Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 37 (2005) 402–416
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Biogeography and divergence times in the mulberry family (Moraceae)

Nyree J.C. Zerega [a,1], Wendy L. Clement [a], Shannon L. Datwyler [b], George D. Weiblen [a,*]

[a] *Department of Plant Biology, University of Minnesota, 1445 Gortner Avenue, Saint Paul, MN 55108, USA*
[b] *California State University, Sacramento, 6000 J Street, Sacramento, CA 95819, USA*

Received 6 December 2004; revised 24 June 2005; accepted 3 July 2005
Available online 19 August 2005

Abstract

The biogeographical history of the mulberry family (Moraceae) was investigated using phylogenetic inferences from nuclear and chloroplast DNA, molecular dating with multiple fossil calibrations, and independent geological evidence. The Moraceae are centered in the tropics which has invited the hypothesis that the family has Gondwanan origins and extant distribution is the result of vicariance due to the break-up of Gondwana. However, the cosmopolitan distribution of Moraceae suggests a more complicated biogeographical history. The timing and location of Moraceae diversification also bears on the origin of the fig pollination mutualism, a model for the study of coevolution and specialization. Recent molecular dating of pollinating fig wasps suggested that an ancient Gondwanan origin coupled with vicariance and dispersal could account for the present day distribution of the mutualism. Here, we provide the first assessment of this hypothesis based on dating of figs and their relatives. Minimum age estimates suggest that the Moraceae had diversified by at least the mid-Cretaceous and major clades including the figs may have radiated during the Tertiary after the break-up of Gondwanaland. Molecular evidence together with Eurasian fossils suggest that the early diversification of Moraceae in Eurasia and subsequent migration into the southern hemisphere is at least as plausible as the Gondwanan hypothesis. These findings invite a reevaluation of the biogeography of fig pollination and highlight the need for incorporating multiple sources of evidence in biogeographical reconstructions.

Keywords: Bayesian analysis; Figs; Fig pollination; Molecular dating; Mutualism; Phylogeny

1. Introduction

Understanding the early evolutionary history of flowering plant families centered in the tropics may be aided by the integration of evidence from biogeography, paleobotany, climatology, and molecular phylogeny. By the time angiosperms originated in the early Cretaceous (Magallón and Sanderson, 2001), the supercontinent of Pangea had already split into Gondwanaland in the southern hemisphere and Laurasia in the north. Many extant plant lineages are distributed across these former landmasses, and vicariance and dispersal events have been invoked to support or reject scenarios of Gondwanan versus Laurasian origin (Bremer, 2002; Manos, 1997; Sequeira and Farrell, 2001).

The mulberry family (Moraceae) is cosmopolitan in distribution with the majority of species occurring in the Old World tropics, particularly Asia and the Indo-Pacific Islands (Berg, 2001; Rohwer, 1993). The distribution of Moraceae across former Gondwanan landmasses led Corner (1967) to hypothesize that disjunctions between extant lineages are the product of vicariance

* Corresponding author. Fax: +1 612 625 1738.
E-mail address: gweiblen@umn.edu (G.D. Weiblen).
[1] Present address: Northwestern University, Program in Plant Biology and Conservation, 2205 Tech Drive, Evanston, IL 60208 and Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, IL 60022, USA.

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.07.004

due to the break-up of Gondwanaland. An alternative hypothesis of a northern hemisphere origin with subsequent migration into the southern hemisphere could account for early Tertiary Moraceae fossils (Collinson, 1989) and the present day distribution of several ancient lineages in Eurasia and North America. The biogeography of Moraceae has yet to be considered in a phylogenetic framework using molecular dating.

Discovering the temporal and geographic origins of the mulberry family is particularly interesting because of the highly specialized mutualism between figs (*Ficus*, Moraceae) and their pollinating fig wasps (Agaoninae: Chalcidoidea). Although this obligate mutualism has informed the study of evolutionary problems ranging from sex ratios (Hamilton, 1967; West and Herre, 1998) to speciation (Weiblen and Bush, 2002), the origin of the interaction has only recently received attention. Molecular phylogenetic studies have identified the closest relatives of the figs (Datwyler and Weiblen, 2004) and have estimated the divergence times of pollinating fig wasps (Machado et al., 2001). Machado et al. (2001) developed a biogeographic hypothesis for *Ficus* on the assumption that fig and fig wasp origins were temporally and geographically congruent owing to the complete reproductive interdependence of the mutualists. This assumption has enabled the ages of other specialized plant–insect interactions to be estimated (Machado et al., 2001; Pellmyr and Leebens-Mack, 1998; Sequeira and Farrell, 2001) but few studies have compared independent divergence times for interacting lineages (Percy et al., 2004).

Our paper compares age estimates for Moraceae with fig wasp divergence times (Machado et al., 2001) to examine the origin of fig pollination based on independent dating of the mutualists. In addition, age estimation for related lineages provides the opportunity to explore biogeographic scenarios for the family as a whole. We present the first phylogenetic estimate for Moraceae based on multiple gene regions and estimates of divergence times based on multiple fossil calibrations. Molecular phylogenies are considered in light of present day distributions, fossil evidence, paleogeographic reconstructions, and theories on ancient climate.

2. Materials and methods

2.1. DNA sequencing

Ninety-five taxa were sequenced for the chloroplast gene *ndhF* (AY289249–AY289287, AY289289–AY289300, AY289302–AY289332, AY289334–AY289335, AY2 89337–AY289344, AY289346–AY289349; Datwyler and Weiblen, 2004) and the 26S nuclear ribosomal subunit (AY686766–AY686860). We included 17 outgroup taxa in four families (Cannabaceae, Cecropiaceae, Celtidaceae, and Urticaceae) and 78 ingroup taxa representing 32 of 37 Moraceae genera (Table 1). Material for five small genera (*Bosqueieospis*, *Hullettia*, *Scyphosyce*, *Treculia*, and *Trilepsium*) was unavailable. Taxa and vouchers are the same as those used in Datwyler and Weiblen (2004) with the exception of six species that could not be sequenced for 26S. DNA was extracted using the Qiagen DNeasy plant extraction kit (Valencia, California, USA) with 10–15 mg of silica gel preserved leaf fragments or herbarium specimens. PCR amplification of the 26S nrDNA region was predominantly achieved in one fragment using forward (5′-GGAGGAAAAGAACTWAC-3′) and reverse (5′-AATGGCCCACTTGGAGCTC-3′) primers designed specifically for Moraceae in this study. In other instances, amplification was achieved in two fragments using Moraceae-specific internal forward (5′-CGTCTTGA AACACGGACC-3′) and reverse primers (5′-C TTCTC ARATCAAGGTCG-3′) paired with the external primers. Amplification of 26S included ~20 ng genomic DNA, 1× *TaKaRa Ex Taq* buffer (2 mM $MgCl_2$; Otsu, Shiga, Japan), 10 μM of each primer, 0.2 mM each dNTP, and 1.25 U *TaKaRa Ex Taq* DNA polymerase. Thermal cycling conditions were 95 °C for 1 min followed by 35 cycles of 95 °C for 1 min, 48 °C for 1 min, and 72 °C for 1 min 30 s, with a final extension of 72 °C for 5 min. PCR products were cleaned with the QIAquick or MinElute PCR purification spin columns (Qiagen). Cleaned PCR products were quantified using Pico Green fluorescent dye designed for quantification of dsDNA (Molecular Probes, Eugene, Oregon, USA) in a Turner Quantech fluorometer (Barnstead-Thermolyne, Dubuque, Iowa, USA). Sequencing was performed in 10 μl reactions using Big Dye sequencing reagents and protocols (Applied Biosystems, Foster City, CA, USA), and data were collected on an ABI 377 automated DNA sequencer (Applied Biosystems). 26S was sequenced in both directions using the four primers above. Sequences were edited using Sequencher version 3.0 (Gene Codes, Ann Arbor, Michigan, USA), and alignment was performed manually.

