Journal of Biogeography (*J. Biogeogr.*) (2009) **36**, 2338–2352

ORIGINAL ARTICLE

From Africa via Europe to South America: migrational route of a species-rich genus of Neotropical lowland rain forest trees (*Guatteria*, Annonaceae)

Roy H. J. Erkens[1]*, Jan. W. Maas[1] and Thomas L. P. Couvreur[2,3]

[1]*Institute of Environmental Biology, Section Plant Ecology and Biodiversity, Utrecht University, Sorbonnelaan, Utrecht, The Netherlands,* [2]*The New York Botanical Garden, Bronx, NY, USA and* [3]*Biosystematics Group, Nationaal Herbarium Nederland – Wageningen branch, Wageningen University, Generaal Foulkesweg, Wageningen, The Netherlands*

*Correspondence: Roy Erkens, Institute of Environmental Biology, Section Plant Ecology and Biodiversity, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands.
E-mail: R.H.J.Erkens@uu.nl

ABSTRACT

Aim Several recent studies have suggested that a substantial portion of today's plant diversity in the Neotropics has resulted from the dispersal of taxa into that region rather than by vicariance. In general, three routes have been documented for the dispersal of taxa onto the South American continent: (1) via the North Atlantic Land Bridge, (2) via the Bering Land Bridge, or (3) from Africa directly onto the continent. Here a species-rich genus of Neotropical lowland rain forest trees (*Guatteria*, Annonaceae) is used as a model to investigate these three hypotheses.

Location The Neotropics.

Methods The phylogenetic relationships within the long-branch clade of Annonaceae were reconstructed (using maximum parsimony, maximum likelihood and Bayesian inference) in order to gain insight in the phylogenetic position of *Guatteria*. Furthermore, Bayesian molecular dating and Bayesian dispersal–vicariance (Bayes-DIVA) analyses were undertaken.

Results Most of the relationships within the long-branch clade of Annonaceae were reconstructed and had high support. However, the relationship between the *Duguetia* clade, the *Xylopia–Artabotrys* clade and *Guatteria* remained unclear. The stem node age estimate of *Guatteria* ranged between 49.2 and 51.3 Ma, whereas the crown node age estimate ranged between 11.4 and 17.8 Ma. For the ancestral area of *Guatteria* and its sister group, the area North America–Africa was reconstructed in 99% of 10,000 DIVA analyses, while South America–North America was found just 1% of the time.

Main conclusions The estimated stem to crown node ages of *Guatteria* in combination with the Bayes-DIVA analyses imply a scenario congruent with an African origin followed by dispersal across the North Atlantic Land Bridge in the early to middle Eocene and further dispersal into North and Central America (and ultimately South America) in the Miocene. The phylogenetically and morphologically isolated position of the genus is probably due to extinction of the North American and European stem lineages in the Tertiary.

Keywords
Ancestral area, Bayes-DIVA, Bering Land Bridge, boreotropic hypothesis, historical biogeography, North Atlantic Land Bridge, plant immigrants, Tertiary.

INTRODUCTION

Plant diversity in South America is among the highest on the planet (Myers *et al.*, 2000). The dominant idea that this high level of diversity arose by *in situ* speciation (e.g. Raven & Axelrod, 1974; Simpson, 1980; Gentry, 1982; Burnham & Graham, 1999; Young *et al.*, 2002) appears to be one explanation amongst others, rendering the picture of Neotropical diversification more complex than expected (Pennington & Dick, 2004; Rull, 2008). Phylogenetic studies of

www.blackwellpublishing.com/jbi
doi:10.1111/j.1365-2699.2009.02162.x

Journal of Biogeography

plants reveal arrivals of immigrants from different origins into South America from the Late Cretaceous and through the Tertiary (Lavin & Luckow, 1993; Doyle & Le Thomas, 1997; Chanderbali *et al.*, 2001; Renner *et al.*, 2001; Davis *et al.*, 2002; Pennington & Dick, 2004; Richardson *et al.*, 2004; Bell & Donoghue, 2005; Hughes & Eastwood, 2006; Nathan, 2006; Pirie *et al.*, 2006; Erkens *et al.*, 2007b). In many cases, these immigrants speciated and became important elements of the Neotropical flora, both in terms of species richness and ecological dominance (Pennington & Dick, 2004). For instance, Chanderbali *et al.* (2001) showed that the Neotropical clade that includes amongst others *Ocotea* (a species-rich Lauraceae genus abundant in lowland South American rain forests) is nested within the less diverse African, Madagascan and Macaronesian lineages. They demonstrated that this trans-Atlantic distribution was the result of migration through Laurasia. This pattern, and the others cited above, can be seen as support for the boreotropical hypothesis (Wolfe, 1975; Tiffney, 1985a,b; Lavin & Luckow, 1993).

In general, three routes have been documented for the migration of taxa into South America (Fig. 1). The first two are directly related to land connections that existed well into the Tertiary (Tiffney & Manchester, 2001; Morley, 2003): (1) the North Atlantic Land Bridge (NALB) connecting Europe and North America in the early Tertiary, and (2) the Bering Land Bridge (BLB or Beringia) connecting Asia and North

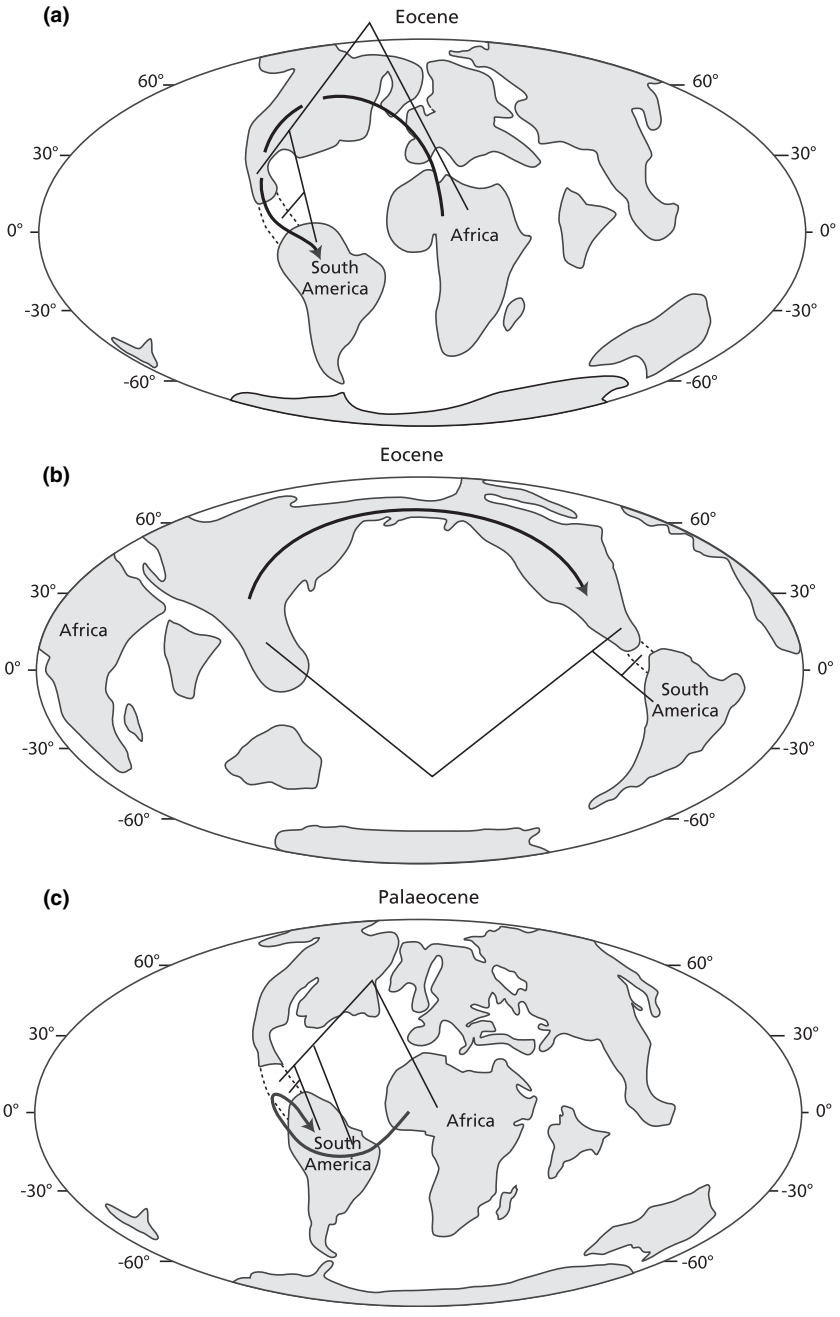

Figure 1 Three alternative hypotheses for the arrival of the most recent common ancestor (MRCA) of *Guatteria* in Central America and the expected phylogenetic patterns derived from them (explanation in the main text). Dotted lines symbolize the present-day outline of Central America. Continental contours are stylized. Information for the base maps of the Eocene was extracted from C. R. Scotese's PALEOMAP project (http://www.scotese.com/). Information for the base map of the Palaeocene was extracted from Morley (2000).

