Naturalis Repository

Age and historical biogeography of thepantropically distributed Spathelioideae(Rutaceae, Sapindales)

Marc S. Appelhans, Paul J.A. Keßler, Erik Smets, Sylvain G. Razafimandimbison, Steven B. Janssens

Downloaded from:
https://doi.org/10.1111/j.1365-2699.2012.02686.x

Article 25fa Dutch Copyright Act (DCA) - End User Rights

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with consent from the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available following a reasonable period after the work was first published, provided that reference is made to the source of the first publication of the work.

This publication is distributed under the Naturalis Biodiversity Center 'Taverne implementation' programme. In this programme, research output of Naturalis researchers and collection managers that complies with the legal requirements of Article 25fa of the Dutch Copyright Act is distributed online and free of barriers in the Naturalis institutional repository. Research output is distributed six months after its first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and copyrights owner(s) of this work. Any use of the publication other than authorized under this license or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the department of Collection Information know, stating your reasons. In case of a legitimate complaint, Collection Information will make the material inaccessible. Please contact us through email: collectie.informatie@naturalis.nl. We will contact you as soon as possible.

Journal of Biogeography (*J. Biogeogr.*) (2012) **39**, 1235–1250

Journal of Biogeography

ORIGINAL ARTICLE

Age and historical biogeography of the pantropically distributed Spathelioideae (Rutaceae, Sapindales)

Marc S. Appelhans[1,2]*, Paul J.A. Keßler[1,2], Erik Smets[1,2,3],
Sylvain G. Razafimandimbison[4] and Steven B. Janssens[3]

[1]Hortus Botanicus Leiden, PO Box 9516, 2300 RA Leiden, The Netherlands, [2]Netherlands Centre for Biodiversity Naturalis (section NHN), Leiden University, PO Box 9514, 2300 RA Leiden, The Netherlands, [3]Laboratory of Plant Systematics, Institute of Botany and Microbiology, KU Leuven, Kasteelpark Arenberg 31, PO Box 2437, BE-3001 Leuven, Belgium, [4]Department of Botany, Bergius Foundation, Stockholm University, SE-10691 Stockholm, Sweden

*Correspondence: Marc Appelhans, Hortus Botanicus Leiden, PO Box 9516, 2300 RA Leiden, The Netherlands.
E-mail: AppelhansM@si.edu

ABSTRACT

Aim The family Rutaceae (rue family) is the largest within the eudicot order Sapindales and is distributed mainly in the tropical and subtropical regions of both the New World and the Old World, with a few genera in temperate zones. The main objective of this study is to present molecular dating and biogeographical analyses of the subfamily Spathelioideae, the earliest branching clade (which includes eight extant genera), to interpret the temporal and spatial origins of this group, ascertaining possible vicariant patterns and dispersal routes and inferring diversification rates through time.

Location Pantropics.

Methods A dataset comprising a complete taxon sampling at generic level (83.3% at species level) of Spathelioideae was used for a Bayesian molecular dating analysis (BEAST). Four fossil calibration points and an age constraint for Sapindales were applied. An ancestral area reconstruction analysis utilizing the dispersal–extinction–cladogenesis model and diversification rate analyses was conducted.

Results Dating analyses indicate that Rutaceae and Spathelioideae are probably of Late Cretaceous origin, after which Spathelioideae split into a Neotropical and a Palaeotropical lineage. The Palaeotropical taxa have their origin inferred in Africa, with postulated dispersal events to the Mediterranean, the Canary Islands, Madagascar and Southeast Asia. The lineages within Spathelioideae evolved at a relatively constant diversification rate. However, abrupt changes in diversification rates are inferred from the beginning of the Miocene and during the Pliocene/Pleistocene.

Main conclusions The geographical origin of Spathelioideae probably lies in Africa. The existence of a Neotropical lineage may be the result of a dispersal event at a time in the Late Cretaceous when South America and Africa were still quite close to each other (assuming that our age estimates are close to the actual ages), or by Gondwanan vicariance (assuming that our age estimates provide minimal ages only). Separation of land masses caused by sea level changes during the Pliocene and Pleistocene may have been triggers for speciation in the Caribbean genus *Spathelia*.

Keywords

Ancestral area reconstruction, diversification rates, LTT plots, molecular dating, pantropical distribution, phylogeny, Rutaceae, *Spathelia*, Spathelioideae.

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/j.1365-2699.2012.02686.x

INTRODUCTION

The family Rutaceae (rue family) is the largest within the eudicot order Sapindales and contains approximately 155 genera and 2100 species (Kubitzki *et al.*, 2011). Members of the family are mainly distributed in the tropical and subtropical regions of both the New World and the Old World, with only a few genera present in temperate zones. The highest species diversity is found in Australasia (Kubitzki *et al.*, 2011).

To date, only two dated phylogenies are available for Rutaceae, and both focus on only a part of the family (Aurantioideae: Pfeil & Crisp, 2008; the *Ruta* and *Cneoridium*/*Haplophyllum* clades: Salvo *et al.*, 2010). A detailed, dated phylogeny and an ancestral area reconstruction (AAR) of the whole family are not yet feasible due to the lack of resolution and support, as well as incomplete taxon sampling in the clade that contains the majority of taxa. This clade contains the former subfamilies Toddalioideae and Flindersioideae, as well as most former Rutoideae, with the exception of the type genus *Ruta* L. and its relatives (Ruteae) (Chase *et al.*, 1999; Groppo *et al.*, 2008; Salvo *et al.*, 2010). This clade is hereinafter named Toddalioideae s.l.

The present study focuses on the subfamily Spathelioideae (= Cneoroideae *sensu* Kubitzki *et al.*, 2011), which is the earliest branching clade of Rutaceae whose generic and tribal limits have recently been addressed (Appelhans *et al.*, 2011). Because it is sister to the rest of Rutaceae, an AAR analysis of Spathelioideae is of particular interest for the whole family.

Spathelioideae is a species-poor subfamily showing considerable morphological diversity (Appelhans *et al.*, 2011). The subfamily consists of 29 species in eight genera: *Bottegoa* Chiov. (one species), *Cedrelopsis* Baill. (eight species), *Cneorum* L. (two species), *Dictyoloma* A.Juss. (one species), *Harrisonia* R.Br. ex A.Juss. (three species), *Ptaeroxylon* Eckl. & Zeyh. (one species), *Sohnreyia* K.Krause (four species) and *Spathelia* L. (nine species) (Appelhans *et al.*, 2011). *Harrisonia* is widespread in tropical Africa and Southeast Asia to northern Australia, while the remaining genera have rather narrow distribution ranges (e.g. *Cedrelopsis*, endemic to Madagascar; *Cneorum*, endemic to the western Mediterranean and the Canary Islands; *Spathelia*, endemic to the Caribbean) (Appelhans *et al.*, 2011). Furthermore, most species of *Spathelia* are also narrow endemics (Beurton, 2008). Despite the small and largely non-overlapping distribution areas of the genera and their low number of species, the subfamily as a whole is distributed pantropically. Spathelioideae is divided into two clades, one being strictly Neotropical and the other being Palaeotropical.

The combination of monotypic genera, narrow endemicity and pantropical distribution of the group makes the Spathelioideae particularly interesting for biogeographical studies. The small number of taxa overall makes a biogeographical analysis at species level feasible. The goals of this study are: (1) to identify when Spathelioideae emerged and to assess whether the split into strictly Neotropical and Palaeotropical subclades indicates a vicariance pattern (break-up of Gondwana) or a long-distance dispersal pattern; (2) to determine the biogeographical history within the clades, especially for taxa on islands (Canary Islands, Caribbean Islands, Madagascar); (3) to investigate whether diversification rates were constant through time or changed, correlating with climatic fluctuations in different geological epochs; and (4) to shed further light on the geographical origin of the Rutaceae family.

MATERIALS AND METHODS

Taxon sampling

The taxon sampling used in this study is largely based on that used to produce the phylogeny of Spathelioideae by Appelhans *et al.* (2011) and it contains all eight genera of Spathelioideae and 83.3% of its species (25 out of 30, 29 described and one undescribed species; see Appendices S1 & S2 in Supporting Information). Only one species of *Sohnreyia* and four species of *Cedrelopsis* were not sampled in this study. Additionally, we included taxa from all families and subfamilies of Sapindales in order to use a published age constraint (Magallón & Castillo, 2009) for Sapindales in the molecular dating analyses. *Theobroma cacao* L. and *Gossypium hirsutum* L. (both Malvaceae) from one of the sister orders of Sapindales (Malvales; Wang *et al.*, 2009) were chosen as outgroups in the phylogenetic analyses and *T. cacao* alone was used as an outgroup in the molecular dating analyses.

Our alignment is based on the chloroplast regions *atp*B, *rbc*L and *trn*L–*trn*F, and we obtained the majority of sequences from GenBank (http://www.ncbi.nlm.nih.gov; see Appendix S1 for accession numbers). Although these three markers have been used in phylogenetic analyses in all Sapindales families, some markers were missing for certain taxa. For *Orixa japonica* Thunb., *Ptelea baldwinii* Torr. & A.Gray (both Rutaceae), *Cedrela odorata* L., *Khaya grandifoliola* C.DC. (both Meliaceae) and *Kirkia acuminata* Oliv. (Kirkiaceae) we had fresh leaf material at our disposal and sequenced the missing markers. The sequences have been deposited in the EMBL Bank (http://www.ebi.ac.uk/embl/) under the accession numbers given in Appendix S1 and voucher information is specified in Appendix S3.

Laboratory work

Total DNA of *Orixa japonica*, *Ptelea baldwinii*, *Cedrela odorata*, *Khaya grandifoliola* and *Kirkia acuminata* was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions and polymerase chain reactions (PCRs) were performed under a standard protocol (see Appelhans *et al.*, 2011) using a primer annealing temperature of 53 °C, and extension at 72 °C for 1.25 min. The chloroplast markers *atp*B, *rbc*L and *trn*L–*trn*F were amplified and sequenced using universal primers (Taberlet *et al.*, 1991; Les *et al.*, 1993; Hoot *et al.*, 1995).

Sequence alignments and phylogenetic analyses

Complementary strands were assembled and edited using Sequencher 4.08 (Gene Codes Corp., Ann Arbor, MI, USA). Separate alignments for each marker were assembled manually in MacClade 4.08 (Sinauer Associates Inc., Sunderland, MA, USA). Indel coding was carried out for 15 sites in the *trn*L–*trn*F alignment, summing to 67 bp. All indels were between 4 and 6 bp long. We used simple indel coding (Simmons & Ochoterena, 2000; Simmons *et al.*, 2007) and treated indels as separate characters. All alignments are available from the corresponding author on request.

The best performing model of sequence evolution was determined separately for each marker using the Akaike information criterion (AIC) as implemented in MrModeltest 2.2 (Nylander, 2004). The AIC proposed the GTR + Γ + I for all markers, but GTR + Γ was selected in all cases, as Γ (Gamma distribution) and I (proportion of invariant sites) are strongly correlated and thus not independent (Ren *et al.*, 2005).

