

## Recent Long-Distance Dispersal Overshadows Ancient Biogeographical Patterns in a Pantropical Angiosperm Family (Simaroubaceae, Sapindales)

JOSHUA W. CLAYTON<sup>1\*</sup>, PAMELA S. SOLTIS<sup>2</sup>, AND DOUGLAS E. SOLTIS<sup>1</sup>

<sup>1</sup>Department of Botany and

<sup>2</sup>Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA;

\*Correspondence to be sent to: Department of Botany, University of Florida, Gainesville, FL 32611-1100, USA; E-mail: joshuk@ufl.edu.

**Abstract.**—Detailed biogeographic studies of pantropical clades are still relatively few, and those conducted to date typically use parsimony or event-based methods to reconstruct ancestral areas. In this study, a recently developed likelihood method for reconstructing ancestral areas (the dispersal–extinction cladogenesis [DEC] model) is applied to the angiosperm family Simaroubaceae, a geographically widespread and ecologically diverse clade of pantropical and temperate trees and shrubs. To estimate divergence dates in the family, Bayesian uncorrelated rates analyses and robust fossil calibrations are applied to the well-sampled and strongly supported phylogeny. For biogeographic analyses, the effects of parameter configurations in the DEC model are assessed for different possible ancestral ranges, and the likelihood method is compared with dispersal–vicariance analysis (DIVA). Regardless of the parameters used, likelihood analyses show a common pattern of multiple recent range shifts that overshadow reconstruction of events deeper in the family’s history. DIVA produced results similar to the DEC model when ancestral ranges were restricted to two areas, but some improbable ancestral ranges were also observed. Simaroubaceae exhibit an early history of range expansion between major continental areas in the Northern Hemisphere, but reconstruction of ancestral areas for lineages diverging in the early Tertiary are sensitive to the parameters of the model used. A North American origin is suggested for the family, with migration via Beringia by ancestral taxa. In contrast to traditional views, long-distance dispersal events are common, particularly in the Late Oligocene and later. Notable dispersals are inferred to have occurred across the Atlantic Ocean in both directions, as well as between Africa and Asia, and around the Indian Ocean basin and Pacific islands. [DEC model; divergence date estimation; historical biogeography; long-distance dispersal; range evolution; Simaroubaceae.]

The number of phylogeny-based biogeographical treatments for tropical groups has grown considerably in recent years (Lavin et al. 2000, 2004; Renner et al. 2001; Davis et al. 2002; Richardson et al. 2004; Weeks et al. 2005; Zerega et al. 2005; Muellner et al. 2006), with improved phylogeny reconstruction and dating techniques, and a diversity of methods for ancestral area reconstruction (AAR). However, resolved and well-supported phylogenies of pantropical clades, that is, clades with taxa distributed in each of the neotropics, tropical Africa and tropical Asia and Australia, are still relatively few (Renner et al. 2001; Davis et al. 2002; Lavin et al. 2004; Weeks et al. 2005; Zerega et al. 2005; Muellner et al. 2006). Simaroubaceae are an ideal pantropical candidate for exploring hypotheses of intercontinental migration, given their well-supported relationships, taxon sampling, and broad geographical distribution (Clayton et al. 2007). Although the family is much smaller in species diversity (~109 species) than many well-studied pantropical clades, a number of intriguing disjunctions warrant a detailed investigation. All three major tropical disjunctions, between Asia and the New World, between Africa and the New World, and between Africa and Southeast Asia, occur in the family, and occur multiple times. Furthermore, Simaroubaceae have tropical, subtropical, and temperate elements, including both mesophytic and dry-adapted taxa. These factors add an ecological dimension to the interpretation of biogeographical patterns (Riddle 1996; Wiens and Donoghue 2004) and may provide evidence for migration and speciation as a result of climatic change. Recent studies have shown the biogeographical history of a number of tropical angiosperm clades (Renner et al. 2001; Davis et al. 2002;

Weeks et al. 2005; Zerega et al. 2005; Muellner et al. 2006) to be the result of interactions between climate change, migration via land-bridges, and long-distance dispersal, and the synthesis of data such as these has provided an analytical framework for the study of the origins and maintenance of tropical biomes as a whole (Davis et al., 2005; Pennington et al. 2004a, 2006). Additionally, the timing of radiations of strictly tropical taxa from subtropical and temperate ancestors may reveal how adaptive shifts contribute to the modern tropical community (Lavin et al. 2004), and Simaroubaceae provide an opportunity for such a study. Although the family does not constitute a significant component of biome diversity in terms of species numbers as in, for example, legumes (Lavin et al. 2004) or Melastomataceae (Renner et al. 2001), it may provide data on how smaller families enhance the phylogenetic diversity (Faith 1992; Barker 2002) of species-rich environments.

Simaroubaceae have a variety of fruit types and morphology; fruits can be small and nut-like (e.g. *Brucea*, *Eurycoma*), small and fleshy (e.g., *Picrasma*, *Castela*), large and fleshy (e.g., *Gymnostemon*, *Pierreodendron*), dry and winged (samaras; e.g., *Ailanthus* and *Soulamea*) and woody and floating (*Samadera indica*). Reconstructing dispersal events in the history of the family may provide some insight into how dispersal capabilities are manifest in fruit type.

The goals of this study were to 1) reconstruct divergence dates within Simaroubaceae using recently developed molecular rates analyses and fossil calibration; 2) propose a center of origin for the Simaroubaceae based on molecular rates analyses and AARs; 3) use suitable inter- and intracontinental dispersal routes, their maintenance and disappearance, to test hypotheses

of vicariant cladogenesis in the history of the family; and 4) identify potential long-distance dispersal events when vicariant hypotheses of speciation seem doubtful, and relate these to fruit morphology. Consideration will be given to the effects of past episodes of climate change, the ecological requirements of extant species, and how these might be reflected in ancestral distributions. Maximum-likelihood analysis of geographic range evolution (Ree et al. 2005; Ree and Smith 2008) has yet to be applied to a pantropical group such as Simaroubaceae. The complex history of inter- and intra-continental migration that is suggested by phylogenetic relationships between the extant species, and the areas they inhabit, should be well suited to a model-based method, rather than parsimony or event-based methods (dispersal–vicariance analysis (DIVA); Ronquist, 1997) that are currently used. A DIVA analysis is included in the study for comparative purposes.

## METHODS

### *Taxon Sampling, DNA Sequencing, and Phylogenetic Analyses*

Taxa used in this study were the same as those of Clayton et al. (2007), with the addition of 4 species of *Simaba* (plus a second accession of *S. cuneata*) and 1 species of *Simarouba*. Sequences for these new taxa were determined using the same methods as Clayton et al. (2007) for 4 gene regions, *rbcl*, *atpB*, *matK* plus partial *trnK* intron, and *phyC* (Genbank accession numbers EU546227–EU546249). These taxa were then added to the 67 accessions used in the previous study, totaling 73 ingroup accessions overall, and 62 of ~109 species. Phylogenetic analyses consisted of Bayesian inference. Sequence data were partitioned by gene region and codon position for coding regions. This partitioning strategy was determined to be the best fitting model using Bayes factors (Clayton et al. 2007). Substitution models for each partition were determined with the Akaike information criterion (AIC) in Modeltest, version 3.6 (Posada and Crandall 1998). Two independent analyses were run for 10 000 000 generations each using 4 Markov chains and default parameters; trees were sampled every thousandth generation, with a burn-in of 150 000 generations, and stationarity of the Markov chain Monte Carlo (MCMC) was determined using the average standard deviation of split frequencies between runs. The data matrix and resulting tree were submitted to TreeBASE (matrix accession number M3859; study accession number S2060). Compared with analyses in Clayton et al. (2007), no changes to the topology, its resolution, or clade support were found with the additional species. A likelihood-ratio test was performed on the combined data in PAUP\* 4.0b10 (Swofford 2002), to determine if the regions sequenced are evolving in a clock-like fashion using maximum likelihood. The substitution model for the combined data set (GTR + I + G) was determined using AIC in Modeltest. Results of this test ( $\chi^2 = 245$  and  $df = 79$ ,  $P < 0.0001$ ) rejected the molec-

ular clock hypothesis, and a relaxed clock method was used.

### *Divergence Date Estimation*

Divergence date estimation was performed using the uncorrelated lognormal (UCLN) model of molecular rates estimation (Drummond et al. 2006), implemented in BEAST (v1.4.7; Drummond and Rambaut 2007), a Bayesian approach to molecular rates analysis. This method uses a global sampling strategy (MCMC) to maximize both phylogenetic inference and molecular rate estimation on branches simultaneously (Drummond et al. 2006), thus eliminating the problems associated with selecting a tree topology and branch lengths a priori. The method also allows for calibration date priors to have a probability distribution, to reflect uncertainties in fossil dates.

*Fossil calibration.*—The phylogeny was calibrated with 4 fossils: samaras of 2 different fossil species of *Ailanthus* (Corbett and Manchester 2004), *Leitneria* fruit endocarps (Dorofeev 1994; Nikitin 2006a) and *Chaneya* fruits (Wang and Manchester 2000; Teodoridis and Kvacek 2005). *Ailanthus* samaras are common in deposits throughout the Tertiary across the Northern Hemisphere (Fig. 1), with earliest occurrences in the Middle to Early Eocene. The earliest example is from the Fossil Butte fish quarries of the Green River formation, Wyoming, dated at 52.2–52.7 Ma. Fossil *Ailanthus* samaras share all the fruit synapomorphies of the extant genus that are discernible in the preserved remains. However, attributing the fossil fruits specifically to one or another extant species is difficult due to the limited number of characters available (Corbett and Manchester 2004). Three fossil species are recognized: *Ailanthus confucii*, the most commonly occurring species in the genus' fossil history, and closest morphologically to extant *Ailanthus altissima*; *Ailanthus tardensis*, known from a single locality in Hungary in the Early Oligocene, and closest in form to extant *Ailanthus triphysa*; *Ailanthus gigas*, known from a single locality in Croatia in the Early Oligocene, and described as being “intriguingly similar” to extant *Ailanthus integrifolia* (Corbett and Manchester 2004). Although the 3 fossil species have affinities to extant species, there is variation in characteristics such as fruit size and position of the ventral vein. *Ailanthus confucii* specimens are the earliest example of the genus; however, *Ailanthus fordii* is the sister to all other extant species (although *Ailanthus excelsa* has not been sampled; Clayton et al. 2007).

