American Journal of Botany 91(11): 1901–1914. 2004.

PHYLOGENY AND INFRAGENERIC CLASSIFICATION OF *SYMPLOCOS* (SYMPLOCACEAE) INFERRED FROM DNA SEQUENCE DATA[1]

YUGUO WANG,[2,5] PETER W. FRITSCH,[3,6] SUHUA SHI,[2,6] FRANK ALMEDA,[3] BONI C. CRUZ,[3] AND LAWRENCE M. KELLY[4]

[2]State Key Laboratory, School of Life Sciences, Zhongshan University, Guangzhou 510275, People's Republic of China; [3]Department of Botany, California Academy of Sciences, 875 Howard Street, San Francisco, California 94103 USA; and [4]The New York Botanical Garden, Cullman Program for Molecular Systematics Studies, 200th St. & Southern Blvd., Bronx, New York 10458 USA

Symplocos comprises ~300 species of woody flowering plants with a disjunct distribution between the warm-temperate to tropical regions of eastern Asia and the Americas. Phylogenetic analyses of 111 species of *Symplocos* based on the nuclear ribosomal internal transcribed spacer (ITS) region and the chloroplast genes *rpl16*, *matK*, and *trnL–trnF* yielded topologies in which only one of the four traditionally recognized subgenera (*Epigenia*; Neotropics) is monophyletic. Section *Cordyloblaste* (subgenus *Symplocos*; eastern Asia) is monophyletic and sister to a group comprising all other samples of *Symplocos*. Section *Palura* (subgenus *Hopea*; eastern Asia) is sister to a group comprising all other samples of *Symplocos* except those of section *Cordyloblaste*. *Symplocos wikstroemiifolia* (eastern Asia) and *S. tinctoria* (southeastern United States), both of subgenus *Hopea*, form a clade that groups with *S. longipes* (tropical North America) and the species of subgenus *Epigenia*. The remaining samples of subgenus *Hopea* (eastern Asia) form a clade. Section *Neosymplocos* (subgenus *Microsymplocos*; Neotropics) is well nested within a clade otherwise comprising the samples of section *Symplocastrum* (subgenus *Symplocos*; Neotropics). Section *Urbaniocharis* (subgenus *Microsymplocos*; Antilles) groups as sister to the clade comprising *Symplocastrum* and *Neosymplocos*. The data support the independent evolution of deciduousness among section *Palura* and *S. tinctoria*. The early initial divergence of sections *Cordyloblaste* and *Palura* from the main group warrants their recognition at taxonomic levels higher than those at which they are currently placed. An inferred eastern Asian origin for *Symplocos* with subsequent dispersal to the Americas is consistent with patterns from other phylogenetic studies of eastern Asian-American disjunct plant groups but contrary to a North American origin inferred from the earliest fossil occurrences of the genus.

Key words: disjunction; ITS; *matK*; phylogeny; *rpl16*; Symplocaceae; *Symplocos*; *trnL–trnF*.

Symplocos Jacq. comprises ~300 species of woody flowering plants distributed in the New World and the lands bordering the western Pacific Rim (Brand, 1901; Nooteboom, 1975; Ståhl, 1995). It is found primarily in humid tropical montane forests. Two taxa reach the north-temperate zone, one in eastern Asia (section *Palura* of subgenus *Hopea*, to 45° N) and one in the southeastern United States (*S. tinctoria*, to 37° N). These are the only taxa with the deciduous condition in the genus. In the Old World, *Symplocos* occurs no farther west than India. Its distribution corresponds to an ''amphi-Pacific tropical'' pattern of disjunction documented in ~100 genera and higher groups of seed plants (Steenis, 1962, 1963; Thorne, 1972).

Symplocos is recognized as the sole genus of Symplocaceae by recent authors (Cronquist, 1981; Takhtajan, 1997; Thorne, 2000), although the family has been divided into many genera in the older literature (e.g., Miers, 1880; Nakai, 1922, 1927; Hatusima, 1936). One of ~24 families comprising order Ericales sensu the Angiosperm Phylogeny Group (1998, 2003), the Symplocaceae are characterized by leaves without stipules; sympetalous, actinomorphic flowers; epipetalous, united, and numerous stamens; an inferior, incompletely loculed ovary; a simple style; unitegmic ovules; and a drupaceous fruit with a hard endocarp (Nooteboom, 1975). No other group of Ericales possesses this combination of characters, leaving little doubt that the family is monophyletic.

In the most recent comprehensive taxonomic revision of *Symplocos*, Brand (1901) recognized 277 species within four subgenera: *Symplocos*, *Epigenia*, *Hopea*, and *Microsymplocos*. In subgenus *Symplocos* (76 species, eastern Asian and American tropics), the androecium is adnate to the corolla for approximately half its total length, the stamens are monadelphous, and the filaments are complanate. In the other subgenera, the androecium is adnate to the corolla only at the base.

[1] Manuscript received 26 August 2003; revision accepted 3 August 2004.

The authors thank F. S. Axelrod, S. L. Dao, H. Li, H.-Y. Wang, J. Wen, Q.-Y. (J.) Xiang, B.-G. Yang, S.-D. Zhang, University of California Botanical Garden, Strybing Arboretum, Jardín Botanico Santo Domingo (JBSD), Hangzhou Botanical Garden, South China Botanical Garden, and the Missouri Botanical Garden (MO) for facilitating the field collection of some of the samples used in this study; D. E. Boufford, C. T. Chen, J. L. Clark, T. Clase, W. A. Haber, X. Gong, J. Homeier, S.-G. Jian, R. Kriebel, G. P. Lewis, A. B. Martins, J. F. Maxwell, S.-J. Ning, D. Penneys, B. Ståhl, C.-W. Wang, B.-Z. Xiao, F.-W. Xing, and W.-N. Ye for providing dried leaf material and herbarium vouchers of their field collections; the curators of KUN, IBSC, and MO for permission to sample leaf material from herbarium specimens; C. Elliger and L. H. Wu for invaluable laboratory assistance; and three anonymous reviewers for helpful comments on the manuscript. A special thanks goes to Y.-F. Deng and Y.-M. Shui who provided critical field collaboration in Hunan, Hainan, and Yunnan provinces and sent their collections to Y. Wang. This work was supported by United States National Science Foundation grant DEB–0126631 to P. W. Fritsch and F. Almeda, National Science Foundation of China grants 39825104 and 30230030 to S. Shi, and a grant from the California Academy of Sciences funded by the Lakeside Foundation to Y. Wang. Most laboratory work was conducted at the Osher Laboratory for Molecular Systematics at CAS.

[5] Current address: Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University (IBSFU), Shanghai 200433, People's Republic of China.

[6] Authors for correspondence (e-mail: pfritsch@calacademy.org; lssssh@zsu.edu.cn).

In subgenus *Epigenia* (20 species, Neotropics), the stamens are partly to completely distinct and the filaments are filamentous. In subgenus *Hopea* (167 species, eastern Asia except *S. tinctoria*), the stamens are pentadelphous and the filaments are filamentous. In subgenus *Microsymplocos* (14 species, Neotropics), the stamens are monadelphous, the filaments are claviform, and the flowers are uniformly small (2 to 5 mm long).

Brand (1901) divided each subgenus into two sections. Within subgenus *Symplocos*, section *Cordyloblaste* (six species, tropical eastern Asia) is distinguished from section *Symplocastrum* (70 species, Neotropics) by the filaments completely connate or distinct only for a short distance distally (vs. distally distinct from the point of separation from the corolla), and the ovary is two-locular (vs. three- to five-locular). Within subgenus *Epigenia*, section *Barberina* (17 species, southern Brazil) differs from section *Pseudosymplocos* (three species, Antilles) by the presence of morphological androdioecy (vs. hermaphroditism), up to 25 (vs. more numerous) stamens per flower, and the cotyledons shorter than the radicle (vs. longer). Within subgenus *Hopea*, section *Palaeosymplocos* (six species, eastern Asia and eastern North America) differs from section *Bobu* (161 species, eastern Asia) by its distinctly (vs. indistinctly) pentadelphous stamens. Within subgenus *Microsymplocos*, section *Neosymplocos* (11 species, southern Brazil) differs from section *Urbaniocharis* (three species, Antilles) by its pubescent (vs. glabrous) filaments and four-locular (vs. two-locular) ovary.

Several authors of regional treatments of *Symplocos* produced since the time of Brand's (1901) monograph have revised the infrageneric classification. In revisions of *Symplocos* of the Old World, Nooteboom (1975, 1977, 1980) pared the number of Asian species to 116 from Brand's 172 (including a reduction of section *Cordyloblaste* to two species) and revised Brand's subgeneric classification. Nooteboom (1975) placed subgenera *Epigenia* and *Microsymplocos* under subgenus *Hopea* based on palynological and scant phytochemical data. In a revision of the Japanese species of *Symplocos*, Nagamasu (1989a, b, 1993) recognized subgenus *Microsymplocos* and placed subgenus *Epigenia* under subgenus *Hopea* based on palynological characters. Bidá (1995), in revising the Brazilian species of the genus, upheld Brand's (1901) classification, citing unspecified floral and fruit characters and the palynological work of Barth (1979) as justification. These works represent significant advances in understanding the systematics of *Symplocos* but are limited by a focus only on the subgenera occurring within the geographic region under consideration.

Of Brand's (1901) sections, only those in subgenus *Hopea* have been disputed. Handel-Mazzetti and Peter-Stibal (1943) recognized sections *Bobu*, *Lodhra*, *Palaeosymplocos*, and *Palura* in a treatment of the Chinese species of *Symplocos* based on the following characters: leaf midrib adaxially impressed (vs. prominent); fruit shape; disk glabrous (vs. pubescent); and ovary two- (vs. three-) locular. Wu (1986a, b, 1987) added sections *Glomeratae* and *Singuliflorae* to the system of Handel-Mazzetti and Peter-Stibal (1943) based on a glomerulate and one-flowered inflorescence, respectively. Based on various morphological characters, Nagamasu (1993) classified the Japanese species of subgenus *Hopea* into eight sections, of which three (*Glaucae*, *Lancifoliae*, and *Okinawenses*) were newly described.

A phylogenetic study of 29 species of eastern Asian *Symplocos* with DNA sequence data from the nuclear ribosomal internal transcribed spacer (ITS) region, and the chloroplast *trnL–trnF* region (comprising the *trnL* [UAA] intron, the *trnL* [UAA] 3′ exon, and *trnL* [UAA]–*trnF* [GAA] spacer) and *trnH–psbA* spacer (Soejima and Nagamasu, 2004) has provided valuable insight into the classification and evolution of *Symplocos* in Japan. Because subdivisional sampling was limited, however, this study could not address the global classification of the genus. Here we test the various subdivisional classifications of *Symplocos* by including representatives from all subgenera and described sections of *Symplocos* in a phylogenetic analysis of the genus. We use DNA sequence data from the ITS region and three regions of the chloroplast genome: the *rpl16* intron, the *matK* gene, and the *trnL–trnF* region. Because of their relative ease of use and comparatively rapid sequence evolution, these regions have become standard sources of nucleotide characters in lower-level phylogenetic studies of plants. From the resultant phylogenetic patterns we explore character evolution in *Symplocos* (particularly that of leaf persistence, carpel number, and pedicel articulation) and the historical biogeography of the genus in the context of its amphi-Pacific disjunct distribution.

