Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae

Charles C. Davis[†‡], Charles D. Bell[§], Sarah Mathews[¶], and Michael J. Donoghue[§]

[†]Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, MA 02138; [§]Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208106, New Haven, CT 06520; and [¶]Division of Biological Sciences, 226 Tucker Hall, University of Missouri, Columbia, MO 65211

Explanations for biogeographic disjunctions involving South America and Africa typically invoke vicariance of western Gondwanan biotas or long distance dispersal. These hypotheses are problematical because many groups originated and diversified well after the last known connection between Africa and South America (≈105 million years ago), and it is unlikely that "sweepstakes" dispersal accounts for many of these disjunctions. Phylogenetic analyses of the angiosperm clade Malpighiaceae, combined with fossil evidence and molecular divergence-time estimates, suggest an alternative hypothesis to account for such distributions. We propose that Malpighiaceae originated in northern South America, and that members of several clades repeatedly migrated into North America and subsequently moved via North Atlantic land connections into the Old World during episodes starting in the Eocene, when climates supported tropical forests. This Laurasian migration route may explain many other extant lineages that exhibit western Gondwanan distributions.

Malpighiaceae contains trees, shrubs, and vines that are distributed widely in tropical and subtropical forests and savannas of the Old and New Worlds (ref. 1; Fig. 1). Approximately 85% of the species occur in the New World where they are pollinated by specialized oil-collecting bees that are absent from the Old World. This distribution has invited the development of alternative theories for the origin and diversification of Malpighiaceae. Vogel (2) proposed the "Gondwanian aborigine" hypothesis, in which the current distribution resulted from the break-up of western Gondwana (the supercontinent comprising Africa and South America). This hypothesis implies that the Malpighiaceae originated before western Gondwana divided [i.e., ≈105 mya (3)], as did several included lineages with disjunct distributions between the New and Old Worlds. In contrast, Anderson (1) and others (4) have favored an "American colonist" hypothesis, according to which Malpighiaceae originated in northern South America, in isolation from Africa, and dispersed eastward across the Atlantic at least twice. This hypothesis predicts that the ages of Malpighiaceae and of divergences between New World and Old World lineages are younger than the last known land connection between South America and Africa.

To elucidate the biogeographic history of Malpighiaceae, we used maximum likelihood to estimate the phylogeny for the group by using DNA nucleotide data from chloroplast *ndhF* and nuclear phytochrome C (*PHYC*) sequences (see *Materials and Methods*). This data set included sequences from one outgroup and 70 species of Malpighiaceae. To test these biogeographic hypotheses, it is necessary to attach a temporal dimension to the phylogeny to infer the timing (*i*) of the origin of Malpighiaceae and (*ii*) of the disjunctions between New and Old World lineages. Although molecular divergence estimates must be viewed with caution, these approaches may identify a window of time for branching events that will help select among competing hypotheses.

Materials and Methods

Our data set included 71 *ndhF* sequences and 71 *PHYC* sequences, which were readily aligned by eye. This data set included sequences from one outgroup and 70 species of Malpighiaceae representing the majority of genera, including multiple accessions from morphologically diverse or putatively non-monophyletic genera. Most *ndhF* sequences were previously obtained by using the methodology published in Davis, Anderson, and Donoghue (5) and in Davis (6). Two additional *ndhF* sequences were also generated for this study (GenBank accession nos. AF500495 and AF500496). The *PHYC* sequences are newly generated (GenBank accession nos. AF500522–AF500582), except for AF436794–AF436804 from Davis (6), and were obtained by using previously detailed (7) PCR, cloning, and sequencing procedures. We screened up to five clones from several species representing most of the major lineages within Malpighiaceae (GenBank accession nos. AF500497–AF500521) and found no evidence of duplication events within *PHYC*, consistent with previous findings by Mathews and Donoghue (7).

To assess the level of congruence between the *ndhF* and *PHYC* data sets, we used the incongruence length difference test (8) implemented in PAUP* V. 4.0B8 (9) for UNIX as the partition-homogeneity test. We used simple taxon addition (saving 10 trees per replicate), tree-bisection-reconnection branch-swapping, and heuristic searches with 999 repartitions of the data. The results (*P* = 0.55) indicated that *ndhF* and *PHYC* were congruent and so they were combined for further analysis. The combined data set included 1,833 aligned sites. Data matrices analyzed in this study are available from the first author and from TreeBASE (http://www.treebase.org).