For these reasons, 3 different *Ailanthus* calibration points were tested in the molecular rates analyses. First, the earliest fossil occurrence (*A. confucii*, 52.7 Ma) was used to date the node at which *Ailanthus* diverges from the rest of the Simaroubaceae (the *Ailanthus* stem lineage). Second, the earliest fossil occurrence (*A. confucii*, 52.7 Ma) was used to date the node at which *A. altissima* diverges from its sister group (*A. integrifolia*+*A. triphysa*). Third, the *A. gigas* fossil (Early Oligocene) was used to

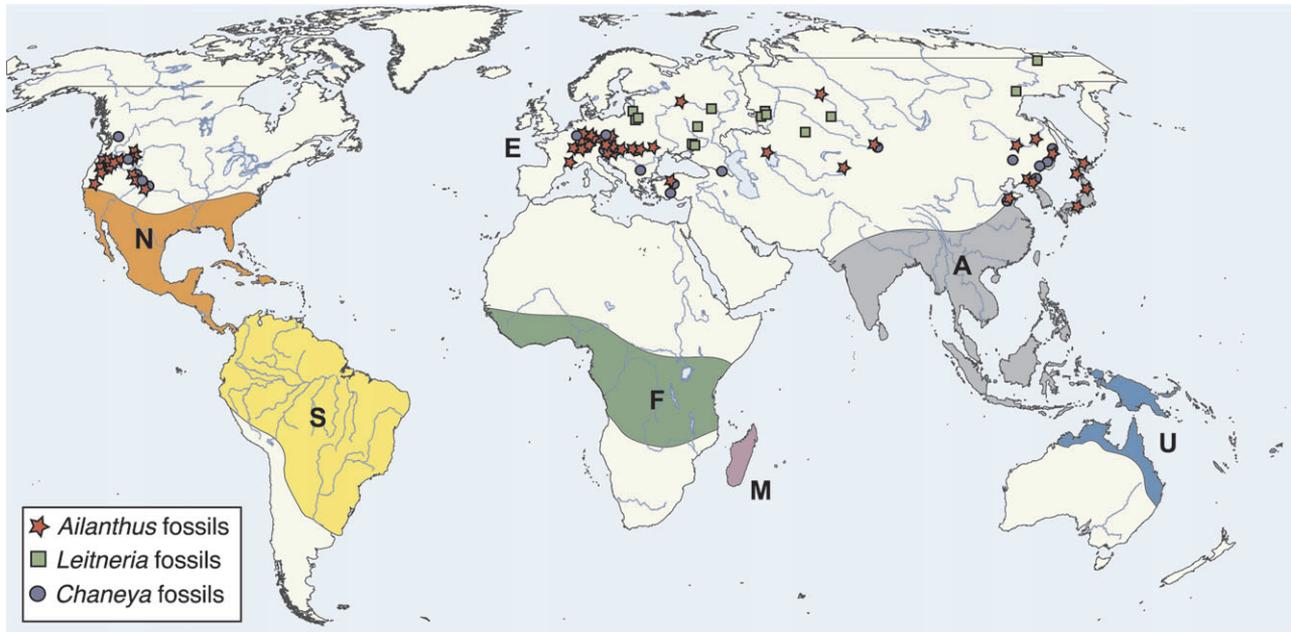


FIGURE 1. Map showing extant geographic distribution and approximate fossil localities for Simaroubaceae. Colors and letters refer to the 7 areas assigned to terminals in biogeographic analyses and shown in Figure 4. Area abbreviations are as follows: N = North and Central America (orange); S = South America (yellow); F = Africa (green); M = Madagascar (purple); A = Mainland and insular southeast Asia (gray); U = Australia (blue); E = Europe (no color).

date the node at which *A. integrifolia* diverges from its sister species (*A. triphysa*). Given the numerous occurrences of *Ailanthus* fossils in deposits across the Northern Hemisphere, and their consistent appearance throughout the Tertiary, up until the Pleistocene of Japan (Corbett and Manchester 2004), the date for *A. confucii* was considered a robust calibration, as in previous dating analyses (Muellner et al. 2006, 2007). This robustness is considered in terms of how closely the earliest known appearance of the fossil approximates to the first appearance of the genus: *Ailanthus* samaras are readily preserved, straightforward to identify, and relatively common, and the sudden appearance of samaras in the fossil record in the Early Eocene may indicate that the genus did not exist before that time. Proximity of the assigned fossil date to the initial cladogenesis event in geological time is critical to accurate divergence date estimation.

*Leitneria* has fossilized endocarps from the Early Oligocene to the Pliocene of Western Siberia and the Miocene and Pliocene of Europe (Fig. 1). Several species are described (Dorofeev 1994; Nikitin 2006a) and show a striking similarity to the endocarp anatomy of extant *Leitneria floridana*. Nikitin (2006a) recorded a putative *Leitneria* sp. from the Late Eocene (Priabonian), but its precise affinities are questioned by the author, therefore Early Oligocene was used in calibrations. As extant *Leitneria* is a monotypic genus, the only logical position for the calibration point is the node at which *L. floridana* diverges from its sister clade.

A third fossil is the extinct genus *Chaneya* (Wang and Manchester 2000; Teodoridis and Kvacek 2005), which has certain affinities with extant *Picrasma*. Although the

fossil appears to have oil cells in the perianth, which are lacking in Simaroubaceae, an accrescent pentamerous calyx surrounding 1, 2, or 3 globose mericarps when mature, is consistent with *Picrasma* (Wang and Manchester 2000). A *Chaneya* calibration point at the *Picrasma* stem group was therefore tested for compatibility with the other fossils, using its earliest occurrences in the Middle Eocene of western North America and Asia. However, the potential utility of *Chaneya* is limited by its uncertain relationship.

Fossil pollen of Simaroubaceae was not considered for this study, given the lack of distinctive morphological characteristics in extant Simaroubaceae pollen (Basak 1963, 1967; Moncada and Machado 1987). Reports of *Ailanthus* fossil pollen (Song et al. 2004) and *Leitneria* fossil pollen (Machen 1971) were considered unreliable.

In BEAST v1.4.7 (Drummond and Rambaut 2007), calibration points take the form of priors with a probability distribution. Each prior tested (Table 1) took a lognormal distribution, with the means and standard deviations designated to reflect our confidence in the fossils used. The *A. confucii* fossil was assigned a narrow lognormal distribution, given its well-characterized fossils and more precise stratigraphic age. *Leitneria* was given a broader probability distribution than *A. confucii* to account for the questionable Priabonian species. Both *A. gigas* and *Chaneya* calibrations had larger standard deviations because of the uncertain affiliations of these fossils. The hard lower bounds represent the youngest dates (Donoghue and Benton 2007) for stratigraphic ages of epoch boundaries (Gradstein and Ogg 2004), and in the case of *A. confucii*, the more precise dating of the Fossil Butte fish quarries at Green River, Wyoming.

TABLE 1. Fossil calibrations used in BEAST analyses, for testing compatibility between fossils, and constructing the final chronogram for biogeographic analyses

Calibration point	Fossil	Prior distribution	Mean/SD of lognormal distribution	Hard lower bound/mean/soft upper bound (95%) (Ma)
A	<i>Ailanthus confucii</i>	Lognormal	1.0/0.5	52.2/55.3/58.4
B	<i>Ailanthus gigas</i>	Lognormal	1.0/1.0	28.4/32.9/42.5
C	<i>Leitneria</i> spp.	Lognormal	1.5/0.75	28.4/34.3/43.8
D	<i>Chaneya</i> spp.	Lognormal	1.0/1.0	40.4/44.9/54.5

SD = standard deviation.

In total, 8 different combinations of fossil placements (here referred to as calibration schemes; Table 2) were tested to determine how well fossils and divergence dates correlated with each other. Calibration at the family level (Wikström et al. 2001, 2004; Muellner et al. 2007) was not used to avoid compounding the error inherent in secondary calibration dates from other molecular rates analyses (Graur and Martin 2004). However, results for the age of the Simaroubaceae are discussed with reference to divergence dates from Muellner et al. (2007).

**Molecular rates analyses.**—Analyses for each calibration scheme (Table 2) were performed in BEAST v1.4.7 (Drummond and Rambaut 2007) using the various calibration points described above. Each analysis was run using the 4-gene, 62-species data set plus outgroups. A combined data set of the 4 coding markers, with the exclusion of the noncoding *trnK* intron, used GTR+I+G as the model of nucleotide substitution, determined with the AIC in Modeltest v3.6 (Posada and Crandall 1998). The noncoding intron was excluded because it was more difficult to align than the coding region (Clayton et al. 2007) and contained large indels in some species; misalignment could cause discrepancies in branch lengths that in turn would affect estimates of molecular rates along those branches. The coding regions were partitioned into 3 codon positions, based on results of partitioning strategies examined in Clayton et al. (2007; partitioning by gene region is not currently implemented in BEAUti v1.4.7, Drummond and Rambaut 2007). Both the substitution model and rate heterogeneity were unlinked across codon positions. Priors other than the calibration points were uniform and set at default values, and the tree root prior was set as a uniform distribution, 0–132 Myr (the estimated age of the angiosperms; Hughes 1994). In calibration scheme 2, this prior produced a tree with 0 likelihood and so was removed. Each BEAST analysis used a starting tree with initial branch lengths satisfying the upper and lower bounds for the calibration priors being used, constructed using pathd8 (Britton et al. 2007). An initial analysis was run using the fossil placement in calibration scheme 1 and no topological constraints, producing a tree with an identical topology to the original phylogenetic analysis (Fig. 2), except that *Quassia* was sister to Clade V (PP = 0.86; also seen in parsimony analyses of Clayton et al. 2007), and all outgroups formed a clade (PP =

