ORIGINAL ARTICLE

Divergence time uncertainty and historical biogeography reconstruction – an example from Urophylleae (Rubiaceae)

Jenny E. E. Smedmark*, Torsten Eriksson and Birgitta Bremer

The Bergius Foundation at the Royal Swedish Academy of Sciences and Department of Botany, Stockholm University, SE-106 91 Stockholm, Sweden

ABSTRACT

Aim When hypotheses of historical biogeography are evaluated, age estimates of individual nodes in a phylogeny often have a direct impact on what explanation is concluded to be most likely. Confidence intervals of estimated divergence times obtained in molecular dating analyses are usually very large, but the uncertainty is rarely incorporated in biogeographical analyses. The aim of this study is to use the group Urophylleae, which has a disjunct pantropical distribution, to explore how the uncertainty in estimated divergence times affects conclusions in biogeographical analysis. Two hypotheses are evaluated: (1) long-distance dispersal from Africa to Asia and the Neotropics, and (2) a continuous distribution in the boreotropics, probably involving migration across the North Atlantic Land Bridge, followed by isolation in equatorial refugia.

Location Tropical and subtropical Asia, tropical Africa, and central and southern tropical America.

Methods This study uses parsimony and Bayesian phylogenetic analyses of chloroplast DNA and nuclear ribosomal DNA data from 56 ingroup species, BEAST molecular dating and a Bayesian approach to dispersal–vicariance analysis (Bayes-DIVA) to reconstruct the ancestral area of the group, and the dispersal–extinction–cladogenesis method to test biogeographical hypotheses.

Results When the two models of geographic range evolution were compared using the maximum likelihood (ML) tree with mean estimates of divergence times, boreotropical migration was indicated to be much more likely than long-distance dispersal. Analyses of a large sample of dated phylogenies did, however, show that this result was not consistent. The age estimate of one specific node had a major impact on likelihood values and on which model performed best. The results show that boreotropical migration provides a slightly better explanation of the geographical distribution patterns of extant Urophylleae than long-distance dispersal.

Main conclusions This study shows that results from biogeographical analyses based on single phylogenetic trees, such as a ML or consensus tree, can be misleading, and that it may be very important to take the uncertainty in age estimates into account. Methods that account for the uncertainty in topology, branch lengths and estimated divergence times are not commonly used in biogeographical inference today but should definitely be preferred in order to avoid unwarranted conclusions.

Keywords
Boreotropics, dispersal–extinction–cladogenesis, historical biogeography, long-distance dispersal, molecular dating, phylogeny, Rubiaceae, Urophylleae.

*Correspondence: Jenny Smedmark, Bergius Foundation, Department of Botany, Stockholm University, SE-106 91 Stockholm, Sweden.
E-mail: jenny@bergianska.se

Journal of Biogeography

INTRODUCTION

Molecular dating is becoming an increasingly common tool for testing hypotheses of historical biogeography. Geological and palaeontological evidence is used to put a time frame on biogeographical events, and, by converting the branch lengths of a phylogenetic tree into absolute time, different biogeographical hypotheses may be evaluated. The uncertainty in obtained age estimates is, however, often very large, and it may be important to take this into consideration when different biogeographical scenarios are compared. Some popular dating methods, such as nonparametric rate smoothing (NPRS) (Sanderson, 1997) and penalized likelihood (PL) (Sanderson, 2002), only give a point estimate of the age of a node. With these methods, confidence intervals of divergence times can be obtained by running multiple dating analyses on bootstrapped data with a fixed tree (Sanderson & Doyle, 2001). Although it is also possible to run multiple dating analyses on a Bayesian sample of phylogenetic trees and thereby obtain a measure of the uncertainty in topology and in estimated divergence times, this has only rarely been performed (e.g. Pfeil & Crisp, 2008; Antonelli *et al.*, 2009). Bayesian dating, as implemented for example in MULTIDIVTIME (Thorne & Kishino, 2002), where the dating analysis is run on a single phylogenetic tree, provides a measure of the uncertainty in estimated divergence times but does not account for the uncertainty in topology or branch lengths. Bayes-DIVA (Nylander *et al.*, 2008) is a recently developed approach to biogeographical analysis, which can account both for the topological uncertainty and for the uncertainty in biogeographical reconstruction. This method takes advantage of a Bayesian sample of phylogenetic trees, but does not allow divergence times to be included in the analyses. In comparison, a full Bayesian approach to molecular dating (as implemented for example in BEAST; Drummond & Rambaut, 2007) has the advantage of taking the uncertainty in topology and branch lengths into account, in addition to the uncertainty in estimated divergence times. The Bayesian sample of dated phylogenies may also be used to run multiple biogeographical analyses and thereby obtain a measure of the uncertainty in the biogeographical reconstruction. To our knowledge this has, however, not been done so far. Instead, a single tree [e.g. maximum likelihood (ML) or consensus tree] is generally used for biogeographal analysis, even when a Bayesian sample of trees is available (e.g. Lo Presti & Oberprieler, 2009; McDill *et al.*, 2009). In this study, we explore how the uncertainty in estimated divergence times affects conclusions about historical biogeography, using a group in the coffee family (Rubiaceae), namely Urophyllyeae Bremek. ex Verdec., as an example.

Urophylleae includes about 240 species (Govaerts, 2009), which are usually woody and have indehiscent, often fleshy, fruits with many seeds. It belongs in the Rubioideae (Bremer & Thulin, 1998) and, together with Ophiorrhizeae, probably constitutes the first branch within this clade (Rydin *et al.*, 2009). Urophylleae occurs in nearly all areas of the world where there is tropical rain forest, with the largest number of species found in Asia, fewer in Africa, and some in central and southern America. The first phylogenetic study of Urophylleae (Smedmark *et al.*, 2008) showed that the group consists of a large Palaeotropical clade and a smaller Neotropical clade, the latter consisting of the genera *Amphidasya* Standl. and *Raritebe* Wernham. That study also indicated, albeit with poor support, that the monotypic African *Temnopteryx* Hook. f. is the sister of these two groups (Smedmark *et al.*, 2008). Furthermore, resolution within the Palaeotropical clade was poor and many questions were left unanswered.

The pantropical distribution of Urophylleae raises a question: how did the group come to occupy such a wide and fragmented geographic range? One possibility is that it began to diversify in the Old World and attained its present distribution by long-distance dispersal to America and Asia. The importance of long-distance dispersal as a cause of geographic disjunctions in plant taxa is controversial. In a study of trans-Atlantic disjunct angiosperm genera it was claimed that it played an important role, especially by water and wind (Renner, 2004), but it has also been dismissed as a significant mechanism for the explanation of disjunct distribution ranges in plant taxa in general (Tiffney & Manchester, 2001). The connection between Africa and South America ceased to exist 96 Ma, but there is evidence from fossil pollen on both sides of the Atlantic indicating that plant dispersals were common for some time following this event (Morley, 2003). Trans-Atlantic dispersal could have been facilitated by some form of land connection, or by series of islands, for example the Walvis ridge or the Rio Grande Rise, which may have existed from the late Cretaceous into the early Cenozoic (Morley, 2003). In more recent times, long-distance dispersal across the Atlantic would have taken place primarily by wind or water (Renner, 2004).

In the region between Africa and Asia there were few connections during the Cenozoic that could have facilitated the dispersal of tropical plant taxa. The Indian plate collided with the Asian plate in the middle Eocene (50–39 Ma) and brought with it plant taxa that had dispersed from Africa, via Madagascar, to India in the late Cretaceous. At the time of the collision, a corridor of tropical climate existed from India to Southeast Asia, allowing the dispersal of many rain forest taxa into the Sunda region (Morley, 2003). It has been claimed, however, that plant dispersal takes place constantly between Africa, Madagascar, India and Malesia, aided by migrating birds, sea currents and wind (Thorne, 1973).

