American Journal of Botany 90(2): 310–320. 2003.

PHYLOGENETIC RELATIONSHIPS IN *GLEDITSIA* (LEGUMINOSAE) BASED ON ITS SEQUENCES[1]

ANDREW SCHNABEL,[2,4] PATRICK E. MCDONEL,[2,5] AND JONATHAN F. WENDEL[3]

[2]Department of Biological Sciences, Indiana University South Bend, South Bend, Indiana 46634 USA; and
[3]Department of Botany, Iowa State University, Ames, Iowa 50011 USA

We used nucleotide sequences from the internal transcribed spacers and 5.8S gene of nuclear ribosomal DNA to test competing phylogenetic and biogeographic hypotheses in *Gleditsia*. Eleven of 13 *Gleditsia* species were sampled, along with two species of its sister genus, *Gymnocladus*. Analyses of ITS data and of a combined data set that included sequences of ITS and two chloroplast genes supported several conclusions that were interpreted in light of fossil data and current legume phylogeny. *Gleditsia* and *Gymnocladus* appear to have originated in eastern Asia during the Eocene. Eastern North American species of both genera most likely evolved from ancestors that migrated across the Bering land bridge, but the eastern Asian/eastern North American disjunction appears to be much older in *Gymnocladus* than in *Gleditsia*. *Gleditsia amorphoides*, from temperate South America, is sister to the rest of the genus, suggesting early long-distance dispersal from Asia. The remainder of *Gleditsia* is divided into three unresolved clades, possibly indicating a split early in the evolution of the genus. Two of those clades contain only Asian species, and one contains Asian and North American species. The North American species, *Gleditsia triacanthos* and *Gleditsia aquatica*, are polymorphic and paraphyletic with respect to their ITS and cpDNA sequences, which suggests recent diversification.

Key words: biogeography; *Gleditsia*; *Gymnocladus*; internal transcribed spacer; Leguminosae; molecular phylogeny; ribosomal DNA.

The well-known floristic similarities between eastern Asia and eastern North America have a long and complicated history (e.g., Chaney, 1947; Graham, 1972; Boufford and Spongberg, 1983; Manchester, 1999; Tiffney, 2000). It is now recognized that floristic interchanges between North America and Asia were dynamic during the Cenozoic, involving at least five distinct periods and migrations across both the Bering and North Atlantic land bridges (Tiffney, 1985a, b, 2000; Manchester, 1999). Consequently, the results of the recent flurry of phylogenetic studies involving taxa with eastern Asian-eastern North American disjunctions have been no less diverse with respect to biogeographic interpretations (e.g., Kelly, 1998; Xiang et al., 1998; Kim and Kim, 1999; Stanford et al., 2000; Soltis et al., 2001; Whitcher and Wen, 2001). As discussed in a review by Wen (1999), phylogenetic analyses have revealed that many of the hypothesized intercontinental, sister-species pairs are not closest relatives and that continued diversification has often occurred in Asia and North America following one or more migration events between the continents (e.g., Qiu et al., 1995; Wen and Shi, 1999; Stanford et al., 2000). Biogeographic explanations for many taxa with significant representation in eastern Asia and eastern North America are further complicated by the presence of disjunct members in other areas, such as western North America, South America, or Europe (e.g., Wen et al., 1998; Cameron and Chase, 1999; Soltis et al., 2001; Whitcher and Wen, 2001).

The genus *Gleditsia* represents one of the best examples of the complexities involved in trying to unravel the biogeographic history of eastern Asian-eastern North American disjunctions. Like most woody genera with Asian-North American connections (Guo and Ricklefs, 2000), the majority of the species diversity in *Gleditsia* is found in eastern Asia, but *Gleditsia* species are also found in eastern North America, east-central South America, and the southern Caucasus. The only thorough review of the genus (Gordon, 1966) recognized 13 species: eight in eastern Asia (*G. australis*, *G. delavayi*, *G. fera*, *G. japonica*, *G. macracantha*, *G. microphylla*, *G. rolfei*, and *G. sinensis*), two in eastern North America (*G. aquatica*, *G. triacanthos*), one in South America (*G. amorphoides*), one in a small area near the southern coast of the Caspian Sea (*G. caspica*), and a final, poorly documented species from northeastern India (*G. assamica*).

These species designations, however, are not universally accepted, especially with respect to the Asian taxa (Larsen et al., 1980; Li, 1982, 1988; Paclt, 1982b). We have used Gordon (1966) as a starting point for our investigations, because more than any other author, he suggested clear hypotheses about phylogenetic and biogeographic relationships. First, Gordon hypothesized a close relationship between *G. triacanthos*, which is widespread throughout eastern North America, and *G. japonica*, *G. delavayi*, and *G. caspica*, a group of morphologically similar species. This hypothesis was later made more specific by Isley (1975, p. 157), who suggested that *G. japonica* was ''the closest oriental equivalent of *G. triacanthos*.'' Second, Gordon hypothesized that *G. aquatica*, a more narrowly distributed species from the eastern United States, was sister to *G. microphylla*, which grows in north and east China. Third, Gordon suggested that the distinctive South American species, *G. amorphoides*, is most closely related either to *G. rolfei* (southeast Asia) or to *G. microphylla*.

[1] Manuscript received 14 May 2002; revision accepted 26 September 2002.

The authors thank K. W. Cheung, M.-G. Chung, E. Hedborn, N. T. Hiep, T. C. Khánh, G.-C. Ling, R. McDonald, B. Middleton, K.-W. Park, J. Quigley, F. Santamour, R. Small, R. Vanni, and Jardín Botánico y Arboretum C. Spegazzini (La Plata, Argentina) for sending plant material; A. Bruneau for helpful comments on the manuscript; C. B. Heiser for helping to collect material from the IU Experimental Garden; and P. S. Herendeen for information on *Gleditsia* and *Gymnocladus* fossils and for helpful comments on the manuscript. This work was supported by a grant from the National Science Foundation (DEB-9203214).

[4] Author for reprint requests (e-mail: aschnabe@iusb.edu).

[5] Current address: Department of Molecular and Cell Biology, University of California, Berkeley, California 94720 USA.

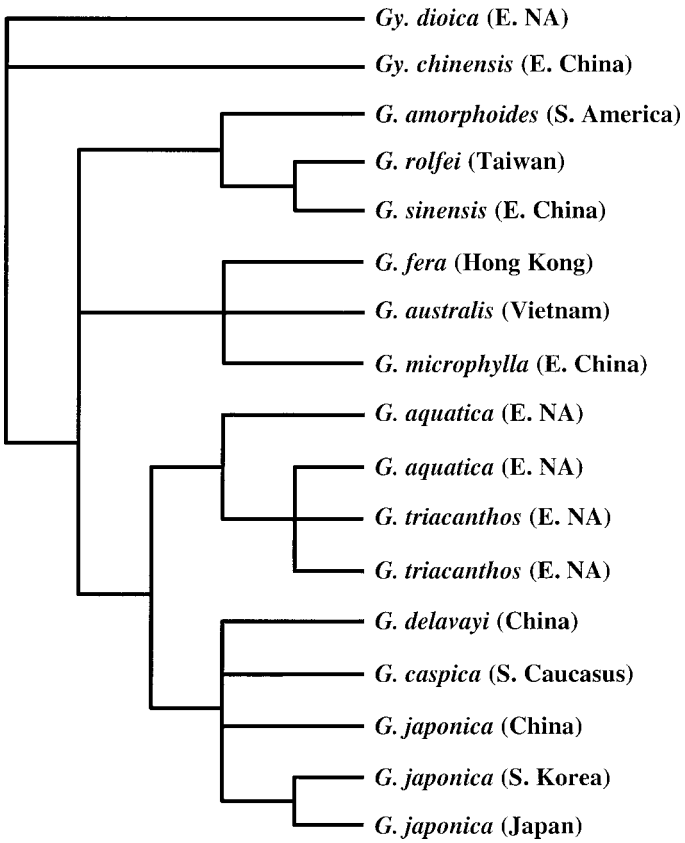

Fig. 1. Hypothesized phylogenetic relationships among 11 *Gleditsia* species, using two *Gymnocladus* species as outgroups, based on maximum parsimony analysis of *ndhF* and *rpl16* cpDNA sequences from 17 accessions (Schnabel and Wendel, 1998).

We previously tested these hypotheses using maximum parsimony analyses of sequences from two chloroplast genes, *ndhF* and *rpl16* (Schnabel and Wendel, 1998). Interpreting our results (Fig. 1) in light of fossil and geologic evidence, we hypothesized that the genus most likely arose in Asia; that the eastern North American lineage probably evolved from a *G. japonica*-like ancestor that migrated across the Bering land bridge during the Miocene (5–23 million years ago [mya]); and that the ancestor of *G. amorphoides* most likely reached South America by dispersal across the Pacific Ocean. In the study described here, we use sequences of the internal transcribed spacers (ITS) of nuclear ribosomal genes to help resolve the differences between earlier phylogenetic and biogeographic hypotheses in *Gleditsia* and those suggested by our cpDNA data. We find that the ITS data largely are congruent with the cpDNA data, thus providing support for nearly all of the conclusions reached in our earlier study.