0.41) sister to Simaroubaceae. We therefore constrained the topology of all further BEAST analyses to have *Samadera* as sister to Clade V, and Rutaceae sister to Simaroubaceae, as per previous Bayesian analyses (Fig. 1; Clayton et al. 2007). Although there is some evidence to suggest application of a relaxed molecular clock leads to improved phylogeny estimation (Drummond et al. 2006), we preferred to follow the more thorough phylogenetic analyses from the previous study (Clayton et al. 2007), particularly with the single outgroup clade seen in the unconstrained analysis. For each calibration scheme, 2 independent analyses ran for 10 000 000 generations each, sampling every 1000 generations. Stationarity of the MCMC was determined by comparing standard deviations of node times between the 2 independent runs, and observing burn-in plots for the posterior and all model parameter values in Tracer v1.4 (Rambaut and Drummond 2003). A burn-in of 1 000 000 generations was removed from each run. For all estimated parameters, an effective sample size (ESS) of >150 was obtained to ensure the parameter space had been sufficiently sampled (except in the case of calibration scheme 2, where 2 further runs were needed for an ESS >150 for all parameters). Once stationarity had been determined, the 2 independent runs for each scheme were combined using LogCombiner v1.4.7 (Drummond and Rambaut 2007). The calibration scheme ultimately selected for biogeographic analyses, of the 8 combinations of fossils tested (see Results), was generated a second time with 4 separate 10-million-generation runs and with scale operators adjusted to recommendations in the BEAST output from the original run for that calibration scheme (BEAST input .xml in Supplementary material, available from <http://www.sysbio.oxfordjournals.org/>). The 4 runs were combined using LogCombiner v1.4.7 (Drummond and Rambaut 2007), ensuring ESS >600 for all parameters.

#### *Biogeographic Analyses Using a Likelihood Approach*

Biogeographic data were compiled from species distributions in the literature (e.g., Cronquist 1944a, 1944b, 1944c, 1944d; Nooteboom 1962a, 1962b), and geographic information system mapping of specimen localities in ArcGIS v9.0 (<http://www.esri.com/software/arcgis/>). The major areas in which Simaroubaceae are distributed were categorized broadly into 7 areas for AAR (Fig. 1). Areas were delimited by continental divisions based on present and past separation of major landmasses,

TABLE 2. Divergence dates (in Ma) resulting from molecular rates analyses for 8 different calibration schemes. Values in bold are dates for calibrated nodes, using priors described in Table 1. Lower and upper 95% highest posterior densities are shown in brackets. The latter 4 nodes are examples taken from the phylogeny for comparative purposes

Calibration scheme	<i>Ailanthus</i> stem (A. confucii fossil)	<i>Ailanthus altissima</i> divergence (A. confucii fossil)	<i>Ailanthus integrifolia</i> divergence (A. gigas)	<i>Leitneria</i> divergence	<i>Picrasma</i> stem (Chaneya fossil)	<i>Picrasma</i> crown	Clade V crown	Family crown	<i>Quassia</i> crown
1	<b>55.1</b> [52.9, 58.1]	22.3 [14.5, 30.1]	17.5 [10.2, 25.1]	42.8 [36.8, 48.0]	68.6 [59.3, 79.3]	14.3 [9.4, 19.9]	20.9 [17.0, 25.4]	71.0 [62.7, 79.8]	13.6 [7.9, 19.9]
2	139.3 [95.0, 190.0]	<b>55.2</b> [52.9, 58.3]	43.3 [30.2, 54.0]	107.6 [71.1, 148.4]	173.6 [112.8, 236.8]	36.2 [20.2, 54.6]	53.0 [35.6, 73.7]	179.5 [118.1, 243.1]	34.1 [16.3, 52.2]
3	68.9 [53.7, 84.3]	35.0 [29.6, 41.9]	<b>30.4</b> [28.5, 33.5]	53.2 [40.5, 66.3]	84.3 [65.0, 104.0]	18.5 [11.1, 26.8]	26.3 [19.4, 33.4]	87.3 [67.9, 105.3]	17.0 [8.4, 25.6]
4	<b>55.8</b> [53.0, 59.7]	33.2 [29.2, 37.7]	<b>30.0</b> [28.5, 32.3]	43.3 [37.3, 49.2]	70.2 [58.4, 81.6]	15.0 [9.9, 21.0]	21.8 [17.9, 26.2]	72.7 [63.5, 83.4]	14.1 [7.9, 20.4]
5	43.1 [35.0, 53.9]	17.3 [10.6, 25.0]	13.5 [7.1, 20.4]	<b>33.3</b> [28.8, 40.4]	53.7 [41.6, 69.0]	11.1 [6.9, 16.0]	16.4 [12.4, 21.5]	55.5 [43.5, 70.3]	10.6 [5.8, 16.2]
6	35.2 [27.8, 42.9]	14.1 [8.8, 19.9]	11.0 [6.2, 16.3]	27.2 [20.9, 33.8]	<b>44.0</b> [40.5, 51.0]	9.2 [5.6, 13.1]	13.5 [10.2, 17.3]	45.4 [40.5, 53.6]	8.6 [4.8, 12.9]
7	<b>54.9</b> [52.8, 57.4]	21.9 [14.3, 29.8]	17.1 [10.3, 25.1]	<b>40.7</b> [34.3, 46.4]	67.4 [57.8, 77.8]	14.2 [9.2, 20.0]	20.6 [16.8, 24.6]	69.8 [62.1, 78.6]	13.4 [7.7, 19.4]
8	<b>55.3</b> [53.0, 58.4]	33.1 [29.1, 37.6]	<b>30.0</b> [28.5, 32.3]	<b>41.2</b> [34.5, 47.4]	69.1 [58.7, 81.5]	15.0 [9.6, 20.6]	21.6 [17.7, 26.2]	71.5 [63.0, 82.2]	13.9 [7.8, 20.4]

with the exception of Asia and Europe. Interpretation of the results, however, incorporates more specific geographic information, especially for island taxa. Figure 4 shows geographical distributions assigned to terminals for AAR, based on the following categories: 1) N—North and Central America (north of the Panamanian Isthmus), and Caribbean Islands; 2) S—South America; 3) F—Tropical Africa; 4) M—Madagascar; 5) A—Mainland Asia, including but not restricted to India, China, Bhutan, Myanmar, Japan, and Russia (fossil localities), plus insular southeast Asia, including but not restricted to Malaysia, Indonesia, and the Philippines; 6) U—Australia, New Guinea, Papua New Guinea, New Caledonia, and the Pacific Islands; 8) E—Europe (fossil localities). Widespread species were assigned to more than one area.

The program Lagrange (Ree and Smith 2008) was used for AAR. Unlike DIVA (Ronquist 1997), it incorporates an explicit model of dispersal routes available at historical intervals, and estimates vagility and extinction parameters, correlating stochastic events with lineage persistence, as part of the dispersal–extinction–cladogenesis (DEC) model (Ree and Smith 2008).

The model of dispersal route availability was developed based on geological history and the presence and dissolution of land bridges and island chains, and climatic data. In particular, geological events such as the presence of the North Atlantic Land Bridge, the Southern track between Australia and South America, island chains between North and South America, the collision of Australian and Asian plates, and the closure of the Panamanian Isthmus (Tiffney 1985a; Scotese 2001; Morley 2003) are accounted for in dispersal probabilities. Dispersal probabilities range from 0.1 for well-separated areas to 1.0 for contiguous landmasses (Table 3). Ancestral ranges are restricted to the following (based on potentially plausible and extant ranges): NS, NF, NE, NA, SU, FM, FA, FE, AU, AE and MAU.

Fossil areas were integrated by manually adding a terminal for *A. confucii* (coded as present in E, N and A) as the sister to extant *Ailanthus*, and a *Leitneria* terminal (coded as present in E and A) as sister to extant *Leitneria*. *Chaneya* was not included in the analysis. The nonoverlapping ranges of the fossil and extant lineages of *Leitneria* led us to place the fossil lineage as diverging from the extant lineage at 28.4 million years, the latest stratigraphic date for the Early Oligocene, when the fossils first appeared. *Ailanthus* fossils were more difficult to integrate, given the presence of fossils in Asia, where extant species also live, and the possibility of these fossils representing part of the *Ailanthus* stem lineage. We tested 3 different points of divergence for the fossil lineage: diverging at 52.2 million years, that is, 2.7 million years from the base of the stem branch, halfway along the stem branch, and 2 million years from the tip of the stem branch. All 3 produced near-identical likelihood scores and AARs (data not shown). Therefore, the divergence at 52.2 million years was chosen as this had the highest likelihood score of the 3 positions. Input files for Lagrange are



TABLE 3. Dispersal network model implemented in Lagrange, showing probabilities of dispersal between areas based on data from Tiffney (1985a), Morley (2003) and Scotese (2001). Probabilities are for movement in both directions

Area connections	Time period (Ma)	Dispersal probabilities
All overwater dispersals without island chains	70–0	0.1
Areas not adjacent between 70 Ma and present	70–0	0.1
North America–South America	70–45	0.25
	45–5	0.75
	5–0	1.0
North America–Europe	70–30	0.75
	30–0	0.1
North America–Mainland and southeast Asia	70–30	0.5
	30–0	0.1
North America–Africa	70–30	0.5
	30–0	0.1
South America–Australia	70–45	0.5
	45–0	0.1
Africa–Madagascar	70–0	0.5
	70–45	0.75
Africa–Mainland and southeast Asia	45–30	0.75
	30–0	0.1
	70–30	0.75
Africa–Europe	30–5	0.5
	5–0	0.25
	70–45	0.1
Mainland and southeast Asia–Australia	45–30	0.5
	30–0	0.75
	70–0	1.0
Mainland Asia–Europe	70–0	1.0

published in Supplementary material (available from <http://www.sysbio.oxfordjournals.org/>).