Phylogenetic analyses were first performed independently for each marker using MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) and the single alignments were then combined into one matrix. The Bayesian analyses of the single markers and the concatenated dataset each included two runs of four chains each, which were monitored for 5 million generations, sampling every thousandth generation. The temperature coefficient of the chain-heating scheme was set to 0.1 in order to ensure sufficient chain swapping. All runs reached stationarity within 5 million generations. Tracer 1.5 (Rambaut & Drummond, 2007) was used to check for convergence of the model likelihood and parameters between the two runs and the first 10–15% of the calculated generations were discarded as burn-in. A 50% majority rule consensus tree was calculated in MrBayes 3.1.2.

Fossil selection and age constraints

The only fossils for Spathelioideae are a seed and a leaf fossil from Brazil (Duarte & Da Conceição Mella Filha, 1980). The fossils clearly resemble the extant genus *Dictyoloma*, but the imprecise age 'Cenozoic' makes the fossils unsuited for a molecular dating analysis.

A number of well-identified and dated fossils are available for Rutaceae and other families of Sapindales, and we used four fossil calibration points in total: three within Rutaceae and one in the closely related Simaroubaceae. We decided not to use fossil calibration points from other Sapindales families because only Rutaceae and Simaroubaceae had a sufficiently high taxon sampling with almost no missing data in our alignments, allowing us to place the fossils. Moreover, a calibration point within Rutaceae and Simaroubaceae would have a much higher impact on our estimates for Spathelioideae than a more distant fossil in the phylogeny.

We used a leaf fossil of *Clausena* Burm.f., dated to 27.36 ± 0.11 Ma (Pan, 2010) for the stem lineage of *Clausena*. A leaf fossil of *Skimmia tortonica* Palamarev & Usunova, described from the Miocene (Tortonian) (Palamarev & Usunova, 1970; Salvo *et al.*, 2010), was used to calibrate the node between *Skimmia* Thunb. and *Dictamnus* L. The third Rutaceae fossil is a seed from the late Palaeocene to late Eocene that has been named *Phellodendron costatum* Chandler, and later *Euodia costata* (Chandler) Tiffney (Tiffney, 1981). The genus *Euodia* J.R.Forst. & G.Forst. has been reduced from about 200 species (Engler, 1931; Hartley, 1981) to seven species (Hartley, 1981, 2001) and many species have been moved to *Tetradium* Lour. and *Melicope* J.R.Forst. & G.Forst. The extant '*Euodia*' species that Tiffney (1981) considered closely related to the *Euodia costata* fossil have all been transferred to *Tetradium*. Apart from Hartley's (1981, 2001) revision work, there is molecular evidence that confirms Hartley's generic boundaries (Harbaugh *et al.*, 2009). *Euodia costata* may therefore not be used as calibration point for *Euodia*. There are two possibilities for the placement of the *Euodia costata* fossil. A conservative calibration scheme places the fossil at the node between the sister genera *Tetradium* and *Phellodendron* Rupr. (calibration scheme 1), considering that the fossil also has similarities with *Phellodendron*, to which it was assigned first (Tiffney, 1981). Following Hartley's (2001) reasoning, the fossil should be placed at the node that leads to *Tetradium* (calibration scheme 2). We performed dating analyses with both possibilities. A seed fossil of *Ailanthus* Desf. (Simaroubaceae) dated to about 52 Ma (Corbett & Manchester, 2004) was used as calibration point outside Rutaceae. *Ailanthus* has an excellent fossil record, beginning at 52 Ma and becoming prominent in all northern continents by the middle Eocene (Corbett & Manchester, 2004), which makes it particularly well suited for molecular dating analyses, and it is commonly used in biogeographical studies within Sapindales (Muellner *et al.*, 2006, 2007; Pfeil & Crisp, 2008; Clayton *et al.*, 2009). In concordance with Pfeil & Crisp (2008) and Clayton *et al.* (2009), the *Ailanthus* fossil was used to calibrate the *Ailanthus* stem.

Fossils of *Ptelea* L. (Call & Dilcher, 1995) are easily recognizable because of their conspicuous samaroid fruits. Nevertheless, we did not use *Ptelea* fossils because the phylogenetic position of this genus within Rutaceae is unclear (Groppo *et al.*, 2008; Kubitzki *et al.*, 2011). The oldest fossil that can be clearly assigned to Rutaceae is a seed named *Rutaspermum biornatum* Knobloch & Mai, dated to the Late Cretaceous (Knobloch & Mai, 1986; Gregor, 1989). We did not use this fossil because it is not clear whether it should be placed as a minimum age for Rutaceae s.s. (without Spathelioideae) or Rutaceae s.l. (including Spathelioideae).

Apart from the fossils, we used an age estimate for Sapindales inferred by a penalized likelihood analysis of angiosperms (Magallón & Castillo, 2009). Magallón & Castillo (2009) performed two analyses: in their 'relaxed' dating they included 125 Ma as a maximum age for the eudicot crown node, and in the 'constrained' dating they used 130 Ma as maximum constraint for the angiosperm crown node. The age constraint we specified for Sapindales covers the range of both analyses by Magallón & Castillo (2009).

Molecular dating analyses

A likelihood ratio test (LRT; Felsenstein, 1988) indicated that our combined dataset did not evolve in a strict clock-like manner ($P < 0.001$ for all markers).

The BEAST 1.6.1. package (Drummond & Rambaut, 2007) was used for the molecular dating analyses. The BEAST input files were created using BEAUTI 1.6.1, in which three partitions, one for each marker, were created; an uncorrelated relaxed clock model assuming a lognormal distribution of rates was used (Drummond *et al.*, 2006); and the GTR + Γ model of sequence evolution was selected. A randomly generated starting tree was used and the tree prior was set to the birth–death process. All fossil calibration points were assigned a lognormal prior. Rutaceae have an abundant fossil record (Gregor, 1989), and fossils of '*Euodia*', *Toddalia* Juss., *Zanthoxylum* L. and *Ailanthus* (Simaroubaceae) have a particularly robust record (Gregor, 1979, 1989; Tiffney, 1980; Corbett & Manchester, 2004). It has been argued that in taxa with a robust fossil record, the dates of the oldest fossils are close to the actual ages of the taxa (Givnish & Renner, 2004; de Queiroz, 2005). Except for *Clausena*, we defined a narrow age range for the fossils accordingly. Only a single fossil of *Clausena* is known and we therefore set a broad age range for this node, allowing a much older age (soft upper boundary of about 45 Ma). We initially planned to apply the same calibration settings in the analyses using two different calibration schemes for the '*Euodia costata*' fossil. However, BEAST crashed repeatedly when analysing calibration scheme 2 and we had to enlarge the standard deviation from 0.75 to 0.95 (Table 1). The prior for the root height (= age constraint for Sapindales; Magallón & Castillo, 2009) was defined with a normal prior. Prior settings for the calibration points are displayed in Table 1. All other priors were kept as defaults.

Two separate analyses (two analyses each for both calibration schemes for the '*Euodia costata*' fossil) with 50 million generations each were carried out. TRACER 1.5 was used to check for convergence between the runs and to determine the amount of burn-in. The two runs were combined, discarding the initial 10% as burn-in, using LOGCOMBINER 1.6.1 (http://code.google.com/p/beast-mcmc/downloads/list) and a maximum clade credibility tree using a posterior probability limit of 0.5 was calculated using TREEANNOTATOR 1.6.1 (http://code.google.com/p/beast-mcmc/downloads/list).

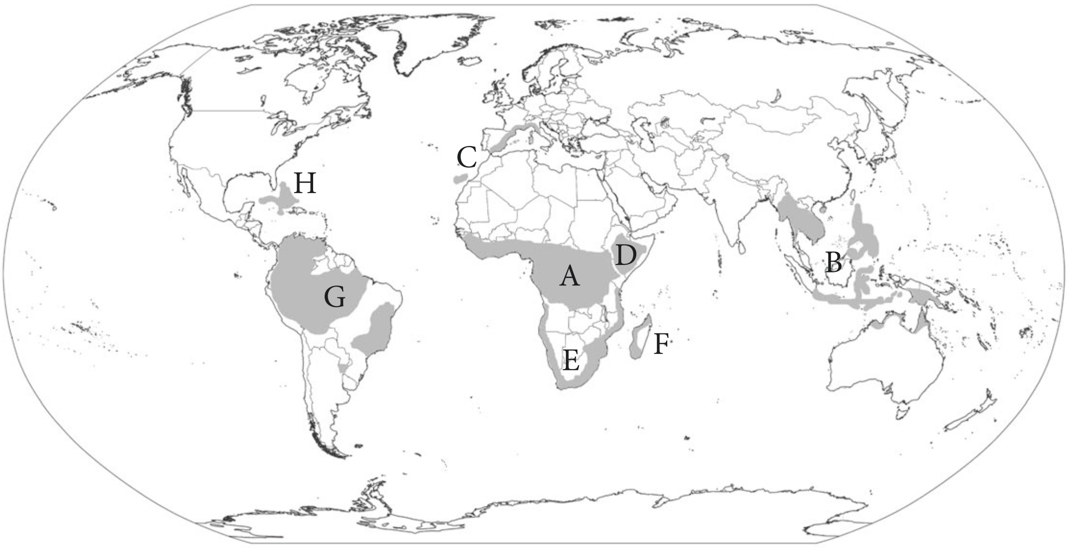
Ancestral area reconstruction

We performed AAR analyses using the dispersal–extinction–cladogenesis model as implemented in LAGRANGE (version from 17 January 2011; Ree *et al.*, 2005; Ree & Smith, 2008). Python scripts were created with the help of the LAGRANGE online configurator (http://www.reelab.net/lagrange/configurator/index). Only an ingroup (Spathelioideae) dataset was used for AAR. A Bayesian analysis was carried out in BEAST based on combined *rbc*L, *atp*B and *trn*L–*trn*F sequences using the settings described above with a maximum age constraint for Spathelioideae taken from our molecular dating analyses (calibration scheme 1). We enforced the monophyly of *Cedrelopsis* in this analysis based on the results of Razafimandimbison *et al.* (2010). *Ptaeroxylon* appeared as nested within *Cedrelopsis* in our previous analyses (Appelhans *et al.*, 2011), but without support. Thus, the results of Razafimandimbison *et al.* (2010), although inferred from a lower taxon sampling, deliver the only statistically supported information about the relationship of the two genera. The resulting maximum clade credibility tree was used as input for LAGRANGE.