MATERIALS AND METHODS

We sampled the same 11 species of *Gleditsia* used in our previous study (Schnabel and Wendel, 1998), and most of the same accessions (accession, source, and voucher information has been archived at the Botanical Society of America website at http://ajbsupp.botany.org/v90/). No material was available for *G. assamica*, and two *G. macracantha* accessions we obtained from botanical gardens proved to be misidentified *G. triacanthos* accessions. No further accessions of *G. macracantha* were available. To better test hypotheses about the evolution of North American *Gleditsia*, we added several accessions of *G. triacanthos* and *G. aquatica*, most of which came from either the Morton Arboretum (Lisle, Illinois, USA) or the Indiana University Experimental Garden (Bloomington, Indiana, USA). All of these trees were grown from seed collected from natural populations. We also added a second accession of *G. amorphoides* from Argentina. The outgroup for this analysis was *Gymnocladus*, which forms a monophyletic ''*Gleditsia* group'' in recent phylogenetic analyses of the Leguminosae (Bruneau et al., 2001; Kajita et al., 2001; Herendeen et al., in press). *Gymnocladus* also has an eastern Asian-eastern North American distribution, with 3–4 species in eastern and southern Asia and one species in eastern North America (Lee, 1976; Larsen et al., 1980). We sampled one accession of *Gymnocladus chinensis* (China) and two accessions of *Gymnocladus dioica* (North America). In total, analyses to test phylogenetic hypotheses about *Gleditsia* as a whole were conducted using three *Gymnocladus* accessions and 23 *Gleditsia* accessions that included one *G. amorphoides* accession, 15 Asian/Caspian accessions, and seven North American accessions.

We extracted total genomic DNA from fresh or dried leaves using the protocol described in Paterson et al. (1993) or Qiagen DNeasy Plant Mini Kits (Qiagen, Valencia, California, USA). Amplification of the internal transcribed spacers 1 and 2 (ITS1 and ITS2) and 5.8S gene of nuclear ribosomal DNA was performed using the ITS5 (forward) and ITS4 (reverse) primers of White et al. (1990) in 50-μL reactions containing 10% dimethylsulfoxide, 0.2 μmol/L of each primer, 0.2 mmol/L of each dNTP, 1.5 mmol/L $MgCl_2$, 2.5 units *Taq* polymerase (Promega, Madison, Wisconsin, USA), and 5 μL Promega Mg-free reaction buffer. Cycling conditions began with 0.5 min at 95°C, continued with 35 cycles of 0.3 min at 95°C, 0.5 min at 55°C, and 1.5 min at 72°C, and finished with 3 min at 72°C.

For most accessions, amplified DNA was cloned following the protocols of the TA cloning kit (Invitrogen, San Diego, California, USA). White colonies were screened by picking cells with sterile toothpicks and dipping the cells directly into a 20-μL polymerase chain reaction (PCR) mix (same ingredient concentrations as above, except for the addition of 0.2 μL of 10× bovine serum albumin). Screening primers were complementary to the regions of the pCR2.1 vector flanking the ITS insert (forward: 5′-GCCGCCAGTGTGCTGGAATT-3′; reverse: 5′-TAGATGCATGCTCGAGCGGC-3′; Qiu et al., 2001). Cycling conditions included 2 min at 96°C, followed by 35 cycles of 0.5 min at 94°C, 1 min at 65°C, and 3.5 min at 72°C, with a final 10 min at 72°C. Colonies showing presence of the insert were screened a second time (50-μL PCR reaction) to generate templates for sequencing. Amplified template was purified using Wizard PCR Preps (Promega) and then sequenced using the ABI Prism BigDye cycle sequencing kit (PE Applied Biosystems, Foster City, California, USA) in a 10-μL reaction that included 1 μL template (250–500 ng), 4 μL reaction mix, and 2 μmol/L primer. We used the ITS2, ITS3, and ITS4 primers from White et al. (1990) for these reactions. Sequenced templates were ethanol precipitated, dried, and sent to the Indiana University (Bloomington, Indiana, USA) DNA facility for electrophoresis. We attempted to sequence 2–3 clones of each accession. This sequencing protocol was followed for all but four of the accessions (*G. amorphoides*-2, *G. japonica*-2, *G. japonica*-4, *G. australis*-1; http://ajbsupp.botany.org/v90/), for which ITS sequences were generated directly from double-stranded PCR product using the Thermo Sequenase cycle sequence kit (Amersham Pharmacia, Piscataway, New Jersey, USA), electrophoresis in Long Ranger (AT Biochem, Malvern, Pennsylvania, USA) gels, followed by autoradiography. Sequences have have been deposited in GenBank (AF509969–AF510034).

We generated an initial alignment of the ITS sequences using CLUSTALW (Thompson et al., 1994; Higgins et al., 1996) and then made several minor adjustments manually. To determine boundaries of ITS1 and ITS2, we compared our sequences with those published on the papilionoid tribe Millettieae (Hu et al., in press). The alignment of all sequences conservatively suggested the presence of 23 gaps, 21 of which were 1–2 base pairs (bp) in length. Of those gaps, 15 were unique to one accession or to one clone within an accession, and five were unique to two or more *Gymnocladus* accessions. Because only three of the hypothesized gaps were potentially phylogenetically informative, all phylogenetic analyses treated gaps as missing data.

Phylogenetic analyses were accomplished using PAUP* 4.0b10 (Swofford,

2002). Pairwise evolutionary distances between accessions were generated under the Kimura 2-parameter model, also using PAUP*. To test hypotheses about the phylogeny of *Gleditsia* as a whole, we used 50 of the ITS sequences, which included 16 of the 32 North American *Gleditsia* sequences and all the complete or nearly complete sequences from the other species. A separate analysis of North American *Gleditsia* was conducted on the full set of 32 clonal sequences from six *G. triacanthos* and nine *G. aquatica* accessions, with two *G. japonica* accessions being used for outgroup comparison.

Both maximum parsimony (MP) and maximum likelihood (ML) optimality criteria were used to evaluate possible tree topologies. The MP analyses used a heuristic search with the accelerated transformation option to optimize the unordered characters and tree bisection-reconnection branch swapping. Starting trees were obtained by stepwise addition with a random addition sequence (10 replicates). Branch support was evaluated through bootstrap analysis (Felsenstein, 1985; Swofford et al., 1996). For the ML analyses, we modeled among-site rate variation using a gamma distribution and used the shortest trees from the MP analyses as a starting point for ML estimation of transition/transversion (ti/tv) ratios and the alpha parameter of the gamma distribution for among-site rate variation. We then followed an iterative procedure described in Swofford et al. (1996), in which the most likely tree from each heuristic search was used to reestimate the ti/tv ratio and alpha parameter. This was repeated until essentially no change occurred in the likelihood estimate between iterations. In these analyses, nucleotide frequencies were assumed to be equal to empirical frequencies, and a molecular clock was not enforced.

Because the results of phylogenetic analyses using ITS sequences were so similar to those found using chloroplast DNA (cpDNA) sequences (Schnabel and Wendel, 1998), differing significantly only in the placement of *G. amorphoides*, we performed another set of analyses on combined ITS and cpDNA data sets for 20 *Gleditsia* and two *Gymnocladus* accessions. For these analyses, composite ITS sequences were generated by merging the 2–3 clonal sequences within an accession and treating multistate nucleotide positions as uncertainties. We followed the same procedures described above for the ITS analyses, except that the MP analysis used a branch-and-bound algorithm, and branch support was evaluated through calculation of both decay indices (Bremer, 1988; Donoghue et al., 1992) and bootstrap values.

To predict the center of origin for *Gleditsia* and *Gymnocladus*, we employed the method of character mapping described by Xiang et al. (1998). Taking the consensus MP tree from the combined ITS/cpDNA analysis as the best hypothesis of *Gleditsia*/*Gymnocladus* phylogeny, we assigned unordered character-state codes (1–5) to the geographical distribution areas of each species (eastern/northeastern Asia, southern Asia, southern Caucasus, eastern North America, South America) and used the Trace Character function in MacClade 4 (Maddison and Maddison, 2000) to map geographic locations on the tree. For this analysis, we assumed that *Gymnocladus* is monophyletic.

RESULTS

Internal transcribed spacer sequence diversity—For the 11 *Gleditsia* and two *Gymnocladus* species sampled, we generated 66 complete or partial ITS sequences from 35 accessions. Each full sequence included 32 nucleotides of the 26S rRNA gene, 38 nucleotides of the 18S rRNA gene, and complete sequences of ITS1, the 5.8S rRNA gene, and ITS2. Partial sequences were obtained for the four accessions that were sequenced directly. The *G. amorphoides*-2 accession was missing over half of the ITS1 and ITS2 sequence, but the 202 nucleotides available were identical to one of the clones from the *G. amorphoides*-1 accession. This partial sequence was not included in subsequent phylogenetic analyses. The *G. japonica*-4 and *G. australis*-1 accessions were missing the 26S fragment and the first 80 bp of ITS1, and those same accessions, as well as the *G. japonica*-2 accession, were missing approximately 50 bp of the 5.8S gene. All complete 5.8S sequences were 160 nucleotides in length, except for the two sequences from *Gymnocladus chinensis*, which were 162 nucleotides long. The ITS1 sequences varied in length from 237 to 242 bp, whereas the ITS2 sequences varied from 214 to 229 bp. Alignment of the ITS1 sequences required the hypothesis of 11 insertion/deletion events of 1–2 bp each and resulted in an aligned length of 246 bp. Alignment of the ITS2 sequences required 10 indels of 1–2 bp each and two longer gaps of 8 bp and 14 bp for a total aligned length of 229 bp. The guanine-cytosine (GC) content averaged 69.8% in ITS1, 74.7% in ITS2, and 54.5% in the 5.8S gene.