2.2. Phylogenetic analysis

2.2.1. Data congruence

Whether the nuclear and plastid data sets should be combined were assessed using the Wilcoxon sum of ranks tests (Larson, 1994; Templeton, 1993) as implemented in PAUP* 4.0b10 (Swofford, 2002). Sum of signed ranks tests have an advantage over alternative test of data combinability, such as the incongruence length difference test (Barker and Lutzoni, 2002; Farris et al., 1994; Hipp et al., 2004) in considering relative clade support. Each data set was analyzed to find a most parsimonious tree (MPT) compatible with constraint trees derived from analyses of the rival data set. For example, *ndhF* data were searched (1000 random addition replicates, TBR branch swapping, and up to 100

Table 1
Classification, species richness, and distribution of Moraceae according to Datwyler and Weiblen (2004)

Tribe	Genus	Spp.	Distribution
Artocarpeae R. Br.	*Artocarpus* J.R. and G. Forster	~50	Asia and Indo-Pacific
	Bagassa Aublet	1	Neotropics
	Batocarpus Karsten	4	Neotropics
	Clarisia Ruiz & Pavón	3	Neotropics
	Hulletia King	2	SE Asia
	Parartocarpus Baillon	3	Indo-Pacific
	Prainea King	4	Indo-Pacific
	Treculia Decne. ex Trécul	3	Afrotropics
Castilleae Berg	*Antiaris* Lesch.	1	Afrotropics
	Antiaropsis K. Sch.	1	New Guinea
	Castilla Sessé	3	Neotropics
	Helicostylis Trécul	7	Neotropics
	Maquira Aublet	5	Neotropics
	Mesogyne Engl.	1	Afrotropics
	Naucleopsis Miquel	~20	Neotropics
	Perebea Aublet	9	Neotropics
	Poulsenia Eggers	1	Neotropics
	Pseudolmedia Trécul	~9	Neotropics
	Sparratosyce Bur.	2	New Caledonia
Dorstenieae Gaudich	*Bosqueiopsis* De Willd. & Th. Dur.	1	Afrotropics
	Brosimum Swartz	13	Neotropics
	Dorstenia L.	~105	Afrotropics and Neotropics
	Helianthostylis Baillon	2	Neotropics
	Scyphosyce Baillon	2	Afrotropics
	Trilepisium Thouars	1	Afrotropics
	Trymatococcus Poeppig & Endl.	3	Neotropics
	Utsetela Pellegrin	1	Afrotropics
Ficeae Gaudich	*Ficus* L.	~750	Pantropical
Moreae Gaudich	*Bleekrodea* Blume	3	Afrotropics and SE Asia
	Broussonettia L'Herit. ex Vent.	8	Madagascar and SE Asia
	Fatoua Gaudich.	3	Asia, Australia and Africa
	Maclura Nutt.	11	Pantropics and N America
	Milicia Sim	2	Afrotropics
	Morus L.	~12	Cosmopolitan
	Sorocea St. Hil.	~20	Neotropics
	Streblus Lour.	~25	SE Asia
	Trophis P. Browne	9	Neotropics

trees per replicate) for the shortest trees compatible with a constraint tree based on the 26S analysis. Bootstrap consensus trees with thresholds of 70, 80, and 90% from the rival data set were used as constraints with the exception of outgroup *Pilea fontana* Lunell (Rydb.), whose incongruent position in separate analyses was left unresolved. A pair of randomly selected trees from constrained and unconstrained searches was compared using Shimodaira–Hasegawa tests (Shimodaira and Hasegawa, 1999). Comparisons of bootstrap values between conflicting nodes in the separate analyses further identified sources of weak and strong phylogenetic incongruence between the nuclear and chloroplast data sets.

2.2.2. Parsimony analysis

Separate and combined maximum parsimony (MP) heuristic searches were performed with uninformative characters excluded, characters equally weighted, 10,000 random addition sequence replicates, TBR branch swapping, and saving up to 50,000 trees per replicate. Bootstrap support was assessed by 1000 replicates with 100 random addition sequence replicates per bootstrap replicate, TBR branch swapping and up to 100 trees saved per addition sequence replicate.

2.2.3. Maximum likelihood analysis

Modeltest version 3.4 (Posada and Crandall, 1998) was used to select substitution models that best fit the separate and combined data sets. Five heuristic searches were performed under likelihood using PAUP* 4.0b10 (Swofford, 2002) with a neighbor joining tree as a starting topology and model parameters obtained with Modeltest under the Akaike Information Criterion (AIC) for model selection (Posada and Buckley, 2004). The tree maximizing the likelihood of the data was used to

re-estimate model parameters. These estimates were then used in an additional five rounds of heuristic searches and TBR branch swapping with the maximum likelihood (ML) tree used as a starting topology. Branch swapping and parameter estimation were iterated in this fashion until subsequent analyses converged on the same likelihood score and model parameters.

2.3. Molecular dating

Likelihood ratio tests (LRT) for rate constancy indicated that neither chloroplast nor nuclear regions evolved in a clocklike manner. Two methods allowing for non-clocklike evolution, penalized likelihood and Bayesian methods, were used to estimate dates of divergence. Penalized likelihood is a semiparametric method that allows for multiple fossil constraints and allows substitution rates to vary among lineages according to a smoothing parameter (Sanderson, 2002a). The optimal smoothing parameter was chosen on the basis of the data by cross-validation (Sanderson, 2003). Cross-validation and initial dates were calculated with multiple fossil constraints (see below). Penalized likelihood search parameters included 2000 maximum iterations, 10 multiple starts, and 30 optimization runs. Confidence intervals were determined in the computer program r8s from the curvature of the likelihood surface around the parameter estimate using the *divtime* command and a cutoff of four (Cutler, 2000; Sanderson, 2003).

Bayesian methods relaxing a strict molecular clock, allowing multiple genes to be parameterized separately, and incorporating multiple fossil calibrations were used to estimate divergence dates in a simultaneous analysis of *ndhF* and 26S data sets. The methods were implemented in MULTIDIVTIME (J. Thorne, North Carolina State University). First, ESTBRANCHES estimated branch lengths from the data and a fixed tree topology given the selected model of sequence evolution. Second, the outgroup (*Celtis philippensis* Blanco) was pruned from the tree and MULTIDIVTIME was used to estimate the prior and posterior ages of branching events, their standard deviations, and the 95% credibility intervals via Markov chain Monte Carlo. The Markov chain was run for 10,000 generations and sampled every 100 generations after an initial burn-in period of 100,000 cycles. The following prior distributions were used in the analysis: 132 my for the expected time between the tips and root, 0.0062 substitutions per site per my for the rate of the root node, and 1000 my for the maximum unit time between the root and the tips.

Minimum ages of clades were estimated on a single topology for direct comparison of divergence dates between the separate and combined data sets. The topology resulting from ML analysis of the combined data was selected because simultaneous analysis of chloroplast and nuclear ribosomal DNA sequences yielded higher resolution and support than analysis of either data set alone. Branch lengths were estimated on this topology under four different assumptions: (1) 26S data under the optimal ML 26S substitution model (26S model), (2) *ndhF* data under the optimal ML *ndhF* substitution model (*ndhF* model), (3) the optimal ML model for the combined data (combined model), and (4) partitioned 26S and *ndhF* models using a Bayesian approach (partitioned model). Analyses under each model were conducted using multiple fossil constraints.

Definitive fossils are scarce in Moraceae, as numerous leaf impressions attributed to the family are fraught with homoplasy. However, a review of fossil reproductive structures provided several conservative minimum age constraints (Collinson, 1989). The oldest fossil achenes with diagnostic features of modern *Ficus* are known from early Eocene formations of southern England. Additional *Ficus* fossils have been recovered from the lower Miocene and upper Oligocene in Germany, the Miocene in the former Czechoslovakia, the Eocene in Bulgaria, and the Tertiary in the former USSR. *Broussonetia* fossil fruits are recorded from the upper Eocene of southern England and Germany, the Miocene of Poland, and from the Tertiary in the former Czechoslovakia, Bulgaria, and the former USSR. *Morus* fruits matching modern mulberries in every detail of structure are recorded from the early Eocene of southern England and the former Czechoslovakia, the Miocene of Germany, and the former USSR. Fossil wood with affinities to modern *Artocarpus* has been described as *Artocarpoxylon deccanensis* from the Deccan Intertrappean beds of Mandla district Madhya Pradesh, India (Mehrotra et al., 1984). This formation has been dated at 54.4 ± 8.1 mya (Srivastava et al., 1986). Outgroup fossils were also reviewed by Collinson (1989) and are represented by Tertiary *Boehmeria* achenes from the lower Miocene in Germany, the Miocene in Poland and the Miocene and Pliocene of the former USSR. *Humulus* fossil fruits are known from the Miocene in Bulgaria and the Tertiary in the former USSR. Based on the fossil evidence, we chose minimum age constraints for the crown nodes of several modern genera. When a fossil was dated to a time period, we used the oldest date from that interval. The following age constraints were used in the penalized likelihood analyses: 55 mya for *Ficus* and *Morus*, 46 mya for *Artocarpus*, 33 mya for *Broussonetia*, and 23 mya for *Boehmeria* and *Humulus*. Additionally, the root node of the tree was set to a maximum age of 132 mya based on the oldest known angiosperm fossil (Magallón and Sanderson, 2001). The following upper and lower age constraints were used in the Bayesian analysis: 40 and 55 mya for *Ficus*, 23 and 65 mya for *Morus*, 33 and 40 mya for *Broussonetia*, 46 and 62 mya for *Artocarpus*, 15 and 23 mya for *Boehmeria*, and 5 and 23 mya for *Humulus*.