Eight areas were delimited based on the distribution of the genera/species: A, central-western and central Africa; B, Southeast Asia (including parts of tropical Australia); C, western Mediterranean and Canary Islands; D, central eastern Africa; E, southern Africa; F, Madagascar; G, northern South America; H, the Caribbean (Fig. 1). Areas A and D were separated because *Bottegoa* is only present in central eastern Africa (D) while *Harrisonia abyssinica* occurs throughout both areas. Area G describes the distribution of *Dictyoloma* and *Sohnreyia*. Their distribution overlaps in the western part, while only *Dictyoloma* occurs in the disjunct areas of eastern Brazil and north-eastern Argentina. We did not separate the

Table 1 Settings of fossil calibration points and root height in the molecular dating analyses of Sapindales. *Euodia* 1 and 2 correspond to the two placements of the *Euodia costata* fossil: at the common stem of *Tetradium* and *Phellodendron* (*Euodia* 1) and at the stem of *Tetradium* (*Euodia* 2). The age ranges fit the time frame given by the authors of the fossils and the authors of earlier molecular dating analyses: Miocene (Tortonian) for *Skimmia* (Palamarev & Usunova, 1970; Salvo *et al.*, 2010), Palaeocene for '*Euodia*' (Tiffney, 1981), and early Eocene for *Ailanthus* (Clayton *et al.*, 2009; Corbett & Manchester, 2004). The age range for *Clausena* has been extended due to the poor fossil record (it has been found only once). The age for the Sapindales includes the age of 98.01–98.51 Ma estimated by Magallón & Castillo (2009).

	Clausena	*Skimmia*	'*Euodia* 1'	'*Euodia* 2'	*Ailanthus*	Sapindales
Prior distribution	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Normal
Offset	26.3	6.3	54.7	54.7	52.0	–
Mean	1.0	1.0	1.0	1.0	1.0	98.3
Standard deviation	1.2	0.55	0.75	0.95	0.55	3.0
Median age (Ma)	29.02	9.02	57.42	57.42	54.72	98.3
Lower and upper boundary (5% and 95%) (Ma)	26.68–45.87	7.4–13.02	55.49–64.09	55.49–64.09	53.1–58.72	93.37–103.2



0 - 23.0 Ma

	A	B	C	D	E	F	G	H
A		0.2	0.5	1.0	0.9	0.5	0.2	0.1
B			0.0	0.3	0.3	0.3	0.0	0.0
C				0.4	0.0	0.0	0.1	0.1
D					0.9	0.6	0.0	0.0
E						0.8	0.1	0.0
F							0.0	0.0
G								1.0
H								

23.0 - 34.0 Ma

	A	B	C	D	E	F	G	H
A		0.2	0.5	1.0	0.9	0.5	0.3	0.1
B			0.0	0.3	0.3	0.3	0.0	0.0
C				0.4	0.0	0.0	0.1	0.1
D					0.9	0.6	0.0	0.0
E						0.8	0.1	0.0
F							0.0	0.0
G								1.0
H								

34.0 - 56.0 Ma

	A	B	C	D	E	F	G	H
A		0.2	0.5	1.0	0.9	0.5	0.3	0.1
B			0.0	0.3	0.3	0.3	0.0	0.0
C				0.4	0.0	0.0	0.2	0.1
D					0.9	0.6	0.0	0.0
E						0.8	0.2	0.0
F							0.0	0.0
G								0.9
H								

56.0 - 75.0 Ma

	A	B	C	D	E	F	G	H
A		0.2	0.5	1.0	0.9	0.5	0.4	0.1
B			0.0	0.3	0.3	0.3	0.0	0.0
C				0.4	0.0	0.0	0.2	0.1
D					0.9	0.6	0.0	0.0
E						0.8	0.2	0.0
F							0.0	0.0
G								0.9
H								

Figure 1 Delimitation of the eight areas used for ancestral area reconstruction of the subfamily Spathelioideae and probability of dispersal throughout four defined time periods (0–23 Ma, Quaternary and Neogene; 23–34 Ma, Oligocene; 34–56 Ma, Eocene; 56–75 Ma, Palaeocene and Late Cretaceous). Area designations are as follows: A, central-western and central Africa; B, Southeast Asia (including the distribution of *Harrisonia brownii* in tropical Australia); C, western Mediterranean and Canary Islands; D, central eastern Africa; E, southern Africa; F, Madagascar; G, northern South America; H, Caribbean region. More detailed information on the distribution of all species is provided in Appendix S2.

disjunct areas from area G, because the same species (*D. vandellianum*) is recognized throughout the whole distribution area (Groppo, 2010) and we were not able to include a specimen from the disjunct areas in our analyses. Detailed information on the distribution of all species is given in Appendix S1.

The maximum number of areas in ancestral ranges was set to three and all geographical ranges that were considered biologically implausible (largely disjunct areas) were excluded from the analyses. Dispersal rates between all areas were configured according to the distance from each other and existing barriers (e.g. oceans), defined as 'symmetric' (meaning that no spatial direction of dispersal is favoured) and allowed to vary through time. For this, we defined four time slices with slightly changed dispersal rates (Fig. 1). The time slices correspond to the Quaternary and Neogene (0–23 Ma; time slice 1), the Oligocene (23–34 Ma; time slice 2), the Eocene (34–56 Ma; time slice 3) and the Palaeocene and Late Cretaceous (56–75 Ma; time slice 4). The dispersal rates between most areas were not varied through time. Only the rates dealing with a trans-Atlantic dispersal and dispersal from South America to the Caribbean were given different values at different time slices (Fig. 1).

Diversification analyses

We generated lineage-through-time (LTT) plots and evaluated the fit of the LTT plots to three generalized models of diversification (Paradis, 1998; see McKenna & Farrell, 2006) based on the maximum clade credibility tree from the molecular dating analyses. We chose to use only the maximum clade credibility tree from calibration scheme 1, as both trees (calibration scheme 1 and 2) were identical in topology and differed only slightly in their estimated node ages. In order to obtain a chronogram that only included the ingroup, we manually deleted all non-Spathelioideae taxa from the maximum clade credibility tree. We also deleted specimens of species of which more than one specimen was included in the previous analyses in order not to artificially enlarge the number of lineages. The resulting tree contained 25 taxa that represent 100% of the Spathelioideae genera and 83.3% of the species (see 'Taxon sampling' above).

The LTT plots represent a schematic visualization of the net diversification rate and the gradient of the curve represents the diversification rate. By means of LTT plots, changes of diversification rates can be determined throughout the evolutionary history of a taxon. The LTT plots were conducted using

LASER 2.3 (Rabosky, 2006) implemented in R. Despite the high taxon sampling at species level, we tested for an effect of incomplete taxon sampling. We used PHYLOGEN 1.1 (Rambaut, 2002) to generate 100 phylogenetic trees of a random dataset of 30 taxa of which 25 were sampled (= 83.3% taxon sampling on species level). The replicates were generated under the assumption of a constant birth–death rate. Based on these trees, a mean LTT curve plus a 95% confidence interval were generated and compared with the LTT plots of the Spathelioideae dataset.

LASER 2.3 was also used to test the fit of the Spathelioideae LTT plot to three generalized diversification models proposed by Paradis (1998; see also McKenna & Farrell, 2006), which are: a model of constant diversification rate (model A), a gradually increasing or decreasing diversification rate (model B) or model of an abrupt change of diversification rates (model C) (McKenna & Farrell, 2006). Likelihood values (calculated by the AIC) for all three models were estimated using LASER 2.3 as implemented in R and plotted through time in intervals of 2.5 Ma.

The LTT plots and the test for the fit of the diversification models were performed on the ingroup tree (= Spathelioideae) and separately on the Neotropical and Palaeotropical groups in order to check for differences between these groups.

RESULTS

Phylogeny

The topologies and support values from the Bayesian analyses carried out in MRBAYES and BEAST are highly similar (Figs 2 & 3). No supported differences are present throughout the consensus trees. Rutaceae are monophyletic with strong support and are sister to Meliaceae and Simaroubaceae: (Rutaceae, (Meliaceae, Simaroubaceae)). Also the subfamily Spathelioideae is monophyletic and highly supported. Spathelioideae are sister to the remainder of Rutaceae and contain a Neotropical subclade consisting of *Dictyoloma*, *Sohnreyia* and *Spathelia*, and a Palaeotropical subclade (extending to subtropical areas) including *Bottegoa*, *Cedrelopsis*, *Cneorum*, *Harrisonia* and *Ptaeroxylon*. The relationships within Spathelioideae are largely congruent with those described by Appelhans *et al.* (2011) and the further grouping within Rutaceae agrees with previous phylogenetic analyses (Chase *et al.*, 1999; Groppo *et al.*, 2008; Salvo *et al.*, 2008, 2010).

The other families of Sapindales appear to be monophyletic in our analyses and their relationships among each other are congruent with previous analyses (e.g. Muellner *et al.*, 2007).

Molecular dating

Only the dates inferred from the first fossil calibration scheme (placement of the '*Euodia costata*' fossil at the stem lineage of *Phellodendron* and *Tetradium*) are discussed here. The age estimates from the second calibration scheme (placement of the '*Euodia costata*' fossil at the stem lineage of *Tetradium*; white star in Fig. 3) are highly similar to those from fossil calibration scheme and therefore a comparison between the dates is only carried out for the most important nodes, namely Rutaceae and Spathelioideae.

The effective sample size (ESS) and the trace of parameters (visualized in TRACER) confirmed that the two runs had converged and that 50 million generations were sufficient. The maximum clade credibility tree from the BEAST analysis is shown in Fig. 3. The age estimates for most families within Sapindales are between 60 and 80 Ma. A mean age of 82 Ma was estimated for Rutaceae with a credible interval ranging from 74 to 87 Ma, suggesting an origin of Rutaceae in the Late Cretaceous. The age of Rutaceae s.s. (without Spathelioideae) was estimated to 70 Ma (Late Cretaceous; 62–73 Ma). Spathelioideae and the split into Neotropical and Palaeotropical subclades were dated to 74 Ma (Palaeocene or Late Cretaceous; 58–78 Ma). The split of *Harrisonia* from the other genera in the Palaeotropical subclade possibly occurred in the early Eocene or Palaeocene (mean age 58 Ma; credible interval 42–65 Ma). Our estimates suggest a split between *Cneorum* and Ptaeroxyleae (Appelhans *et al.*, 2010; *Bottegoa*, *Cedrelopsis*, *Ptaeroxylon*) in the Oligocene or late Eocene (40 Ma; 24–42 Ma). The split of the two species of *Cneorum* was dated to the Miocene (17 Ma; 9–24 Ma) and the split between *Bottegoa* and *Cedrelopsis*/*Ptaeroxylon* dated to the Miocene or late Oligocene (27 Ma; 14–29 Ma).

Within the Neotropical lineage, divergences occurred later than in the Palaeotropical subclade. *Sohnreyia* possibly split from *Dictyoloma* and *Spathelia* in the early Miocene, Oligocene or late Eocene (31 Ma; 22–41 Ma). The divergence of *Dictyoloma* and *Spathelia* may have occurred in the early Miocene or Oligocene (19 Ma; 16–33 Ma), and *Spathelia brittonii* potentially diverged from the other *Spathelia* species in the late or mid-Miocene (9 Ma; 5–14 Ma).