TABLE 1. Summary of molecular information for the 50 ITS sequences (45 *Gleditsia* and five *Gymnocladus*) used in the main phylogenetic analysis. Numbers in parentheses are estimates for the 45 *Gleditsia* sequences alone. Totals include variable sites from 32 base pairs (bp) of the 26S gene and 38 bp of the 18S gene flanking the ITS1/5.8S/ITS2 region.

rDNA region	Aligned length (bp)	No. of variable sites	No. of informative sites
ITS1	246	83 (52)	68 (36)
5.8S	162	20 (18)	7 (4)
ITS2	229	90 (54)	77 (41)
Total	707	202 (131)	154 (82)

Because studies have shown that some plant genomes harbor multiple, and in some cases, highly divergent ITS sequences (e.g., Buckler and Holtsford, 1996), we sampled more than one clone for most accessions studied. When we considered only substitution polymorphisms, no two of the cloned accessions were identical to each other, and overall, we found 58 unique sequences among the 62 cloned sequences. For the 50 sequences included in the main phylogenetic analysis, 202 of 707 characters (28.6%) exhibited substitution polymorphisms, and 154 (21.8%) of those were parsimony-informative. For the 45 *Gleditsia* sequences alone, 131 sites were polymorphic, 82 of which were parsimony-informative. The informative sites were approximately equally distributed between ITS1 and ITS2, with much less information present in the 5.8S gene (Table 1).

We used six *G. triacanthos* (13 sequences) and nine *G. aquatica* (19 sequences) accessions to investigate intraspecific ITS variation in North American *Gleditsia*. In *G. triacanthos*, 12 of 13 sequences were unique, sequence divergence (Kimura 2-parameter distances) ranged from 0.000 to 0.020 (mean $\pm$ 1 SD = 0.009 $\pm$ 0.005), and 24 of 26 polymorphic sites were autapomorphies. In *G. aquatica*, 18 of 19 sequences were unique, sequence divergence ranged from 0.000 to 0.024 (mean $\pm$ 1 SD = 0.012 $\pm$ 0.005), and 19 of 28 polymorphic sites were autapomorphies.

Mean sequence divergence (Kimura 2-parameter distances) between *Gymnocladus dioica* and *Gy. chinensis* (0.163) was considerably greater than average divergence among *Gleditsia* species (range 0.004–0.100) (Table 2). *Gymnocladus* showed 18–24% sequence divergence from *Gleditsia*. Within *Gleditsia*, *G. amorphoides* was the most divergent from all other species (range 0.070–0.102). Four groups of species, *G. japonica*/*G. delavayi*/*G. caspica* (range 0.004–0.007), *G. sinensis*/*G. rolfei* (0.013), *G. triacanthos*/*G. aquatica* (0.014), and *G. fera*/*G. australis*/*G. microphylla* (range 0.016–0.033) showed low within-group sequence divergence. Among the remaining Asian *Gleditsia*, mean sequence divergence ranged from 0.040 to 0.073, whereas mean sequence divergence between Asian and North American species was 0.063–0.087.

In summary, sequence variation was observed at all levels in the ITS data set, including among clonal sequences within all but one of the accessions (*Gymnocladus chinensis*). The amount of intra-individual variation, however, was low and not indicative of the several potential problems that may cause phylogenetically misleading signal when using ITS sequences, such as presence of divergent paralogs or pseudogenes (Buckler et al., 1997; Wendel, 2000). Confidence in the phylogenetic utility of the ITS data set was further increased by congruence with the independently generated cpDNA data (see below).

Phylogenetic analyses—The MP analysis using 50 sequences identified 1435 most parsimonious trees. One of those trees, with the same topology as the single best tree from the ML analysis, was selected to illustrate overall topology and branch support (Fig. 2). The strict consensus of these 1435 trees suggested several insights into the history of the genus, many of which were strongly supported by high bootstrap percentages (Fig. 2). First, the South American *G. amorphoides* sequences formed a clade that is sister to the rest of the genus. Second, the remaining sequences were divided into two large clades, one of which was relatively weakly supported and wholly Asian, and one of which was more strongly supported and contained both Asian and North American species. Both major clades bifurcated into two strongly supported clades. In the wholly Asian clade, *G. sinensis* and *G. rolfei* formed a clade (hereafter called the *G. sinensis* clade) that was sister to a clade containing *G. australis*, *G. fera*, and *G. microphylla* (hereafter called the *G. australis* clade). In the second major clade, the North American species, *G. triacanthos* and *G. aquatica*, formed a monophyletic group (hereafter called the *G. triacanthos* clade) that was sister to a clade containing the Asian species, *G. japonica*, *G. delavayi*, and *G. caspica* (hereafter called the *G. japonica* clade). In about 50% of the 1435 trees, the *G. sinensis* clade fell sister to the *G. japonica*/*G. triacanthos* clade instead of being sister to the *G. australis* clade. Each of these two arrangements was supported by a single synapomorphy. Third, within the *G. australis* clade, *G. fera* and *G. australis* always formed a pair of sister species. Fourth, the *G. japonica* sequences did not form a monophyletic group. Within this clade, the *G. japonica* sequences from Korea formed a clade with the single *G. delavayi* sequence, while the *G. japonica* sequences from China and Japan formed a clade with the *G. caspica* sequences. Fifth, the North American sequences of *G. triacanthos* and *G. aquatica* did not form separate monophyletic groups. Most of the topological variability within the 1435 shortest trees was due to the numerous equally parsimonious arrangements of these 16 sequences. The strict consensus suggested, albeit with very little support, that *G. triacanthos* is derived from a *G. aquatica*-like ancestor, but the separate analysis of all 32 North American sequences did not support this hypothesis. The 32 sequences contained 63 variable characters, only 10 of which were parsimony informative. Consequently, the MP analysis uncovered 500 shortest trees, each with length 127, a consistency index (excluding uniformative characters) of 0.746, and a retention index of 0.911. The strict consensus of these trees was a star phylogeny (not shown), in which all phylogenetic structure was lost, except for a small clade of four *G. aquatica* sequences.

We combined ITS and cpDNA sequences for each of 22 accessions to produce a data set with an aligned length of 3998 bp. A total of 153 of these characters were variable within *Gleditsia*, and 103 of those were parsimony informative. An MP analysis produced 16 shortest trees, the topologies of which differed only in the relationships among the North American accessions (eight arrangements of the subclade containing the *G. triacanthos* accessions) and among the accessions of the *G. japonica* clade (two possible arrangements). All of the results described above for the ITS analysis were supported, except that here all trees showed the *G. australis* clade being sister to the clade containing the *G. triacanthos* and *G. japonica* clades (Fig. 3A). This arrangement was weakly supported, however, with a bootstrap value of 54% and decay index of 1. The sister relationship of *G. amorphoides* to the rest of the genus was again supported (86% bootstrap; decay index = 4), but was contradicted by the results of the ML analyses, which suggested instead that *G. amorphoides*, *G. sinensis*, and *G. rolfei* are a monophyletic group that is sister to all other *Gleditsia* accessions (Fig. 3B). This arrangement was four steps longer than the most parsimonious trees. Mapping the geographical distribution areas of each species onto Fig. 3A suggested that *Gymnocladus* and *Gleditsia* have a common origin in eastern Asia (Fig. 4).

DISCUSSION

Broad congruence between molecular data sets—The phylogenies for *Gleditsia* generated from separate and combined analyses of ITS and cpDNA sequences show several well-supported areas of agreement (Figs. 1–3). Four clades appear in all analyses: the *G. japonica* clade, the *G. triacanthos* clade, the *G. australis* clade, and the *G. sinensis* clade. The *G. japonica* clade and the *G. triacanthos* clade always form a larger monophyletic group, and within the *G. australis* clade, *G. australis* and *G. fera* are always sister species. Despite this constancy, two major areas of ambiguity remain. The MP analysis of ITS data suggests that the closest lineage to the *G. triacanthos*/*G. japonica* clade is equally likely to be either the *G. australis* or *G. sinensis* clade, whereas the ML analysis favors a third topology, in which these two latter clades are joined to form a large Asian clade. In contrast, both MP and ML analyses of combined ITS/cpDNA data place the *G. australis* clade sister to the *G. triacanthos*/*G. japonica* clade. The most conservative resolution of this phylogenetic uncertainty is to place the three clades in an unresolved trichotomy, which is the same result reached with the cpDNA data alone, and we interpret this as an indication of rapid divergence of *Gleditsia* into three main lineages early in its evolution. Also in doubt is the placement of *G. amorphoides*. Our original cpDNA study suggested that *G. amorphoides* was sister to the *G. sinensis* clade, but all MP analyses of ITS or combined data suggest that the *G. amorphoides* lineage is sister to the rest of the genus. Only the ML analysis of the combined data set returns the original result obtained with cpDNA data alone. The broad areas of congruence between data sets suggest resolutions to several competing phylogenetic and biogeographic hypotheses (Gordon, 1966; Isely, 1975; Larsen et al., 1980; Paclt, 1982a, b, 1984; Schnabel and Wendel, 1998), whereas the incongruities and ambiguities suggest alternative hypotheses about the early evolution of the genus.