America from the early Palaeogene into the late Miocene/early Pliocene. The third hypothetical migration route is dispersal from Africa directly onto the continent, possibly via some form of land connection or stepping stones (Morley, 2003). These three routes can be seen as alternative hypotheses for the *ex situ* origin of elements of South American diversity. Phylogenetic trees can be used to test the likelihood of these different hypotheses, especially when temporal data and inferences on directionality are incorporated. It is possible to infer the time at which a taxon underwent migration and to determine the source and time of entry of a taxon onto the continent (Donoghue & Moore, 2003).

A taxon that is comparable to *Ocotea* in species richness in the Neotropical lowlands is *Guatteria* (Annonaceae), a genus consisting of *c.* 300 mainly tropical lowland tree species (Erkens *et al.*, 2008b). In an earlier infrageneric biogeographical study (Erkens *et al.*, 2007b) it was inferred that the most recent common ancestor (MRCA) of all extant *Guatteria* species could be found in Central America and originated at least some 11 Ma (Erkens *et al.*, 2007b). However, whether the ancestors of *Guatteria* migrated from South America into Central America or came from Asia or via Europe from Africa could not be determined on the basis of that data set. Furthermore, on the basis of other studies, *Guatteria* as a whole was ambiguously placed with respect to its sister clades (Richardson *et al.*, 2004; Couvreur *et al.*, 2008b). Also, in morphological analyses of the family its position was unclear (Doyle & Le Thomas, 1996) because many of the character states for *Guatteria* are unique within the family or, when not unique, are homoplastic. The development of a biogeographical hypothesis for its origin is therefore hampered.

The clade to which *Guatteria* belongs contains genera with African, Asian and South American origins (Doyle *et al.*, 2004; Richardson *et al.*, 2004; Couvreur *et al.*, 2008b). Without any further information, any of the three above-mentioned hypotheses could apply to the Central American origin of its MRCA. Firstly, the scenario could be as in *Ocotea*, with the ancestors of *Guatteria* having migrated from Africa via North America into Central America (Fig. 1a). If this were the case a split from its sister group should be dated around a time interval when a warm and moist climate allowed migration of tropical plants through the NALB (*c.* 45–55 Ma; Tiffney & Manchester, 2001; Morley, 2003). Migration outside this climatic optimum can be regarded as unlikely because *Guatteria* is a megathermal taxon (also called thermophilic or frost intolerant). It has been shown that megathermal Annonaceae are sensitive to temperature (Punyasena *et al.*, 2008) and small temperature differences might be of great importance for the distribution of lowland taxa (Janzen, 1967) such as *Guatteria*. Furthermore, *Guatteria* is a taxon associated with wetter conditions (Butt *et al.*, 2008), illustrating its moisture dependence. In this scenario Africa would be reconstructed as the ancestral area for *Guatteria* and its sister group.

Secondly, the MRCA could have come from Asia (Fig. 1b). This would have involved migration over the BLB, which connected Asia and North America from the early Palaeogene until its closure between 7.4 and 4.8 Ma (Tiffney & Manchester, 2001). The BLB has been a viable route for terrestrial plants since the Palaeocene. Moreover, fossils of paratropical taxa are known from southern Alaska (although they might be on displaced terrain; McKenna, 1983), but they are less commonly reported from the Siberian side. This may indicate that the BLB route was less frequently used by paratropical taxa than the NALB route (Tiffney & Manchester, 2001). A reason for this could be that although there is some evidence for a Middle Eocene climatic optimum (Zachos *et al.*, 2001), daylight might have been a limiting factor for migration of megathermal plant taxa over the BLB (which was probably too far north at 75–80°; Tiffney & Manchester, 2001; Morley, 2003). As for the first hypothesis, the timing of the migration would be estimated between Lower and Middle Eocene during the climatic optimum; however, the ancestral area should be reconstructed as Asian.

The third possible migration route is dispersal from Africa directly onto the South American continent (Morley, 2003). Direct land connections between the two continents ceased to exist at the end of the Albian (*c.* 96 Ma), but some form of land connection or a series of islands acting as stepping stones possibly facilitated dispersal throughout the late Cretaceous and into the earliest Tertiary (Morley, 2003). Since the ancestor of *Guatteria* could be found in Central America (Erkens *et al.*, 2007b) an additional migration from South into Central America followed by extinction of the first immigrants in South America would have to be postulated (Fig. 1c). In this scenario the age for the stem node would be older than for the other two hypotheses, with a date around the Late Cretaceous or earliest Tertiary. In addition, South America would be reconstructed as the ancestral area for the stem node of *Guatteria*. Africa would show as the ancestral area for the stem node of the clade containing *Guatteria* and its sister group.

In this study we reconstruct the phylogenetic relationships within the long-branch clade of Annonaceae in order to gain insight in the phylogenetic position of *Guatteria*. Furthermore, Bayesian molecular dating and Bayesian dispersal–vicariance (Bayes-DIVA) analyses (Nylander *et al.*, 2008) are undertaken in order to discriminate between the hypotheses shown in Fig. 1.

MATERIALS AND METHODS

Taxon sampling

All clades in the long-branch clade (LBC) of Annonaceae are represented in this study (Richardson *et al.*, 2004; Couvreur *et al.*, 2008b; see Appendix S1 in Supporting Information for vouchers). Taxa from the short-branch clade, the ambavioids and *Anaxagorea*, were chosen to represent the other major Annonaceae lineages (Mols *et al.*, 2004; Richardson *et al.*, 2004; Pirie *et al.*, 2006). A representative of the sister family of Annonaceae (*Eupomatia bennettii*, Eupomatiaceae) as well as three more distant lineages (*Degeneria vitiensis*, Degeneriaceae;

Galbulimima belgraveana, Himantandraceae; *Magnolia kobus*, Magnoliaceae) were added as outgroups in order to facilitate molecular dating. Because not all outgroups had data for all regions sampled, composite taxa were created (i.e. sequences of the same species, but originating from different accessions, were combined). *Guatteria anomala* was included because it is the first branching lineage in *Guatteria* (Erkens *et al.*, 2007a,b).

DNA extraction, amplification and sequencing

Total genomic DNA from fresh silica-dried leaves was extracted using a modified cetyl trimethyl ammonium bromide (CTAB) method (Doyle & Doyle, 1987; Erkens *et al.*, 2008a) or the GenElute® Plant Genomic DNA Miniprep Kit (Sigma-Aldrich Co., St Louis, MO, USA). Seven coding and non-coding plastid regions were targeted in this study: *trnL–F*, *rbcL*, *matK*, *psbA–trnH*, *trnS–G*, *ndh*F and *atpB–rbcL* (for primers see Table 1). In general a standard polymerase chain reaction (PCR) protocol (35 cycles; 30 s,: 94°C; 30 s, 53°C; 1 min, 72°C; with an initial 5 min, 94°C and a final 10 min, 72°C) was used and 0.4% bovine serum albumin was added to the mixes. For long fragments (> 500 bp) a programme with longer cycles and sometimes a lower annealing temperature was applied (28 cycles; 1 min, 94°C; 1 min, 50°C or 53°C; 2 min, 72°C; with an initial 5 min, 94°C and a final 10 min, 72°C). When necessary, PCR products of low concentration were reamplified in order to obtain sufficient material for sequencing. PCR products were purified using the QIAquick PCR purification kit (Qiagen Benelux B.V., Venlo, the Netherlands). For cycle-sequence reactions the same primers were used as for amplifying the particular region, except for the *trnS–G* intergenic spacer (Table 1). Cycle-sequencing was done with DYE-ET (GE Healthcare, Chalfont St Giles, UK) or BIGDYE terminators (Applied Biosystems (ABI), Foster City, CA, USA) and run on an ABI 3730XL automated DNA sequencer.

Phylogenetic analyses

Sequences were edited and assembled in SeqMan pro (DNASTAR Inc., Madison, WI, USA), while alignment was done manually using PAUP* 4.0b10 (Swofford, 2003). Gaps were coded following the simple indel coding method (Simmons & Ochoterena, 2000). All data partitions were combined for analyses because no bootstrap-supported incongruence existed between the separate data partitions (data not shown). A maximum parsimony (MP) analysis was carried out using PAUP* 4.0b10 (Swofford, 2003) from 100,000 replicates of random taxon addition and swapped using tree bisection–reconnection (TBR) and equal weights. Bootstrap analysis (Felsenstein, 1985) with 100,000 replicates of 25 additional sequence replicates was performed with equal weights and TBR branch swapping.

A maximum likelihood (ML) analysis with RAxML version 7.0.4 (Stamatakis, 2006) was also undertaken using the web-server version (Stamatakis *et al.*, 2008) at the CIPRES portal (http://8ball.sdsc.edu:8889/cipres-web/Home.do). The 'Maxi-

Table 1 Primers used for PCR and sequencing reactions.