The age of Rutaceae s.s. (excluding Spathelioideae) as inferred from the second calibration scheme was about 2 million years (Myr) older (84 Ma; 75–88 Ma) than that from the first calibration scheme and the credible intervals were almost identical. A similar observation was made for the inferred age of Spathelioideae, for which the age estimate from the second calibration scheme was about 4 Myr older and the credible intervals were also largely congruent (78 Ma; 59–80 Ma). Only minor differences between the two calibration schemes were observed for the age estimates within Spathelioideae.

Ancestral area reconstruction

Most of the ancestral areas and splitting of areas were unambiguous and, with the exception of two nodes, the relative probability values for the areas at the nodes shown in Fig. 4 and Table 2 were at least 10% higher than the second option of area combinations suggested by LAGRANGE. Out of the 39 nodes in the tree, 35 had AARs supported by more than 50% relative probability and 29 nodes had AARs supported by at least 90% relative probability. The AARs for nodes (1) and (2) (Fig. 4, Table 2) were the least congruent, but the AARs

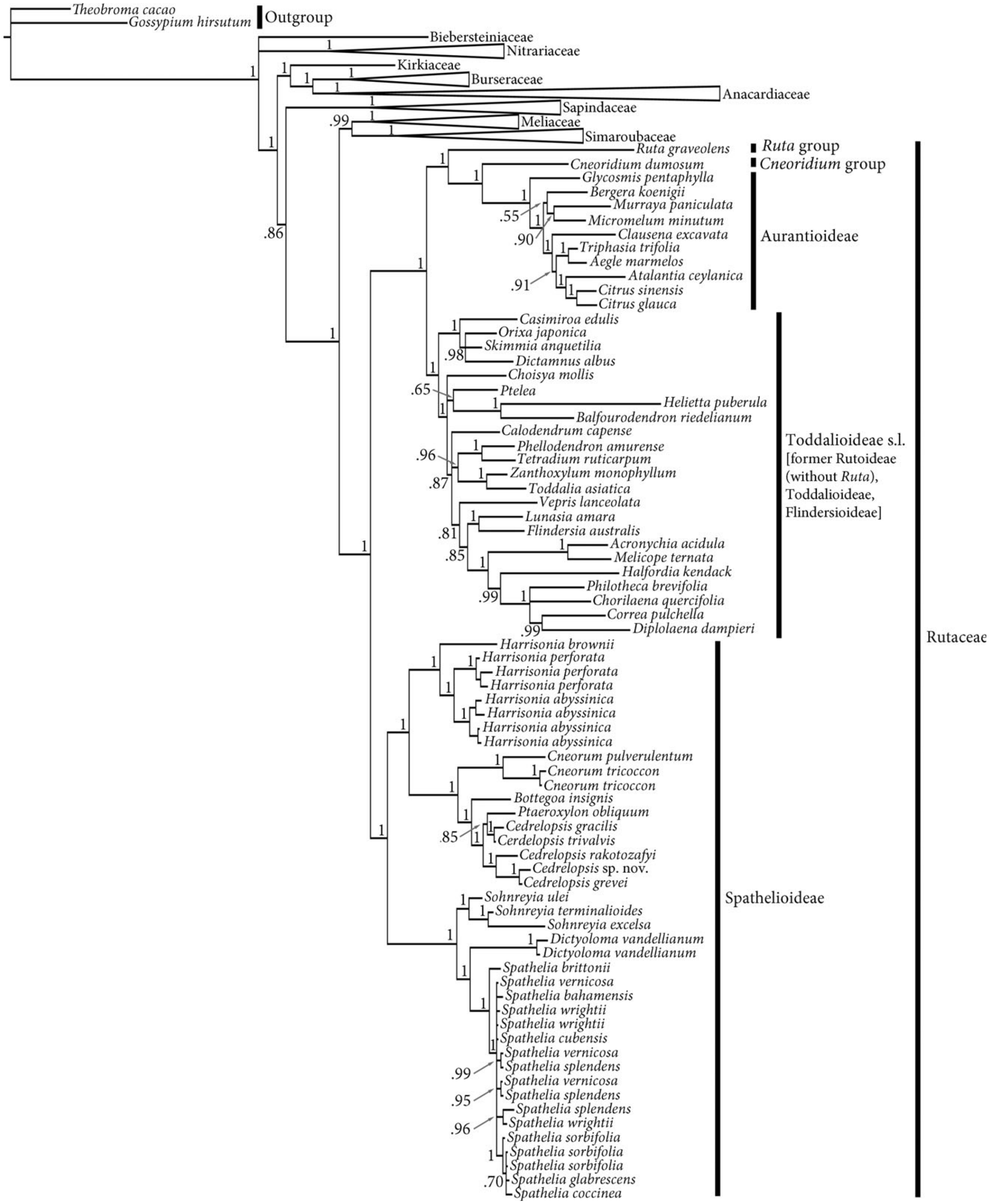


Figure 2 Phylogeny of the order Sapindales. The 50% majority rule consensus tree of the combined *atp*B, *rbc*L and *trn*L–*trn*F dataset from the Bayesian analysis is shown. Posterior probability values are indicated above the branches (next to the branches or marked with an arrow in the case of limited space). The branches for all families except Rutaceae were collapsed. Information about included taxa is given in Appendix S1.

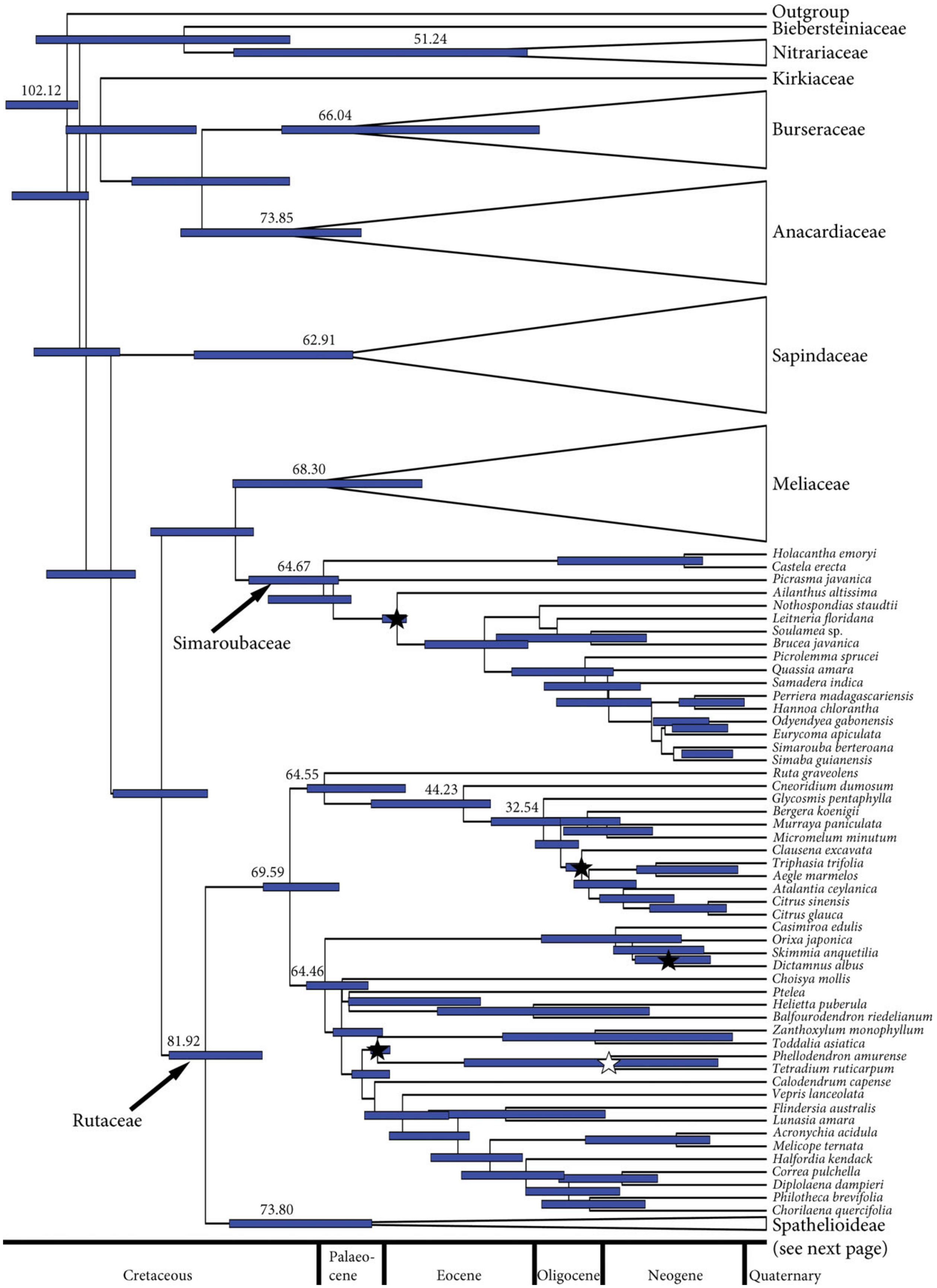

Outgroup
Biebersteiniaceae
51.24
Nitrariaceae
102.12
Kirkiaceae
66.04
Burseraceae
73.85
Anacardiaceae
62.91
Sapindaceae
68.30
Meliaceae
64.67
Simaroubaceae
Holacantha emoryi
Castela erecta
Picrasma javanica
Ailanthus altissima
Nothospondias staudtii
Leitneria floridana
Soulamea sp.
Brucea javanica
Picrolemma sprucei
Quassia amara
Samadera indica
Perriera madagascariensis
Hannoa chlorantha
Odyendyea gabonensis
Eurycoma apiculata
Simarouba berteroana
Simaba guianensis
64.55
44.23
32.54
69.59
64.46
81.92
Rutaceae
73.80
Ruta graveolens
Cneoridium dumosum
Glycosmis pentaphylla
Bergera koenigii
Murraya paniculata
Micromelum minutum
Clausena excavata
Triphasia trifolia
Aegle marmelos
Atalantia ceylanica
Citrus sinensis
Citrus glauca
Casimiroa edulis
Orixa japonica
Skimmia anquetilia
Dictamnus albus
Choisya mollis
Ptelea
Helietta puberula
Balfourodendron riedelianum
Zanthoxylum monophyllum
Toddalia asiatica
Phellodendron amurense
Tetradium ruticarpum
Calodendrum capense
Vepris lanceolata
Flindersia australis
Lunasia amara
Acronychia acidula
Melicope ternata
Halfordia kendack
Correa pulchella
Diplolaena dampieri
Philotheca brevifolia
Chorilaena quercifolia
Spathelioideae
(see next page)
Cretaceous
Palaeo-cene
Eocene
Oligocene
Neogene
Quaternary

Figure 3 Maximum clade credibility tree of the combined *atp*B, *rbc*L and *trn*L–*trn*F dataset of the order Sapindales from the BEAST analysis. The fossil calibration points are indicated with a black star. The alternative position of the '*Euodia costata*' fossil is pictured as a white star. Families in which there are no fossil calibration points are shown by 'triangles' formed by collapsed branches. Mean age estimates for Sapindales, the families (except for Biebersteiniaceae and Kirkiaceae of which only one taxon was sampled), major lineages within Rutaceae, as well as the nodes within Spathelioideae, are displayed next to the branches. The bars indicate age intervals (credible intervals). Geological periods: Cretaceous (145.5–65.5 Ma), Palaeocene (65.5–55.8 Ma), Eocene (55.8–33.9 Ma), Oligocene (33.9–23.03 Ma), Neogene (23.03–2.588 Ma), Quaternary (2.588–0 Ma).

