durand	Terry M. Hardig, California 1994.	GBAN-AF297537
Subgen. *Arctocrania*		
C. canadensis L.	Xiang et al. 198, WS.	GBAN-AF297530
C. unalaschkensis Ledeb.	Xiang 210, WS.	GBAN-AF297534
C. suecica L.	Chris Brochmann, 94-388, Norway.	GBAN-AF297531
Subgen. *Cynoxylon*		
C. florida L.	Xiang 250, WS.	GBAN-AF297532
Subgen. *Syncarpea*		
C. kousa Hance	Xiang 310, Ohio State University compus.	GBAN-AF297533
Outgroups		
Nyssa ogeche Marsh.	U.S. National Arbortum, s.n.	GBAN-AF297545
Alangium platanifolium (Sieb. & Zucc.) Harms	Soltis 2543, Japan.	GBAN-AF297544
Schizophragma hydrangeoides Sieb. et Zucc.	Soltis 2516, Japan.	GBAN-AF297543

[a] The prefix GBAN- has been added to link the online version of *American Journal of Botany* to GenBank but it is not part of the actual accession number.

to gain insights into the evolution of some subgroup-diagnostic morphological characters in *Cornus*.

MATERIALS AND METHODS

Sampling—Thirteen species of *Cornus* L. representing all major morphological diversity of the genus and the five major lineages identified through earlier phylogenetic analyses of cpDNA were sampled in the 26S rDNA sequencing study (Table 2; also see Table 1). All subgenera except subg. *Discocrania* and subg. *Sinocornus* were included. These two subgenera, each represented by a single species, were not sampled because of lack of DNA. Previous cpDNA studies indicated that *C. disciflora* (the only member of subg. *Discocrania*) is a member of the big-bracted dogwoods and *C. chinensis* (the only member of subg. *Sinocornus*) is a member of the cornelian cherries (see Table 1, Fig. 1; Xiang et al., 1996; Xiang, Soltis, and Soltis, 1998). Thus, even without these two species our sampling represented well all major subgroups within the genus. Although subgenus *Kraniopsis* is the largest subgenus in *Cornus* (~30 spp.), all species in the subgenus are morphologically very similar, and they formed a strongly supported monophyletic group in the cpDNA study (Xiang et al., 1996). Therefore, only two species were sampled from this subgenus. All of the species sampled were included in previous *matK* and *rbcL* sequencing and chloroplast DNA restriction site analyses, except *C. suecica* (a dwarf dogwood) and *C. volkensii* (the single species from Africa). These two species had not been included in the cpDNA studies because of the lack of materials of these taxa at the time the studies were completed. Three genera, *Nyssa*, *Alangium*, and *Schizophragma* (Hydrangeaceae), were chosen as the outgroups for the present study based on the results of broad phylogenetic analyses of *matK* amd *rbcL* sequences for Cornaceae (Xiang, Soltis, and Soltis, 1998), which suggested that *Alangium* is the sister of *Cornus*; *Nyssa* and Hydrangeaceae are close relatives of *Cornus*.

DNA isolation, PCR amplification, and DNA sequencing of 26S rDNA—To maximize comparability between the present and previous studies, DNAs used herein were those isolated for previous *rbcL* and *matK* sequencing studies. The procedures of isolating DNA were described elsewhere (Xiang et al., 1993; Xiang, Soltis, and Soltis, 1998). The 26S rDNA sequences were amplified via PCR (polymerase chain reaction) from total DNA aliquots using the forward primer N-nc26S1 (5′-CGACCCCAGGTCAGGCG-3′) and the reverse primer 3331rev (5′-ATCTCAGTGGATCG TGGCAG-3′) following Kuzoff et al. (1998) with slight modifications. Our PCR reactions contained the following: 5 μL of 10× Mg free buffer, 6 μL of 25 mmol/L $MgCl_2$, 10 μL of 2.5 mmol/L dNTP, 1.0 μL of 20 μmol/L N-nc26S1 (forward primer), 1.0 μL of 20 μmol/L 3331 rev (reverse primer), 5 μL of DMSO (dimethyl sulfoxide), 0.3 μL of *Taq* polymerase (Promega), 2.0 μL of 20 ng/μL total DNA extract, and 19.7 μL of deionized water. The PCR reaction mix was covered with two drops of mineral oil and run on a Robocycle PCR machine as the following: (1) 94°C for 3 min for one cycle; (2) 30 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 3.5 min; and (3) a terminal phase at 72°C for 5 min.

The double-stranded (DS) PCR products were subsequently purified via precipitation with 60 μL of 20% PEG 8000/2.5 mol/L NaCl on ice for at least 1 h (Morgan and Soltis, 1993; Soltis and Soltis, 1997). The precipitated DS products were centrifuged for 15 min at 14 000 rpm at 4°C. The DNA pellets were then washed with 1000 μL of 75% ethanol (prechilled to 4°C) and centrifuged for 3 min. The DNA pellets were washed again with chilled 95% ethanol and centrifuged for 3 min at 4°C. The ethanol was decanted, and the pellets were dried in a vacuum. The dried DNA products were resuspended in 20–30 μL of ddH_2O. One microliter of the clean PCR products was electrophoresed in a 1% agarose mini-gel for quantification.

