Evolution, 52(4), 1998, pp. 988–997

ORIGIN AND BIOGEOGRAPHY OF *AESCULUS* L. (HIPPOCASTANACEAE): A MOLECULAR PHYLOGENETIC PERSPECTIVE

QIU-YUN XIANG,[1] DANIEL J. CRAWFORD,[2,3] ANDREA D. WOLFE,[2] YAN-CHENG TANG,[4] AND CLAUDE W. DEPAMPHILIS[5]

[1]*Department of Biological Sciences, Idaho State University, Pocatello, Idaho 83209-8007*
[2]*Department of Plant Biology, The Ohio State University, 1735 Neil Avenue, Columbus, Ohio 43210-1293*
[4]*Laboratory of Systematic and Evolutionary Botany, Chinese Academy of Sciences, Xiangshan, Beijing 100093, China*
[5]*Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802*

Abstract.—Sequences of chloroplast gene *matK* and internal transcribed spacers of nuclear ribosomal RNA genes were used for phylogenetic analyses of *Aesculus*, a genus currently distributed in eastern Asia, eastern and western North America, and southeastern Europe. Phylogenetic relationships inferred from these molecular data are highly correlated with the geographic distributions of species. The identified lineages closely correspond to the five sections previously recognized on the basis of morphology. Ancestral character-state reconstruction, a molecular clock, and fossil evidence were used to infer the origin and biogeographic history of the genus within a phylogenetic framework. Based on the molecular phylogenetic reconstruction of the genus, sequence divergence, and paleontological evidence, we infer that the genus originated during the transition from the Cretaceous to the Tertiary (~65 M.Y.B.P.) at a high latitude in eastern Asia and spread into North America and Europe as an element of the "boreotropical flora"; the current disjunct distribution of the genus resulted from geological and climatic changes during the Tertiary.

Key words.—*Aesculus*, biogeography, Hippocastanaceae, internal transcribed spacers, *matK*, origin, phylogeny.

Received July 3, 1997. Accepted March 16, 1998.

The buckeye genus (*Aesculus*) represents one of the most remarkable examples of intercontinental disjunction of plants in the Northern Hemisphere. The genus has 13–19 species that are currently found in eastern Asia, eastern and western North America, and Europe (Hardin 1957a, 1960; Fang 1981). The taxonomy of *Aesculus* has long been controversial and its history was reviewed thoroughly by Hardin (1957a, 1960). In the most recent monographic study of *Aesculus* by Hardin (1957a, 1960), 13 species, divided among five sections, were recognized within the genus (see Table 1; Fig. 1). More recently, six new species of *Aesculus* were described from China (Fang 1981). These are *A. polyneura* Hu et Fang, *A. lantsangensis* Hu et Fang, *A. tsiangii* Hu et Fang, *A. wangii* Hu et Fang, *A. chuniana* Hu et Fang, and *A. megaphylla* Hu et Fang. These new species are nearly indistinguishable from previously described species and their recognition remains tentative.

A phylogenetic scheme was proposed for *Aesculus* by Hardin (1957a), based on a noncladistic analysis of morphological data (Fig. 1). Hardin (1957a) believed that *Aesculus* diverged into two major groups, one consisting of the eastern North American species (section *Pavia* or Sect. A in figures) with the exception of *A. parviflora*, and a second group that subsequently diverged into a Mexican species (section *Parryana* or Sect. B in figures) and a large group that later split into three subgroups: section *Aesculus* (or Sect. D in figures) containing two species, one from Japan and one from southeastern Europe; section *Calothyrsus* (or Sect. E in figures) containing five species, one from California and four from eastern Asia; and section *Macrothyrsus* (or Sect. C in figures) containing a single species (*A. parviflora*) from Alabama and Georgia (Fig. 1). Based on this phylogenetic hypothesis and the fact that the genus *Billia*, the other member of Hippocastanaceae and the putative sister group of *Aesculus*, occurs in tropical America (Central and South America), Hardin (1957a) proposed that *Aesculus* originated in Central or South America from a *Billia*-like ancestor with subsequent migration northward into North America in the early Tertiary or earlier, with one part to the Appalachian area and a second element up the west coast of North America and on to Asia and Europe via the Bering Strait. In contrast, Raven and Axelrod (1974, 1978), using paleontological evidence, suggested that *Aesculus* as well as the family Hippocastanaceae originated in North America and subsequently migrated to other continents. They suggested that *Billia* did not attain its distribution in South America until the Pliocene or perhaps even more recently.

A clear understanding of the phylogenetic relationships of *Aesculus* is the prerequisite for evaluating these differing biogeographic hypotheses. Thus, our goals in this study were (1) to conduct phylogenetic analyses for *Aesculus* using molecular data from both chloroplast (sequences of *matK*) and nuclear (sequences of internal transcribed spacer [ITS] of rDNAs) genomes; and (2) to gain insights into the origin and biogeographic history of the genus from a molecular phylogenetic perspective.

MATERIALS AND METHODS

All species of *Aesculus* recognized in Hardin's (1957a, 1960) revision of the genus were included in this study with the exception of one species from eastern Asia (*A. assamica* Griffith) for which leaf material was not available (Table 1). One of the six newly described species, *A. wangii*, was also included. The bitypic neotropical genus *Billia* (Hippocastanaceae) and the monotypic genus *Handeliodendron* (Sapindaceae) from southern China were chosen as outgroups based on a recent phylogenetic analysis for Sapindaceae and putative relatives by Judd et al. (1994) using morphological

[3] Corresponding author. E-mail: crawford.13@osu.edu.

TABLE 1. Leaf materials used for internal transcribed spacer and *matK* sequencing. Sections of *Aesculus* are listed following the sequence of Hardin (1957a). RSABG, Rancho Santa Ana Botanical Garaden. Sequences and vouchers without information for herbaria where they are housed are available from QYX.

Taxon	Voucher	Locality
Ingroup:		
Aesculus L.		
Sect. *Parryanae* Wiggins		
A. parryi Gray	RSABG, BC-17230	Baja California, Mexico
Sect. *Aesculus*		
A. turbinata Blume	J. Wen, sene no.	Nikko Botanical Garden, Japan
A. hippocastanum L.	Kew, 000-69.11289-263	
Sect. *Calothyrsus* (Spach) K. Koch		
A. californica (Spach) Nuttall-1	D. J. Crawford, 406, OS	Alameda Co., California
A. californica (Spach) Nuttall-2	T. M. Hardig, 2795, WS	Calaveras Co., California
A. chinensis Bunge	Q.-Y. Xiang, 305	Mts. Qingling, China
A. indica (Camb.) Hook	Q.-Y. Xiang, 301	Godawari, Kathmandu, Nepal
A. wangii Hu et Fang	Q.-Y. Xiang, 300	Kunming, Yunnan, China
A. wilsonii Rehder	Q.-Y. Xiang, 303	Mao Co., Sichuan, China
Sect. *Macrothyrsus* (Spach) K. Koch		
A. parviflora Walter	Kew, 000-69.10442-265	
Sect. *Pavia* (Mill.) Persoon		
A. flava Solander-1	D. J. Crawford, 408, OS	Vinton Co., Ohio
A. flava Solander-2	C. W. DePamphilis, F-Ml-4	Mantair Lake, North Carolina
A. glabra Willd.-1	D. J. Crawford, 413, OS	Miami Co., Ohio
A. glabra Willd.-2	C. W. DePamphilis, 931	USA
A. pavia L.-1	D. J. Crawford, 404, OS	Franklin Co., Ohio
A. pavia L.-2	C. W. DePamphilis, P-Tu-2	Tuscaloosa Co, Alabama
A. sylvatica Bartram-1	C. W. DePamphilis, 50	Clarke Co., Georgia
A. sylvatica Bartram-2	C. W. DePamphilis, S-Ga-2	Gaston Co., North Carolina
Outgroups:		
Billia sp.	B. Hammel, 20075, OS	Costa Rico
Handeliodendron bodinieri (Levl.)	Rehder Q.-Y. Xiang, 302	Kuiling, Guangxi, China