Another possibility is that Urophylleae, like several other groups of tropical angiosperms, were part of the boreotropical flora (Wolfe, 1975), which covered a large part of the Northern Hemisphere in the late Cretaceous and early Palaeogene (e.g. Melastomataceae, Renner *et al.*, 2001; Malpighiaceae, Davis *et al.*, 2004; Annonaceae, Richardson *et al.*, 2004; Burseraceae, Weeks *et al.*, 2005). During this time, the global climate was very warm (Zachos *et al.*, 2001), and tropical vegetation was present at much higher latitudes than it is today (Wolfe, 1975; Tiffney, 1985a). The existence of a connection between the Eurasian and North American continental plates, the North

Atlantic Land Bridge (NALB), during the late Palaeocene and early Eocene allowed boreotropical taxa to expand their ranges between these two continents (Tiffney, 1985b; Morley, 2003). There was also a connection between Asia and North America, the Bering Land Bridge (BLB), which was open at least from the early Palaeocene until the end of Miocene. Its role as a route of migration for tropical taxa has, however, been disputed (Tiffney & Manchester, 2001). It may have been located too far north (perhaps as far as 80 °N in the early Cenozoic), where a long period of winter darkness would have made it impossible for evergreen angiosperms to cross, and fossil evidence showing that tropical taxa have existed in the region is indeed scarce and questioned (Tiffney & Manchester, 2001).

The cooling climate during the late Eocene (Zachos *et al.*, 2001) caused tropical taxa to recede towards the equator, or to go extinct, and it has been suggested that boreotropical forests were replaced by frost-tolerant temperate vegetation about 40 Ma (Wolfe, 1978, 1980; Morley, 2000). Fossil and molecular dating evidence from Malpighiaceae (Davis *et al.*, 2004) does, however, show that this group probably crossed the north Atlantic in the Oligocene (34–31 Ma), which may be an indication that tropical taxa could in fact migrate between the continents as late as 31 Ma. Today, descendants of boreotropical taxa are found primarily in refugia in China, Southeast Asia and Central America, and a few species are also found in Macaronesia (Morley, 2000).
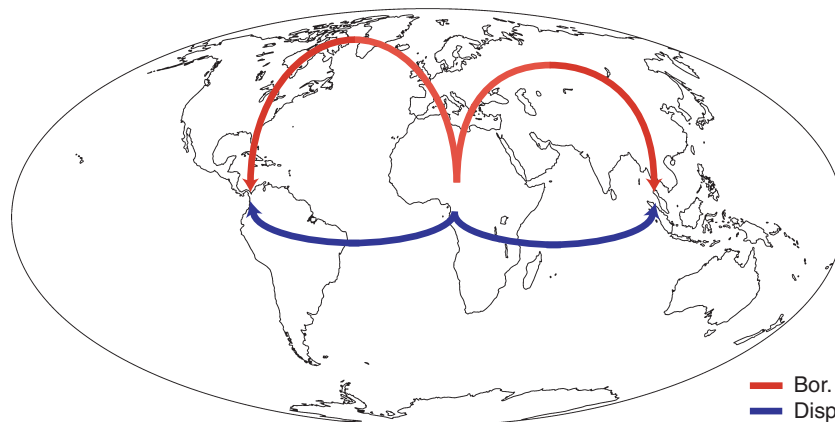
There are three aims of this study. The first is to test which of the two hypotheses chance dispersal or boreotropical migration (Fig. 1) best explains the distribution patterns of extant species of Urophylleae. Explicit models corresponding to these two hypotheses are constructed to test which provides a better explanation of the data using the dispersal–extinction–cladogenesis (DEC) method in the computer software LAGRANGE (Ree & Smith, 2008). The second aim is to evaluate how the uncertainty in estimated divergence times might affect conclusions by running multiple biogeographical analyses on a sample of dated phylogenies. The third aim is to resolve phylogenetic relationships among major lineages in Urophylleae, which is necessary in order to understand the historical biogeography of the group. An important question is whether African and Asian taxa within the large Palaeotropical clade belong to two separate clades that are sister groups, or if taxa from one region have evolved inside a group from the other region. Previous phylogenetic studies including more than one or two representatives from Urophylleae were based on chloroplast DNA (cpDNA) data alone (Smedmark *et al.*, 2008; Sonké *et al.*, 2008). In this study, both nuclear ribosomal (nrDNA) and cpDNA data are used, and, further to the previous phylogenetic study of the group (Smedmark *et al.*, 2008), 17 additional Urophylleae species are included.

MATERIALS AND METHODS

Material

Fifty-six Urophylleae species, 17 of which have not been included in any previous phylogenetic study, were selected to represent as many as possible of the genera in Urophylleae. The sample was limited by access to well-preserved herbarium material from which DNA could be extracted and amplified, and represents 23% of the species in the group and 61% of the genera. The missing genera are nearly all monotypic and all but one are Asian. All missing genera, as well as 96% of the missing species, have been classified in Urophylleae s. str. or Pauridiantheae s. str. (Bremekamp, 1966) based on morphological characters and are likely to belong in the Palaeotropical clade (Smedmark *et al.*, 2008). For nine of the eleven represented genera, type species are included. Fifteen outgroup species were also selected to obtain a good representation from the rest of Rubiaceae, especially from taxa in Rubioideae that have been shown to be closely related to Urophylleae, for example its sister group, the entirely Asian clade Ophiorrhizeae (Rydin *et al.*, 2009). For most species, leaf material was sampled from herbarium specimens but in some cases silica gel-dried material was used. Voucher specimens are listed in Table 1.

Figure 1 The two proposed hypotheses of historical biogeography for Urophylleae. Blue arrows represent long-distance dispersal (Disp.) and red arrows boreotropical migration (Bor.).

Table 1 Plant material of species from the study group Urophylleae, and of outgroup species from the remainder of Rubiaceae. Herbaria are designated by acronyms according to the Index Herbariorum (Thiers, 2009). EMBL accession numbers of DNA sequences produced in this study begin with HM. Regions for which no sequence could be obtained are marked with a dash.