The purified DS DNA products were used as the templates for sequencing on an ABI-377 automated sequencer following the standard protocol recommended by the company (Applied Biosystems, Foster City, CA 94404 USA).

TABLE 3. Twenty insertions and deletions (lettered A–T) inferred from 26S rDNA sequences of *Cornus* and outgroups. An asterisk indicates indels that are potentially phylogenetically informative within *Cornus* (the indels can be present on one or more than one lineage).

Indel	Sequence involving the indels	Base position in the aligned sequences	Species with the indel sequence present
A*	C	8	*C. walteri, C. controversa, Nyssa*
B	C	432	*C. volkensii, Nyssa*
C	C/T	468	*C. canadensis, C. suecica, C. florida, C. kousa, C. unalaschkensis, C. mas, C. officinalis, C. sessilis, C. racemosa, C. oblonga, C. walteri, C. controversa, C. volkensii, Nyssa*
D*	C/G	591	*C. kousa, C. mas, C. officinalis, C. sessilis, C. racemosa, C. oblonga, C. walteri, C. controversa, C. volkensii*
E*	G	600	*C. canadensis, C. suecica, C. unalaschkensis*
F	G	606	*Nyssa*
G*	T	638	*C. canadensis, C. suecica, C. unalaschkensis*
H	CCCC	751–754	*C. volkensii*
I	GGG	761–763	*C. volkensii*
J	RK	1002–1003	*Alangium, Schizophragma, Nyssa*
K	C/T	2083	*Schizophragma, Nyssa*
L	AGG	2095–2097	*Schizophragma*
M*	T	2156	*C. unalaschkensis, C. mas, C. controversa, Schizophragma*
N*	G	2209	*C. kousa, C. sessilis, C. oblonga, C. walteri, C. controversa, C. volkensii, Alangium, Schizophragma, Nyssa*
O*	A	2238	*C. sessilis, C. volkensii*
P*	C	2596	*C. sessilis, C. volkensii*
Q	CGC	3182–3184	*Alangium*
R	C	3257	*C. volkensii, Alangium, Schizophragma, Nyssa*
S	GGTGC	3270–3274	*Alangium*
T*	T	3358	*C. suecica, C. florida, C. recemosa, C. walteri, C. controversa, C. volkensii*

For some species, DMSO was added to the sequencing reactions to obtain clear sequences. Sixteen sequencing primers (N-nc26S1, N-nc26S3, N-nc26S4, N-nc26S5, N-nc26S6, N-nc26S8, N-nc26S10, N-nc26S12, N-nc26S14, 268rev, 641rev, 950rev, 1449rev, 2134rev, 2782rev, and 3331rev) described in Kuzoff et al. (1998) were used to obtain the entire sequence of 26S rDNA. The sequence chromatogram output files for all species were checked and edited base by base before being aligned manually.

Phylogenetic analysis—The 26S rDNA sequences representing 16 taxa comprise 3380 aligned bp with small gaps (see RESULTS). The data matrix was analyzed with both parsimony and maximum likelihood (ML) methods using PAUP 4.0b4 (Swofford, 2000). For parsimony analysis, gaps were coded as missing data; branch-and-bound search was conducted. Branch-and-bound search was performed using furthest taxon addition sequence and initial upper bound computed via stepwise addition. To evaluate relative robustness of the clades found in the most parsimonious trees, bootstrap and decay analyses were conducted. Bootstrap analysis (Felsenstein, 1985) of 1000 replicates was conducted using fast heuristic search and TBR branch-swapping. The method for decay analysis (Bremer, 1988) was described in Xiang et al. (1996), which followed Eernisse and Kluge (1993). This method involves examining each node of interest in turn using a constraint statement that specifies only the node of interest being monophyletic and saving the shortest trees that do not satisfy this criterion. The difference between the length of these trees and the true shortest trees is used as the decay value for that node.

Maximum likelihood analyses were first conducted using heuristic searches with random taxon addition of five replicates with default settings of the ML program (i.e., empirical base frequency, HKY [Hasegawa, Kishino, and Yano, 1985] two-parameter model variant for unequal base frequency, ti/tv = 2, and equal rates for all sites). The resulting tree was simultaneously used to estimate values for base frequency, ti/tv ratio, proportion of invariable sites, and shape of the discrete gamma distribution of rates across sites, through maximum likelihood. A subsequent ML analysis using the estimated values of these parameters (A = 0.220264, C = 0.272649, G = 0.307582, T = 0.199505; ti/tv ratio = 1.343463 [kappa = 2.741222]; proportion of invariable sites = 0.704318; shape value for discrete gamma distribution = 0.848768) was then conducted to see whether adjusting these parameters to the estimated values resulted in significant differences in tree topologies.