We optimized the model by examining the effects of different ancestral ranges between the Old World and New World (namely NA, NE, NF, and SU) on the likelihood score of the model. These potential migration routes are important to pantropical clades, and in Simaroubaceae, disjunctions between the Old World and New World are common throughout the family. Each ancestral range was tested alone, along with all possible combinations (Table 5). It is not clear how to compare likelihood scores between models with changes in the specified ancestral ranges because the models are not nested, and the additional parameters involved in having fewer ancestral ranges are fixed at a boundary value of 0. The resulting likelihood scores were therefore compared directly.

A DIVA analysis (Ronquist 1997) was also performed on the data to compare AARs between this event-based approach and the explicit model-based approach of Lagrange (Ree and Smith 2008). Tree topology and area assignments from Lagrange were also used in DIVA, which ran with default program parameters. Two analyses were conducted, with maxareas unlimited and constrained to 2.

## RESULTS

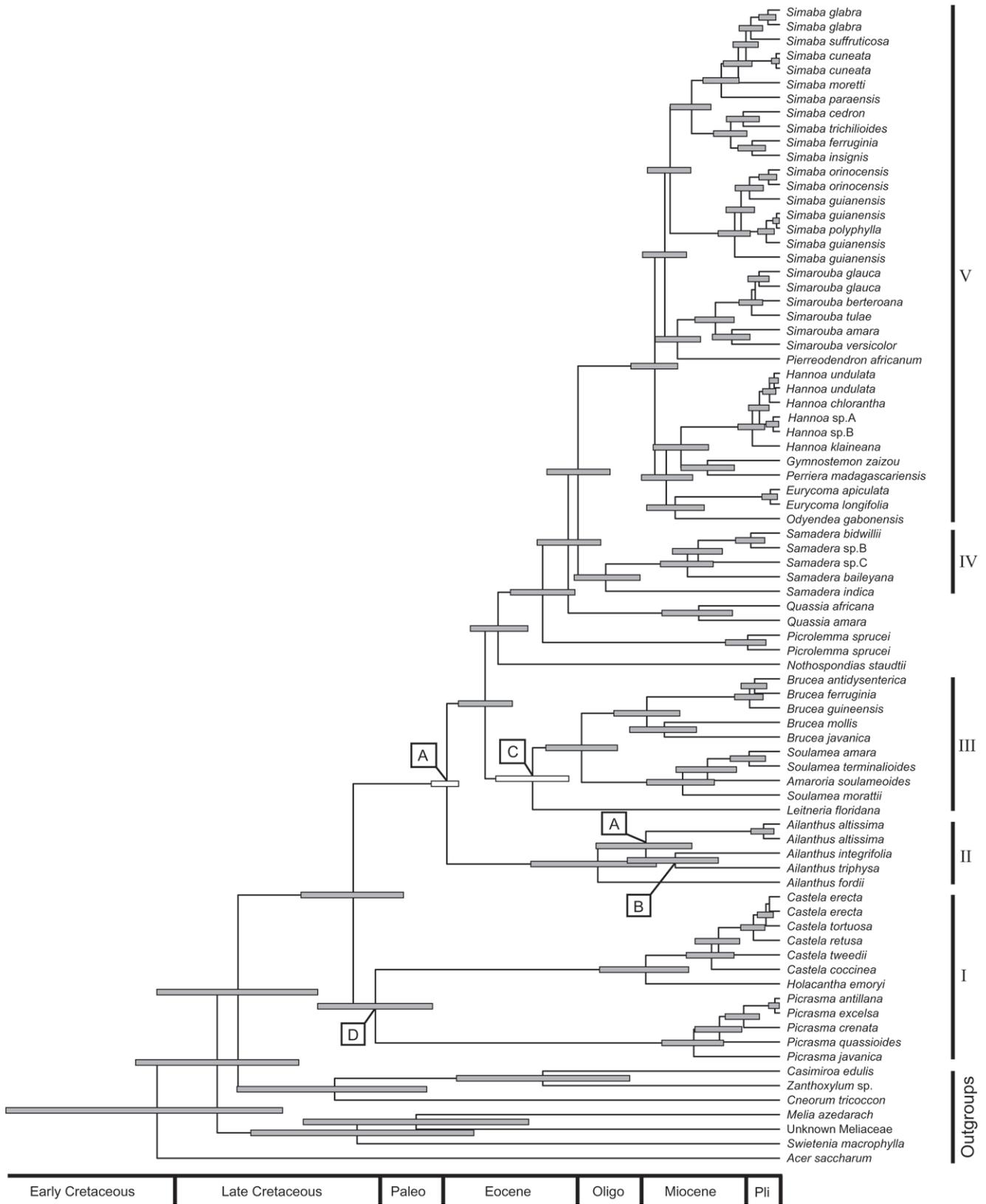
### *Divergence Date Estimation*

Tracer output statistics for BEAST analyses show the data are not evolving in a clock-like manner (co-

efficient of variation for calibration scheme 7 = 0.31 [0.22, 0.40]), and substitution rates between ancestor-descendant lineages are uncorrelated (covariance for calibration scheme 7 = 0.0071[−0.16, 0.15]). Therefore, the UCLN model of rate variation was the most appropriate method for estimating divergence dates as opposed to methods such as r8s (Sanderson 1997, 2002) and multidivtime (Thorne and Kishino 2002), which use an a priori assumption of autocorrelation between ancestor-descendant lineages. However, the covariance statistic appears to bias results against finding autocorrelation. Results of the different calibration schemes are shown with dates for fossil nodes as well as other example nodes from the tree (Table 2), and lower and upper 95% highest posterior densities are given in square brackets. Fossils of *Leitneria* and *Ailanthus confucii* (calibrating the *Ailanthus* stem group) corresponded reasonably well with each other (calibration schemes 1 and 5); therefore, these two fossils were used together (calibration scheme 7) to produce the chronogram for biogeographic analyses (Fig. 3). In calibration scheme 7, as well as in scheme 1, the divergence date of *Leitneria* was estimated to be older than the fossil dates. Using *A. confucii* at the divergence of *A. altissima* (scheme 2) produced very old dates throughout the tree, with the genus diverging from its sister clade at 139.3 [95.0,190.0] Ma. Use of *A. gigas* gave higher age estimates on its own (family crown group 87.3 [67.9,105.3] Ma), and in the final chronogram (calibration scheme 7) the divergence of *A. integrifolia* was considerably younger than the *A. gigas* fossil, hence its exclusion from the final analysis. *Chaneya* alone produced younger dates than the other calibrations, estimating the divergence of *Ailanthus* at 35.2 [27.8,42.9] Ma.

### *Biogeographic Analyses*

Lagrange results for the model including all Old World–New World ancestral areas were similar to DIVA with ancestral ranges restricted to two areas (Table 4). DIVA results showed unconstrained analyses had numerous equally optimal reconstructions at deeper nodes in the phylogeny. Optimal reconstruction required 30 dispersal events for the unconstrained analysis, and 32 dispersals when maxareas was limited to 2. The effects of different ancestral range combinations on likelihood reconstructions are shown in Table 5, ordered by decreasing likelihood score. Analyses including SU and NE typically had lower likelihoods and higher dispersal and extinction rates than other combinations. Although the likelihood scores could not be compared using an objective criterion such as the AIC (Akaike 1974), the top 4 ancestral range permutations had scores over 8 log likelihood units better than the next best model, and all produced near-identical AARs (Table 5), except for the root node. The combination with the highest likelihood was NA + NF, and this is shown in Figure 4. For clarity, the chronogram has been pruned of selected terminals that have a sister taxon in the same area. For



Downloaded from https://academic.oup.com/sysbio/article/58/4/395/1681797 by guest on 16 August 2022

FIGURE 3. Chronogram with 95% highest posterior density bars, based on BEAST analyses using two fossil calibrations, *Ailanthus confucii* and *Leitneria* sp., at nodes indicated with an open bar. Nodes labelled A, B, C, and D refer to the positions of the 4 fossils used to assess different calibration schemes (Table 1). Geological time scale (Gradstein and Ogg 2004) is shown at the bottom, and major clades of Simaroubaceae are indicated on the right-hand side.

TABLE 4. Results for biogeographic models tested, including likelihood scores (-lnL) and estimates of dispersal (D) and extinction (E) rates (events per Myr) for the likelihood method. In AARs, for example, nodes, areas shown indicate the range inherited by the daughter lineages. In cases where 2 ranges are separated by a bar, the first area is inherited by the upper branch on Figure 4, the second area is inherited by the lower branch. relative probability of each AAR is shown in parentheses for multiple reconstructions. For DIVA results, all equally optimal reconstructions are shown, separated by commas

Model	Constraints	-lnL	D	E	Examples of reconstructions for ancestral taxon of the following clades:						
					Whole family	Clade I	Sister clade to clade I	<i>Ailanthus</i>	Clade III	<i>Nothospondius</i>	Clade V
Likelihood	Ancestral ranges: NS, NF, NE, NA, SU, FM, FA, FE, AU, AE, MAU; dispersal probabilities fixed (table 3) Unconstrained	143.9	0.045	0.005	NA   N (0.31), N (0.28), A   NA (0.11), N   NA (0.07) <sup>a</sup>	N (0.65), N   NA (0.23)	NA   A (0.34), N   NA (0.21), N   A (0.10), A (0.09) <sup>a</sup>	A (0.51), A   NA (0.28), A   N (0.13)	A   NA	N (0.53), NF   F (0.18), F (0.10)	F   FA (0.63), N   A (0.14), NF   F (0.12)
DIVA 1	Unconstrained	—	—	—	NSFAUE	N, S, NA, SA, NSA	FAUE, NFAUE, SEFAUE, NSFAUE	A	UE, NUE, FUE, NFUE, AUE, NAUE, FAUE, NFAUE	F, SF	F, SF
DIVA 2	Maxareas = 2	—	—	—	A, NA, SA	N, S, NA, SA	A	A	A	F, SF	F, SF

<sup>a</sup>Further alternative reconstructions (not shown) are within 2 log likelihood units of the optimal reconstruction.

nodes where likelihood scores were not significantly different between multiple reconstructions within the confidence window of a 2 log likelihood unit difference (Edwards, 1992; Ree and Smith, 2008), the relative probability of the global likelihood is shown. Figure 4 shows 3 circumstances of range inheritance. First, there are instances of dispersal resulting in range expansion, which are common throughout the tree, especially in the Miocene. Second, there are local extinction events (3 are hypothesized), which are inferred when a daughter lineage inherits a different range from its parent (a range expansion prior to extinction is inferred). Third, there is vicariance by cladogenesis, where the ancestral range encompassing 2 or more areas subdivides between daughter lineages. Also seen in Figure 4 are numerous occurrences of incipient speciation, whereby an ancestral range of 2 areas subdivides, with 1 daughter inheriting the widespread ancestral range and 1 daughter inheriting just 1 of the 2 ancestral areas. An example is the root node of Clade III (Fig. 4).