Species	Origin	Voucher (collector, collection number and herbarium)	EMBL accession numbers			
			rps16	trnT–F	ITS	ETS
Amphidasya ambigua (Standl.) Standl.	Ecuador	Stähl et al. 3542 (GB)	AF129271[1]	EU145576[2]	EU145383[2]	–
Amphidasya colombiana (Standl.) Steyerm.	Ecuador	Stähl et al. 3542 (GB)	AF242906[3]	–	–	–
Amphidasya colombiana (Standl.) Steyerm.	Ecuador	Persson 120 (GB)	–	HM042591	HM042459	HM042517
Amphidasya longicalycina (Dwyer) C.M.Taylor	Costa Rica	Huber 2963 (CR)	AM900637[4]	–	HM042461	HM042518
Amphidasya sp.	Peru	Morawetz & Wallnöfer 114–11888 (GB)	AM900636[4]	HM042592	HM042460	HM042519
Cinchona pubescens Vahl	Cult. Göteborg Botanical Garden	Andersson 2214 (GB)	AF004035[5]	AY538451[6]	AY538356[6]	HM042505
Coccocypselum condalia Pers.	Brazil	Pirani & Bremer 4891 (SPF)	EU145499[2]	EU145547[2]	EU145358[2]	HM042509
Coussarea ilheotica Müll.Arg.	Brazil	De Carvalho et al. 4081 (K)	AM900597[4]	HM042586	HM042454	HM042510
Faramea crassifolia Benth.	Guyana	Jansen-Jacobs et al. 3882 (GB)	HM042567	HM042587	HM042463	HM042511
Faramea torquata Müll.Arg.	Ecuador	Stähl 3021 (GB)	HM042568	HM042588	HM042455	HM042512
Faramea trinervia K.Schum. & Donn.Sm.	Costa Rica	Gomez-Laurito 8374 (CR)	AM900598[4]	HM042589	–	HM042513
Lasianthus chevalieri Pit.	Vietnam	Averyanov et al VH2673 (AAU)	AM900596[4]	–	HM042453	HM042508
Luculia grandifolia Ghose	Cult. Stockholm Univ.	Bremer 2713 (S)	AM900593[4]	AJ346929[7]	AJ346896[7]	–
Luculia grandifolia Ghose	Cult. Uppsala University	Bremer 88–7 (UPS)	–	–	–	HM042504
Mussaenda scratchleyi Wernham	New Guinea	Drozd & Molem 118–11–13	AJ320079[8]	AJ847412	HM042451	HM042506
Neurocalyx zeylanicus Hook.	Sri Lanka	B & K Bremer 937 (S)	AM900594[4]	EU145562[2]	HM042457	–
Ophiorrhiza mungos L.	Cult. Uppsala University	Bremer 3301 (UPS)	AF004064[5]	DQ662151[9]	EU145377[2]	HM042514
Pauridiantha brideloides Verdc.	Tanzania	Bridson 584 (BR)	HM042573	HM042616	HM042486	HM042543
Pauridiantha canthiiflora Hook.f.	Gabon	Breteler 6746 (AAU)	AM900604[4]	HM042618	HM042488	HM042545
Pauridiantha callicarpoides (Hiern) Bremek.	Congo	Lejoly 86/347 (BR)	HM042574	HM042617	HM042487	HM042544
Pauridiantha dewevrei (De Wild. & T.Durand) Bremek.	Congo	Champluvier 5083 (K)	AM900603[4]	HM042619	HM042489	HM042549
Pauridiantha efferata N.Hallé	Gabon	van Valkenburg 2552 (BR)	HM042575	HM042620	HM042490	HM042550
Pauridiantha floribunda (K.Schum. & K.Krause) Bremek.	Nigeria	van Meer 1556 (BR)	HM042576	HM042621	HM042491	HM042551
Pauridiantha hirtella (Benth.) Bremek.	Liberia	Jacques-Georges 27404 (BR)	HM042577	HM042622	HM042492	HM042552
Pauridiantha liebrechtsiana (De Wild. & T.Durand) Bremek.	Cameroon	Breteler 1782 (BR)	HM042569	HM042615	HM042485	HM042532
Pauridiantha microphylla R.D.Good	Gabon	van Valkenburg 3009 (BR)	HM042578	HM042625	HM042494	HM042553
Pauridiantha mayumbensis (R.D.Good) Bremek.	Gabon	Diabata 1158 (UPS)	AM900601[4]	HM042624	HM042493	HM042554

Table 1 Continued

Species	Origin	Voucher (collector, collection number and herbarium)	EMBL accession numbers			
			*rps*16	*trn*T–F	ITS	ETS
Pauridiantha paucinervis (Hiern) Bremek.	Tanzania	Bremer 3090 (UPS)	AM900600[4]	EU145578[2]	EU145385[2]	HM042557
Pauridiantha pleiantha Ntore & Dessein	Gabon	van Valkenburg 2856 (BR)	HM042579	HM042626	HM042495	HM042558
Pauridiantha pyramidata (K.Krause) Bremek.	Central African Republic	Harris & Fay 644 (K)	AM900607[4]	HM042627	HM042496	HM042555
Pauridiantha schnellii N.Hallé	Liberia	Adanes 829 (UPS)	AM900606[4]	HM042628	HM042497	HM042560
Pauridiantha siderophila N.Hallé	Equatorial Guinea	Parmentier & Esono 2854 (BR)	HM042580	HM042629	HM042498	HM042562
Pauridiantha sp. 1	Congo	Devred 2297 (UPS)	AM900605[4]	HM042630	HM042499	HM042559
Pauridiantha sp. 2	Liberia	Jongkind 5454 (BR)	HM042581	HM042623	HM042500	HM042561
Pauridiantha sylvicola (Hutch. & Dalziel) Bremek.	Cameroon	Leeuwenberg 9711 (UPS)	AM900602[4]	HM042631	HM042501	HM042563
Pauridiantha symplocoides (S.Moore) Bremek.	Malawi	Lantz 123 (UPS)	AM900599[4]	HM042632	EU145386[2]	HM042546
Pauridiantha talbotii (Wernham) Ntore & Dessein	Cameroon	Davis 3057 (BR)	HM042583	HM042634	HM042502	HM042564
Pauridiantha udzungwaensis Ntore & Dessein	Tanzania	Gereau & Kayombo 4702 (BR)	HM042582	HM042633	–	HM042556
Pauridiantha viridiflora (Schweinf. ex Hiern) Hepper	Tanzania	Bidgood *et al.* 4888 (BR)	HM042584	HM042635	HM042503	HM042565
Pentaloncha humilis Hook.f.	Gabon	Breteler *et al.* 10985 (WAG)	AM900632[4]	HM042614	HM042476	HM042533
Pleiocarpidia enneandra	Malaysia	Wiart & Teo 44664 (P)	–	–	–	HM042535
Pleiocarpidia kinabaluensis Bremek.	Borneo	Beaman 8841 (S)	AM900610[4]	HM042608	HM042477	HM042534
Pocilocalyx crystallinus N.Hallé	Gabon	Breteler 6859 (BR)	–	HM042609	HM042478	HM042536
Pocilocalyx schumannii Bremek.	Cameroon	de Wilde 1573 (BR)	HM042571	HM042611	HM042482	HM042537
Pocilocalyx stipulosa (Hutch. & Dalziel) N.Hallé	Ivory coast	Wilde & Leeuwenberg 3515 (UPS)	AM900608[4]	HM042610	HM042479	HM042538
Praravinia verruculosa Bremek.	Borneo	Beaman 9670 (S)	AM900612[4]	HM042612	HM042480	HM042539
Pravinaria cf. *endertii* Bremek.	Brunei	Sands 5262 (K)	AM900614[4]	HM042613	HM042481	HM042547
Pravinaria leucocarpa Bremek.	Borneo	Beaman 7950 (S)	AM900613[4]	EU145580[2]	EU145388[2]	HM042548
Psychotria pittieri Standl.	New Guinea	Drozd & Molem 13 Nov 1998	AF002746[5]	AF152614[10]	AF071998[11]	–
Rariebe palicoureoides Wernham	Panama	Antonio 1697 (AAU)	AM900635[4]	HM042593	HM042481	HM042520
Spermacoce princeae (K.Schum.) Verdc.	Kenya	Luke & Luke (UPS?)	HM042566	HM042585	HM042452	HM042507
Stelechantha arcuata S.E.Dawson	Cameroon	de Wilde 8009 (BR)	HM042571	HM042594	HM042483	HM042540
Stelechantha cauliflora (R.D.Good) Bremek.	Gabon	van Valkenburg 2769 (BR)	HM042572	HM042595	–	HM042541
Stelechantha ziamaeana (Jacq.-Fél.) N.Hallé	Liberia	Adam 20999 (UPS)	AM900609[4]	HM042596	HM042484	HM042542
Temnopteryx sericea Hook.f.	Gabon	Tabak 99 (WAG)	AM900633[4]	HM042590	HM042458	HM042516
Trichostachys sp.	Cameroon	Sonké 1725 (UPS)	AM900595[4]	EU145560[2]	EU145373[2]	–
Urophyllum arboreum (Reinw. ex Blume) Korth.	Sumatra	Boeea 7887 (S)	AM900617[4]	EU145582[2]	–	–
Urophyllum blumeanum (Wight) Hook.f.	Thailand	Puff & Sridith 930724 (AAU)	AM900629[4]	–	HM042463	HM042521
Urophyllum britannicum Wernham	Papua New Guinea	Gideon *et al.* 76915 (K)	AM900623[4]	–	HM042464	HM042523
Urophyllum ceylanicum (Wight) Thwaites	Sri Lanka (S)	Klackenberg 214 (S)	AM900620[4]	HM042598	HM042465	HM042522
Urophyllum congestiflorum Ridl.	Brunei	Wong 1057 (K)	AM900621[4]	HM042599	HM042466	HM042524