Since the results of analyses of 26S rDNA sequences suggested phylogenetic relationships within *Cornus* are highly congruent with those achieved via cpDNA data (combined *matK, rbcL* sequences, and restriction sites), the 26S rDNA sequences were combined with cpDNA data for further parsimony and ML analyses to obtain a comprehensive view of relationships. The combined data matrix included ten species of *Cornus* and two outgroups (*Alangium* and *Nyssa*), each of 6338 characters, among which 3380 bp were from 26S rDNA, 1212 bp from *matK*, 1504 bp from *rbcL*, and 242 from restriction sites. Phylogenetic analyses of the combined 26S rDNA and cpDNA were conducted as above.

RESULTS

Sequence divergence—The 26S rDNA sequences generated for the 16 species of *Cornus* varied from 3340 to 3370 bp in length before alignment. All sequences can be aligned easily by sight against the reference sequences from two angiosperm species, *Saxifraga mertensiana* Bong. (AFO36498, Kuzoff et al., 1998) and *Tragopogon dubius* Scop. (AFO36493, Kuzoff et al., 1998), obtained from GenBank. The aligned 26S rDNA sequences in *Cornus* and outgroups contained a total length of 3380 bp. Sequences for all species are complete except for *C. controversa* and *C. sessilis.* In *C. controversa*, 166 bp (bases 1542–1707) are missing and in *C. sessilis*, 476 bp (bases 1217–1692) are missing. Despite repeated efforts, these missing data in these two species could not be obtained due to potential primer divergence or heterogeneity of temperate DNA. Among the 16 sequences of *Cornus* and outgroups, 391 of the 3380 sites are variable (11.56%) and 137 sites (4.05%) are phylogenetically informative.

The alignment sequences of *Cornus* and outgroups required the addition of 20 small alignment gaps (1–5 bp in length).

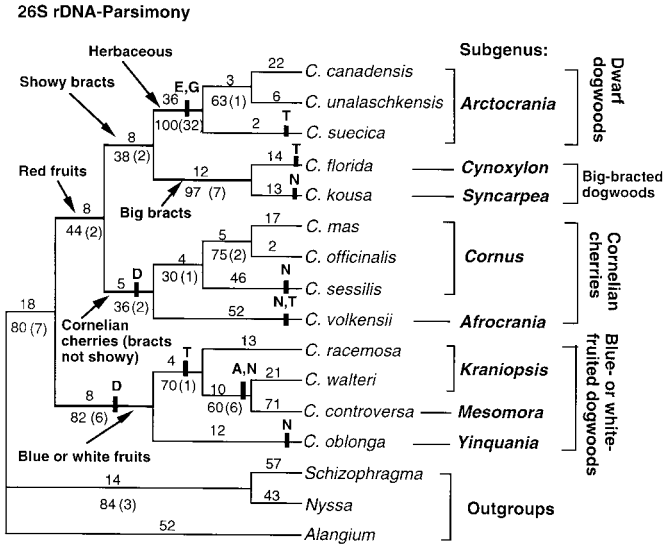

Fig. 3. The single most parsimonious tree resulting from analysis of the entire 26S rDNA sequences (tree length = 577 steps, consistency index = 0.781 excluding uninformative characters, retention index = 0.619). Base substitutions are indicated above branches; bootstrap values are indicated below branches; decay values are indicated by numbers in parentheses. Indels were coded as missing. Six of phylogenetically informative indels (A, D, E, G, N, T) within *Cornus* are mapped on the branches. Sequence information regarding these indels is shown in Table 3.

Eleven of these gaps appear to be autapomorphies, and nine of them are potentially phylogenetically informative within *Cornus* (Table 3, also see Fig. 3).

It is noteworthy that a majority of these indels (15 of 20) occur in the expansion segments, regions of gene that evolve more rapidly (Clark et al., 1984; Dover and Flavell, 1984; Flavell, 1986). The location of expansion segments of 26S rDNA sequences appears to be highly conserved over a wide range of taxa (Bult, Sweere, and Zimmer, 1995). According to the coordinates for expansion segment positions in the sequence of *Oryza sativa* (Kuzoff et al. 1998), 12 expansion segments were identified in the 26S rDNA sequence data of *Cornus* and outgroups (Table 4). These expansion segments of 26S rDNA in *Cornus* and outgroups span a total of 1034 bp, among which 233 bp (22.53%) are variable and 97 sites (9.38%) are phylogenetically informative. These values are much higher than those for the conserved core regions, which contain 2346 bp, of which 158 (6.73%) are variable and have only 40 sites (1.71%) that are phylogenetically informative. The expansion segments also have a higher average $G + C$ content (68.5%) than the entire 26S rDNA (58.1%) and the conserved core regions (52.6%).