3. Results

3.1. Phylogenetic analyses

3.1.1. Separate analyses

The aligned 26S data included 1014 base pairs, 185 of which were parsimony informative, and 150 of which were autapomorphic. Parsimony searches recovered 50,000 MPTs of 819 steps (Fig. 1; consistency index, CI = 0.3541; retention index, RI = 0.7117). The monophyly of the Moraceae was not recovered and the only monophyletic tribes recovered with 26S sequence data were Ficeae and Castilleae, with 97 and 84% bootstrap support, respectively. The sister relationship of these tribes had very weak support (<50% bootstrap). In contrast, the split between Old and New World *Maclura* was well supported (88%). Dorstenieae were paraphyletic because of the position of *Bleekrodea*, *Streblus*, and *Fatoua*. As in the previously published *ndhF* phylogeny, Moreae were polyphyletic and *Artocarpus, Batocarpus*, *Brosimum*, *Perebea*, *Streblus*, and *Trophis* were not monophyletic (Datwyler and Weiblen, 2004). In addition, *Castilla* and *Pseudolmedia* were not monophyletic and relationships within Artocarpeae were unresolved.

Aligned *ndhF* sequences from the same individuals sampled for 26S included 2092 base pairs, 631 of which were parsimony informative and 387 of which represented autapomorphies. Parsimony searches recovered 50,000 MPTs of 1862 steps (CI = 0.5322; RI = 0.8506). There was strong support for a monophyletic Moraceae (Fig. 1) and for the monophyly of four of the five Moraceaous tribes (Artocarpeae, Castilleae, Dorstenieae, and Ficeae) while the Moreae were polyphyletic. The sister relationship between Ficeae and Castilleae was well supported (88%), as was the split between the Old and New World Artocarpeae (100%) and *Maclura* (89%). This analysis also recovered a split between the Old and New World Castilleae, but bootstrap support was lacking.

Modeltest (Posada and Crandall, 1998) identified a general time reversible model with a gamma distribution and proportion of invariable sites (GTR + I + Γ; Rodriguez et al., 1990) as the best fitting model of sequence evolution for each of the separate data sets according to the AIC. The single most likely 26S tree resulting from heuristic searches had a score of $-\ln L = 6953.19888$ with a rate matrix of AC = 1.35075, AG = 2.36409, AT = 1.99 467, CG = 0.65374, and CT = 10.53404, Γ = 0.538634, I = 0.4252715, and base frequencies of A = 0.2137359, C = 0.2485911, and G = 0.3397735. Several minor conflicts were apparent between the 26S ML tree and the parsimony strict consensus. Under ML, *Antiaropsis* plus *Sparattosyce* collapsed, *Trymatococcus* was paraphyletic with *Helianthostylis* nested within it, *Sorocea* was paraphyletic with *Parartocarpus* nested within it, and the position of *Brosimum* species were slightly different. Additionally, under ML, Moraceae was paraphyletic because Artocarpeae was sister to the rest of the Moraceae plus Cecropiaceae and Urticaceae. However, a Shimodaira–Hasegawa test comparing the 26S ML tree with a tree from a search constraining the monophyly of Moraceae indicated that there was no significant difference in the likelihood of the data between a monophyletic and a paraphyletic Moraceae ($p = 0.475$).

The single most likely *ndhF* tree had a score of $-\ln L = 16{,}864.8073$ with a rate matrix of AC = 1.66827, AG = 2.02428, AT = 0.37772, CG = 1.11846, CT = 2.037 71, Γ = 0.960542, I = 0.1385749, and base frequencies A = 0.3025505, C = 0.1416032, and G = 0.1588692. The only topological conflict between the *ndhF* ML tree and the parsimony strict consensus was the placement of the outgroup *Pourouma* as sister to either *Cecropia* (MP) or *Coussapoa* (ML).

3.1.2. Phylogenetic congruence

Wilcoxon sum of ranks tests (Larson, 1994; Templeton, 1993) detected no significant incongruence when the 26S 90% bootstrap consensus was enforced as a constraint in a search of the *ndhF* data ($p = 0.3938$). When the *ndhF* 90% bootstrap consensus was enforced as a constraint in a search of the 26S data, the MP trees were significantly longer than unconstrained trees ($p = 0.0004$). Inspection of *ndhF* and 26S bootstrap consensus trees identified the source of this incongruence (Fig. 1). The most severe conflicts involved two outgroup taxa, *Pilea fontana* and *Coussapoa latifolia* Aubl., with strongly supported (>80% bootstrap) and conflicting placements within the outgroup. No conflicting relationships in the ingroup were strongly supported by both data sets. Four conflicts had strong support from one data set but not the other and traversed more than one node (*Fatoua pilosa* Gaudich., *Parartocarpus venenosus* Becc., *Streblus smithii* (Cheeseman) Corner, and *Utsetela neglecta* Jongkind). Two conflicts had strong support from one data set and traversed a single node (*Naucleopsis ulei* Warb. Ducke and *Streblus elongatus* (Miq.) Corner).

Incongruence was also significant when the 26S 80% bootstrap consensus was used as a constraint in searches of the *ndhF* data ($p < 0.0001$). Inspection of the *ndhF* strict consensus and the 26S 80% bootstrap tree revealed four local conflicts. The conflicting positions of *Naucleopsis guianensis* (Mildbr.) C.C. Berg and *N. naga* Pittier had no bootstrap support and traversed only a single node. The position of *Streblus smithii* traversed several nodes but the *ndhF* position had low support (<50% bootstrap). Local conflict involving *Pilea fontana* within the outgroup had high support in both data sets and traversed multiple nodes. When conflict within the outgroup was collapsed and constraint searches of *ndhF* were repeated with the 80 and 70% 26S bootstrap consensus trees, phylogenetic incongruence was not significant ($p = 0.1893$ and 0.1512, respectively). Furthermore,