Phylogenetic analyses by using maximum likelihood were conducted on the combined data set with PAUP*. Tree searches were conducted with 300 random sequence addition replicates and tree-bisection-reconnection branch swapping. We performed 300 bootstrap replicates to assess clade support. To choose a model of sequence evolution we performed likelihood ratio tests (10) with likelihood trees generated by using a series of models with increasing complexity. The GTR+I+Γ model had a higher likelihood than other models and was used to evaluate molecular rate constancy.

To infer the location of disjunctions between New and Old World lineages, ancestral areas were reconstructed by using dispersal-vicariance analysis (DIVA; ref. 11). DIVA reconstructs ancestral areas by minimizing the number of dispersal and extinction events needed to explain the distribution pattern. Vicariance is the default mode of speciation in DIVA, and such events are not counted as steps in identifying optimal solutions; inferred dispersal and extinction events are counted as one step each. Our data matrix used to assess ancestral areas was con-
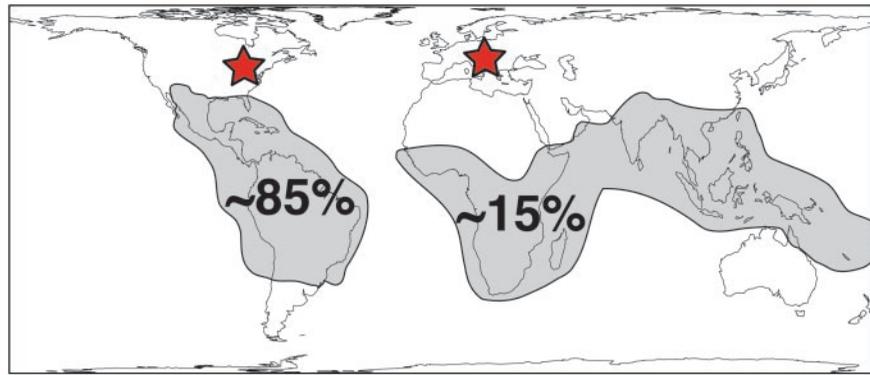
EVOLUTION

Fig. 1. Current geographic distribution of Malpighiaceae estimated from Arénes (22). Malpighiaceae, with ≈1,250 species (1), are most diverse in northern South America (22). The ≈180 Old World species, belonging to six lineages (see Fig. 2) are variously represented in Africa (47 species), Australasia (17 species), India (43 species), and Madagascar (80 species) [from Arénes (22)]. Red stars indicate fossil localities from Hungary and Slovenia (17), and Tennessee (21).

structed by scoring terminals for presence in either the New World or the Old World, and the analysis was carried out by using DIVA V. 1.1 [ref. 12; available by anonymous FTP from Uppsala University (ftp.uu.se or ftp.systbot.uu.se)]. The outgroup for Malpighiaceae was scored as New World based on the close phylogenetic relationship of the neotropical endemic *Whittonia guianensis* Sandw (Peridiscaceae) (13).

Branch lengths and an associated likelihood score were calculated for the respective models of sequence evolution in the absence of a molecular clock. We used trees resulting from maximum likelihood searches to test for rate constancy among lineages under the assumption of a molecular clock by using $2(-\ln L_1 - \ln L_0)$ as a test statistic. This test statistic was compared with a χ^2 distribution (with $n-2$ degrees of freedom; $n =$ number of taxa) to assess significance. Finding that a clock was rejected ($P < 0.05$), we used a nonparametric rate-smoothing algorithm (14) to estimate divergence times, focusing specifically on New World–Old World disjunctions inferred from our DIVA reconstructions.