MATERIALS AND METHODS

Taxon sampling and DNA sequencing—Subdivisional classification of *Symplocos* follows Brand (1901) except for the sectional classification of the Chinese and Japanese species of subgenus *Hopea*, which follows Wu (1987) and Nagamasu (1993; Appendix, see Supplemental Data with the online version of this article). DNA sequences from 120 accessions of *Symplocos* were newly generated for this study (Appendix). This sample includes representatives of 90 *Symplocos* species, all four subgenera sensu Brand, and all sections recognized by Brand (1901), Wu (1987), and (in combination with the ITS data of Soejima and Nagamasu [2004]) Nagamasu (1993).

The sister group of Symplocaceae has long remained uncertain (Anderberg et al., 2002) but a recent molecular study of Ericales supports a clade consisting of Symplocaceae, Styracaceae, and Diapensiaceae and a more-inclusive clade comprising Theaceae, Roridulaceae, Actinidiaceae, Sarraceniaceae, Clethraceae, Cyrillaceae, and Ericaceae (Schöneberger et al., 2004). We chose the genus *Ternstroemia* (Theaceae [Cronquist, 1981] or Ternstroemiaceae [Angiosperm Phylogeny Group, 1998, 2003]) as outgroup for the combined chloroplast DNA and combined four-gene analyses because it is one of the few genera within Ericales for which sequences of all three chloroplast genes are available. Moreover, the genetic distance between *Symplocos* and *Ternstroemia* is low relative to those between *Symplocos* and representatives of 10 other families of Ericales (see Fritsch et al., 2001). The ITS analysis included 31 samples of *Symplocos* generated from Soejima and Nagamasu (2004). To further confirm the root of the *Symplocos* tree, we conducted phylogenetic analyses on separate ITS and *matK* data sets that included representatives of various families of Ericales as outgroups (Appendix).

Total DNAs were isolated with DNeasy Plant Mini DNA extraction kits (Qiagen, Inc., Valencia, California, USA) or with the cetyltrimethyl ammonium bromide (CTAB) method of Doyle and Doyle (1987) from desiccant- or air-dried leaf tissue. Some leaf samples (~20 mg) were obtained from herbarium specimens (Appendix). Prior to extraction, dried leaf tissue was pulverized by high-speed action of the Wig-L-Bug grinding mill (REFLEX Analytical Corp., Ridgewood, New Jersey, USA). PCR amplification was performed with standard methods (Dieffenbach and Dveksler, 1995) and BIOLASE (Bioline USA, Inc., Randolph, Massachusetts, USA) as the DNA polymerase. PCR products were purified with the Wizard PCR Preps DNA purification system (Promega, Madison, Wisconsin, USA). Cycle sequencing was performed with the ABI Prism BigDye Terminators v2.0 cycle sequencing reaction kit (Applied Biosystems, Foster City, California, USA) by using 1/4-scale reaction mixtures in a model 9600 PCR system thermal cycler (Perkin-Elmer, Boston, Massachusetts, USA). Sequences were determined with an ABI Prism 3100 genetic analyzer (Applied Biosystems) by obtaining forward

and reverse reads for all samples. Sequences were edited with the computer program Sequencher (3.0 and 4.1; Gene Codes, Ann Arbor, Michigan, USA) and all sequences have been deposited in GenBank (Appendix).

Amplification and sequencing of the ITS region employed primers ITS–4, ITS–5p, ITS–2p, and ITS–3p from Swensen et al. (1998). The *rpl16* intron was amplified with the forward primer F71 (5′–GCTATGCTTAGTGTGTGACTCGTTG–3′; Jordon et al., 1996) and the reverse primer L16 exon2 (Downie et al., 2000). Some samples were amplified with the forward primer rps3 (Downie et al., 2000) instead of F71. *rpl16* sequencing was performed with F71 and L16exon2, and the following internal primers designed specifically for *Symplocos*: 221F (5′–CTGATTCTAAGTTGTGAAGC–3′), 618F (5′–GCCGGGAAGCAATTAATCTA–3′), 191R (5′–TATTTCAGTTGTTACAATTA–3′), and 609R (5′–CCATCCCGACCAATGAATCA–3′). The *matK* gene was amplified with the primers *matK*–1F and *matK*–1R (Sang et al., 1997). The amplified region includes the entire *matK* coding region and a total of ~300 bp of the *trnK* intron that flanks the ends of *matK* (~70 bp at the 5′ end and ~230 bp at the 3′ end). *matK* sequencing was performed with the two amplification primers and the internal primers symp–1176F (5′–CAATTCATTCAMTATTTCCTT–3′), symp–1540F (5′–GTTCAAGGATCCTTTCATGC–3′), symp–2030F (5′–CTTCGACTTTCTTGTGCTAG–3′), symp–1988R (5′–ACGCCCGAATCGGTCAATAA–3′), symp–1470R (5′–AAGATATTAATCGTAAATGA–3′), and symp–866R (5′–CTATGATCATGAGCAAGTGC–3′) designed specifically for *Symplocos*. The *trnL–trnF* region was amplified and sequenced with the universal primers c, d, e, and f (Taberlet et al., 1991). Target sequences unsuccessfully amplified with the external primers were often successfully amplified in two fragments with an external and one of the internal primers.

Data analysis—Sequence alignment was manual except in the ITS analysis. The computer program ClustalX (Thompson et al., 1997) was used for ITS alignment because high sequence divergence among the outgroups and between the outgroups and ingroup made manual alignment problematic. Gap opening and gap extension parameters were 10 and 0.2, respectively. Aligned sequence matrices are available from the authors upon request. The computer program MacClade version 4.0 (Maddison and Maddison, 2000) was used to translate DNA sequences into protein sequences to aid the alignment of *matK* sequences. Gaps introduced into the alignment were treated as missing data. Various preliminary analyses with unambiguously aligned parsimony-informative gaps included had little effect on topology and clade support; these characters were therefore not included in final analyses.

Phylogenetic analyses employed maximum parsimony (MP) and Bayesian inference (BI). The MP analyses were conducted with the heuristic search option in PAUP* version 4.0b10 (Swofford, 2002). Searches were conducted over 100 random-taxon-addition replicates with tree bisection-reconnection branch-swapping, steepest descent, and MulTrees in effect. All characters and states were weighted equally and unordered. All trees from the replicates were swapped to completion, all shortest trees were saved, and a strict consensus tree was computed. To provide a complete search for the shortest tree in the ITS analysis, MulTrees was not enforced and no more than 3000 trees were saved per replicate. Relative support for individual clades was estimated with the parsimony bootstrap (bt) method (Felsenstein, 1985). One thousand pseudoreplicates were performed with uninformative characters excluded. Ten random-taxon-addition heuristic searches for each pseudoreplicate were performed and all minimum-length trees were saved per search. To expedite the search, MulTrees was not enforced in the ITS bootstrap analysis.

Bayesian analyses were conducted with MrBayes version 3.0b4 (Huelsenbeck and Ronquist, 2001) by using uniform prior probabilities and estimating base frequencies and the parameters for the HKY + Γ model. We ran four chains of the Markov chain Monte Carlo by beginning with a random tree and sampling one tree every 100 generations for 1 000 000 generations. The first 30 000 generations of the chain were used as ''burn in'' after stationarity was reached, and the phylogenetic estimate was based on trees sampled after generation 30 000. To estimate the posterior probability (pP) of recovered branches, 50% majority-rule consensus trees were created. An ITS phylogram was produced as average-branch-length consensus trees with MrBayes.

Data set congruence was determined with incongruence length difference (ILD) tests (Farris et al., 1994, 1995; implemented in PAUP* as the partition-homogeneity test) by using 500 heuristic-search randomizations, simple addition sequence, and TBR swapping, with uninformative characters excluded, one tree held at each step, MulTrees not enforced, and no more than 2000 trees saved per randomization. We conducted pairwise comparisons and a three-way comparison with the chloroplast data sets. We also conducted a pairwise comparison with the ITS data set and the combined chloroplast data set, and a four-way comparison of all data sets. After assessing data congruence, data sets were combined to provide a total-evidence phylogenetic estimate. Samples missing all data in one or more of the individual data sets (see Appendix) were excluded from combined analyses except in some analyses involving *Symplocos longipes* and *S. tenuifolia*.

We used Fitch parsimony optimization (Maddison et al., 1992) to assess the historical biogeography of *Symplocos* with the molecular topology. This method assumes that geographic distributions are solely the result of dispersal (as opposed to vicariance) events. Thus, polymorphic area states are restricted to terminal nodes. Ancestral states are inferred through minimizing the number of character state changes on the tree. The data matrix was constructed by coding ''area'' as a single multistate character, and the analysis was performed with MacClade 4.0 (Maddison and Maddison, 2000). The four areas circumscribed for the analysis were eastern Asia, North America, South America, and the Antilles. The combined four-gene tree was used for optimization, with all clades comprising operational taxonomic units (OTUs) from the same area reduced to individual OTUs. This reduction will have no effect on the interpretation of the results in the context of our biogeographic objectives. Analyses were conducted with *S. longipes* in either of two possible artificially resolved positions.

The evolution of ovary cell number and leaf persistence was inferred with Fitch parsimony optimization onto the combined four-gene reduced topology. Character states were scored for each OTU and the characters were optimized with MacClade 4.0 (Maddison and Maddison, 2000).

RESULTS

The 50% majority-rule consensus tree from BI analysis of each data set matches the strict consensus tree from the corresponding MP analysis in the topology of all major clades. Clades from each BI analysis not found in the corresponding MP analysis and vice-versa (either unambiguously or ambiguously through non-resolution if soft polytomies are assumed) are always poorly supported. For brevity, we therefore depict only the BI tree from each analysis but show both pP and bt values on the tree.

ITS analysis—The ITS data set included 143 OTUs, including 31 samples of *Symplocos* from Soejima and Nagamasu (2004) and eight outgroups, representing a total of 109 *Symplocos* species. The ITS region is the most length-variable of all gene regions analyzed in this study. Total sequence length in *Symplocos* ranges from 609 base pairs (bp; e.g., *S. celastrinea*) to 635 bp (e.g., *S. lanata*). The length of ITS 1 ranges from 246 bp (e.g., *S. berteroi*) to 253 bp (*S. chinensis* and *S. paniculata*). The 5.8S gene is consistently 164 bp long. The length of ITS 2 in *Symplocos* ranges from 198 bp (e.g., *S. celastrinea*) to 220 bp (e.g., *S. confusa*). All samples of subgenus *Epigenia* have a 20-bp deletion in ITS 2. Alignment with all outgroups resulted in a data set of 694 characters, of which 447 (64.4%) were variable and 339 (48.8%) were parsimony-informative.

