Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 49–59
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Molecular evidence for an Asian origin and a unique westward migration of species in the genus *Castanea* via Europe to North America ☆

Ping Lang [a,1], Fenny Dane [a,*], Thomas L. Kubisiak [b], Hongwen Huang [c]

[a] *Department of Horticulture, Auburn University, Auburn, AL, USA*
[b] *USDA-Forest Service, Southern Institute of Forest Genetics, Saucier, MS, USA*
[c] *Chinese Academy of Science, Wuhan Institute of Botany, Wuhan City, Hubei, PR China*

Received 25 February 2006; revised 30 June 2006; accepted 31 July 2006
Available online 26 August 2006

Abstract

The genus *Castanea* (Fagaceae) is widely distributed in the deciduous forests of the Northern Hemisphere. The striking similarity between the floras of eastern Asia and those of eastern North America and the difference in chestnut blight resistance among species has been of interest to botanists for a century. To infer the biogeographical history of the genus, the phylogeny of *Castanea* was estimated using DNA sequence data from different regions of the chloroplast genome. Sequencing results support the genus *Castanea* as a monophyletic group with *Castanea crenata* as basal. The three Chinese species form a strongly supported sister clade to the North American and European clade. A unique westward expansion of extant *Castanea* species is hypothesized with *Castanea* originating in eastern Asia, an initial diversification within Asia during the Eocene followed by intercontinental dispersion and divergence between the Chinese and the European/North American species during the middle Eocene and a split between the European and the North American species in the late Eocene. The differentiation within North America and China might have occurred in early or late Miocene. The North America species are supported as a clade with *C. pumila* var. *ozarkensis*, the Ozark chinkapin, as the basal lineage, sister to the group comprising *C. pumila* var. *pumila*, the Allegheny chinkapin, and *Castanea dentata*, the American chestnut. Morphological evolution of one nut per bur in the genus may have occurred independently on two continents.

Keywords: Biogeography; *Castanea*; *ndh*F; *ycf*6-*psb*M; *ycf*9-*trn*GM; *rpl*16

1. Introduction

The genus *Castanea* Miller is in the same angiosperm family of Fagaceae as oak (*Quercus*) and beech (*Fagus*) trees. Fagaceae are widely distributed in the deciduous forests of the temperate regions of the Northern Hemisphere. The distribution was continuous during the Oligocene (38–25 mybp), leading to evolutionary scenarios including a hypothesized center of origin (Manos and Stanford, 2001). Their modern distribution reflects disjunctions and different biogeographic histories. Chestnut species in the genus *Castanea*, characterized by three nuts per bur (section *Eucastanon*), can be found in China (*C. mollissima* BL. and the precocious *C. seguinii* Dode.), in Japan (*C. crenata* Sieb. & Zucc.), in North America (*C. dentata* (Marsh.) Brokh), and in Europe (*C. sativa* Mill.) (Jaynes, 1975) (Fig. 1). They have varying economic importance as nut tree crops and consequently have been influenced by human colonization.

☆ The sequences reported in the paper have been deposited in the Genbank database (Accession Nos. AY586319-AY586335 for *trn*T-L spacer, AY586285-AY586301 for *trn*L intron, AY586302-AY586318 for *trn*L-F spacer. AY526885-AY526902 for *rpl*16 intron, AY586341-AY586361 for *ndh*F, AY525345-AY525366 for *ycf*6-*psb*M, AY496081-AY496091 for *ycf*9–*trn*GM, and DQ386688-DQ386695).

* Corresponding author. Fax: +1 334 844 3131.
E-mail address: danefen@auburn.edu (F. Dane).
[1] Present address: Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY, USA.

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.07.022

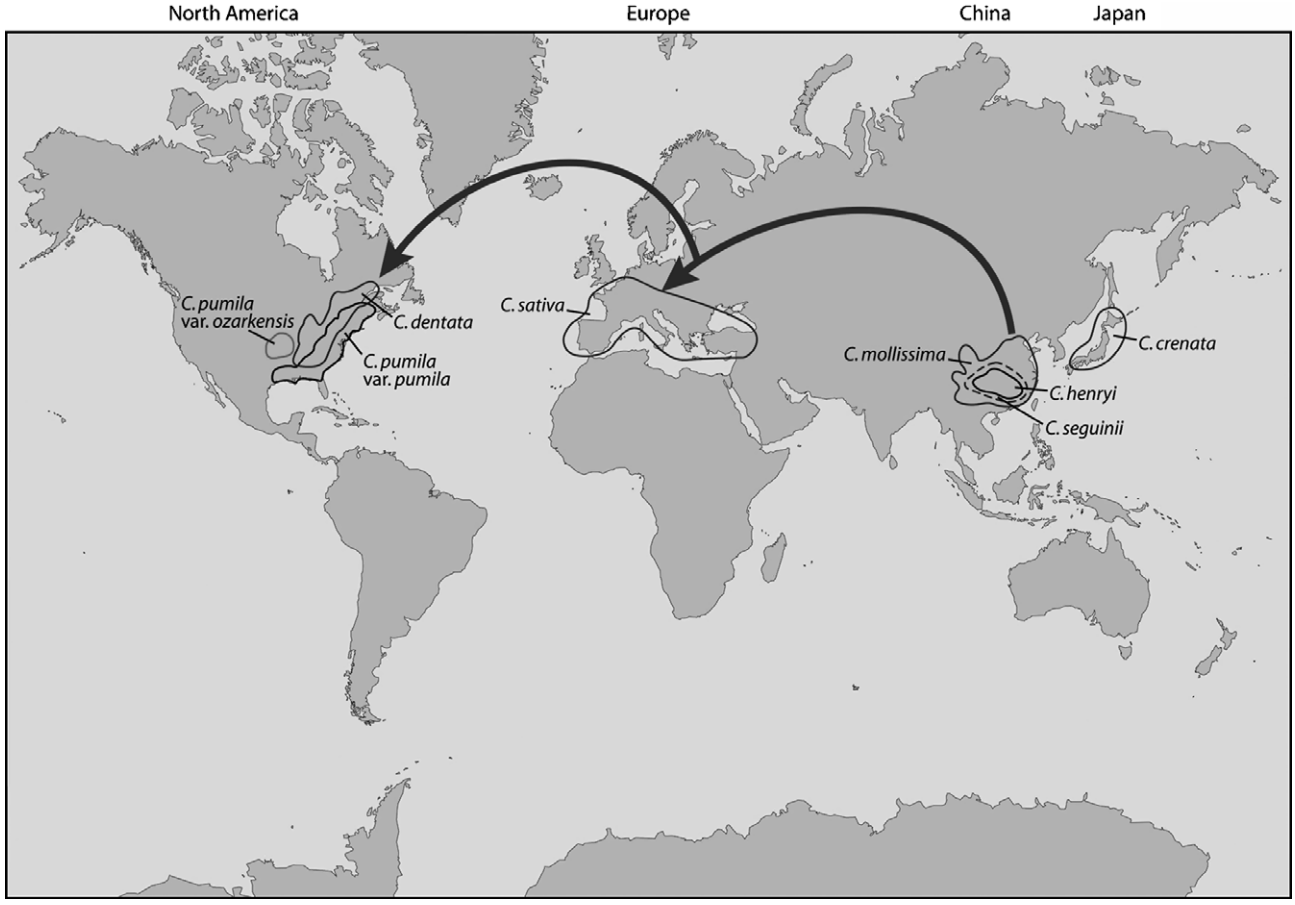

Fig. 1. Distribution of the genus *Castanea* and a hypothesized westward migration route indicated by arrows based on cpDNA sequence information from extant species.

