Phylogeny and biogeography of Spathelioideae (Rutaceae)
Appelhans, M.S.

Citation
Appelhans, M. S. (2011, November 15). *Phylogeny and biogeography of Spathelioideae (Rutaceae)*. Retrieved from https://hdl.handle.net/1887/18076

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

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

Abstract

The main objective of this study is to present the first molecular dating and biogeographic analyses of the subfamily Spathelioideae (Rutaceae), which allow us to unravel the temporal and spatial origins of this group, ascertaining possible vicariant patterns and dispersal routes and determining diversification rates through time.

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 utilising the dispersal-extinction-cladogenesis model and diversification rate analyses were conducted.

Rutaceae and Spathelioideae are probably of Late Cretaceous origin, whereafter Spathelioideae split into a Neotropical and a Paleotropical lineage. The Paleotropical taxa have their origin in Africa with dispersal events to the Mediterranean, the Canary Islands, Madagascar and South-East Asia. The lineages within Spathelioideae evolved at a relatively constant diversification rate. However, abrupt changes in diversification rates became more evident at the beginning of the Miocene and during the Pliocene/Pleistocene.

Western and central tropical Africa are likely ancestral areas for Spathelioideae. The existence of a Neotropical lineage might 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 landmasses caused by sea level changes during the Pliocene and Pleistocene might have been triggers for the current species composition of the Caribbean genus *Spathelia*.

Keywords: Ancestral area reconstruction; Diversification rates; LTT plots; Molecular dating; Pantropical distribution; Phylogeny; Rutaceae; *Spathelia*; Spathelioideae

Introduction

Rutaceae (Rue family) is the largest family within the eudicot order Sapindales and contains approximately 161 genera and 2070 species (Stevens, 2001 onwards). Members of the family are mainly distributed in the tropical and subtropical regions of both the New 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).

So far, only two dated phylogenies are available for Rutaceae, and both focus on 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 is not yet feasible due to the lack of resolution and support, as well as an 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; Chapter 3). 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; Chapter 3). The subfamily consists of 29 species in eight genera: *Bottegoa* Chiov. (1 spp.), *Cedrelopsis* Baill. (8 spp.), *Cneorum* L. (2 spp.), *Dictyoloma* A.Juss. (1 spp.), *Harrisonia* R.Br. ex A.Juss. (3 spp.), *Ptaeroxylon* Eckl. & Zeyh. (1 spp.), *Sohnreyia* K.Krause (4 spp.), and *Spathelia* L. (9 spp.) (Appelhans *et al*., 2011; Chapter 3). *Harrisonia* is widespread in tropical Africa and Australasia, 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; Chapter 3). 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 pantropically distributed. Spathelioideae is divided into two clades, one being strictly Neotropical, the other being Paleotropical.

The combination of monotypic genera, narrow endemicity, and pantropical distribution of the group makes Spathelioideae a particularly interesting group for biogeographical studies. The small number of overall taxa makes a biogeographic 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 Paleotropical subclades indicates a vicariance pattern (break-up of Gondwana) or a long-distance dispersal pattern; (2) to determine the distribution patterns within the clades, especially the dispersal over large distances and the colonisation of islands (Canary Islands, Caribbean Islands, Madagascar); (3) to investigate whether diversification rates were constant through time or if climatic fluctuations in different geological epochs caused change in diversification rates; and (4) to shed further light on the geographical origin of the Rutaceae family.

Materials & 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; Chapter 3) and it contains all eight genera of Spathelioideae, and 83,3% of its species (25 out of 30 [29 described and one undescribed species]). Only one species of *Sohnreyia* and four species of *Cedrelopsis* are not sampled in this study. Additionally, we include taxa from all families and subfamilies of Sapindales in order to be able 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, Malvales) were chosen as outgroups in the phylogenetic analyses and *T. cacao* alone was used as outgroup in the molecular dating analyses.

Our alignment is based on the chloroplast regions *atpB*, *rbcL*, and *trnL-trnF*, and we obtained the majority of sequences from Genbank (www.ncbi.nlm.nih.gov; see Table 5-1 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 EMBL Bank (http://www.ebi.ac.uk/embl/) under the accession numbers given in Table 5-1 and voucher information is specified in Table 5-2.

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 PCR reactions were performed under a standard protocol (see Appelhans *et al.*, 2011; Chapter 3) using a primer annealing temperature of 53°C, and extension at 72°C for 1.25 min. The chloroplast markers, *atpB*, *rbcL*, and *trnL-trnF* 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 SequencherTM (Gene Codes, Ann Arbor, Michigan, USA). Separate alignments for each marker were assembled manually in MacClade 4.08 (Sinauer Associates Inc., Sunderland, Massachusetts, USA). Indel coding was done in 15 sites in the *trnL-trnF* alignment, summing to 67bp. All indels were between four and six 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). 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).

Taxon	*rbcL*	*atpB*	*trnL-trnF*
Spathelioideae (Rutaceae)			
Bottegoa insignis	AJ402931	FR747871	FR747905
Cedrelopsis gracilis	FR747839	FR747873	HM637911
Cedrelopsis grevei	FR747842	FR747876	FR747908
Cedrelopsis rakotozafyi	FR747841	FR747875	HM637909
Cedrelopsis sp. nov.	FR747843	FR747877	FR747909
Cedrelopsis trivalvis	FR747840	FR747874	FR747907
Cneorum pulverulentum	FR747836	AF209567	EU853787
Cneorum tricoccon	FR747837	GU178995	GU178987
Cneorum tricoccon	-	GU178994	GU178988
Dictyoloma vandellianum	FR747846	FR747880	FR747912
Dictyoloma vandellianum	FR747845	FR747879	FR747911
Harrisonia abyssinica ssp. occidentalis	FR747833	FR747869	FR747904
Harrisonia abyssinica ssp. occidentalis	FR747832	FR747868	FR747903
Harrisonia abyssinica ssp. abyssinica	FR747835	GU178993	GU178986
Harrisonia abyssinica ssp. abyssinica	FR747834	FR747870	FR747930
Harrisonia brownii	FR747828	FR747864	FR747899
Harrisonia perforata	FR747831	FR747867	FR747902
Harrisonia perforata	FR747829	FR747865	FR747900
Harrisonia perforata	FR747830	FR747866	FR747901
Ptaeroxylon obliquum	FR747838	FR747872	FR747906
Spathelia bahamensis	FR747855	FR747889	FR747921
Spathelia brittonii	FR747847	FR747881	FR747913
Spathelia coccinea	FR747852	FR747886	FR747918
Spathelia cubensis	FR747856	FR747890	FR747922
Spathelia excelsa	AF066798	AF066854	EU853820
Spathelia glabrescens	FR747849	FR747883	FR747915
Spathelia sorbifolia	FR747848	FR747882	FR747914
Spathelia sorbifolia	FR747850	FR747884	FR747916
Spathelia sorbifolia	FR747851	FR747885	FR747917
Spathelia splendens	FR747853	FR747887	trN747919
Spathelia splendens	FR747857	FR747891	FR747923
Spathelia splendens	FR747860	FR747894	FR747926
Spathelia terminalioides	FR747844	FR747878	HM637910
Spathelia ulei	-	FR747898	FR747931
Spathelia vernicosa	FR747859	FR747893	FR747925
Spathelia vernicosa	FR747863	FR747897	FR747929