Symplocos is monophyletic in both MP (1700 equally parsimonious trees of length = 1409, consistency index [CI] = 0.45, retention index [RI] = 0.78; bt = 100) and BI analyses of ITS sequences (pP = 0.93; Fig. 1A). The species of section *Cordyloblaste* sampled form a clade (pP = 0.96; bt = 100; Fig. 1B) that is sister to the clade (pP = 0.93; bt = 100)

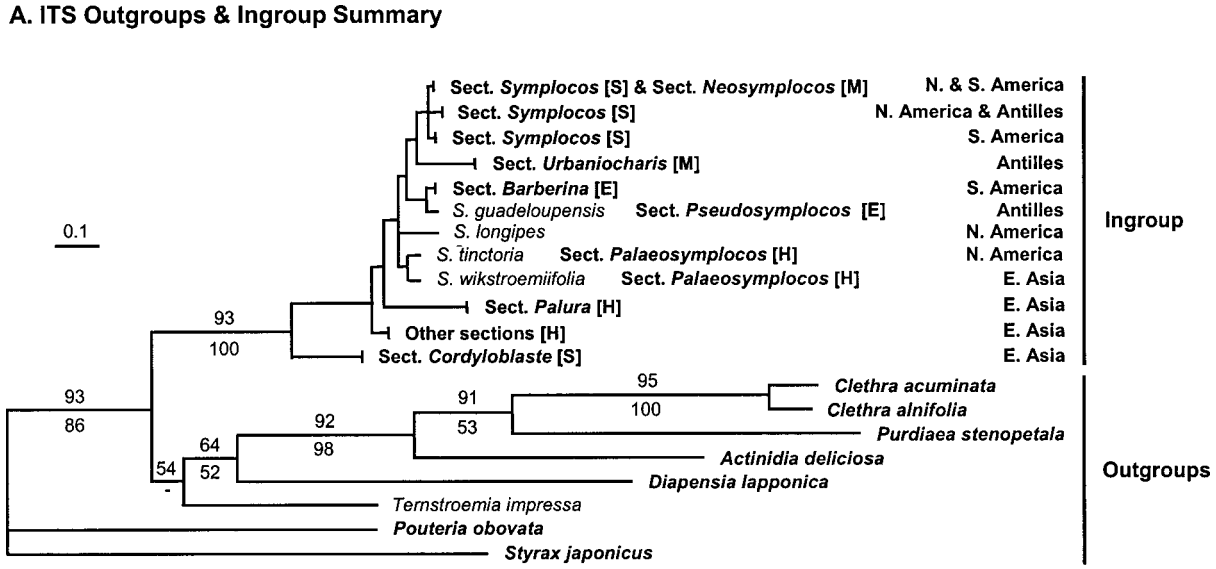

Fig. 1. Bayesian inference (BI) tree of *Symplocos* from analysis of ITS region DNA sequences. Numerals above each branch are BI posterior clade probabilities >50%. Numerals below each branch are maximum parsimony bootstrap values >50%. E = subgenus *Epigenia*; H = subgenus *Hopea*; M = subgenus *Microsymplocos*; S = subgenus *Symplocos*; U = subgenus uncertain; E. Asia = eastern Asia; N. America = North America; S. America = South America. A. ITS outgroups and ingroup summary. Geographic distribution is indicated at right. B. ITS ingroup. The main Asian subgenus *Hopea* clade has been depicted as a single branch for illustration purposes. Subgenera and sections are indicated at right. C. ITS main Asian subgenus *Hopea* clade. Sections are indicated at right. Boldface species names indicate samples from Soejima and Nagamasu (2004).

comprising all other *Symplocos* species. Other well-supported major clades comprise (1) section *Palura* (pP = 0.95; bt = 100); (2) the species of subgenus *Epigenia* (pP = 0.94; bt = 100); (3) the species of section *Barberina* (pP = 0.94; bt = 88); (4) the species of section *Symplocastrum* and *Urbaniocharis* (pP = 0.95; bt = 78); (5) most Asian species of subgenus *Hopea* (''main Asian *Hopea*'' clade; pP = 0.94; bt = 81; Fig. 1B, 1C); and (6) *S. tinctoria* and *S. wikstroemiifolia* (pP = 0.96; bt = 68; Fig. 1B). Other than the first divergence, the basal nodes of the *Symplocos* ITS topology are not well supported.

cpDNA analysis—Analysis of *matK* sequences with all outgroups yielded a single *Symplocos* clade (not shown). MP and BI analyses with either *Ternstroemia* alone or all outgroups included resulted in identical placements of the *Symplocos* root and a similar *Symplocos* topology. This root was also the same as that from the ITS analyses with all outgroups included. On this basis, we used *Ternstroemia* as the outgroup for the combined three-gene chloroplast analysis. None of the chloroplast data partitions (either pairwise or three-way) were significantly incongruent at the 0.05 level (ILD *P*-values range from 0.35 to 0.56). We therefore combined the three chloroplast data sets into a single data set for all subsequent analyses.

The combined chloroplast data set (*matK*, *rpl16*, and *trnL–trnF*) included 72 OTUs. The length of the *rpl16* intron in *Symplocos* ranges from 1008 bp in *S. confusa* to 1057 bp in *S. celastrifolia*. The aligned sequences comprised 1095 nucleotide characters (excluding the first 13 bp, which were often not readable or constant in those that were), of which 138 (12.6%) were variable and 49 (4.5%) were parsimony-informative. The length of the *matK* amplified fragment in *Symplocos* ranges from 1810 bp (e.g., *S. wikstroemiifolia*; including a coding region of 1506 bp) to 1831 bp (*S. confusa*; including a coding region of 1527 bp). The aligned sequences comprised 1855 nucleotide characters, of which 268 (14.4%) were variable and 78 (4.2%) were parsimony-informative. The length of the *trnL–trnF* region ranges from 905 bp in *S. paniculata* to 939 bp in *S. quitensis*. The length of the *trnL* intron ranges from 484 bp in *S. paniculata* to 518 bp in *S. quitensis*; that of the *trnL* [UAA] 3′ exon is consistently 50 bp; and that of the *trnL–trnF* spacer ranges from 354 bp in, e.g., *S. tortuosa* to 365 bp in *S. macrophylla*. The aligned sequences comprised 974 nucleotide characters, of which 115 were variable (11.8%) and 19 (2.0%) were parsimony-informative.

In both the MP (1899 equally parsimonious trees of length = 234, CI = 0.68, RI = 0.89) and BI analyses of combined chloroplast sequences, *Symplocos confusa*, representing section *Cordyloblaste*, is sister to the clade (pP = 0.99; bt = 100) comprising all other *Symplocos* species (Fig. 2). Other well-supported major clades comprise (1) section *Palura* (pP = 1.00; bt = 100); (2) the sister clade to section *Palura* (pP = 98; bt = 85); (3) the species of subgenus *Epigenia* (pP = 0.97; bt = 92); (4) the species of section *Barberina* (pP = 0.97; bt = 98); (5) subgenus *Epigenia*, *S. tinctoria*, and *S. wikstroemiifolia* (''WTLE'' clade; pP = 0.97; bt = 94); (6) section *Urbaniocharis* (pP = 0.99; bt = 71); and (7) the main Asian *Hopea* clade (pP = 0.99; bt = 99).

Combined four-gene analysis—The data partitions ITS, *matK*, *rpl16*, and *trnL–trnF* were not significantly incongruent in a four-way ILD test (*P* = 0.44). The ITS and combined chloroplast partitions were, however, significantly incongruent in a pairwise test (*P* = 0.02). After experimenting with various taxon exclusion sets, we recovered a nonsignificant *P*-value (*P* = 0.07) when just *Symplocos austrosinensis* and *S. quitensis* were removed from the ITS vs. combined chloroplast partitions. This suggests that lineage incongruence involving these species has resulted in conflict between the nuclear and chloroplast genomes, although there appears to be nothing about

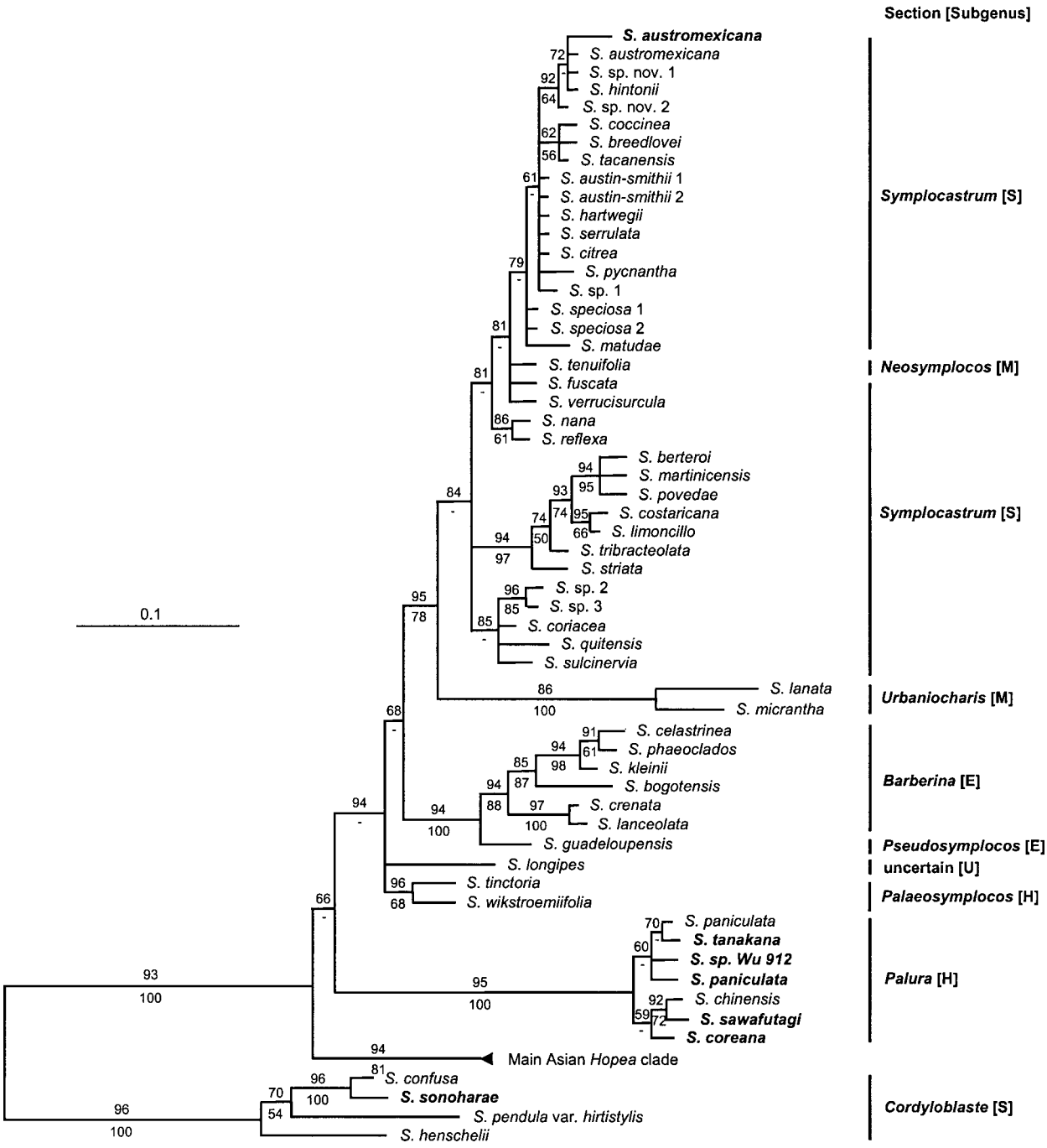

Fig. 1. Continued.

their sequence data, such as an unusually high number of polymorphic sites, to further support this. Visual comparison of the ITS and chloroplast topologies (Figs. 1 and 2) revealed no strong areas of conflict. We therefore combined all four data sets into a single analysis to provide a total-evidence phylogenetic estimate.