Chinkapins with one nut per bur are distributed in more limited areas of China [*C. henryi* (Skan) Rehder & Wilson, Section *Hypocastanon*] and the southeastern United States (*C. pumila* Mill., Section *Balanocastanon*) (Johnson, 1988). *C. pumila* has two recognized varieties, the Allegheny chinkapin (var. *pumila*) which is distributed throughout much of the southeastern United States and the Ozark chinkapin (var. *ozarkensis*) which has a limited distribution in the Ozark mountains (Johnson, 1988).

Castanea is a Tertiary genus with a disjunct distribution pattern in eastern Asia and eastern North America (NA) and species with floristic similarity but differences in resistance to *Cryphonectria parasitica* (Murrill) Barr, causal agent of chestnut blight. The geographic distributions of *Castanea* in China and NA have similar patterns in that geographically widespread species (*C. mollissima* in China and *C. dentata* in NA) overlap restricted species (*C. seguinii* and *C. henryi* in China and *C. pumila* in NA). The NA species are known to be highly blight susceptible, while the Asian species are relatively blight resistant (Jaynes, 1975). Since the introduction of chestnut blight into eastern NA at the beginning of the 20th century, the American chestnut has been reduced from the status of an important timber- and nut-producing tree to that of an understory shrub (Graves, 1950; Huang et al., 1996).

Although the accuracy of fossil dating needs to be considered, it is reasonable to assume that *Castanea* trees were more widespread in Tertiary times than they are today. Macrofossils of *Castanea* have been reported in the Rocky Mountain region of North America from the Paleocene to the Early Eocene (http://www.nhm.uio.no/palmus/galleri/montre/english/a31696.htm; Tidwell, 1998; Graham, 1999) and throughout the Tertiary in European floras (International Organization of Palaeobotany, 1997; www.paleodb.org). In Asia, fossils of *Castanea* have been found in Japan (Oligocene), North Korea (Miocene), northeastern, eastern and southwestern China (late Eocene, Miocene, and Pliocene), and western Asia (Oligocene) (Zhou, 1999). In eastern North America, Eocene leaves and cupules with strong affinities to modern species suggest long-term presence (Crepet and Daighlian, 1980; Graham, 1999; Manos and Stanford, 2001). Although many fossil records are available, attempts to reconstruct phylogeographic history based on fossil distribution patterns may be misleading, because distinguishing characters such as locule and style number have not been found and pollen and leaf characters are too general to be of diagnostic value (Graham, 1999; Manos and Stanford, 2001). The origin of the genus *Castanea* needs to be considered carefully, since little is known about the evolutionary and genetic relationship of

these morphologically similar but intercontinentally isolated species. A clear phylogenetic framework is needed to elucidate complex biogeographic patterns. It was hypothesized that the genus *Castanea* originated in Asia, the westward extension of the genus gave rise to the European chestnut, and the eastward migration gave rise to the NA species (Jaynes, 1975). However, this hypothesis has not been adequately tested. Systematic studies of the genus *Castanea* on a worldwide basis have so far been limited or inconclusive (Manos and Stanford, 2001; Stanford, 1998; Manos et al., 2001). Population genetic studies of *Castanea* species have been conducted using allozymes (Dane et al., 1999, 2003; Villani et al., 1991; Huang et al., 1994; Lang and Huang, 1999), random amplified polymorphic DNA (Huang et al., 1998), and PCR-restriction fragment length polymorphism (RFLP) analysis of chloroplast (cp) and mitochondrial genomes of *C. sativa* (Fineschi et al., 2000). The only comprehensive molecular study suggests a complicated history for the genus (Stanford, 1998). Unfortunately, relationships within the genus *Castanea* were poorly resolved due to low levels of informative characters in the *mat*K region and large amounts of missing data. Therefore, further investigations to examine the phylogenetic and biogeographic relationships within the genus *Castanea* were conducted.

Chloroplast DNA (cpDNA) is well suited to evolutionary and phylogenetic study (Clegg and Zurawaki, 1992; Olmstead and Palmer, 1994). The maternal inheritance of cpDNA in most angiosperms allows for the direct examination of seed-mediated dispersal and gene flow and hence can be used for inferring colonization routes (Dumolin-Lapegue et al., 1997). Noncoding regions of cpDNA (introns and intergenic spacers) were found to be more variable than coding regions of cpDNA due to the presence of fewer selective constraints (Taberlet et al., 1991). Sequencing analysis of the variable *trn*T-L-F region supported *C. crenata* as most basal and placed the Chinese species in a monophyletic clade with the North American and European species as a sister group (Lang et al., 2006). Sequences of other variable regions such as *rpl*16 intron, *ndh*F, *ycf*6-*psb*M, and *ycf*9-*trn*GM were chosen to further study the sister group relationships between North American and European species. The specific objectives of this study were to further elucidate (1) the phylogenetic relationships among species of the genus and (2) the origin and formation of the present disjunct distribution pattern of *Castanea* in eastern Asia and eastern North America.

2. Material and methods

2.1. Plant material, DNA extraction, PCR, and nucleotide sequencing

Sampling information of 18 taxa representing all seven species of *Castanea* and outgroups *Quercus falcata* and *Fagus grandifolia* are shown in Table 1. Genomic DNAs were extracted from fresh plant material using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). In addition to published cpDNA sequences (Lang et al., 2006), the variable regions *rpl*16 intron, *ycf*6-*psb*M, and *ycf*9-*trn*GM in the large single copy (LSC) and *ndh*F 3′ coding and 3′ flanking region in the small single copy (SSC) of the cp genome were chosen for the analysis. Universal primers were used to amplify the corresponding cp regions of interest (Lang, 2004).

Double-stranded DNA amplification was performed in a 50-μl volume containing the respective 1× PCR buffer (Invitrogen Life Technologies, Carlsbad, CA), 2 mmol/L $MgCl_2$, 200 μmol/L each dNTP, 0.2 μmol/L each primer, 2

Table 1
Investigated taxa

Taxon	Sample ID	DNA source/voucher
C. dentata	D2	CA-44: Ontario, Canada
C. dentata	D12	BR-34: Asheville, North Carolina
C. pumila var. *pumila*	P4	FLE-10: Eglin AirForce Base, Florida
C. pumila var. *pumila*	P6	HH R5T2: Lafayette County, Florida
C. pumila var. *pumila*	Pa	VA3: Iron Mountain, Virginia
C. pumila var. *pumila*	Pb	VB1: Iron Mountain, Virginia
C. pumila var. *ozarkensis*	P7	CC R-gammaT3: Russelville, Arkansas, planted 1936
C. pumila var. *ozarkensis*	P9	CC R-etaT3: G. Miller, Carrollton, Ohio, planted 2003
C. sativa	Sa1	HH R3T2: from the Cavcas Biosphere Reserve, planted 1994
C. sativa	Sa7	24°E 10′ and 46°N 80′, Baia Mare, northern Romania, 2001
C. mollissima	M1	CH-MAH: Chinese Mahogany
C. mollissima	M4	WL R37T7: USDA #104061, from China, planted 1935
C. seguinii	S3	SL R3T8: USDA #70317, seed 1926, seedling planted 1930
C. seguinii	S6	HH R4T2: cross of SL R8T4 (female) × SL R2T16 (male)
C. henryi	H1	Hubei Province, P.R. China, 1997
C. henryi	H3	Rock R3T9: from Liu Liu, Nanjing Botanical Garden
C. crenata	C1	WL R34T6: USDA#104016
C. crenata	C7	92-JH-2a-2: collected on 9-25-92, Bare of Mt. Tougo, Kuki-cho
Quercus falcata		Saucier, Mississippi, 2002
Fagus grandifolia		Saucier, Mississippi, 2002

U *Taq* polymerase (Invitrogen), and 2.5 μl template DNA (25 ng/μl). Following an activation step of three cycles of 1 min at 94^0C and 30 s at 72 °C, the PCR mixture underwent 34 cycles of 1 min at 94 °C, 50 s at 54–56 °C, 2 min at 72 °C, with a final extension of 7 min at 72 °C. PCR products were purified using Qiaquick PCR purification kit (Qiagen) and sequenced by Auburn Genomics and Sequencing Lab with the ABI3100 sequencer (Applied Biosystems Inc., Foster City, CA). To verify complementary stands, both forward and reverse PCR primers were used to generate the entire sequences.