Taxon	*rbcL*	*atpB*	*trnL-trnF*
Spathelioideae (Rutaceae) - continued			
Spathelia vernicosa	FR747858	FR747892	FR747924
Spathelia wrightii	FR747854	FR747888	FR747920
Spathelia wrightii	FR747862	FR747896	FR747928
Spathelia wrightii	FR747861	FR747895	FR747927
Rutaceae (other subfamilies)			
Acronychia acidula	U38862	-	AF026025
Aegle marmelos	AF066811	AF066839	AY295294
Atalantia ceylanica	AF066812	AF066840	AY295288
Balfourodendron riedelianum	-	-	EU853779
Bergera koenigii	AB505905	EF118832	EF126637
Calodendrum capense	AF066805	AF066834	EF489250 + AF025511
Casimiroa edulis	AF066808	EU042767	GU593003
Choisya mollis	AF066800	AF066829	EU853784
Chorilaena quercifolia	AF066810	AF066838	EU853785
Citrus sinensis	AB505951	EF118866	EU369570
Clausena excavata	AF066813	AF066841	AY295284
Cneoridium dumosum	FN552678	-	EF489256
Correa pulchella	AF066816	AF066844	EU853790
Dictamnus albus	AF066801	AF066830	EU853792
Diplolaena dampieri	AF066807	AF066836	EU853794
Eremocitrus glauca	AF066819	AF066847	AY295293
Eriostemon brevifolius	AF156883	AF156882	FJ716787
Flindersia australis	FAU38861	EF118872	EF126677
Glycosmis pentaphylla	AF066820	AF066849	AY295279
Halfordia kendack	-	-	EU853798
Helietta puberula	-	-	EU853799
Lunasia amara	AF066814	AF066842	EU853805
Melicope ternata	AF116271	AF066826	EU853808
Micromelum minutum	AB505902	EF118889	EF126691
Murraya paniculata	U38860	EF118891	AY295280
Orixa japonica	HE588085*	HE588080*	DQ225930 + DQ225875
Phellodendron amurense	AF066804	AF066833	FJ716781 + AF025523
Ptelea trifoliata	-	-	EU853813
Ptelea baldwinii	HE588086*	HE588081*	-

Taxon	*rbcL*	*atpB*	*trnL-trnF*
Rutaceae (other subfamilies) - continued			
Ruta graveolens	RGU39281	AF035913	EU853815
Skimmia anquetilia	AF066818	AF066846	EF126698
Tetradium ruticarpum	GQ436747	-	DQ225983 + DQ225912
Toddalia asiatica	-		DQ226011 + DQ225923
Triphasia trifolia	AB505911	EF118902	AY295297
Vepris lanceolata	-	-	EU853823
Zanthoxylum monophyllum	ZMU39282	AF035919	EF655855
Simaroubaceae			
Ailanthus altissima	AY128247	AF035895	GU593006
Brucea javanica	EU042986	EU042778	GU593011
Castela erecta	EU042990	EU042781	GU593013
Eurycoma apiculata	EU042995	EU042786	GU593014
Hannoa chlorantha	EU042998	EU042789	GU593015
Holacantha emoryi	EU043002	EU042793	GU593016
Leitneria floridana	AF062003	EU042794	GU593017
Nothospondias staudtii	EU043004	EU042795	GU593018
Odyendyea gabonensis	EU043005	EU042796	GU593019
Perriera madagascariensis	EU043007	EU042798	GU593020
Picrasma javanica	EU043011	EU042802	GU593021
Picrolemma sprucei	EU043014	EU042804	GU593023
Quassia amara	EU043017	EU042807	GU593026
Samadera indica	EU043020	EU042810	GU593028
Simaba guianensis	EU043034	EU042824	GU593030
Simarouba berteroana	EU546231	EU546249	GU593032
Soulamea sp.	EU043042	EU042832	GU593033
Meliaceae			
Aglaia sp.	AB586406	-	-
Azadirachta indica	AJ402917	-	EF489263
Cedrela odorata	AJ402938	HE588082*	AB057509 + AB057455
Dysoxylum gaudichaudianum	AY128227	-	-
Dysoxylum caulostachyum	-	-	AB057530 + AB057476
Khaya anthotheca	AJ402964	-	-
Khaya grandifoliola	-	HE588083*	HE588087 *

Taxon	*rbcL*	*atpB*	*trnL-trnF*
Meliaceae - continued			
Melia azedarach	EU042973	EU042764	FM179536
Nymania capensis	AY128238	AF066855	-
Swietenia macrophylla	AY128241	AF066857	EF489262
Toona ciliata	-	EF118901	EF126701
Toona sp.	AY128243	-	-
Trichilia emetica	TEU39082	AF066851	-
Trichilia pallida	-	-	FJ039159
Sapindaceae / Xanthoceraceae			
Acer campestre	DQ978399	-	AF401189
Acer saccharum	EU676897	AF035893	AF401173
Aesculus pavia	U39277	AF035894	EU721462 + EU721274
Cupaniopsis anacardioides	L13182	AF035903	EU721387 + EU721199
Dodonea viscosa	DQ978445	-	DQ978578
Handeliodendron bodinieri	DQ978446	-	EF186776
Litchi chinensis	AY724361	-	EU721341 + EU721152
Koelreuteria paniculata	KPU39283	AJ235513	EU721506 + EU721318
Melicoccus pedicellaris	FJ038160	-	FJ039343
Xanthoceras sorbifolium	AF206833	AF209697	EU721337 + EU721148
Anacardiaceae			
Anacardium occidentale	AY462008	-	DQ131556
Cyrtocarpa procera	U39272	-	-
Dobinea delavayi	EU123469	-	-
Pistacia vera	AJ235786	AJ132282	AY677209
Poupartia minor	-	-	AY594530
Rhus copallina	U00440	AF035912	AY640438
Schinus molle	U39270	AF035914	AY640463
Searsia leptodictya	EU213507	-	AY640466
Spondias radlkoferi	GQ981883	-	-
Burseraceae			
Beiselia mexicana	AJ402925	-	-
Bursera inaguensis	L01890	AF035899	-
Bursera tecomaca	-	-	FJ466463

Taxon	*rbcL*	*atpB*	*trnL-trnF*
Burseraceae - continued			
Canarium harveyi	FJ466631	-	FJ466468
Commiphora edulis	FJ466630	-	FJ466480
Dacryodes buettneri	FN796555	-	FM162285
Protium sagotianum	FJ037983	-	FJ039109
Santiria trimera	FN796551	-	FN796604
Kirkiaceae			
Kirkia acuminata	-	HE588084*	HE588088 *
Kirkia wilmsii	KWU38857	-	-
Nitrariaceae			
Nitraria tangutorum	DQ267158	-	DQ267166
Peganum harmala	DQ267164	-	DQ267173
Tetradiclis tenella	AJ403009	-	-
Biebersteiniaceae			
Biebersteinia heterostemon	DQ408667	EF431915	-
Outgroups (Malvales)			
Gossypium hirsutum	M77700	AJ233063	AF031434
Theobroma cacao	AF022125	AJ233090	EF010969

Table 5-1. Taxa studied in molecular phylogenetic analyses. Voucher information for the specimens sequenced here and EMBL/ GenBank accessions for the three markers are displayed. '-' indicates that there is no sequence available for that marker. An asterisk (*) marks sequences that were generated in this study.

Species	Herbarium voucher	Date	Location
Cedrela odorata	M. Appelhans, MA 299 (L)	16.04.2009	Nationale Plantentuin Meise, Belgium
Khaya grandifoliola	M. Appelhans, MA 308 (L)	16.04.2009	Nationale Plantentuin Meise, Belgium
Kirkia acuminata	M. Appelhans, MA 393 (L)	22.06.2011	Cultivated by the first author
Orixa japonica	M. Appelhans, MA 246 (L)	21.07.2006	Botanical Garden Marburg, Germany
Ptelea baldwinii	M. Appelhans, MA 247 (L)	21.07.2006	Botanical Garden Marburg, Germany

Table 5-2. Voucher information for the additionally sequenced species.

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 1000th 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 indicated 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, as 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 has a much higher impact on our estimates about Spathelioideae than a more distant fossil in the phylogeny.