Marker	Primer name (and sequence if new)	Reference
trnL–F	c, d, e, f	Taberlet *et al.* (1991)
rbcL	1F, 724R	Olmstead *et al.* (1992)
	636F, 1460R	Fay *et al.* (1997, 1998)
	217F, 922F, 536R, 1104R	Pirie *et al.* (2005)
matK	390F, 1326R	Cuénoud *et al.* (2002)
	390F-2	Erkens *et al.* (2007a)
	MintF, MintR	Pirie *et al.* (2005)
psbA–trnH	psbA, trnH	Hamilton (1999)
trnS–G	trnS (GCU), trnG (UCC)	Hamilton (1999)
	trnSintF: 5′-GTTTGARCRCTTGAGTCC-3′	This study*
	trnGintR: 5′-CCAAAWTTTATGAATTTKGGTCA-3′	This study*
ndhF	1F, 972F, 2210R	Olmstead & Sweere (1994)
	-47F, MF561F, 972R, 1165R	Kim *et al.* (2001)
	689R: 5′-GGCATCRGGYAACCATACATGAAG-3′	This study
	LBC-intF: 5′-TCAATAYCTATATGGGGGAAAG-3′	This study
	LBC-intR: 5′-TTCGAAAGGAATTCCTATGRAYCC-3′	This study
atpB–rbcL	atprbc2	Scharaschkin & Doyle (2005)
	atprbc3	Complementary to S20 of Hoot *et al.* (1995)

*Most lineages appeared to have a poly-A/T run in the *trnG* intron near the 3′ *trnG* end. Usually this did not affect the sequence reaction. However, for some taxa it was necessary to apply this primer to overcome the sequencing problem.

mum likelihood search' and 'Estimate proportion of invariable sites' boxes were selected and a total of 100 bootstrap replicates were performed.

Bayesian estimation of tree topology and divergence dates were simultaneously carried out using the software program BEAST 1.4.8 (Drummond & Rambaut, 2007). The data set was partitioned following Couvreur (http://tlpcouvreur.googlepages.com/) with each gene representing one partition. For each gene partition the best fitting evolutionary model was identified (Table 2) under two different model selection criteria, the hierarchical likelihood ratio test (hLRT) and the Akaike information criterion (AIC; Akaike, 1973) as implemented in MRMODELTEST version 2.3 (Nylander, 2004). The hLRT and AIC gave identical results except for *psbA–trnH*. A total of 12 independent runs of 8 million generations each were undertaken on the online cluster of the Computational Biology Service Unit (Cornell University; http://cbsuapps.tc.cornell.edu/beast.aspx). These independent runs were then combined into a single longer run after all runs had reached stationarity (likelihood plateau) and had convergence between them. The ML tree found with RAxML, rendered ultrametric using the program r8s (Sanderson, 2003), was used as a starting tree for all independent runs. Analyses were undertaken by sampling every 1000th generation. Because the sequence data set deviated from a strict molecular clock and rates between adjacent branches were uncorrelated, a lognormal non-correlated relaxed clock method (Drummond *et al.*, 2007) was specified as well as the birth–death speciation process (Gernhard, 2008). TRACER 1.4 (http://beast.bio.ed.ac.uk/) was used to check for convergence of the model likelihood and parameters between each run and to check that each run had reached stationarity. The resulting log files were combined using the program LOGCOMBINER that is part of the BEAST package (Drummond & Rambaut, 2007). Results were considered reliable once effective sample size (ESS) values of all parameters were above 200 (Drummond *et al.*, 2007).

Table 2 Best fitting evolutionary model of seven plastid DNA partitions for the Annonaceae under two different model selection criteria, the hierarchical likelihood ratio test (hLRT) and the Akaike information criterion (AIC; Akaike, 1973) as implemented in MRMODELTEST v2.3 (Nylander, 2004). Γ denotes the shape parameter of the gamma distribution and I is the proportion of invariable sites.

Plastid partition	hLRT	AIC
matK	GTR + Γ	= hLRT
rbcL	GTR + I + Γ	= hLRT
trnL–F	GTR + Γ	= hLRT
trnS–G	GTR + Γ	= hLRT
aptB–rbcL	GTR + Γ	= hLRT
ndhF	GTR + I + Γ	= hLRT
psbA–trnH	HKY + Γ	GTR + Γ

Molecular dating

Three calibration points were applied in this study. The first one is an early Cenomanian fossil from North America (Dilcher & Crane, 1984) which was assigned to the stem of Magnoliaceae (Fig. 2, marked 'Fo1'; 98 Ma, based on stratigraphy). This fossil taxon is commonly used in molecular dating of Annonaceae (Doyle *et al.*, 2004; Richardson *et al.*, 2004; Pirie *et al.*, 2006; Erkens *et al.*, 2007b; Couvreur *et al.*, 2008a). The second is a fossil seed from the Maastrichtian in Nigeria (Chesters, 1955; our Fig. 2, marked 'Fo2'; 68 Ma, based on stratigraphy), placed at the split between the so-called ambavioids and the combined short- and long-branch clades. The placement and accuracy of both calibration points are discussed extensively elsewhere (Doyle *et al.*, 2004; Richardson *et al.*, 2004; Couvreur *et al.*, 2008a,b). The third calibration point was the crown node age for *Guatteria* (Fig. 2, marked 'Gua'; 11.4 ± 1.4 Ma; Erkens *et al.*, 2007b).

In contrast to previous statements (Erkens *et al.*, 2007b) two annonaceous fossil leaves have been attributed to *Guatteria*: *Anonaceaephyllum culebrensis* and *Anonaceaephyllum cretaceum* (originally described as *Guatteria culebrensis* and *Guatteria cretacea*, respectively). The use of fossils as calibration points should not be done without close scrutiny. The original description, taxonomic assignment and geological dating should not be accepted without further investigation (Hermsen & Gandolfo, 2004). Gandolfo *et al.* (2008) mention 11 conditions with respect to taxon selection, fossil source and age of the fossil that should be met before a fossil can be used as a trustworthy calibration point. Both *A. culebrensis* and *A. cretaceum* did not meet several of these criteria, but especially those regarding the evaluation of the original diagnosis (lacking in the case of *Anonaceaephyllum cretaceum*) and the check for synapomorphies that provide unequivocal taxonomic assignment of the fossil. This is especially important since leaves are less diagnostic for Annonaceae than, for instance, are seeds (Doyle *et al.*, 2004). For these reasons, these fossils were not used in the current study.

Four separate analyses were run: only Fo1; Fo1 and Fo2; Fo1, Fo2 and Gua together; and Fo1 and Gua. These alternative calibration strategies were used in order to test the robustness of the obtained ages. Finally, in order to account for the uncertainty in the age of the calibration points itself a normal prior distribution was specified with a standard deviation of 1 Myr for Fo1 and Fo2, and 1.4 Myr for Gua (Erkens *et al.*, 2007b).

Biogeography

In order to infer ancestral areas in the LBC a DIVA analysis was carried out (Ronquist, 1997). However, given the uncertainty surrounding the exact placement of *Guatteria* (Fig. 3) it would be questionable to use a single fully resolved topology as is normally done. Therefore the Bayes-DIVA approach as presented by Nylander *et al.* (2008) was used. In this method phylogenetic uncertainty is accommodated by utilizing the posterior distribution of trees resulting from a BEAST analysis