Phylogenetic questions within Asian species—Most of the diversity in *Gleditsia* currently is found in Asia, where nine of the 13 species recognized by Gordon (1966) occur. Because many of these species are distributed over large areas and show great morphological variability (http://ajbsupp.botany.

TABLE 2. Kimura 2-parameter distances (means ± 1 SD) within and between two *Gymnocladus* and 11 *Gleditsia* species. Numbers in parentheses following species names are the total numbers of sequences from all accessions sampled of that species.

Taxon name	*Gy. dioica*	*Gy. chinensis*	*G. amorphoides*	*G. rolfei*	*G. sinensis*	*G. australis*
Gy. dioica	0.010 ± 0.005					
Gy. chinensis	0.163 ± 0.007	0.000				
G. amorphoides	0.210 ± 0.005	0.242 ± 0.002	0.006			
G. rolfei	0.206 ± 0.004	0.240 ± 0.004	0.082 ± 0.003	0.009		
G. sinensis (2)	0.191 ± 0.004	0.215 ± 0.000	0.070 ± 0.001	0.013 ± 0.000	0.000	
G. australis (4)	0.188 ± 0.006	0.213 ± 0.002	0.102 ± 0.002	0.061 ± 0.001	0.053 ± 0.001	0.004 ± 0.004
G. fera (2)	0.190 ± 0.006	0.215 ± 0.002	0.096 ± 0.003	0.053 ± 0.003	0.045 ± 0.003	0.016 ± 0.003
G. microphylla (3)	0.183 ± 0.006	0.226 ± 0.005	0.086 ± 0.003	0.048 ± 0.002	0.040 ± 0.002	0.033 ± 0.003
G. triacanthos (13)	0.177 ± 0.005	0.224 ± 0.005	0.100 ± 0.005	0.066 ± 0.004	0.063 ± 0.004	0.083 ± 0.004
G. aquatica (19)	0.204 ± 0.007	0.236 ± 0.005	0.100 ± 0.005	0.071 ± 0.004	0.067 ± 0.004	0.087 ± 0.004
G. delavayi (1)	0.199 ± 0.005	0.234 ± 0.000	0.087 ± 0.002	0.049 ± 0.000	0.041 ± 0.000	0.072 ± 0.001
G. caspica (5)	0.200 ± 0.004	0.238 ± 0.002	0.094 ± 0.002	0.051 ± 0.002	0.043 ± 0.002	0.072 ± 0.002
G. japonica (7)	0.187 ± 0.004	0.234 ± 0.002	0.092 ± 0.002	0.051 ± 0.002	0.043 ± 0.002	0.073 ± 0.003

[a] nc, no comparison possible.

org/v90/; Gordon, 1966; Isely, 1975; Paclt, 1982b, 1984; Tucker, 1991), species definitions and relationships within Asian *Gleditsia* have been unclear. The best example of this is *G. japonica*, which is widespread in Korea, Japan, and eastern China. Gordon (1966) recognized *G. japonica* var. *japonica* as the taxon distributed over most of the range, but also recognized *G. japonica* var. *stenocarpa* as a Korean endemic. Gordon (1966) also expressed some doubts about the status of *G. delavayi* and *G. caspica* as separate species but noted several morphological characters that set them apart. *Gleditsia delavayi*, which grows in southwestern China, is the only species in the genus with a woody inflorescence axis, and it has the largest number of ovules per ovary (35–42 ovules vs. 23–32 ovules for *G. japonica*) and, consequently, the longest fruits (27–70 cm vs. 23–42 cm for *G. japonica*) in the genus. Li (1982, 1988), however, recognized *G. delavayi* simply as one subspecies of *G. japonica*, along with *G. japonica* var. *japonica* and a newly defined *G. japonica* var. *velutina*. Paclt (1982a, 1984) went one step further and suggested that *G. caspica* also is not a distinct species from *G. japonica*. Neither of these authors recognized Gordon's *G. japonica* var. *stenocarpa* as a valid taxon. Moreover, all *Gleditsia* in Korea is often recognized as *G. japonica* var. *koraiensis* (e.g., Huh et al., 1999).

Our results are most supportive of Paclt's (1982a, 1984) interpretation of variability in the *G. japonica* clade. In contrast to the cpDNA data sets, which showed three characters unique to *G. delavayi* and five characters unique to *G. caspica*, no unique ITS characters define *G. delavayi* and 0–2 characters define each of the four *G. caspica* sequences (Fig. 2). This result, in itself, is somewhat surprising, given that ITS in *Gleditsia* appears to be evolving considerably faster than either *ndhF* or *rpl16*. For example, approximately 11% of the nucleotide sites were polymorphic within the *Gleditsia* ITS sequences, whereas only 2.1% of the *ndhF* sites and 2.9% of the *rpl16* sites were polymorphic (Schnabel and Wendel, 1998). All analyses of ITS data alone and the ML analysis for the combined data, however, show that *G. delavayi* and *G. caspica* fall into separate (but weakly supported) clades and that each is sister to different *G. japonica* accessions. Thus, although our molecular data can separate *G. delavayi* from *G. caspica*, we cannot separate either of those from *G. japonica*. In addition, our two molecular data sets show different relationships among the *G. japonica* accessions. The cpDNA data suggest that the South Korean and Japanese accessions cluster separately from the Chinese accessions. In contrast, all the analyses of ITS data alone and the ML analysis of combined data place the South Korean accessions in one clade and the Japanese and Chinese accessions in a second clade. More extensive sampling of *G. japonica*, especially in China, will be necessary to determine whether any of the subspecific designations of *G. japonica* are phylogenetically justified and to clarify the taxonomic complexity of this group.

Additional conflict between our results and previous hypotheses involves the relationship between *G. rolfei* and *G. fera*. Both species are found in southeastern China, including Hong Kong, and Vietnam (Gordon, 1966; Larsen et al., 1980), but the range of *G. rolfei* extends also into Taiwan, the Philippines, and purportedly, Sulawesi (Gordon, 1966). Gordon (1966) hypothesized a close relationship between *G. rolfei* and *G. fera*, and Larsen et al. (1980), Li (1988), and Larsen (1989) suggested that *G. rolfei* is conspecific with *G. fera*. Our results indicate that these two species have no close relationship, implying that the morphological similarities between them result from convergence. We note, however, that our *G. rolfei* accession came from Taiwan, and it is not inconceivable that *G. rolfei* in southeast Asia is a different lineage more closely related to *G. fera*.

Other vexing phylogenetic and taxonomic questions still remain within the Asian taxa and can only be answered with more complete sampling and analysis of multiple data sets. First, although early publications on *Gleditsia* in China recognize *G. sinensis* and *G. macracantha* as separate species (Woon, 1921; Steward, 1958), subsequent authors have questioned this distinction (Isley, 1975; Paclt, 1982b), and more recent Chinese floras have united the two under *G. sinensis* (Li, 1988). Gordon (1966) maintained them as separate species, but noted a long list of specimens that he was not able to assign definitively to either species. In our original study, we sequenced two accessions purported to be *G. macracantha* only to find that the sequences were identical to our *G. triacanthos* sequences. Neither of these accessions came directly from China and therefore probably were mislabeled in their respective arboreta. A recent check of one of those growing in the Royal Botanic Garden at Kew, UK, has confirmed this hypothesis (A. Schnabel, personal observations). Second, the relationship of *G. assamica*, from northeastern India, to the rest of the genus is unresolved, because of the difficulty of sampling this poorly documented species. Gordon (1966) was able to view an illustration of a single specimen for his mono-

TABLE 2. Extended.

G. fera	*G. microphylla*	*G. triacanthos*	*G. aquatica*	*G. delavayi*	*G. caspica*	*G. japonica*
0.004						
0.025 ± 0.003	0.004 ± 0.003					
0.077 ± 0.005	0.077 ± 0.005	0.009 ± 0.005				
0.082 ± 0.005	0.082 ± 0.005	0.014 ± 0.005	0.012 ± 0.005			
0.065 ± 0.003	0.060 ± 0.003	0.065 ± 0.003	0.067 ± 0.004	nc[a]		
0.065 ± 0.003	0.060 ± 0.002	0.067 ± 0.004	0.069 ± 0.004	0.004 ± 0.002	0.004 ± 0.002	
0.067 ± 0.004	0.061 ± 0.003	0.067 ± 0.004	0.070 ± 0.004	0.007 ± 0.004	0.007 ± 0.004	0.007 ± 0.002