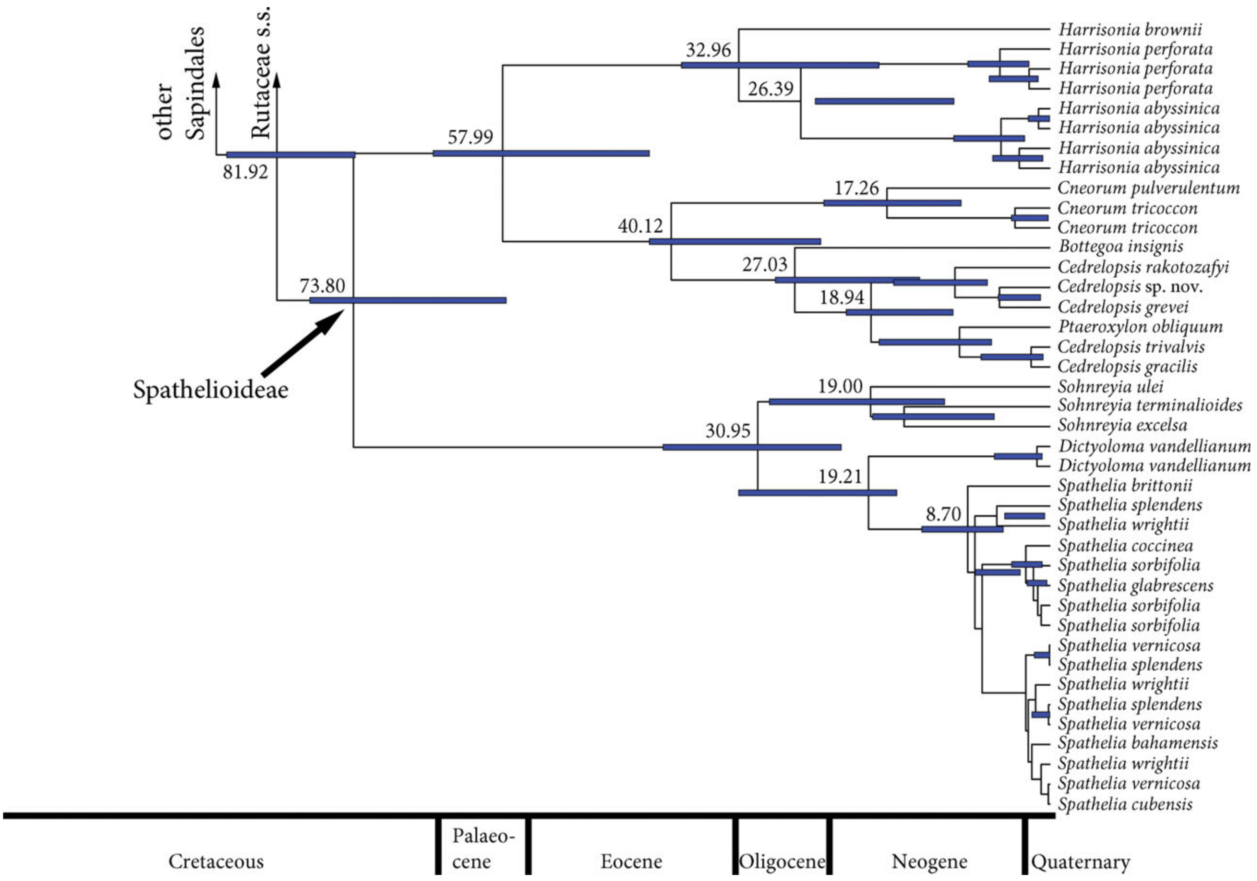


Figure 3 Continued

remained fairly similar. For node 1, the alternative to a central and southern African lineage (AE) is a central and eastern African lineage (AD). For node 2, the alternative ancestral areas are different combinations of the adjacent areas central-western and central Africa (A), central eastern Africa (D) and southern Africa (E).

The base of Fig. 4 shows the split into an African lineage and a South American lineage (A–G). The splitting into the *Harrisonia* – and the '*Cneorum* & Ptaeroxyleae' – lineages is inferred to have occurred in the African area A. The LAGRANGE results suggest a wide ancestral distribution area (ABD) for *Harrisonia* from which two Southeast Asian lineages [*H. brownii*, *H. perforata*, (B)] and one widespread African lineage [*H. abyssinica*, (AD)] diverged. The '*Cneorum* & Ptaeroxyleae' lineage split into a northern [*Cneorum*, (C)] and a more eastern and/or southern lineage [(AE or AD), Table 2], and the latter split into an eastern African [*Bottegoa*, (D)] and later a southern African [*Ptaeroxylon*, (E)] and a Madagascan [*Cedrelopsis*, (F)] lineage.

Within the South American subclade, dispersal from the South American mainland (G) to the Caribbean Islands is inferred [*Spathelia*, (H)]. A split between a western Cuban lineage (*S. brittonii*) and a combined eastern Cuban, Jamaican and Bahamian lineage (all other species) can be observed.

Diversification analyses

The curve of the empirical LTT plot is approximately parallel to the curve of the simulated LTT plots (Fig. 5a), and runs slightly below the curve of the simulated LTT plots and its 95% confidence interval throughout the whole range. The gradient of the curve of the empirical LTT plot is generally lower, between 35 and 75 Ma, than in the following time span (35 Ma until present). Between about 18 and 32 Ma, the gradient of

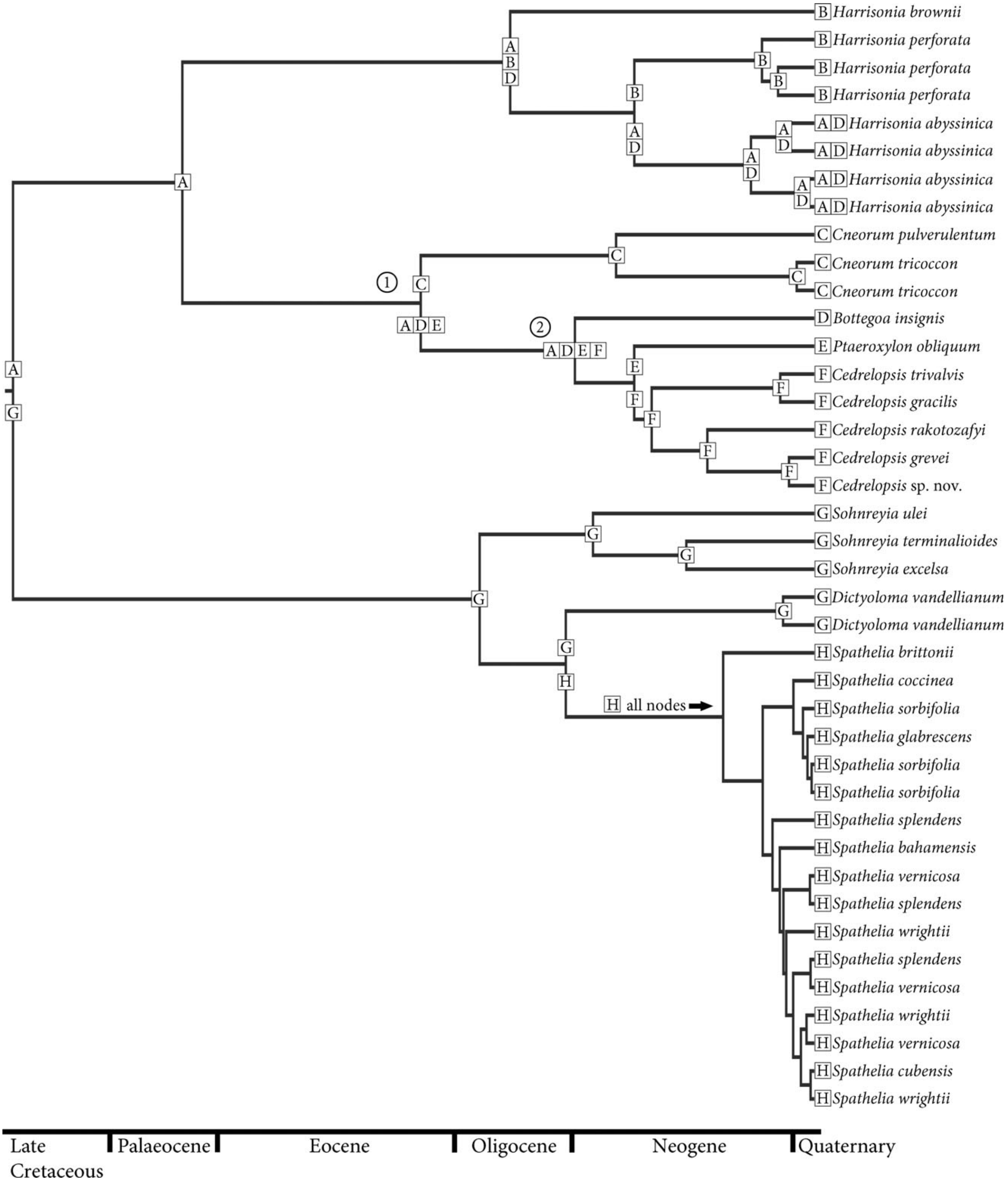


Figure 4 Ingroup (Spathelioideae) chronogram showing the results of the ancestral area reconstruction (AAR) analysis using LAGRANGE. The AARs with the highest likelihood values are shown as boxes at each node. Single boxes or combined boxes indicate ancestral ranges confined to a single or two or three areas. Boxes separated by a space indicate a split of areas. For the nodes marked with (1) and (2), the likelihood values were low and alternative AARs within a 10% range of relative probability exist (see Table 2). For these nodes, all alternative areas are displayed. Area designations are as follows: A, central-western and central Africa; B, Southeast Asia (including the distribution of *Harrisonia brownii* in tropical Australia); C, western Mediterranean and Canary Islands; D, central eastern Africa; E, southern Africa; F, Madagascar; G, northern South America; H, Caribbean region. Geological periods: Cretaceous (145.5–65.5 Ma), Palaeocene (65.5–55.8 Ma), Eocene (55.8–33.9 Ma), Oligocene (33.9–23.03 Ma), Neogene (23.03–2.588 Ma), Quaternary (2.588–0 Ma).