We used a leaf fossil of *Clausena* Burm.f., dated 27,36 ± 0,11Ma (Pan, 2010) for the stem lineage-age 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 Upper Paleocene to Upper 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 to *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 52Ma (Corbett & Manchester, 2004) was used as calibration point outside Rutaceae. *Ailanthus* has an excellent fossil record, beginning at 52Ma 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 biogeographic 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 recognisable because of their conspicuous samaroid fruits. Nevertheless, we did not use *Ptelea* fossils because its phylogenetic position 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 as 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 penalised likelihood analysis of the whole angiosperms (Magallón & Castillo, 2009). Magallón & Castillo (2009) performed two analyses: in their ´relaxed´ dating, they included 125Ma as a maximum age for the eudicot crown node, and in the ´constrained´ dating, they used 130Ma 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 birth-death process. All fossil calibration points were assigned a lognormal prior. Rutaceae have an abundant fossil record (Gregor, 1989), and especially fossils of "*Euodia*", *Toddalia* Juss., *Zanthoxylum* L., and *Ailanthus* (Simaroubaceae) have a robust record (Gregor, 1979; Tiffney, 1980; Gregor, 1989; 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 45Ma). 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 when analysing calibration scheme 2 and we had to enlarge the standard deviation from 0.75 to 0.95 (Table 5-3). 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 5-3. 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 and 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 v.1.6.1. and a maximum clade credibility tree using a posterior probability limit of 0.5 was calculated using TreeAnnotator v.1.6.1.

Ancestral area reconstruction (AAR)

AAR analyses were performed using the dispersal-extinction-cladogenesis model as implemented in Lagrange (version from 17. Jan. 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 *rbcL*, *atpB*, and *trnL-trnF* using the settings described above with an enforced monophyly of *Cedrelopsis* and a maximum age constraint for Spathelioideae taken from our molecular dating analyses (calibration scheme 1). 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 (incl. 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. 5-1). Areas A and D were separated because *Bottegoa* is only present in Central Eastern Africa (D) whilst *Harrisonia abyssinica* occurs throughout both areas. Area G describes the distribution of *Dictyoloma* and *Sohnreyia*. Their distribution overlaps in the western part, whilst only *Dictyoloma* occurs the disjunct areas in Eastern Brazil and North-eastern Argentina. We did not separate the disjunct areas from area G, because the same species (*D. vandellianum*) is recognised throughout the whole distribution area (Groppo, 2010) and we were not able to include a specimen from the disjunct areas

	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 bound (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

Table 5-3. Settings of fossil calibration points and root height in the molecular dating analyses. *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, 2010), Palaeocene for "*Euodia*" (Tiffney, 1981), and early Eocene for *Ailanthus* (Corbett & Manchester, 2004; Clayton *et al.*, 2009). The age range for *Clausena* has been extended due to poor the fossil record (has been found only once). The age for the Sapindales includes the age of 98.01 – 98.51 estimated by Magallón & Castillo (2009).

into our analyses. Detailed information on the distribution of all species is given in Appendix 5-1.

Maximum number of areas in ancestral ranges was set to three and all geographic ranges that were considered biologically implausible (largely disjunct areas) were excluded from the analyses. Dispersal rates between all areas were 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. 5-1). The time slices correspond to the Quaternary & Neogene (0-23Ma; Time slice 1), the Oligocene (23-34Ma; Time slice 2), the Eocene (34-56Ma; Time slice 3) and the Paleocene & Late Cretaceous (56-75Ma; Time slice 4). The dispersal rates between most areas were not varied through time. Only the rates dealing with a transatlantic dispersal and dispersal from South America to the Caribbean were given different values at different time slices (Fig. 5-1).

Diversification analyses

We generated lineage through time plots (LTT plots) and evaluated the fit of the LTT plots to three generalised models of diversification (Paradis, 1998; see McKenna & Farrell, 2006) based on the maximum clade credibility tree from the molecular dating analyses. We chose
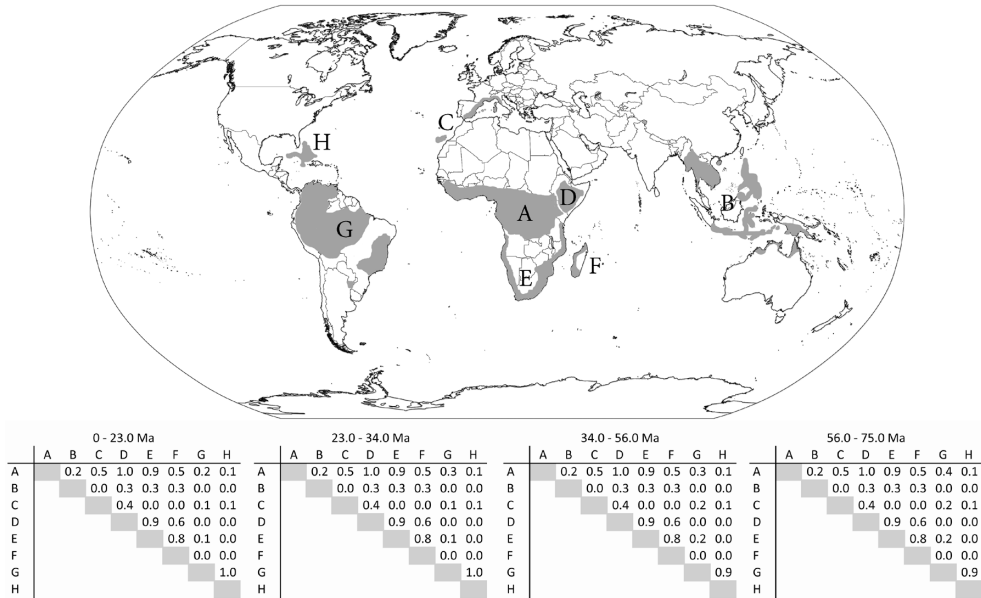

0 – 23.0 Ma

	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	0.0	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	0.0	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	0.0	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	0.0	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								

Fig. 5-1. Delimitation of the eight areas used for Ancestral Area Reconstruction and probability of dispersal throughout four defined time periods (0 – 23Ma: Quaternary & Neogene, 23 – 34Ma: Oligocene, 34 – 56Ma: Eocene, 56 – 75Ma: Palaeocene & Late Cretaceous). Area designations are as follows: A = Central-western and central Africa, B = Southeast Asia (incl. 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 given in Appendix 1.

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 the 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).

LTT plots represent a schematic visualisation 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. LTT plots were conducted using Laser version 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 v.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 to the LTT plots of the Spathelioideae dataset.

Laser version 2.3 was also used to test the fit of the Spathelioideae LTT plot to three generalised 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 AIC) for all three models were estimated using Laser version 2.3 as implemented in R and plotted through time in intervals of 2.5Ma.