The combined four-gene data set included 70 OTUs. The aligned sequences comprised 4594 nucleotide characters, of which 777 (17.0%) were variable and 303 (6.6%) were parsimony-informative. Both the MP (128 equally parsimonious trees of length = 778, CI = 0.61, RI = 0.84) and BI analyses of the four-gene combined sequences yielded trees with higher clade resolution and support than those from ITS or chloroplast analyses alone (Fig. 3). Section *Cordyloblaste* is sister to the clade comprising all other *Symplocos* species (pP = 0.99; bt = 100). Within the latter, section *Palura* forms a clade (pP = 0.99; bt = 100) that is sister to the clade comprising the remaining species (pP = 0.99; bt = 94).

The remaining species form two well-supported clades. One is the main Asian *Hopea* clade (pP = 0.99; bt = 100). The other

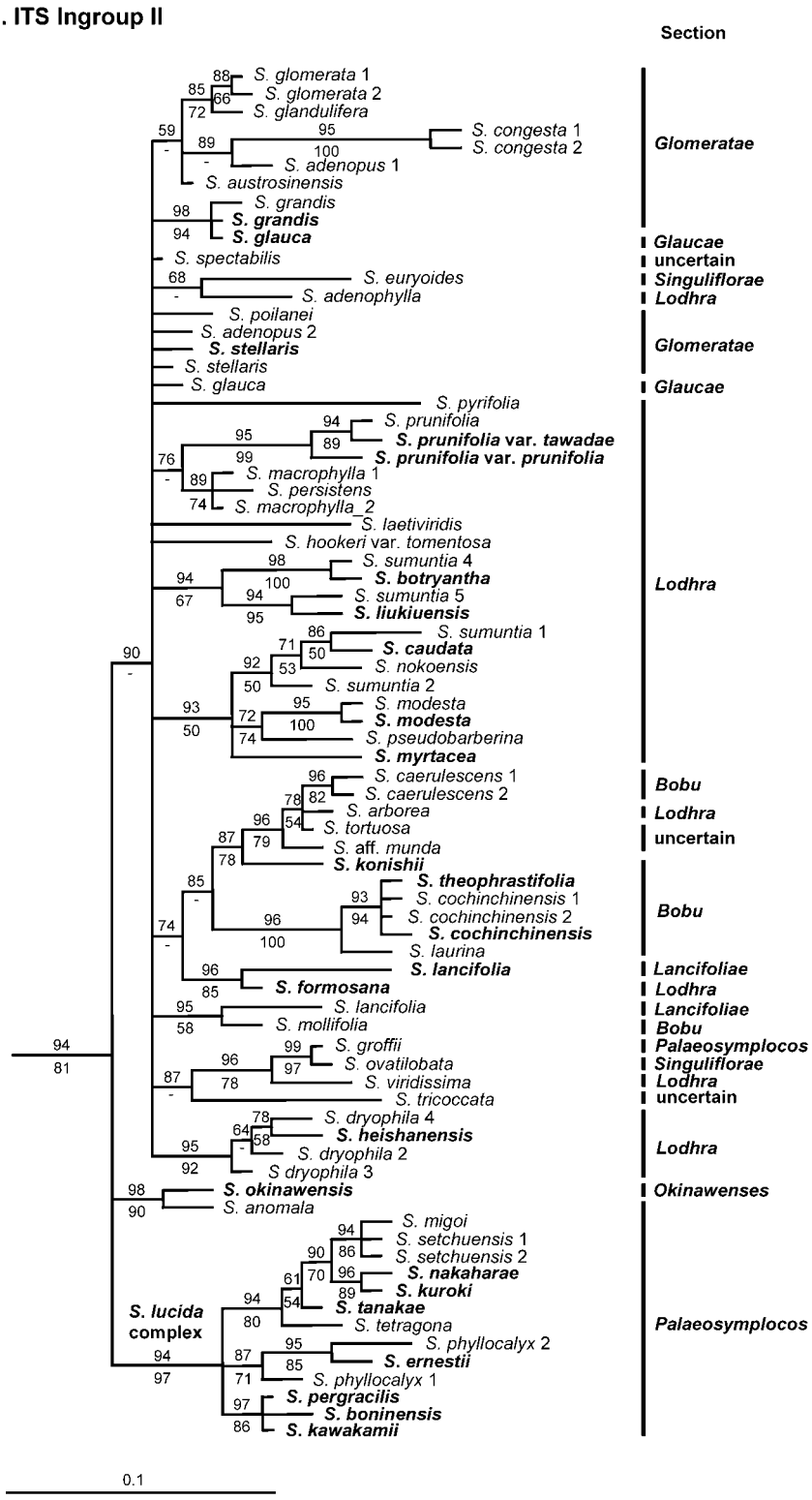

Fig. 1. Continued.

(pP = 0.98; bt = 64) consists of the WTLE clade (pP = 0.99; bt = 90), a clade comprising section *Symplocastrum* (pP = 0.99; bt = 90), and a clade comprising section *Urbaniocharis* of subgenus *Microsymplocos* (pP = 0.99; bt = 100). Section *Urbaniocharis* is sister to section *Symplocastrum* (pP = 0.99; bt = 87). Within the WTLE clade, *Symplocos tinctoria* and *S. wikstroemiifolia* form a monophyletic group (pP = 0.99; bt = 95), as do the species of subgenus *Epigenia* (pP = 0.99; bt = 100).

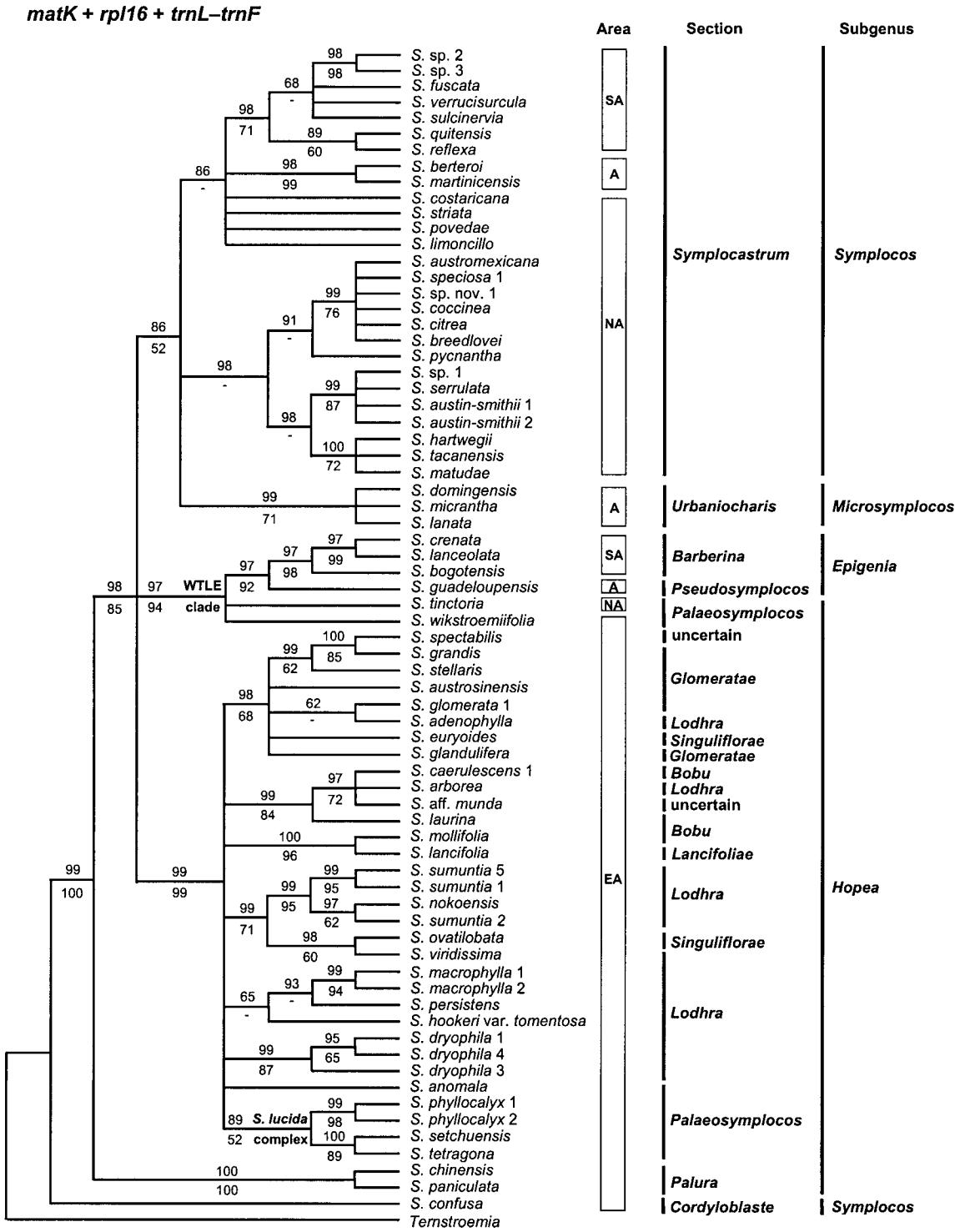

Fig. 2. Bayesian inference (BI) tree of *Symplocos* from combined analysis of *matK, rpl16,* and *trnL–trnF* cpDNA sequences. Numerals above each branch are BI posterior clade probabilities >50%. Numerals below each branch are maximum parsimony bootstrap values >50%. Geographic distribution, subgenus, and section of each species are indicated at right. The WTLE clade and that of the *S. lucida* complex (see text) are indicated. A = Antilles; EA = eastern Asia; NA = North America; SA = South America.