Table 2 Alternative ancestral areas and area splits for nodes 1 and 2 from the ancestral area reconstruction of the subfamily Spathelioideae (Fig. 4). For all other areas at the nodes of Fig. 4, the relative probability from the LAGRANGE analysis was at least 10% higher than the second suggestion for an area combination. Area designations are as follows: A, central-western and central Africa; C, western Mediterranean and Canary Islands; D, central eastern Africa; E, southern Africa; F, Madagascar.

	Split of areas	Relative probability (%)
Node 1	A,E/C	26.59
	A,D/C	19.71
Node 2	E/A	11.32
	D/D	8.61
	E/D	8.10
	E/E	7.77
	A,E/D	6.02
	F/D	5.23
	E,F/A	5.20

the empirical curve increases during three separate small time periods. A significant increase of the gradient in the empirical curve can also be observed in the last 2–3 Myr.

Figure 5(b) shows the testing of a model of constant diversification rate (model A), a model of gradually increasing or decreasing diversification rate (model B) and a model of an abrupt change of diversification rates (model C) (McKenna & Farrell, 2006). Apart from the very beginning, a model with a constant rate of diversification best fits the empirical data. Models B and C have very similar likelihood values through time and the curve for model C winds around the line for model B until about 10 Ma. It is interesting that within the last 5 Myr, the curve for model C approaches the values for model A. Also, the curve for model C forms a valley at about 20 Ma, which stands for an increase of the likelihood for an abrupt change in diversification rates.

The separate analyses for the Neotropical and Palaeotropical clades (results not shown) delivered very similar results to the combined analyses. Unlike the curve for the combined and the Neotropical analyses, the curve of the Palaeotropical clade does not show a further increase in diversification rate in the Pliocene–Pleistocene period. The model testing analyses for the Neotropical and the Palaeotropical clades (not shown) also show similar results to the combined analysis. In both cases, model A is suggested to be most likely throughout time and the likelihood for model C increases at about 20 Ma. In accordance with the empirical LTT plots, an increase of likelihood for model C in the Pliocene–Pleistocene is not observed in the Palaeotropical clade.

DISCUSSION

Spathelioideae: age and biogeographical patterns

Our analyses (Fig. 3) reveal an age of 74 Ma (58–78 Ma) and therefore point to a Late Cretaceous origin of Spathelioideae. The age of Spathelioideae also marks the divergence of the Neotropical and Palaeotropical lineages and the AAR reveals a split into a South American and an African lineage [(A–G), Fig. 4]. The origin of the Spathelioideae stem lineage remains unclear because we did not include outgroups in the AAR analyses. The outgroup in our analysis would have been Rutaceae s.s., but it was not feasible to include Rutaceae s.s. because the ancestral distribution of the family is not known

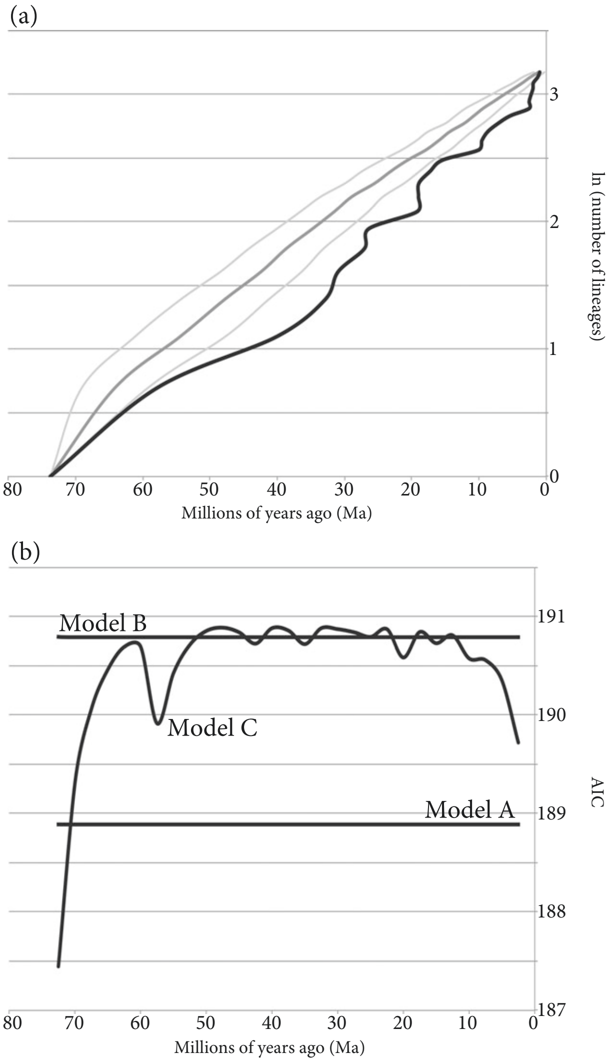


Figure 5 Lineage-through-time (LTT) plot and test for the fit of three diversification models of the ingroup (Spathelioideae). (a) The LTT plot. The black line shows the empirical LTT plot for Spathelioideae, the dark grey line equates to the simulated LLT plot and the light grey lines delimit the 95% confidence interval of the simulated LTT curve. Values on the *x*-axis are given in million years ago (Ma), those on the *y*-axis are logarithmic values of the numbers of lineages. (b) The fit of three diversification models to the Spathelioideae dataset. Model A describes a constant diversification rate; in model B, the diversification rate is gradually increasing or decreasing; model C is a model of abrupt changes of diversification rates. The values on the *x*-axis are in Ma and the *y*-axis shows likelihood values calculated by Akaike's information criterion (AIC).

(Kubitzki *et al.*, 2011). This is primarily due to the lack of resolution and support in the Toddalioideae s.l. clade in which most American taxa are nested. American Rutaceae outside Toddalioideae s.l. and Spathelioideae only include *Cneoridium* Hook.f. and *Thamnosma* Torr. & Frém. Both genera belong to otherwise Palaeotropical clades (*Ruta* clade, *Cneoridium*/*Haplophyllum* clade; Groppo *et al.*, 2008; Salvo *et al.*, 2010) and are probably cases of long-distance dispersal. *Cneoridium* might have diverged from its sister group (*Haplophyllum* or *Haplophyllum* + Aurantioideae) in the Eocene (Salvo *et al.*, 2010, 2011) and *Thamnosma* possibly originated in the Miocene (Thiv *et al.*, 2011). A Palaeotropical origin of all main clades except Toddalioideae s.l. and Spathelioideae is evident, so there is some evidence that Rutaceae have a Palaeotropical origin, possibly a North Tethyan origin (Kubitzki *et al.*, 2011). Assuming Rutaceae to be Palaeotropical, an origin of the Spathelioideae stem lineage in Africa and a dispersal event to South America is more probable than an origin in South America with subsequent dispersal to Africa. The last connections between South America and Africa existed between 119 and 105 Ma (McLoughlin, 2001), which is at least 25 Myr older than the upper limit of our credible interval (Fig. 3). Based on this and the assumption that molecular dating on taxa with a robust fossil record results in age estimates that might be close to actual ages, a transoceanic dispersal event at a time when South America and Africa were still quite close to each other may have brought Spathelioideae to South America.

Alternatively, assuming that age estimates based on fossil evidence provide only minimal ages, we cannot exclude the possibility that the actual ages are significantly older than our estimates. Given this assumption, a Gondwanan origin of Spathelioideae is possible and the split of the major lineages would be explained by vicariance (the break-up of South America and Africa).

The ancestor of *Harrisonia* may have diverged from the other Palaeotropical members either in the Eocene or Palaeocene (Fig. 3) and has an African origin. By the time of the first diversification of *Harrisonia* in the Miocene, Oligocene or late Eocene, the genus may already have had a broad distribution in Africa and Asia [Fig. 4, (ABD)]. Assuming a broad distribution of the ancestor, we suppose that *Harrisonia* dispersed to Asia only once. *Harrisonia brownii* is distributed in the eastern part of Malesia and northern Australia and it is possible that its ancestor separated from the remainder of *Harrisonia* by dispersal in an eastward direction. The presumed widespread ancestor of *H. abyssinica* and *H. perforata* might have gone extinct in Arabia, western Asia and India, so that the African and Southeast Asian populations were isolated. An alternative to this scenario is that the ancestor of *Harrisonia* dispersed to Asia and the ancestor of *H. abyssinica* dispersed back to Africa after the separation from *H. perforata*. Nothing is known about the dispersal vector(s) of *Harrisonia*, but judging from its wide distribution and its fleshy fruits, it seems to be a good disperser and both described scenarios of ancestral dispersal seem possible.

Cneorum has usually been regarded as a very old genus because of its isolated position as inferred from morphology and due to the description of a *Cneorum* species from Cuba (Borhidi, 1991; Riera *et al.*, 2002). However, it has been shown that the Cuban *C. trimerum* is conspecific with *C. tricoccon* L. (Mediterranean) and is apparently a recent introduction by humans and became extinct again in Cuba soon after its introduction (Lobreau-Callen & Jérémie, 1986; Oviedo *et al.*, 2009; Appelhans *et al.*, 2010). Based on this, we did not include a Caribbean distribution of *Cneorum* in our analyses. However, there has been no dated phylogeny of *Cneorum* prior to this study and we suggest that, based on molecular dating, *Cneorum* is not a relict genus. Our results (Fig. 3) show that the common ancestors of *Cneorum* and Ptaeroxyleae diverged in the Oligocene or late Eocene and that the ancestors of the two extant species of the genus split in the Miocene. This split is consistent with the age of the Canary Islands, which is inferred to be around 20 Ma (Hoernle & Carracedo, 2009). *Cneorum pulverulentum* is endemic to the Canary Islands and our age estimates deliver evidence that the species is a neoendemic to the Canaries. If the split was older, a distribution of *C. pulverulentum* in north-western Africa or the Palaeo Canaries (Fernández-Palacios *et al.*, 2011) prior to the emergence of the Canary Islands might be conceivable. Based on our results and the current distribution of *Cneorum*, we are unable to draw conclusions about a former occurrence of the genus in north-western Africa, although the dispersal of *Cneorum* seeds (both species) by lizards (Traveset, 1995a,b; Riera *et al.*, 2002) would make a direct dispersal from the Mediterranean to the Canary Islands unlikely. It is noteworthy that the lizard genus *Gallotia* Arribas, which is endemic to the Canary Islands and disperses *C. pulverulentum* seeds, originated in the same period (Miocene) as the *C. pulverulentum* lineage (Cox *et al.*, 2010).