2.2. Sequence alignment and phylogenetic analyses

Multiple alignment of the sequences was obtained using the AlignX program implemented in the Vector NTI software (Informax, Invitrogen, Frederick, MD) and adjusted manually. Gaps were introduced in the alignment to optimize positional homology. Single-base indels were cross-checked to the original chromatograms to verify that they were not sequencing artifacts missed during base calling. Indels that were potentially parsimonious (e.g., shared by two or more taxa) were scored and added to the end of the data sets as present (1) or absent (0) type characters. Gaps with overlaps were considered nested and treated as a single multi state character according to Simmons and Ochotreana (2000). Areas of ambiguous alignment or with poly-n strings were excluded from all analyses. Based on earlier results obtained for the *trn*L intron and *trn*L-F spacer regions of *F. grandifolia* (Lang et al., 2006), only *Q. falcata* was used as outgroup for combined phylogenetic analyses (Table 1).

Phylogenies were reconstructed by maximum parsimony (MP) and maximum likelihood (ML) implemented in PAUP 4.0b10 (Swofford, 2003) and Bayesian package MrBayes 3.1 (Huelsenbeck and Ronquist, 2001). To explore the possibility of multiple islands of most parsimonious trees, searches were conducted over 1000 random-taxon-addition replicates with tree bisection-reconnection (TBR) branch swapping, with the save multiple trees (MulTrees) option selected and with all characters and states weighted equally and unordered. A 50% majority-rule consensus tree was computed. The robustness of nodes was inferred by bootstrap analysis using 1000 replicates (Felsenstein, 1985). Before performing ML analyses, the best-fitting model of sequence evolution and the corresponding values for the rate matrix, shape of the gamma distribution, and the proportion of invariant sites were estimated for all data sets using the program Modeltest 3.7 (Posada and Crandall, 1998). Bootstrap support values were obtained with 100 replicates for ML analyses. In addition, Bayesian phylogenetic reconstruction was performed using MrBayes with 10,000,000 generations, a sampling frequency of every 100 generations, GTR+I+G, four Markov chains, random starting trees, and a burn-in of 2,000,000 generations (Huelsenbeck and Ronquist, 2001). Multiple Bayesian analyses were performed to ensure convergence. To determine data set combinability, the structure of the data sets was assessed with a series of incongruence length difference (ILD) tests (implemented in PAUP* as the partition-homogeneity test) (Farris et al., 1994). Individual and combined data topological congruence was evaluated by the conservative Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 2001) as implemented in the CONSEL package (http://www.is.titech.ac.jp/~shimo/prog/consel/; Shimodaira and Hasegawa, 2001).

2.3. Biogeographic analyses

To infer the biogeographic history of the genus of *Castanea*, dispersal–vicariance analysis (DIVA) was used (Ronquist, 1996, 1997). DIVA is a pattern-based method used for reconstructing the distribution history of a group of organisms from the distribution areas of extant species and their phylogeny. This method considers vicariance, dispersal, and extinction explicitly (Ronquist, 1996, 1997; Xiang and Soltis, 2001). DIVA had been commonly used for reconstructing ancestral areas of angiosperms in the Northern Hemisphere (Xiang and Soltis, 2001; Xiang et al., 2005, 2006). The ML tree derived from combined phylogenetic study was used as the framework for the reconstruction of optimal ancestral distribution using DIVA 1.1. Four unit areas were defined to cover distributions of all lineages and the outgroup: North America (A), Europe (B), China (C), and Japan (D). A species was coded for presence or absence in each unit area of the genus distribution. DIVA assumes optimal solutions that minimize dispersal and extinction events under a parsimony criterion. To better search for ancestral distributions, optimize command "maxareas" was used to impose a constraint on the number of unit areas allowed in ancestral distributions. The number was set to three or two. Physical relationships among geographic areas were used to choose among alternative optimal solutions.

2.4. Estimating divergence times

To examine rate constancy among lineages in the data set, the baseml program implemented in PAML was used (Yang, 1997). The null hypothesis of a constant evolutionary rate was rejected. To estimate divergence time, the Bayesian relaxed-clock method varying the molecular rate among lineages in an autocorrected manner was used (Thorne et al., 1998). The Bayesian analyses were done with the estbranches and multidivtime programs available at http://statgen.ncsu.edu/thorne/multidivtime.html.

Megafossils from the Paleocene/Eocene Buchanan locality in western Tennessee represent the earliest unequivocal evidence of subfamily Castaneoideae of Fagaceae (Crepet, 1989; Graham, 1999). To time the separation between the disjunct species, the divergence time of *Castanea* and *Quercus* based on the fossil record, 60 million year before present (mybp), was used as landmark to calibrate the phylogenetic tree.

3. Results and discussion

3.1. Molecular variations of the studied cpDNA regions

Characteristics of the studied cpDNA regions are listed in Table 2. Within the *Castanea* matrix, 98 variable sites were detected, 46 of which are parsimony informative. Each *Castanea* species shows distinct haplotypes, with some species harboring several haplotypes. All of the examined cp regions in *Castanea* are AT rich, with an average A + T content of approximately 70%. A + T-rich regions have been found to be highly prone to replication slippage (Levinson and Gutman, 1987; Cummings et al., 1994), a mechanism that involves local intrahelical denaturation and displacement of replicating strands, which can lead to further increases in A + T content, possibly enhancing transversion bias (Bakker et al., 2000). It was found that transversions are more prevalent than transitions in the *Castanea* chloroplast genome. The patterns follow the observations of Morton (1995), in that positions flanked by A or T are more likely to undergo transversions. Insertion and deletions (indels) ranged in length from 1 to 44 bases, making sequence alignment more problematic in many cases, which occurred at varied frequency (Table 2). Most of the indels are duplications or deletions of adjacent sequences, but some may involve secondary structure. Indel events in the *ycf*9-*trn*GM region are more frequent than indels in other regions. This might be related to the high A+T content (78%) of this region, which is the highest value detected in *Castanea*. Most of the indels detected in this region were caused by slipped-strand mispairing.