Analysis of Symplocos longipes and S. tenuifolia—*Symplocos longipes* and *S. tenuifolia* are significant for both classification and historical biogeography, because *S. longipes* has not been classified as to subdivision, and *S. tenuifolia* is the only representative of section *Neosymplocos* (subgenus *Microsymplocos*). The positions of *S. longipes* and *S. tenuifolia* sampled for a subset of the four genes (through lack of amplification for the missing gene regions) were assessed in separate analyses. In a combined analysis of the ITS and *rpl16* data sets, *S. longipes* forms a trichotomy with the clade comprising

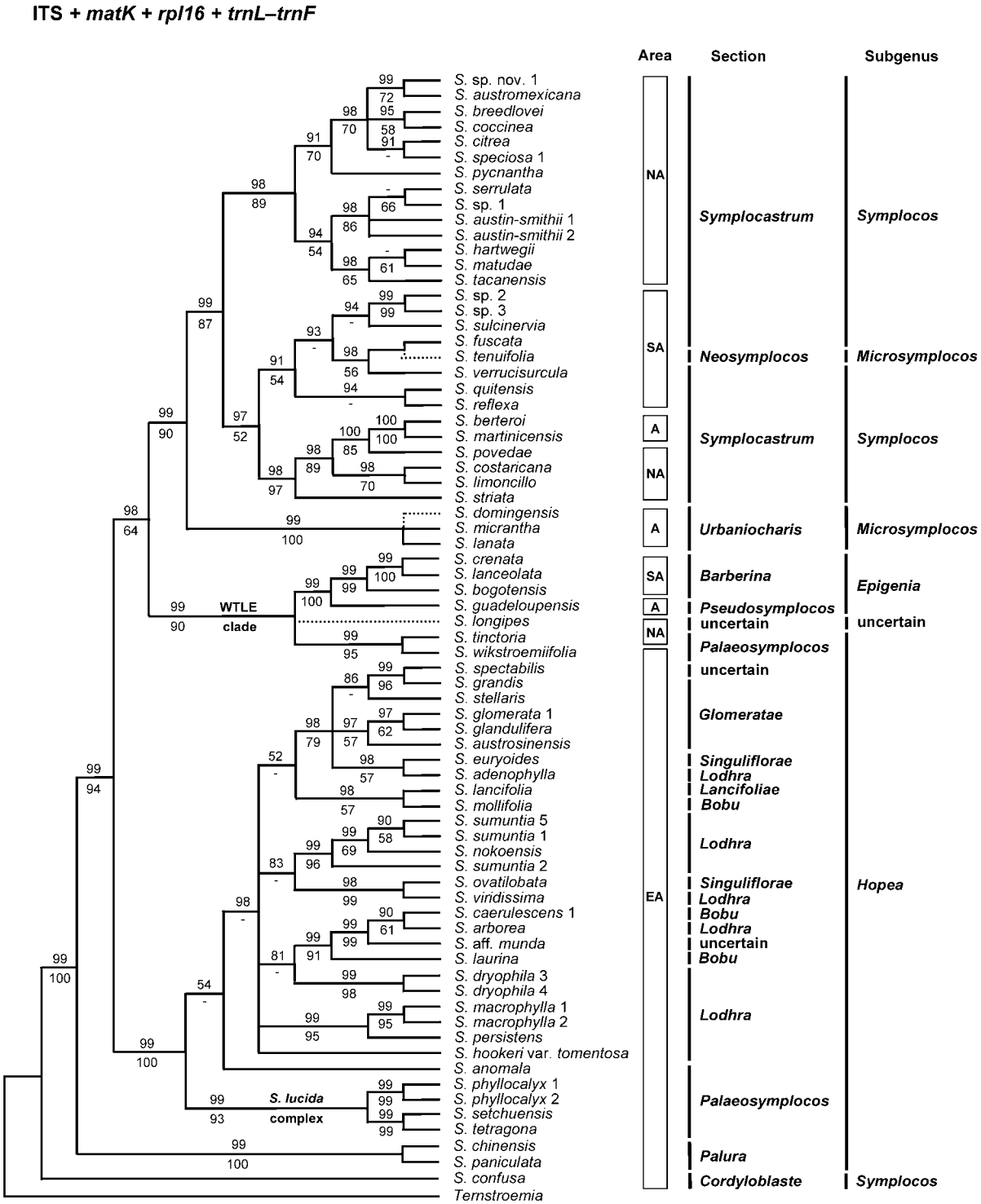

Fig. 3. Bayesian inference (BI) tree of *Symplocos* from combined analysis of ITS, *matK*, *rpl16*, and *trnL–trnF* DNA sequences. Numerals above each branch are BI posterior clade probabilities >50%. Numerals below each branch are maximum parsimony bootstrap values >50%. Geographic distribution, subgenus, and section of each species are indicated at right. The WTLE clade and that of the *S. lucida* complex (see text) are indicated. Broken lines indicate branches leading to taxa that lack sequence data from one or more individual data sets and have been placed in the topology on the basis of separate analyses (see text). A = Antilles; EA = eastern Asia; NA = North America; SA = South America.

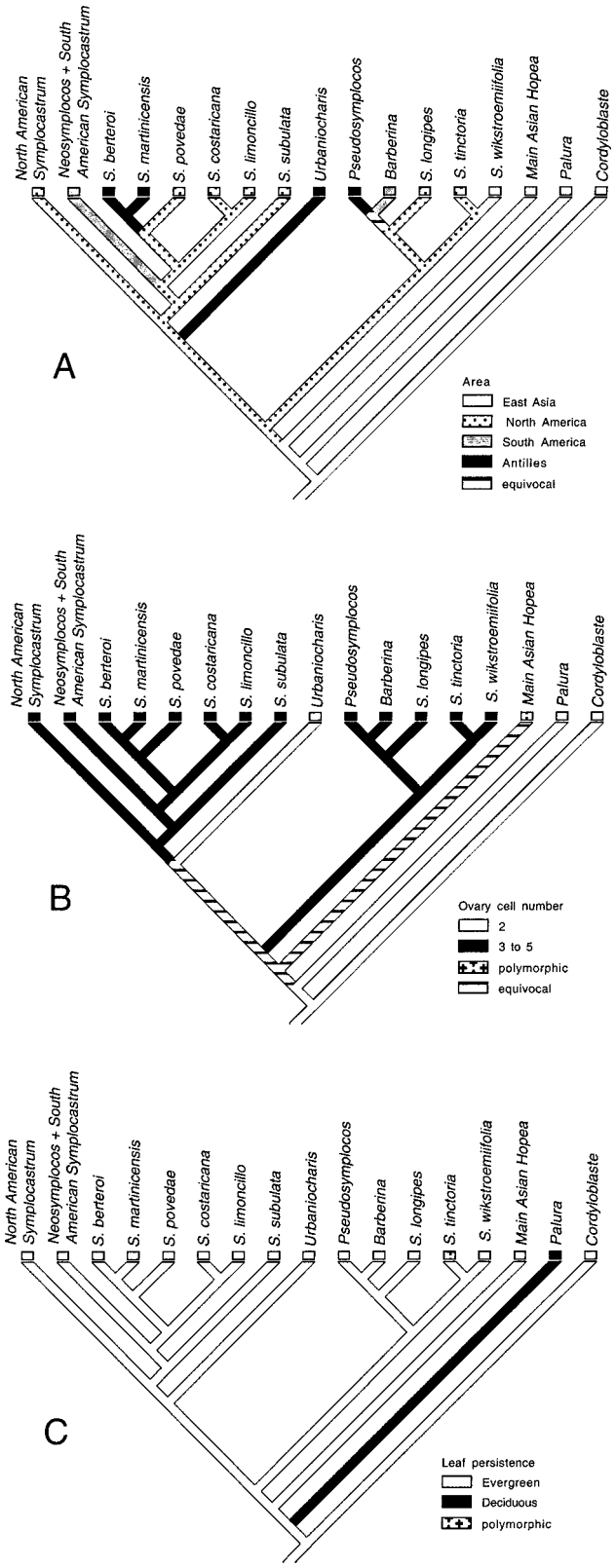

Fig. 4. Fitch parsimony optimization of several characters onto the *Symplocos* four-gene combined topology (see Fig. 3) with all clades comprising operational taxonomic units (OTUs) from the same area reduced to individual OTUs. *Symplocos longipes* is artificially resolved as sister to the clade comprising sections *Pseudosymplocos* and *Barberina* (subgenus *Epigenia*; see text). A. Area. There are three equally optimal reconstructions of seven steps in the topology displayed. The three reconstructions differ only in the assignment of the stem lineage of the clade comprising *Pseudosymplocos* and *Barberina* (indicated with a horizontal line; Antilles, North America, or South America). See text for additional analyses. B. Ovary cell number. There are three equally optimal reconstructions of three steps (the main Asian *Hopea* clade is polymorphic). C. Leaf persistence. There is one optimal reconstruction of two steps (*S. tinctoria* is polymorphic).

subgenus *Epigenia* and that comprising *S. tinctoria* and *S. wikstroemiifolia* (Fig. 3; pP = 0.96, bt = 63 [not shown]). In a combined analysis of the ITS and *matK* data sets, *S. tenuifolia* groups in a clade with all samples of section *Symplocastrum* (*S. tenuifolia* + section *Symplocastrum*, pP = 0.99; bt = 0.91). In an additional parsimony analysis, *S. longipes* and *S. tenuifolia* were included with all OTUs having complete four-gene data. This analysis yielded a topology with *S. longipes* in the same position as in the combined ITS and *rpl16* analysis (bt = 84) and *S. tenuifolia* as highly nested within the rest of South American subgenus *Symplocastrum* as sister to *S. fuscata* (Fig. 3; bt = 89 [not shown]).

Biogeographical analysis—Fitch parsimony optimization with *Symplocos longipes* resolved as sister to subgenus *Epigenia* resulted in three equally optimal reconstructions of seven steps (Fig. 4A). The three reconstructions differ only in the assignment of the stem lineage of the clade comprising sections *Pseudosymplocos* and *Barberina* (Antilles, North America, or South America).

The most recent common ancestor of *Symplocos* is assigned as eastern Asia, as are the two nodes above it. Above these nodes, a dispersal event from eastern Asia to North America is inferred, followed by back-dispersal to eastern Asia along the *S. wikstroemiifolia* stem lineage. In one of the three reconstructions, dispersals from North America to the Antilles and from North America to South America are inferred along the branches leading to sections *Pseudosymplocos* and *Barberina*, respectively. Three other dispersals are inferred, two to the Antilles (one along the stem of section *Urbaniocharis*, the other along the stem of the clade comprising *S. berteroi* and *S. martinicensis*) and one to South America (along the stem of the clade comprising the South American species of subgenus *Symplocastrum* and section *Neosymplocos*).