data. The results of that analysis indicated that *Aesculus, Billia*, and *Handeliodendron* formed a strongly supported trichotomical clade and suggested that *Handeliodendron* is more closely related to *Aesculus* and *Billia* than to other members of Sapindaceae.

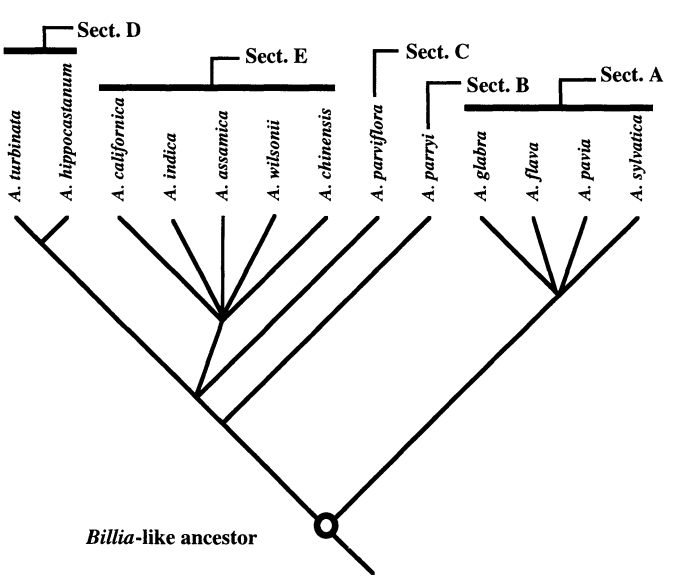

FIG. 1. Phylogenetic and classification schemes of *Aesculus* proposed by Hardin (1957). Sect. A, section *Pavia*; Sect. B, section *Parryana*; Sect. C, section *Macrothyrsus*; Sect. D, section *Aesculus*; and Sect. E, section *Calothyrsus*.

DNAs were isolated from silica-gel dried or fresh leaf material using the miniprep protocol of Cullings (1992), which we modified by grinding a piece of 2 × 2 cm leaf tissue in 1 ml of 2× CTAB with 4% PVP and 0.5% of β-mercaptoethanol with a mortar and pestle. The complete ITS region of nuclear rDNA and the chloroplast gene *matK* were amplified using PCR. Primers ITS-5m (forward; Sang et al. 1995) and ITS-4 (reverse; White et al. 1990) were used for ITS amplification, which covers the 5.8S as well as the ITS-1 and ITS-2 regions. Primers *matK*-1F and *matK*-1R (Sang et al. 1997) were used for *matK* amplification. Double-stranded PCR products were used as templates for sequencing of *matK* and single-stranded PCR products were used as templates for sequencing of ITS. Single-stranded ITS sequences were amplified by asymmetric PCR using the double-stranded PCR products as templates and only one PCR primer, ITS-4 or ITS-5m. PCR products were cleaned following the method of Morgan and Soltis (1993) and Xiang et al. (1993). Methods for *matK* sequencing followed Sang et al. (1997), and the ITS sequencing followed Baldwin (1992) and Soltis and Kuzoff (1995) using the internal primers ITS-2 (reverse) for ITS-1 region and ITS-3 (forward) for ITS-2 region.

Sequences of ITS were aligned first using the computer program CLUSTAL V (Higgins et al. 1992) and then were adjusted manually. Experiments with various alignment parameters (gap opening cost/gap extension cost of 50/8, 30/7, 20/5, 10/5, 5/3, and 3/1) using CLUSTAL V gave highly similar alignments that yielded identical or nearly identical tree topologies to that derived from the manually adjusted

alignment. Sequences of *matK* were aligned easily, without ambiguity (see Results below), using the *matK* of *Alangium* (Cornaceae; Xiang et al. 1998) and *Saxifraga* (Saxifragaceae; Johnson and Soltis 1994) as references. For both ITS and *matK* sequences, nucleotide compositions and sequence divergence between species were examined using MEGA 1.0 (Kumar et al. 1993). The manually adjusted alignment of the ITS sequences was used for these analyses and the following phylogenetic analyses. Phylogenetic analyses were conducted for both ITS and *matK* sequences separately using standard Fitch parsimony on PAUP 3.1.1 (Swofford 1993). The heuristic search parameters included the MULPARS option and random taxon addition of 100 replicates, and tree-bisection-reconnection (TBR) branch-swapping. Characters were defined as unordered with equal weight. Analyses of ITS sequences excluding regions with uncertain alignment were also conducted and the same tree topology was found.

The ITS and *matK* tree topologies obtained were tested for incongruence using the Templeton test (Templeton 1983) following Larson (1994). Results indicated that the *matK* tree does not significantly reject the ITS tree ($n = 14$, $Ts = 40 > T_{0.05} = 21$), but the ITS tree significantly rejects the *matK* tree ($n = 37$, $Ts = 91 < T_{0.001} = 140$), suggesting the two trees are incongruent. Therefore, the two datasets were not combined for further analysis. Relative nodal support for both ITS and *matK* shortest trees was estimated by bootstrap and Bremer support (decay) analyses (Felsenstein 1985; Bremer 1988). The boostrap analysis of 100 replicates was conducted using the same parameters as above. The largest number of trees that can be saved by PAUP 3.1.1 (32,767) was reached at 60 replicates of bootstrap analysis for the *matK* sequence data matrix. The Bremer support analysis was conducted following Eernisse and Kluge (1993). A skewness test (Hillis and Heulsenbeck 1992) was performed in PAUP using 10,000 random trees to examine the structures of both ITS and *matK* data matrices.