To estimate standard errors associated with divergence dates, we used a parametric bootstrapping strategy similar to the three-step procedure of Baldwin and Sanderson (15). Their method was a nonparametric procedure to estimate the stochastic error associated with sampling a finite number of nucleotide characters. We adopted a parametric bootstrapping approach, as follows: (*i*) 100 data sets were simulated on the maximum likelihood tree with the computer software SEQ-GEN V. 1.2.3 (16); (*ii*) the resulting simulated data sets were imported into PAUP*, and divergence times were estimated on the original tree topology with the sequence model and parameters estimated from the original data; and (*iii*) the resulting branch length estimates from the simulated data sets were used to calculate the variance in divergence time estimates. These error estimates reflect the stochastic error associated with the molecular evolution process.

A fossil species of *Tetrapterys* recovered from the early Oligocene (33 mya; ref. 17) from Hungary and Slovenia provides a reliable divergence estimate for the *Tetrapterys* clade and was used to calibrate our tree. Several other excellent Malpighiaceae fossils are known from the Eocene and provide important information on its distribution in the past (see *Results and Discussion*). However, because these fossils are not readily assignable to any extant genus, they are difficult to include as calibration points.

Results and Discussion

Our phylogeny (Fig. 2) is robust and congruent with previous parsimony analyses based on four chloroplast genes (5, 18).

The basal split in Malpighiaceae separates the wholly New World subfamily Byrsonimoideae from a major clade containing both New and Old World species. Within this clade, the Old World Malpighiaceae belong to six separate lineages (Fig. 2), all of which contain species with winged-fruits. A fossil species of *Tetrapterys* recovered from the early Oligocene (33 mya; ref. 17) suggests that the divergence between *Tetrapterys* and its sister lineage had occurred by this time (Fig. 2; green star). Based on this assumption, the minimum age of Malpighiaceae is estimated to be 63.6 ± 5.8 mya, and the six New World-Old World divergences would have occurred between the early Eocene (55.1 ± 6.0 mya) and the middle Miocene (12.9 ± 0.85 mya). If, instead, we assume that the age of the fossil corresponds to the basal split within *Tetrapterys* (Fig. 2; yellow star), we obtain maximum age estimates that are ≈4 Myr older than the estimated minimum ages. Similar results are obtained by placing the fossil at less likely positions deeper in the tree.

Our divergence estimates indicate that Malpighiaceae originated well after the last known connection between Africa and South America (≈105 mya). These estimates parallel those reported by Magallón and Sanderson (19), which suggest that many major angiosperm lineages have radiated during more recent times in the Tertiary, and do not extend back to the Cretaceous. In fact, our fossil calibration points would have to be approximately 20 Myr older to make the ancestral node within Malpighiacecae consistent with a Gondwanan origin (i.e., 105 mya). Divergence estimates for the New World-Old World disjunctions are even younger, rendering the "Gondwanian aborigine" explanation untenable (Fig. 2; nodes 1–6).

In contrast, our divergence estimates are potentially consistent with a series of episodic long-distance dispersal events as suggested by Anderson (1) and others (4). However, our results indicate that the "American colonist" scenario would require at least six dispersal events across the Atlantic. The plausibility of this scenario is further diminished by excellent Eocene and Oligocene fossils of Malpighiaceae from several localities throughout the Northern Hemisphere and Africa: *Perisyncolporites* from northern South America and Nigeria (20), *Eoglandulosa* from Tennessee (21), and *Tetrapterys* from Hungary and Slovenia (17). Our phylogenetic results, together with the fossil evidence, suggest migration through Laurasia as a new explanation for the present distribution of Malpighiaceae.