Fitch parsimony optimization with *Symplocos longipes* resolved as sister to the clade comprising *S. tinctoria* and *S. wikstroemiifolia* resulted in four equally optimal reconstructions of seven steps (not shown). Three of the four optimizations are identical to those reconstructed from the analysis with *S. longipes* resolved as sister to subgenus *Epigenia*. In the fourth, a dispersal from eastern Asia to the Antilles is inferred along the stem lineage immediately above the divergence of the main Asian *Hopea* clade. This is followed by two dispersals to North America, one along the stem of the clade comprising *S. longipes*, *S. tinctoria*, and *S. wikstroemiifolia*, the other along the stem immediately above the divergence of section *Urbaniocharis*. The remaining dispersals are identical to those inferred from other optimization analyses.

DISCUSSION

Sequence analysis—In the percentage of phylogenetically informative characters and the resolution of major clades, the cpDNA sequence data consistently exhibit lower levels of phy-

logenetic information than the nuclear ITS region. The ITS region resolved most clades except WTLE, whereas *matK* and *rpl16* resolved WTLE but not the relationship between sections *Urbaniocharis* and *Symplocastrum* (not shown). Among cpDNA data sets, *rpl16* displays only a slightly higher number of parsimony informative sites (4.5%) than does the *matK* region (4.2%). Separate analyses of each data set resulted in similar topologies, the only major difference being that sections *Urbaniocharis* and *Symplocastrum* form an unresolved clade with low bootstrap support in the *rpl16* tree, but are unresolved with respect to each other (and section *Neosymplocos*) in the *matK* tree (not shown). The *trnL–trnF* region exhibits the lowest level of phylogenetic information among the four genes (2.0% parsimony informative sites). Nonetheless, in analyses with and without the *trnL–trnF* region included, we observed that the use of this region in combined analyses substantially increased bootstrap support values for some major clades of the tree. The combined four-gene data set demonstrated a greater ability to recover phylogenetic relationships than did data from single gene regions. As in other work involving the combination of large molecular data sets (e.g., Soltis et al., 1998; Hoot et al., 1999), the four-gene parsimony analysis of *Symplocos* had a shorter computer run time and resulted in more highly resolved trees and higher support values than in analyses with separate data sets, although we caution that the higher number of taxa in some of the data sets vs. the four-gene combined data set could have also affected run times.

Subgeneric classification and character evolution—The combined analyses support only one of the four subgenera of *Symplocos* (subgenus *Epigenia*) recognized by Brand (1901). The data support neither subgenus *Epigenia* (Nooteboom, 1975) nor subgenus *Microsymplocos* (Nooteboom, 1975; Nagamasu, 1989a, b, 1993) as synonyms of subgenus *Hopea*. The characters on which previous classifications have been based are assessed below in the context of their use in classification and their evolution. This is done with the understanding that the best way to resolve optimizations of morphological characters is to incorporate a morphological data matrix into studies of *Symplocos*, the construction of which is in progress (L. M. Kelly et al., unpublished data).

Subgenus Symplocos—The widely separated placement of sections *Cordyloblaste* and *Symplocastrum* in the topologies suggests that some or all of the characters used to define subgenus *Symplocos* are homoplasious or plesiomorphic. These characters include the androecium adnate to the corolla for approximately half its total length (vs. only at the base), stamens monadelphous (vs. pentadelphous), stamen filaments complanate (vs. filamentous), and (a character mentioned in Nooteboom [1975]) stamen filaments constricted (vs. gradually tapered) distally. Upon further investigation, we have observed several characters that appear to distinguish section *Cordyloblaste* from all other species of the genus: calyx teeth apices truncate (vs. rounded to acute), corolla lobes papillate (vs. smooth) adaxially, ovary semi-inferior (vs. completely inferior), and an articulation between the hypanthium and the pedicel absent (vs. present). Whereas the first three characters may be autapomorphic for the section, the last is probably plesiomorphic in *Symplocos* because articulated pedicels are both uncommon and scattered throughout the Ericales clade, and nonarticulated pedicels therefore are likely to be the plesiomorphic state of the entire order (Fritsch et al., 2001). As such, the presence of articulation is a synapomorphy for the non-*Cordyloblaste* clade in *Symplocos*.

Chromosome and pollen data further distinguish section *Cordyloblaste* from the rest of the genus. The chromosome number of section *Cordyloblaste* is $2n = 90$ vs. $2n = 22$, 24, and possibly 28 for the rest of the species of the family that have been sampled (Nooteboom, 1975; Nagamasu, 1993), suggesting that this section is octoploid. No taxa of section *Symplocastrum* have yet been sampled cytologically for comparison. The three-colporate pollen grains of section *Cordyloblaste* with a massive tectum and a smooth perforate exine (*Henschelii* subtype; Meijden, 1970; Liang, 1986) differ from those of section *Symplocastrum* (type 2: two-colporate, exine smooth or more or less punctate; and type 4: three-colporate, exine verrucose; Barth, 1979, 1982). They also differ from those of subgenus *Epigenia* (three-colporate, tectum reduced, exine baculate and finely or coarsely verrucose) and subgenus *Hopea* (tectum thin, exine with supratectal structure; Meijden, 1970; Liang, 1986).

Section *Cordyloblaste* has not been accorded generic status since its initial treatment at the generic level (Moritizi, 1848; Miers, 1880). Nonetheless, the early diversification of this group in combination with its unique morphological features and chromosome number provide justification for reassessing the level at which this taxon is best recognized. Variation in pedicel articulation may prove to be particularly informative in this regard. The presence of articulated pedicels is a synapomorphy of a major clade above the genus level within Styracaceae (Fritsch et al., 2001) and appears to be genus-specific throughout most of the order Ericales (P. W. Fritsch, unpublished data). The recognition of *Cordyloblaste* at the genus level is therefore generally consistent with overall generic concepts in Ericales.

Subgenus Hopea—Section *Palura* of subgenus *Hopea* sensu Wu (1987) is treated by various authors as constituting anywhere from one to five species (Nooteboom, 1975; Wu, 1987; Wu and Nooteboom, 1996; Nagamasu, 1993). It is the only group in *Symplocos* that possesses exclusively terminal panicles and suprareticulate pollen exine (Barth, 1979, 1982; Liang, 1986; Nagamasu, 1993), corroborating its monophyly as supported here. The deep divergence of section *Palura* in the topology was not predicted from the classifications of Brand (1901), who placed the complex in section *Bobu* subsection *Palura* with many other species, or Wu (1987) and Nagamasu (1993), who placed it in its own section within subgenus *Hopea*. Two morphological characters, however, corroborate our results. Section *Palura* has a chromosome number of $n = 11$, a probable synapomorphy for all *Symplocos* species excluding those of section *Cordyloblaste*. It also has a two-locular ovary (vs. three- to five-locular in the rest of the genus). Two-locular ovaries are shared among the members of sections *Palura*, *Cordyloblaste*, and *Urbaniocharis*, and some members of the main Asian *Hopea* clade (Brand, 1901; Nagamasu, 1993). Contrary to the hypothesis of Nagamasu (1993), our results unambiguously support the ancestral two-locular condition in *Symplocos* followed by various equally optimal scenarios that include a reversal to the two-locular condition (Fig. 4B).

Miers (1880) grouped the Asian deciduous species (section *Palura* sensu Wu, 1987) with *Symplocos pendula* of section *Cordyloblaste* sensu Brand (1901) as the sole members of the

genus *Palura*. Although our data do not support the genus *Palura* sensu Miers (1880), the deep divergence of the Asian deciduous clade within *Symplocos* does support Miers' recognition of this group at a level higher than section. If section *Cordyloblaste* were raised to the generic level, then it would be reasonable to recognize section *Palura* sensu Wu (1987) as one of two subgenera of *Symplocos*, the other comprising the members of its sister group.

Section *Palura* is one of only two deciduous elements in *Symplocos*, the other being *S. tinctoria* (also in subgenus *Hopea*) of the southeastern United States. *Symplocos tinctoria* is deciduous in the northern portion of its range. In warmer areas to the south, the leaves from the previous season remain through the spring flowering period and the unfolding of new leaves (Elias, 1980); thus, *S. tinctoria* in these regions is considered evergreen or nearly so (Peattie, 1950; Radford et al., 1964; Elias, 1980; Little, 1980). Section *Palura* appears to be consistently deciduous throughout its range. Because section *Palura* and *S. tinctoria* are placed in widely separated regions of the topology, the data support the independent evolution of deciduousness from the evergreen condition in these lineages (Fig. 4C).

The WTLE clade does not accord with classifications of *Symplocos*. Some aspects of pollen morphology, however, appear to support this clade. The three-colporate pollen type with verrucose tectum occurs in both *S. wikstroemiifolia* and subgenus *Epigenia* (Barth, 1982; Liang, 1986). The pollen of *S. tinctoria* is also three-colporate but has a thin perforate tectum, distinct supratectal ornamentation, and a thin but distinct columella layer (Nagamasu, 1989a). This wall structure is the same as Barth's type 2 *celastrinea* type characterizing subgenus *Epigenia* (Nagamasu, 1989a), suggesting a close relationship between *S. tinctoria* and subgenus *Epigenia*. The pollen of *S. longipes* has not been described. The combined data place the WTLE clade as sister to the clade comprising section *Symplocastrum* and subgenus *Microsymplocos*. Clade support for this placement, however, is inconclusive, with a high posterior clade probability (0.98) but a fairly low bootstrap percentage (64; Fig. 3). The alternative position of the WTLE clade as sister to the main Asian *Hopea* clade would accord better with the placement of subgenus *Epigenia* as a synonym of subgenus *Hopea*, as in Nooteboom (1975) and Nagamasu (1993). Higher confidence regarding the placement of this clade will require more phylogenetic data.

The sister-group relationship between *Symplocos tinctoria* and *S. wikstroemiifolia* (continental southeastern Asia, including southern China) conflicts with previous research on the systematics of *Symplocos*. Brand (1901) and Nooteboom (1975) suggested an affinity of *S. tinctoria* with *S. lucida*, a complex treated as one (Nooteboom, 1975; Wu and Nooteboom, 1996) or seven (Wu, 1987) species. Wu (1987) placed *S. wikstroemiifolia* in the same section (*Palaeosymplocos*) as a member of the *S. lucida* complex but made no mention of *S. tinctoria*. Our data place the *S. lucida* complex within the main Asian *Hopea* clade instead of with *S. tinctoria* (Figs. 2, 3). Like molecular data from other groups with a distribution in eastern Asia and eastern North America (see Wen, 1999), those from *Symplocos* appear to reject a disjunct species-pair hypothesis based on morphology in favor of an alternative set of relationships (in our case, a different species pair). Whether this is due to morphological stasis (symplesiomorphy), convergent evolution, or another process has yet to be assessed.