RESULTS

Rate and Pattern of Sequence Variation

ITS

The aligned ITS sequence data matrix with manual adjustment contains a total of 528 bp. Of the 528 bp, 256 bp are from ITS-1 and 272 bp from ITS-2. Of the 528 sites, 239 (45.3%) are variable and 109 (20.6%) are potentially informative. Nucleotide composition of ITS sequences in *Aesculus* and its sister groups is skewed to a high G-C content (66.8% G-C vs. 33.2% A-T). The sequence divergence of ITS between species of *Aesculus* ranges from 0.040 to 0.2752, with an average of 0.142 nucleotide substitutions per site. The g_1-value from the skewness test (Hillis and Huelsenbeck 1992) is -0.677978, indicating that the ITS data matrix contains phylogenetic signal ($P = 0.01$).

matK

Only approximately two-thirds of the *matK* sequence was obtained for all species in this study due to sequence divergence between one of the primers, *matK*-2R (designed for *Paeonia*), and the corresponding region of *matK* in *Aesculus*. The *matK* data matrix contains 1210 bp including 1150-bp of the 5′ coding region of the gene and 60 bp from the *trnK* intron flanking the 5′ end of *matK*. The aligned 5′ portion of *matK* coding region has four gaps, and none is phylogentically informative within *Aesculus*. Of the 1150-bp 5′ coding region of *matK*, 129 (11.2%) sites are variable with a ratio of 1.37:1.00:1.31 (48:35:46) at three different codon positions, similar to that found in Cornales (1.0:1.0:1.3; Xiang et al. 1998). Of the 129 variable sites in the coding region, 45 (3.9%) are potentially phylogenetically informative. The nucleotide composition of *matK* is, in contrast to the ITS sequences, skewed to a high A-T content (66.9% A-T vs. 33.1% G-C), identical to that observed in Cornales (Xiang et al. 1998). The sequence divergence of *matK* between species of *Aesculus* ranges from 0.0019 to 0.0465, with an average of 0.0172 nucleotide substitutions per site. The g_1-value from skewness test (Hillis and Huelsenbeck, 1992) for the *matK*-data matrix is -0.468539, indicating the *matK* data matrix is also phylogenetically structured ($P < 0.01$).

Results of Phylogenetic Analyses

Parsimony analysis of the ITS sequence data found a single minimum- length tree of 432 steps with a CI of 0.588 (excluding uninformative characters) and a RI of 0.619 (Fig. 2). This shortest tree shows that the Chinese and Himalayan species, *A. chinensis, A. wilsonii, A. wangii*, and *A. indica*, form a strongly supported monophyletic group (supported by 13 base substitutions, a bootstrap value of 93%, and a Bremer support value of 5 steps). This clade is the sister of the remainder of the genus (17, 66%, 5). The Japanese species, *A. turbinata*, is the sister of the European species, *A. hippocastanum*, (22, 100%, 10); this group is, in turn, the sister of all the North American species (16, 47%, 2). Within the North American clade, the eastern North American species (excluding *A. parviflora*), *A. glabra, A. flava, A. pavia*, and *A. sylvatica*, form a strongly supported monophyletic group (16, 96%, 6) that is sister to the Mexican species, *A. parryi* (14, 75%, 2). The Californian species, *A. californica*, and the southeastern North American species, *A. parviflora*, are united as sister species, but supported by a relatively low bootstrap value (12, 38%, 2; Fig. 2).

Parsimony analyses of *mat*K sequences found 80 shortest trees, each of 176 steps with a CI of 0.612 (excluding uninformative characters) and a RI of 0.663. All shortest *matK* trees resolve two major clades, a New World clade and an Old World clade that also includes the Californian species. The southeastern North American species, *A. parviflora*, is the sister of the remainder of the clade. The Mexican species, *A. parryi*, is again the sister of the remainder of the eastern North American species (Fig. 3) as in the ITS tree. Relationships within the eastern North American subclade and within the Old World clade vary among shortest trees and are poorly resolved in the strict-consensus tree. Most clades identified by the *matK* trees are supported by low bootstrap and Bremer support values (Fig. 3).

DISCUSSION

Phylogenetic Relationships, Congruence, and Discordance between Gene Trees

A comparison of Figures 1 and 2 reveals that the clades found in the ITS tree correspond to the geographic distri-

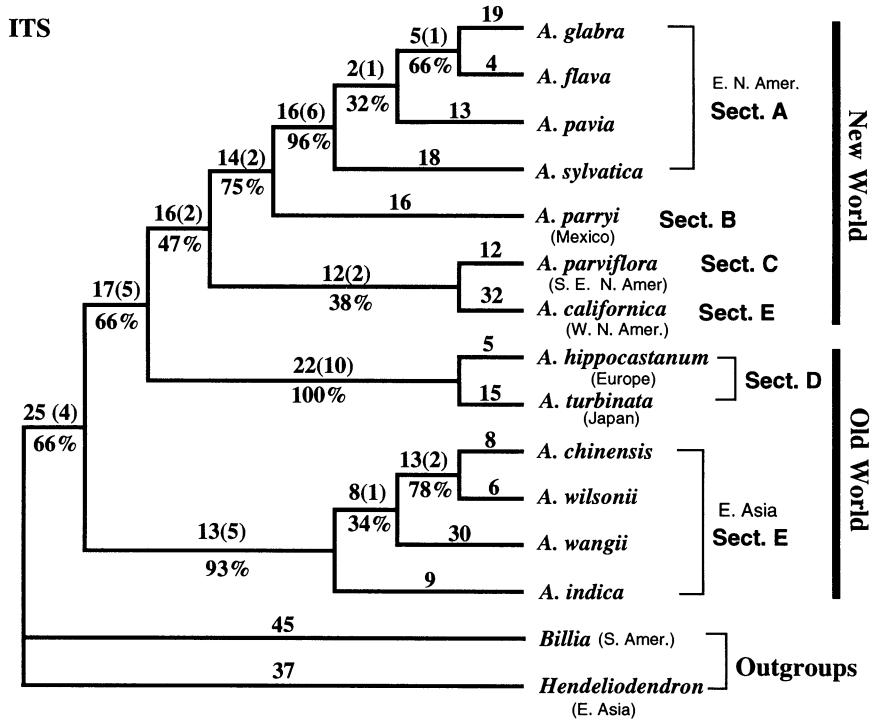

FIG. 2. The single most-parsimonious tree resulting from analysis of internal transcribed spacer (ITS) sequences of *Aesculus* (432 steps, CI = 0.588, RI = 0.610). Numbers above branches indicate number of nucleotide changes supporting the branch. Percentages are bootstrap values (bootstrap values of branches below 50% are not given). Numbers in parentheses are Bremer support. E. N. Amer., eastern North America; S. E. N. Amer., southeastern North America; W. N. Amer., western North America; S. Amer., South America; and E. Asia, eastern Asia. Letters indicate sections of *Aesculus* as in Figure 1.