Ptaeroxyleae unambiguously have an African origin; however, the ancestral areas for the Ptaeroxyleae lineage are not well resolved (nodes 1 and 2; Fig. 4, Table 2). The presence of area A at theses nodes appears counterintuitive, because no taxon of Ptaeroxyleae occurs in this area. The explanation for this is that we defined high dispersal rates between the adjacent areas A, D and E and that a taxon could move from area D to area A simply by range expansion. Given the low likelihood values for a particular area or combination of areas, nodes 1 and 2 (Fig. 4) should simply be treated as 'tropical Africa' or 'sub-Saharan Africa'. The relationship between *Ptaeroxylon* and *Cedrelopsis* is not clear from our analyses (*Ptaeroxylon* nested within *Cedrelopsis* but without support; Figs 2 & 3). However, Razafimandimbison *et al.* (2010) resolved a sister group relationship of both genera. Our results (Fig. 4) show that the stem lineage of *Cedrelopsis* split from a southern African ancestor. Madagascar has had no direct connection to Africa for the past 160 Myr (Goodman, 2009) while our results suggest a split in the Miocene, so a long-distance dispersal event to Madagascar is the most likely scenario explaining the present distribution. Dispersal might have occurred by air due to the winged seeds of *Ptaeroxylon* and *Cedrelopsis*. Alternatively, assuming that *Ptaeroxylon* is indeed nested within

Cedrelopsis, dispersal from central eastern Africa (area D) might have brought the ancestor of *Ptaeroxylon* and *Cedrelopsis* to Madagascar and the ancestor of *Ptaeroxylon* dispersed back to Africa subsequently.

The ancestral area for the Neotropical clade of Spathelioideae is inferred to be northern South America (area G). Conclusions about biogeographical patterns within this area are not feasible, especially because of incomplete knowledge of *Sohnreyia*. Two of the four species are known from their type locality only, so the actual area of distribution of the genus cannot be determined reliably. *Dictyoloma* has a large distribution, which may be explained by its light and winged seeds (da Silva & Paoli, 2006).

From area G, one dispersal event to the Caribbean Islands is inferred, potentially in the Miocene or Oligocene. The fruits of *Spathelia* are winged (Appelhans *et al.*, 2011), so we would assume that the ancestor of *Spathelia* colonized the Caribbean Islands by anemochorous dispersal. However, the fruits are relatively heavy and the wings are rather narrow, so they might not be suited for transportation over such a long distance. Nevertheless, Caribbean hurricanes have been hypothesized as vectors for *Spathelia* (Parra-O, 2005). Another possibility for dispersal to the Caribbean would be island hopping or dispersal via a land bridge. A connection between Cuba and South America via a series of islands functioning as stepping-stones was available from the early Miocene onwards (Heinicke *et al.*, 2007) and is congruent with the splitting of the *Dictyoloma*/*Spathelia* lineage (19 Ma; credible interval 16–33 Ma). Remnants of these connections are the Lesser Antilles, Puerto Rico and Hispaniola, and dispersal via this route would explain the absence of *Spathelia* from Central America and Mexico. Iturralde-Vincent & MacPhee (1999) proposed a land bridge, called the GAARlandia land span, similar to the scenario of Heinicke *et al.* (2007). They postulated a continuous land connection between South America and Cuba between 33 and 35 Ma. The importance of this land bridge has recently been questioned (Ali, 2011) and it seems to be too old for the *Spathelia* lineage considering our age estimates. However, the GAARlandia period is still within the upper boundary of our credible interval for the split of *Dictyoloma* and *Spathelia* (Fig. 3).

The resolution and support for the relationships among *Spathelia* species is low (Fig. 2), but *S. brittonii*, the only species from western Cuba, is clearly sister to the other species, which are distributed in eastern Cuba, Jamaica and the Bahamas. Cuba is characterized by mountainous areas surrounded by lowlands (Woods & Sergile, 2009), and *S. brittonii* may have been isolated during fragmentation of the island through rising sea levels (Woods & Sergile, 2009). Lowering sea levels in the late Pliocene and Pleistocene (McNeill *et al.*, 2008; Woods & Sergile, 2009) might have facilitated dispersal to Jamaica. As the Jamaican group (2.5 Ma, 1–4 Ma; Fig. 3) is monophyletic (Fig. 2), we conclude that a single dispersal event took place. A fairly recent colonization event brought *Spathelia* to the Bahamas (*Spathelia bahamensis*). The subaerial exposure of the Bahamas started at 2.5–3 Ma (McNeill *et al.*, 2008), and since then many elements of the flora and fauna have dispersed from Cuba and Hispaniola to the Bahamas (Woods & Sergile, 2009).

Diversification rates through time

Although the model-fit test globally suggests a constant diversification rate through the evolution of Spathelioideae (Fig. 5b), the empirical LTT shows that the diversification rates from 35 to 75 Ma were slightly lower than in the following time periods (Fig. 5a). Between 18 and 32 Ma there are three time periods with increased diversification rates. The last and most abrupt of these periods (± 18–20 Ma) corresponds to the beginning of the Miocene, so climate change may have triggered speciation. The distinct increase of the diversification rate in the last 2–3 Ma observed in the empirical LTT plot mainly relates to speciation within the genus *Spathelia* (Fig. 3). In the curves of the model-fit test, a tendency towards a change in diversification rates is already apparent from about 8 Ma onwards. In addition to the speciation in *Spathelia*, the biggest percentage of missing lineages through missing taxa (mainly *Cedrelopsis*) is expected to fall into this period as well, given the estimated age of the genus (Fig. 3). Thus, the likelihood for model C would probably be even higher in a complete dataset (100% of the species). The increase in diversification rate might correspond to Pliocene and Pleistocene sea-level changes (McNeill *et al.*, 2008; Woods & Sergile, 2009).

CONCLUSIONS

Our results suggest an origin of Rutaceae and its subfamily Spathelioideae in the Late Cretaceous. This view is consistent with previous molecular dating studies of sapindalean taxa (Weeks *et al.*, 2005; Muellner *et al.*, 2006, 2007; Clayton *et al.*, 2009; Nie *et al.*, 2009). The ancestral distribution of Rutaceae is not evident because the backbone phylogeny of Toddalioideae s.l. is not resolved and supported. However, most early branching clades are inferred to be of Palaeotropical origin, suggesting a Palaeotropical origin of the whole family. Based on that, Africa (rather than South America) would be a likely ancestral area for Spathelioideae, with dispersal events to the Mediterranean, the Canary Islands, Madagascar and Southeast Asia–Australia. The occurrence of Spathelioideae in the Neotropics may be explained by a transoceanic dispersal event at a time when South America and Africa were still quite close to each other (assuming that our age estimates are close to the actual ages according to the robust fossil record of Rutaceae), or by vicariance (the break-up of South America and Africa, assuming that our age estimates provide minimal ages).

Diversification analyses show that the lineages within Spathelioideae probably evolved with a constant diversification rate throughout their evolution. Only during the early Miocene and the sea level changes in the Pliocene and Pleistocene are abrupt changes in diversification rates more probable. These changes mainly relate to speciation events within *Spathelia*, the largest genus of the subfamily.

ACKNOWLEDGEMENTS

We thank the National Botanical Garden of Belgium in Meise for plant material. Randall Bayer (University of Memphis) and Cynthia Morton (Carnegie Museum of Natural History, Pittsburgh) are thanked for sending us *Toddalia* sequences. We thank Pauline Ladiges and two referees for very helpful comments and suggestions on the manuscript. We thank Liberty Blank and Atiyo Ghosh (both Leiden University) for proofreading and editing the English. The first author would like to thank Niklas Walberg (University of Turku) and Johan Nylander (Stockholm University) for very inspiring lectures on Bayesian phylogenetics and molecular dating during a course in Copenhagen, as well as the Leiden University Funds (LUF; 9103/27-1-09/N) and the Alberta Mennega Stichting for supporting fieldwork.

REFERENCES

Ali, J.R. (2011) Colonizing the Caribbean: is the GAARlandia land-bridge hypothesis gaining a foothold? *Journal of Biogeography*, **39,** 431–433.

Appelhans, M.S., Smets, E., Baas, P. & Keßler, P.J.A. (2010) *Cneorum* (Rutaceae) in Cuba? The solution to a 150 year old mystery. *Taxon*, **59,** 1126–1134.

Appelhans, M.S., Smets, E., Razafimandimbison, S.G., Haevermans, T., van Marle, E.J., Couloux, A., Rabarison, H., Randrianarivelojosia, M. & Keßler, P.J.A. (2011) Phylogeny, evolutionary trends, and classification of the *Spathelia–Ptaeroxylon* clade: morphological and molecular insights. *Annals of Botany*, **107,** 1259–1277.

Beurton, C. (2008) Rutaceae. *Flora de la República de Cuba*, Vol. 14 (ed. by W. Greuter and R. Rankin Rodriguez), pp. 1–134. A.R. Gantner Verlag KG, Ruggell, Liechtenstein.

Borhidi, A. (1991) *Phytogeography and vegetation ecology of Cuba*. Akadémiai Kiadó, Budapest.

Call, V.B. & Dilcher, D.L. (1995) Fossil *Ptelea* samaras (Rutaceae) in North America. *American Journal of Botany*, **82,** 1069–1073.

Chase, M.W., Morton, C.M. & Kallunki, J.A. (1999) Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from RBC and ATP sequence variation. *American Journal of Botany*, **86,** 1191–1199.

Clayton, J.W., Soltis, P.S. & Soltis, D.E. (2009) Recent long-distance dispersal overshadows ancient biogeographical patterns in a pantropical angiosperm family (Simaroubaceae, Sapindales). *Systematic Biology*, **58,** 395–410.

Corbett, S.L. & Manchester, S.R. (2004) Phytogeography and fossil history of *Ailanthus* (Simaroubaceae). *International Journal of Plant Sciences*, **165,** 671–690.

Cox, S.C., Carranza, S. & Brown, R.P. (2010) Divergence times and colonization of the Canary Islands by *Gallotia* lizards. *Molecular Phylogenetics and Evolution*, **56,** 747–757.

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., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4,** 699–710.

Duarte, L. & Da Conceição Mella Filha, M. (1980) Flórula Cenozóica de Gandarela, MG. I. *Anais da Academia Brasileira de Ciencias*, **52,** 77–91.

Engler, A. (1931) Rutaceae. *Die natürlichen Pflanzenfamilien*, Band 19a (ed. by A. Engler and H. Harms), pp. 187–359. Verlag von Wilhelm Engelmann, Leipzig.

Felsenstein, J. (1988) Phylogenies inferred from molecular sequences: inference and reliability. *Annual Review of Genetics*, **22,** 521–565.

Fernández-Palacios, J.M., de Nascimento, L., Otto, R., Delgado, J.D., García-del-Rey, E., Arévalo, J.R. & Whittaker, R.J. (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38,** 226–246.

Givnish, T.J. & Renner, S.S. (2004) Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *International Journal of Plant Sciences*, **165**(4 Suppl.)**,** S1–S6.