Neither Brand's (1901) nor Wu's (1987) sectional classifications are supported by the data. Of Wu's (1987) seven sections in subgenus *Hopea*, only section *Palura* (see above) is unambiguously supported as monophyletic. Within section *Palaeosymplocos*, the data support the monophyly of the *Symplocos lucida* complex, but *S. anomala* groups as the first-diverging lineage within the main Asian *Hopea* clade (four-gene analysis), *S. wikstroemiifolia* groups in the WTLE clade (cpDNA and four-gene), and *S. groffii* is nested within a clade otherwise comprising species of sections *Lodhra* (*S. nokoensis*, *S. sumuntia*, and *S. viridissima*) and *Singuliflorae* (*S. ovatilobata*; cpDNA and four-gene). *Symplocos euryoides*, the other species in section *Singuliflorae*, groups strongly with the species of section *Glomeratae* and *S. adenophylla* of section *Lodhra* (four-gene analysis). The samples of section *Bobu* occur in three separate clades (all analyses). The samples of section *Glomeratae* form two clades that are unresolved with respect to a clade comprising *S. adenophylla* and *S. euryoides* in the highest-resolution analysis (four-gene analysis). Neither of Brand's (1901) sections of subgenus *Hopea* are supported because section *Palaeosymplocos* includes *S. tinctoria* and section *Bobu* includes the Asian deciduous species (Wu's section *Palura*).

Because Nagamasu's (1993) sections of subgenus *Hopea* were delimited only in the context of the Japanese species of *Symplocos*, they are difficult to assess with more-inclusive phylogenetic data. The three new sections described by Nagamasu each consist of one or two species. Section *Okinawenses* (containing only *S. okinawensis*) groups with S. *anomala*. Two samples of section *Lancifoliae* were analyzed (both *S. lancifolia*; the only other species in the section, *S. microcalyx*, was not sampled). The sample from Soejima and Nagamasu (2004) groups with *S. formosana* (section *Lodhra*), whereas ours groups with *S. mollifolia* (section *Bobu*). Similarly, the two samples of section *Glaucae* (both *S. glauca*) occur in different positions. Because we have not seen vouchers from the study of Soejima and Nagamasu (2004), we cannot discern the reason for the different positions of the samples of these sections.

Although our data provide resolution of some of the relationships within the main Asian clade of subgenus *Hopea* (i.e., section *Glomeratae*, the *Symplocos lucida* complex, and several small clades), the clade is internally neither well-resolved nor strongly supported in any analysis. If clades within this group are to be formally recognized, parts of the topology suggest possible recircumscriptions. For example, section *Glomeratae* sensu Wu (1987) may be expanded to accommodate all species with glomerulate inflorescences (e.g., *S. adenophylla*, *S. adenopus*, *S. austrosinensis*, *S. congesta*, *S. euryoides*, *S. glandulifera*, *S. glauca*, *S. glomerata*, *S. grandis*, *S. poilanei*, *S. spectabilis*, and *S. stellaris*, as sampled here). Also, the small sections *Glaucae*, *Lancifoliae*, and *Okinawenses* may be best treated as synonyms because they each group strongly with species from other sections. Although visual inspection of the four-gene data set suggests that homoplasy accounts for some of the lack of resolution, most clade ambiguity appears to be from low sequence divergence, suggesting rapid speciation in this clade relative to others that exhibit higher internal resolution. In this context, the data support Nooteboom's (1975) view of avoiding sectional delimitation in subgenus *Hopea* altogether as the most appropriate taxonomic approach to the main Asian *Hopea* clade.

Subgenus Microsymplocos—The apparent polyphyly of subgenus *Microsymplocos* conflicts with the characters used by Brand (1901) to define this subgenus. Observations on the species of section *Urbaniocharis* (P. W. Fritsch, unpublished data), however, contradict two of Brand's character descriptions for subgenus *Microsymplocos*. All species examined possess a pentadelphous rather than monadelphous arrangement of the stamens, and the filaments are filamentous, not claviform. Because more inclusive clades from the molecular data exclusively have these states as well (i.e., all except those of subgenus *Symplocos*), the placement of section *Urbaniocharis* as sister to section *Symplocastrum* supports the ancestral condition of these characters in section *Urbaniocharis*. From field and herbarium specimen observations, characters that appear to be unique to section *Urbaniocharis* (and possibly subgenus *Microsymplocos*) are the cupuliform corolla, stamens more or less appressed to the inner corolla surface, and the stamens shorter than the corolla, all of which support the monophyly of the section.

Field observations are currently lacking for species of section *Neosymplocos*. From examination of herbarium specimens, however, the filaments appear flattened and at least sometimes strongly constricted at the apex, as in section *Symplocastrum*. The nested placement of section *Neosymplocos* within section *Symplocastrum* supports the shared derived status of these characters throughout this clade. The data support the independent evolution of small flowers common to all species of sections *Neosymplocos* and *Urbaniocharis*. This convergence may have been influenced by the predominance of minute endemic insects that serve as pollinators for much of the flora of the Greater Antilles (Borhidi, 1996).

The pollen of section *Neosymplocos* is similar to that of section *Symplocastrum* in its indistinct columellae and massive tectum without supratectal ornamentation (Barth, 1979, 1982), consistent with our results that group these two sections together. Also consistent with our data, Mai (1986) has distinguished a ''tenuifolia'' type of pollen characterizing section *Neosymplocos* from the ''celastrinea'' and ''lanata'' types that are found in section *Urbaniocharis*. The tenuifolia type has a triangular-suboblate grain with reticulate-punctate exine, whereas those of the celastrinea and lanata types have a circular-oblate or -suboblate grain with verrucose or areolate exine, respectively (Mai, 1986).

The only chromosome number known for subgenus *Microsymplocos* is $n = 12$ (*Symplocos micrantha* of section *Urbaniocharis*). This number differs from all non-*Cordyloblaste Symplocos* sampled ($n = 11$ and possibly $n = 14$). It is unclear whether the interpretation of the *S. micrantha* chromosome data is correct or, as Nooteboom (1975) surmised, erroneous from the presence of B chromosomes.

Biogeographical implications—The earliest record of *Symplocos* known is fossil pollen from the western United States dating from the Late Cretaceous (Campanian-Maastrichtian, 84–65 million years ago), but there remain doubts as to whether this pollen should be classified as *Symplocos* (Krutzsch, 1989). Fossil endocarps and pollen grains are known from the Paleocene to the Miocene in North America, from the Lower Eocene to the Pliocene in middle Europe, and from the Upper Eocene and the Pliocene in Japan (Kirchheimer, 1949; Nooteboom, 1975; Krutzsch, 1989; Manchester, 1999). From the fossil record, Krutzsch (1989) has synthesized the following scenario for the historical biogeography of *Symplocos*. After a Late Cretaceous or Paleocene North American origin, *Symplocos* migrated across the North Atlantic and by the Eocene was widespread throughout the Northern Hemisphere. Range restrictions in the Miocene displaced the group southward and dispersal to South America occurred in the Pliocene. No early Tertiary fossils of *Symplocos* are yet known from Africa, India, Australia, or southern Southeast Asia (Krutzsch, 1989).

Because the first-diverging lineages are all from eastern Asia, the data support an eastern Asian origin for the geographic origin of *Symplocos* rather than the North American origin hypothesized by Krutzsch (1989; Fig. 4A). Our inferences are based on molecular phylogeny, whereas those of Krutzsch (1989) are based primarily on spatiotemporal fossil data. Phylogenies based on extant taxa in which the basal nodes are exclusively eastern Asian are proving to be a predominant pattern among disjuncts between eastern Asia and eastern North America (Donoghue et al., 2001; Xiang and Soltis, 2001; Wen et al., 2002). This implies a geographic origin in eastern Asia for many Northern Hemisphere disjuncts. From perspectives based on fossil data, however, it has been proposed that eastern Asia might best be considered a large refugium for many taxa of mixed-mesophytic forests once distributed widely across the Northern Hemisphere that have since become extinct in other areas (Tiffney, 1985a, b; Manchester, 1999). Although the earliest *Symplocos* fossils are from North America, those from the Eocene of Japan suggest that the rich diversity of Symplocaceae in eastern Asia today resulted from a long and continuous presence of the family in eastern Asia throughout most of the Tertiary.

The six area optimizations of *Symplocos* that infer dispersal from eastern Asia to North America (followed by dispersals to the Antilles and South America) are consistent with a scenario in which one of the Northern Hemisphere high-latitude land bridges (Beringian or North Atlantic) provided a means of overland migration between eastern Asia and North America. The sole area optimization that infers dispersal from Asia to the Antilles can only be accommodated through extinction in North America, or long-distance dispersal. A pattern of dispersal from an eastern Asian origin through North America to the Antilles and South America is also the most likely migration route for the amphi-Pacific tropical group *Styrax* section *Valvatae* (Styracaceae; Fritsch, 2001, 2003). Like *Symplocos*, this is an ericalean group in which the vast majority of species are montane. The identical pattern of dispersal inferred from the phylogeny of *Styrax* and *Symplocos* suggests a common geohistorical explanation for both, and should prompt further research into the historical biogeography of other montane ericalean amphi-Pacific tropical genera (e.g., *Gaultheria* [Ericaceae], *Saurauia* [Actinidiaceae], *Ternstroemia* [Ternstroemiaceae]) to search for shared biogeographical patterns among these and other amphi-Pacific tropical disjuncts.

Further insight into the historical biogeography of *Symplocos* will require a comprehensive examination of its fossil record, more extensive taxon sampling for molecular phylogenetic data, and divergence time estimates of intercontinentally disjunct clades. Sampling should focus particularly on tropical South American species of subgenus *Symplocastrum*, many of which have been described since Brand's (1901) comprehensive treatment of the genus (see, e.g., Ståhl, 1991, 1995). These species have not been sampled extensively, yet have the most potential to affect the outcome of biogeographic analyses of the Neotropics. Complete four-gene data are also desirable from samples of section *Neosymplocos* to further resolve the

position of this section within the clade otherwise corresponding to section *Symplocastrum*.

LITERATURE CITED

ANDERBERG, A. A., AND X.-P. ZHANG. 2002. Phylogenetic relationships of Cyrillaceae and Clethraceae (Ericales) with special emphasis on the genus *Purdiaea* Planch. *Organisms, Diversity and Evolution* 2: 127–137.

ANDERBERG, A. A., C. RYDIN, AND M. KÄLLERSJÖ. 2002. Phylogenetic relationships in the order Ericales s.l.: analyses of molecular data from five genes from the plastid and mitochondrial genomes. *American Journal of Botany* 89: 677–687.

ANGIOSPERM PHYLOGENY GROUP. 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.

ANGIOSPERM PHYLOGENY GROUP. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

BARTH, O. M. 1979. Pollen morphology of Brazilian *Symplocos* species (Symplocaceae). *Grana* 18: 99–107.

BARTH, O. M. 1982. The sporoderm of Brazilian *Symplocos* pollen types (Symplocaceae). *Grana* 21: 65–69.

BIDÁ, A. 1995. Revisão taxonômica das espécies de *Symplocos* Jacq. (Symplocaceae) do Brasil. Ph.D. dissertation, Universidade de São Paulo, São Paulo, Brazil.