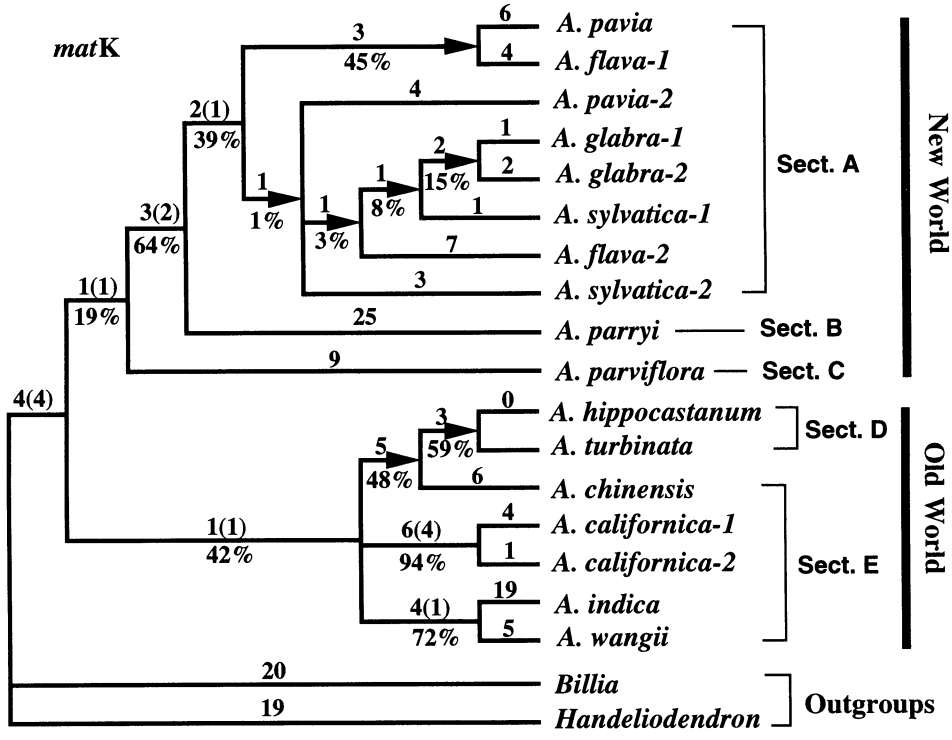

FIG. 3. One of the 80 most-parsimonious trees resulting from analysis of *matK* sequences of *Aesculus* (176 steps, CI = 0.612, RI = 0.663). Numbers above branches indicate number of nucleotide changes supporting the branches. Percentages are bootstrap values above 50%. Numbers in parentheses are Bremer support. Arrows point to nodes that are not present in all most-parsimonious trees. Letters indicate sections of *Aesculus* as in Figure 1.

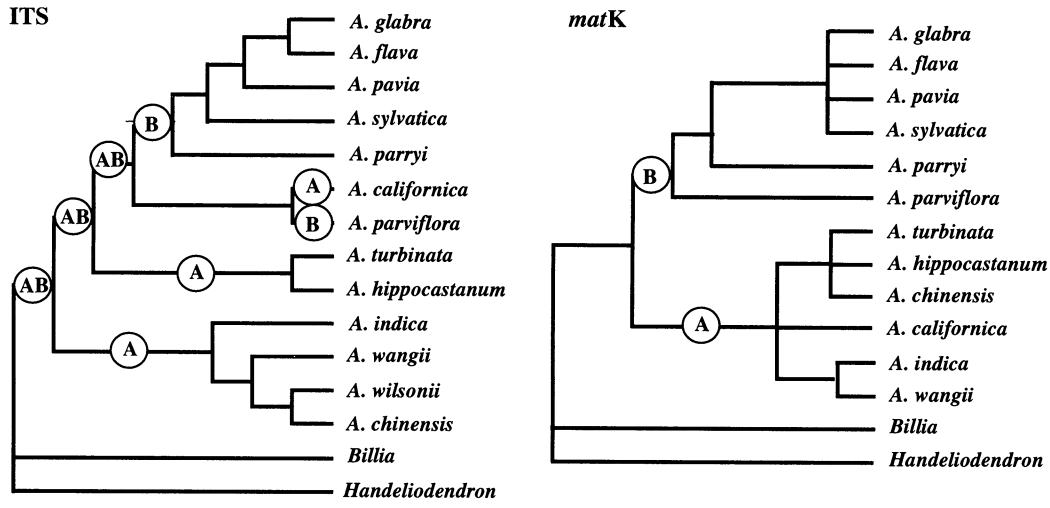

FIG. 4. Comparison between internal transcribed spacer (ITS) and *matK* strict consensus trees, showing the lineage sorting scenario that explains the discordance between the ITS and *matK* trees. "A" and "B" are hypothesized chloroplast haplotypes.

bution of species and the morphological groups recognized by Hardin (1957a, 1960). Each of the five morphological sections of Hardin are resolved as monophyletic groups in the shortest ITS tree except for section *Calothyrsus*, in which *A. californica* is united with the southeastern North American species *A. parviflora* rather than the eastern Asian species of the same section. *Aesculus californica* was placed with the eastern Asian species in section *Calothyrsus* by Hardin mainly because it has resinous bud scales and a smooth fruit surface like the other members of the section.

The eastern North American section *Pavia*, the Eurasian section *Aesculus*, and the eastern Asian section *Calothyrsus* (excluding *A. californica*) are all supported by high bootstrap and Bremer indices (Fig. 2). Relationships among these sections are also well resolved in the ITS tree. The eastern Asian species, section *Calothyrsus*, is the most basally divergent taxon in the tree. Section *Aesculus* is the sister of all the North American species. Within the North American clade, section *Parryana* is sister to the eastern North American species, sect. *Pavia*; the western North American species (*A. californica*) and the Alabama-Georgia endemic species (*A. parviflora*, Sect. *Macrothyrsus*) are united as sisters, but this relationship is not strongly supported by the bootstrap analysis (Fig. 2).

Trees resulting from analyses of *matK* sequence data have lower nodal support than the ITS trees as measured by bootstrap and Bremer analyses (Fig. 3). However, despite the relatively poor nodal support, groups identified by *matK* trees are very similar to those shown on the ITS tree. There is also a strong correlation between the clades found in the *matK* trees and the geographic distributions of species (Fig. 3). The low bootstrap values and Bremer support in the *matK* tree may be an effect of low variability of the gene. The average sequence divergence between species in *mat*K is 8.3 times lower than that in ITS (0.142 in ITS vs. 0.0172 in *matK*). Lack of resolution of relationships within section *Pavia* may be a result of frequent hybridzation among species in the section (Hardin 1957b; DePamphilis and Wyatt 1989, 1990).

Differences between the ITS and *matK* shortest trees involve the placements of *A. hippocastanum, A. turbinata*, and *A. californica* (Fig. 4). The ITS tree shows section *Aesculus* (*A. turbinata* and *A. hippocastanum*) as the sister of all of the North American species and *A. californica* as the sister of the southeastern North American species *A. parviflora* (Figs. 2, 4); the *matK* trees place all three species in the same clade with all the eastern Asian species (Fig. 3, 4). This incongruence between the ITS and *matK* trees is statistically significant according to the results from the Templeton test (1983). If the ITS and *matK* trees represent the nuclear and chloroplast phylogenies of *Aesculus*, respectively, the discordance between the two trees may imply that organellar lineage sorting (extinction of organellar lineages following speciation from a polymorphic ancestor) has occurred in *Aesculus* (Fig. 4). In this hypothesis we assume (1) that the *matK* sequence represents the chloroplast haplotype, and (2) the ancestor of *Aesculus* had a polymorphic chloroplast genome, with "A" and "B" haplotypes. The "B" haplotype was subsequently lost in all the Old World species as well as in *A. californica*, and the "A" haplotype was lost in all the New World species (except in *A. californica*), resulting in the phylogenetic pattern shown on the *matK* tree (Fig. 4). Alternately, it could be hypothesized that the "A" chloroplast haplotype is the ancestral state retained in all the Old World species. The cpDNA of the ancestor giving rise to all New World species diverged into the "B" type. The only problem with the latter hypothesis is *A. californica*, which has the "Asian type" or "A" chloroplast genome (Fig. 4). There are two alternative hypotheses to explain this discrepancy. One is that a reversal occurred in *A. californica*. Only one or two mutations in the *matK* sequence are required to put the species in the New World clade (Fig. 3). The other hypothesis is that *A. californica* obtained the "A" chloroplast genome from an Asian species through hybridization, although this is difficult to visualize given the present allopatric distribution of *A. californica* and the Asian species. We must emphasize that the above hypotheses are based on the ITS and *matK* sequences under the assumption that these data represent the nuclear and chloroplast genomes of *Aesculus*.