Consistent with the observation of Manos and Stanford (2001), the sequence divergence among and within Fagaceae genera is very low. Pairwise sequence divergence values within *Castanea* ranged 0–2.421%. Possible explanations for the low molecular divergence could be (1) the genus *Castanea* evolves relatively slowly, or (2) divergence in *Castanea* occurred relatively recently such that neutral markers would not have had sufficient time to become fixed in different species. The sequence divergence between genera is slightly larger, since divergences of 0.214–2.913% between *Castanea* and *Quercus* and 5.183–7.229% between *Castanea* and *Fagus* were observed. Higher levels of cpDNA diversity were found in the Asian species, as evidenced by the accumulation of many unique substitutions and indels in *C. crenata*, than in the European and American species (Figs. 2 and 3). Intraspecific polymorphisms were detected in all Chinese species, whereas no differentiation was found among the *C. dentata* or *C. crenata* taxa. A further examination of 16 samples of *C. dentata* collected from throughout its natural range failed to uncover any detectable variability using PCR-RFLP of five cp regions and four restriction enzymes or using sequencing analysis of the highly variable *ycf*9-*trn*GM region (P. Lang and F. Dane, unpublished data). The highest levels of intraspecific polymorphism were detected in *C. pumila* in all regions and in *C. mollissima* in the *ndh*F3′ flanking region and the *ycf*9-*trn*GM and *ycf*6-*psb*M spacer regions. Intraspecific sequence polymorphism was detected only in the *ycf*9-*trn*GM spaces region in *C. seguinii*, in the *ndhF* 3′ flanking region in *C. henryi*, and in the *ycf*6-*psb*M and *ndhF* 3′ flanking regions in *C. sativa*.

In this study, the levels of cpDNA diversity observed among the *Castanea* species are similar to levels previously reported for allozyme diversity, with *C. mollissima* as the most diverse *Castanea* species and *C. dentata* as the least diverse (Dane et al., 2003; Huang et al., 1994; Lang and Huang, 1999). There are several possible explanations for the apparent lack of genetic diversity within *C. dentata*: this species might be a young taxon despite its wide geographic range or a genetic bottleneck may have occurred during the last glacial event or in the last century as a result of chestnut blight. Due to the effects of the Pleistocene glaciation, the decrease in species richness and diversity during the Quaternary appears to have been more pronounced in eastern NA than in eastern Asia (Qian and Ricklefs, 2000; Wen, 1999).

Unlike the divergence revealed by allozyme analysis, cpDNA diversity in *Castanea* is highly structured. Haplotypes are not randomly distributed; each haplotype is circumscribed to a particular geographic area. Moreover, haplotypes belonging to the same lineage often occupy a similar range. Seven haplotypes were found in Asia, six in China, one unique and diverse haplotype in Japan, seven in

Table 2
Characterization of cpDNA regions (coding and noncoding regions) in *Castanea* using *Quercus falcata* and *Fagus grandifolia*

	*trn*T-L-F[a]	*rpl*16	*ndh*F 3′ coding region	*ndh*F 3′ flanking region	*ycf*6-*psb*M	*ycf*9-*trn*GM
Aligned length	1770	1241	1709	679	1202	903
Mean A+T content (%)	71%	67.5%	69.4%	66.7%	70%	78%
No. parsimony informative sites (proportion)	12 (0.65%)	9 (0.73%)	9 (0.68%)	2 (0.35%)	6 (0.50%)	8 (1.32%)
Ts:Tv:Indels within Castanea	10:14:6	5:8:3	9:13:4	1:0:9	2:7:3	1:4:13
Average (range) pairwise divergence within *Castanea*	0.333% (0.046–0.851%)	0.314% (0–1.081%)	0.369% (0.12–1.019%)	0.383% (0.09–1.010%)	0.353% (0–0.791%)	1.388% (0–2.421%)
Average (range) pairwise divergence from *Q. falcate*	0.893% (0.436–1.121%)	0.979% (0.832–1.082%)	1.439% (1.319–1.559%)	1.349% (1.230–1.493%)	0.718% (0.452–0.904%)	2.256% (1.937–2.913%)
Average (range) pairwise divergence from *F. grandifolia*		6.821% (6.775–6.6861%)	6.916% (6.304–7.038%)	5.710% (5.183–5.813%)	6.339% (6.20–6.544%)	6.701% (6.325–7.229%)

[a] Data from Lang et al. (2006).

NA, and two in Europe (Fig. 2). The haplotype from Japan shares many unique substitutions with the *Quercus* and *Fagus* haplotype especially at the *rpl*16 region (Fig. 3). This haplotype divergence in Asia is consistent with species richness and points to the center of origin for *Castanea*. It is known that the entire flora of eastern Asia contains approximately one third more species than the flora of NA (Qian, 2002). Woody genera exhibit a strong diversity bias favoring eastern Asia (Qian and Ricklefs, 2000). Contemporary patterns of diversity suggest that the effects of climate change in the late Tertiary were less severe in eastern Asia and promoted diversification but were more severe in NA and may have caused widespread extinctions (Wen, 1999; Qian, 2002; Guo and Ricklefs, 2000).

3.2. Relationships within the genus Castanea

Trees generated from the individual data sets are highly congruent. Many clades have high bootstrap support based on separate analyses and support did increase, in general,

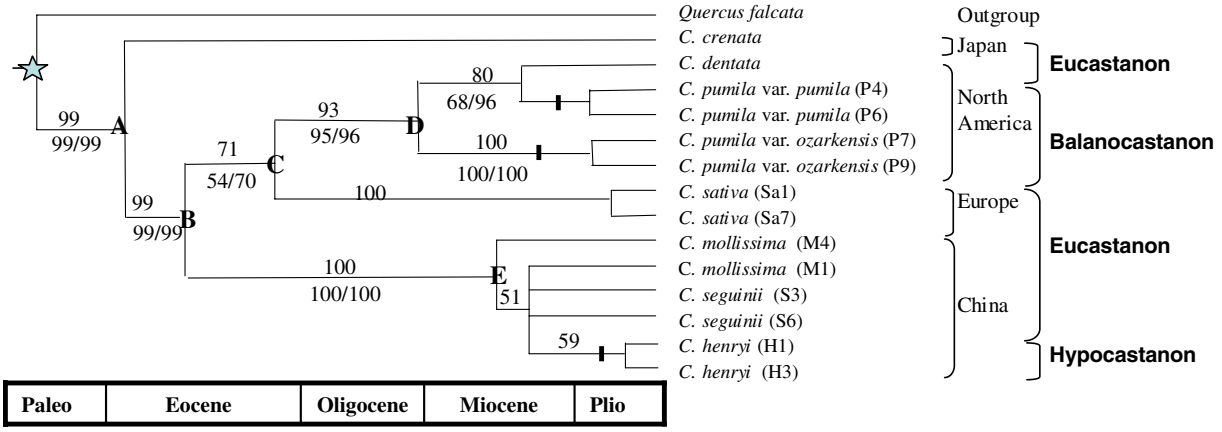

Fig. 2. The 50% majority-rule consensus of 13,451 MP trees inferred from combined data sets of *trn*T-L-F (Lang et al., 2006), *rpl*16, *ycf*6-*psb*M, and*ycf*9-*trn*GM (117 steps, CI = 0.8889, RI = 0.9030) in genus *Castanea* using *Q. falcata* as outgroup. The tree from ML and Bayesian analyses has identical topology. Numbers above branches indicate MP bootstrap value support for clade; numbers below branches are ML bootstrap value (left) and Bayesian posterior probabilities (right). Number of changes to one nut per bur species also indicated, assuming that three nuts per bur was the ancestor. Ages estimated for nodes A–E (calibrated with the fossil placed at the star) are, respectively: 54.43 mybp (95% credibility intervals 38.04–68.24), 42.55 (38.09–56.17), 39.14 (24.09–53.02), 24.41 (8.93–42.07), 10.7 (2.51–26.44). The scale bar indicates major Tertiary epochs (Paleo, Paleocene; Plio, Pliocene).