BORHIDI, A. 1996. Phytogeography and vegetation ecology of Cuba, 2nd ed. Akadémiai Kiadó, Budapest, Hungary.

BRAND, A. 1901. Symplocaceae. *In* A. Engler [ed.], Pflanzenreich IV, 242. Engelmann, Leipzig, Germany.

CRONQUIST, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, New York, New York, USA.

DIEFFENBACH, C. W., AND G. S. DVEKSLER [EDS.]. 1995. PCR primer: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, New York, USA.

DONOGHUE, M. J., C. D. BELL, AND J. H. LI. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. *International Journal of Plant Sciences* 162 (6 Supplement): S41–S52.

DOWNIE, S. R., D. S. KATZ-DOWNIE, AND M. F. WATSON. 2000. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA *rpl16* and *rpoC1* intron sequences: towards a suprageneric classification of subfamily Apioideae. *American Journal of Botany* 87: 273–292.

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.

ELIAS, T. S. 1980. The complete trees of North America. Gramercy Publishing, New York, New York, USA.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1994. Testing significance of incongruence. *Cladistics* 10: 315–319.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1995. Constructing a significance test for incongruence. *Systematic Biology* 44: 570–572.

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

FRITSCH, P. W. 2001. Phylogeny and biogeography of the flowering plant genus *Styrax* (Styracaceae) based on chloroplast DNA restriction sites and DNA sequences of the internal transcribed spacer region. *Molecular Phylogenetics and Evolution* 19: 387–408.

FRITSCH, P. W. 2003. Multiple geographic origins of Antillean *Styrax*. *Systematic Botany* 28: 421–430.

FRITSCH, P. W., C. M. MORTON, T. CHEN, AND C. MELDRUM. 2001. Phylogeny and biogeography of the Styracaceae. *International Journal of Plant Sciences* 162 (6 Supplement): S95–S116.

HANDEL-MAZZETTI, H., AND E. PETER-STIBAL. 1943. Eine revision der chinesischen Arten der Gattung *Symplocos* Jacq. *Beihefte zum Botanischen Centralblatt* 62-B: 1–42.

HATUSIMA, S. 1936. Miscellaneous notes on the Symplocaceae of Eastern Asia. *Journal of Japanese Botany* 12: 279–283.

HOLMGREN, P. K., N. H. HOLMGREN, AND L. C. BARNETT. 1990. Index herbariorum, part I: the herbaria of the world, 8th ed. International Association for Plant Taxonomy, Bronx, New York, USA.

HOOT, S. B., S. MAGALLÓN, AND P. R. CRANE. 1999. Phylogeny of basal eudicots based on three molecular data sets: *atpB, rbcL,* and 18S nuclear ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 86: 1–32.

HUELSENBECK, J. P., AND F. R. RONQUIST. 2001. MrBayes: bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.

JORDON, W. C., W. M. COURTNEY, AND E. J. NEIJEL. 1996. Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in North American duckweeds (Lemnaceae). *American Journal of Botany* 83: 430–439.

KELLY, L. M., AND F. ALMEDA. In press. Symplocaceae. *In* G. Davidse, M. Sousa S., and S. Knapp [eds.], Flora Mesoamericana. Universidad Nacional Autónoma de México, México, D. F., México.

KIRCHHEIMER, F. 1949. Die Symplocaceae der erdgeschichtlichen Vergangenheit. *Palaeontographica* 90B: 1–52.

KRUTZSCH, W. 1989. Paleogeography and historical phytogeography (paleochorology) in the Neophyticum. *Plant Systematics and Evolution* 162: 5–61.

LIANG, Y. H. 1986. Pollen morphology of the family Symplocaceae from China. *Acta Botanica Austro Sinica* 2: 111–121.

LITTLE, E. L. 1980. The Audubon Society field guide to North American trees. Eastern region. Alfred A. Knopf, New York, New York, USA.

MADDISON, D. R., AND W. P. MADDISON. 2000. MacClade 4: analysis of phylogeny and character evolution. Version 4.0. Sinauer Associates, Sunderland, Massachusetts, USA.

MADDISON, D. R., M. RUVOLO, AND D. L. SWOFFORD. 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Systematic Biology* 41: 111–124.

MAI, D. 1986. Über die Antillanische Symplocaceae. *Feddes Repertorium* 97: 1–28.

MANCHESTER, S. R. 1999. Biogeographical relationships of North American Tertiary floras. *Annals of the Missouri Botanical Garden* 86: 472–522.

MEIJDEN, R. VAN DER. 1970. A survey of the pollenmorphology of the Indo-Pacific species of *Symplocos* (Symplocaceae). *Pollen et Spores* 12: 513–551.

MIERS, J. 1880. On the Symplocaceae. *Journal of the Linnean Society of London, Botany* 17: 283–306.

MORITIZI, A. 1848. *Cordyloblaste* (Henschel). *Botanische Zeitung (Berlin)* 6: 604–606.

NAGAMASU, H. 1989a. Pollen morphology of Japanese *Symplocos* (Symplocaceae). *Botanical Magazine, Tokyo* 102: 149–164.

NAGAMASU, H. 1989b. Pollen morphology and relationship of *Symplocos tinctoria* (L. f.) L'Her. (Symplocaceae). *Botanical Gazette* 150: 314–318.

NAGAMASU, H. 1993. The Symplocaceae of Japan. *Contributions from the Biological Laboratory, Kyoto University* 28: 173–260.

NAKAI, T. 1922. Trees and shrubs indigenous in Japan proper, vol. 1. Seibido Shoten, Nihombashi, Tokyo, Japan.

NAKAI, T. 1927. Trees and shrubs indigenous in Japan proper, vol. 1, revised ed. Seibido Shoten, Nihombashi, Tokyo, Japan.

NOOTEBOOM, H. P. 1975. Revision of the Symplocaceae of the Old World, New Caledonia excepted. Leiden University Press, Leiden, Netherlands.

NOOTEBOOM, H. P. 1977. Symplocaceae. *In* C. G. G. J. van Steenis [ed.], Flora Malesiana. Series I. Spermatophyta. Flowering plants, vol. 8, 205–274. Sijthoff & Noordhoff, Alphen aan den Rijn, Netherlands.

NOOTEBOOM, H. P. 1980. Symplocaceae. *In* A. Aubréville and J.-F. LeRoy [eds.], Flore de la Nouvelle Calédonia et dépendances, vol. 9. Múseum National d'Histoire Naturelle, Paris, France.

PEATTIE, D. C. 1950. A natural history of trees of eastern and central North America. Houghton Mifflin, Boston, Massachusetts, USA.

RADFORD, A. E., H. E. AHLES, AND C. RITCHIE BELL. 1964. Manual of the vascular flora of the Carolinas. University of North Carolina Press, Chapel Hill, North Carolina, USA.

SANG, T., D. J. CRAWFORD, AND T. F. STUESSY. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120–1136.

SCHÖNEBERGER, J., A. A. ANDERBERG, AND K. J. SYTSMA. 2004. Molecular phylogenetics and patterns of floral evolution in the Ericales. Botany 2004 online abstract No. 271 at http://www.botanyconference.org/engine/search/index.php?func=detail&aid=271.

SOEJIMA, A., AND H. NAGAMASU. 2004. Phylogenetic analysis of Asian *Symplocos* (Symplocaceae) based on nuclear and chloroplast DNA sequences. *Journal of Plant Research* 117: 199–207.

SOLTIS, D. E., P. E. SOLTIS, M. E. MORT, M. W. CHASE, V. SAVOLAINEN, S. B. HOOT, AND C. M. MORTON. 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. *Systematic Biology* 47: 32–42.

STÅHL, B. 1991. Symplocaceae, Oleaceae. *In* G. Harling and L. Andersson [eds.], Flora of Ecuador 43. Berlings, Arlöv, Sweden.

STÅHL, B. 1995. Diversity and distribution of Andean Symplocaceae. *In* S. P. Churchill et al. [eds.], Biodiversity and conservation of Neotropical montane forests, 397–405. New York Botanical Garden, Bronx, New York, New York, USA.

STEENIS, C. G. G. J. VAN. 1962. The land-bridge theory in botany. *Blumea* 11: 235–542.

STEENIS, C. G. G. J. VAN. 1963. Transpacific floristic affinities, particularly in the tropical zone. *In* J. L. Gressitt [ed.], Pacific basin biogeography, 219–231. Bishop Museum Press, Honolulu, Hawaii, USA.

SWENSEN, S. M., J. N. LUTHI, AND L. H. RIESEBERG. 1998. Datiscaceae revisited: monophyly and the sequence of breeding system evolution. *Systematic Botany* 23: 157–169.

SWOFFORD, D. L. 2002. PAUP*, phylogenetic analysis using parsimony (*and other methods), Version 4. Computer program and documentation. Sinauer Associates, Sunderland, Massachusetts, USA.

TABERLET, P., L. GIELLY, G. PAUTOU, AND J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

TAKHTAJAN, A. 1997. Diversity and classification of flowering plants. Columbia University Press, New York, New York, USA.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876–4882.

THORNE, R. F. 1972. Major disjunctions in the geographic ranges of seed plants. *Quarterly Review of Biology* 47: 365–411.

THORNE, R. F. 2000. The classification and geography of the flowering plants: dicotyledons of the class Angiospermae (subclasses Magnoliidae, Ranunculidae, Caryophyllidae, Dilleniidae, Rosidae, Asteridae, and Lamiidae). *Botanical Review* 66: 441–647.

TIFFNEY, B. H. 1985a. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. *Journal of the Arnold Arboretum* 66: 73–94.

TIFFNEY, B. H. 1985b. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* 66: 243–273.

WEN, J. 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology and Systematics* 30: 421–455.

WEN, J., P. P. LOWRY II, J. L. WALCK, AND K.-O. YOO. 2002. Phylogenetic and biogeographic diversification in *Osmorhiza* (Apiaceae). *Annals of the Missouri Botanical Garden* 89: 414–428.

WU, R.-F. 1986a. A preliminary study on *Symplocos* of China. *Acta Phytotaxonomica Sinica* 24: 193–202.

WU, R.-F. 1986b. A preliminary study on *Symplocos* of China (cont.). *Acta Phytotaxonomica Sinica* 24: 275–291.

WU, R.-F. 1987. Symplocaceae. *In* Flora reipublicae popularis sinicae, vol. 60, 1–77. Science Press, Beijing, China.

WU, R.-F., AND H. P. NOOTEBOOM. 1996. Symplocaceae. *In* Z.-Y. Wu and P. H. Raven [eds.], Flora of China: Myrsinaceae through Loganiaceae, vol. 15, 235–252. Science Press, Beijing, China.

XIANG, Q.-Y. (J.), AND D. E. SOLTIS. 2001. Dispersal-vicariance analyses of intercontinental disjuncts: historical biogeographical implications for angiosperms in the Northern Hemisphere. *International Journal of Plant Sciences* 162 (6 Supplement): S29–S39.