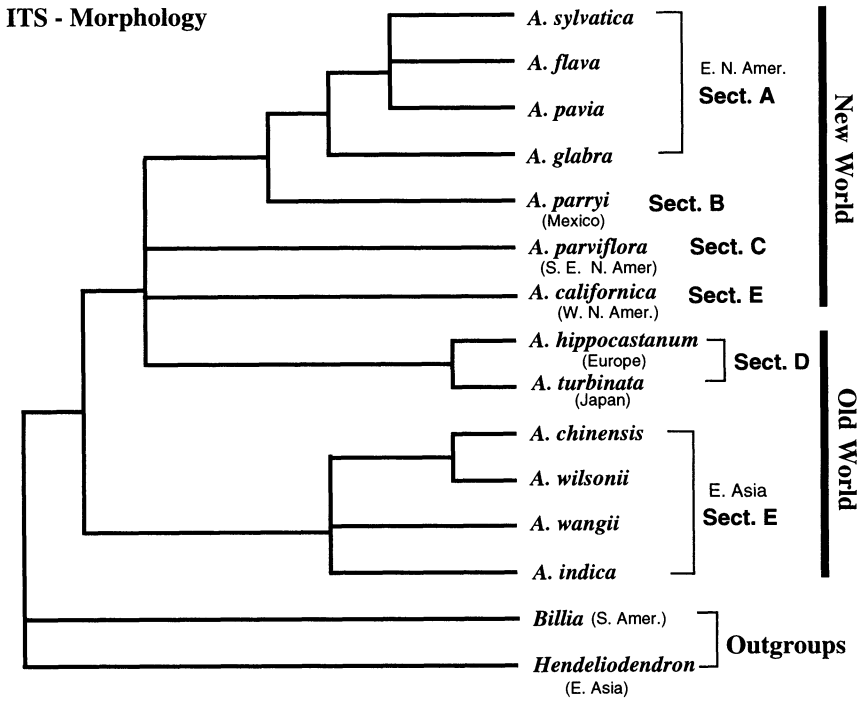

FIG. 5. The strict consensus tree of the five most-parsimonious trees resulting from analysis of the combined internal transcribed spacer (ITS) sequence and morphological dataset using branch-and-bound search (483 steps, CI = 0.574, RI = 0.624). Letters indicate sections of *Aesculus* as in Figure 1.

The phylogenies of *Aesculus* derived from ITS and *matK* sequence data disagree with that derived from analysis of morphological characters using the "Groundplan" method by Hardin (1957a; cf. Figs. 1–3). To gain insights into this incongruence, parsimony analysis of an expanded morphological dataset of Hardin (1957a; with the addition of the Old World species; containing the same 20 characters coded using *Billia* as the outgroup) was conducted. Results of this analysis suggested relationships within *Aesculus* similar to those suggested by the ITS tree, except that *A. californica* appears within the Old World clade. The morphological data and the ITS sequences were then combined for further parsimony analyses, with characters equally weighted. The phylogenetic relationships within the genus suggested by this analysis are nearly identical to those inferred from analysis of the ITS sequences alone, except that there is less resolution among the New World species and Section *Aesculus* (Fig. 5). These results suggest that the discrepancy between the phylogeny derived from the Wagner's Groundplan method and that derived from molecular data may be attributed to the nonparsimonious interpretation of the morphological data of the Groundplan method.

Origin and Biogeography from a Phylogenetic Perspective

The oldest fossils of *Aesculus* are reported from northeastern Asia, Alaska, and the North Atlantic islands in the Paleocene (~ 65 M.Y.B.P.; Budantsev 1992; IOP/PFR on ibs.uel.ac.uk/ibs; M. C. Boulter, pers. comm.). Younger ones are reported from northeastern Asia, British Columbia, Nevada, and north Atlantic islands in the Eocene; from Europe and northwestern United States in the Oligocene; from eastern China, Japan, Washington state, and Europe in the Miocene; from Japan, Europe, and the Pacific Northwest of North America in the Pliocene; and from Europe in the Pleistocene (Axelrod 1966; Leopold 1969; Wolfe 1969; BBI-NGPI 1978; Rouse and Mathews 1979; Budantsev 1992; IOP/PFR on ibs.uel.ac.uk/ibs; M. C. Boulter, pers. comm.). Fossil evidence supports the hypothesis that *Aesculus* evolved at a high latitude of the Northern Hemisphere in the Paleocene or earlier, but it is not informative as to the continent on which the genus originated.

To gain insights into the center of origin and biogeographic history of *Aesculus* in a phylogenetic perspective, we employed character-state-reconstruction and a molecular clock approach. These analyses are based on the assumption that no lineage extinction occurred.

To infer the center of origin of *Aesculus* based on the ITS phylogeny, we conducted character-state mapping using MacClade 3.05 (Swofford and Maddison 1992; Maddison and Maddison 1995). The geographic distributional areas of species were considered as unordered character states and represented by different numbers (Fig. 6). A data matrix containing this single character (the geographic area) was constructed. By choosing the "Tree Display Window" under "Display" on MacClade, a random tree was generated based on this data matrix, which was then modified into the ITS tree manually by dragging branches around. The "area" character states were then optimized (mapped) onto the ITS topology when choosing "Chasing Character" option. Transformations of the character states were then inferred and displayed on the ITS tree (Fig. 6). This character-state mapping exercise on the ITS tree, using either *Billia* or *Handeliod-*