DISCUSSION

*Divergence Date Estimation*

The different placements of the 3 fossil constraints strongly affected the divergence dates estimated, especially dates prior to the divergence of *Ailanthus*, that is, the divergence of *Picrasma*, *Castela*, and *Holacantha*, and the age of the family. Before the divergence of *Ailanthus*, however, only minor differences were seen between different fossil placements in Clades III, IV, and V. There are many sources of error in divergence time estimation (Sanderson 2002; Graur and Martin 2004; Near and Sanderson 2004). These include, but are not limited to: accuracy of dating fossil strata; correctly identifying fossils (e.g., the ambiguity around the *Chaneya* assignment); assigning fossil dates to the most appropriate nodes in the phylogeny; poorly supported nodes in the phylogeny; justifying the nature of the prior probability distribution of dates around the fossils; errors in estimating rates of molecular evolution, such as the problems inherent in decoupling rates and time, correctly estimating branch lengths on the topology, and assuming certain characteristics of molecular evolution (e.g., the UCLN distribution prior for substitution rates). However, these potential sources of error do not necessarily inhibit our ability to make general statements about migration patterns and dispersal events, in a system that encompasses large-scale and relatively slow processes, such as continental drift and land bridge formation and dissolution, over long periods of time.

Cross-calibration of fossils showed some compatibility between *A. confucii* and *Leitneria* fossils, but only when *A. confucii* was associated with the *Ailanthus* stem group. Calibration scheme 7 was chosen as the best chronogram for biogeographic analyses as it included the 2 fossil placements that had the best fossil records and showed some overlap in the probability distributions of node ages. In calibration schemes with

TABLE 5. Effects of Old World–New World ancestral range permutations, ordered by decreasing likelihood. Likelihood scores (–lnL), and estimates of dispersal (D) and extinction (E) rates (events per million years) are shown. For example nodes, AARs follow the format of Table 4; only the AAR with the highest relative probability is shown

Ancestral ranges	Examples of reconstructions for ancestral taxon of the following clades:									
	–lnL	D	E	Whole family	Clade I	Sister clade to Clade I	<i>Ailanthus</i>	Clade III	<i>Notiospondias</i>	Clade V
NA + NF	142.8	0.047	0.005	N (0.37)	N (0.74)	NA	A (0.55)	A   NA	N (0.63)	F   FA (0.61)
NA + NF + SU	142.9	0.047	0.005	N (0.37)	N (0.74)	NA	A (0.55)	A NA	N (0.63)	F   FA (0.61)
NA + NF + NE	143.8	0.045	0.005	NA   N (0.31)	N (0.66)	NA	A (0.51)	A NA	N (0.53)	F   FA (0.63)
All possible <sup>a</sup>	143.9	0.045	0.005	NA   N (0.31)	N (0.65)	NA	A (0.51)	A NA	N (0.53)	F   FA (0.63)
NF + NE + SU	152.0	0.072	0.011	N   S (0.44)	S (0.48)	NE	E (0.31)	E   NE (0.66)	N (0.53)	F   FA (0.48)
NF + NE	154.4	0.074	0.013	N (0.37)	N (0.58)	NE	E (0.34)	E   NE (0.62)	N (0.47)	F   FA (0.44)
NA + NE	156.1	0.063	0.010	N (0.31)	N (0.63)	NA	A (0.40)	A   NA	N (0.20)	N   A (0.25)
NA + NE + SU	156.2	0.063	0.010	N (0.31)	N (0.62)	NA	A (0.40)	A NA (0.70)	N (0.20)	N   A (0.25)
NA	158.8	0.056	0.008	N (0.37)	N (0.74)	NA	A (0.59)	A NA (0.83)	A (0.53)	NA   A (0.46)
NA + SU	158.9	0.055	0.008	N (0.37)	N (0.73)	NA	A (0.59)	A NA (0.83)	A (0.53)	NA   A (0.45)
NE + SU	161.4	0.092	0.016	N   S (0.39)	S (0.44)	NE	E (0.33)	E   NE	E (0.34)	NE   E (0.49)
NE	163.3	0.095	0.018	N (0.44)	N (0.60)	NE	E (0.35)	E   NE	E (0.34)	NE   E (0.50)
NF + SU	170.6	0.072	0.011	A   U (0.31)	S (0.28)	NF	F (0.38)	F   NF (0.53)	F (0.47)	F   FA (0.48)
NF	177.5	0.069	0.012	F (0.29)	F (0.35)	F   FA	A (0.34)	F   NF (0.38)	F (0.63)	F   FA (0.49)
SU	206.0	0.121	0.022	S (0.24)	S (0.44)	U (0.48)	U (0.76)	U   SU (0.50)	U (0.51)	U   A

<sup>a</sup>Equivalent to model used in Table 4.

multiple calibrations, *A. confucii* at the *Ailanthus* stem lineage was the most influential calibration point. The older ages for the *Leitneria* divergence produced in calibration schemes 1, 4, 7, and 8 could be accounted for by the *Leitneria* sp. described by Nikitin (2006a) as questionably Priabonian, that is, older than the conservative lower bound at the end of the Early Oligocene. Although the *Chaneya* calibration (scheme 6) produced dates younger than those derived from the more extensive fossil records of *Ailanthus* and *Leitneria*, the mean dates produced for the *Picrasma* stem and crown in calibration scheme 7 were 67.4 [57.8,77.8] Ma and 14.2 [9.2,20.0] Ma, respectively, and the *Chaneya* fossils fall between these dates. Therefore, an association between *Chaneya* and *Picrasma* remains a possibility.

The final chronogram (Fig. 3) illustrates when the major clades of Simaroubaceae arose. Muellner et al. (2007) estimated the age of the Simaroubaceae crown group at 52 million years, as this was the most conservative placement for the *Ailanthus* fossil, given their limited taxon sampling. With the addition of *Picrasma*, *Castela*, and *Holacantha* in this study, the age of the crown group is estimated to be Maastrichtian. The mean age of the stem group (88.4 [75.9, 101.4] Ma) is difficult to compare to dates provided by Muellner et al. (2007), given that the sister relationship of the family remains poorly resolved.

### Biogeographic Analyses

**Likelihood models.**—Reconstructions have a tendency to become more ambiguous deeper in the phylogeny (the so-called widespread ancestor problem; Ree et al. 2005), which is seen in the number of ambiguous nodes in the AARs. This is an inevitable result given the amount of time that has passed since the Simaroubaceae first arose, and the extrapolations made from geographic distributions of the extant tips. The restriction of ancestral ranges to 1 or 2 (and in 1 case 3) areas helps to counterbalance the widespread ancestor problem (Ree et al. 2005) and is an assumption based on the low likelihood of an ancestral species having a range across 3 or more of the geographic areas described above. Given the extant species distributions, only *S. indica* has a presence in 3 areas. Specification of ancestral ranges was an influential factor in how ancestral areas were reconstructed, particularly the inclusion or exclusion of NA, NF, or NE. These differences indicate the critical nature of the path between the Old World and the New World, and which is more plausible for Simaroubaceae. Comparison of likelihood scores between NA-, NE-, NF-, or SU-constrained models (Table 5) shows that NA has the highest likelihood for a single ancestral range, and in combination NA and NF have the highest likelihood (Fig. 4). This suggests trans-Beringial migration to be an important determining factor in Simaroubaceae biogeography, more so than the North Atlantic Land Bridge (NALB). Exclusively trans-Atlantic disjunctions (excluding *Ailanthus* fossil species that could have migrated via Beringia or the

NALB) were only hypothesized for 3 nodes, 1 of which arose after the NALB was available (the root node of *Simaba*, *Simarouba*, and *Pierreodendron*), and 1 of which the timing is unknown (the dispersal of *Nothospondias* to Africa). Models that included NF and NE, but excluded NA, increased probability of migration via the NALB as opposed to Beringia a priori and may be more biologically realistic for tropical groups. However, the early-diverging clades of Simaroubaceae, including genera such as *Picrasma*, *Ailanthus*, and *Leitneria*, continue to maintain a temperate element, supporting the Beringial track. The Southern track (SU) produced significantly lower likelihoods than the other 3 ranges due to how disconnected South America and Australia were through much of the Tertiary. North America was typically the optimal reconstruction for the root node in the ancestral range permutations with the highest likelihoods. Ree and Smith (2008) note that in their analysis of *Psychotria*, widespread ranges are rapidly reduced to single areas by cladogenesis or local extinction. However, in Simaroubaceae, many nodes retain widespread distributions, especially deeper in the phylogeny. Most of these widespread distributions break down through vicariance, with few lineages going extinct locally (*Picrolemma* and *Nothospondias*). Whether these ancestral ranges are plausible over a significant period of time is debatable, but the assumption of trans-Atlantic and trans-Beringial disjunctions in the early history of the family has been suggested for other tropical angiosperm families. The likelihood model could be streamlined further by introducing stratified ancestral ranges, such that ancestral distributions can vary with different periods in the history of the group.