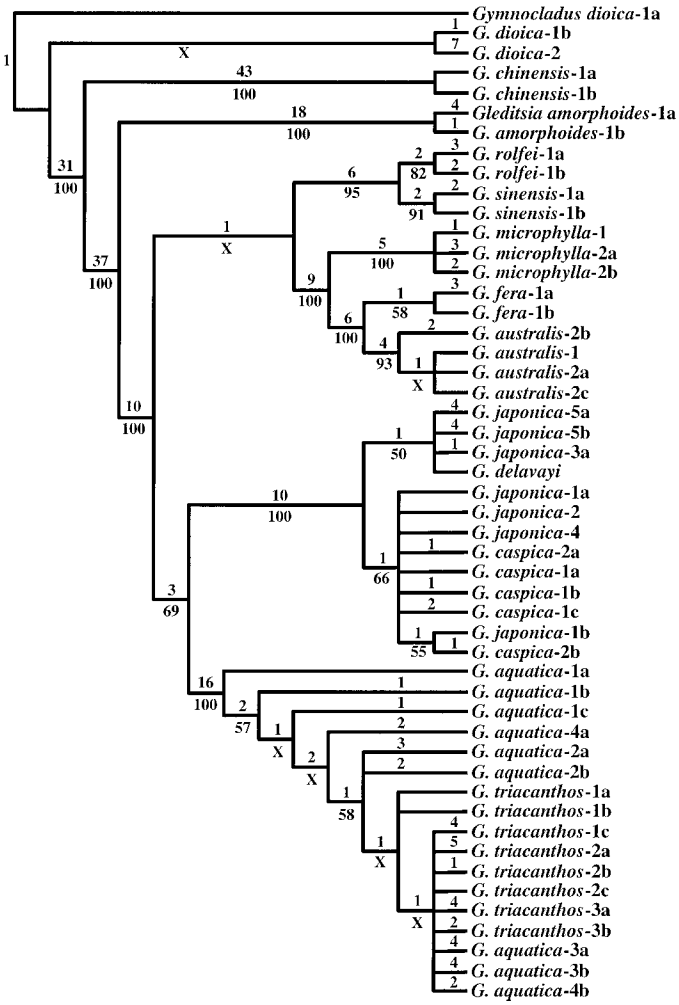

Fig. 2. One of 1435 most parsimonious trees based on analysis of ITS data (707 base pairs [bp]) from five *Gymnocladus* and 45 *Gleditsia* sequences (297 steps, consistency index with autapomorphies excluded = 0.787, retention index = 0.946). Maximum likelihood analysis produced this same topology. Numbers above branches are lengths (>0). Numbers below branches are bootstrap percentages (>50%; 100 replicates). Branches that collapse in the strict consensus are noted by an ''X.'' Numbers and letters following species names (e.g., 1a, 2b) indicate numbered accessions and individual clonal sequences, respectively. See Botanical Society of America website (http://ajbsupp.botany.org/v90/) for details on each numbered accession.

graph. The species is now considered to be threatened with extinction in the wild (Sanjappa, 1990), but is known to survive under cultivation (Das and Thapliyal, 1999).

North American Gleditsia—*Gleditsia triacanthos* and *G. aquatica* are the only extant *Gleditsia* species in North America. Due in large part to its being planted for windbreaks, fodder, and as an ornamental and shade tree, *G. triacanthos* has greatly expanded its range in the United States and now covers nearly the whole of the eastern United States and adjacent Canada, as well as much of the Great Plains (Stephens, 1973; Blair, 1989). Its original distribution, however, is thought to have been in the east-central United States, with maximum abundance in southern Illinois and Indiana as a part of streamside and bottomland forests (Gordon, 1966; Blair, 1989). In contrast, *G. aquatica* ranges more narrowly in floodplains and swamps from the southeastern coastal plain of South Carolina, Georgia, and northern Florida west to eastern Texas and north along the Mississippi and Ohio river drainages as far as southern Illinois and Indiana (Gordon, 1966; Roberston and Lee, 1976; Godfrey and Wooten, 1981; Brown and Kirkman, 1990). Hybrids between the two species (*G.* × *texana*) have been reported in areas where the ranges overlap (Sargent, 1922; Vines, 1960; Gordon, 1966).

Gordon (1966) examined a large number of specimens of these two species and concluded that each was more closely related to Asian species than to each other. Our original cpDNA data suggested otherwise (Schnabel and Wendel, 1998). In all the analyses of cpDNA data, *G. triacanthos* and *G. aquatica* formed a well-supported clade, in which two *G. aquatica* accessions were sister to a clade containing both *G. triacanthos* and *G. aquatica* accessions. The relationships among those accessions, however, were based on a small number of characters, because the *ndhF* sequences of the two species were identical, and the *rpl16* intron sequences contained only seven shared polymorphisms and one species-specific indel.

We proposed two hypotheses to explain these results. First, we hypothesized that the cpDNA data were misleading, due to extensive hybridization and cytoplasmic introgression, and that Gordon's suggested relationships might emerge upon analysis of a nuclear gene data set. As an alternative, we hypothesized that Gordon's conclusions about relationships were incorrect and that the cpDNA data were pointing to the true relationship between these two species. This scenario envi-

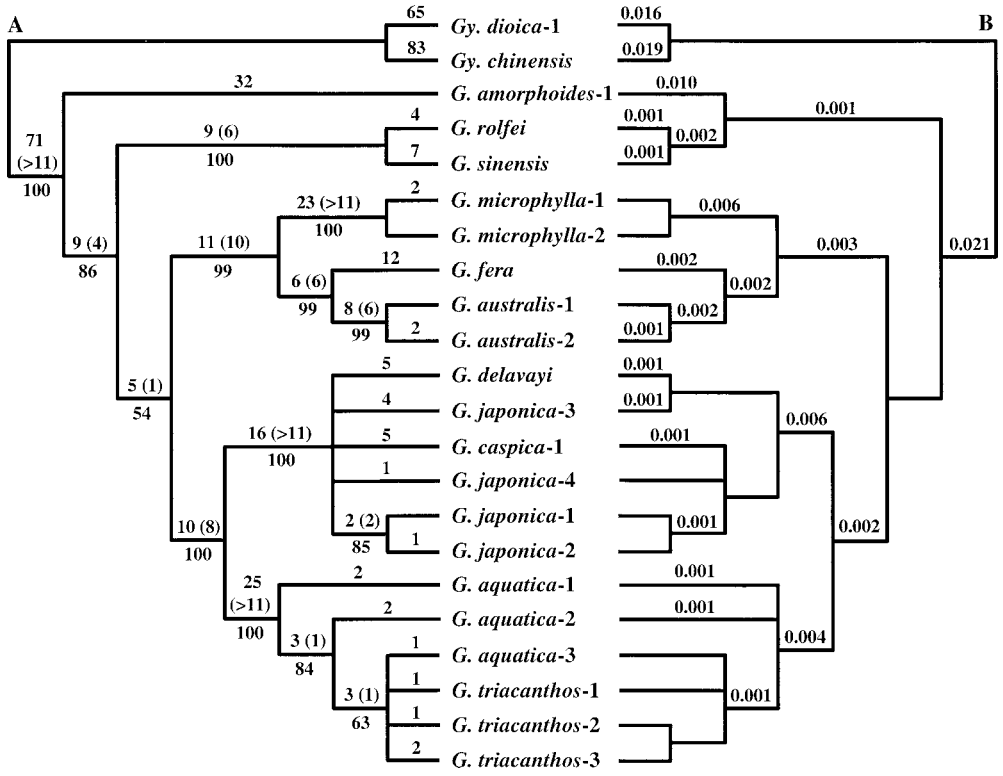

Fig. 3. Phylogenetic trees resulting from analysis of combined ITS and cpDNA sequence data (3998 bp) from two *Gymnocladus* and 20 *Gleditsia* accessions. (A) Strict consensus of 16 most parsimonious trees (430 steps, consistency index with autapomorphies excluded = 0.791, retention index = 0.900). Numbers above branches are lengths (>0) and decay indices (in parentheses). Numbers below branches are bootstrap percentages (>50%; 1000 replicates). (B) Maximum likelihood tree based on HKY+Γ model (Swofford et al., 1996), where −ln L = 7942.29377, transition/transversion ratio = 1.298, alpha = 0.014. Branch lengths ≥0.001 are indicated.

sions a recent divergence, in which shared polymorphisms are ascribed to incomplete lineage sorting.

The ITS data strongly support the monophyly of the North American lineage (16 synapomorphies, 100% bootstrap value), but again suggest that *G. triacanthos* and *G. aquatica* are paraphyletic. Although considerable nucleotide variability was observed within each species, most polymorphisms were autapomorphies. Among the 10 phylogenetically useful polymorphisms, we found no fixed differences and four shared polymorphisms. With such a small amount of information, relationships among the sampled accessions were not well defined. Total molecular evidence therefore suggests that Gordon (1966) was incorrect in his assessment of relationships between North American and Asian *Gleditsia* and that *G. triacanthos* and *G. aquatica* represent a pair of recently diverged species.

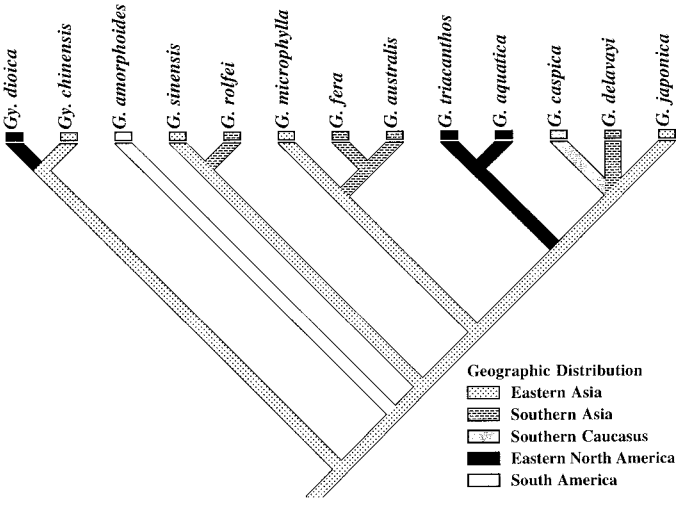

Fig. 4. Most parsimonious mapping of geographic distribution on the strict consensus MP tree from combined ITS/cpDNA data for *Gymnocladus* and *Gleditsia*. See Botanical Society of America website (http://ajbsupp.botany.org/v90/) for a more complete description of the geographic distribution for each species.