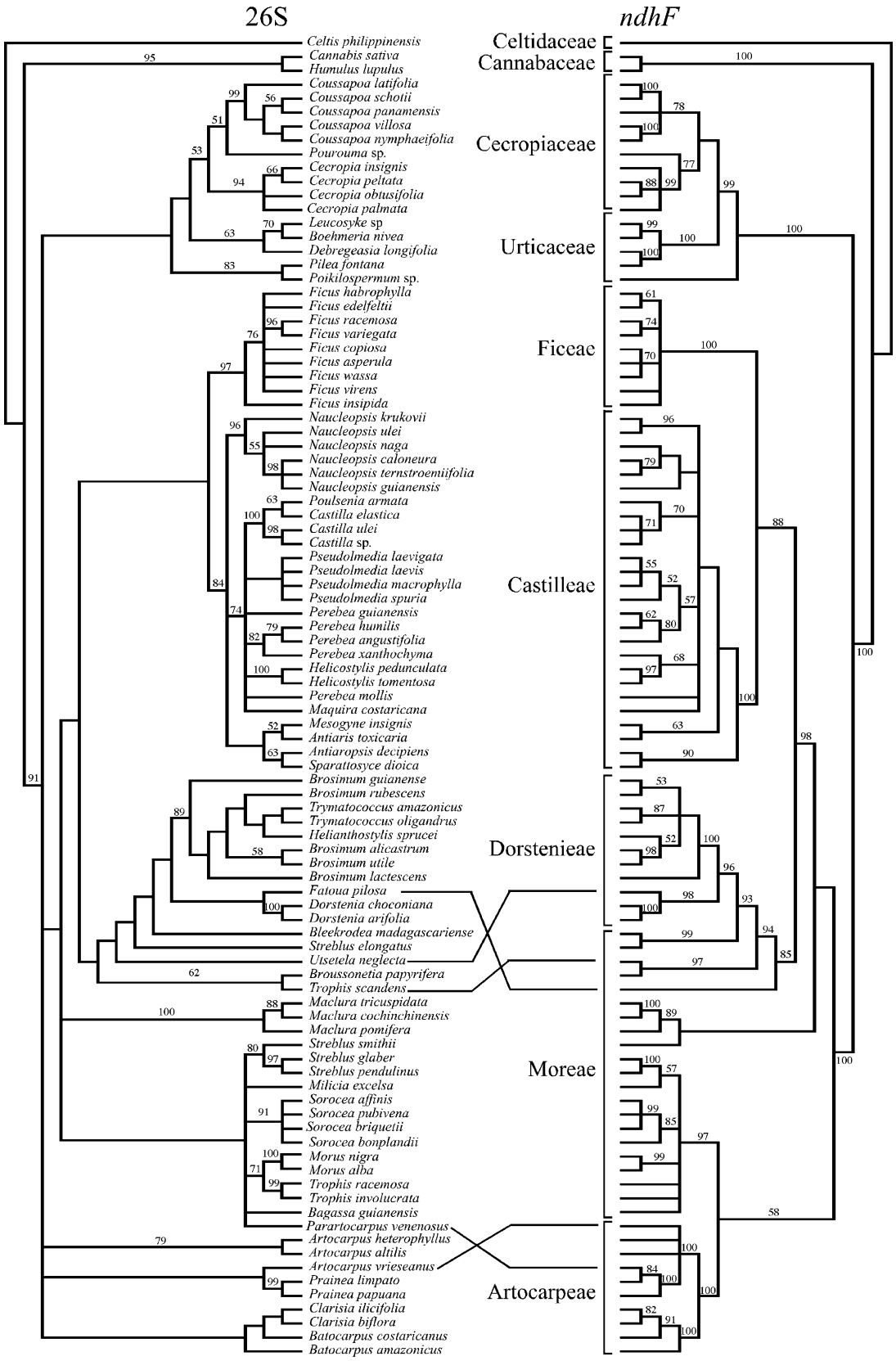

Fig. 1. Strict consensus trees based on separate parsimony analyses of Moraceae 26S nuclear ribosomal DNA and *ndhF* chloroplast gene sequences. Bootstrap support values greater than or equal to 50% are listed above the branches. Outgroup families and Moraceae tribes are indicated in brackets.

tree topologies resulting from separate analyses of 26S and *ndhF* did not differ significantly according to Shimodaira–Hasegawa tests when only well-supported clades (>75% boostrap support) were considered. We interpret this as evidence of phylogenetic signal indicating shared history.

3.1.3. Combined analysis

Maximum parsimony searches of the combined data sets yielded 15,388 MPTs with 2738 steps (CI = 0.47; RI = 0.81). In this analysis, there was unequivocal support for Moraceae (100% bootstrap support), and only 17 nodes collapsed in the strict consensus tree. Support for the four monophyletic tribes was generally high, with 100% for Artocarpeae, 100% for Ficeae, 100% for Castilleae, and 68% for Dorstenieae (Fig. 2). As in separate analyses, the Moreae were polyphyletic. Well-supported splits between New and Old World clades were detected in the Artocarpeae (100%), and *Maclura* (100%). Compared to separate analyses where monophyly of many genera was not detected, only six genera were not monophyletic in the combined analysis: *Artocarpus*, *Batocarpus*, *Brosimum*, *Castilla*, *Streblus*, and *Trophis*.

The combined analyses yielded both greater resolution and higher bootstrap support than either data set alone. In the strict consensus, 41 nodes had higher support than in the *ndhF* data set while only 19 nodes had slightly lower support. Comparison of the combined strict consensus with the 26S strict consensus revealed 56 nodes with greater support only 14 nodes with reduced support.

Modeltest (Posada and Crandall, 1998) identified a general time reversible model with a gamma distribution and proportion of invariable sites (GTR + I + Γ; Rodriguez et al., 1990) as the best fitting model of sequence evolution for the combined data sets according to the AIC. The single most likely combined 26S rDNA and *ndhF* ML tree had a score of $-\ln L = 24{,}802.98094$ with a rate matrix of AC = 1.67021, AG = 2.06849, AT = 0.60556, CG = 1.00212, and CT = 3.50368, Γ = 0.828064, I = 0.26869001, and base frequencies of A = 0.28072927, C = 0.16692689, and G = 0.21333819. There was no significant difference between the combined ML and randomly selected combined MPTs according to the Shimodaira and Hasegawa (1999) test ($p = 0.426$).

3.2. Molecular dating

Likelihood ratio tests for rate constancy rejected the assumption of a molecular clock for both data sets ($p < 0.001$). Cross-validation analyses identified the optimal smoothing parameters for the 26S, *ndhF*, and combined data sets as 5.62, 1000, and 10, respectively. When all available fossils were employed as minimum age constraints, estimates of divergence dates based on the 26S data were older on average than those based on the *ndhF* data set (Fig. 4A). A paired *t* test comparing dates for each node rejected the null hypothesis that 26S and *ndhF* ages are equal ($p < 0.001$). Ages based on the combined data sets were generally intermediate between those of the separate data sets, with the dates of the combined model being slightly older than those of the partitioned model (Fig. 4B). A paired *t* test comparing dates for each node rejected the null hypothesis that the two models are equal ($p < 0.001$). Estimates across the four models suggest that Moraceae diversified at least 89.1–103.4 mya (Table 2). This falls in the range of a recent study using minimum age node mapping (a method based on mapping of reliable fossils onto phylogenies) that placed the Urticales, of which Moraceae is a member, at a minimum age of 94 mya (Crepet et al., 2004). We focus our discussion on the results of the partitioned model because it has the advantage over other models of accommodating different substitution rates for separate gene regions.

4. Discussion

4.1. Moraceae phylogeny

Simultaneous analysis of multiple gene regions improved overall clade support and resolution compared to a single gene (Datwyler and Weiblen, 2004). There were only two instances in which incongruence between gene regions was highly supported, each involving relationships within the outgroup (Fig. 1). Parsimony and likelihood analyses of the combined data strongly support the monophyly of Moraceae and four of the five tribes in the family. The fifth tribe, Moreae, comprised a basal grade, which is not surprising given its delimitation on plesiomorphic characters such as inflexed stamens and wind pollination (Berg, 2001; Datwyler and Weiblen, 2004). Chloroplast and nuclear genes together resolved intergeneric relationships in Castilleae, with increased support for a neotropical clade and the monophyly of two genera that were weakly paraphyletic based on a single marker. We interpret these findings as evidence that Moraceae phylogeny reconstruction is improved by the inclusion of more characters from the nuclear and chloroplast genomes that appear to share a common evolutionary history (de Queiroz et al., 1995; Nixon and Carpenter, 1996).

4.2. Divergence times

Divergence time estimates based on 26S data alone were slightly older on average than those based on *ndhF* (Fig. 4A), with most of the discordance occurring in the tribe Castilleae. The position of fossil constraints may have played a role, as Castilleae is the only tribe with no