The LTT plots and the test for the fit of the diversification models were done on the ingroup tree (= Spathelioideae) and separately on the Neotropical and Paleotropical 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 (Fig. 5-2 and 5-3). 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 Paleotropical subclade (extending to subtropical areas) including *Bottegoa*, *Cedrelopsis*, *Cneorum*, *Harrisonia*, and *Ptaeroxylon*. The relationships within Spathelioideae are largely congruent to those described by Appelhans *et al.* (2011; Chapter 3) and the further grouping within Rutaceae agrees with previous phylogenetic analyses (Chase *et al.*, 1999; Groppo *et al.*, 2008; Salvo *et al.*, 2008, 2010).
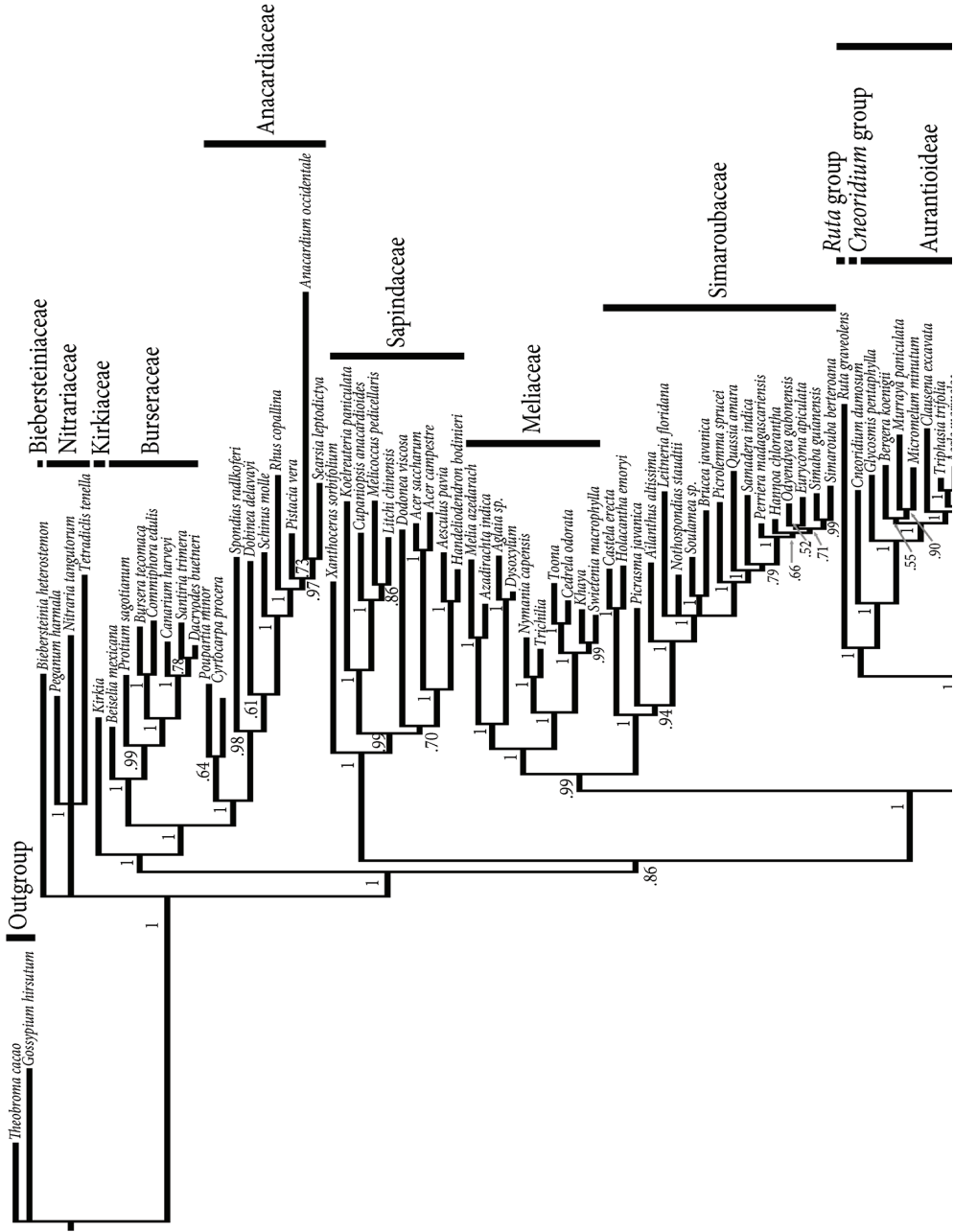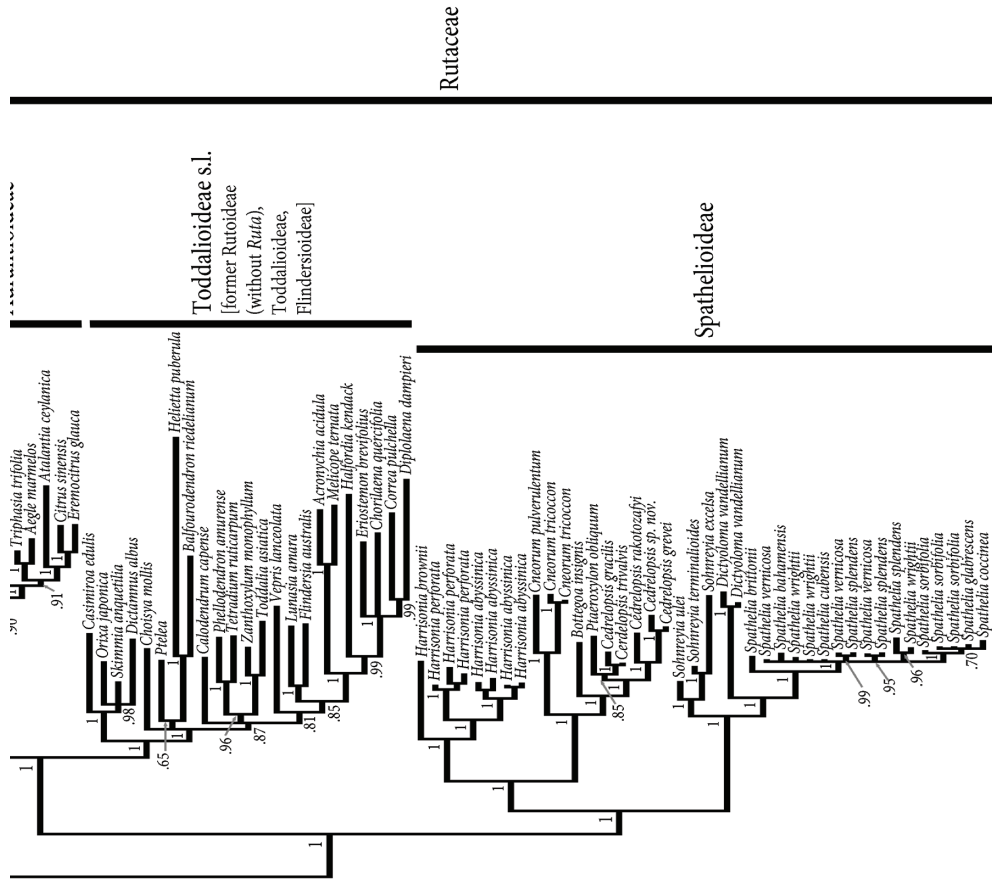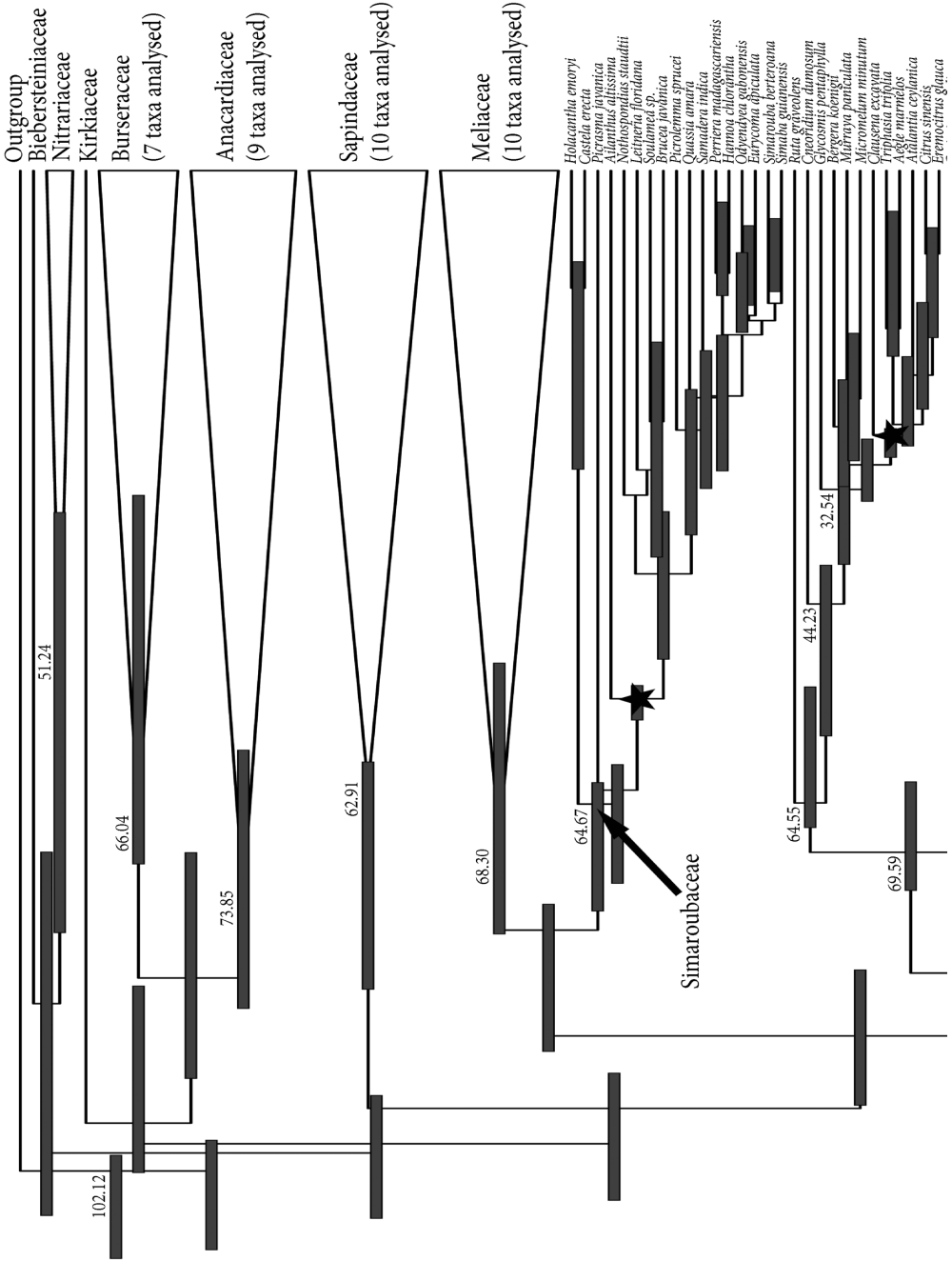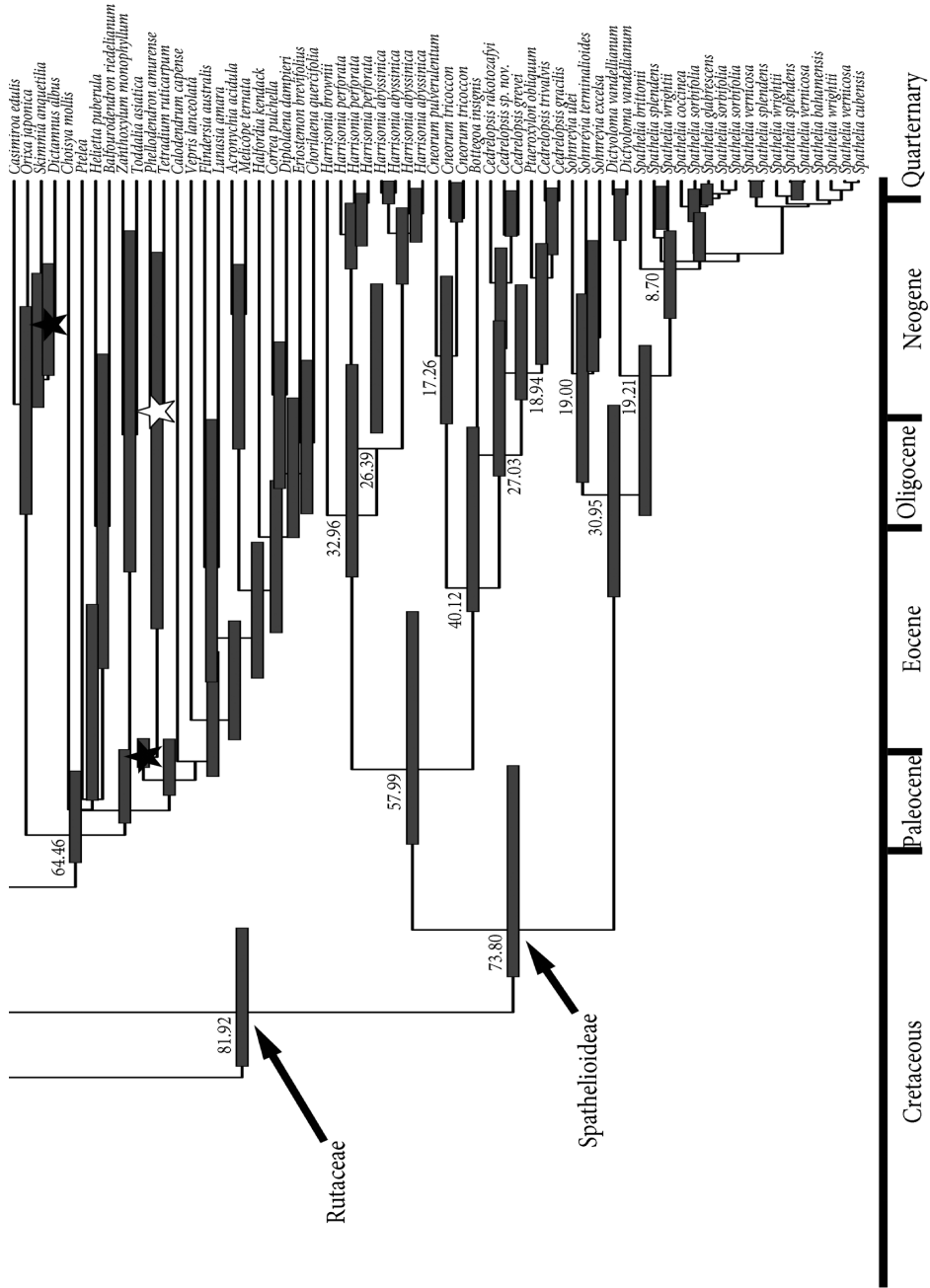
Fig. 5-2. 50% majority rule consensus tree of the combined *atpB*, *rbcL* and *trnL-trnF* dataset from the Bayesian analysis. Posterior probability values are indicated above the branches (next to the branches or marked with an arrow in case of limited space).