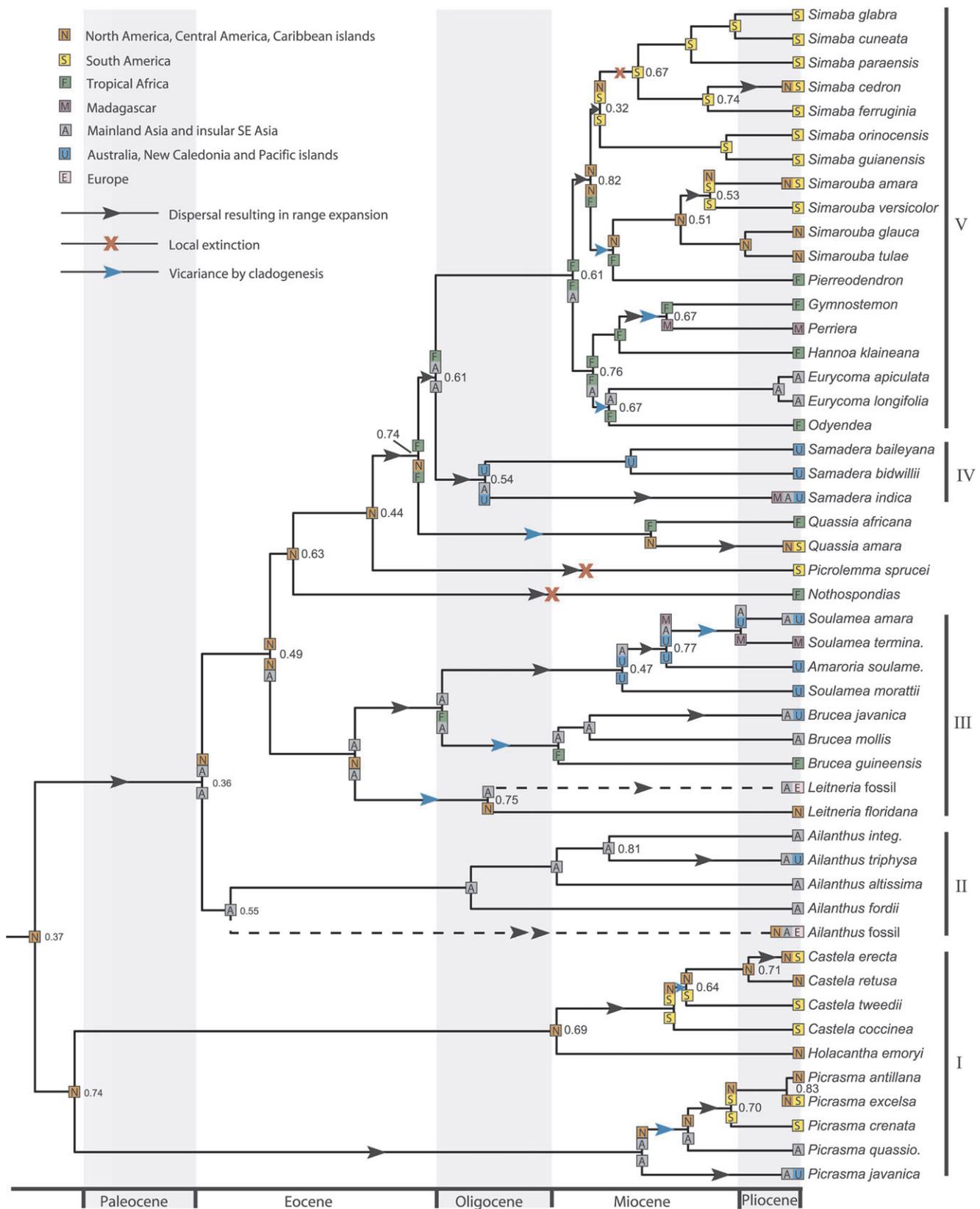
Ultimately we must assess whether the model-based approach is a valid one, given the ambiguity in the results and the number of model parameters that have been introduced to fit the data. In any likelihood analysis, there is a trade-off between a model too simplistic to describe the data accurately, and one that is overly complex, leading to overconfidence in one particular result. Ambiguity is seen in AARs for all models, and there are differences among ancestral range permutations (Table 5), especially for deeper nodes. This ambiguity is likely a reflection of a simplistic model attempting to describe the multitude of historical events that have shaped the family's history. Despite this limitation, there is an overall biogeographical signal concurrent among the different models tested. Optimal AARs are congruent among models for lineages diverging in the Late Oligocene and later, and the signal is one of multiple recent range shifts. These recent shifts appear to overshadow reconstruction of events deeper in the family's history, and as a result, Paleocene, Eocene, and Early Oligocene reconstructions are much more sensitive to the specification of the model. However, when alternative scenarios within each model have very similar likelihoods, the areas involved are typically the same. It is how these areas are inherited, whether through vicariance, incipient speciation, or dispersal followed by extinction, that distinguishes the alternative AARs. The

likelihood method encourages a more fluid approach to geographic range evolution and historical biogeography, a discipline that has traditionally been dominated by the vicariance–dispersal dichotomy.

DIVA produced results similar to Lagrange when ancestral ranges were confined to 2 areas, but South America was more commonly involved in DIVA AARs than in likelihood analyses (Table 4). Without the maxareas constraint, most deeper nodes in the phylogeny had implausibly widespread distributions. Some improbable ancestral ranges were observed, such as SA and FU, demonstrating the limitation of this method in a complex system. Also there is no objective method to favor one reconstruction over another at any particular node. Freedom from temporal constraints (Donoghue and Moore 2003) and limited ancestral ranges removes extinction as a factor in shaping distributions, as alternative vicariant events will always be “cheaper” in the DIVA cost matrix.

*Origin and early history of Simaroubaceae.*—Given the small size of Simaroubaceae (~109 spp.), which is more comparable with a large genus than previously studied tropical families (Melastomataceae: 3000 spp.; Malpighiaceae: 1200 spp.; Meliaceae: 550 spp.; Burseraceae: 500 spp.; Moraceae: 1500 spp. including temperate elements; Annonaceae: 2300 spp.; Dalbergioid legumes: 1100 spp. [species numbers from Lavin et al. 2000; Judd et al. 2007]), the complex reconstruction of biogeographical events proposed, and the resulting extant geographical heterogeneity, is perhaps more surprising. This is especially true given the origin of the family in the Late Cretaceous, a similar timeframe to many of the families mentioned above. Additionally, although attempts were made to include extinct species in AAR, further extinctions in the family are hypothesized, especially for depauperate and isolated lineages, such as *Nothospondias* (arising 46.7 [41.8, 51.5] Ma, a single species in Africa) and *Picrolemma* (diverging 39.1 [33.7, 44.4] Ma, 2 species in South America).

Based on the results of likelihood analyses, the family originated in North America, with expansion to Asia via Beringia early on in the Tertiary. This is hypothesized if we assume ancestral species were adapted to temperate climates, and several extant species from early-diverging clades retain this characteristic. However, if we assume ancestral species were tropical, we could rule out NA and SU as ancestral ranges, and the results for NF + NE (the highest likelihood ancestral range permutation without NA or SU) support movement via the NALB, which has been proposed for other tropical groups (Lavin et al. 2000; Renner et al. 2001; Davis et al. 2002; Richardson et al. 2004; Weeks et al. 2005; Zerega et al. 2005; Muellner et al. 2006). The origin of Clade I is North American only; however, the relationship of *Picrasma* to *Castela* and *Holacantha* is not strongly supported in phylogenetic analyses (Clayton et al. 2007), which could impact AAR for the earliest diverging nodes in the phylogeny. Clade III is



hypothesized to have originated in Asia and North America, with subsequent dispersals to Africa and southeast Asia. The presence of *Nothospondias* in Africa is the result of a migration via the NALB followed by local extinction in North America, although a later episode of long-distance dispersal cannot be ruled out. Similarly, *Picrolemma* dispersed to South America with subsequent local extinction, possibly via island chains, or potentially much later after the closure of the Panamanian Isthmus.

*Long-distance dispersal.*—The breakup of Gondwanaland, traditionally the most parsimonious explanation for tropical disjunctions (Raven and Axelrod 1974), has now been replaced by hypotheses incorporating knowledge of phylogenetic relationships, more realistic divergence dates, availability of dispersal routes (Morley, 2003), and a general acceptance of long-distance dispersal as a major driving force in extant plant distributions (Givnish and Renner 2004; Lavin et al. 2004; Renner 2004a; De Queiroz 2005). The African-Mesoamerican/South American disjunctions in Simaroubaceae mostly occurred after the NALB was viable for tropical groups (Tiffney 1985a). The ancestor of *Quassia amara* and *Quassia africana* supposedly maintained a trans-Atlantic disjunction for 22 million years, which underwent vicariance, an unlikely scenario without the NALB for much of that period. NF and NE ancestors are unlikely in the Miocene when tropical vegetation was restricted to lower latitudes; therefore, long-distance dispersal is the best explanation for these disjunctions, and dispersals between Africa and South America are hypothesized for a number of well-studied clades (Lavin et al. 2004; Renner 2004a; De Queiroz 2005). *Leitneria* may be another candidate for trans-Atlantic or trans-Pacific long-distance dispersal, but this is dependent on how the fossils are integrated, and if they are considered part of the extant lineage. In our interpretation, migration via Beringia is most likely, but the presence of fossils in Europe and Russia, with no definitive fossils in North America, might suggest a late arrival to North America via long-distance dispersal.

Disjunctions between mainland Asia and North America occur in a number of temperate genera (Gray 1859; Li 1952; Donoghue et al. 2001; Xiang and Soltis 2001; Donoghue and Smith 2004), with phylogenetic relationships between the 2 areas showing a variety of patterns. The disjunction seen in *Picrasma*, between 2

Asian/southeast Asian species and 6 Central and South American and Caribbean species, is less common (e.g., *Magnolia*, Azuma et al. 2001; *Hedyosmum*, Zhang and Renner 2003; *Ehretia*, Gottschling et al. 2004). Reconstructions suggest *Picrasma* arrived in Asia from North America, but the timing of this migration is unknown, as ancestral *Picrasma* could have maintained a North American Asian ancestral range beginning at any time along the stem branch of the genus. This ancestral distribution may be reflected in fossils of *Chaneya*, which are widespread across the Northern Hemisphere (Fig. 1). Trans-Beringial migration is plausible for cool-adapted species (Donoghue et al. 2001; Sanmartin et al. 2001), and *Picrasma quassioides* is a temperate species that occurs as far north as northern Japan. Range expansion into the Asian tropics may have prompted the incipient speciation of *Picrasma javanica*, with the ancestral NA range eventually undergoing a vicariant split ~9.9 [6.1, 14.0] Ma, and the New World species moving south. Alternatively, a long-distance dispersal event by the ancestor of *P. quassioides* could be inferred, but this would be more plausible if it were adapted to a climate similar to that inhabited by the New World species.