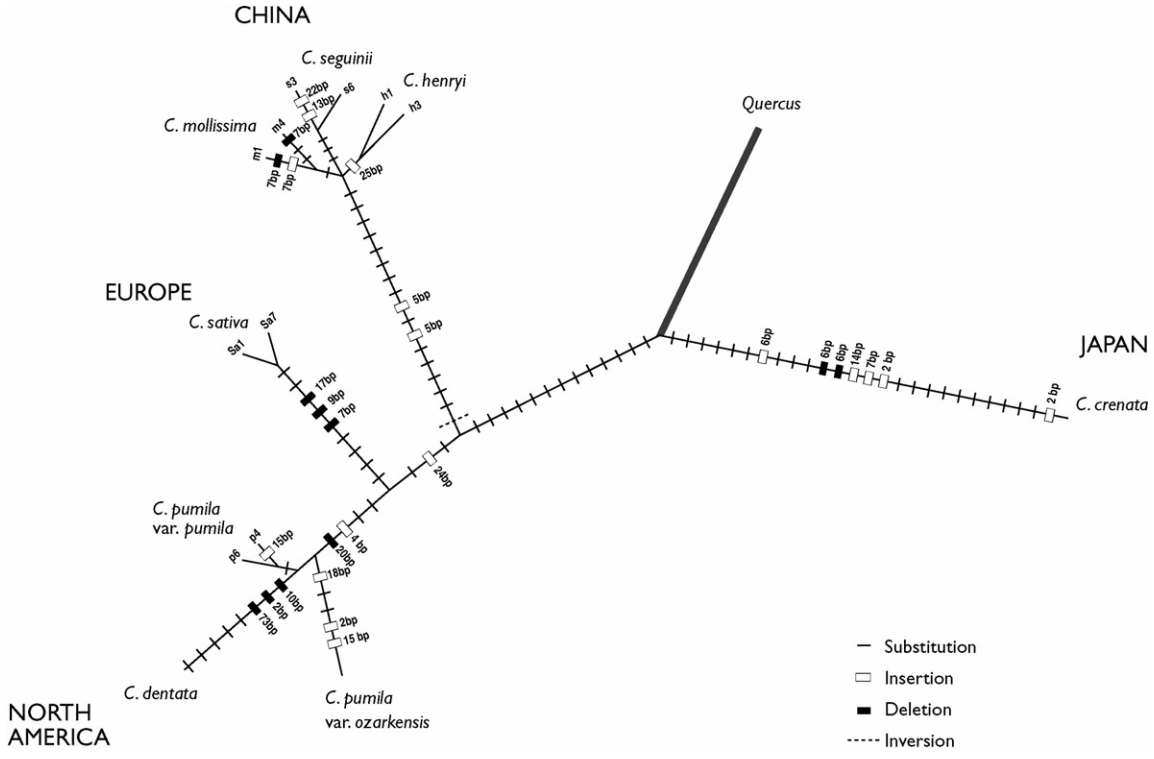

Fig. 3. Relationships among *Castanea* accessions indicating unique substitutions, insertions, deletions, and inversions detected at studied cpDNA regions using *Quercus falcata* as outgroup. The placement of indels along the branch is arbitrary.

when a combined analysis was performed. For the combined analysis, data from the *trn*T-L-F region (Lang et al., 2006) and accessions with sequence information from all studied regions were used, while accessions with identical sequences were eliminated. This combined data set contained 7721 characters, with sequences ranging in length from 6545 bp (*C. dentata*) to 6550 bp (*C. crenata*). Results from ILD tests suggest that patterns of character state variation among the studied regions differ significantly ($p = 0.02$). However, when the *ndh*F 3′ coding and 3′ flanking regions were excluded, ILD tests on all pairwise combinations of the other chloroplast sequences were not significantly incongruent ($p = 0.42$). For this reason the *ndh*F sequences were excluded from the combined analysis and a total of 5324 aligned characters were used.

The topology derived from the combined data set was identical to that generated from the separate data analyses, although the position of *C. sativa* was further strengthened. The genus *Castanea* is monophyletic, with *C. crenata* basal and sister to the remainder of the genus. The Chinese species are supported as a single monophyletic clade and are sister to the NA and European species. The most notable effect of the combined analysis is the strong support for European chestnut and NA species (71% bootstrap) and the Chinese clade at 100% bootstrap value (Fig. 2). The most parsimonious tree has 117 steps, a consistency index (CI) of 0.889 and a retention index (RI) of 0.903. Trees resulting from maximum likelihood analysis and Bayesian analyses are identical to the MP trees described above, although the bootstrap support level from ML analysis and the posterior probability from Bayesian analysis at the node of European and North American species are slightly lower (54 and 70%, respectively) than those in the MP tree (Fig. 2). Since the effect of different outgroups (*Castanopsis* or *Fagus*) and additional *C. pumila* chlorotypes on tree topology was negligible, *Quercus* was used to root the phylogenetic tree in the combined data set.

The taxonomic status of the Allegheny and Ozark chinkapins has been disputed (Johnson, 1988; Nixon, 1997). Although they are currently considered two taxonomically different varieties of *C. pumila* (Johnson, 1988), Nixon considered them separate species: *C. pumila* and *C. ozarkensis* (Nixon, 1997), respectively. The combined phylogenetic tree clearly shows *C. pumila* var. *ozarkensis* as basal but sister to the group of *C. pumila* var. *pumila* and *C. dentata*. A study of Allegheny chinkapin populations from across the Appalachian range led to the detection of the var. *ozarkensis* haplotype in some Virginia Allegheny chinkapin populations (taxon Pb, Table 1), indicative of a complicated biogeographical history of *C. pumila* and the possibility of multiple colonization routes following glaciation events.

3.3. Biogeography and evolution of the Castanea

Biogeographic explanations for taxa with significant representation in eastern Asia and eastern NA are complicated by the presence of disjunct members in Europe (Schnabel and Wendel, 1998; Schnabel et al., 2003). The oldest known *Castanea* fossils and the western NA fossil records show a pattern of geographic distribution inconsistent with the patterns from the analysis of extant taxa. However, the location of the oldest fossil record does not necessarily mark the origin of a species or genus because of the differential availability of pertinent stratigraphic sequences on different continents and the differences in the intensity of paleobotanical study in different regions (Graham, 1999; Manchester and Tiffney, 2001). Also, the accuracy of fossil records needs to be carefully considered since fossil leaf and pollen characteristics often lack diagnostic value.

Results of DIVA suggest several alternative ancestral distributions for the root of *Castanea.* When North America (A), Europe (B), China (C), and Japan (D) were defined as four different unit areas, optimization without constraint resulted in two alternative distributions (BCD or ABCD). When the number of maximum areas was reduced to three, a total of five alternative solutions were obtained (BD, ABD, CD, ACD, or BCD). When the number of maximum areas was limited to two, only three alternative solutions (AD, BD, or CD) were obtained. The most logical ancestral area appears to be located in eastern Asia (CD) since (1) the pattern (Japan (China, (Europe, North America))) suggests some diversification within Asia prior to the Old World–New World split, (2) eastern Asia has greater species richness and genetic diversity based on cpDNA and allozyme analysis, and (3) China and Japan are geographically closely linked. Geographically distant units such as North America and Japan or Europe and Japan can be eliminated because these areas require long-distance dispersal or extinction of ancestors from one unit area to another. We thus hypothesize that the common ancestor of extant *Castanea* originated in East Asia. The genus first diverged into two lineages: Japan and China. The lineage in China spread westward into Europe and then onto North America, followed by a subsequent vicariance event that fragmented Chinese, European, and North American flora, resulting in the isolation of species in China, Europe, and North America. At least three vicariance events and two dispersal events can be used to explain the distribution pattern of *Castanea* (Fig. 4).