In general, species of *Gleditsia* show considerable morphological variability that makes them difficult to distinguish using either vegetative or floral characters (Gordon, 1966; Isely, 1975; Tucker, 1991), and *G. triacanthos* and *G. aquatica* are no exception. Gordon's (1966) descriptions of branches, spines, and flowers for these two species show overlapping variability in nearly all characters. Later work by Tucker (1991) on floral organogenesis demonstrated that *G. triacanthos* and *G. aquatica* have similar ranges of floral character variability. Thus, most available morphological data appear to mirror our molecular results.

Despite these morphological and molecular similarities, the two species are easily differentiated on the basis of fruit and seed morphology. Like most of the genus, *G. triacanthos* has a many-ovuled ovary (>25 ovules/fruit; Gordon, 1966) and elongate fruits with a thick pericarp and a sweet pulp. In contrast, the fruits of *G. aquatica* are short (elliptically shaped), contain 1–3 ovules, have a thinner, lighter pericarp, and lack

pulp. Seeds in *G. triacanthos* are oblong and 7–8 mm wide, whereas *G. aquatica* seeds are oval and 9–10.5 mm wide (Gordon, 1966). The striking differences between *G. aquatica* fruits and those of most other *Gleditsia* species have been assumed to result from natural selection for better fruit dispersal in a wetland habitat (Gordon, 1966; Heiser, 1985). Whatever the reason, the rapid evolution of this character stands in strong contrast to the much slower rate of evolution in other morphological characters, not only between *G. aquatica* and *G. triacanthos*, but also between members of the genus (e.g., *G. fera* and *G. rolfei*) that appear to have been evolving separately for many millions of years longer than have the North American species.

A biogeographic picture of Gleditsia and Gymnocladus—The close congruence between cpDNA and ITS data sets provides a solid phylogenetic framework for investigating the origins and biogeographic history of *Gleditsia* and *Gymnocladus*. Two major questions need to be addressed. First, when and where did the *Gleditsia*/*Gymnocladus* lineage evolve? Second, when and how did the major disjunctions in the two genera originate? To answer these questions, we interpret our phylogenetic results in light of current understandings about the phylogeny and fossil history of the Leguminosae, the fossil history of *Gleditsia* and *Gymnocladus*, and patterns of species migrations in the Northern Hemisphere during the Cenozoic.

Origin of Gymnocladus and Gleditsia—Gymnocladus and *Gleditsia* have long been considered a small monophyletic group within the tribe Caesalpinieae (Lee, 1976; Polhill and Vidal, 1981). Because of a suite of unspecialized floral features (Cowan, 1981; Polhill et al., 1981; Tucker, 1991), the *Gleditsia* group was originally considered to be among the most primitive of extant legumes (Dickison, 1981; Polhill and Vidal, 1981). This position was supported by the morphological cladistic analysis of Tucker and Douglas (1994), in which floral ontogenetic characters were used extensively, but not by a second morphological analysis performed by Chappill (1995). Subsequent molecular analyses have confirmed the monophyly of the group and have placed it as a lineage separate from the majority of the Caesalpinieae within a clade composed of several other small genera (e.g., *Acrocarpus*, *Ceratonia*) along with a large clade that includes the Mimosoideae and the rest of the Caesalpinieae (Käss and Wink, 1996; Doyle et al., 1997, 2000; Bruneau et al., 2001; Kajita et al., 2001; Herendeen et al., in press). The position of the *Gleditsia* group in these phylogenetic analyses indicates that it is older than the Mimosoideae but approximately the same age or younger than the Papilionoideae. Herendeen et al. (1992) emphasize that all major lineages of legumes were present during the Eocene (35–56 mya), and Magallón et al. (1999) place the minimum ages for the Mimosoideae and Papilionoideae at the middle (44.3 mya) and lower (53.3 mya) Eocene, respectively. The minimum age of the *Gleditsia*/*Gymnocladus* lineage must therefore be within these bounds. Moreover, the character mapping analysis (Fig. 4) supported an eastern Asian origin for both *Gymnocladus* and *Gleditsia*.

Unfortunately, the fossil record for *Gleditsia* and *Gymnocladus* is relatively slim compared to some other genera for which detailed biogeographic reconstructions have recently been attempted (cf. Xiang et al., 1998; Stanford et al., 2000; Whitcher and Wen, 2001). Although some Eocene legume fossils from North America show resemblence to *Gymnocladus* and *Gleditsia* (Axelrod, 1992; Herendeen, 1992b), the oldest reliable fossils for this lineage are from Oligocene sediments (23–35 mya) (P. S. Herendeen, George Washington University, personal communication). The best documentation of both genera, however, comes from the Miocene (5–23 mya) and Pliocene (2–5 mya). Miocene fossils of *Gleditsia* and *Gymnocladus* have been found in western North America (Prakash and Barghoorn, 1961; Axelrod, 1992), and Miocene and Pliocene *Gleditsia* fossils are known from eastern North America (McCartan et al., 1990; Wheeler and Baas, 1992). All of these *Gleditsia* fossils (wood and leaves) show strong affinities with *G. triacanthos*. In Asia, Miocene fossils document the presence of *Gleditsia* in Japan (Tanai, 1972; Wheeler and Baas, 1992), *Gleditsia* and *Gymnocladus* in eastern China (Hsü, 1983; Guo and Zhou, 1992; Tao, 1992), and *Gleditsia* in the Caucasus (Shakryl, 1992). Both genera are also found among Pliocene fossils from the Caucasus (Shakryl, 1992). The Chinese Miocene *Gleditsia* have been divided into two species, *G. parajaponica* Guo & Zhou and *G. miosinensis* Hu & Chaney, that are considered to be most similar to the extant *G. japonica* and *G. sinensis* (Guo and Zhou, 1992). *Gleditsia miosinensis* has been found in east-central (Shandong Province) and southwestern China (Yunnan Province), whereas *G. parajaponica* is known only from a single east-central locality (Zhejiang Province). No reliable fossils of either genus have been described from Europe (e.g., Herendeen, 1992a).

In summary, the combination of fossil and phylogenetic analyses supports an eastern Asian origin of both *Gleditsia* and *Gymnocladus* during the early to middle Eocene. The relatively sparse fossil record, however, leaves many questions unanswered about diversity and distribution of these genera during the Eocene and Oligocene, but it clearly indicates the presence of several lineages in Asia and North America by the Miocene and a widespread Asian distribution from eastern China to the Caucasus by the Pliocene.

Origins of eastern Asian-eastern North American disjunctions—Several authors have discussed the mechanisms by which Asian-North American disjunctions could arise (see reviews in Manchester, 1999; Tiffney, 2000). Both the North Atlantic and Bering land bridges have been important migration routes between Eurasia and North America (Manchester, 1999; Tiffney, 2000). Because the North Atlantic land bridge was broken by the late Eocene (Tiffney, 2000), and because there is no fossil evidence of either *Gleditsia* or *Gymnocladus* in Europe, it seems most likely that current Asian-North American disjunctions in these genera arose through migrations across the Bering land bridge. Figure 4 further supports a hypothesis of migration from Asia to North America and not vice versa. The Bering land bridge is known to have been operational until 5–7 mya (Marincovich and Gladenkov, 1999), although at that time (the end of the Miocene), the Alaskan climate may have been too cool to support warm temperate taxa like *Gymnocladus* and *Gleditsia* (White et al., 1997; Tiffney, 2000).

Although the migration route taken to reach North America seems clear, the timing of these events is not. Because molecular divergence at ITS and the cpDNA genes is more than twice as great between *Gymnocladus dioica* and *Gy. chinensis* than it is between members of the *Gleditsia triacanthos* and *G. japonica* clades, we suspect that the current eastern Asian-eastern North American disjunction in *Gymnocladus* arose considerably earlier than the equivalent disjunction in *Gledit-*

sia. The paucity of fossil evidence from the Eocene and Oligocene, however, makes any hypotheses about diversification and migration during those epochs merely speculative. If estimates of divergence times based on molecular clocks are to be believed, the Miocene was an important period for flowering plant migrations between Asia and North America (reviewed in Wen, 1999), and evidence of both *Gymnocladus* and *Gleditsia* in North America during that time suggests that migration must have occurred at least as early as 15–20 mya.

Regardless of the absolute timing of these migration events, our molecular data refute Gordon's (1966) hypothesis of two intercontinental sister-species pairs in *Gleditsia*. The ITS and cpDNA data sets both strongly indicate that *Gleditsia aquatica* and *G. microphylla* are only distantly related and that no intercontinental species pairs exist. Following what appears to be a general pattern for east Asian-eastern North American disjuncts (Wen, 1999; Stanford et al., 2000), our data point to species diversification occurring within each of the disjunct lineages subsequent to migration of the *G. japonica*-like ancestor across the Bering land bridge. The evidence of Miocene fossils resembling *G. japonica* and *G. triacanthos*, combined with the observation of extremely low levels of genetic divergence within each of these lineages (for ITS, $<1\%$ mean divergence within the *G. japonica* clade and 1.4% mean divergence within the *G. triacanthos* clade) compared to divergence between lineages (6.5–7% for ITS), suggests that these lineages existed long before the currently ongoing diversification. Thus, these two lineages were widely distributed on their respective continents during the Miocene, but diversification is apparently more recent, perhaps beginning only during the Pliocene.