Phylogenetic relationships inferred from 26S rDNA sequences—Phylogenetic analyses of the entire 26S rDNA sequences of *Cornus* using parsimony found a single shortest tree of 577 steps (CI [consistency index] = 0.780, excluding the uninformative characters, RI [retention index] = 0.619; Fig. 3); this single shortest tree is completely resolved. Major clades identical to those recognized by earlier phylogenetic analyses of cpDNA data were identified in the 26S rDNA tree: (1) all of the blue- or white-fruited dogwoods (represented by *C. racemosa, C. walteri, C. controversa,* and including *C. oblonga*); (2) the cornelian cherries (*C. mas, C. officinalis, C. sessilis,* and *C. volkensii*); (3) the big-bracted dogwoods (represented by *C. florida* and *C. kousa*); and (4) the dwarf dogwoods (*C. canadensis, C. suecica,* and *C. unalaschkensis*). The dwarf dogwoods (clade 4) and the big-bracted dogwoods (clade 3) (all producing showy bracts on the inflorescence) were recognized as sisters. This showy-bracted dogwood clade was, in turn, sister to the cornelian cherries (clade 2). All of the blue- or white-fruited dogwoods formed a monophyletic group (clade 1) sister to the large clade consisting of the dwarf dogwoods, big-bracted dogwoods, and the cornelian cherries. Although all of the major clades (except the cornelian cherries) are strongly supported (with >82% bootstrap value and >6 decay index), the relationships among major groups described above are not strongly supported (with bootstrap value <50% and decay index <3) (Fig. 3). Within the dwarf dogwood lineage (with all of the three species sampled), *C. canadensis* and *C. unalaschkensis* are sisters. Within the cornelian cherry group (four of the five species were sampled, see Table 1), *C. mas* and *C. officinalis* are sisters, with *C. sessilis* sister to them, and *C. volkensii* is at base within the clade. Within the blue- or white-fruited lineage, *C. oblonga* is sister to the remainder of the lineage (Fig. 3).

TABLE 4. Locations and lengths of the 12 expansion segments (D1–D12) in 26S rDNA sequence of *Cornus*. Location positions are expressed with reference to the coordinates for expansion segment positions in the sequence of *Oryza sativa* (Sugiura et al., 1985; Kuzoff et al., 1998). All sequences are written 5′ → 3′.

Expansion-segment	Position	Length	Sequence before ES in *Cornus* and outgroups	Sequence after ES in *Cornus* and outgroups
D1	113–261	149	GAANAGCCCA	ACGAGTCGGG
D2	422–652	231	GGGAGGGAAG	GCCCGTYTTG
D3	694–809	116	ACATGTGTGC	AGCATGCCTG
D4	1004–1010	7	GCTGGAGCCC	TTMTATCGGG
D5	1091–1129	39	ATAGGTAGGA	AGCTCCAAGT
D6	1162–1188	27	GTAAGCAGAA	GGKTTACCGT
D7a	1562–1605	44	TCGATCCTAA	AAAGGGAATC
D7b	1648–1673	26	ACGTGRCGGC	ACGTYGGCGG
D8	1970–2115	146	GCTCTGAGGG	CARCTGACTC
D9	2490–2515	26	GGATAAGTGG	CCACTACTTT
D10	2541–2619	79	TTATTTTACT	GACATTGTCA
D11	3021–3024	4	CCCTACTGAT	GTGYCGCAAT
D12	3179–3318	140	AGYSACGCAT	AGAATCCTTT

The ML analysis using default setting (see MATERIALS AND METHODS) found a tree with the best score of -Ln = 8531 with a topology identical to the parsimony tree (Fig. 4). The ML analysis using estimated values found a tree with a higher likelihood (-Ln = 8219), but showing the same relationships as described above.

Analyses of combined cpDNA and 26S rDNA sequence data—The combined 26S rDNA and cpDNA data set contains 939 (14.82%) variable sites of which 347 are from 26S rDNA, 232 from *matK*, 153 from *rbcL*, and 207 from restriction sites; and 327 (5.16%) are phylogenetically informative sites, of which 98 are from 26S rDNA, 59 from *matK*, 65 from *rbcL*, and 105 from restriction sites. Parsimony analysis of the combined data set found a single most parsimonious tree of 1229 steps, with a topology nearly identical to that of 26S rDNA tree (Fig. 5). The only difference between the two trees involves the placement of *C. volkensii* within the cornelian cherries. In the combined tree, *C. volkensii* is placed as the sister

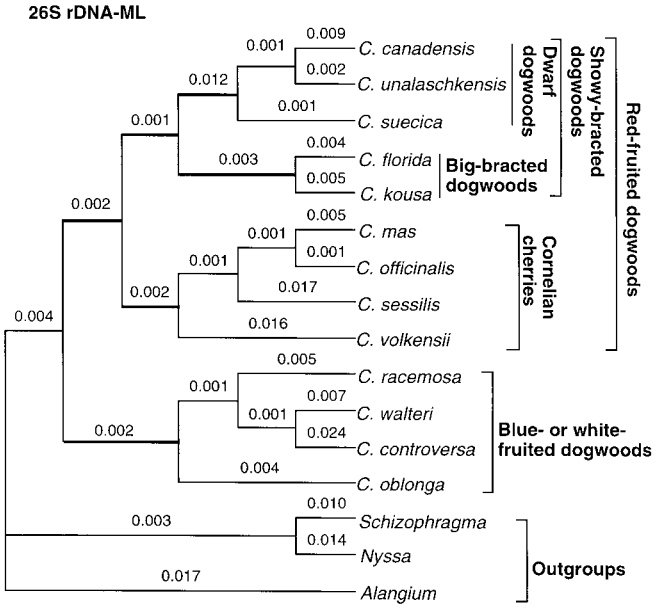

Fig. 4. The maximum likelihood (ML) tree resulting from analysis of the 26S rDNA sequences using heuristic searches with random taxon addition of five replicates, empirical base frequency, HKY model requested two parameter model variant for unequal base frequency, ti/tv = 2 [kappa = 2.402791], and equal rates for all sites; branch lengths are indicated above the nodes (-Ln likelihood = 8531.9677). See MATERIALS AND METHODS: PHYLOGENETIC ANALYSIS for an explanation of abbreviations.