Biogeography analyses also support an Asian origin for *Fagus* (Manos and Stanford, 2001), although this was recently disputed by Denk et al. (2005) who support a North Pacific origin of the genus. It is possible that many plant clades originated and radiated within Asia before spreading to other regions, owing to Asia's retention of the greatest number of species. Conversely, many plant species went extinct in western North America probably due to drying climates (Donoghue and Smith, 2004). CpDNA sequence data of *Castanea* clearly reveal fragmentation of the species into two main lineages, with the accumulation of unique substitutions in one main lineage of shared ancestry of Chinese, European, and NA species (Fig. 3). The Chinese species fragmented early and accumulated >16 unique substitutions, while the European and NA species, which

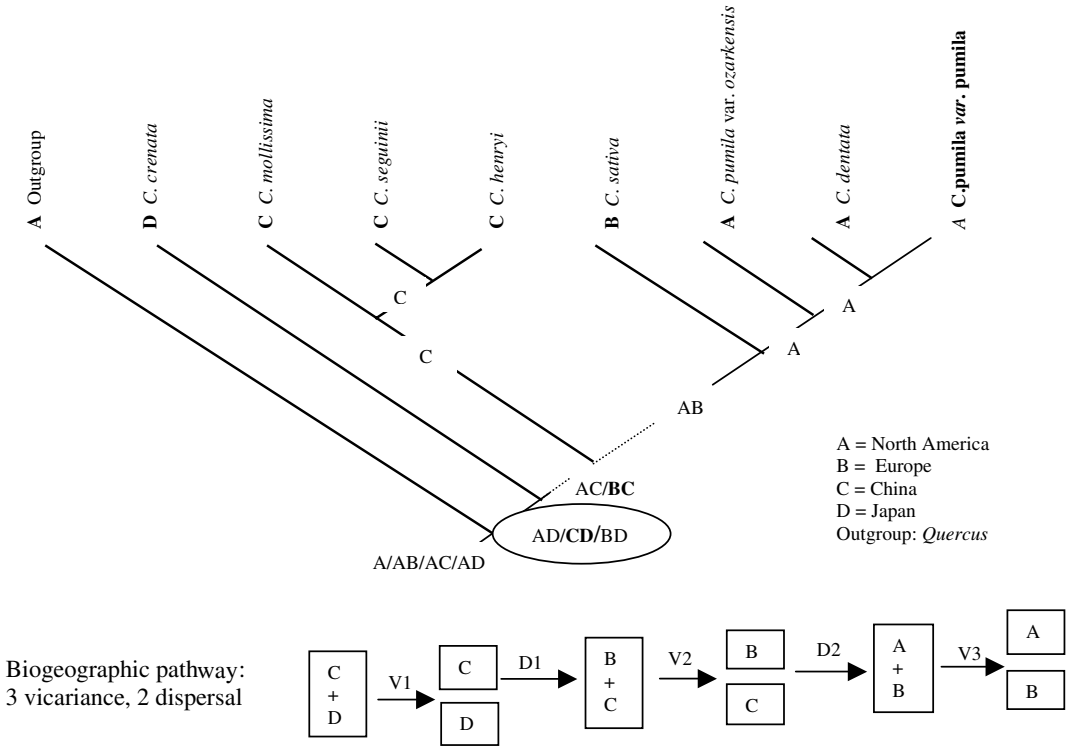

Fig. 4. Results of dispersal–vicariance analysis and the inferred biogeographic pathways of the disjunct lineages of *Castanea*. DIVA optimization was conducted using constraints of maxareas = 2 over the most resolved topology derived from the combined cpDNA sequences. Ancestral node in question is circled and the dispersal events are indicated with dashed lines.

share 2 unique substitutions and two unique insertions, are more closely related to each other than to the Asian species (Figs. 2 and 3). This clearly indicates the direction of dispersal and evolution in extant *Castanea* from Asia to Europe and onto North America. Similar results were obtained in a study of the eastern nearctic–eastern palaeoarctic disjunction of a large set of animal groups (Sanmartin et al., 2001) and several woody angiosperms such as *Liquidambar* (Li and Donoghue, 1999) and *Cercis* (Davis et al., 2002).

According to the Bayesian analyses of the combined cpDNA data, the split of eastern Asia occurred in the early Eocene (54.43 mybp, 95% credibility interval: 38.04–68.24); the divergence of European/NA species and Chinese species can be estimated to have occurred in the middle Eocene (42.55 mybp; 38.09–56.17); the split in *Castanea* between the European and the NA species is estimated to have occurred at 39.14 mybp (24.09–53.02) during the late Eocene. Diversification in NA was estimated at 24.41 mybp (8.93–42.07) and divergence of Chinese species was estimated at 10.7 mybp (2.51–26.44). Although the estimated divergence dates for *Castanea* seems older than most intercontinental disjunctions (Donoghue et al., 2001; Xiang et al., 2000), the first two splits are very similar to the age estimates of *Liquidambar* with the divergence of Europe and China at 55 mybp and the disjunction between Europe and eastern North America at 34.99 ± 5.5 mybp (Donoghue et al., 2001).

Flora exchange between Asia and eastern North America could have taken place via two different routes: either across the Bering Land Bridge (BLB) or across the North Atlantic European–North American land bridge (NALB). Palynological and foliar evidence suggest that the Eocene floras near the BLB were dominated by deciduous angiosperm genera (as reviewed by Graham (1999) & Tiffney and Manchester (2001)). Although botanists have considered the BLB to be more important (Wen, 1999), Tiffney (1985a,b) argued that the boreotropical flora dispersed mainly across the North Atlantic land bridge. An important difference between these two hypotheses is in the predicted age of the disjunction. The North Atlantic European–North American land bridge is speculated to have existed from the early Eocene (55 mybp) until the late Miocene, although it was possibly interrupted during the Oligocene, 40–30 mybp (Manchester, 1999). Geological evidence suggests that physical connections between North America, Greenland, and Europe became increasingly uncommon through the Paleogene and Neogene, although phytogeographic similarities between Oligocene and Miocene floras of North America and Europe (Hably et al., 2000) and molecular studies (Wen, 1999) indicate a Miocene divergence time for North American–European clades. These links appear especially strong between Europe and western North America (Hably et al., 2000), although this conclusion may partially be a function of the relative scarcity of eastern North American Tertiary deposits. This pattern may reflect the action of "morphological stasis" of taxa that achieved distribution earlier in time (Wen, 1999). The continuity of Arctic–alpine taxa shared

between the Old and the New Worlds in the present day attests to the exchange of cold-adapted taxa in the latest Tertiary and Quaternary (Tiffney and Manchester, 2001). Although more investigations are needed for the accurate estimation of the divergence times of the disjunct *Castanea* species, current age estimates are readily accommodated by prevailing ideas on the availability of the NALB pathway.

The Japanese chestnut represents descendants of the earliest radiations in *Castanea*. It was estimated that an initial diversification happened within eastern Asia in the early Eocene (54.43 mybp). Following the formation of the Japanese sea, the unique island climate resulted in the allopatric speciation of the Japanese chestnut. Hence, *C. crenata* has the most divergent haplotype in relation to the other *Castanea* species. Almost all autapomorphic characters at the studied cpDNA regions are unique to *C. crenata*, possibly as a result of haplotype isolation. The common finding in existing phylogenetic studies is that east Asia taxa are basal to derived New World taxa, leading to the conclusion that east Asia is a cradle for taxa with the classic east Asia–eastern North American disjunction (Tiffney and Manchester, 2001). However, the final geographical disposition of the most primitive surviving fossil may well be a function of differential extinction rather than of place of origin. Europe and North America have suffered extensive climatic changes in the Tertiary, which would favor both extinction and allopatric speciation. Southeastern and eastern Asia offer a great diversity of land forms, many of which evolved in the early and mid Tertiary. The diversity of topography and east–west orientation of mountains created habitats suitable for a refugium (Tiffney and Manchester, 2001). After the split within Asia, the Chinese lineage spread westward into Europe and then onto North America and resulted in the radiation of species in China, Europe, and North America. The Quaternary history of European *C. sativa* has been well documented, indicating that the species retreated back to western Asia and survived as a relict in a few locations of southeastern Europe and Turkey (Villani et al., 1991, 1994; Fineschi et al., 2000). Because of the complete coverage of Europe with ice during the glacial periods and the east–west orientation of the mountains, which effectively prevented organisms from migrating to southern refugia, Europe lacks species diversification (Latham and Ricklefs, 1993). This might explain why only chestnuts (*C. sativa*) evolved on the European continent, while chinkapins evolved on the American and Asian continents. However, because of effects of climate changes, extinction might be another possible reason for the lack of species diversity in Europe. Chinkapins might have existed in Europe but were eliminated by recent Pleistocene glacial events.