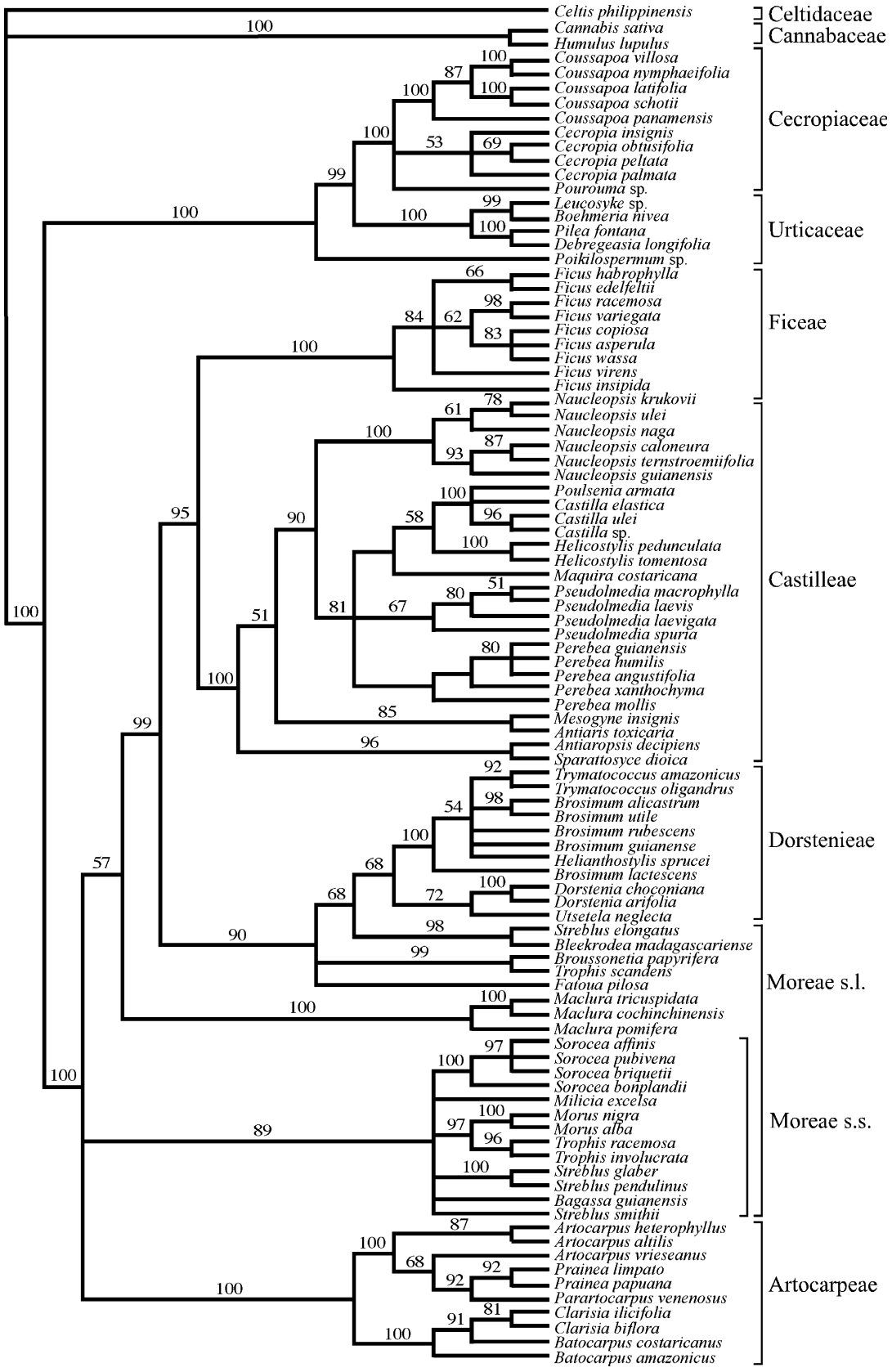

Fig. 2. The strict consensus of 15,388 most parsimonious trees resulting from a combined analysis of 26S and *ndhF* sequences. Bootstrap support values greater than or equal to 50% are indicated above branches. Outgroup families and Moraceae tribes are indicated in brackets.

Table 2
Minimum age estimates (and confidence intervals) for selected clades in the mulberry family based on four different substitution models (see Section 2)

	26S	*ndhF*	Combined model	Partitioned model
Moraceae	101.8 (101.0–109.7)	97.2 (89.7–104.9)	103.4 (96.8–109.6)	89.1 (72.6–110.0)
Ficus	77.8*	55*	55*	43.3 (40.1–51.0)*
Ficus and Castilleae	98.9 (80.2–112.8)	76.0 (68.5–84.2)	84.8 (77.1–92.9)	72.0 (59.6–88.2)
Castilleae	81.9 (57.2–101.6)	35.7 (26.6–46.2)	63.7 (51.2–75.3)	53.3 (39.8–68.5)
Artocarpeae	78.9 (46.7–125.4)	61.2 (53.9–70.5)	72.1 (61.5–84.3)	65.1 (52.2–80.6)
Dorstenieae	83.3 (52.7–104.4)	50.2 (40.9–60.5)	63.0 (48.9–76.1)	48.1 (35.3–62.9)
Moreae s.s.	75.5 (60.1–88.3)	59.1 (51.4–68.7)	78.9 (70.0–88.9)	58.6 (44.2–75.2)

An asterisk marks nodes constrained by fossils.

known fossils, and the accuracy of age estimates decreases with distance from the nearest calibration point (Arbogast et al., 2002). A possible explanation for the discrepancy in divergence date estimates between the different gene regions is the taxonomic level at which each locus is informative and how nucleotide substitutions are distributed across branches of the tree (Fig. 3). The majority of 26S substitutions occurred closer to the tips, favoring older divergence dates while most *ndhF* substitutions occurred along internal branches and favored more recent dates. The application of multiple fossil constraints should help to minimize such variability in divergence time estimates based on multiple gene regions. Similarly, data from multiple regions should help to minimize erroneous inferences based on a single locus (Arbogast et al., 2002; Yang and Yoder, 2003; Yoder and Yang, 2004). Although divergence time estimates from the combined and partitioned models were not identical, variance in divergence times between these models was less than that of models based on a single locus (Fig. 4).

4.3. Historical biogeography

Evaluating the historical biogeography of Moraceae involves a comparison of age estimates derived from fossil-calibrated molecular phylogenies with major events in ancient continental evolution. The supercontinent of Pangea existed from the close of the Paleozoic into the Mesozoic, splitting into Laurasia in the northern hemisphere and Gondwanaland in the south during the Triassic. These events were followed in the Jurassic by the separation of Laurasia into North America and Eurasia and the breakup of Gondwanaland into South America, Africa, India, Madagascar, Australia, Antarctica, New Zealand, and New Caledonia beginning about 180 mya (McLoughlin, 2001). Africa had rifted from Antarctica and India–Madagascar by 165 mya but remained attached to South America until ~105 mya, eventually colliding with Eurasia about 60 mya (McLoughlin, 2001). There are competing views on the timing of the separation of India from Gondwanaland. The prevailing view is that the Madagascar–India–Seychelles land mass rifted northward from Antarctica ~132 mya, that Madagascar separated from India 84–95 mya, and that India drifted northward to collide with Asia approximately 43–50 mya (McLoughlin, 2001; Sanmartín and Ronquist, 2004). An alternate view proposes a land connection between Madagascar–India–Seychelles and Antarctica until as late as 80 mya (Hay et al., 1999). Dates for the separation of Australia from Antarctica range from 35 to 50 mya (McLoughlin, 2001; Woodburne and Case, 1996), followed by the split between South America and Antarctica (~35 mya). Finally, the Australian and Pacific plates collided approximately 25–30 mya bringing the landmasses of Southeast Asia and Australia into close proximity during periods of low sea level.

4.3.1. Gondwanan origin?

We interpret the historical biogeography of Moraceae according to divergence times estimated from nuclear and chloroplast DNA sequences based on the partitioned model, noting that the combined model does not contradict these interpretations (Table 2). We begin with a discussion of dates in light of the prevailing paleogeographic view (summarized by McLoughlin, 2001). It has been suggested that extant Moraceae distribution is the result of vicariance from the break-up of Gondwana (Corner, 1967), However, according to our analyses, Moraceae diversified 89.1 (72.6–110.0) mya, well after Africa and India had separated from Gondwanaland (Fig. 5; Table 2). Although the rifting of Africa from South America is believed to have occurred around 105 mya and the confidence intervals extend to ~110 mya, this split would not likely explain present day Moraceae distribution by vicariance because the extant African lineages included in the analysis (*Utsetela*, *Mesogyne*, *Antiaris*, *Bleekrodea*), appear to be of more recent Eocene or Oligocene origin (Figs. 5 and 6).

Given an estimate of 73–110 mya for the radiation of Moraceae, the distribution of the family could be the result of vicariance following the break-up of the landmass comprising South America, Antarctica, and Australia, with subsequent migrations into Africa and the northern hemisphere. However, there are at least two problems with this explanation for the disjunct distributions of extant lineages. First, northern temperate lineages retaining pleisomorphic features such as *Morus*