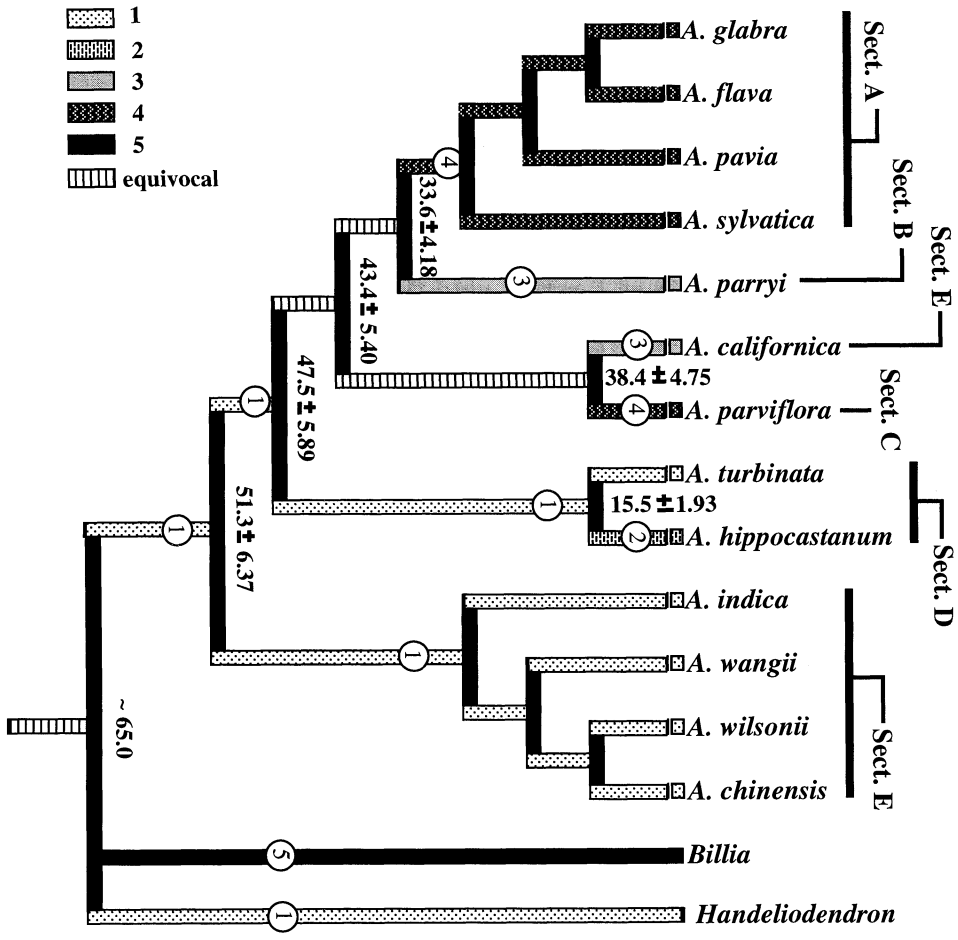

FIG. 6. Area-state reconstruction of *Aesculus* using MacClade 3.0. 1, eastern Asia; 2, Europe; 3, western North America; 4, eastern North America; and 5, Central and South America. Numbers on nodes indicate the time of divergence M.Y.B.P.) between the two lineages on its two sides estimated using internal transcribed spacer (ITS) clocks. Letters indicate sections of *Aesculus* as in Figure 1.

endron as the outgroup, both put the root node (or ancestor) of *Aesculus* in eastern Asia (Fig. 6). This implies that the genus evolved in eastern Asia and subsequently spread into North America. However, the root node of the North American clade was indicated as "equivocal." Thus, whether the genus arrived first in western North America or in eastern North America remains uncertain from the molecular phylogenetic perspective.

Molecular clocks are based on the neutral theory of molecular evolution and assume constant rates of nucleotide substitutions across lineages. There is controversy regarding the neutral theory and many studies have revealed heterogeneous rates of sequence evolution across lineages (Wilson et al. 1990; Gaut et al. 1992, 1993, 1996; Brunsfeld et al. 1994; Xiang et al. 1998; see also Soltis and Soltis 1995), but when the clock can be calibrated with some confidence, it is still useful for estimating divergence times and for testing contrasting biogeographic hypotheses when fossil evidence is not available or inadequate (Parks and Wendel 1990; Cunningham and Collins 1994; Savard et al. 1994; Wen and Jansen 1995). Because the ITS and *matK* phylogenies for *Aesculus* are not congruent, the molecular-clock approach discussed below is applied to only the ITS phylogeny and the ITS sequence data. Assuming a divergence time of 65 M.Y.B.P. (beginning of the Paleocene) between *Aesculus* and its sister based on the fossil evidence, a rate of $1.722 \pm 0.21 \times 10^{-9}$ per site per year is obtained for ITS evolution, based on a 0.2239 ± 0.0276 sequence divergence value between *Aesculus* and *Billia*. If a molecular clock exists for the ITS of *Aesculus*, this rate can be used to estimate the divergence times for geographically disjunct lineages and thus to gain insights into the origins of disjunctions in the genus. The divergence time between two lineages can be calculated by dividing the sequence divergence value between the two lineages by twice the rate of divergence (Li 1997).

To test the molecular-clock hypothesis in *Aesculus*, we performed relative rate tests for ITS following the method of Li (1997; pp. 216–221). Results show that rates of ITS evolution do not significantly differ among lineages within *Aesculus* and the molecular clock hypothesis cannot be rejected (Table 2). Using the ITS molecular clock, divergence times within the genus are estimated at 51.3 ± 6.37 M.Y.B.P. or earlier (beginning of the Eocene) between the eastern Asian section (*Carothyrsus*) and the remainder of the genus; at 47.5 ± 5.89 M.Y.B.P. or earlier (middle Eocene) between the North American clade and its sister Eurasian species, secttion *Aesculus*;

TABLE 2. Results of relative rate tests in internal transcribed spacers of *Aesculus*. The rate of base substitutions in each lineage (K) was estimated as the average number of total base substitutions per site between each species constituting the lineage and *Handeliodendron* using the Jukes-Cantor distance method. A difference in K between two lineages greater than twice the standard error is considered significant at the 0.05 level (Li, 1997). 1, section *Pavia*; 2, *A. parviflora–A. californica*; 3, section *Aesculus*; 4, section *Calothyrsus*.

$K1 = 0.2056 \pm 0.02544$	$K1 - K2 = 0.02800 \pm 0.02425$
$K2 = 0.1776 \pm 0.02305$	$K1 - K3 = 0.00420 \pm 0.02497$
$K3 = 0.2014 \pm 0.02450$	$K1 - K4 = 0.02250 \pm 0.02462$
$K4 = 0.1831 \pm 0.02380$	$K2 - K3 = 0.02375 \pm 0.02380$
	$K2 - K4 = 0.00550 \pm 0.02343$
	$K3 - K4 = 0.01825 \pm 0.02415$

at 43.4 ± 5.40 M.Y.B.P. or earlier (late Eocene) between section *Pavia–A. parryi* clade and the *A. californica–A. parviflora* clade; at 33.6 ± 4.18 and 38.4 ± 4.75 M.Y.B.P. or earlier (beginning of Oligocene) between the eastern North American section *Pavia* and *A. parryi* and between *A. californica* and *A. parviflora*, respectively; at 15.5 ± 1.93 M.Y.B.P. or earlier (middle Miocene) between the Japanese species *A. turbinata* and the southeastern European species *A. hippocastanum* (see Fig. 6). Although these values may be subject to some uncertainties, they are well correlated with the paleontological history of the Northern Hemisphere (see below).

A brief examination of the paleontological history of the Northern Hemisphere since the late Cretaceous helps to understand the biogeographic history of *Aesculus*. Paleontological evidence indicates that during the late Cretaceous, the Midcontinent Seaway of North America and the Asian Turgai Strait divided the Laurasian landmass into two distinctive provinces, the "Normapolles Province" of eastern North America and Europe and the "Aquillapollenites Province" of the eastern Asia and western North America (Raven and Axelrod 1974; Wolfe 1975; Tiffney 1985). The Midcontinent Seaway of North America retreated during the end of Cretaceous and the Paleocene and disappeared by 60 M.Y.B.P., and the Asian Turgai Strait closed during the Oligocene (Tiffney 1985). The regression of the Midcontinent Seaway occurred concomitantly with the uplift of the Rocky Mountains.