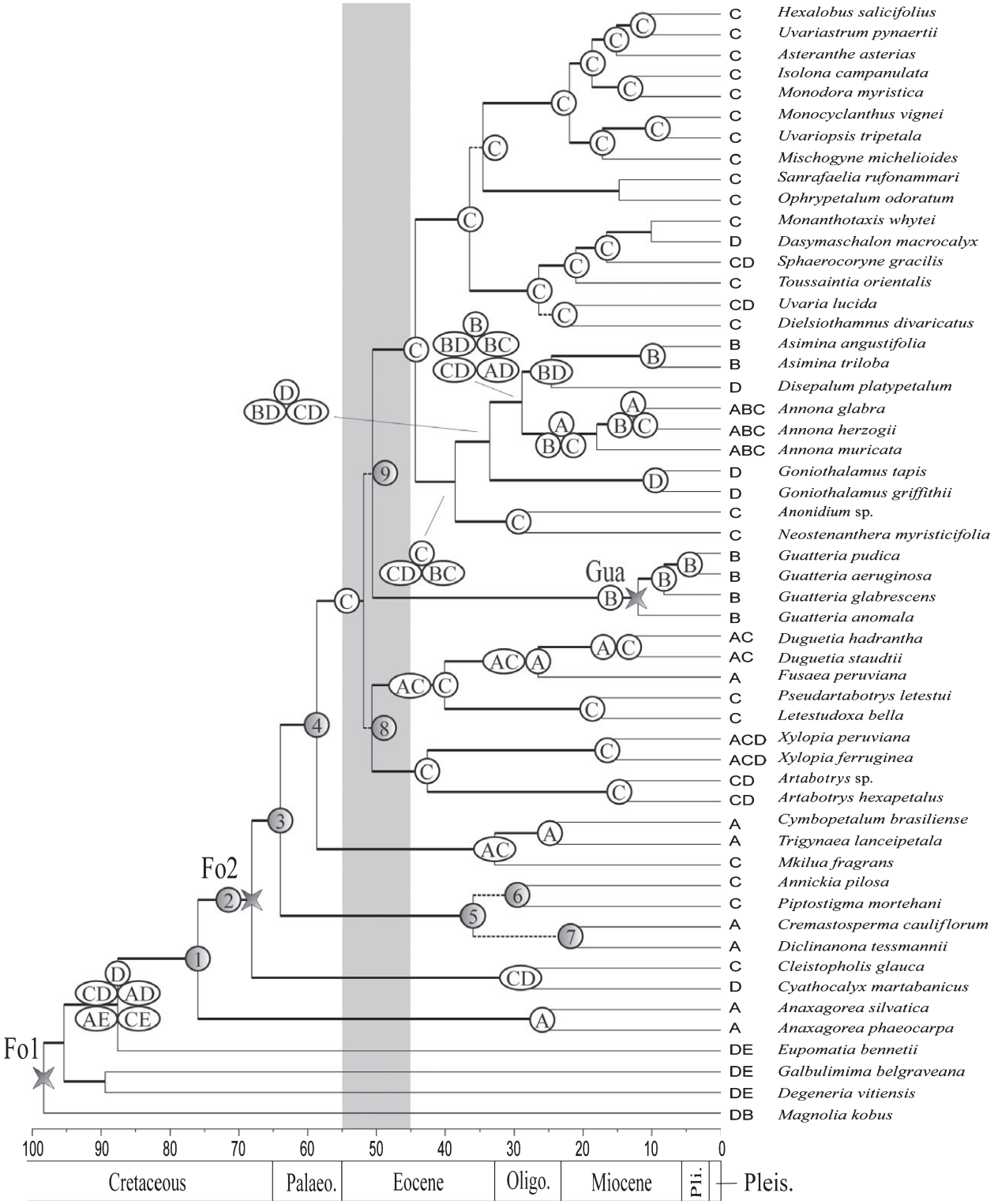

Figure 2 Maximum clade credibility (MCC) tree with mean estimated ages for selected clades in the Annonaceae using three calibration points (Fo1, early Cenomanian fossil (98 Ma); Fo2, fossil seed from the Maastrichtian (68 Ma); Gua, crown node age estimate for *Guatteria* (11.4 Ma)). Ancestral areas (A, South America; B, North America, including Central America and the Caribbean; C, Africa, including Madagascar; D, Asia; E, Australia) are scored for genera as a whole and do not represent the ancestral areas of the individual species. Nodes for which the marginal posterior probabilities of the ancestral area(s) were equal to 1 are plotted (i.e. no variation in the reconstructions after the 10,000 dispersal–vicariance analyses (DIVA) on the last 10,000 posterior trees). The nine nodes for which several alternative reconstructions were found are given in Table 4. The grey band indicates the timing of the North Atlantic Land Bridge. Thick branches have a posterior probability > 0.85; dashes branches have a posterior probability < 0.85.

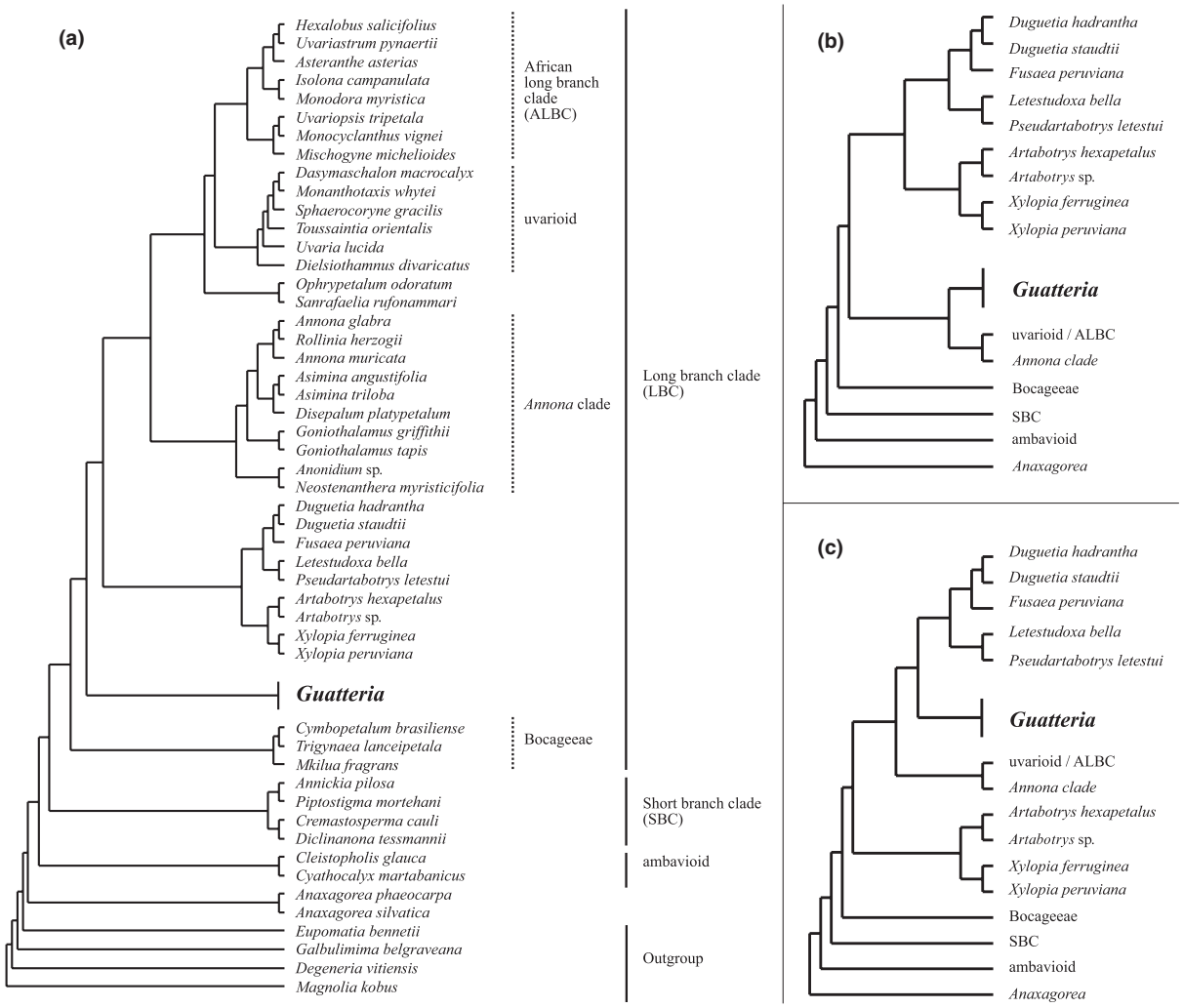

Figure 3 Recovered alternative placements of *Guatteria* by different methods of analysis: (a) most parsimonious tree as recovered by maximum parsimony, (b) maximum clade credibility tree as obtained from Bayesian analysis, (c) maximum likelihood tree obtained from RAxML.

(Nylander *et al.*, 2008). Ancestral areas are then displayed as marginal posterior distributions, providing an indication of the uncertainty of the reconstructions at nodes. The last 10,000 trees from the combined tree file were used to reconstruct ancestral areas using the program DIVA 1.2 (Ronquist, 1997) constraining ancestral distributions to contain at most four unit areas (maxareas = 4). Because DIVA only accepts one tree at a time, a Perl script allowing the automation of the procedure was used. The frequency of ancestral areas for clades was then plotted on the maximum clade credibility (MCC) tree. When multiple ancestral areas for a node on any given tree were reconstructed they were recorded as a fraction (e.g. A, 1/2; B, 1/2; Nylander *et al.*, 2008). The likely ancestral distribution of genera was scored following Doyle *et al.* (2004; see Appendix S1 and Fig. 2). Source of modifications for individual genera or of ancestral areas for newly added ones are also indicated there. One noteworthy change is the scoring of the ancestral area of *Guatteria*. It was not scored as South American, but as North/Central American (following scoring by Doyle *et al.*, 2004; for endemic Central American genera such as *Sapranthus* and *Desmopsis*) because the ancestral distribution of the genus is shown to be Central American (Erkens *et al.*, 2007b). This is important because a higher taxon should be coded for the likely ancestral distribution, not as being distributed in all areas where descendants occur (Drummond *et al.*, 2007). Ancestral areas (A, South America; B, North America, including Central America and the Caribbean; C, Africa, including Madagascar; D, Asia; E, Australia; Fig. 2) are scored for genera as a whole and do not represent the ancestral areas of the individual species. This could be done because the aim was not to determine the ancestral areas of individual genera, but the ancestral areas of nodes deeper down into the tree.