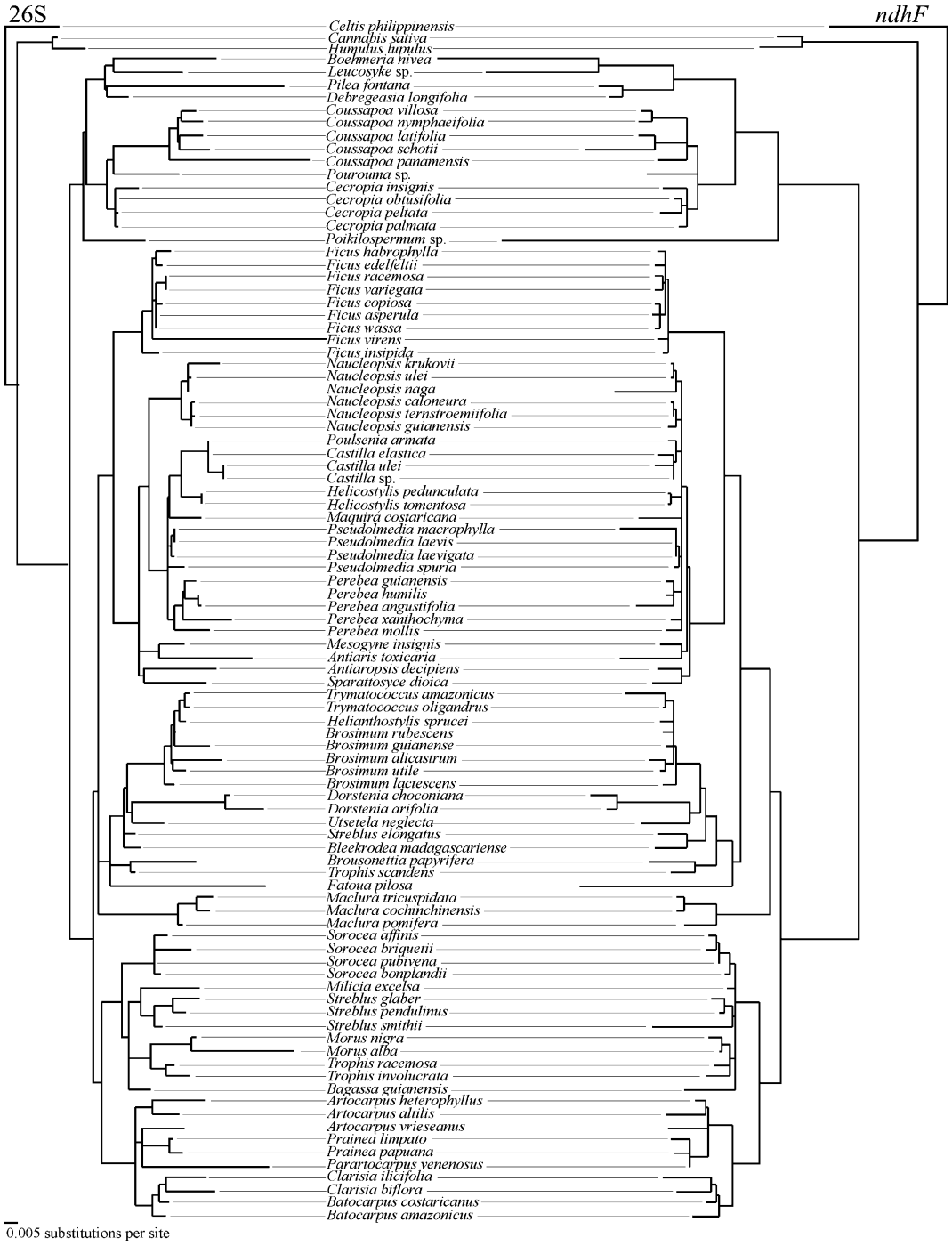

Fig. 3. Maximum likelihood phylogram based on the combined ML analysis comparing branch lengths estimated under separate substitution models for 26S and *ndhF* sequences.

appear to be at least as old as their southern tropical relatives (Fig. 5). Second, the best examples of this phytogeographic pattern come from warm temperate and tropical montane lineages such as Nothofagaceae (Manos, 1997) and Araucariacae (Sequeira and Farrell, 2001) while the Moraceae are dominated by lowland tropical lineages. In order for a southern hemisphere origin and vicariance to be plausible after Africa and India had separated from Gondwanaland, a trans-Antarctic pathway must be invoked during the early Tertiary at the time when the New and Old World lineages diverged (Fig. 5). While Antarctica was important for the migration of some organisms, the local climate was not likely to support tropical Moraceae. Models taking into account seasonal temperature fluctuations suggest that maximum temperatures near sea level in Antarctica were

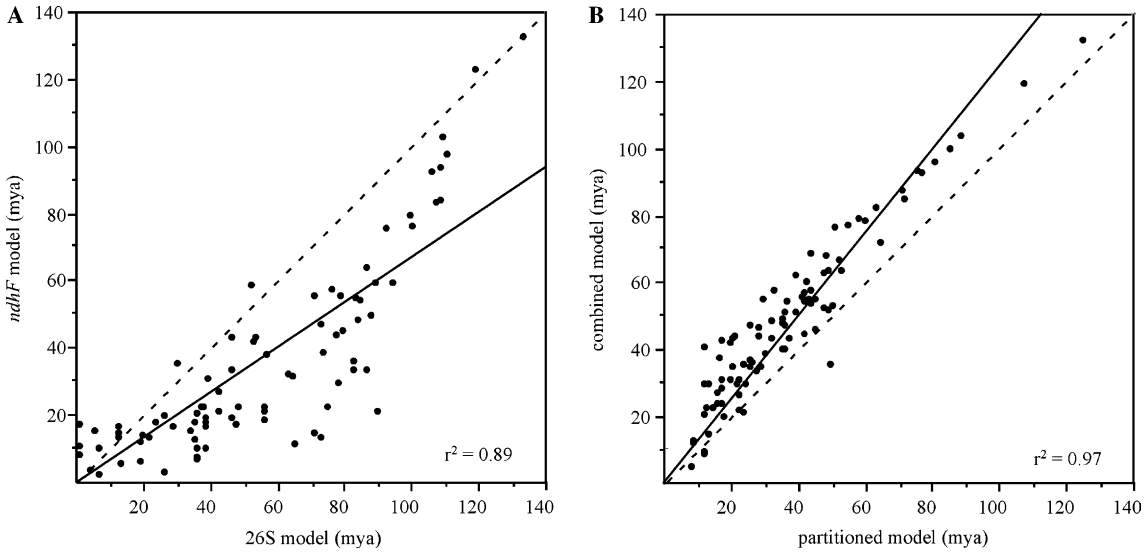

Fig. 4. Minimum age estimates for Moraceae clades derived from different DNA substitution models with the same fossil constraints. (A) 26S substitution model versus the *ndhF* model. (B) Combined data model versus a partitioned data model. Solid and dashed lines indicate the simple regression and the expectation of perfect congruence, respectively.

13–14 °C around 80 mya (Crowley et al., 1986). At 60 mya there was pronounced cooling, followed by a global warming trend from the late Paleocene into the Eocene (Crowley et al., 1986). Antarctica remained temperate during this period and widespread ice accumulation had begun by 38–42 mya (Dallai et al., 2001). It therefore seems unlikely that the climate of Antarctica would have supported lineages adapted to humid tropical conditions and that the divergence of Old and New World Moraceae lineages could have been associated with a trans-Antarctic pathway. Admittedly, our findings are based on minimum age estimates that could change as new fossils are accumulated, and the oldest known fossils are unlikely to coincide with the earliest appearance of a lineage. Nonetheless, the available evidence suggests that an alternative explanation for Moraceae origins should be considered.

4.3.2. Laurasian migration

Due to the predominantly southern tropical distribution of extant Moraceae, it has been supposed that the family originated in Gondwanaland and diversified through vicariance and/or dispersal. Molecular dating is equivocal and suggests that a northern origin in Eurasia is at least as plausible as the hypothesis of Gondwanan origin. If Moraceae diversified in Eurasia during the mid-Cretaceous, there would have been multiple opportunities for migration into the southern hemisphere. The collisions of Africa and India with Eurasia at ~60 mya and ~43–50 mya, respectively, coincided with the radiations of four tribes on these landmasses (Fig. 6). A North Atlantic land connection existed at warm latitudes during most of the Eocene (Tiffney, 1985), and the Americas were in close proximity due to the presence of scattered islands (Iturralde-Venent and MacPhee, 1957). Sanmartín and Ronquist (2004) demonstrated that flora and fauna in northern South America often have greater affinities to northern temperate groups than to southern temperate groups, supporting the notion of a migration route for tropical lineages through North America. Moraceae might have originated in South America and migrated over these land routes but molecular dating suggests that paleotropical Moraceae lineages are generally older than neotropical ones (Fig. 6). Long distance dispersal over water seems unlikely for Moraceae owing to large, short-lived seeds for the most part (Berg, 2001). The absence of tropical Moraceae lineages in the northern hemisphere can be attributed to the current temperate climate and Eocene fossils of tropical lineages are in fact known from temperate localities (Collinson, 1989). In summary, minimum age estimates for the disjunctions of Old and New World Moraceae (Fig. 6) fall within the period when there existed overland routes linking these regions, and the wet tropical climate of southern Eurasia would have been more favorable for Moraceae migration than the warm temperate climate of Antarctica during the same period.