These data imply that the Malpighiaceae originated and began to diversify in northern South America in the early Paleocene (≈64 mya). Anderson (1) also postulated a South American origin, which is consistent with the high levels of diversity and
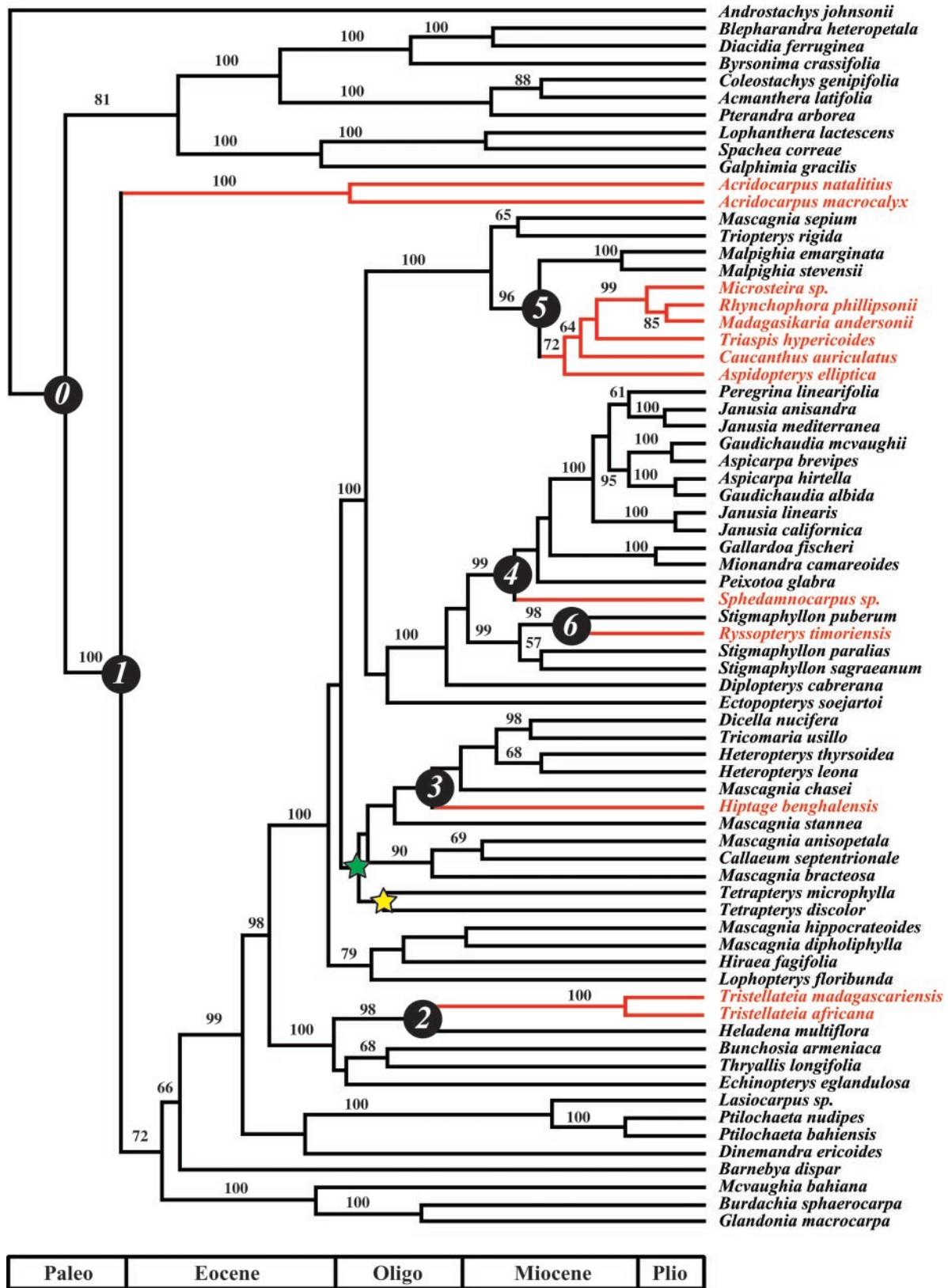
Fig. 2. Maximum likelihood tree topology ($-\ln L = 17{,}333.97$) for combined *ndhF* and *PHYC* data with bootstrap values ($>50\%$) indicated on branches. Old World lineages are shaded red. Divergence times were calculated on the rate-smoothed topology by calibrating the two nodes indicated by the colored stars with a fossil of *Tetrapterys* (33 mya; ref. 17). The origin of Malpighiaceae (node 0) is reconstructed as New World and is estimated at 63.6 ± 5.8 mya. The nodes labeled 1–6 correspond to the New/Old World disjunctions estimated from the DIVA reconstruction. Age estimates for nodes 1 through 6 (calibrated with the fossil placed at the green star) are, respectively: 55.1 ± 6.0 mya, 30.4 ± 2.6, 29.4 ± 2.1, 19.1 ± 1.5, 15.1 ± 1.2, and 12.9 ± 0.85. The scale bar indicates major Tertiary epochs (Paleo = Paleocene; Oligo = Oligocene; Plio = Pliocene).
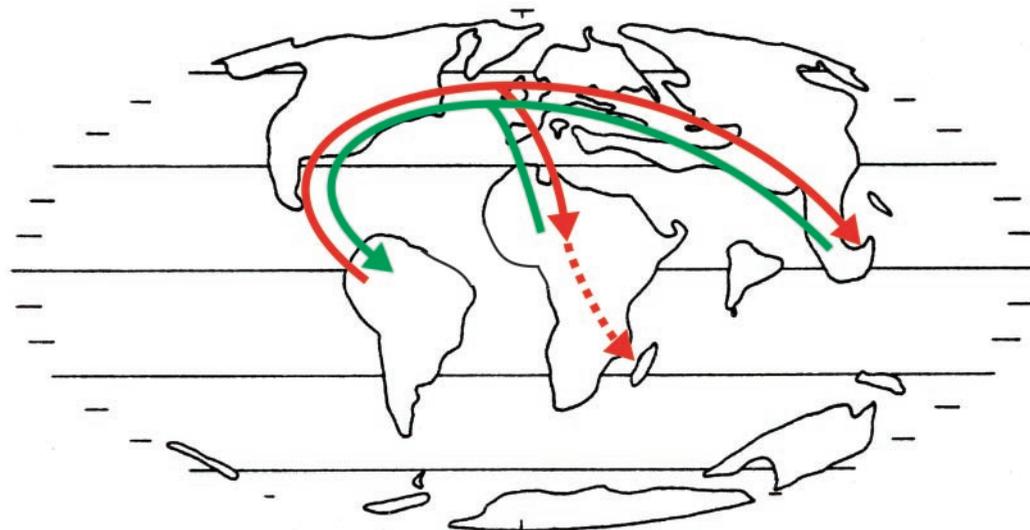
Fig. 3. Eocene paleogeographic map [after Parrish (41); modified from Doyle and Le Thomas (35)] summarizing hypothesized Laurasian migration of western Gondwana taxa. The distribution of Malpighiaceae is proposed to have resulted by migration from South America to the Old World (red arrow); green arrow shows a possible pattern of dispersion from the Old World to the New World taken by other tropical angiosperm clades.