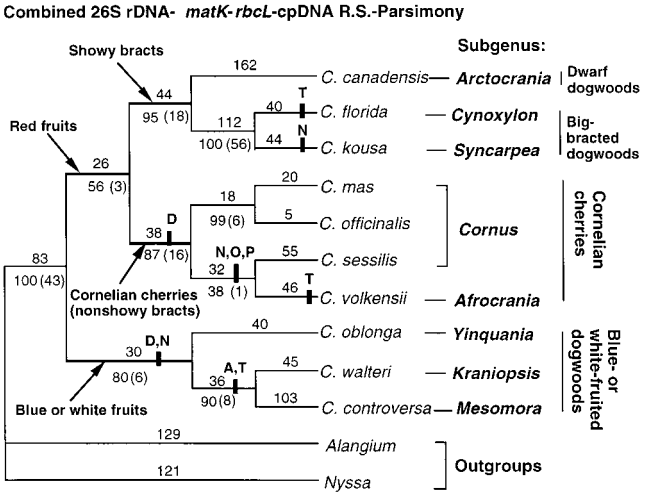

Fig. 5. The single most parsimonious tree resulting from analysis of combined 26S rDNA sequences and cpDNA data (*matK-rbcL* sequences and restriction sites) (tree length = 1229 steps, consistency index = 0.830 excluding uninformative characters, retention index = 0.627). Base substitutions are indicated above branches; bootstrap values are indicated below branches; decay values are indicated by numbers in parentheses. Six phylogenetically informative indels (A, D, N, O, P, T) from 26S rDNA sequences are mapped on the branches. Sequence information regarding these indels is shown in Table 3.

of *C. sessilis*, whereas, in the 26S rDNA tree, the species is recognized as a distinct lineage sister to the remainder of the cornelian cherries. Significantly, higher bootstrap and decay values for all clades were obtained for the tree derived from the combined data (Fig. 5).

DISCUSSION

Phylogenetic potential of 26S rDNA sequences below the genus level—The phylogenetic utility of 26S rDNA sequences in seed plants has been demonstrated only recently by a few studies of taxa at higher taxonomic levels (above the generic level) (e.g., Ro, Keener, and McPheron, 1997; Kuzoff et al., 1998; Stefanovic et al., 1998; Ashworth, 2000). Our phylogenetic study of *Cornus* using 26S rDNA sequences provides an example of phylogenetic utility of this gene at an intrageneric level. Although the rate of evolution of 26S rDNA is relatively conservative (comparable to *rbcL*), the gene contains several expansion segments that evolve faster than the conserved core regions (see Bult, Sweere, and Zimmer, 1995; Kuzoff et al., 1998). The variable rate of evolution in the expansion segments and the conserved core regions makes the gene useful for phylogenetic analyses at different taxonomic levels depending on the regions employed (see Kuzoff et al., 1998). The expansion segments can be used at lower taxonomic levels, whereas the conserved core regions are suitable at higher levels. In addition, the large size of the gene provides more variable characters for phylogenetic analysis than *rbcL*; the low sequence variation as a result of the conservative rate of the gene is well compensated by its large size. Although only 11.56% of the sites are variable, and 4.05% is phylogenetically informative in the 26S rDNA matrix of *Cornus*, the total number of variable sites from the entire sequence is 391, and the number of total informative sites is 137. These numbers are nearly two times higher than those for *rbcL* and *matK*, respectively, in the combined data matrix (see RESULTS). Analyses of the 26S rDNA sequences of *Cornus* resulted in a completely resolved phylogeny of the genus and suggested relationships congruent with the cpDNA-based phylogeny (Figs. 3, 4). This suggests that the 26S rDNA sequences as a whole contain sufficient phylogenetic information at the intrageneric level of *Cornus* and are useful for elucidating phylogenetic relationships within the genus.