RESULTS

Phylogenetic analyses

All analyses produced highly congruent trees with no supported incongruence between clades as judged by eye (see below and Fig. 3). The parsimony analysis of all molecular data

combined (8804 aligned characters, of which 1666 were excluded due to ambiguities in the alignment, 1835 (20.8%) were parsimony informative and 34 were coded as gaps) yielded one most parsimonious tree of 6234 steps.

All 12 independent runs of the Bayesian analyses for each calibration strategy reached stationarity very quickly (within the first 10,000 generations) and all converged to the same likelihood values for all parameters. These runs were combined after a small burn-in of 200 generations each into one long run of *c.* 75 million generations. Most parameters had reached ESS values over 200 and were thus deemed reliable.

The tree files were also combined after a small burn-in of 200 generations each. However, this resulted in a large unmanageable file. Thus, the original tree files were once again combined, but this time trees were resampled every 5000 instead of 1000 generations using LogCombiner. This combined tree file totalled 16,000 sampled trees and was used to generate the MCC tree (Fig. 2) with the program TreeAnnotator that is part of the beast package (Drummond & Rambaut, 2007). All resulting MCC tree topologies as well as support values were identical under the four alternative calibration strategies.

Phylogenetic relationships

The long-branch clade (LBC) is recovered as a strongly supported monophyletic group (bootstrap support (BS) 100%; maximum likelihood support (ML) 100; posterior probability (PP) 1.00). The most basal group within the LBC is the Bocageeae (BS 100%; ML 100; PP 1). Three larger and previously recognized groups are also recovered with high support: the *Annona* clade (BS 100%; ML 100; PP 1), the uvarioids (BS 100%; ML 100; PP 1) and the African LBC excluding the genera *Ophrypetalum* and *Sanrafaelia* (BS 100%; ML 100; PP 1). The relationship between the *Duguetia* clade, the *Xylopia–Artabotrys* clade and *Guatteria* is unclear (the position of *Guatteria* as sister to the remaining genera of the LBC has no significant support: BS < 50%; ML < 50%; PP 0.40). Finally, the position of *Diclinanona tessmannii* is surprising. It is placed with strong support within the so-called short-branch clade (SBC) of Annonaceae (BS 100%; ML 100; PP 1).

Age estimates

All four different calibration strategies yielded age estimates that did not significantly differ (Table 3). The crown node estimates for Annonaceae ranged from 72.7 to 76.0 Ma. For the split of the ambavioids from the clade comprising the SBC and LBC, estimates are between 64.5 and 65.7 Ma, a younger age than the 68 Ma assigned to that node on the basis of the fossil calibration point. However, these ages fall within each other's confidence intervals. The MRCA of the LBC and SBC together is estimated between 60.6 and 63.8 Ma. The LBC minus the Bocageeae has an estimated age of 55.8–58.5 Ma. The stem node age estimate of *Guatteria* ranges between 49.2 and 51.3 Ma whereas the crown node age estimate ranges between 17.0 and 17.8 Ma. Again, the estimated age is older than the age assigned to that node for calibration, but all estimates fall within each other's confidence intervals. In short, the different calibration strategies do not significantly differ in their age estimates for the relevant nodes.

Biogeography

For a large number of nodes the marginal probabilities were maximum for the set of areas (i.e. no variation in the reconstruction). These areas are represented in Fig. 2. However, for nine nodes, including the one connecting *Guatteria*

Table 3 Age estimates (million years ago, Ma) for selected nodes in the phylogeny of the Annonaceae with 95% lower and upper highest posterior distribution (i.e. 95% confidence interval).

Calibration strategy		Magnoliaceae	Annonaceae	Ambavioids–(LBC–SBC) split	LBC–SBC split	LBC	*Guatteria*–(*Annona*–uvarioids–aLBC) split	*Guatteria*
Fo1	Mean	98 (cal.)	74.0	65.7	61.8	56.8	49.9	17.0
	95% lower	96.0	62.6	54.3	50.9	46.2	39.7	8.6
	95% upper	99.9	84.7	76.6	72.9	67.4	59.6	25.8
Fo1 and Fo2	Mean	98 (cal.)	76.0	68 (cal.)	63.8	58.5	51.3	17.8
	95% lower	96.1	70.2	66.0	59.0	52.4	43.4	8.9
	95% upper	100.0	82.7	69.9	67.9	64.0	58.3	27.0
Fo1 and Gua	Mean	98 (cal.)	72.7	64.5	60.6	55.8	49.2	11.4 (cal.)
	95% lower	96.0	60.1	53.0	49.6	45.5	39.8	9.5
	95% upper	99.9	83.9	75.7	71.4	65.8	58.7	14.5
Fo1, Fo2 and Gua	Mean	98 (cal.)	75.7	68 (cal.)	63.7	58.5	51.1	11.4 (cal.)
	95% lower	96.0	70.3	65.9	58.9	52.3	43.9	9.6
	95% upper	99.9	81.9	69.8	67.9	64.1	58.1	14.6

Fo1, early Cenomanian fossil (98 Ma); Fo2, fossil seed from the Maastrichtian (68 Ma); Gua, crown node age estimate for *Guatteria* (11.4 Ma); (cal.), indicates age set as calibration point; LBC, long branch clade; SBC, short branch clade; aLBC, African long branch clade.

Table 4 Marginal probability of reconstructed ancestral areas for selected nodes in the phylogeny of the Annonaceae. Node numbers correspond to those in Fig. 2.

Node number	Marginal probability of ancestral area			
1	AC = 0.48	AD = 0.48	A = 0.04	
2	C = 0.46	CD = 0.46	AC = 0.04	AD = 0.04
3	C = 0.92	A = 0.04	AC = 0.04	
4	C = 0.92	AC = 0.055	AB = 0.015	A = 0.01
5	C = 0.97	AC = 0.03		
6	C = 0.97	AC = 0.03		
7	A = 0.80	AC = 0.20		
8	C = 0.985	AC = 0.014	AB = 0.001	
9 (*Guatteria*)	BC = 0.99	AB = 0.01		

A, South America; B, North America (including Central America and the Caribbean); C, Africa (including Madagascar); D, Asia; E, Australia.

with its sister group, several areas were reconstructed as probable after the Bayes-DIVA analysis (Table 4). For the ancestral area of *Guatteria* and its sister group the area North America–Africa was reconstructed for 99% after 10,000 DIVA analyses, while South America–North America was found just 1% of the time. The clade containing *Anonidium* and *Annona* appeared particularly complex in terms of reconstructed ancestral areas.

DISCUSSION

Uncertain phylogenetic position of *Guatteria*

The results presented here for the LBC of Annonaceae (Figs 2 & 3) are congruent with several topologies found before (Richardson *et al.*, 2004; Couvreur *et al.*, 2008b). Most relationships are resolved and highly supported (Fig. 3), but two aspects of the tree draw attention. First, despite the addition of yet another plastid marker, ambiguity remains with respect to the placement of four lineages: *Guatteria*, the *Duguetia* clade, the *Xylopia–Artabotrys* clade and the remainder of the species of the LBC (the African–LBC, the *Annona*-group and the uvarioids; Fig. 3). This phylogenetic uncertainty is mainly due to different placements of *Guatteria* (Fig. 3) and might be the result of the fact that taxa with particularly long stem lineages can be difficult to place (Felsenstein, 1978; Steel, 1993). This so-called long-branch attraction might influence the reconstruction of relationships in the LBC because of the long branch subtending *Guatteria*, especially in combination with the insignificantly supported short branch that groups the *Xylopia–Artabotrys* clade and the *Duguetia* clade, and the short branch that subtends *Guatteria* and its sister group (Fig. 2). There might have been little opportunity for evolutionary changes to accumulate in the short time span during which the base of the LBC diversified. Long-branch attraction then results, because a relatively large amount of divergence time separates members of crown groups compared with the relatively small amount of divergence time that separates the ancestors of their stem lineages (Whitfield & Lockhart, 2007). These long stem lineages provide significant opportunity for the loss of phylogenetic signal through substitutional saturation, a situation that has been documented for the *psbA–trnH* marker in Annonaceae (Erkens, 2007). Furthermore, rate shift analyses uncover a small (but non-significant) increase in diversification rate at the base of the LBC (Erkens, 2007). However, for definitive insight into the cause of this ambiguity further analyses should be done that go beyond the scope of this paper.

The second interesting aspect in the trees of Figs 2 and 3 is that *D. tessmannii* has a completely different placement when compared to the analysis of Richardson *et al.* (2004). There it was placed in an unsupported clade together with, amongst others, *Goniothalamus* and *Annona*, but here it is found to belong to the short-branch clade of Annonaceae. This misplacement by Richardson *et al.* (2004) was the result of dubious *trnL–F* and *rbcL* sequences. Resequencing of these regions yielded the position of *Diclinanona* found here, a position also supported by the other plastid regions sampled in this study.