endemism in Malpighiaceae of Guyana (22). From South America, various wing-fruited species may have migrated into North America via scattered continental and/or volcanic islands that connected North and South America at various times through the Tertiary, starting in the Eocene (23). This route apparently played a central role in facilitating the distribution of Antillean land-mammals (24), and would most likely have been available to Malpighiaceae during this timeframe. From North America, these lineages could migrate eastward across Laurasia into the Old World via a series of connections across the North Atlantic (the "North Atlantic Land Bridge"; ref. 25), subsequently diversifying in Africa (and later in Madagascar) and in some cases in Asia (see Fig. 3; red arrow). Populations in eastern North America, Europe, and other northern regions were presumably eliminated as tropical/subtropical climates disappeared from these areas during the Miocene and later epochs.

In addition to the "North Atlantic Land Bridge," Beringia facilitated the spread of the northern hemisphere biota throughout much of the Tertiary, and may represent an alternative migration pathway for Malpighiaceae. It apparently played an important role in determining the distribution of mammals of Asian origin that migrated to North America during periods of climatic warming in the late Palaeocene-early Eocene (26). Similarly, there may have been repeated dispersal of temperate and boreal plants through Beringia from the Eocene to the present (27). Although most of our estimates of divergence times indicate that Malpighiaceae could have migrated via Beringia, its high latitude (between 69° and 75° N; ref. 28) almost certainly prohibited its use by thermophilic lineages for most of this time, and fossil localities do not support this route for the dispersion of Malpighiaceae. Moreover, Old World Malpighiaceae are almost entirely restricted to Africa and Madagascar with a few, relatively minor, excursions into Asia. Together, paleoclimatic, phylogenetic, and fossil data are most consistent with an easterly migration across the Atlantic rather than a westerly movement via Beringia.