Diversification in NA was estimated at 24.41 mybp (8.93–42.07) during the early Miocene and resulted in the chinkapins. Although *C. dentata* has the ancestral cupule type, the general lack of the dense, stellate pubescence on its leaves as found in most members of sect. *Eucastanon* and chinkapins may indicate that *C. dentata* is a derived condition. Johnson (1988) considered the Ozark chinkapin as ancestral and less highly evolved than the Allegheny chinkapin. NA species might have a common extinct ancestor with three nuts per cupule. Based on the cpDNA sequence data, evolution of one nut per cupule may have evolved two times in North America resulting in C. *pumila* var. *ozarkensis* first and *C. pumila* var. *pumila* later (Fig. 2). The distribution of NA species was affected by the Pleistocene glaciations. In eastern North America, the American species were driven southward to refugia in the Gulf Coastal region and Florida during the Wisconsin glacial maximum 18,000–20,000 years ago. *C. dentata* and *C. pumila* migrated northward along the Appalachia after the glacier's retreat (Huang et al., 1998; Davis, 1983; Graham, 1999; Delcourt, 2002).

Low sequence divergence among three Chinese species suggests a recent divergence. It was estimated at 10.7 mybp (2.51–26.44) between *C. mollissima* and the other two Chinese species. *C. seguinii* is the only species within the genus to have a precocious growth habit, indicating that *C. seguinii* evolved from an ancestor closely resembling *C. mollissima.* Similarly, *C. henryi* is derived from the chestnuts via a reduction of nuts per cupule. Eastern Asia is known to have been less affected by the Pleistocene glaciations (Tiffney, 1985a,b). It was hypothesized that the Changjiang river area, particularly in Shennongjia, provided an ideal refugium for Chinese *Castanea* species during the glacial maximum, which migrated northward after the glacier's retreat (Lang and Huang, 1999).

4. Conclusion

This study illustrates the great potential of cpDNA sequence data to reconstruct the phylogenetic relationships and directions of migration of *Castanea* species. Unique westward expansion of extant *Castanea* species via NALB, which does not exclude expansion via BLS of extinct species, is hypothesized based on cpDNA sequences. Our biogeographic reconstruction and molecular dating suggest that *Castanea* originated in eastern Asia. An initial diversification within Asia (Japan and China) during the early Eocene is postulated followed by intercontinental dispersion and divergence between the Chinese and the European/North American species during the middle Eocene and a split between the European and the North American species in the early Oligocene. The differentiation in North America and China occurred in the early and late Miocene, respectively. Because of difficulties with the integration of the fossil record with studies of modern species, our analyses with DIVA were based solely on phylogenies of extant species. More information from the geological and climatic history of the earth and single-copy nuclear DNA sequences are needed in combination with our cpDNA data to provide a comprehensive view of evolutionary relationships and biogeography of this ecologically and economically important genus.

Acknowledgments

We thank Cliff Parks of the University of North Carolina, Fred Hebard from The American Chestnut Foundation, Sandy Anagnostakis of the Connecticut Agricultural Research Station, and Alexandru-Lucian Curtu for providing samples.

References

Bakker, F.T., Culham, A., Gomez-Martinez, R., Carvalho, J., Compton, J., Dawtrey, R., Gibby, M., 2000. Patterns of nucleotide substitution in angiosperm cpDNA *trn*L (UAA)-*trn*F (GAA) regions. Mol. Biol. Evol. 17, 1146–1155.

Clegg, M.T., Zurawaki, G., 1992. Chloroplast DNA and the study of plant phylogeny: Present status and future prospects. In: Doyle, J.J. (Ed.), Molecular Systematics of Plants. Chapman Hall, New York, pp. 275–294.

Crepet, W.L., Daighlian, C.P., 1980. Castaneoid inflorescences from the Middle Eocene of Tennessee and the diagnostic value of pollen (at the subfamily level) in the Fagaceae. Am. J. Bot. 67, 739–757.

Crepet, W.L., 1989. History and implications of the early North American fossil record of Fagaceae. In: Blackmore, S. (Ed.), Evolution, systematics, and Fossil History of the Hamamelidae. Oxford University Press, New York, pp. 45–66.

Cummings, M., King, L., Kellogg, E., 1994. Slipped-strand mispairing in a plastid gene: *rpo*C2 in grasses (Poaceae). Mol. Biol. Evol. 11, 1–8.

Dane, F., Hawkins, L.K., Huang, H., 1999. Genetic variation and population structure of *Castanea pumila* var. *ozarkensis*. J. Am. Soc. Hort. Sci. 124, 666–670.

Dane, F., Lang, P., Huang, H., Fu, Y., 2003. Intercontinental genetic divergence of *Castanea* species in eastern Asia and eastern North America. Heredity 91, 314–321.

Davis, M.B., 1983. Quaternary history of deciduous forests in eastern North America and Europe. Ann. Mo. Bot. Gard. 70, 550–563.

Davis, C.C., Fritsch, P.W., Li, J., Donoghue, M.J., 2002. Phylogeny and biogeography of *Cercis* (Fabaceae): evidence from nuclear ribosomal ITS and chloroplast *ndh*F sequence data. Syst. Bot. 27, 289–302.

Delcourt, H.R., 2002. Forests in peril. Tracking Deciduous Trees from Ice-age Refuges into the Greenhouse World. The McDonald and Woodward Publishing Company.

Denk, T., Grimm, G.W., Hemleben, V., 2005. Patterns of molecular and mophological differentiation in *Fagus* (Fagaceae): phylogenetic implications. Am. J. Bot. 92, 1006–1016.

Donoghue, M.J., Bell, C.D., Li, J., 2001. Phylogenetic patterns in Northern Hemisphere geography. Int. J. Plant Sci. 162, S41–S60.

Donoghue, M.J., Smith, S.A., 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. Phil. Trans. R. Soc. Lond. B. 359, 1633–1644.

Dumolin-Lapegue, S., Demesure, B., Fineschi, S., Le, C.V., Petit, R.J., 1997. Phylogeographic structure of white oaks throughout the European continent. Genetics 146, 1475–1487.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, B.C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.

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

Fineschi, S., Taurchini, D., Villani, F., Vendramin, G.G., 2000. Chloroplast DNA polymorphism reveals little geographical structure in *Castanea sativa* mill. (Fagaceae) throughout southern European countries. Mol. Ecol. 9, 1495–1503.

Graham, A., 1999. Late Cretaceous and Cenozoic history of North American Vegetation. Oxford University Press, NY.

Graves, A.H., 1950. Relative blight resistance in species and hybrids of *Castanea*. Phytopathology 49, 1125–1131.