Table 1 Continued

Species	Origin	Voucher (collector, collection number and herbarium)	EMBL accession numbers			
			rps16	trnT–F	ITS	ETS
Urophyllum corymbosum (Blume) Bremek.	Sabah	Ridsdale 2471 (L)	AM900611[4]	EU145577[2]	EU145384[2]	–
Urophyllum cyphandrum Stapf	Borneo	Beaman *et al.* 9549 (K)	AM900618[4]	HM042600	HM042467	–
Urophyllum ellipticum (Wight) Thwaites	Sri Lanka	Lundqvist 11085 (UPS)	AM900619[4]	EU145581[2]	EU145389[2]	–
Urophyllum glaucescens Valeton	Papua New Guinea	Damas 58925 (K)	AM900625[4]	HM042601	HM042468	HM042525
Urophyllum leucophlaeum Ridl.	Malaysia	Stone 12658 (AAU)	AM900626[4]	–	HM042469	–
Urophyllum longifolium (Wight) Hook.f.	Thailand	Larsen & Larsen 33340 (AAU)	AM900616[4]	HM042602	HM042470	HM042526
Urophyllum salicifolium Stapf	Borneo	Beaman 11523 (K)	AM900624[4]	HM042603	HM042471	HM042527
Urophyllum schmidtii C.B.Clarke	Thailand	Geesink *et al.* 6592 (AAU)	AM900627[4]	HM042604	HM042472	HM042528
Urophyllum sp. 1	Borneo	Beaman 7501 (S)	AM900630[4]	HM042605	HM042473	HM042530
Urophyllum sp. 2	Philippines	Ingle 725 (AAU)	AM900615[4]	HM042606	HM042474	HM042529
Urophyllum streptopodium Wall. ex Hook.f.	Sumatra	Boea 7766 (S)	AM900622[4]	HM042607	HM042475	HM042531
Xanthophytum borneense (Valeton) Axelius	Borneo	Axelius 316 (S)	EU145513[2]	EU145567[2]	EU145381[2]	HM042515

[1]Piesschaert *et al.* (2000), [2]Rydin *et al.* (2008), [3]J.H.E. Rova (Länstyrelsen, Jönköping, Sweden, unpublished data), [4]Smedmark *et al.* (2008), [5]Andersson & Rova (1999), [6]Andersson & Antonelli (2005), [7]Razafimandimbison & Bremer (2002), [8]Novotny *et al.* (2002), [9]Backlund *et al.* (2007), [10]Rova *et al.* (2002), [11]Nepokroeff *et al.* (1999).

Molecular methods and dataset construction

DNA extractions were carried out using the cetyl trimethyl ammonium bromide (CTAB) extraction method (Doyle & Doyle, 1990), and amplifications followed standard polymerase chain reaction (PCR) procedures. The *rps*16 intron was amplified using the rpsF and rpsR2 primers (Oxelman *et al.*, 1997), and sequenced using the same two primers. The *trn*T–F region was amplified in two parts. The *trn*T–L spacer and *trn*L intron were amplified with the a1F and iR primers (Razafi-mandimbison & Bremer, 2002), and for sequencing two additional primers, 820F and 940R (Rydin *et al.*, 2008), were also used. The *trn*L–F spacer was amplified with the 1250F and 2670R primers (Rydin *et al.*, 2008) and sequenced with the same two primers in combination with 1880F (Rydin *et al.*, 2008) and d (Taberlet *et al.*, 1991). The internal transcribed spacer (ITS) region was amplified using the primers P17 (Popp & Oxelman, 2001) and P25R (Oxelman, 1996) and sequenced with the same two primers. In some cases, when readings were poor, ITS2 and ITS3 (White *et al.*, 1990) were also used for sequencing. For amplification and sequencing of the external transcribed spacer (ETS) region, the primers Erit-F (Negrón-Ortiz & Watson, 2002) and 18S-E (Baldwin & Markos, 1998) were used. Sequences were assembled and edited using the PHRED (Green & Ewing, 2002) and PHRAP (Green, 1999) modules in Pregap4 and Gap4 (Staden *et al.*, 1998). All new sequences have been submitted to EMBL, and accession numbers are presented in Table 1. Sequence alignment was performed by eye, in the sequence alignment editor SE-AL (Rambaut, 1996).

Model selection and phylogenetic analyses

Evolutionary models were selected based on the Akaike information criterion (AIC; Akaike, 1973), which was calculated with MrAIC (Nylander, 2004). For the *trn*T–F dataset, the general time-reversible substitution model (GTR; Tavaré, 1986) with gamma-distributed rate variation among sites ($+\wp$) and a portion of sites invariable, was selected. For *rps*16, as well as for ITS, the GTR$+\wp$ model was selected, and for ETS the Hasegawa–Kishino–Yano substitution model (Hasegawa *et al.*, 1985) $+\wp$ was selected. In the analyses of cpDNA (*trn*T–F and *rps*16) and nrDNA (ITS and ETS), mixed models incorporating the models selected for each individual region by the AIC were used. In the Bayesian inference (BI) analysis of the combined dataset, the data were divided into two partitions, cpDNA and nrDNA, each evolving according to the GTR$+\wp$ model. Bayesian phylogenetic analyses were performed in MRBAYES 3.2 (Huelsenbeck & Ronquist, 2000; Ronquist & Huelsenbeck, 2003). In the BI analyses of the separate datasets, the Markov chain was run for 4 million generations, and in the combined analyses it was run for 10 million generations, sampling every 1000 generations. The Markov chain was assumed to have reached convergence when plots of the overall likelihood, as well as individual parameters of the model, were fluctuating around stable values. At this point the average standard deviation of split frequencies for two analyses run in

parallel was <0.01, and the potential scale reduction factor was 1.00 for all parameters. Phylogenetic hypotheses sampled during the first 1 million generations were discarded as 'burn-in', and the remaining trees from the two parallel analyses were used to construct a majority-rule consensus tree and calculate posterior probabilities (PPs) of clades. A maximum parsimony (MP) bootstrap analysis (Felsenstein, 1985) with 10,000 pseudo-replicates was performed with PAUP* (Swofford, 2002). Each pseudo-replicate was analysed with heuristic search, creating 10 start trees by random addition and improving these by tree bisection–reconnection (TBR) branch swapping. One tree was saved for each bootstrap replicate and used to construct a majority-rule consensus tree and calculate bootstrap proportions (BPs). All phylogenetic analyses were repeated three times to make sure that independent runs converged on similar topologies containing the same nodes with PPs ≥ 0.95 or BPs ≥ 70. Trees were rooted on *Luculia*, because this taxon has been shown to be a basal lineage in Rubiaceae (Bremer, 1996).