The North Atlantic pathway apparently served as a conduit for migration of several lineages of Malpighiaceae at different times. Dispersion episodes leading to nodes 1, 2, and 3 (Fig. 2) broadly coincide with Eocene/Oligocene thermal maxima and the dominance of a "boreotropical flora" across the North Atlantic (29). Significant cooling during the Oligocene (30) resulted in south-

ward retreats and the extirpation of some lineages comprising this flora (31–33). Warming during the Miocene resulted in the return of several thermophilic lineages to North America and Europe (28), probably including lineages of Malpighiaceae labeled 4, 5, and 6 (Fig. 2).

These findings have general implications for the origin of disjunctions between South America and Africa (and elsewhere in the Old World tropics). Vicariance resulting from the break-up of western Gondwana cannot explain those disjunct angiosperm lineages (of which there may be many) that originated and diversified after the last direct continental connection. Dispersal directly across the Atlantic Ocean, although perhaps feasible in some cases, seems unlikely to be a general explanation because vagility varies greatly among lineages. Additionally, dispersal through Antarctica into southern Africa seems unlikely in view of the ancient split between Africa and Australasia (162–165 mya; ref. 3). Moreover, none of these theories explain the existence of undisputed Laurasian fossils for some Gondwanan disjuncts. Dispersion through Laurasia at times during the Tertiary when climatic conditions supported tropical vegetation may provide the best explanation for many organisms that show the classic western Gondwanan disjunction pattern. Malpighiaceae are here interpreted as a case of iterative spread from South America to the Old World through Laurasia. Recent phylogenetic analyses of several other tropical angiosperm clades appear to provide examples of spread either in the same direction (Melastomeae within Melastomataceae; ref. 34) or in the opposite direction, from the Old to the New World [Annonaceae (35); Lauraceae (36); Fig. 3, green arrow]. Other groups may have originated within subtropical North America and migrated into the Old World, and perhaps much later into South America [dichrostachioid Leguminosae (37); dalbergioid Leguminosae (38)].

The existence of a "boreotropical" connection across the North Atlantic during the Eocene has long been viewed as a key to understanding patterns of disjunction around the Northern Hemisphere [e.g., close relationships between Eastern Asian and Eastern North American plants (25, 27, 28, 39)]. Our analysis indicates that this pathway may also have played an important role in explaining the global distribution of tropical groups. Were it not for such Laurasian dispersion,

many familiar pantropical clades might now be restricted to one hemisphere or the other. Conversely, the restriction of some major plant lineages to one hemisphere [e.g., Cactaceae in the New World (40)] may reflect their origin and radiation during a time when Laurasia was no longer a viable conduit for migration.