Guo, Q., Ricklefs, R., 2000. Species richness in plant genera disjunct between temperate eastern Asia and North America. Bot. J. Linn. Soc. 134, 401–423.

Hably, L., Kvacek, Z., Manchester, S.R., 2000. Shared taxa of land plants in the Oligocene of Europe and North America in context of Holarctic phytogeography. Acta Univ. Carol. Geol. 44, 59–74.

Huang, H., Carey, W.A., Dane, F., Norton, J.D., 1996. Evaluation of Chinese chestnut cultivars for resistance to *Cryphonectria parasitica*. Plant Dis. 80, 45–47.

Huang, H., Dane, F., Kubisiak, T.L., 1998. Allozyme and RAPD analysis of the genetic diversity and geographic variation in wild populations of the American chestnut (Fagaceae). Am. J. Bot. 85, 1013–1021.

Huang, H., Dane, F., Norton, J.D., 1994. Allozyme diversity in Chinese, Seguin and American chestnut (*Castanea* spp.). Theor. Appl. Genet. 88, 981–985.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Jaynes, R., 1975. Chestnut. In: Moore, J. (Ed.), Advances in Fruit Breeding. Purdue University Press, West Lafayette, pp. 490–503.

Johnson, G.P., 1988. Revision of *Castanea* Sect. Balanocastanon (Fagaceae). J. Arnold Arboretum 69, 25–49.

Lang, P. 2004. Phylogeny and Biogeography of *Castanea* (Fagaceae) Based on Chloroplast Sequence Data. Ph.D. Dissertation. Auburn University, AL.

Lang, P., Huang, H., 1999. Genetic diversity and geographic variation in natural populations of the endemic *Castanea* species in China. Bot. Sinica 41, 651–657.

Lang, P., Dane, F., Kubisiak, T.L., 2006. Phylogeny of *Castanea* (Fagaceae) based on chloroplast *trn*T-L-F sequence data. Tree Gene. Genomes. 2, 132–139.

Latham, R.E., Ricklefs, R.E., 1993. Continental comparisons of temperate-zone tree species diversity. In: Schluter, D. (Ed.), Species Diversity in Ecological Communities. The University of Chicago Press, Chicago and London, pp. 294–314.

Levinson, G., Gutman, G., 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol Evol. 4, 203–221.

Li, J., Donoghue, M.J., 1999. More molecular evidence for interspecific relationships in *Liquidambar* (Hamamelidaceae). Rhodora 101, 87–91.

Manchester, S.R., 1999. Biogeographical relationships of North American Tertiary floras. Ann. Miss. Bot. Gard. 86, 472–522.

Manchester, S.R., Tiffney, B.H., 2001. Integration of Paleobotanical and Neobotanical Data in the Assessment of Phytogeographic History of Holarctic Angiosperm Clades. Int. J. Plant Sci. 162, S19–S27.

Manos, P.S., Stanford, A.M., 2001. The historical biogeography of Fagaceae: tracking the Tertiary history of temperate and subtropical forests of the Northern Hemisphere. Int. J. Plant Sci. 162, S77–S93.

Morton, B.R., 1995. Neighboring base composition and transversion/transition bias in a comparison of rice and maize chloroplast noncoding regions. Proc. Natl. Acad. Sci. USA 92, 9717–9721.

Manos, P.S., Zhou, Z.K., Cannon, C.H., 2001. Systematics of Fagaceae: Phylogenetic tests of reproductive trait evolution. Int. J. Plant Sci. 162, 1361–1379.

Nixon, K.C., 1997. Flora of North America Editorial Committee (Eds.), Flora of North America North of Mexico, Vol. 3. Oxford University Press, NY, pp. 436–437.

Olmstead, R., Palmer, J.D., 1994. Chloroplast DNA systematics: a review of methods and data analysis. Am. J. Bot. 81, 1205–1224.

Palaeobotany, I.O.O. 1997. The plant fossil record (PFR v 2.2) database. Available from <http://www.biodiversity.org.uk/ibs/palaeo/pfr2/pfr.htm/>.

Posada, D., Crandall, K.A., 1998. Model test: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Qian, H., 2002. A comparison of the taxonomic richness of temperate plants in East Asia and North America. Am. J. Bot. 89, 1818–1825.

Qian, H., Ricklefs, R.E., 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. Nature (London) 407, 180–186.

Ronquist, F. 1996. DIVA, ver. 1.1. Computer program Available from <http://www.ebc.uu.se/systzoo/research/diva/diva.html/>.

Ronquist, F., 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. Syst. Biol. 46, 195–203.

Sanmartin, I., Enghoff, H., Ronquist, F., 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linn. Soc. 73, 345–390.

Schnabel, A., McDonel, P.E., Wendel, J.F., 2003. Phylogenetic relationships in *Gleditsia* (Leguminosae) based on ITS sequences. Am. J. Bot. 90, 310–320.

Schnabel, A., Wendel, J.F., 1998. Cladistic biogeography of *Gleditsia* (Leguminosae) based on *ndh*F and *rpl*16 chloroplast gene sequences. Am. J. Bot. 85, 1753–1765.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Simmons, M.P., Ochotreana, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.

Stanford, A., 1998.The Biogeography and Phylogeny of *Castanea*, *Fagus*, and *Juglans* based on *mat*K and ITS sequence data Biology UNC, Chapel Hill.

Swofford, D.L., 2003. PAUP* Ver. 4.0b10. Sinauer Associates, Sunderland, MA.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.

Tidwell, W.D., 1998. Common fossil plants of western North America. Smithsonian Institution Press, Washington.

Tiffney, B.H., 1985a. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the northern hemisphere. J. Arnold Arboretum 66, 243–273.

Tiffney, B.H., 1985b. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. J. Arnold Arboretum 66, 73–94.

Tiffney, B.H., Manchester, S.R., 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemipshere Tertiary. Int. J. Plant Sci. 162, S3–S17.

Villani, F., Pigliucci, M., Benedettelli, S., Cherubini, M., 1991. Genetic differentiation among Turkish chestnut (*Castanea sativa* Mill.) populations. Heredity 66, 131–136.

Villani, F., Pigliucci, M., Cherubini, M. 1994. Evolution of *Castanea sativa* Mill. in Turkey and Europe. Genet. Res. 63, 109–116.

Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu. Rev. Ecol. Syst. 30, 421–455.

Xiang, Q.-Y., Soltis, D.E., Soltis, P.S., Manchester, S.R., Crawford, D.J., 2000. Timing the eastern Asian-eastern North American floristic disjunction: molecular clock corroborates paleontological data. Mol. Phylogenet. Evol. 15, 462–472.

Xiang, Q.-Y.J., Soltis, D.E., 2001. Dispersal-vicariance analyses of intercontinental disjuncts: historical biogeographical implications for angiosperms in the Northern Hemisphere. Int. J. Plant Sci. 162, S29–S39.

Xiang, Q.-Y., Manchester, S.R., Thomas, D.T., Zhang, W.H., Fan, C.Z., 2005. Phylogeny, biogeography, and molecular dating of Cornelian cherries (*Cornus*, Cornaceae): tracking tertiary plant migration. Evolution 59, 1685–1700.

Xiang, Q.-Y., Thomas, D.T., Zhang, W.H., Manchester, S.R., Murrell, Z., 2006. Species level phylogeny of the genus *Cornus* (Cornaceae) based on molecular and morphological evidence-implications for taxonomy and Tertiary intercontinental migration. Taxon 55, 9–30.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.

Zhou, Z., 1999. Fossils of the Fagaceae and their implications in systematics and biogeography. Acta Phytotax. Sin. 37, 369–385.