Dispersal between Africa and Asia has been proposed for clades such as Crypteroniaceae (Conti et al. 2002) under the hypothesis of rafting on the Indian subcontinent, and *Exacum* (Yuan et al. 2005) and Melastomataceae (Renner 2004b) through long-distance dispersal. *Eurycoma*, nested within the African species, and sister clades *Brucea* and *Soulamea*, fit with the idea of long-distance dispersal, given the significant ocean barrier between Africa and mainland and southeast Asia during the Oligocene (Scotese 2001; Morley 2003), coupled with climatic cooling. Long-distance dispersal to Madagascar has been suggested for many taxa, both animals (Poux et al. 2005) and plants (e.g., *Adansonia*, Baum et al. 1998; *Nepenthes*, Meimberg et al. 2001; Melastomataceae, Renner 2004b; *Exacum*, Yuan et al. 2005). Simaroubaceae are represented on Madagascar by 2 species, both of which arrived there by long-distance dispersal in the Miocene. With the presence of numerous island chains among mainland Asia, the southeast Asian islands and Australia during later epochs (Meimberg et al. 2001; Cannon and Manos 2003; Morley 2003), island-hopping dispersal explains the widespread distributions of species of *Ailanthus*, *Picrasma*, *Brucea*, and *Samadera*. *Soulamea*, like a number of other genera, for example, New Caledonian Sapotaceae

FIGURE 4. Ancestral area reconstruction for Simaroubaceae using Lagrange, with ancestral ranges NS, NA, NF, FM, FA, FE, AU, AE, and MAU, and stratified dispersal probabilities between areas (Table 3). AARs with highest likelihood are shown as colored boxes at each node. Single area boxes indicate an ancestor confined to a single geographic area; combined boxes indicate an ancestor with a distribution encompassing 2 or more areas; 2 boxes separated by a space indicate the ancestral ranges inherited by each of the daughter lineages arising from the node. For nodes with alternative reconstructions (within 2 log likelihood units of the maximum), the relative probability of the global likelihood for the optimal reconstruction is given. Three modes of range inheritance (range expansion, local extinction, and vicariance by cladogenesis) are indicated as symbols on branches in the phylogeny (see Results). For clarity, selected sister terminals from the same single area have been pruned from the chronogram. Fossil lineages are shown with a dashed line. Geological time scale (Gradstein and Ogg 2004) is shown at the bottom, and major clades of Simaroubaceae are indicated on the right-hand side. Area abbreviations are as follows: N = North and Central America (orange); S = South America (yellow); F = Africa (green); M = Madagascar (purple); A = Mainland and insular southeast Asia (gray); U = Australia (blue); E = Europe (pink).

(Bartish et al. 2005), *Cyrtandra* (Cronk et al. 2005), and *Weinmannia* (Bradford 2002), has readily dispersed between New Caledonia and surrounding islands, notably Fiji and Tuvalu, reaching as far as the Seychelles (*Soulamea terminalioides*).

In the New World, *Castela* and *Holacantha* show evidence of multiple dispersals between North and Central and South America, with speciation on several Caribbean islands and the Galápagos. This pattern reflects the amphitropical disjunctions of other dry-adapted clades such as *Tiquilia* subgenus *Tiquilia* (Moore et al. 2006) and *Hoffmannseggia* (Simpson et al. 2005). New World disjunctions are also seen in *Simarouba* and *Picrasma*, and both genera have at least 3 species distributed on Caribbean islands, implying one or more overwater dispersal events within the last 10 million years (Lavin et al. 2003; Santiago-Valentin and Olmstead 2004; Morris et al. 2007).

There is no obvious link between dispersability and diaspore size in Simaroubaceae, especially in the tropical clades, which tend to have drupaceous, bird- or mammal-dispersed fruits (e.g., Hardesty et al. 2005). The fruits of *Castela*, *Picrasma*, and *Simarouba* are small, bird-dispersed drupes, and so north-south dispersal may be facilitated by the migratory patterns of fruit-eating birds in the New World. Similarly, dispersal of the large fleshy drupes of *Gymnostemon* and *Perriera* between mainland Africa and Madagascar, may involve migrating birds (Renner 2004b). Wind-dispersed *Ailanthus* was widespread across the Northern Hemisphere based on fossil evidence (Corbett and Manchester 2004), and extant *A. altissima* is a weedy species well known for its dispersability (Benvenuti, 2007). The fruits of *Soulamea* and *Samadera indica* are dry, with an air cavity allowing them to float (Nooteboom 1962b), which potentially explains the widespread distribution of these taxa around the Indian Ocean basin, in conjunction with ocean currents, such as monsoon circulation seen today (Schott and McCreary 2001). *Leitneria*, which inhabits swamp and coastal forests, also has floating fruits, which could have facilitated an oceanic migration to the New World from Europe or Asia.

#### SUPPLEMENTARY MATERIAL

Supplementary material can be found at: <http://www.sysbio.oxfordjournals.org>.

#### FUNDING

Funding was provided by the National Science Foundation (angiosperm AtoL ÉF-0431266 to D.E.S., P.S.S. et al.; DDIG DEB-0710202 to J.W.C. and D.E.S.), Botanical Society of America, and American Society of Plant Taxonomists.

#### ACKNOWLEDGEMENTS

We thank Rick Ree (Field Museum, Chicago, USA) and Stephen Smith (Yale University, USA) for help with

Lagrange, Alexei Drummond (University of Auckland, New Zealand) for comments regarding BEAST, Aleksej Hvalj (Komorov Institute, St. Petersburg, Russia) for providing information on *Leitneria* fossils, Steve Manchester (FLMNH), Susanne Renner and 2 anonymous reviewers.

#### REFERENCES

- Akaike H. 1974. A new look at statistical model identification. *IEEE Trans. Automat. Contr.* 19:716–723.
- Azuma H., Garcia-Franco J.G., Rico-Gray V., Thien L.B. 2001. Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *Am. J. Bot.* 88:2275–2285.
- Barker G.M. 2002. Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. *Bot. J. Linn. Soc.* 76:165–194.
- Bartish I.V., Swenson U., Munzinger J., Anderberg A.A. 2005. Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): molecular evidence for generic polyphyly and repeated dispersal. *Am. J. Bot.* 92:667–673.
- Basak R.K. 1963. Pollen morphology of Indian Simaroubaceae. *Bull. Bot. Surv. India* 5:381–397.
- Basak R.K. 1967. Studies on the pollen morphology of the Simaroubaceae. *Bull. Bot. Surv. India* 9:63–67.
- Baum D.A., Small R.L., Wendel J.F. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Syst. Biol.* 47:181–207.
- Benvenuti S. 2007. Weed seed movement and dispersal strategies in the agricultural environment. *Weed Biol. Manag.* 7:141–157.
- Bradford J.C. 2002. Molecular phylogenetics and morphological evolution in Cunoniaceae (Cunoniaceae). *Ann. Mo. Bot. Gard.* 89:491–503.
- Britton T., Anderson C.L., Jacquet D., Lundqvist S., Bremer K. 2007. Estimating divergence times in large phylogenetic trees. *Syst. Biol.* 56:741–752.
- Cannon C.H., Manos P.S. 2003. Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). *J. Biogeogr.* 30:211–226.
- Clayton J.W., Fernando E.S., Soltis P.S., Soltis D.E. 2007. Molecular phylogeny of the Tree-of-Heaven family (Simaroubaceae) based on chloroplast and nuclear markers. *Int. J. Plant Sci.* 168:1325–1339.
- Conti E., Eriksson T., Schönenberger J., Sytsma K.J., Baum D.A. 2002. Early tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. *Evolution* 56:1931–1942.
- Corbett S., Manchester S. 2004. Phytogeography and fossil history of *Ailanthus* (Simaroubaceae). *Int. J. Plant Sci.* 165:671–690.
- Cronk Q.C.B., Kiehn M., Wagner W.L., Smith J.F. 2005. Evolution of *Cyrtandra* (Gesneriaceae) in the Pacific Ocean: the origin of a super-tramp clade. *Am. J. Bot.* 92:1017–1024.
- Cronquist A. 1944a. Studies in the Simaroubaceae I: the genus *Castela*. *J. Arnold Arbor.* 25:122–128.
- Cronquist A. 1944b. Studies in the Simaroubaceae II: the genus *Simarouba*. *Bull. Torrey Bot. Club* 71:226–234.
- Cronquist A. 1944c. Studies in the Simaroubaceae III: the genus *Simaba*. *Lloydia* 7:81–92.
- Cronquist A. 1944d. Studies in the Simaroubaceae IV: resume of American genera. *Brittonia* 5:128–147.
- Davis C.C., Bell C.D., Mathews S., Donoghue M.J. 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *Proc. Natl. Acad. Sci. U.S.A.* 99:6833–6837.
- Davis C.C., Webb C.O., Wurdack K.J., Jaramillo C.A., Donoghue M.J. 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. *Am. Nat.* 165: E36–E65.
- De Queiroz A. 2005. The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* 20:68–73.
- Donoghue M.J., Bell C.D., Li J. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. *Int. J. Plant Sci.* 162:541–552.
- Donoghue M.J., Moore B.R. 2003. Toward an integrative historical biogeography. *Integr. Comp. Biol.* 43:261–270.