4.3.3. Age of fig pollination

It was not the focus of this investigation to reconstruct the biogeography of *Ficus* in detail, but our findings at the family level have implications for the origin of fig pollination. The obligate mutualism between figs and their specialized pollinating fig wasps is a textbook example of plant–insect coevolution (Cook and Rasplus, 2003; Weiblen, 2002), and understanding the temporal congruence of the association could shed light on

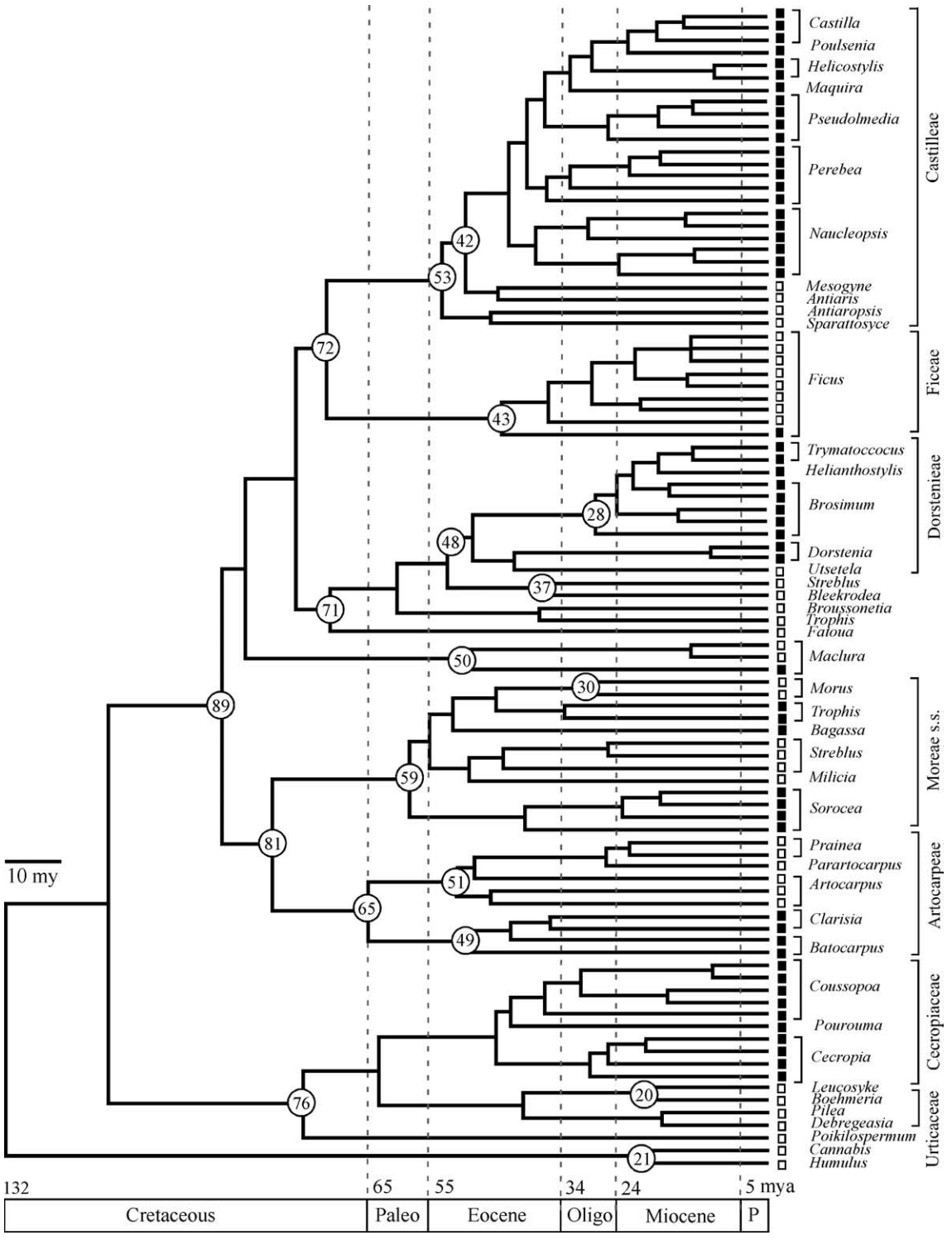

Fig. 5. Chronogram of Moraceae divergence resulting from a Bayesian analysis with separate chloroplast and nuclear gene partitions plus six fossil constraints. The estimated minimum ages of selected clades are shown in circles. Black and white boxes at the terminals mark New and Old World taxa, respectively. Moraceae tribes and outgroup families are indicated in brackets. Geological epochs are abbreviated as follows: Paleo, Paleocene; Oligo, Oligocene; and P, Pliocene.

whether cocladogenesis has occurred. Here, we consider our estimates for the diversification of *Ficus* with the findings for fig wasps (Machado et al., 2001). Dual phylogenies indicate that the oldest split in fig pollination is between neotropical *Ficus* section *Pharmacosycea*, the associated *Tetrapus* pollinators, and their respective

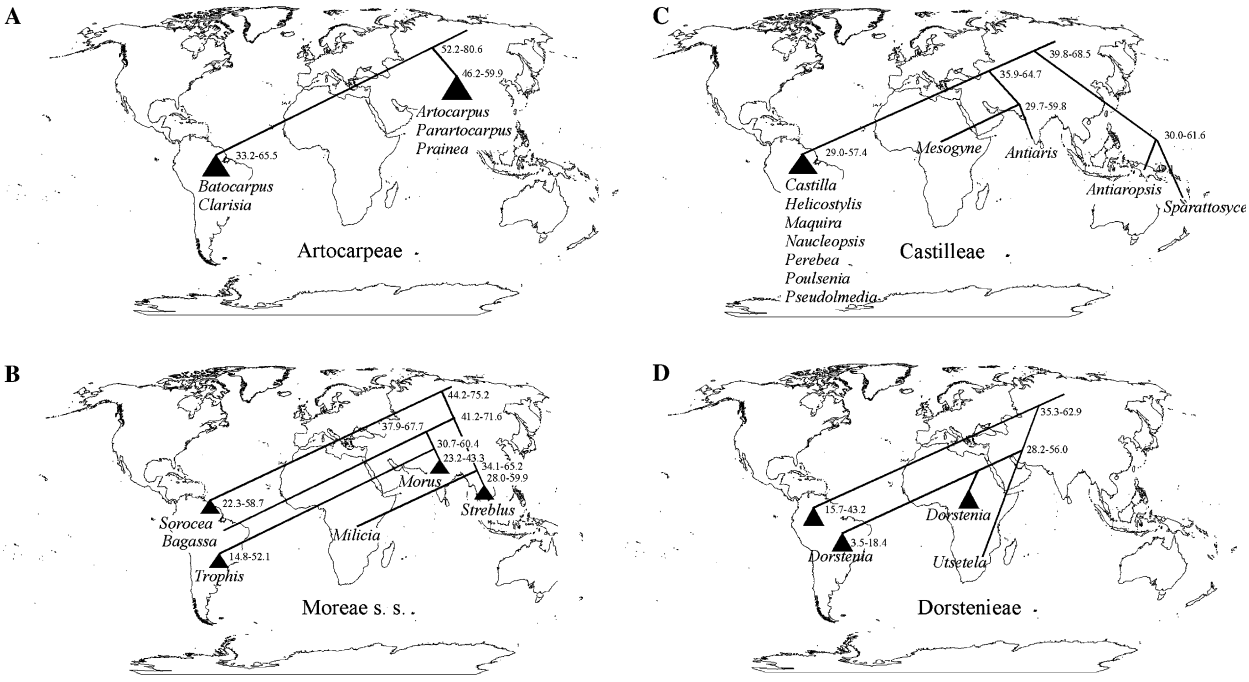

Fig. 6. Biogeographical hypotheses for Moraceae tribes. (A) Artocarpeae, (B) Moreae, (C) Castilleae, and (D) Dorstenieae. Confidence intervals for minimum ages (mya) for selected stem lineages and crown radiations are shown for the partitioned model.

lineages (Datwyler and Weiblen, 2004; Machado et al., 2001). Minimum ages for this split in *Ficus* ranged from 40 to 51 mya (Table 2). Older dates for the diversification of fig pollinators (75–100 mya; Machado et al., 2001) do not support the prediction of simultaneous cladogenesis. However, the discrepancy could have a methodological explanation given that pollinator divergence dates were based on a single gene region, a single fossil calibration, and the assumption of a molecular clock that involved pruning taxa with substitution rate heterogeneity from the analysis. It is interesting that estimates for the crown radiation of pollinators (75–100 mya) coincide better with the divergence of *Ficus* from Castilleae (77–93 mya), but not with fig diversification. The further possibility of a decelerated rate of molecular evolution in the fig lineage or an accelerated rate in the pollinator lineage could make it difficult to reconcile divergence time estimates for the two lineages and to test hypotheses of temporal congruence. In any event, our estimates from Moraceae DNA sequences concerning the age of fig pollination invite molecular dating of the associated insects with a relaxed clock and with increased sampling of fossils, genes, and taxa.