Molecular dating

In order to obtain estimates of lineage divergence times, we used a Bayesian approach as implemented in BEAST 1.4.8 (Drummond & Rambaut, 2007). The molecular data were analysed with a GTR substitution model with ℘-distributed rate variation, an uncorrelated relaxed lognormal clock, and a Yule tree prior with a constant speciation rate per lineage. Two calibration points, one fossil and one from a previous molecular dating analysis, were applied in the analysis. The fossil is a *Faramea* pollen from the late Eocene found in Puerto Rico (Graham & Jarzen, 1969). It has been argued that the only logically consistent way of assigning fossil calibration points to phylogenies is to use them as minimum ages of stem nodes (Renner, 2005). This conservative approach will, however, lead to a consistent underestimation of divergence times. The underestimation is caused first by the fact that fossils always will be younger than the taxa they represent, and second by their being assigned to nodes that are too old. The placement of fossil calibration points is a difficult issue because there is usually no way of knowing where in the phylogeny the fossil actually belongs. It may even represent a branch inside the crown group. In recent studies it seems to be more common to constrain the age of the stem node of a clade that shares some of the characters that are present in the fossil (e.g. Davis *et al.*, 2004; Richardson *et al.*, 2004; Pirie *et al.*, 2006), although in some studies constraints are instead placed on the age of the crown group (e.g. Zerega *et al.*, 2005; Muellner *et al.*, 2006). In this study, we use the age of the *Faramea* pollen to set a minimum age of the *Faramea* crown node. For this calibration point, a uniform prior with a lower limit of 37 Ma and an upper limit of 145 Ma was used. The latter age corresponds to the beginning of the Cretaceous, the time when the first triaperturate pollen appears (Crane *et al.*, 1995). For comparison, we also ran the analysis with the fossil calibration placed at the stem node of *Faramea*. Of the two analyses, the one

calibrated on the crown node of *Faramea* yielded an age of clade g that was closer to the estimate obtained for the same node in a dating analysis covering the entire Rubiaceae (Bremer & Eriksson, 2009). Therefore we chose to use the results from the analysis calibrated on the crown node in this study. Furthermore, the differences in estimated divergence times between the two analyses were so small compared to the confidence intervals that it did not seem important to include them both. The second calibration point is the age of the Rubiaceae crown group, which was set to 86.5 Myr based on a molecular dating analysis (Bremer & Eriksson, 2009). For this node (i.e. the tree model root height), a prior with a normal distribution and a standard deviation of 8 Myr was used. This corresponds to the 95% highest posterior density (HPD) interval (73–101 Ma) obtained in the previous analysis (Bremer & Eriksson, 2009). The Markov chain Monte Carlo (MCMC) was run for 50 million generations, sampling hypotheses every 1000 generations. The output was visualized using TRACER (Rambaut & Drummond, 2003), making sure that parameter values were fluctuating at stable levels. Based on these results, the first 5 million generations were discarded as burn-in, and the remaining samples were summarized as a maximum clade credibility tree with mean divergence times and HPD intervals of age estimates in TreeAnnotator.

Biogeographical inference

The three geographic areas used in the analyses were: tropical Africa (A), tropical and subtropical Asia (B), and central and southern tropical America (C). Because an accurate taxon sample outside the group of interest is necessary when reconstructing its ancestral area, and a Rubiaceae-wide taxon sample would make it difficult to set up a correct DEC model, we did not want to use LAGRANGE to reconstruct the ancestral area of Urophylleae (node d). Instead, we used Bayes-DIVA analysis (Nylander *et al.*, 2008), where ancestral reconstructions are averaged over all trees in a Bayesian sample to account for phylogenetic uncertainty in the biogeographical analysis, for this purpose. The DIVA program (Ronquist, 2001) was run on 10,000 trees from the MCMC output from the MrBayes analysis of the complete dataset using a perl script that automated the procedure (Nylander *et al.*, 2008).

To test which of the two hypotheses of long-distance dispersal and boreotropical migration provides the better explanation of extant range data for Urophylleae, two DEC models, DISP and BOR, respectively, were set up and compared using the computer software LAGRANGE 2.0.1 (Ree & Smith, 2008). In this program, the rate of dispersal is set to 1.0 during periods when dispersal is assumed to have been possible and to 0.0 during periods when it is thought not to have been possible (Ree & Smith, 2008). Thus, in the BOR model rates of dispersal are set to 0.0 during periods when regions are assumed to have been unconnected and to 1.0 during periods when they are assumed to have been connected (Ree & Smith, 2008). In the period from the late Palaeocene (60 Ma) until the mid-Eocene (40 Ma), when the climate in

large parts of the Northern Hemisphere is thought to have been tropical (Zachos et al., 2001) and the NALB, and possibly also the BLB, were open to migration for tropical taxa (Tiffney, 1985b; Morley, 2003), the rate of dispersal is set to 1.0 between all areas. This is not strictly correct, because the Tethys probably constituted a barrier to dispersal between Africa and Europe during this period (Morley, 2003), but because the hypothesis relies on the assumption of a continuous range in the Northern Hemisphere, spanning Europe, where there are no extant representatives of the group, this is a necessary simplification. Migration across the NALB was allowed until 32 Ma, which is the latest that a tropical taxon has been shown to use this route of migration (Davis et al., 2004). After this there was no connection between the Old and New Worlds, and the rate of dispersal to and from tropical America is set to 0.0. In the DISP model, on the other hand, dispersal is always allowed to take place between any regions. For regions that are completely unconnected there is a very low rate of dispersal, representing dispersal by chance, for example by wind or water. This rate was set to 0.001, but higher values (0.01 and 0.1) were also tested. For periods when dispersal is assumed to have taken place more easily between certain regions, the rate of dispersal was set to 1.0. This higher rate was used for dispersal from Africa to Asia before 65 Ma, when dispersal was possible via the Indian subcontinent (Morley, 2003), and between Africa and the Neotropics before 50 Ma, when it has been suggested that islands in the Atlantic facilitated dispersal (Morley, 2003).

The two models were compared using DEC, and to avoid likelihood scores being influenced by biogeographical events outside Urophylleae, trees were pruned to include only taxa from this group. First, the analysis was run with a point estimate, namely the tree with the highest likelihood score from the BEAST dating analysis. Then, to see whether biogeographical conclusions were affected by the uncertainty in estimated divergence times, the analysis was run with 1000 chronograms from the BEAST analysis, randomly sampled after burn-in. To do this we wrote two perl scripts: one to sample the trees from BEAST, remove outgroup taxa, and run the

LAGRANGE analyses; and another to compile the results from the separate analyses. Only nodes that were free from conflict between cpDNA and nrDNA trees were included in the analysis. The root node was constrained to Africa based on the result from the Bayes-DIVA analysis.

RESULTS

Data

In this study, 19 new rps16 DNA sequences, 50 new sequences from the trnT–F region, 53 from the ITS region, and 61 ETS sequences were produced. All EMBL accession numbers of sequences are shown in Table 1. The information content of each amplified DNA region is given in Table 2.