▲   **Fig. 5-3.** Maximum clade credibility tree of the combined *rbcL, atpB* and *trnL-trnF* dataset 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 as triangles. 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).

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. 5-3) are highly similar to those from fossil calibration scheme one and therefore a comparison between the dates is only done for the most important nodes: Rutaceae and Spathelioideae.

The effective sample size (ESS) and the trace of parameters (visualised 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. 5-3.
The age estimates for most families within Sapindales are between 60 and 80Ma. A mean age of 82Ma was estimated for Rutaceae with a credible interval ranging from 74 – 87Ma, suggesting an origin of Rutaceae in the Late Cretaceous. The age of Rutaceae s.s. (without Spathelioideae) was estimated to 70Ma (Late Cretaceous; 62 – 73Ma). Spathelioideae and the split into a Neotropic- and a Paleotropic subclade were dated to 74Ma (Paleocene or Late Cretaceous; 58 – 78Ma). The split of *Harrisonia* from the other genera in the Paleotropic subclade possibly occurred in the Early Eocene or Paleocene (Mean age: 58Ma; Credible interval: 42 – 65Ma). Our estimates suggest a split between *Cneorum* and Ptaeroxyleae (Appelhans *et al.*, 2010; Chapter 3; *Bottegoa, Cedrelopsis, Ptaeroxylon*) in Oligocene or Late Eocene (40Ma; 24 – 42Ma). The split of the two species of *Cneorum* was dated to the Miocene (17; 9 – 24Ma) and the split between *Bottegoa* and *Cedrelopsis/Ptaeroxylon* dates to the Miocene or Late Oligocene (27Ma ; 14 – 29Ma).
Within the Neotropical lineage, a further splitting occurred much later than in the Paleotropical subclade. *Sohnreyia* possibly split from *Dictyoloma* and *Spathelia* in the Early Miocene, Oligocene or Late Eocene (31Ma; 22 – 41Ma). The splitting of *Dictyoloma* from *Spathelia* might have occurred in or Early Miocene or Oligocene (19Ma; 16 – 33Ma) and *Spathelia brittonii* potentially diverged from the other *Spathelia* species in Late or mid Miocene (9; 5 – 14Ma).
The age of Rutaceae s.s. (excluding Spathelioideae) as inferred from the second calibration scheme was about 2Ma older (84Ma; 75 – 88Ma) 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 4Ma older and the credible intervals were also largely congruent (78Ma; 59 – 80Ma). 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 except for two nodes (Fig. 5-4, Table 5-4), the relative probability values for the areas at the nodes shown in Fig. 5-4 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. 5-4, Table 5-4) were the least congruent, but the AARs 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 combination of the adjacent areas Central-western and central Africa [A], Central Eastern Africa [D], and Southern Africa [E].