- Donoghue M.J., Smith S.A. 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philos T. Roy. Soc. B* 359:1633–1644.
- Donoghue P.C.J., Benton M.J. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends Ecol. Evol.* 22: 424–431.
- Dorofeev P.I. 1994. *Leitneria*. In: Budantsev L., editor. Fossil flowering plants of Russia and adjacent states. Vol 3. Leitneriaceae-Juglandaceae. St. Petersburg (Russia): Kamarov Botanical Institute, Russian Academy of Sciences. p. 8–12, plates 47–55.
- Drummond A., Rambaut A. 2007. BEAST: bayesian evolutionary analysis sampling trees. *BMC Evol. Biol.* 7:214.
- Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
- Edwards A.W.F. 1992. Likelihood. Baltimore (MD): John Hopkins University Press.
- Faith D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61:1–10.
- Givnish T., Renner S. 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *Int. J. Plant Sci.* 165:S1–S6.
- Gottschling M., Diane N., Hilger H.H., Weigend M. 2004. Testing hypotheses on disjunctions present in the primarily woody Boraginales: Ehretiaceae, Cordiaceae, and Heliotropiaceae, inferred from ITS1 sequence data. *Int. J. Plant Sci.* 165:S123–S135.
- Gradstein F.M., Ogg J.G. 2004. Geologic time scale 2004—why, how, and where next! *Lethaia* 37:175–181.
- Graur D., Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20:80–86.
- Gray A. 1859. Diagnostic characters of phanerogamous plants collected by Charles Wright, botanist of the U.S. North Pacific Exploring Expedition, with observations upon the relations of the Japanese flora to that of North northern temperate zone. *Mem. Am. Acad. Arts Sci. N.S.* 6:377–453.
- Hardesty B.D., Dick C.W., Kremer A., Hubbell S., Bermingham E. 2005. Spatial genetic structure of *Simarouba amara* Aubl. (Simaroubaceae), a dioecious, animal-dispersed Neotropical tree, on Barro Colorado Island, Panama. *Heredity* 95:290–297.
- Hughes N.F. 1994. The enigma of angiosperm origins. Cambridge (UK): Cambridge University Press.
- Judd W.S., Campbell C.S., Kellogg E.A., Stevens P.F., Donoghue M.J. 2007. Plant systematics: a phylogenetic approach. 3rd ed. Sunderland (MA): Sinauer Associates Inc.
- Lavin M., Schrire B.P., Lewis G., Pennington R.T., Delgado-Salinas A., Thulin M., Hughes C.E., Matos A.B., Wojciechowski M.F. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philos T. Roy. Soc. B* 359:1509–1522.
- Lavin M., Thulin M., Labat J.-N., Pennington R.T. 2000. Africa, the odd man out: molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. *Syst. Bot.* 25:449–467.
- Lavin M., Wojciechowski M.F., Gasson P., Hughes C.H., Wheeler E. 2003. Phylogeny of robinoid legumes (Fabaceae) revisited: *Course-tia* and *Gliricidia* circumscribed, and a biogeographical appraisal of the Caribbean endemics. *Syst. Bot.* 28:387–409.
- Li H.-L. 1952. Floristic relationships between eastern Asia and eastern North America. *Trans. Am. Phil. Soc.* 42:371–429.
- Machen J. 1971. Plant microfossils from Tertiary deposits of the Isle of Wight. *New Phytol.* 70:851–872.
- Meimberg H., Wistuba A., Dittrich P., Heubl G. 2001. Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequence data. *Plant Biol.* 3:164–175.
- Moncada M., Machado S. 1987. Los granos de polen de Simarubaceae. *Acta. Bot. Cub.* 45:1–7.
- Moore M.J., Tye A., Jansen R.K. 2006. Patterns of long-distance dispersal in *Tiquilia* subg. *Tiquilia* (Boraginaceae): implications for the origins of amphitropical disjuncts and Galapagos Islands endemics. *Am. J. Bot.* 93:1163–1177.
- Morley R.J. 2003. Interplate dispersal paths for megathermal angiosperms. *Perspect. Plant Ecol. Evol. Syst.* 6:5–20.
- Morris A.B., Bell C.D., Clayton J.W., Judd W.S., Soltis D.E., Soltis P.S. 2007. Phylogeny and divergence time estimation in *Illicium* with implications for New World biogeography. *Syst. Bot.* 32: 236–249.
- Muellner A.N., Savolainen V., Samuel R., Chase M.W. 2006. The mahogany family “out-of-Africa”: divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. *Mol. Phylogenet. Evol.* 40:236–250.
- Muellner A.N., Vassiliades D.D., Renner S. 2007. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. *Plant Syst. Evol.* 266:233–252.
- Near T.J., Sanderson M.J. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philos T. Roy. Soc. B* 359:1477–1483.
- Nikitin V.P. 2006. [Paleocarpology and stratigraphy of the Paleogene and Neogene strata in Asian Russia]. Novosibirsk (Russia): Academic Publishing House Geo. Russian.
- Nooteboom H.P. 1962a. Generic delimitation in Simaroubaceae Tribus Simaroubeae and a conspectus of the genus *Quassia* L. *Blumea* 11:509–528.
- Nooteboom H.P. 1962b. Simaroubaceae. *Flora Malesiana Ser. I* 6: 193–226.
- Pennington R.T., Cronk Q.C.B., Richardson J.E. 2004a. Introduction and synthesis: plant phylogeny and the origin of major biomes. *Philos T. Roy. Soc. B* 359:1455–1464.
- Pennington R.T., Richardson J.E., Lavin M. 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytol.* 172:605–616.
- Posada D., Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Poux C., Madsen O., Marquard E., Vieites D.R., de Jong W.W., Vences M. 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. *Syst. Biol.* 54:719–730.
- Rambaut A., Drummond A.J. 2003. Tracer. Version 1.3. Available from: URL <http://evolve.zoo.ox.ac.uk/>.
- Raven P.H., Axelrod D.I. 1974. Angiosperm biogeography and past continental movements. *Ann. Mo. Bot. Gard.* 61:539–673.
- Ree R.H., Moore B.R., Webb C.O., Donoghue M.J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.
- Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14.
- Renner S.S. 2004a. Multiple miocene Melastomataceae dispersal between Madagascar, Africa and India. *Philos T. Roy. Soc. B* 359: 1485–1494.
- Renner S.S. 2004b. Plant dispersal across the tropical Atlantic by wind and sea currents. *Int. J. Plant Sci.* 165:S23–S33.
- Renner S.S., Clausing G., Meyer K. 2001. Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *Am. J. Bot.* 88:1290–1300.
- Richardson J.E., Chatrou L.W., Mols J.B., Erkens R.H.J., Pirie M.D. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philos T. Roy. Soc. B* 359:1495–1508.
- Riddle B.R. 1996. The molecular phylogeographic bridge between deep and shallow history in continental biotas. *Trends Ecol. Evol.* 11:207–211.
- Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46:195–203.
- Sanderson M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* 14:1218–1231.
- Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.
- Sanmartin I., Enghoff H., Ronquist F. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linn. Soc. Lond.* 73:345–390.
- Santiago-Valentin E., Olmstead R.G. 2004. Historical biogeography of Caribbean plants: introduction to current knowledge and possibilities from a phylogenetic perspective. *Taxon* 53:299–319.

- Schott F.A., McCreary J.P. 2001. The monsoon circulation of the Indian Ocean. *Prog. Oceanogr.* 51:1–123.
- Scotese C.R. 2001. Atlas of earth history. PALEOMAP project. Arlington (TX). Available from: URL <http://www.scotese.com>.
- Simpson B.B., Tate J.A., Weeks A. 2005. The biogeography of *Hoffmannseggia* (Leguminosae, Caesalpinoideae, Caesalpinieae): a tale of many travels. *J. Biogeogr.* 32:15–27.
- Song Z.-C., Wang W.-M., Huang F. 2004. Fossil pollen records of extant angiosperms in China. *Bot. Rev.* 70:425–458.
- Swofford D.S. 2002. Phylogenetic analysis using parsimony. Version 4.0b10. Sunderland (MA): Sinauer Associates Inc.
- Teodoridis V., Kvacek Z. 2005. The extinct genus *Chaneya* Wang et Manchester in the Tertiary of Europe—a revision of *Porana*-like fruit remains from Öhningen and Bohemia. *Rev. Palaeobot. Paly-nol.* 134:85–103.
- Thorne J.L., Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51:689–702.
- Tiffney B.H. 1985a. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the northern hemisphere. *J. Arnold Arbor.* 66:243–273.
- Wang Y., Manchester S.R. 2000. *Chaneya*, a new genus of winged fruit from the Tertiary of North America and Eastern Asia. *Int. J. Plant Sci.* 161:167–178.
- Weeks A., Daly D.C., Simpson B.B. 2005. The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. *Mol. Phylogenet. Evol.* 35:85–101.
- Wiens J.J., Donoghue M.J. 2004. Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* 19:639–644.
- Wikström N., Savolainen V., Chase M.W. 2001. Evolution of the angiosperms: calibrating the family tree. *Proc. R. Soc. B* 268:2211–2220.
- Wikström N., Savolainen V., Chase M.W. 2004. Angiosperm divergence times: congruence and incongruence between fossils and sequence divergence estimates. In: Donoghue P.C.J., Smith M.P., editors. *Telling the evolutionary time: molecular clocks and the fossil record*. London: Taylor and Francis. p. 142–165.
- Xiang Q.Y., Soltis D.E. 2001. Dispersal-vicariance analysis of intercontinental disjuncts: historical biogeographical implications for angiosperms in the Northern Hemisphere. *Int. J. Plant Sci.* 162: S29–S39.
- Yuan Y.M., Wohlhauser S., Möller M., Klackenberg J., Callmander M., Küpfer P. 2005. Phylogeny and biogeography of *Exacum* (Gentianaceae): a disjunctive distribution in the Indian Ocean basin resulted from long-distance dispersal and extensive radiation. *Syst. Biol.* 54:21–34.
- Zerega N.J.C., Clement W.L., Datwyler S.L., Weiblen G.D. 2005. Biogeography and divergence times in the mulberry family (Moraceae). *Mol. Phylogenet. Evol.* 37:402–416.
- Zhang L.-B., Renner S.S. 2003. The deepest splits in Chloranthaceae as resolved by chloroplast sequences. *Int. J. Plant Sci.* 164: S383–S392.

Received 6 March 2008; reviews returned 28 April 2008, 27 June 2009;  
accepted 30 June 2009

Associate Editor: Susanne Renner