The age estimates obtained here are for a large part congruent with estimates obtained before (Table 3; Doyle *et al.*, 2004; Richardson *et al.*, 2004; Pirie *et al.*, 2006). It is interesting to note, however, that the crown node of the short-branch clade is estimated at 35.9 Ma (19.6–53.8 Ma), *c.* 20 Myr earlier than previous estimates. The two nodes of importance here are the stem and crown node of *Guatteria.* The stem node age estimates for *Guatteria* range between 49.2 and 51.3 Ma and are in the same range as the early Eocene estimate of Richardson *et al.* (2004), the only estimate so far for this node. Age estimates for the crown node of *Guatteria* are older when obtained from Annonaceae-wide analyses (Richardson *et al.*, 2004; Pirie *et al.*, 2005) than when obtained from an analysis of the genus itself and some outgroups (Erkens *et al.*, 2007b; although ages in the range found here were also reported). Erkens *et al.* (2007b) estimated the crown node age of *Guatteria* based on penalized likelihood (PL) and nonparametric rate smoothing (NPRS), finding that estimates calculated with PL were always younger than those calculated with NPRS. This effect was attributed to the fact that NPRS tends to over-smooth short branches (Sanderson, 1997) and many short branches were present in the recovered tree. Their minimum ages, although younger than the estimated ages here, fall within the 95% confidence interval of the estimates obtained in the present study. It has been shown that taxon sampling can influence age estimates, especially when only a few species of a species-rich taxon are included (as is done here; Pirie *et al.*, 2005). Furthermore, sampling error in age estimation is bad in data sets with isolated long branches, but possibly equally bad in data sets with very short branches (Renner, 2004b). This means that we cannot choose between dates obtained in studies on the LBC as a whole (including only a few species per genus) and the dates obtained from the *Guatteria*-centred study, since many short branches were present in the study by Erkens *et al.* (2007b). However, it can

be concluded that the MRCA of *Guatteria* originated somewhere in the Miocene.

Biogeographical scenario

The Bayes-DIVA analysis shows that the most likely distribution of the stem lineage of *Guatteria* is the African–North American region. This result is robust and does not change even when accounting for phylogenetic uncertainty in the placement of *Guatteria*. Furthermore, not only was the ancestral area determined, but also a time window in which the biogeographical scenario could have taken place, at *c.* 49–52 and 11–18 Ma. This allows a distinction between three rival hypotheses: migration of the MRCA of *Guatteria* from Africa via Europe, from Asia or from South America. Both timing and biogeography support the first hypothesis (Fig. 1a). It is known that migration from eastern Laurasia (Europe) to western Laurasia (North America) may have been facilitated by the boreotropical NALB (Tiffney, 1985a; Tiffney & Manchester, 2001), which spanned the North Atlantic during the early to mid Eocene, and that paratropical lineages could be freely exchanged between the New World and the Old World in the Northern Hemisphere. Global temperatures during the Eocene were highest during this time period (Zachos *et al.*, 2001) and it is known that tropical vegetation occurred on this land corridor (Wolfe, 1975; Zachos *et al.*, 2001). Megathermal taxa have been documented up to 50–60° N (Chandler, 1964; Jolly, 1998). Lineages originating in northern Gondwana crossed into Laurasia and became established along the southern coast of Eurasia, which had a tropical climate in the Eocene (Doyle & Le Thomas, 1997) as can be seen from London Clay Annonaceae fossil occurrences (Chandler, 1964). This is further supported by fossils from Palaeocene Egypt (Chandler, 1954). Migration between Gondwana and Laurasia in the late Cretaceous–early Tertiary could have been possible as the Tethys narrowed (Richardson *et al.*, 2004). There is much evidence to support the claim that the pathway via the NALB plays an important role in explaining the global distribution of tropical plant families (Annonaceae, Doyle & Le Thomas, 1997; Lauraceae, Chanderbali *et al.*, 2001; Melastomataceae, Renner *et al.*, 2001; Malpighiaceae, Davis *et al.*, 2002; Burseraceae, Weeks *et al.*, 2005; Meliaceae, Muellner *et al.*, 2006). The groups with tropical amphi-Atlantic disjunctions split up at a time consistent with the disruption of boreotropical ranges around the Eocene–Oligocene boundary (Richardson *et al.*, 2004). Many familiar groups might now be restricted to one hemisphere or the other were it not for this Laurasian–North American (or vice versa) migration (Muellner *et al.*, 2006). Interestingly, findings support a Laurasian ancestry of most of today's tropical American diversity of Lauraceae, with estimates for the arrival of the *Ocotea* complex in South America somewhere during the Miocene (Chanderbali *et al.*, 2001; Renner, 2004b). This pattern seems the same as documented here for *Guatteria*, indicating that the evolutionary histories of *Ocotea* and *Guatteria* show comparable phylogenetic patterns and timing, as suggested by Erkens *et al.* (2007b).

A rival hypothesis is that *Guatteria* migrated into Central America via the BLB (Fig. 1b). Although the timing of the split between *Guatteria* and its sister clade (irrespective of which taxon that is) cannot rule out a BLB hypothesis, the Bayes-DIVA analysis excludes Asia as the ancestral area for *Guatteria*. Additionally, from a palaeontological point of view it would be unlikely that this route was taken. Deciduous angiosperms are inferred to have crossed the BLB in the early Tertiary (e.g. *Aesculus*) and deciduous taxa used the BLB later in the Tertiary (Wen, 1999). Although paratropical taxa, such as Annonaceae, might be known from southern Alaska (if they are not displaced; McKenna, 1983), they are less commonly reported from the Siberian side and from further south in Kamchatka (although putative annonaceous leaves from the Middle Eocene have been reported from this location; Budantsev, 1994, 1997). This may indicate that the BLB was less travelled by paratropical taxa than was the NALB (Tiffney & Manchester, 2001). Data also suggest that the BLB lay perhaps at 75–80° N in the early Tertiary (Tiffney & Manchester, 2001; Morley, 2003) and that this would place it in an area of extended winter darkness. Given the known dependence of Annonaceae on high temperature and precipitation (Punyasena *et al.*, 2008) it is highly unlikely that these evergreen angiosperms were able to cross this land bridge.

Finally, there is no evidence to support the third hypothesis, dispersal of the ancestor of *Guatteria* out of Africa via South America into Central America (Fig. 1c). Timing as well as biogeography are not in line with this hypothesis. Also, it would require extinction of all the predicted stem relatives of *Guatteria* in South America. Richardson *et al.* (2004) suggested that the current distribution patterns of Annonaceae in general may be best explained by an origin in western Gondwana, followed by interchange between Africa and South America by short-distance dispersal across the opening Atlantic and dispersal via the Walvis Ridge or the Sierra Leone Rise. However, it seems that this route was not taken by the ancestors of *Guatteria.*

A fourth, alternative, hypothesis to the three hypotheses discussed above is long-distance (oceanic) dispersal between Africa and South America. That such dispersal events occur, even in genera in which such an event was deemed unlikely, can be seen from the literature (e.g. Dick *et al.*, 2003, 2007; Renner, 2004a). However, the fossil record in general suggests that plants are unlikely to disperse from one continent to another unless an appropriate land connection exists and the climate is appropriate (Morley, 2003). Furthermore, long-distance dispersal is an *ad hoc* explanation that can always be inferred when a disjunct distribution is seen that cannot be explained by other factors. Unfortunately, it is non-testable due to its random nature. In the specific case of *Guatteria*, several other factors rule against this hypothesis. Genera for which long-distance dispersal has been documented show a trans-Atlantic or trans-Pacific distribution of extant species (Renner, 2004a). *Guatteria* does not have such a trans-Atlantic distribution (as can be seen in other Annonaceae genera such as *Annona* or *Duguetia*). Also, the DIVA analysis uncovers a

high degree of geographical structure (Fig. 2), suggesting an explanation other than undirected long-distance dispersal. In fact, it has already been suggested that long-distance transoceanic dispersal has had a relatively minor role in determining present distribution patterns in Annonaceae in general (Richardson *et al.*, 2004). In conclusion, although long-distance dispersal cannot be ruled out as an explanation (since it never can), our favoured hypothesis for the origin of the MRCA of *Guatteria* is based on a migration across the NALB, as described above.

Isolated position of *Guatteria*

The scenario of migration from Africa via Europe into Central and South America could explain the phylogenetically (long branch) and morphologically (unique character states) isolated position of *Guatteria*. The late Palaeocene/early Eocene warm interval allowed megathermal angiosperms to extend their ranges further poleward than at any time during the period in which they were dominant and these taxa could disperse unhindered over remarkably wide areas (Morley, 2003). However, from the Middle Eocene onwards global climates underwent stepwise cooling (Zachos *et al.*, 2001). Frost-intolerant taxa were removed from much of North America (many families became extinct, at least in the latitudes between 40 and 60° N), unless they dispersed to lower latitudes. By this mechanism the northern mid-latitude Boreotropical forests were replaced with frost-tolerant temperate vegetation (Morley, 2003). Around 11 Ma cool temperate plants first appeared in northern South America, as elements that dispersed from the north (Cox & Moore, 2005).