The base of Fig. 5-4 shows the split into a Central-western & central African lineage and a South American lineage [A – G]. The splitting into the *Harrisonia-* and the "*Cneorum* & Ptaeroxyleae"-lineages happened in the African area A. At the base of *Harrisonia* the Lagrange results suggest a wide distribution area [ABD] of which two South East Asian lineages (*H. brownii*, *H. perforata*, [B]) and one widespread African lineage (*H. abyssinica*, [AD]) emerged. The "*Cneorum* & Ptaeroxyleae"-lineage split into a Northern (*Cneorum*, [C]) and a Central & Southern lineage ([AE or AD], Table 5-4) and the latter splits into an Eastern African (*Bottegoa*, [D]) and later a Southern African (*Ptaeroxylon*, [E]) and a Madagascan (*Cedrelopsis*, [F]) lineage.

Within the South American subclade, a dispersal from the South American mainland [G] to the Caribbean Islands occurred (*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.

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

Table 5-4. Alternative ancestral areas and area splits for nodes (1) and (2) from Fig. 5-4. For all other areas at the nodes of Fig. 5-4, the relative probability from the Lagrange analysis was at least 10% higher than the second suggestion for an area combination.
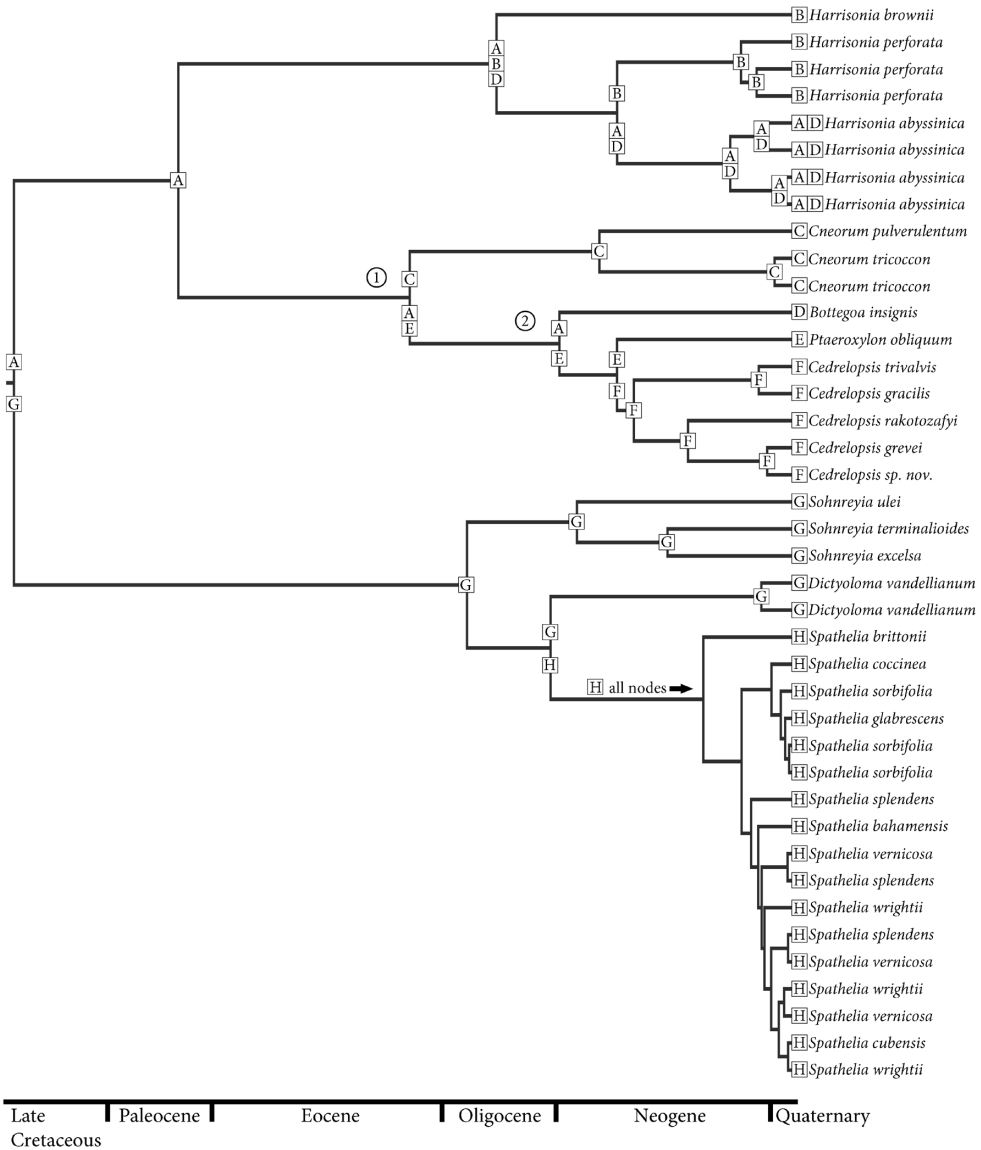
Fig. 5-4. Ingroup chronogram showing the results of the Ancestral Area Reconstruction analysis using Lagrange. AARs with highest likelihood values are shown as boxes at each node. Single boxes or combined boxes indicate ancestral ranges confined to a single or two or three areas. Boxes separated by a 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 5-4). Area designations are as follows: A = Central-western and central Africa, B = Southeast Asia (incl. parts of tropical Australia), C = Western Mediterranean and Canary Islands, D = Central Eastern Africa, E = Southern Africa, F = Madagascar, G = Northern South America, H = Caribbean region.

Diversification analyses

The curve of the empirical LTT plot is approximately parallel to the curve of the simulated LTT plots (Fig 5-5; upper diagram), 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 35Ma and 75Ma than in the following time span (35Ma until present). Between about 18Ma and 32Ma, the gradient of 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 to 3Ma.

The lower diagram in Fig. 5-5 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 10Ma ago. It is interesting that within the last 5Ma, the curve for Model C approaches the values for Model A. Also, the curve for Model C forms a valley at about 20Ma, which stands for an increase of the likelihood for an abrupt change in diversification rates.

The separate analyses for the Neotropical and the Paleotropical clades (results not shown) delivered very similar results as the combined analyses. Unlike the curve for the combined and the Neotropical analyses, the curve of the Paleotropical clade does not show a further increase in diversification rate in the Pliocene-Pleistocene period. The model testing analyses for the Neotropical and the Paleotropical clades (not shown) also show similar results compared 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 20Ma. In accordance with the empirical LTT plots, an increase of likelihood for Model C in the Pliocene/Pleistocene is not observed in the Paleotropical clade.

Discussion

Spathelioideae – Age and biogeographic patterns

Our analyses (Fig. 5-3) reveal an age of 74Ma (58 – 78Ma) and therefore also point to a Late Cretaceous origin of Spathelioideae. The age of Spathelioideae also marks the divergence of the Neotropical and Paleotropical lineages and the AAR reveals a split into a South American and a Central-western & central African lineage ([A – G], Fig. 5-4). 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 geographic origin of the family is not known (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 Paleotropic clades (*Ruta*-clade, *Cneoridium*/*Haplophyllum*-clade; Groppo *et al.*, 2008; Salvo *et al.*, 2010). Both are probably cases of long-distance dispersal as *Cneoridium* might have diverged from its sister group (*Haplophyllum* or *Haplophyllum* +
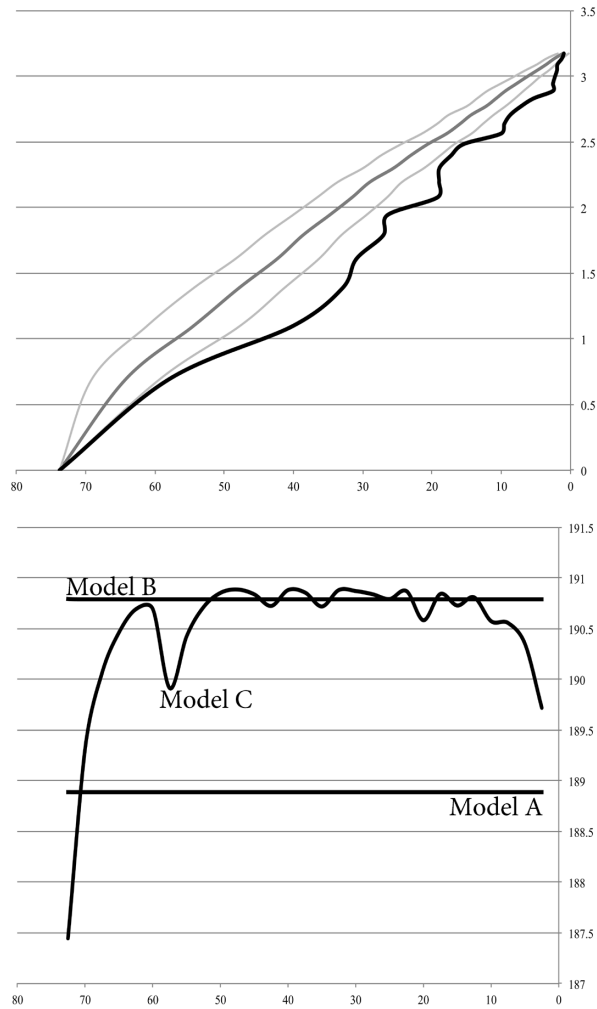
Fig. 5-5. Lineage-through-time (LTT) plot and test for the fit of three diversification models of the ingroup (Spathelioideae). The upper chart shows the LTT plot. The black line shows the empirical LTT plot for Spathelioideae, the dark grey line equates to the simulated LLT plot and the light grey lines delimit the 95% confidence interval of the simulated LTT curve. Values on the x-axis are given in Ma, those on the y-axis are logarithmic values of the numbers of lineages. The lower chart displays 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 AIC.