To explore differential utilities of the expansion segments and conserved core regions of 26S rDNA sequences in *Cornus*, we conducted phylogenetic analyses of the sequences from the expansion segments and conserved core regions separately. The analyses of these portions of 26S rDNA did not produce topologies that were as fully resolved. The results indicated that the analysis of the expansion segments alone produced a weakly supported phylogeny inconsistent with both the cpDNA-based phylogeny and the morphology-based phylogeny. In this phylogeny, the monophyly of the cornelian cherries was unsupported and species of the group appear in different clades. Similar results were obtained from the analysis of the conserved core regions. In the trees resulting from the analysis of the conserved core sequences, the monophyly of *Cornus* was even unsupported. These odd results may be explained as the effect of insufficient variable characters in either the expansion segments or the conserved core regions. Although the expansion segments contain a higher percentage of phylogenetically informative sites (9.38%) than that of the entire 26S rDNA (4.05%), the total number of phylogenetically informative sites in the expansion segments is only 97, 40 sites less compared to 137 sites in the entire gene. The conserved core regions contain only 40 potentially phylogenetically informative sites. Moreover, the expansion segments have a significantly higher average $G + C$ content (68.5%)

than that of the entire 26S rDNA (58.1%) and that of the conserved core regions (52.6%). The higher $G + C$ content in the 26S rDNA expansion segments poses a potential problem for phylogenetic analysis if the algorithm used to construct a tree assumes equally abundant nucleotides (see Kuzoff et al., 1998).

Phylogenetic relationships within Cornus—The phylogenetic tree derived from the analysis of 26S rDNA sequences of *Cornus* shows a topology identical to the tree derived from cpDNA data (Fig. 1; Xiang et al., 1996; Xiang, Soltis, and Soltis, 1998), although some nodes in the 26S rDNA tree are not strongly supported by bootstrap and decay analyses (Figs. 3, 4). This 26S rDNA tree is also congruent with the tree derived from analyses of the combined 26S rDNA and cpDNA data (Fig. 5). Thus all of the molecular data from both nuclear and chloroplast genomes of *Cornus* are concordant. These data suggest that the genus diverged early into two large lineages: (1) the blue- or white-fruited group and (2) the red-fruited group. The red-fruited group subsequently separated into the cornelian cherries and a clade bearing showy bracts, which then diverged into the big-bracted dogwoods and the dwarf dogwoods. Within the blue- or white-fruited group, *C. oblonga* was the first lineage to branch off. This phylogenetic pattern is consistent with the scheme proposed by Eyde (1988), but at odds with the morphology-based phylogeny (Murrell, 1993). Phylogenetic analysis of morphological characters by Murrell (1993) placed the cornelian cherries (rather than the dwarf dogwoods as in the molecular tree) as the sister of the big-bracted dogwoods. This relationship was supported by five inflorescence characters that were treated as independent characters, including protective bracts, precocious peduncle, inflorescence preformed in the previous fall and developed from a mixed bud, reduced primary inflorescence branches, and reduced secondary inflorescence branches. However, three of these characters (protective bracts, precocious peduncle, and reduced secondary inflorescence branches) are homoplasious on the morphological trees, whereas they are reversed one or two times in some other clades or terminal taxa. In addition, these five inflorescence characters may not be independent because these characters may be developmentally correlated. Thus, differences in inflorescence may be overweighted in the morphological analysis, resulting in an incongruence between the morphological and molecular phylogeny. Based on the molecular phylogeny derived from multiple molecular data sets, the synapomorphies used to unite the cornelian cherries and the big-bracted dogwoods appear to be plesiomorphic characters evolved in the ancestor of the red-fruited group, but were later lost in the dwarf dogwoods (see discussion in Xiang et al., 1996). Alternatively, these features may have evolved independently in the big-bracted dogwoods and the cornelian cherries. However, these two hypotheses cannot be distinguished without fossils representing the ancestor of the red-fruited group.

The second major incongruence between the molecular and morphological phylogenies involves the placement of *C. oblonga* (subg. *Yinquania*). *Cornus oblonga* was placed as the basal group within *Cornus* in the morphological tree (Fig. 2; Murrell, 1993), whereas in the molecular phologeny, *C. oblonga* is a distinct lineage sister to the remainder of the blue- or white-fruited dogwoods, a relationship strongly supported by bootstrap and decay analyses (Figs. 3, 5). A single morphological character (displaced bracts, present in the remainder of the genus, absent in *C. oblonga*) separated *C. oblonga* and the remainder of the genus in the morphological tree (Murrell, 1993). Based on the molecular phylogenies, this character is better explained as a plesiomorphic condition, and the nondisplaced bract in *C. oblonga* is a derived state (perhaps a reversal). Several morphological, anatomical, biochemistry characters (blue fruits, lack of iridoids, crassinucellate ovule, open cyme with minute bracts) also support the placement of *C. oblonga* as a member of the blue- or white-fruited group.