The estimated stem to crown node ages of *Guatteria* (49–52 to 11–18 Ma) are in line with this southward movement of megathermal forests (Wolfe, 1975; Hallam, 1994; Morley, 2003). Climatic changes might have caused extinction of any intermediate taxa present at higher latitudes. In this scenario, the ancestors of *Guatteria* reached Central America in the mid/early Miocene (between 18 and 11 Ma; Table 3), just before the northern climate became too hostile. The configuration of Central America was favourable to its ancestors so that they could establish there and speciate. Because of the extinction of any intermediate North American and European taxa, a phylogenetically isolated genus remains sitting on a long branch. Also, all taxa with intermediate morphological character states were removed and therefore it is now impossible to connect extant *Guatteria* to other Annonaceae genera morphologically. The short branches subtending the *Duguetia* clade and the *Xylopia–Artabotrys* clade suggest that extinction was of less importance for their stem lineages.

Current distribution and diversity pattern of *Guatteria* and the boreotropical hypothesis

The boreotropical biogeographical history of *Guatteria* (as presented here and in Erkens *et al.*, 2007b) is reflected in its current distribution and diversity pattern. The migration of its ancestors from Africa via the NALB into North and Central America was followed by extinction of its European and North American lineages as explained above. As a result, its oldest extant lineages can be found in Central America (Erkens *et al.*, 2007b). These lineages have diversified and given rise to mainly endemic Central American species, which are morphologically clearly distinguishable. In the Miocene a transoceanic migration of one of these lineages took place from Central into South America (Erkens *et al.*, 2007b). Subsequently, a major diversification occurred in South America (mainly in the Amazon, where more than half of all extant species diversity now occurs). More recently, several small migrations occurred from lineages out of South America into Central America (Erkens *et al.*, 2007b). These species are morphologically difficult to distinguish from each other, as is to be expected of young species.

As underlined by Pennington & Dick (2004), the contribution of immigrant taxa to the South American rainforest flora needs a re-evaluation in order to properly understand the historical assembly of its biodiversity. The importance of such immigrant taxa has already been suggested on the basis of fossil (e.g. Leopold & MacGinitie, 1972) and pollen (e.g. Simpson, 1975) data but is further strengthened by phylogenetic data (e.g. Bell & Donoghue, 2005; Hughes & Eastwood, 2006; Nathan, 2006) and by the analysis of geographical diversity patterns (e.g. Bjorholm *et al.*, 2006). *Guatteria* provides another case in which migration from North America contributed substantially to South American speciation and diversity.

CONCLUSIONS

The availability and timing of appropriate dispersal pathways limits the number of dispersal opportunities open to any particular taxon. Knowledge of these pathways allows the taxonomist to make better judgements regarding biogeographical history (Morley, 2003). Here it is shown that such data can indeed provide a plausible scenario for the history of *Guatteria* despite the fossil record being deficient. The estimated stem to crown node ages in combination with the Bayes-DIVA analysis imply a scenario congruent with an origin in Africa followed by dispersal across the NALB into North and Central America (and ultimately South America). The phylogenetically and morphologically isolated position of the genus is probably due to extinction in the Tertiary.

ACKNOWLEDGEMENTS

We thank Marc Florence and Jonathan Wingerath (both at the Department of Paleobiology, Smithsonian Institution, Washington) for making the effort to locate and photograph specimens of *Anonaceaephyllum culebrensis*. Also, we would like to thank Linda Klise (Division of Paleobotany, Yale Peabody Museum of Natural History, New Haven) for sending collection data and photographs of *Anonaceaephyllum cretaceum*. Johan Nylander is thanked for providing the Perl script

used for the Bayes-DIVA approach. Also, Jim Doyle is thanked for critically reading the manuscript and suggesting many useful improvements. Finally, the authors thank the referees for providing comments on the manuscript.

REFERENCES

Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. *Second international symposium on information theory* (ed. by B.N. Petrov and F. Csaki), pp. 267–281. Akademiai Kiado, Budapest.

Bell, C.D. & Donoghue, M.J. (2005) Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms, Diversity & Evolution*, **5,** 147–159.

Bjorholm, S., Svenning, J.-C., Baker, W.J., Skov, F. & Balslev, H. (2006) Historical legacies in the geographical diversity patterns of New World palm (Arecaceae) subfamilies. *Botanical Journal of the Linnean Society*, **151,** 113–125.

Budantsev, L.Y. (1994) The fossil flora of the Paleogene climatic optimum in north eastern Asia. *Cenozoic plants and climates of the Arctic* (ed. by M.C. Boulter and H.V. Fisher), pp. 297–313. Springer, Berlin, Germany.

Budantsev, L.Y. (1997) Late Eocene flora of western Kamchatka. *Proceedings of the Komarov Botanical Institute, Russian Academy of Sciences*, **19,** 3–115.

Burnham, R.J. & Graham, A. (1999) The history of Neotropical vegetation: new developments and status. *Annals of the Missouri Botanical Garden*, **86,** 546–589.

Butt, N., Malhi, Y., Phillips, O. & New, M. (2008) Floristic and functional affiliations of woody plants with climate in western Amazonia. *Journal of Biogeography*, **35,** 939–950.

Chanderbali, A.S., Van Der Werff, H. & Renner, S.S. (2001) Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. *Annals of the Missouri Botanical Garden*, **88,** 104–134.

Chandler, M.E.J. (1954) Some Upper Cretaceous and Eocene fruits from Egypt. Bulletin of the British Museum (Natural History). *Geology*, **2,** 149–187.

Chandler, M.E.J. (1964) *The lower Tertiary floras of southern England.* Vol. 4. *A summary and survey of findings in the light of recent botanical observations*. British Museum (Natural History), London.

Chesters, K.I.M. (1955) Some plant remains from the Upper Cretaceous and Tertiary of West Africa. *Annals and Magazine of Natural History*, **12,** 489–504.

Couvreur, T.L.P., Chatrou, L.W., Sosef, M.S.M. & Richardson, J.E. (2008a) Molecular phylogenetics reveal multiple Tertiary vicariance origins of the African rain forest trees. *BMC Biology*, **6,** 54.

Couvreur, T.L.P., Richardson, J.E., Sosef, M.S.M., Erkens, R.H.J. & Chatrou, L.W. (2008b) Evolution of syncarpy and other morphological characters in African Annonaceae: a posterior mapping approach. *Molecular Phylogenetics and Evolution*, **47,** 302–318.

Cox, C.B. & Moore, P.D. (2005) *Biogeography: an ecological and evolutionary approach*, 7th edn. Blackwell Publishing, Oxford.

Cuénoud, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J. & Chase, M.W. (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, a*tpB*, and *matK* DNA sequences. *American Journal of Botany*, **89,** 132–144.

Davis, C.C., Bell, C.D., Mathews, S. & Donoghue, M.J. (2002) Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *Proceedings of the National Academy of Sciences USA*, **99,** 6833–6837.

Dick, C.W., Kobinah, A.-S. & Bermingham, E. (2003) Molecular systematic analysis reveals cryptic Tertiary diversification of a widespread tropical rain forest tree. *The American Naturalist*, **162,** 691–703.

Dick, C.W., Bermingham, E., Lemes, M.R. & Gribel, R. (2007) Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology*, **16,** 3039–3049.

Dilcher, D.L. & Crane, P.R. (1984) *Archaeanthus*: an early angiosperm from the Cenomanian of the Western Interior North America. *Annals of the Missouri Botanical Garden*, **71,** 351–383.

Donoghue, M.J. & Moore, B.R. (2003) Toward an integrative historical biogeography. *Integrative and Comparative Biology*, **43,** 261–270.

Doyle, J.J. & Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19,** 11–15.

Doyle, J.A. & Le Thomas, A. (1996) Phylogenetic analysis and character evolution in Annonaceae. *Adansonia*, **18,** 279–334.

Doyle, J.A. & Le Thomas, A. (1997) Phylogeny and geographic history of Annonaceae. *Géographie physique et Quarternaire*, **51,** 353–361.

Doyle, J.A., Sauquet, H., Scharaschkin, T. & Le Thomas, A. (2004) Phylogeny, molecular and fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliales). *International Journal of Plant Sciences*, **165,** S55–S67.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7,** 214.

Drummond, A.J., Ho, S.Y.W., Rawlence, N. & Rambaut, A. (2007) *A rough guide to BEAST 1.4*. Available at: http://beast.bio.ed.ac.uk/Main_Page (last accessed 16 April 2009).

Erkens, R.H.J. (2007) *From morphological nightmare to molecular conundrum. Phylogenetic, evolutionary and taxonomic studies on* Guatteria *(Annonaceae)*. PhD Thesis, Utrecht University, Utrecht. Available at: http://igitur-archive.library.uu.nl/dissertations/2007-0227-200317/UUindex.html (last accessed 16 April 2009).

Erkens, R.H.J., Chatrou, L.W., Koek-Noorman, J., Maas, J.W. & Maas, P.J.M. (2007a) Classification of a large and widespread genus of Neotropical trees, *Guatteria* (Annonaceae) and its three satellite genera *Guatteriella*, *Guatteriopsis* and *Heteropetalum*. *Taxon*, **56,** 757–774.