Aurantioideae) in the Eocene (Salvo *et al.*, 2010, 2011) and *Thamnosma* possibly originated in the Miocene with the 11 species of the genus having a disjunct distribution area in south-western North America, southern Africa, Somalia and the southern Arabian Peninsular (Oman, Yemen incl. Socotra) (Thiv *et al.*, 2011). A Paleotropic origin of all main clades except Toddalioideae s.l. and Spathelioideae is evident, so there is some evidence that Rutaceae have a Paleotropic origin, possibly a North-Tethyan origin (Kubitzki *et al.*, 2011). Assuming Ruta-ceae to be Paleotropic, an origin of the Spathelioideae stem lineage in Central-western & cen-tral Africa [area A] 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 119-105Ma (McLoughlin, 2001), which is at least 25 Ma older than the upper limit of our credible interval (Fig. 5-3). 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 might 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 (break-up of South America and Africa).

The ancestor of *Harrisonia* might have diverged from the other Paleotropical members either in the Eocene or Paleocene (Fig. 5-3) and has an African origin [A]. By the time of the first diversification of *Harrisonia* in the Miocene, Oligocene or Late Eocene, the genus might have had a broad distribution already in Africa and Asia (Fig. 5-4, [ABD]). Assuming a broad distribution of the ancestor, we suppose that *Harrisonia* dispersed to Asia only once. *Har-risonia brownii* is distributed in the Eastern part of Australasia and it is possible that it´s 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 South East 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. per-forata*. Nothing is known about the dispersal vector(s) of *Harrisonia*. Judging from it´s wide distribution and it´s fleshy fruits, *Harrisonia* 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 in-ferred from morphology and due to the description of a *Cneorum* species from Cuba (Borhi-di, 1991; Riera *et al.*, 2002). However, it has been shown that the Cuban *C. trimerum* is not a distinct species, but 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; Chapter 4).

Based on this, we did not include a Caribbean distribution of *Cneorum* in our analyses. How-ever, there has been no dated phylogeny of *Cneorum* prior to this study and we deliver the first arguments inferred from molecular dating that *Cneorum* is not a relict genus. Our results (Fig. 5-3) show that the common ancestors of *Cneorum* and Ptaeroxyleae might have di-verged 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 are inferred to be around 20Ma old (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 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 which 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 (Fig 5-4, Table 5-4, node 2). The relationship between *Ptaeroxylon* and *Cedrelopsis* is not clear from our analyses (*Ptaeroxylon* nested within *Cedrelopsis* but without support; Fig 5-2, 5-3). However, Razafimandimbison *et al.* (2010; Chapter 2) resolved a sister group relationship between both genera and we constrained the monophyly of *Cedrelopsis* for the Lagrange input tree accordingly. Our results (Fig. 5-4) show that the stem lineage of *Cedrelopsis* split from a southern African ancestor. Madagascar has had no direct connection to Africa for the past 160Ma (Goodman, 2009) whilst 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*.

The ancestral area for the Neotropical clade is Northern South America [area G]. Conclusions about biogeographic patterns within this area are not feasible, especially because of the 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, 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; Chapter 3), so we would assume that the ancestor of *Spathelia* colonised 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 along such a long distance. Still Caribbean hurricanes have been hypothesised 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 in congruence with the splitting of the *Dictyoloma*/*Spathelia*-lineage (19Ma, 16 – 33Ma). 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. A similar land bridge scenario as that of Heinicke *et al.* (2007) has been proposed by Iturralde-Vincent & MacPhee (1999), called the GAARlandia land span. Iturralde-Vincent & MacPhee (1999) postulated a continuous land-connection between South America and Cuba between 33 and 35Ma. This land bridge seems to be rather too old considering our age estimates, but the period is still within the upper boundary of our credible interval for the split of *Dictyoloma* and *Spathelia* (Fig. 5-3).

The resolution and support for the relationships among *Spathelia* species is low (Fig. 5-2), but *S. brittonii*, the only species from western Cuba, is clearly sister to the rest of the species, which are distributed in eastern Cuba, Jamaica and the Bahamas. The ancestor of *S. brittonii* and the rest of the genus split in Pliocene or Late Miocene (9; 5 – 14Ma; Fig. 5-3). Cuba is characterised by mountainous areas, which are surrounded by lowlands (Woods & Sergile, 2009). Cuban *Spathelia* species are distributed in the areas of the Western and the Eastern mountain ranges (Beurton, 2008) and an isolation of *S. brittonii* during a fragmentation of the island through rising sea levels (Woods & Sergile, 2009) might have occurred. The further splitting of the Eastern lineage (all species except *S. brittonii*) of *Spathelia* may have taken place in Late Miocene or Pliocene (3 – 8Ma; Fig. 5-3). Higher sea levels in the Caribbean during the Early Pliocene (McNeill *et al.*, 2008) might have isolated the Eastern lineage of *Spathelia* in Eastern Cuba. The current species in Eastern Cuba have overlapping distributions and appear in similar habitats (Beurton, 2008; Appendix 5-1). Speciation might therefore be sympatric/parapatric, and also hybridisation might (have) occur(ed) in the East Cuban species. Lowering sea levels in the Late Pliocene and Pleistocene (McNeill *et al.*, 2008; Woods & Sergile, 2009) have facilitated dispersal to Jamaica. As the Jamaican group (2.5Ma, 1-4Ma; Fig. 5-3) is monophyletic (Fig. 5-2), we conclude that a single dispersal event took place. A fairly recent colonisation event brought *Spathelia* to the Bahamas (*S. bahamensis*). The subaerial exposure of the Bahamas started at 2.5 to 3Ma (McNeill *et al.*, 2008) and since then, many flora and fauna elements 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 5-5; lower diagram), the empirical LTT shows that the diversification rates from 35 – 75Ma were slightly lower than in the following time periods (Fig. 5-5; upper diagram). Between 18Ma and 32Ma, there are three time periods with increased diversification rates. The last and most abrupt of these periods (+/- 18-20Ma) corresponds with the beginning of the Miocene, so the temperature increase at the early Miocene might have been a trigger for speciation. A constant diversification rate is suggested also for the period between 18Ma and 32Ma (Fig 5-5; lower diagram). However, the curve for Model C, which stands for an abrupt change in diversification rate, clearly forms a valley at about 20Ma, which stands for an increase of the likelihood for an abrupt change in diversification rates. The distinct increase of the diversification rate in the last 2-3Ma before present observed in the empirical LTT plot mainly corresponds to the speciation within the genus *Spathelia* (Fig. 5-3). In the curves of the model-fit test, a tendency towards a change in diversification rates (increased likelihood for Model C) is already apparent from about 8Ma 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. 5-3). Thus, the likelihood for Model C would probably be even higher in a complete dataset (100% of the species). The increase in diversification rate corresponds to the Pliocene and Pleistocene. Sea level changes in the Late Pliocene and Pleistocene (McNeill *et al.*, 2008; Woods & Sergile, 2009) might have been triggers for speciation in *Spathelia* as described above.