1. Anderson, W. R. (1990) *Mem. N. Y. Bot. Gard.* **64,** 210–224.
2. Vogel, S. (1990) *Mem. N. Y. Bot. Gard.* **55,** 130–142.
3. McLoughlin, S. (2001) *Aust. J. Bot.* **49,** 271–300.
4. Raven, P. H. & Axelrod, D. I. (1974) *Ann. Missouri Bot. Gard.* **61,** 539–673.
5. Davis, C. C., Anderson, W. R. & Donoghue, M. J. (2001) *Am. J. Bot.* **88,** 1830–1846.
6. Davis, C. C. (2002) *Am. J. Bot.* **89,** 723–730.
7. Mathews, S. & Donoghue, M. J. (1999) *Science* **286,** 947–950.
8. Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. (1994) *Cladistics* **10,** 315–319.
9. Swofford, D. L. (1999) PAUP* (Sinauer, Sunderland, MA), Version 4.0b8.
10. Felsenstein, J. (1981) *J. Mol. Evol.* **17,** 368–376.
11. Ronquist, F. (1997) *Syst. Biol.* **46,** 195–203.
12. Ronquist, F. (1996) DIVA (Uppsala Univ., Uppsala), Version 1.1.
13. Savolainen, V., Fay, M. F., Albach, D. C., Backlund, A., van der Bank, M., Cameron, K. M., Johnson, S. A., Lledó, M. D., Pintaud, J. C., Powell, M., *et al.* (2000) *Kew Bull.* **55,** 257–309.
14. Sanderson, M. J. (1997) *Mol. Biol. Evol.* **14,** 1218–1231.
15. Baldwin B. G. & Sanderson, M. J. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9402–9406.
16. Rambaut, A. & Grassly, N. C. (1997) *Comput. Appl. Biosci.* **13,** 235–238.
17. Hably, L. & Manchester, S. R. (2000) *Rev. Palaeobot. Palynol.* **111,** 93–101.
18. Cameron, K. M., Chase, M. W., Anderson, W. R. & Hills, H. G. (2001) *Am. J. Bot.* **88,** 1847–1862.
19. Magallón, S. & Sanderson, M. J. (2001) *Evolution* **55,** 1762–1780.
20. Germeraad, J. H., Hopping, C. A. & Muller, J. (1968) *Rev. Palaeobot. Palynol.* **6,** 189–348.
21. Taylor, D. W. & Crepet, W. (1987) *Am. J. Bot.* **74,** 274–286.
22. Arènes, J. (1957) *Compt. Rend. Sommaire Séances Soc. Biogéogr.* **290,** 81–108.
23. Iturralde-Vinent, M. A. & MacPhee, R. D. E. (1999) *Bull. Am. Mus. Nat. Hist.* **238,** 1–95.
24. MacPhee, R. D. E. & Iturralde-Vinent, M. A. (1995) *Am. Mus. Novit.* **3141,** 1–31.
25. Tiffney, B. H. (1985) *J. Arnold Arbor.* **66,** 243–273.
26. Beard, K. C. (1998) *Bull. Carnegie Mus. Nat. Hist.* **34,** 5–39.
27. Donoghue, M. J., Bell, C. D. & Li, J. (2001) *Int. J. Plant Sci.* **162,** S41–S57.
28. Tiffney, B. H. (1985) *J. Arnold Arbor.* **66,** 73–94.
29. Wolfe, J. A. (1975) *Ann. Missouri Bot. Gard.* **62,** 264–279.
30. Berggren, W. A. & Prothero, D. R. (1992) in *Eocene–Oligocene Climatic and Biotic Evolution*, eds. Prothero, D. R. & Berggren, W. A. (Princeton Univ. Press, Princeton), pp. 1–28.
31. Collinson, M. E. (1992) in *Eocene–Oligocene Climatic and Biotic Evolution*, eds. Prothero, D. R. & Berggren, W. A. (Princeton Univ. Press, Princeton), pp. 437–450.
32. Wolfe, J. A. (1992) in *Eocene–Oligocene Climatic and Biotic Evolution*, eds. Prothero, D. R. & Berggren, W. A. (Princeton Univ. Press, Princeton), pp. 421–436.
33. Manchester, S. R. (1999) *Ann. Missouri Bot. Gard.* **86,** 472–522.
34. Renner, S. S., Clausing, G. & Meyer, K. (2001) *Am. J. Bot.* **88,** 1290–1300.
35. Doyle, J. A. & Le Thomas, A. (1997) *Géogr. Phys. Quatern.* **51,** 353–361.
36. Chanderbali, A. S., van der Werff, H. & Renner, S. S. (2001) *Ann. Missouri Bot. Gard.* **88,** 104–134.
37. Lavin, M. & Luckow, M. (1993) *Am. J. Bot.* **80,** 1–14.
38. Lavin, M., Thulin, M., Labat, J.-N. & Pennington, R. T. (2000) *Syst. Bot.* **25,** 449–467.
39. Wen, J. (1999) *Ann. Rev. Ecol. Syst.* **30,** 421–455.
40. Nyffeler, R. (2002) *Am. J. Bot.* **89,** 312–326.
41. Parish, J. T. (1987) in *The Origins of Angiosperms and Their Biological Consequences*, eds. Friis, E. M., Chaloner, W. G. & Crane, P. R. (Cambridge Univ. Press, Cambridge, U.K.), pp. 51–73.

EVOLUTION