5. Conclusions

It is important to consider as many lines of evidence as possible when evaluating biogeographical scenarios. Sources of evidence include divergence dates based on multiple gene regions, multiple unambiguous fossils when available, geological events, and ancient climate. Minimum age estimates, fossil distributions, paleogeographical reconstructions, and climate considerations make a vicariant Gondwanan scenario less plausible than migration to explain Moraceae distribution. A northern hemisphere origin is at least as plausible as a southern origin. Although not considered previously, multiple sources of evidence are consistent with a hypothesis of Laurasian migration. Most ancient Moraceae fossils are Eurasian and most neotropical lineages are younger than paleotropical and northern temperate lineages (Fig. 6). At the very least, the available data point to a mid-Cretaceous origin of Moraceae with subsequent migration facilitated by multiple land routes.

Acknowledgments

We thank C.C. Berg, N. Rønsted, and F.K. Barker for helpful discussion, two anonymous reviewers for comments, and S.J. Swenson for laboratory assistance. We also acknowledge the US National Science Foundation, National Research Institute of Papua New Guinea, Forest Research Institute of Papua New Guinea, New Guinea Binatang Research Center, Organization for Tropical Studies, Centre National de la Recherche Scientifique, National Museum of Natural History, Smithsonian Tropical Research Institute, Field Museum of

Natural History, Harvard University Herbaria, Nationaal Herbarium Nederland (Leiden), Bell Museum of Natural History, and the Supercomputing Institute at the University of Minnesota. This material is based upon work supported by the National Science Foundation under Grant No. DEB 0128833.

References

Arbogast, B.S., Edwards, S.V., Wakely, J., Beerli, P., Slowinski, J.B., 2002. Estimating divergence times from molecular data on phylogenetic and population timescales. Annu. Rev. Ecol. Syst. 33, 707–740.

Barker, F.K., Lutzoni, F.M., 2002. The utility of the incongruence length difference test. Syst. Biol. 51, 625–637.

Berg, C.C., 2001. *Moreae, Artocarpeae*, and *Dorstenia* (Moraceae). New York Botanical Garden Press, New York.

Bremer, K., 2002. Gondwanan evolution of the grass alliance of families (Poales). Evolution 56, 1374–1387.

Collinson, M.E., 1989. The fossil history of the Moraceae, Urticaceae (including Cecropiaceae), and Cannabaceae. In: Crane, P.R., Blackmore, S. (Eds.), Evolution, Systematics, and Fossil History of the Hamamelidae, vol. 2: 'Higher' Hamamelidae. Clarendon Press, Oxford, pp. 319–339.

Cook, J.M., Rasplus, J.-Y., 2003. Mutualists with attitude: coevolving fig wasps and figs. Trends Ecol. Evol. 18, 241–248.

Corner, E.J.H., 1967. *Ficus* in the Solomon Islands and its bearing on the Post-Jurassic history of Melanesia. Phil. Trans. R. Soc. Lond., Ser. B. 253, 23–159.

Crepet, W.L., Nixon, K.C., Gandolfo, M.A., 2004. Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. Am. J. Bot. 91, 1666–1682.

Crowley, T.J., Short, D.A., Mengel, J.G., North, G.R., 1986. Role of seasonality in the evolution of climate during the last 100 million years. Science 231, 579–584.

Cutler, D.J., 2000. Estimating divergence times in the presence of an overdispersed molecular clock. Mol. Biol. Evol. 17, 1647–1660.

Dallai, L., Ghezzo, C., Longinelli, A., 2001. Fossil hydrothermal systems tracking Eocene climate change in Antarctica. Geology 29, 931–934.

Datwyler, S.L., Weiblen, G.D., 2004. On the origin of the fig: phylogenetic relationships of Moraceae from ndhF sequences. Am. J. Bot. 91, 767–777.

de Queiroz, A., Donoghue, M.J., Kim, J., 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. 26, 657–681.

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

Hamilton, W.D., 1967. Extraordinary sex ratios. Science 156, 477–488.

Hay, W.W., DeConto, R.M., Wold, C.N., Wilson, K.M., Voigt, S., Schulz, M., Rossby Wold, A., Dullo, W.-Chr., Ronov, A.B., Balukhovsky, A.N., Söding, E., 1999. Alternative global Cretaceous paleogeography. In: Barrera, E., Johnson, C.C. (Eds.), Evolution of the Cretaceous Ocean-Climate System. Geological Society of America. Boulder, Colorado, pp. 1–48.

Hipp, A.L., Hall, J.C., Sytsma, K.J., 2004. Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Syst. Biol. 53, 81–89.

Larson, A., 1994. The comparison of morphological and molecular data in phylogenetic systematics. In: Schierwater, B., Streit, B., Wagner, G.P., DeSalle, R. (Eds.), Molecular Ecology and Evolution: Approaches and Applications. Birkhauser, Basel, Switzerland, pp. 371–390.

Machado, C.A., Jousselin, E., Kjellberg, F., Compton, S.G., Herre, E.A., 2001. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. Proc. R. Soc. Lond. B Biol. Sci. 268, 1–10.

Magallón, S., Sanderson, M.J., 2001. Absolute diversification rates in angiosperm clades. Evolution 55, 1762–1780.

Manos, P.S., 1997. Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. Am. J. Bot. 84, 1137–1155.

McLoughlin, S., 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. 49, 271–300.

Mehrotra, R.C., Prakash, U., Bande, M.B., 1984. Fossil woods of *Lophopetalum* and *Artocarpus* from the Deccan Intertrappean Beds of Mandla district, Madhya Pradesh, India. Palaeobotanist 32, 310–320.

Nixon, K.C., Carpenter, J.M., 1996. On simultaneous analysis. Cladistics 12, 221–241.

Pellmyr, O., Leebens-Mack, J., 1998. Herbivores and molecular clocks as tools in plant biogeography. Biol. J. Linn. Soc. 63, 367–378.

Percy, D.M., Page, R.D.M., Cronk, Q.C.B., 2004. Plant–insect interactions: double-dating associated insect and plant lineage reveals asynchrous radiations. Syst. Biol. 53, 120–127.

Posada, D., Buckley, T., 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

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

Rodriguez, F., Oliver, J.F., Marín, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142, 485–501.

Rohwer, J.G., 1993. Moraceae. In: Kubitzki, K., Rohwer, J.G., Bittrich, V. (Eds.), The Families and Genera of Vascular Plants. Springer-Verlag, Berlin, pp. 438–453.

Sanderson, M.J., 2002a. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Sanderson, M.J., 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.

Sanmartín, I., Ronquist, F., 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53, 216–243.

Sequeira, A.S., Farrell, B.D., 2001. Evolutionary origins of Gondwanan interactions: How old are *Araucaria* beetle herbivores?. Biol. J. Linn. Soc. 74, 459–474.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Srivastava, A.P., Rafagopalan, G., Ambwani, K., 1986. Fission-track dating of fossil palm wood from Shahpura, Mandla district, Madhya Pradesh. Geophytology 16, 136–137.

Swofford, D.L., 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer Associates, Sunderland, MA, USA.

Templeton, A.R., 1993. The "Eve" hypothesis: a genetic critique and reanalysis. Am. Anthropol. 95, 51–72.

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

Weiblen, G.D., 2002. How to be a fig wasp. Annu. Rev. Entomol. 47, 299–330.

Weiblen, G.D., Bush, G.L., 2002. Speciation in fig pollinators and parasites. Mol. Ecol. 11, 1573–1578.

West, S.A., Herre, E.A., 1998. Stabilizing selection and variance in fig wasp sex ratios. Evolution 52, 475–485.

Woodburne, M.O., Case, J.A., 1996. Dispersal, vicariance, and the late Cretaceous to early Tertiary land mammal biogeography from South America to Australia. J. Mamm. Evol. 3, 121–161.

Yang, Z., Yoder, A.D., 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst. Biol. 52, 705–716.

Yoder, A.D., Yang, Z., 2004. Divergence dates for Malagasy lemurs estimated from multiple gene loci: geological and evolutionary context. Mol. Ecol. 13, 757–773.