Origin of Asian-South American disjunction in Gleditsia—One of the most confounding aspects of *Gleditsia* biogeography is the origin of *G. amorphoides*, which is currently found in temperate, east-central South America and represents a rare and poorly studied type of floristic disjunction (Thorne, 1972). The placement of *G. amorphoides* within the *Gleditsia* phylogeny was the only significant incongruity between the cpDNA trees and the ITS trees. Our cpDNA phylogeny placed *G. amorphoides* as a close relative of *G. sinensis*/*G. rolfei*, and using a molecular clock hypothesis, we suggested that *G. amorphoides* diverged from an Asian ancestor 5–7 mya (Schnabel and Wendel, 1998). In contrast, the ITS data indicate that *G. amorphoides* is the most genetically divergent species in the genus, and all ITS trees show *G. amorphoides* to be sister to the rest of the genus. The results from combined analysis are equivocal, with the MP trees supporting this basal split and the ML tree supporting a sister relationship with the *G. sinensis* clade.

Whichever of these topologies is correct, *G. amorphoides* clearly has its closest extant relative in eastern Asia. Moreover, the many molecular and morphological (Gordon, 1966) differences between this species and the rest of the genus suggest that *G. amorphoides* has been isolated from the rest of *Gleditsia* for millions of years. Because South America was an island throughout most of the Tertiary, the ancestor of *G. amorphoides* must have reached South America by long-distance dispersal. This dispersal could have occurred directly from Asia across the Pacific Ocean or indirectly by means of migration from Asia to North America followed by dispersal from North America to South America. The latter hypothesis assumes that the North American ancestor is now extinct, leaving a phylogeny that shows a direct link with Asian *Gleditsia*. This same problem was faced by Soltis et al. (2001) for *Chrysosplenium*, which also possesses a temperate South American disjunct with a sister species in Asia, but in neither case can current data eliminate one or the other possible explanations. As discussed by Soltis et al. (2001), Asian-North American-South American disjunctions probably originated in multiple ways, and each case may be different. Cameron and Chase (1999), for example, argue that a similar disjunction in the orchid subtribe Pogoniinae arose in the opposite direction, through migration from South America to Asia by way of North America and the Bering land bridge.

In conclusion, *Gleditsia* and *Gymnocladus* most likely evolved in eastern Asia in the middle or early Eocene. Species of both genera apparently migrated across the Bering land bridge to North America, but the much greater molecular divergence between Asian and North American *Gymnocladus* than between Asian and North American *Gleditsia* suggests that *Gymnocladus* may have migrated much earlier than *Gleditsia*. Fossil evidence from Miocene strata clearly show that migration throughout Asia and North America had occurred by 15–20 mya. The lone *Gleditsia* species in South America most likely descended from an ancestor that reached that continent by long-distance dispersal in the Oligocene or Miocene. Within Asia, *Gymnocladus* has exhibited little diversification, whereas *Gleditsia* has branched into three main lineages, two of which include significant intracontinental disjunctions, and one of which gave rise to the extant North American clade. The North American species, although morphologically and ecologically distinct from one another, are polymorphic and paraphyletic with respect to molecular characters. Speciation within this clade therefore appears to be recent, possibly beginning during the Pliocene.

LITERATURE CITED

AXELROD, D. I. 1992. Climatic pulses, a major factor in legume evolution. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 259–279. Royal Botanic Gardens, Kew, UK.

BLAIR, R. M. 1989. *Gleditsia triacanthos* L. (honeylocust). *In* R. M. Burns and B. H. Honkala [technical coordinators], Silvics of North America, 2, hardwoods. Agricultural Handbook 654, 358–364. U.S. Department of Agriculture, Forest Service, Washington, D.C., USA.

BOUFFORD, D. E., AND S. A. SPONGBERG. 1983. Eastern Asian-eastern North American phytogeographical relationships: a history from the time of Linnaeus to the twentieth century. *Annals of the Missouri Botanical Garden* 70: 423–439.

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

BROWN, C. L., AND L. K. KIRKMAN. 1990. Trees of Georgia and adjacent states. Timber Press, Portland, Oregon, USA.

BRUNEAU, A., F. FOREST, P. S. HERENDEEN, B. B. KLITGAARD, AND G. P. LEWIS. 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Systematic Botany* 26: 487–514.

BUCKLER, E. S., IV, AND T. P. HOLTSFORD. 1996. *Zea* systematics: ribosomal ITS evidence. *Molecular Biology and Evolution* 13: 612–622.

BUCKLER, E. S., IV, A. IPPOLITO, AND T. P. HOLTSFORD. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145: 821–832.

CAMERON, K. M., AND M. W. CHASE. 1999. Phylogenetic relationships of Pogoniinae (Vanilloideae, Orchidaceae): an herbaceous example of the eastern North America-eastern Asia phytogeographic disjunction. *Journal of Plant Research* 112: 317–329.

CHANEY, R. W. 1947. Tertiary centers and migration routes. *Ecological Monographs* 17: 139–148.

CHAPPILL, J. A. 1995. Cladistic analysis of the Leguminosae: the develop-

ment of an explicit phylogenetic hypothesis. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, phylogeny, 1–9. Royal Botanic Gardens, Kew, UK.

COWAN, R. S. 1981. Caesalpinioideae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 57–65. Royal Botanic Gardens, Kew, UK.

DAS, P., AND R. C. THAPLIYAL. 1999. Enhancement of germination of *Adenanthera pavonina* and *Gleditsia assamica* by pretreatment. *Indian Journal of Forestry* 22: 37–41.

DICKISON, W. C. 1981. The evolutionary relationships of the Leguminosae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 35–54. Royal Botanic Gardens, Kew, UK.

DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, AND J. D. PALMER. 1992. Phylogenetic relationships of Dipsicales based on *rbc*L sequences. *Annals of the Missouri Botanical Garden* 79: 249–265.

DOYLE, J. J., J. A. CHAPPILL, C. D. BAILEY, AND T. KAJITA. 2000. Towards a comprehensive phylogeny of legumes: evidence from *rbcL* sequences and non-molecular data. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 1–20. Royal Botanic Gardens, Kew, UK.

DOYLE, J. J., J. L. DOYLE, J. A. BALLENGER, E. E. DICKSON, T. KAJITA, AND H. OHASHI. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541–554.

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

GODFREY, R. K., AND J. W. WOOTEN. 1981. Aquatic and wetland plants of southeastern United States, dicotyledons. University of Georgia Press, Athens, Georgia, USA.

GORDON, D. 1966. A revision of the genus *Gleditsia* (Leguminosae). Ph.D. dissertation, Indiana University, Bloomington, Indiana, USA.

GRAHAM, A. 1972. Outline of the origin and historical recognition of floristic affinities between Asia and eastern North America. *In* A. Graham [ed.], Floristics and paleofloristics of Asia and eastern North America, 1–18. Elsevier, Amsterdam, Netherlands.

GUO, Q., AND R. E. RICKLEFS. 2000. Species richness in plant genera disjunct between temperate eastern Asia and North America. *Botanical Journal of the Linnean Society* 134: 401–423.

GUO, S.-X., AND Z.-K. ZHOU. 1992. The megafossil legumes from China. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 207–223. Royal Botanic Gardens, Kew, UK.

HEISER, C. B., JR. 1985. Of plants and people. University of Oklahoma Press, Norman, Oklahoma, USA.

HERENDEEN, P. S. 1992a. A re-evaluation of the fossil genus *Podogonium* Heer. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 3–18. Royal Botanic Gardens, Kew, UK.

HERENDEEN, P. S. 1992b. The fossil history of the Leguminosae from the Eocene of southeastern North America. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 85–160. Royal Botanic Gardens, Kew, UK.

HERENDEEN, P. S., A. BRUNEAU, AND G. P. LEWIS. In press. Phylogenetic relationships in Caesalpinioid legumes: a preliminary analysis based on morphological and molecular data. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10. Royal Botanic Gardens, Kew, UK.

HERENDEEN, P. S., W. L. CREPET, AND D. L. DILCHER. 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 303–316. Royal Botanic Gardens, Kew, UK.

HIGGINS, D. G., J. D. THOMPSON, AND T. J. GIBSON. 1996. Using CLUSTAL for multiple sequence alignments. *Methods in Enzymology* 266: 383–402.

HSÜ, J. 1983. Late Cretaceous and Cenozoic vegetation in China, emphasizing their connections with North America. *Annals of the Missouri Botanical Garden* 70: 490–508.

HU, J.-M., M. LAVIN, M. F. WOJCIECHOWSKI, AND M. J. SANDERSON. In press. Phylogenetic analysis of nuclear ribosomal ITS/5.8S sequences in the tribe Millettieae (Fabaceae): *Poecilanthe-Cyclolobium*, the core Millettieae, and the Callerya group. *Systematic Botany*.

HUH, M. K., H. Y. LEE, AND H. W. HUH. 1999. Genetic diversity and population structure of *Gleditsia japonica* var. *koraiensis* in Korea. *Botanical Bulletin of Academia Sinica* 40: 159–164.

ISELY, D. 1975. Leguminosae of the United States: II. Subfamily Caesalpinioideae. *Memoirs of the New York Botanical Garden* 25: 1–228.