Phylogenetic relationships

Trees from the separate analyses of cpDNA (trnT–F and rps16) and nrDNA (ITS and ETS) are congruent concerning the relationships among major lineages within Urophylleae, but neither fully resolves the topology with good support. The two datasets both resolve six lineages within Urophylleae that correspond to recognized genera: Temnopteryx, Amphidasya, Raritebe, Urophyllum, Pauridiantha and Pentaloncha. Regarding species-level relationships within the genera Pauridiantha, Urophyllum and Amphidasya there are, however, supported conflicts between the datasets (not shown). In the BI analysis of the combined dataset (Fig. 2a), Temnopteryx is the sister of the rest of Urophylleae (node d, PP 0.95), but this relationship receives minimal support in the MP analysis (BP 18). A Neotropical clade (Fig. 2a, node f), consisting of the genera Amphidasya and Raritebe, and a large Palaeotropical clade (Fig. 2a, node g) are supported as sister groups. Within the Palaeotropical clade, Urophyllum (Fig. 2a, node h) is found to be the sister of a clade consisting of the African Pentaloncha and Pauridiantha (Fig. 2a, node i). Both separate and combined analyses show that several small genera are ingroups in Urophyllum and Pauridiantha, but detailed intergeneric relationships will be presented elsewhere.

Biogeography

For the group as a whole (node d), the Bayes-DIVA reconstruction of ancestral areas shows that an origin in Africa (B) is most likely (0.76), while Africa and Asia (AB) and a pantropical distribution (ABC) are equally probable (0.12). For divergences within Urophylleae, the ML reconstruction of geographic range evolution from the LAGRANGE analysis is shown in Fig. 2(b). The support for alternative solutions found among the 1000 sampled trees is shown as a pie chart for each divergence.
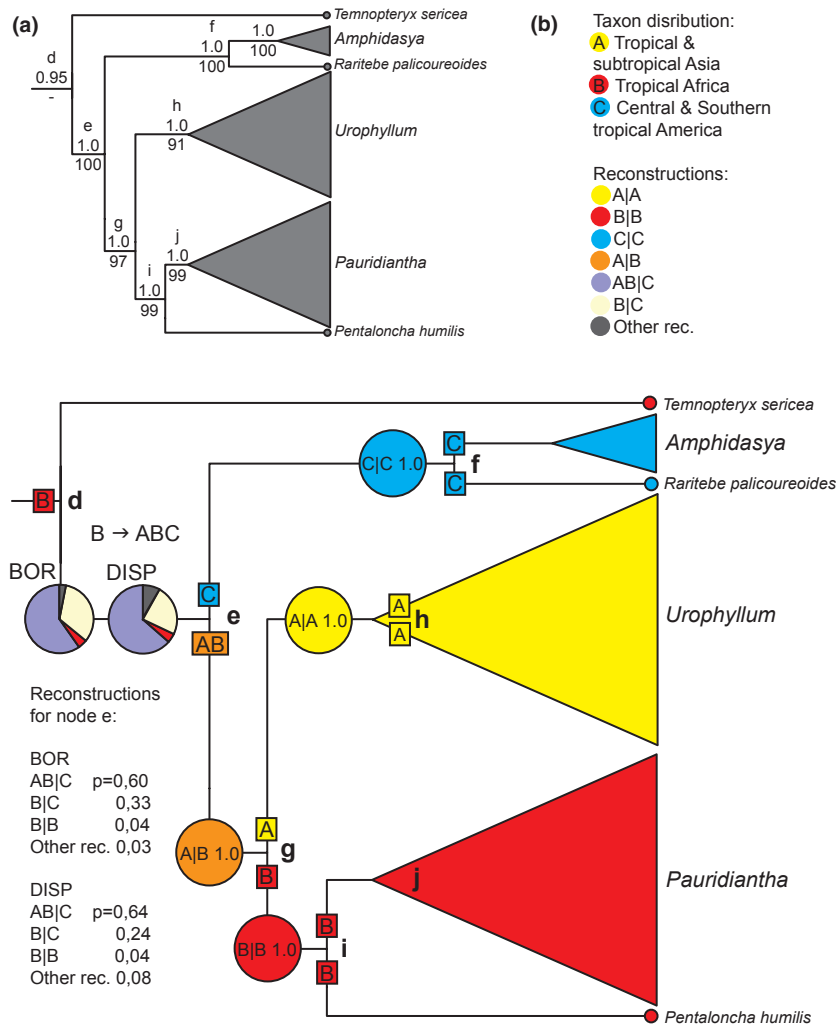
When LAGRANGE was run with the ML tree from the BEAST dating analysis, the log-likelihood score was considerably higher with the BOR model (lnL = −11.5) than with the DISP model (lnL = −20.7). The likelihood scores of the DISP model increased with higher rates of long-distance dispersal (0.01;

Table 2 Amount of phylogenetic information in each of the molecular datasets used in the study of phylogenetic relationships within Urophylleae. For each DNA region, the number of terminals, the number of aligned DNA characters, the number of variable and parsimony-informative characters, and the amount of missing data are given.

DNA region	Terminals	Aligned DNA characters	Variable characters	Parsimony-informative characters	Missing data (%)
rps16	69	1432	424	190	1.2
trnT–F	67	2706	905	380	29
ITS	66	766	408	294	5.5
ETS	62	497	393	298	3.9

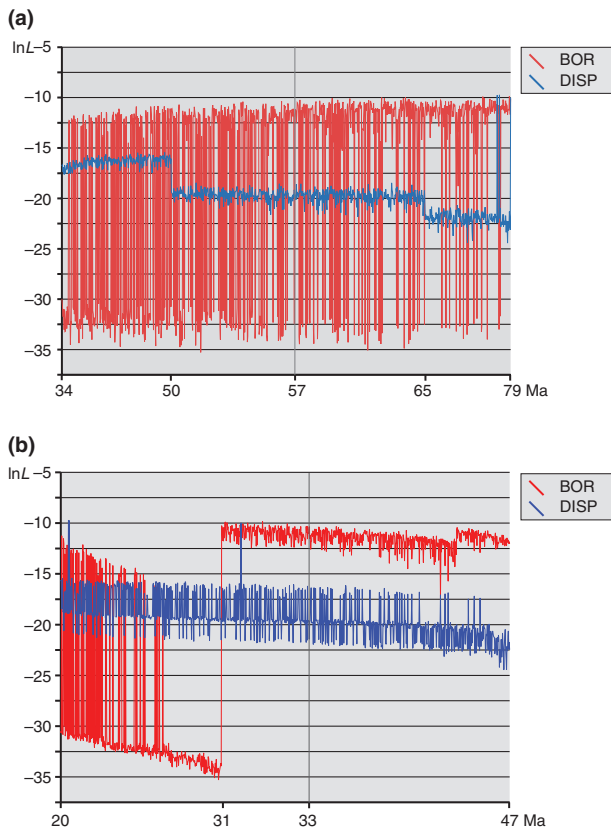ITS, internal transcribed spacer; ETS, external transcribed spacer.

Figure 2 Simplified 95% majority-rule tree of Urophylleae from the Bayesian phylogenetic analysis of combined cpDNA and nrDNA data. Nodes discussed in the text are indicated by lower-case letters. (a) Bayesian posterior probabilities are shown above branches and parsimony bootstrap proportions (BPs) below. (b) Geographic origin is indicated by a coloured circle or triangle in front of each taxon name. For each divergence, the maximum likelihood reconstruction of geographic range evolution from the LAGRANGE analysis is shown. Pie charts represent the probabilities of reconstructions. For all nodes except node e, reconstructions were identical under the two models: long-distance dispersal (DISP) and boreotropical migration (BOR).

lnL = −15.4, 0.1; lnL = −12.1), but were still lower than under the BOR model.