According to fossil and palynological evidence, a wide range of extant flowering plants first appeared during the transition from the Cretaceous to Tertiary. The newly evolved taxa spread rapidly over the Northern Hemisphere during the temperate global climates of the early Tertiary, resulting in the formation of a hemispheric "boreotropical flora" (Wolfe 1975, 1977). The early Eocene North Atlantic land bridge and the Bering land bridge (connecting Alaska and Siberia in the early Eocene and late Eocene) provided intercontinental land connections for the spreading of the boreotropical flora between Eurasia and North America (Tiffney 1985). Within the North American continent, spreading of the boreotropical flora across the central portion of the continent was possible until the early Oligocene, and within Eurasia the connected islands in the Tethys Seaway (a marine system that extended from southeastern Asia to westernmost Europe that was disrupted about 60 M.Y.B.P.; Brown and Gibson 1983) served as a main route for the spreading of the boreotropical flora (Tiffney 1985). Most deciduous taxa of the flora (now the boreal temperate taxa) were formed in vast areas of Angaro-Beringia and went eastward through high-latitude marginal sites of North America, Greenland, and Spitsbergen at the end of Cretaceous–early Paleocene (Budantsev 1992).

A climatic deterioration (cooling) started at the Eocene-Oligocene boundary and extended into the Oligocene, resulting in the contraction of the boreotropical flora (Raven and Axelrod 1974; Wolfe 1975; Tiffney 1985). Further subsequent climatic fluctuation in the later Oligocene and Miocene shaped the boreotropical flora into a "Mixed Mesophytic Forest" (Tiffney 1985). Extreme cooling of the climate occurred in the later Tertiary and Quaternary, together with the effect of the uplift of the Rocky Mountains, eliminated elements of this mixed mesophytic forest from many areas. The survivors of this forest are now found in a few "refugia" including eastern Asia, southeastern Europe, eastern and western North America, and western Asia (Graham 1972; Leopold and McGinitie 1972; Wolfe 1975; Tiffney 1985).

Using the molecular phylogenetic reconstructions, the estimated divergence times between lineages, fossil evidence, and the current geographic distributions of species (see Fig. 6), we suggest the following biogeographic hypotheses for *Aesculus*. The genus evolved during the transition from Cretaceous to Tertiary at a high latitude of eastern Asia, as an element of the boreotropical flora. *Aesculus* then diverged into two lineages in the late Paleocene or at the beginning of the Eocene (~51 M.Y.B.P.; see Fig. 6). During the climatic cooling period in the late Tertiary and Quaternary, one lineage moved southward into the Himalayan region and survived in areas south of the Qinglin Mountain of China and subsequently diverged into several species constituting section *Calothyrsus* (excluding *A. californica*). The other lineage expanded its range both eastward and westward. One part of this lineage perhaps spread eastward circumpolarly into North America and Europe. This lineage was isolated into two parts, the Eurasian part and the North American part, during the middle Eocene (~48 M.Y.B.P.) when the Bering land bridge and the North Atlantic bridge were both disrupted (see Tiffney 1985). The continuous distribution of the Eurasian part was first interrupted around 15.5 million yr BP (or in the middle Miocene) due to climatic cooling. This part of the lineage was eventually eliminated during the late Tertiary and Quaternary from most areas of the Eurasian continent except those in southeastern Europe and Japan, which survived and diverged into two species (*A. hippocastanum* and *A. turbinata*).

The North American part of this lineage diverged into two groups (*A. pavia–A. flava--A. glabra–A. sylvatica–A. parryi* and *A. californica–A. parviflora*) in the middle Eocene (~43 M.Y.B.P.) and expanded their range southward. The eastern and western parts of these two groups were isolated in the early Oligocene (34–38 M.Y.B.P.) when the drought caused by the eastern rain shadow cast by the Rocky Mountains made the central portion of North America unable to support the boreotropical taxa. This isolation resulted in the divergence of the eastern and western species in each group (see Fig. 6; Leopold and MacGinitie 1972; Tiffney 1985).

In conclusion, the molecular data from ITS and *matK* identify groups in *Aesculus* that correspond to those recognized by Hardin (1957a) using morphological characters with the exception of section *Calothyrsus*, in which *A. californica* may be misplaced. The phylogenies inferred from ITS and *mat*K sequence data are discordant with the phylogeny derived from the Wagner Groundplan method (Hardin 1957a). The phylogeny inferred from ITS sequence data does not support the biogeographic hypotheses proposed by either Hardin (1957a) or Raven and Axelrod (1978) that the genus had an origin in tropical America or North America.

The molecular data from ITS sequences and fossil evidence suggest a north temperate eastern Asian origin for *Aesculus* during the transition from Cretaceous to the Tertiary, when it was an element of the boreotropical flora. Climatic and geological changes in the Tertiary and Quaternary resulted in the disjunct distribution of *Aesculus* observed today. The disjunction between Eurasian and North American species occurred first in the genus, and the disjunction between the eastern and western North American species occurred before the separation of the European and Japanese species. The estimate of time of origin for the genus based on molecular data and fossil evidence is close to the hypothesis of Hardin (1957a) that *Aesculus* evolved in the early Tertiary or earlier.

Acknowledgments

We thank D. E. Soltis and P. S. Soltis for providing some *matK* primers for initial experiments and facilities for isolation of several DNA samples, S. Boyd, Y.-L. Cao, S.-K. Chen, B. Hammel, M. T. Hardig, L. Hufford, Y. Tang, M. Wall, F.-N. Wei, J. Wen, G-H. Zhu, and H.-Q. Wen for help in collecting leaf materials, and K. N. Gandhi and D. E. Boufford for verifying the validity of *Aesculus flava* Solander. This work was supported by a postdoctoral fellowship from the Graduate School at The Ohio State University and a National Science Foundation grant DEB-9500499.

Literature Cited

Axelrod, D. I. 1966. The Eocene Copper Basin flora of northeastern Nevada. Univ. Calif. Publ. Geol. Sci. 59:1–124.

Baldwin, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Phyl. Evol. 1:3–16.

BBI-NGPI (Beijing Botanical Institute and Nanjing Geological and Paleobiological Institute). 1978. Plant fossils of China. Vol. 3. Cenozoic plants of China. Science Press, Beijing. In Chinese.

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

Brown, J. H., and A. C. Gibson. 1983. Biogeography. C. V. Mosby Company, St. Louis, MO.

Brunsfeld, S. J., P. S. Soltis, D. E. Soltis, P. A. Gadek, C. J. Quinn, D. D. Strenge, and T. A. Ranker. 1994. Phylogenetic relationships among genera of the conifer families Taxodiaceae and Cupressaceae: evidence from *rbcL* sequences. Syst. Bot. 19: 53–262.