Erkens, R.H.J., Chatrou, L.W., Maas, J.W., Van Der Niet, T. & Savolainen, V. (2007b) A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. *Molecular Phylogenetics and Evolution*, **44,** 399–411.

Erkens, R.H.J., Cross, H., Maas, J.W., Hoenselaar, K. & Chatrou, L.W. (2008a) Age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. *Blumea*, **53,** 407–428.

Erkens, R.H.J., Westra, L.Y.T. & Maas, P.J.M. (2008b) Increasing diversity in the species-rich genus *Guatteria* (Annonaceae). *Blumea*, **53,** 467–514.

Fay, M.F., Swensen, S.M. & Chase, M.W. (1997) Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae). *Kew Bulletin*, **52,** 111–120.

Fay, M.F., Bayer, C., Alverson, S., de Bruijn, A.Y. & Chase, M.W. (1998) Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon*, **47,** 43–50.

Felsenstein, J. (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27,** 401–410.

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

Gandolfo, M.A., Nixon, K.C. & Crepet, W.L. (2008) Selection of fossils for calibration of molecular dating models. *Annals of the Missouri Botanical Garden*, **95,** 34–42.

Gentry, A.H. (1982) Phytogeographic patterns as evidence for a Chocó refuge. *Biological diversification in the tropics* (ed. by G.T. Prance), pp. 112–136. Columbia University Press, New York.

Gernhard, T. (2008) New analytic results for speciation times in neutral models. *Bulletin of Mathematical Biology*, **70,** 1082–1097.

Hallam, A. (1994) *An outline of Phanerozoic biogeography*. Oxford University Press, New York.

Hamilton, M.B. (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8,** 521–523.

Hermsen, E.J. & Gandolfo, M.A. (2004) Fossil identification used in molecular dating studies should be justified. *American Journal of Botany*, http://www.amjbot.org/cgi/eletters/91/6/881#33 (last accessed 3 October 2008).

Hoot, S.B., Culham, A. & Crane, P.R. (1995) The utility of *atpB* gene sequences in resolving phylogenetic relationships: comparison with *rbcL* and 18S ribosomal DNA sequences in the Lardizabalaceae. *Annals of the Missouri Botanical Garden*, **82,** 194–207.

Hughes, C. & Eastwood, R. (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences USA*, **103,** 10334–10339.

Janzen, D.H. (1967) Why mountain passes are higher in the tropics. *The American Naturalist*, **101,** 233–249.

Jolly, D.W. (1998) Early Eocene palynofloras from holes 915A, 916A, 917A, and 918D, east Greenland. *Proceedings of the Ocean Drilling Program – Scientific Results*, **152,** 221–231.

Kim, S., Park, C.-W., Kim, Y.-D. & Suh, Y. (2001) Phylogenetic relationships in family Magnoliaceae inferred from *ndhF* sequences. *American Journal of Botany*, **88,** 717–728.

Lavin, M. & Luckow, M. (1993) Origins and relationships of tropical North America in the context of the boreotropics hypothesis. *American Journal of Botany*, **80,** 1–14.

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

McKenna, M.C. (1983) Holarctic landmass rearrangement, cosmic events, and Cenozoic terrestrial organisms. *Annals of the Missouri Botanical Garden*, **70,** 459–489.

Mols, J.B., Gravendeel, B., Chatrou, L.W., Pirie, M.D., Bygrave, P.C., Chase, M.W. & Kessler, P.J.A. (2004) Identifying clades in Asian Annonaceae: monophyletic genera in the polyphyletic Miliuseae. *American Journal of Botany*, **91,** 590–600.

Morley, R.J. (2000) *Origin and evolution of tropical rain forests*. Wiley, Chichester.

Morley, R.J. (2003) Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics*, **6,** 5–20.

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. *Molecular Phylogenetics and Evolution*, **40,** 236–250.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403,** 853–858.

Nathan, R. (2006) Long-distance dispersal of plants. *Science*, **313,** 786–788.

Nylander, J.A.A. (2004) *MrModeltest v2*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander, J.A.A., Olsson, U., Alström, P. & Sanmartín, I. (2008) Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal-vicariance analysis of the thrushes (Aves: *Turdus*). *Systematic Biology*, **57,** 257–268.

Olmstead, R.G. & Sweere, J.A. (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology*, **43,** 467–481.

Olmstead, R.G., Michaels, H.J., Scott, K.M. & Palmer, J.D. (1992) Monophyly of the Asteridae and identification of major lineages inferred from DNA sequences of *rbc*L. *Annals of the Missouri Botanical Garden*, **79,** 249–265.

Pennington, R.T. & Dick, C.W. (2004) The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 1611–1622.

Pirie, M.D., Chatrou, L.W., Erkens, R.H.J., Maas, J.W., Van Der Niet, T., Mols, J.B. & Richardson, J.E. (2005) Phylogeny reconstruction and molecular dating in four Neotropical genera of Annonaceae: the effect of taxon sampling in age

estimations. *Plant species-level systematics: new perspectives on pattern and process*, Regnum Vegetabile 143 (ed. by F.T. Bakker, L.W. Chatrou, B. Gravendeel and P.B. Pelser), pp. 149–174. A. R. G. Gantner Verlag, Liechtenstein.

Pirie, M.D., Chatrou, L.W., Mols, J.B., Erkens, R.H.J. & Oosterhof, J. (2006) 'Andean-centred' genera in the short-branch clade of Annonaceae: testing biogeographic hypotheses using phylogeny reconstruction and molecular dating. *Journal of Biogeography*, **33,** 31–46.

Punyasena, S.W., Eshel, G. & McElwain, J.C. (2008) The influence of climate on the spatial patterning of Neotropical plant families. *Journal of Biogeography*, **35,** 117–130.

Raven, P.H. & Axelrod, D.L. (1974) Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden*, **61,** 539–673.

Renner, S. (2004a) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, **165,** S23–S33.

Renner, S.S. (2004b) Variation in diversity among Laurales, Early Cretaceous to Present. *Biologiske Skrifter*, **55,** 441–458.

Renner, S.S., Clausing, G. & Meyer, K. (2001) Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany*, **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. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 1495–1508.

Ronquist, F. (1997) Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46,** 195–203.

Rull, V. (2008) Speciation timing and neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17,** 2722–2729.

Sanderson, M.J. (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, **14,** 1218–1231.

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.

Scharaschkin, T. & Doyle, J.A. (2005) Phylogeny and historical biogeography of *Anaxagorea* (Annonaceae) using morphology and non-coding chloroplast sequence data. *Systematic Botany*, **30,** 712–735.

Simmons, M.P. & Ochoterena, H. (2000) Gaps as characters in sequence-based phylogenetic analysis. *Systematic Biology*, **49,** 369–381.

Simpson, B.B. (1975) Pleistocene changes in the flora of the high tropical Andes. *Paleobiology*, **1,** 273–294.

Simpson, B.B. (1980) *Splendid isolation: the curious history of South American mammals*. Yale University Press, New Haven, CT.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22,** 2688–2690.

Stamatakis, A., Hoover, P. & Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology*, **57,** 758–771.

Steel, M.A. (1993) Confidence in evolutionary trees from biological sequence data. *Nature*, **364,** 440–442.

Swofford, D.L. (2003) *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10*. Sinauer Associates, Sunderland, MA.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991) Universal primers for amplification of the three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17,** 1105–1109.

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

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

Tiffney, B.H. & Manchester, S.R. (2001) The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the northern hemisphere Tertiary. *International Journal of Plant Sciences*, **162,** S3–S17.

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. *Molecular Phylogenetics and Evolution*, **35,** 85–101.

Wen, J. (1999) Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology and Systematics*, **30,** 421–455.

Whitfield, J.B. & Lockhart, P.J. (2007) Deciphering ancient rapid radiations. *Trends in Ecology and Evolution*, **22,** 258–265.

Wolfe, J.A. (1975) Some aspects of plant geography of the northern hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden*, **62,** 264–279.

Young, K.R., Ulloa, C.U., Luteyn, J.L. & Knapp, S. (2002) Plant evolution and endemism in Andean South America: an introduction. *Botanical Review*, **38,** 4–21.

Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292,** 686–693.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Taxa, ancestral areas, voucher information and GenBank/EBI accession numbers of specimens used in this study.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BIOSKETCHES

Roy H. J. Erkens is an assistant professor at Utrecht University. He is currently engaged in phylogeographical and population genetic work on the colonization of the Dutch reclaimed lands by mosses and ferns. He also works on the phylogenetics, biogeography and systematics of the family Annonaceae, in particular the species-rich genus *Guatteria*.

Jan W. Maas is a research assistant at the Department of Biology, Utrecht University, and supports several projects with his molecular work.

Thomas L. P. Couvreur is a post-doctoral researcher at the New York Botanical Garden where he works on the systematics and evolution of woody Boraginaceae. He is also actively engaged in systematic and biogeographical studies in the families Annonaceae and Arecaceae.

Editor: Jens-Christian Svenning