The results from the analyses of 1000 randomly sampled trees from the posterior distribution do, however, show that this result is not consistent. The BOR model performs better than the DISP model in 69% of the analyses. As the estimated age of the group increases, the likelihood under the DISP model gradually decreases (Fig. 3a), while under the BOR model the likelihood varies, regardless of the age of the group, between scores that are higher than under the DISP model and scores that are much lower. The estimated age of node e affects which model is performing best, such that for age estimates older than 31 Ma higher likelihood values are obtained under the BOR model than under the DISP model, while for ages younger than 31 Ma the DISP model generally performs better (Fig. 3b).

For node e, the DEC reconstruction of geographic range evolution is uncertain, but the two models give similar results (Fig. 2b). For the remaining nodes, the two models favour the same reconstructions with a probability of 1.0. Two dispersal events, one from Africa to Asia, and one from the Old World to the New World, are inferred to have taken place between nodes d and e (Fig. 2b, +BC). The divergence at node e is indicated to have been concomitant with a vicariance event between the Old World (AB) and New World (C). The descendant at node f then underwent radiation in central and southern tropical America. The split between the ancestral lineages of *Urophyllum* s. lat. and the *Pentaloncha–Pauridiantha* clade was accompanied by vicariance between Asia (A, node h) and Africa (B, node i). The descendants then diversified in tropical Africa and tropical and subtropical Asia, respectively.

Figure 3 The log likelihood (lnL) of the data under each of the two models, long-distance migration (DISP, blue) and boreotropical migration (BOR, red) plotted against (a) the age of Urophylleae (Fig. 2, node d) and (b) the age of node e (Fig. 2), for 1000 randomly sampled trees from the BEAST analysis. Note that the scale on the x-axis is not proportional because each sampled tree is plotted and sampling is denser closer to the mean. The mean of the age estimates is indicated by a vertical line in each diagram.

Molecular dating

The chronogram from the dating analysis is shown in Fig. 4, and estimated divergence times with their respective 95% HPDs for major clades are presented in Table 3. The mean age estimate for the origin of Urophylleae (Fig. 4, node d) is in the late Palaeocene (57.7 Ma), but as for all nodes the 95% HPD is large (76.6–39.2 Ma) and extends from the late Cretaceous until the mid-Eocene. The three major lineages within the group, that of the Neotropical clade (Fig. 4, node f) and those of the two Palaeotropical clades (Fig. 4, nodes h and i), are estimated to have arisen during the Oligocene (33.1 and 27.1 Ma), with 95% HPDs covering a period from the mid-Eocene until the early Miocene (44.2–19.3 Ma, Table 3). Most of the extant species diversity in the group is indicated to have evolved from about 20 Ma until the present (Fig. 4).

The two inferred dispersals (see above) are estimated to have taken place in the interval from 57.7 Ma in the late Palaeocene (node d, 95% HPD 76.6–39.2 Ma) to 33.1 Ma in the late Eocene (node e, 95% HPD 44.2–22.8 Ma). The first of the two

inferred vicariance events (see above) is estimated to have taken place at 33.1 Ma in the late Eocene (node e, 95% HPD 44.2–22.8 Ma), and the second at 27.1 Ma in the mid-Oligocene (node g, 95% HPD 35.9–19.3 Ma).
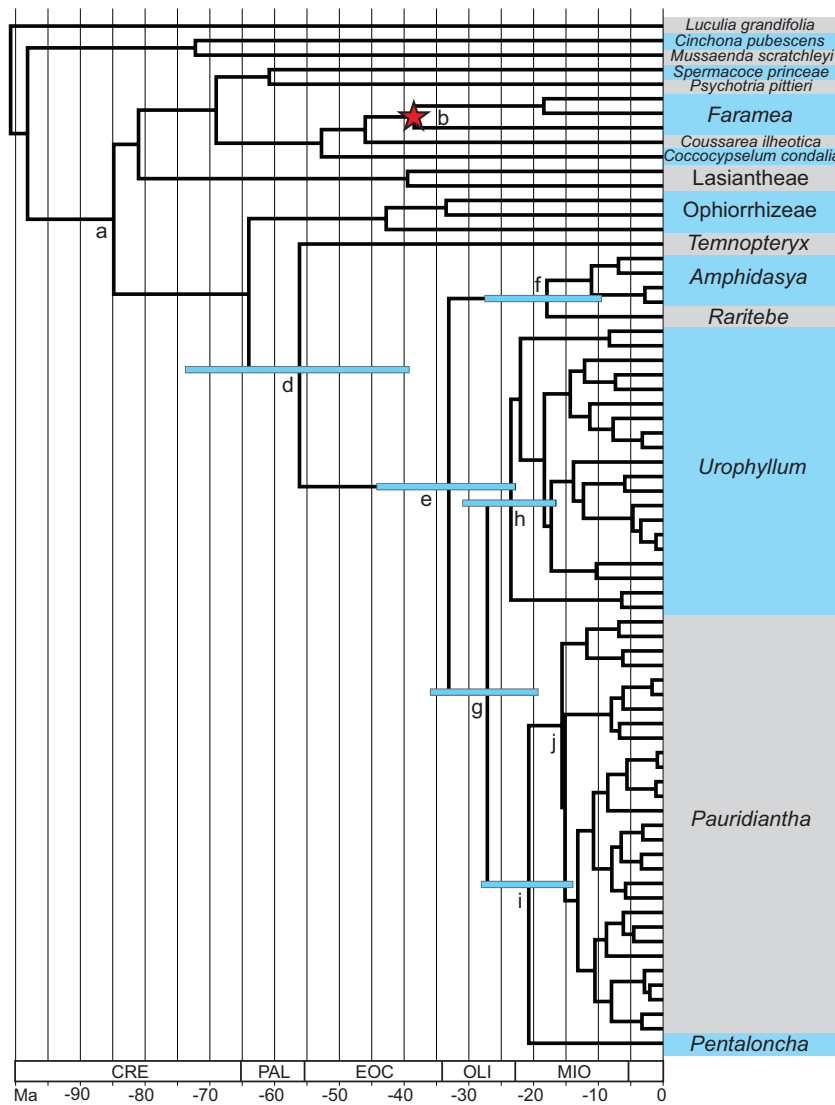
DISCUSSION

Phylogeny

There has been much uncertainty and disagreement regarding the affinities of the monotypic genus *Temnopteryx* (Bremekamp, 1952, 1966; Verdcourt, 1958; Hallé, 1966; Robbrecht, 1988; Andersson, 1996). In contrast to previous phylogenetic studies (Smedmark *et al.*, 2008; Sonké *et al.*, 2008), this study resolves the phylogenetic position of this genus with good support. Here, it is shown by BI to be the sister of the remainder of Urophylleae (PP 0.95), a topology that was reported previously but with low support by Smedmark *et al.* (2008). In that study, the support for this position did, however, increase substantially when taxa outside Rubioideae were removed. The *Temnopteryx* branch is very long in proportion to those of other species in Urophylleae (not shown). The low support by MP in this study (BP 18) is therefore likely to be a long-branch attraction phenomenon (Felsenstein, 1978), wherein *Temnopteryx* is incorrectly inferred to be more closely related to taxa outside Urophylleae that also have long branches. Another basal Rubioideae taxon that has been difficult to place phylogenetically is *Colletoecema* (Piesschaert *et al.*, 2000; Rydin *et al.*, 2008). This genus was not included in the present study, but the topology reported here does not change if *Colletoecema dewevrei* is included in the analysis (results not shown), and neither does the support for *Temnopteryx* as the sister of other Urophylleae.

This study confirms that the remainder of Urophylleae is divided into two clades (Smedmark *et al.*, 2008), one restricted to the Neotropics and one to the Palaeotropics (Fig. 2), and shows that the Palaeotropical clade is divided into the Asian *Urophyllum* s. lat. and an African clade consisting of *Pauridiantha* s. lat. and *Pentaloncha*. Support for an African and an Asian clade was also found by Sonké *et al.* (2008), although that study included only two Asian representatives.