Budantsev, L. Y. 1992. Early stages of formation and dispersal of the temperate flora in the Boreal region. Bot. Rev. 58:1–48.

Cullings, K. W. 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1: 233–240.

Cunningham, C. W., and T. M. Collins. 1994. Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. Pp. 405–433 *in* B. Schierwater, B. Streit, G. P. Wangner and R. DeSalle, eds. Molecular ecology and evolution: applications and approaches. Burkhauser Verlag, Basel, Switzerland.

Depamphilis, C. W., and R. Wyatt. 1989. Hybridization and introgression in buckeyes (*Aesculus*: Hippocastanaceae): a review of the evidence and a hypothesis to explain long-distance gene flow. Syst. Bot. 14:593–611.

———. 1990. Electrophoretic confirmation of interspecific hybridization in *Aesculus* (Hippocastanaceae) and the genetic structure of a broad hybrid zone. Evolution 44:1295–1317.

Eernisse, D. J., and A. G. Kluge. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. Mol. Biol. Evol. 10:1170–1195.

Fang, W.-P. 1981. Hippocastanaceae. Pp. 274–289 *in* W.-P. Fang, ed. Flora Reipublicae Popularis Sinicae. Vol. 46. Science Press, Beijing.

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

Gaut, B. S., S. V. Muse, W. D. Clark, and M. T. Clegg. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. J. Mol. Evol. 35:292–303.

Gaut, B. S., S. V. Muse, and M. T. Clegg. 1993. Relative rates of nucleotide substitution in the chloroplast genome. Mol. Phyl. Evol. 2:89–96.

Gaut, B. S., B. Morton, B. C. McCaig, and M. T. Clegg. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proc. Nat. Acad. Sci. USA 93:10274–10279.

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

Hardin, J. W. 1957a. A revision of the American Hippocastanaceae. Brittonia 9:145–171, 173–195.

———. 1957b. Studies in the Hippocastanaceae. IV. Hybridization in *Aesculus*. Rhodora 59:185–203.

———. 1960. Studies in the Hippocastanaceae, V. Species of the Old World. Brittonia 12:26–38.

Higgins, D. M., A. J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: improved software for multiple sequence alignment. Comp. Appl. Biosci. 8:189–191.

Hillis, D. M., and J. P. Huelsenbeck. 1992. Signal, noise, and reliability in molecular phylogenetic analysis. J. Hered. 83:189–195.

Johnson, L. A., and D. E. Soltis. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst. Bot. 19:143–156.

Judd, W. S., R. W. Sanders, and M. J. Donoghue. 1994. Angiosperm family pairs: preliminary phylogenetic analyses. Harv. Pap. Bot. 1:1–15.

Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: molecular evolutionary genetic analysis. Vers. 1.0. Pennsylvania State Univ., University Park, PA.

Larson, A. 1994. The comparison of morphological and molecular data in phylogenetic systematics. Pp. 371–390 *in* B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds. Molecular ecology and evolution: approaches and applications. Birkhauser Verlag, Basel, Switzerland.

Leopold, E. B. 1969. Late Cenozoic palynology. Pp. 377–438 *in* R. H. Tschudy and R. A. Scott, eds. Aspects of palynology. Wiley-Interscience, New York.

Leopold, E. B., and H. D. MacGinitie. 1972. Development and affinities of Tertiary floras in the Rocky Mountains. Pp. 147–200 *in* A. Graham, ed. Floristics and paleofloristics of Asia and eastern North America. Elsevier, Amsterdam.

Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.

Maddison, W. P., and D. R. Maddison. 1995. MacClade 3.05. Interactive analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.

Morgan, D. R., and D. E. Soltis. 1993. Phylogenetic relationships

among members of Saxifragaceae sensu lato based on *rbcL* sequence data. Ann. Mo. Bot. Gard. 80:631–660.

PARKS, C. R., AND J. F. WENDEL. 1990. Molecular divergence between Asian and North American species of *Liriodendron* (Magnoliaceae) with implications for interpretation of fossil floras. Am. J. Bot. 77:1243–1256.

RAVEN, P. H., AND D. I. AXELROD. 1974. Angiosperm biogeography and past continental movements. Ann. Mo. Bot. Gard. 61:539–673.

———. 1978. Origin and relationships of the California flora. Univ. Calif. Publ. Bot. 72:1–134.

ROUSE, G. E., AND W. H. MATHEWS. 1979. Tertiary geology and palynology of the Quesnel area, British Columbia. Can. Petrol. Geol. 27:418–445.

SANG, T., D. J. CRAWFORD, AND T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc. Nat. Acad. Sci. USA 92:6813–6817.

———. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84: 1120–1136.

SAVARD, L., P. LI, S. H. STRAUSS, M. W. CHASE, M. MICHAUD, AND J. BOUSQUET. 1994. Chloroplast and nuclear gene sequences indicate Late Pennsylvatican time for the last common ancestor of extant seed plants. Proc. Nat. Acad. Sci. USA 91:5163–5167.

SOLTIS, D. E., AND R. KUZOFF. 1995. Discordance between nuclear and chloroplast phylogenies in the Heuchera Group (Saxifragaceae). Evolution 49:727–742.

SOLTIS, P. S., AND D. E. SOLTIS. 1995. Plant molecular systematics: inferences of phylogeny and evolutionary processes. Evol. Biol. 28:139–194.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Vers. 3.1.1. Illinois Natural History Survey, Champaign, IL.

SWOFFORD, D. L., AND W. P. MADDISON. 1992. Parsimony, character-state reconstructions, and evolutionary inferences. Pp. 186–223 *in* R. L. Mayden, ed. Systematics, historical ecology, and North American freshwater fishes. Stanford Univ. Press, Stanford, CA.

TEMPLETON, A. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

TIFFNEY, B. H. 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J. Arn. Arb. 66:243–273.

WEN, J., AND R. K. JANSEN. 1995. Morphological and molecular comparisons of *Campsis grandiflora* and *C. radicans* (Biognoniaceae), an eastern Asian and eastern North American vicariad species pair. Plant Syst. Evol. 196:173–183.

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

WILSON, M. A., B. GAUT, AND M. T. CLEGG. 1990. Chloroplast DNA evolves slowly in the palm family. Mol. Biol. Evol. 7: 303–314.

WOLFE, J. A. 1969. Neogene floristic and vegetational history of the Pacific Northwest. Madroño 20:83–110.

———. 1975. Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. Ann. Mo. Bot. Gard. 62:264–279.

———. 1977. Paleogene floras from the Gulf of Alaska region. U.S. Geol. Surv. Prof. Pap. 997:1–108.

XIANG, Q.-Y., D. E. SOLTIS, D. R. MORGAN, AND P. S. SOLTIS. 1993. Phylogenetic relationships of *Cornus* L. sensu lato and putative relatives inferred from *rbcL* sequence data. Ann. Mo. Bot. Gard. 80:723–734.

XIANG, Q.-Y., D. E. SOLTIS, AND P. S. SOLTIS. 1998. Phylogenetic relationships of Cornaceae and close relatives inferred from *matK* and *rbcL* sequences. Am. J. Bot. 85:285–297.

Corresponding Editor: R. DeSalle