KAJITA, T., H. OHASHI, Y. TATEISHI, C. D. BAILEY, AND J. J. DOYLE. 2001. *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. *Systematic Botany* 26: 515–536.

KÄSS, E., AND M. WINK. 1996. Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcL*-sequences. *Biochemical Systematics and Ecology* 24: 365–378.

KELLY, L. M. 1998. Phylogenetic relationships in *Asarum* (Aristolochiaceae) based on morphological and ITS sequences. *American Journal of Botany* 85: 1454–1467.

KIM, Y.-D., AND S.-H. KIM. 1999. Phylogeny of *Weigela* and *Diervilla* (Caprifoliaceae) based on nuclear rDNA ITS sequences: biogeographic and taxonomic implications. *Journal of Plant Research* 112: 331–341.

LARSEN, K. 1989. *Gleditsia* Lin. (Leguminosae-Caesalpinioideae) a genus new to Thailand. *Thai Forest Bulletin* 18: 84–87.

LARSEN, K., S. S. LARSEN, AND J. E. VIDAL. 1980. Flore du Cambodge, du Laos, et du Viêt-Nam. 18. Légumineuses-Césalpinioïdées. Muséum National D'Histoire Naturelle, Paris, France.

LEE, Y.-T. 1976. The genus *Gymnocladus* and its tropical affinity. *Journal of the Arnold Arboretum* 57: 91–112.

LI, L.-C. 1982. Two new varieties of *Gleditsia japonica* Miq. *Acta Phytotaxonomica Sinica* 20: 228–229.

LI, L.-C. 1988. *Gleditsia*. *In* Flora reipublicae popularis sinicae, tomus 39 (Leguminosae, Caesalpinioideae), 80–90. Science Press, Beijing, China.

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

MAGALLON, S., P. R. CRANE, AND P. S. HERENDEEN. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. *Annals of the Missouri Botanical Garden* 86: 297–372.

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

MARINCOVICH, L., JR., AND A. Y. GLADENKOV. 1999. Evidence of an early opening of the Bering Strait. *Nature* 397: 149–151.

MCCARTAN, L., B. H. TIFFNEY, J. A. WOLFE, S. L. WING, L. A. SIRKIN, L. W. WARD, AND J. BROOKS. 1990. Late Tertiary floral assemblage from upland gravel deposits of the southern Maryland Coastal Plain. *Geology* 18: 311–314.

PACLT, J. 1982a. *Gleditsia caspia*, not a distinct species (Leguminosae). *Taxon* 31: 336–339.

PACLT, J. 1982b. On the repeatedly confused nomenclature of Chinese species of *Gleditsia* (Caesalpiniaceae). *Taxon* 31: 551–553.

PACLT, J. 1984. A note on *Triaenodendron*, with new combinations in *Gleditsia* (Caesalpiniaceae). *Taxon* 33: 100–101.

PATERSON, A. H., C. L. BRUBAKER, AND J. F. WENDEL. 1993. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Molecular Biology Reporter* 11: 112–127.

POLHILL, R. M., P. H. RAVEN, AND C. H. STIRTON. 1981. Evolution and systematics of the Leguminosae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 1–26. Royal Botanic Gardens, Kew, UK.

POLHILL, R. M., AND J. E. VIDAL. 1981. Caesalpinieae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 81–96. Royal Botanic Gardens, Kew, UK.

PRAKASH, U., AND E. S. BARGHOORN. 1961. Miocene fossil woods from the Columbia basalts of central Washington. *Journal of the Arnold Arboretum* 42: 165–203.

QIU, X., L. WU, H. HUANG, P. E. MCDONEL, A. V. PALUMBO, J. M. TIEDJE, AND J. ZHOU. 2001. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S ribosomal RNA gene-based cloning. *Applied and Environmental Microbiology* 67: 880–887.

QIU, Y.-L., M. W. CHASE, AND C. R. PARKS. 1995. A chloroplast DNA phylogenetic study of the eastern Asian-eastern North American disjunct section *Rytidospermum* of *Magnolia* (Magnoliaceae). *American Journal of Botany* 82: 1582–1588.

ROBERTSON, K. R., AND Y.-T. LEE. 1976. The genera of Caesalpinioideae (Leguminosae) in the southeastern United States. *Journal of the Arnold Arboretum* 57: 1–31.

SANJAPPA, M. 1990. *Gleditsia assamica* Bor. *In* M. P. Nayar and A. R. K.

Sastry [eds.], Red data book of Indian plants, 139–140. Botanical Survey of India, Calcutta, India.

SARGENT, C. S. 1922. Notes on North American trees. X. New species and varieties of *Crataegus*. *Journal of the Arnold Arboretum* 3: 182–207.

SCHNABEL, A., AND J. F. WENDEL. 1998. Cladistic biogeography of *Gleditsia* (Leguminosae) based on *ndhF* and *rpl16* chloroplast gene sequences. *American Journal of Botany* 85: 1753–1765.

SHAKRYL, A. K. 1992. Leguminosae species from the Tertiary of Abkhazia. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 189–206. Royal Botanic Gardens, Kew, UK.

SOLTIS, D. E., M. TAGO-NAKAZAWA, Q.-Y. XIANG, S. KAWANO, J. MURATA, M. WAKABAYASHI, AND C. HIBSCH-JETTER. 2001. Phylogenetic relationships and evolution in *Chrysosplenium* (Saxifragaceae) based on *matK* sequence data. *American Journal of Botany* 88: 883–893.

STANFORD, A. M., R. HARDEN, AND C. R. PARKS. 2000. Phylogeny and biogeography of *Juglans* (Juglandaceae) based on *matK* and ITS sequence data. *American Journal of Botany* 87: 872–882.

STEPHENS, H. A. 1973. Woody plants of the north central plains. University of Kansas Press, Lawrence, Kansas, USA.

STEWARD, X. 1958. Manual of vascular plants of the lower Yangtze Valley, China. Oregon State College, Corvallis, Oregon, USA.

SWOFFORD, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0b10. Sinauer, Sunderland, Massachusetts, USA.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inference. *In* D. M. Hillis, C. Moritz, and B. K. Mable [eds.], Molecular systematics, 407–514. Sinauer, Sunderland, Massachusetts, USA.

TANAI, T. 1972. Tertiary history of vegetation in Japan. *In* A. Graham [ed.], Floristics and paleofloristics of Asia and eastern North America, 235–255. Elsevier, Amsterdam, Netherlands.

TAO, J.-R. 1992. The tertiary vegetation and flora and floristic regions in China. *Acta Phytotaxonomica Sinica* 30: 25–43.

THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

THORNE, R. F. 1972. Major disjunctions in the geographic ranges of seed plants. *Quarterly Review of Biology* 47: 365–410.

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.

TIFFNEY, B. H. 2000. Geographic and climatic influences on the Cretaceous and Tertiary history of Euramerican floristic similarity. *Acta Universitatis Carolinae—Geologica* 44: 5–16.

TUCKER, S. C. 1991. Helical floral organogenesis in *Gleditsia*, a primitive caesalpinioid legume. *American Journal of Botany* 78: 1130–1149.

TUCKER, S. C., AND A. W. DOUGLAS. 1994. Ontogenetic evidence and phylogenetic relationships among basal taxa of legumes. *In* I. K. Ferguson and S. C. Tucker [eds.], Advances in legume systematics, part 6: structural botany, 11–32. Royal Botanic Gardens, Kew, UK.

VINES, R. A. 1960. Trees, shrubs, and woody vines of the Southwest. University of Texas Press, Austin, Texas, USA.

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., AND S. SHI. 1999. A phylogenetic and biogeographic study of *Hamamelis* (Hamamelidaceae), an eastern Asian and eastern North American disjunct genus. *Biochemical Systematics and Ecology* 27: 55–66.

WEN, J., S. SHI, R. K. JANSEN, AND E. A. ZIMMER. 1998. Phylogeny and biogeography of *Aralia* sect. *Aralia* (Araliaceae). *American Journal of Botany* 85: 866–875.

WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–249.

WHEELER, E., AND P. BAAS. 1992. Fossil wood of the Leguminosae: a case study in xylem evolution and ecological anatomy. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, the fossil record, 281–301. Royal Botanic Gardens, Kew, UK.

WHITCHER, I. N., AND J. WEN. 2001. Phylogeny and biogeography of *Corylus* (Betulaceae): inferences from ITS sequences. *Systematic Botany* 26: 283–298.

WHITE, J. M., T. A. AGER, D. P. ADAM, E. B. LEOPOLD, G. LIU, H. JETTE, AND C. E. SCHWEGER. 1997. An 18 million year record of vegetation and climate change in northwestern Canada and Alaska: tectonic and global correlates. *Palaeogeography, Palaeoclimatology, Palaeoecology* 130: 293–306.

WHITE, T. J., T. BRUNS, S. LEE, AND J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. Innis, D. Gelfand, J. Sninsky, and T. White [eds.], PCR protocols: a guide to methods and applications, 315–322. Academic Press, San Diego, California, USA.

WOON, Y. C. 1921. Chinese economic trees. Commercial Press, Shanghai, China.

XIANG, Q.-Y., D. J. CRAWFORD, A. D. WOLFE, Y.-C. TANG, AND C. W. DEPAMPHILIS. 1998. Origin and biogeography of *Aesculus* L. (Hippocastanaceae): a molecular phylogenetic perspective. *Evolution* 52: 988–997.