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). A clear picture about the spatial origin of Rutaceae does not become apparent yet especially because the backbone phylogeny of Toddalioideae s.l. (as defined in the introduction) is not resolved and supported. However, most early branching clades are of Paleotropic origin suggesting a Paleotropic origin of the whole family. Based on that, Central-western & central Africa would be a likely ancestral area for Spathelioideae. The occurrence of Spathelioideae in the Neotropics might 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 (break-up of South America and Africa; assuming that our age estimates provide minimal ages). The Paleotropical taxa have their origin in Africa with dispersal events to the Mediterranean, the Canary Islands, Madagascar and Australasia.

Diversification analyses show that the lineages within Spathelioideae evolved most likely with a quite constant diversification rate throughout their evolution. Only during the increasing temperatures at the beginning of the Miocene and the vast climatic changes in the Pliocene and Pleistocene, abrupt changes in diversification rates became more probable.

The speciation of *Spathelia*, the biggest genus of the subfamily, probably occurred in the late Miocene, Pliocene and Pleistocene.

Acknowledgements

Appendix

Species	Area (Fig. 1)	Distribution	Vegetation type and altitude	Included in analysis
Bottegoa insignis	D	Ethiopia, Kenya, Somalia	*Acacia* and *Commiphora* bush-land; sandy soils, rocky sites on limestone, granite, basalt, sand-stone; 250-1200m	X
Cedrelopsis ambanjensis	F	Northern Madagascar	Only known from a few locations. Semi-deciduous forests	-
Cedrelopsis gracilis	F	Western Madagascar (Moron-dava region)	Only known from a few loca-tions. No further information in literature	X
Cedrelopsis grevei	F	Madagascar (throughout distribution of the genus)	Abundant; Didieraceae forest, shrubby vegetation with Euphor-biaceae and *Didiera*, on slopes and plateau, arid, subhumid and humid vegetation, on siliceous, calcareous, sandy, clayey soils and on riverine vegetation; 0-900m	X
Cedrelopsis longibracteata	F	South-Eastern Madagascar	Wet forests on laterite	-
Cedrelopsis microfoliolata	F	Madagascar (throughout distribution of the genus)	Xeromorphic shrubland, disturbed forests, xeromorphic forests, on rocky soil	-
Cedrelopsis procera	F	North-Eastern Madagascar	Wet forests on Gneiss; about 500m	-
Cedrelopsis rakotozafyi	F	Northern Madagascar	Dry forest or shrubland, on white, siliceous sands; sea level	X
Cedrelopsis trivalvis	F	Madagascar (throughout distribution of the genus)	Dry and disturbed forests, on rocky, volcanic, and basaltic soil	X
Cneorum pulverulentum	C	Canary Islands; Gran Ca-naria, Tenerife, La Palma, La Gomera, El Hierro	Xerophylic shrubby vegetation; 200-600m	X
Cneorum tricoccon	C	Western Mediterranean (Southern Spain, Southern France, Italy, Western Mediter-ranean Islands)	Maquis shrubland; 50-500(-1000) m	X
Dictyoloma vandellianum	G	Ecuador, Peru, Bolivia, West-ern Brazil (Acre, Amazonas, Rondônia, Pará), Eastern Brazil (Bahia, Minas Gerais, Espirito Santo, Rio de Janeiro, Sao Paulo), North-Western Argentina (Corrientes)	Disturbed areas, roadsides, sea-sonally dry forest; to 1500m	X
Harrisonia abyssinica	A, D	Widespread in tropical Africa; From Angola to Tanzania in the South to Guinea and Ethiopia in the North	Open forests, woodland, thicket, riverine vegetation; medium-rainfall regions	X

Species	Area (Fig. 1)	Distribution	Vegetation type and altitude	Included in analysis
Harrisonia brownii	B	Australasia; Philippines, Sulawesi, Java, Lesser Sunda Islands, southern Moluccas, New Guinea, Northern Australia (Kimberley region, Islands of Northern Territory, Gulf of Carpentaria and Cape York Peninsula)	Dry, open, hot places under distinctly seasonal conditions, locally common in thickets; often on limestone; up to 700m	X
Harrisonia perforata	B	SE Asia; Burma, Thailand, Laos, Vietnam, China [Hainan], Cambodia, the Malay Peninsula, northern Borneo, Sulawesi, Sumatra, Java, Bali, Philippines	Dry, open, hot places under distinctly seasonal conditions, locally common in thickets; often on limestone; up to 700m	X
Ptaeroxylon obliquum	E	South Africa (mainly Eastern). Disjunct occurrences in Angola and the Usambara Mountains (Tanzania)	Usually in drier habitats, semi-evergreen forests, bushland and thicket, semi-evergreen scrub-forest dominated by *Commiphora* and *Euphorbia*, subdesert rocks, *Colophospermum* woodland; on rocky soils, limestone, sandstone	X
Sohnreyia excelsa	G	Brazil (Amazonas)	Forest in the Rio Negro area	X
Sohnreyia giraldiana	G	North-East Colombia (close to Brazilian border)	Only known from the type location; rocky vegetation; 80-130m	-
Sohnreyia terminalioides	G	North-East Peru	Only known from a few locations (Río Nanay, near Iquitos); primary forest on white sand soil below 400m	X
Sohnreyia ulei	G	Venezuela (states Bolívar and Amazonas)	Shrubby, forested, or savannah habitats, often along watercourses, Bonnetia thickets; on sandstone; 100-2500m	X
Spathelia bahamensis	H	Bahamas (New Providence, Eleuthera, Cat Island)	Sandy vegetation, shrubland, rocky thickets; sea level	X
Spathelia brittonii	H	Eastern Cuba (Pinar del Rio province)	Limestone hills (Mogotes); 600m	X
Spathelia coccinea	H	Jamaica	Only known from one location in Trelawny parish	X
Spathelia cubensis	H	Western Cuba (Holguin and Santiago de Cuba provinces)	Xeromorphic shrubby vegetation on serpentine, pine woodlands, Mogotes, evergreen woodlands; 300-1000m	X
Spathelia glabrescens	H	Jamaica	Hilly parts of central parishes; 380-730m	X

Species	Area (Fig. 1)	Distribution	Vegetation type and altitude	Included in analysis
Spathelia sorbifolia	H	Jamaica	Open thickets and woodlands; well drained shale or limestone; 0-450m	X
Spathelia splendens	H	Western Cuba (Guantanamo and Holguin provinces)	Xeromorphic shrubby vegetation on serpentine, pine woodlands, disturbed gallery forests; 0-700m	X
Spathelia vernicosa	H	Western Cuba (Guantanamo, Holguin, and Santiago de Cuba provinces)	Humid mountainous woodlands on iron rich soils, xeromorphic shrubby vegetation on serpentine, pine woodlands, evergreen woodlands, occasionally on disturbed areas; 100-850m	X
Spathelia wrightii	H	Western Cuba (Guantanamo and Holguin provinces)	Xeromorphic shrubby vegetation on serpentine, pine woodlands, humid mountainous woodlands on iron rich soils, disturbed gallery forests; 200-900m	X

Appendix 5-1. Detailed information about the distribution of Spathelioideae taxa and the inclusion of taxa in our analyses. "X" = species included in analysis; "-" = species not included in analysis. Information taken from: Krause, 1914; Nooteboom, 1962; Wild & Phipps, 1963; Adams, 1972; Straka *et al.*, 1976; Correll & Correll, 1982; Hewson, 1985; White, 1990; Leroy & Lescot, 1991; Van der Ham *et al.*, 1995; Vásquez Martínez, 1997; Pennington *et al.*, 2004; Kallunki, 2005; Parra-O, 2005; Schönfelder & Schönfelder 2005; Beurton, 2008; Groppo, 2010.