Goodman, S.M. (2009) Madagascar. *Encyclopedia of islands* (ed. by R.D. Gillespie and D.A. Clague), pp. 577–582. University of California Press, Berkeley, CA.

Gregor, H.J. (1979) Systematics, biostratigraphy and paleoecology of the genus *Toddalia* Jussieu (Rutaceae) in the European Tertiary. *Review of Paleobotany and Palynology*, **28,** 311–363.

Gregor, H.J. (1989) Aspects of the fossil record and phylogeny of the family Rutaceae (Zanthoxyleae, Toddalioideae). *Plant Systematics and Evolution*, **162,** 251–265.

Groppo, M. (2010) New synonyms in *Hortia* and *Dictyoloma* (Rutaceae), with validation of the name *Hortia badinii*. *Novon*, **20,** 163–165.

Groppo, M., Pirani, J.R., Salatino, M.L.F., Blanco, S.R. & Kallunki, J.A. (2008) Phylogeny of Rutaceae based on two noncoding regions from cpDNA. *American Journal of Botany*, **95,** 985–1005.

Harbaugh, D.T., Wagner, W.L., Allan, G.J. & Zimmer, E.A. (2009) The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *Journal of Biogeography*, **36,** 230–241.

Hartley, T.G. (1981) A revision of the genus *Tetradium* (Rutaceae). *Garden's Bulletin, Singapore*, **34,** 91–131.

Hartley, T.G. (2001) On the taxonomy and biogeography of *Euodia* and *Melicope* (Rutaceae). *Allertonia*, **8,** 1–341.

Heinicke, M.P., Duellman, W.E. & Hedges, S.B. (2007) Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal. *Proceedings of the National Academy of Sciences USA*, **104,** 10092–10097.

Hoernle, K. & Carracedo, J.C. (2009) Canary Islands, geology. *Encyclopedia of islands* (ed. by R.D. Gillespie and D.A. Clague), pp. 133–143. University of California Press, Berkeley, CA.

Hoot, S.B., Culham, A. & Crane, P.R. (1995) The utility of *atp*B gene sequences in resolving phylogenetic relationships:

comparisons with *rbc*L and 18S ribosomal DNA sequences in the Lardizabalaceae. *Annals of the Missouri Botanical Garden*, **82,** 194–207.

Iturralde-Vincent, M.A. & MacPhee, R.D.E. (1999) Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History*, **238,** 1–95.

Knobloch, E. & Mai, D.H. (1986) Monographie der Früchte und Samen in der Kreide von Mitteleuropa. *Rozpravy Ústředního ústavu Geologického*, **47,** 1–219.

Kubitzki, K., Kallunki, J.A., Duretto, M. & Wilson, P.G. (2011) Rutaceae. *The families and genera of vascular plants*, Vol. 10 (ed. by K. Kubitzki), pp. 276–356. Springer Verlag, Berlin.

Les, D.H., Garvin, D.K. & Wimpee, C.F. (1993) Phylogenetic studies in the monocot subclass Alismatidae: evidence for a reappraisal of the aquatic order Najadales. *Molecular Phylogenetics and Evolution*, **2,** 304–314.

Lobreau-Callen, D. & Jérémie, J. (1986) L'espéce *Cneorum tricoccon* (Cneoraceae, Rutales) représentée à Cuba. *Grana*, **25,** 155–158.

Magallón, S. & Castillo, A. (2009) Angiosperm diversification through time. *American Journal of Botany*, **96,** 349–365.

McKenna, D.D. & Farrell, B.D. (2006) Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. *Proceedings of the National Academy of Sciences USA*, **103,** 10947–10951.

McLoughlin, S. (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany*, **49,** 271–300.

McNeill, D.F., Klaus, J.S., Evans, C.C. & Budd, A.F. (2008) An overview of the regional geology and stratigraphy of the Neogene deposits of the Cibao Valley, Dominican Republic. *Evolutionary stasis and change in the Dominican Republic Neogene* (ed. by R.H. Nehm and A.F. Budd), pp. 21–46. Springer Science + Business Media B.V., New York.

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 *rbc*L DNA sequences, extant, and fossil distribution of diversity. *Molecular Phylogenetics and Evolution*, **40,** 236–250.

Muellner, A.N., Vassiliades, D.D. & Renner, S. (2007) Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. *Plant Systematics and Evolution*, **266,** 233–252.

Nie, Z.L., Sun, H., Meng, Y. & Wen, J. (2009) Phylogenetic analysis of *Toxicodendron* (Anacardiaceae) and its biogeographic implications on the evolution of north temperate and tropical intercontinental disjunctions. *Journal of Systematics and Evolution*, **47,** 416–430.

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

Oviedo, R., Traveset, A., Valido, A. & Brull, G. (2009) Sobre la presencia de *Cneorum* (Cneoraceae) en Cuba: ejemplo de disyunción biogeographical Mediterráneo-Caribe? *Anales del Jardín Botánico de Madrid*, **66,** 25–33.

Palamarev, D. & Usunova, K. (1970) Morphologisch-anatomischer Nachweis der Gattung *Skimmia* in der Tertiärflora Bulgariens. *Dokladi na Bulgarskata Akademiya Naukite*, **23,** 835–838.

Pan, A.D. (2010) Rutaceae leaf fossils from the Late Oligocene (27.23 Ma) Guang River flora of northwestern Ethiopia. *Review of Palaeobotany and Palynology*, **159,** 188–194.

Paradis, E. (1998) Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. *Molecular Biology and Evolution*, **15,** 476–479.

Parra-O, C. (2005) Primer registro de *Spathelia* L. (Rutaceae) y una nueva especie del género para Colombia. *Caldasia*, **27,** 17–23.

Pfeil, B.E. & Crisp, M.D. (2008) The age and biogeography of *Citrus* and the orange subfamily (Rutaceae: Aurantioideae) in Australasia and New Caledonia. *American Journal of Botany*, **95,** 1621–1631.

de Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution*, **20,** 68–73.

Rabosky, D.L. (2006) LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evolutionary Bioinformatics Online*, **2,** 257–260.

Rambaut, A. (2002) *PHYLOGEN: phylogenetic tree simulation package*. University of Oxford, Oxford.

Rambaut, A. & Drummond, A.J. (2007) *Tracer v1.4*. Available at: http://beast.bio.ed.ac.uk/Tracer.

Razafimandimbison, S.G., Appelhans, M.S., Rabarison, H., Haevermans, T., Rakotondrafara, A., Rakotonandrasana, S.R., Ratsimbason, M., Labat, J.N., Keßler, P.J.A., Smets, E., Cruaud, C., Couloux, A. & Randrianarivelojosia, M. (2010) Implications of a molecular phylogenetic study of the Malagasy genus *Cedrelopsis* and its relatives (Ptaeroxylaceae). *Molecular Phylogenetics and Evolution*, **57,** 258–265.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57,** 4–14.

Ree, R.H., Moore, B.R., Webb, C.O. & Donoghue, M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59,** 2299–2311.

Ren, F., Tanaka, H. & Yang, Z. (2005) An empirical examination of the utility of codon-substitution models in phylogeny reconstruction. *Systematic Biology*, **54,** 808–818.

Riera, N., Traveset, A. & Garcia, O. (2002) Breakage of mutualisms by exotic species: the case of *Cneorum tricoccon* L. in the Balearic Islands (western Mediterranean Sea). *Journal of Biogeography*, **29,** 713–719.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19,** 1572–1574.

Salvo, G., Bacchetta, G., Ghahremaninejad, F. & Conti, E. (2008) Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data. *Molecular Phylogenetics and Evolution*, **49,** 736–748.

Salvo, G., Ho, S.Y.W., Rosenbaum, G., Ree, R. & Conti, E. (2010) Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the *Citrus* family (*Ruta* L., Rutaceae). *Systematic Biology*, **59,** 705–722.

Salvo, G., Manafzadeh, S., Ghahremaninejad, F., Tojibaev, K., Zeltner, L. & Conti, E. (2011) Phylogeny, morphology, and biogeography of *Haplophyllum* (Rutaceae), a species-rich genus of the Irano-Turanian floristic region. *Taxon*, **60,** 513–527.

da Silva, L.L. & Paoli, A.A.S. (2006) Morfologiae e anatomia da semente de *Dictyoloma vandellianum* Adr. Juss. (Rutaceae). *Revista Brasileira de Sementes*, **28,** 116–120.

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

Simmons, M.P., Müller, K. & Norton, A.P. (2007) The relative performance of indel-coding methods in simulations. *Molecular Phylogenetics and Evolution*, **44,** 724–740.

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

Thiv, M., van der Niet, T., Rutschmann, F., Thulin, M., Brune, T. & Linder, H.P. (2011) Old–New World and trans-African disjunctions of *Thamnosma* (Rutaceae): intercontinental long-distance dispersal and local differentiation in the succulent biome. *American Journal of Botany*, **98,** 76–87.

Tiffney, B.H. (1980) Fruits and seeds of the Brandon Lignite, V. Rutaceae. *Journal of the Arnold Arboretum*, **61,** 1–40.

Tiffney, B.H. (1981) *Euodia costata*, new combination (Rutaceae) from the Eocene of southern England. *Paläontologische Zeitschrift*, **55,** 185–190.

Traveset, A. (1995a) Reproductive ecology of *Cneorum tricoccon* L. (Cneoraceae) in the Balearic Islands. *Botanical Journal of the Linnaean Society*, **117,** 221–232.

Traveset, A. (1995b) Seed dispersal of *Cneorum tricoccon* L. (Cneoraceae) by lizards and mammals in the Balearic Islands. *Acta Oecologica–Oecologia Plantarum*, **16,** 171–178.

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R. & Soltis, D.E. (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences USA*, **106,** 3853–3858.

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.

Woods, C.A. & Sergile, F.E. (2009) Antilles, biology. *Encyclopedia of islands* (ed. by R.D. Gillespie and D.A. Clague), pp. 20–29. University of California Press, Berkeley, CA.

SUPPORTING INFORMATION

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

Appendix S1 Taxa studied in the molecular phylogenetic analyses.

Appendix S2 Detailed information about the distribution of Spathelioideae taxa and the inclusion of taxa in our analyses.

Appendix S3 Voucher information for the additionally sequenced species.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Marc Appelhans completed his doctorate from the Botanic Garden (Hortus Botanicus) and the Netherlands Centre for Biodiversity Naturalis in Leiden in November 2011. He started a post-doctoral fellowship at the Smithsonian Institution in Washington DC, USA, in February 2012. His main research interests are the evolution, biogeography and systematics of Rutaceae and related families. Parts of this publication were presented at the Biosystematics conference in Berlin, Germany, in February 2011 and received the first prize for the best student presentation.

Author contributions: M.S.A., P.J.A.K., E.S., S.G.R. and S.B.J. conceived the ideas; M.S.A. and S.B.J. collected and analysed the data; M.S.A. led the writing and all authors made significant comments on and improvements to the manuscript.

Editor: Pauline Ladiges