The tropical African species *C. volkensii* is the only dioecious species in *Cornus.* Due to its morphological uniqueness, the species was treated as a separate subgenus (subg. *Afrocrania* Harms) or as a distinct genus (*Afrocrania* Hutch.) (Hutchinson, 1942; Ferguson, 1966; also see Murrell, 1993; Xiang et al., 1993). Eyde (1988) considered *C. volkensii* as a member of the cornelian cherries based mainly on similarities in fruit morphology and inflorescence type between *C. volkensii* and other cornelian cherries. Our 26S rDNA sequence data and the combined 26S rDNA-cpDNA data strongly supported *C. volkensii* being a member of the cornelian cherries (Figs. 3, 5) and recognized it as either the sister of all of the other cornelian cherries (Fig. 3) or as the sister of *C. sessilis* (Fig. 5). However, neither of these placements of *C. volkensii* within the cornelian cherries are highly supported by bootstrap and decay analyses (Figs. 3, 5). At present, no synapomorphic morphological characters can be identified to support either of these relationships of *C. volkensii* in the cornelian cherries. Both Eyde (1988) and Murrell (1993) suggested that *C. volkensii* is sister to all of the other cornelian cherries (also Fig. 2) based on the autopomorphies found in *C. volkensii* (e.g., dioecy, reduced secondary inflorescence branches). Neither Eyde nor Murrell identified any synapomorphies to unite the rest of the cornelian cherries. Eyde (1988) further proposed that the divergence of *C. volkensii* from the other cornelian cherries might have occurred in the Paleocene or early Eocene based on morphological and fossil evidence. For example, *C. volkensii* has dioecious breeding system, spiny pollen, up to nine or ten ripe fruits per umbel, a broad apical depression in the fruit-stones, whereas all the other cornelian cherries have a synoecious breeding system, smooth pollen, only one to five ripe fruits per umbel, and no apical depression in the fruit-stones. According to Eyde (1988) fossil fruits of an ancient, extinct lineage of cornelian cherries, *C. multilocularis,* were found in the early Tertiary in the London Clay, suggesting that the divergence of this lineage and the extant cornelian cherries occurred in or before the early Tertiary; the divergence of *C. volkensii* from the other cornelian cherries was subsequent to this event. Therefore, the basal placement of *C. volkensii* within the cornelian cherries as suggested by 26S rDNA sequence data is concordant with hypotheses of Eyde (1988) and Murrell (1993).

The dwarf dogwoods are the only herbaceous members of the dogwood genus *Cornus.* This group comprises three species, *C. canadensis, C. suecica,* and *C. unalaschkensis.* Evidence from cytology, phytogeography, and morphology suggests that *C. unalaschkensis* may be an allotetroploid species derived from past hybridization between *C. canadensis* and *C. suecica* followed by chromosomal doubling (Love and Love, 1975; Bain and Denford, 1979; Murrell, 1994). The 26S rDNA sequence data suggested that *C. unalaschkensis* is more closely related to *C. canadensis* than it is to *C. suecica.* This result may indicate that, if *C. unalaschkensis* is indeed an allotetroploid, as has been hypothesized, the 26S rDNA in *C. unalas-*

chkensis has converted to the type of *C. canadensis*. However, more extensive analyses (e.g., analyses of both nuclear and cpDNA data including all three species with more extensive sampling) are needed to test the hypothesis and to determine whether *C. unalaschkensis* is of hybrid origin.

In summary, the entire 26S rDNA sequence of *Cornus* shows a low level of sequence divergence, but because of its great length, provides sufficient variable characters (compared to *rbcL* and *matK*) to resolve the phylogenetic relationships within *Cornus*. Analyses of 26S rDNA sequence data result in a phylogeny of *Cornus* that is congruent with that inferred from combined cpDNA data, but at odds with the phylogeny derived from morphological analyses. Combining data for phylogenetic analysis can minimize sampling error and maximize the explanatory power of the data if congruent hypotheses are generated from separate analyses (see review by Johnson and Soltis, 1998). This was also demonstrated in our analysis of the combined 26S rDNA and cpDNA data in *Cornus*, which not only suggested congruent phylogenetic relationships to those inferred from separate data analyses within *Cornus*, but also significantly increased supports for most of the clades recognized in the tree (compare Figs. 3 and 5).

LITERATURE CITED

ASHWORTH, V. E. T. M. 2000. Phylogenetic relationship in Phoradendreae (Viscaceae) inferred from three regions of the nuclear ribosomal cistron. I. Major lineages and paraphyly of *Phoradendron*. *Systematic Botany* 25: 349–370.

BAIN, J. F., AND K. E. DENFORD. 1979. The herbaceous members of the genus *Cornus* in NW North America. *Botanical Notiser* 132: 121–129.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, AND M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247–277.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

BULT, C. J., J. A. SWEERE, AND E. A. ZIMMER. 1995. Cryptic sequence simplicity, nucleotide composition bias, and molecular coevolution in the large subunit of ribosomal DNA in plants: implications for phylogenetic analysis. *Annals of the Missouri Botanical Garden* 82: 235–246.

CLARK, G. B., B. W. TAGUE, V. C. WARE, AND S. A. GERBI. 1984. *Xenopus Laevis* 28S ribosomal RNA: a secondary structure and its evolutionary and functional implications. *Nucleic Acid Researches* 12: 6197–6220.

DOVER, G. A., AND R. B. FLAVELL. 1984. Molecular coevolution: DNA divergence and the maintenance of function. *Cell* 38: 622–623.

EERNISSE, D. J., AND A. G. KLUGE. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Molecular Biology and Evolution* 10: 1170–1195.

EYDE, R. H. 1987. The case for keeping *Cornus* in the broad Linnaean sense. *Systematic Botany* 12: 505–518.

———. 1988. Comprehending *Cornus*: puzzles and progress in the systematics of dogwoods. *Botanical Review* 54: 233–351.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

FERGUSON, I. K. 1966. Notes on the nomenclature of *Cornus*. *Journal of the Arnold Arboretum* 47: 100–105.

FLAVELL, R. B. 1986. Structure and control of expression of ribosomal RNA genes. *Oxford Survey of Plant Molecular Cell Biology* 3: 252–274.

HARMS, H. 1898. Cornaceae. *In* A. Engler and K. Prantl [eds.], Die naturalichen Pflanzenfamilien, Teil III, Abteilung 8, 250–270. W. Engelmann, Leipzig, Germany.

HARDIN, J. W., AND Z. E. MURRELL. 1997. Foliar micromorphology of *Cornus*. *Journal of the Torrey Botanical Society* 124: 124–139.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 21: 160–174.

HUTCHINSON, J. 1942. Neglected genetic characters in the family Cornaceae. *Annals of Botany* (London), series 2(6): 83–93.

JOHNSON, L. A., AND D. E. SOLTIS. 1998. Assessing congruence: empirical examples from molecular data. *In* D. E. Soltis, P. E. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants II, 297–347. Chapman and Hall, New York, New York, USA.

KUZOFF, R. K., J. A. SWEERE, D. E. SOLTIS, P. S. SOLTIS, AND E. A. ZIMMER. 1998. The phylogenetic potential of entire 26S rDNA sequences in plant. *Molecular Biology Evolution* 15: 251–263.

LÖVE, A., AND D. LÖVE. 1975. Cytotaxonomic atlas of the Arctic flora. J. Cramer, Vaduz, Germany.

MISHLER, B. D., L. A. LEWIS, M. A. BUCHHEIM, K. S. RENZAGLIA, D. J. GARBARY, C. F. DELWIICHE, F. W. ZECHMAN, T. S. KANTZ, AND R. L. CHAPMAN. 1994. Phylogenetic relationships of the ''green algae'' and ''bryophytes.'' *Annals of the Missouri Botanical Garden* 81: 451–483.

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 relationship in *Cornus* (Cornaceae). *Systematic Botany* 18: 469–495.

———. 1994. Dwarf dogwoods: intermediacy and the morphological landscape. *Systematic Botany* 19: 539–556.

NAKAI, T. 1909. Cornaceae in Japan. *Botanical Magazine Tokyo* 23: 35–45.

POJARKOVA, A. I. 1950. K voprosu o sistematicheskikh otnosheniyakh vnutri linneevskogo roda *Cornus* L. *Bot. Mater. Gerb. Bot. Inst. Komarova Akad. Nauk. SSSR* 12: 164–180.

RAFINESQUE, C. S. 1838. Alsograph. C. S. Rafinesque, Philadelphia, Pennsylvania, USA.

RO, K., C. S. KEENER, AND B. A. MCPHERON. 1997. Molecular phylogenetic study of the Ranunculaceae: utility of the nuclear 26S ribosomal DNA in inferring intrafamilial relationships. *Molecular Phylogenetics and Evolution* 8: 117–127.

SOLTIS, D. E., AND P. S. SOLTIS. 1997. Phylogenetic relationships among Saxifragaceae sensu lato: a comparison of topologies based in 18S rDNA and *rbc*L sequences. *American Journal of Botany* 84: 504–522.

———, AND ———. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. *In* D. E. Soltis, P. E. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants II, 1–42. Chapman and Hall, New York, New York, USA.

SPACH, E. 1839. Les Cornaceae. *In* Histoive naturelle des vegetaux, vol. 8, 86–110. Librarie Encyclopedique de Roret, Paris, France.

STEFANOVIC, S., M. JAGER, J. DEUTSCH, J. BROUTIN, AND M. MASSELOT. 1998. Phylogenetic relationships of conifers inferred from partial 28S rDNA gene sequences. *American Journal of Botany* 85: 688–697.

SUGIURA, M., Y. IIDA, K. OONO, AND F. TAKAIWA. 1985. The complete nucleotide sequence of a rice 25S rRNA gene. *Gene* 37: 255–259.

SWOFFORD, D. L. 2000. PAUP: phylogenetic analysis using parsimony, version 4.0b4. Sinauer, Sunderland, Massachusetts, USA.

VON BERCHTOLD, F., AND P. M. OPIZ. 1838. Cornaceae. *In* Oekonomisch-technische Flora Bohmens, Band 1, Teil 2, 167–180. T. Thabor, Prague, Czechoslovakia.

WANGERIN, W. 1910. Cornaceae. *In* A. Engler [ed.], Das Pflanzenreich, series IV, family 229 (Heft 41). W. Engelmann, Leipzig, Germany.

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

———, D. E. SOLTIS, 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.

———, ———, AND P. S. SOLTIS. 1998. Phylogenetic relationships of Cornaceae and close relatives inferred from *matK* and *rbcL* sequences. *American Journal of Botany* 85: 285–297.