Biogeography

When the performance of the two models of geographic range evolution was compared using the ML tree, the log-likelihood score under the BOR hypothesis was distinctly better than that under the DISP model (lnL: −11.5 > −20.7). The common practice in biogeographical inference today would be to accept this result and conclude that boreotropical migration provides a much more likely explanation of the present-day distribution patterns of Urophylleae than does long-distance dispersal. Repeating the analysis over a random tree sample from the posterior distribution did, however, show that this is not a general result (Fig. 3a). The strong fluctuation of log-likelihood scores obtained under the BOR model shows that there

Figure 4 Chronogram of Urophylleae based on a 50% majority-rule consensus tree from the Bayesian dating analysis of the combined cpDNA and nrDNA dataset. For selected nodes, 95% highest posterior density intervals are indicated by scale bars. The node where the fossil used to calibrate the tree was placed is indicated by a red star. Ages of the geological time scale (Gradstein & Ogg, 2004) are shown at the base of the figure: CRE, Cretaceous; PAL, Palaeogene; EOC, Eocene; OLI, Oligocene and MIO, Miocene. Nodes indicated by lower-case letters are discussed in the text.

is some factor, for example the topology, relative length of branches within Urophylleae, or estimated divergence times, that has a pronounced impact on the performance of this model (Fig. 3a). Exploration of the results showed that the age

Table 3 Results from the BEAST estimation of the divergence time of Urophylleae and clades therein. Uncertainty, measured as 95% highest posterior density (HPD) intervals, is also given for each age estimate.

Node	Clade	Age estimate (Ma)	95% HPD lower (Ma)	95% HPD upper (Ma)
d	Urophylleae	57.7	39.2	76.6
e	Neo- + Palaeotropical clades	33.1	22.8	44.2
f	Neotropical clade	17.9	9.5	27.5
g	Palaeotropical clade	27.1	19.3	35.9
h	*Urophyllum* s. lat.	23.5	16.5	30.9
i	African clade	20.7	13.9	28.1

of node e is such a factor (Fig. 3b). For age estimates of node e older than 31 Ma, higher likelihood values are obtained under the BOR model (ln$L \approx -11.5$) than under the DISP model (ln$L \approx -20$), while for ages younger than 31 Ma the DISP model usually (in 84% of the cases) performs better (ln$L \approx -18$) than the BOR model (ln$L \approx -33$). The reason why the estimated age of node e has such an impact on likelihood scores obtained under the BOR model is that the reconstruction of geographic range evolution showed that Urophylleae probably expanded its range from the Old World to the New World along the ancestral lineage of node e (Fig. 2b), and 31 Ma is the latest that migration across the NALB is allowed to take place according to the BOR model (Tiffney & Manchester, 2001; Davis *et al.*, 2004). The mean age estimate of node e is slightly older than this (33.1 Ma, Fig. 3b), which is probably why the BOR model performs better than the DISP model in 69% of the DEC analyses.

The choice between the two models does not affect the preferred reconstruction of geographic range evolution (Fig. 2b). Under both models, Urophylleae is inferred to have

attained a wide geographic range early in the history of the group, a range that was later split by two successive instances of vicariance. The BOR model implies that the wide distribution spanned the Northern Hemisphere and that the vicariance events at nodes e and g were caused by the drastic cooling of the climate, forcing a previously continuous population in the Northern Hemisphere to be pushed back to refugia in Central America, Southeast Asia and Africa. While the ancestor at node e is implied to have existed at the same latitude as the NALB, node g probably would have occurred at lower latitudes in Eurasia, and it is reasonable to expect the divergence between the African and Southeast Asian lineages to have taken place later than that between the New and Old Worlds, which is what our results show. The more or less contemporary radiations of the Southeast Asian and African sister groups (Fig. 4, nodes h and i), beginning somewhat earlier in Southeast Asia (23.5 Ma) than in Africa (20.7 Ma), are also consistent with a boreotropical ancestry.

Latitudinal barriers in Europe (such as the Alps and the Mediterranean) and Africa (a strong latitudinal zonation of the vegetation) seem to have made it difficult for boreotropical taxa to migrate south to Africa, and boreotropical relicts are typically not found in African rain forests today (Morley, 2003). Therefore, Urophylleae, with the large African Pentaloncha–Pauridiantha s. lat. clade, is perhaps not a typical candidate for a boreotropical taxon. In the late Cretaceous and early Cenozoic, the African continent was quite isolated from other tropical land masses (Morley, 2000). The African plate came close to the Iberian peninsula in the late Eocene (39–36 Ma), and in the mid-Miocene (16–10 Ma) the Tethys closed, allowing African and Eurasian mammals to intermingle. Records of tropical plant taxa migrating from Eurasia to Africa during this whole period are, however, scarce (Morley, 2000), but it has been suggested that migration took place within several lineages of Malpighiaceae (Davis et al., 2004). The fruits of Urophyllum s. lat., Pentaloncha and Pauridiantha s. lat. are fleshy, and at least those of Urophyllum have been shown to be dispersed by birds (Ingle, 2003). One possibility is that birds, which migrate between Eurasia and Africa in large numbers, could have brought seeds.

After the closure of the Turgai Straits at the end of the Eocene, the route from western to eastern Asia was probably open for the dispersal of tropical plant taxa and may have remained open to the migration of boreotropical taxa into east Asia during the Oligocene (Morley, 2000). Our results imply that migration along this route would have taken place between nodes g and h (Fig. 4, 27.1–23.5 Ma, or 35.9–16.5 Ma including the 95% HPDs), which partly coincides with the period of favourable conditions.

If Urophylleae had a boreotropical distribution in the Eocene this could provide an explanation for the huge time span of 40 Myr following the origin of the group when the phylogeny indicates a very depauperate group in terms of species richness (Fig. 4). The group could have been much more species-rich during the Eocene and Oligocene than indicated by the phylogeny (Fig. 4), but a lot of that diversity

may have been lost by extinction owing to the cooling climate in the Northern Hemisphere. There are no known Urophylleae fossils from the Northern Hemisphere, but three different types of fossil pollen from other Rubiaceae taxa have been found in the Palaeogene of the Paris basin (Morley, 2000), showing that at least some Rubiaceae taxa were part of the boreotropical flora.

If the group attained its present distribution by long-distance dispersal there is no obvious explanation for the inferred instances of vicariance (Fig. 2b). Had the range expansions instead been inferred to have taken place later, namely to America between nodes e and f, and to Asia between nodes g and h (Fig. 2b), long-distance dispersal would have seemed a likely explanation. The reason why biogeographical events are reconstructed the way they are under the DISP model probably has to do with details of the model. For example, the rate of dispersal between Africa and the Neotropics was set to have been high before 50 Ma, when it has been suggested that dispersal was facilitated by islands in the Atlantic (Morley, 2003), and after this the rate was set to have been low. This will strongly favour dispersal from Africa to the Neotropics along the branch between nodes d and e, because the age estimate of node d is 57.7 Ma (95% HPD 76.6–39.2 Ma) and the lower limit of the 95% HPD of node e is 44.2 Ma (Table 2).

The rate of dispersal from Africa to Asia of the DISP model was set to have been high before 65 Ma, when it has been suggested that dispersal was possible from Africa to the Indian subcontinent via the Malagasy/Mascarene islands (Morley, 2003). This should, however, be too early to have had any major impact on the reconstruction, because the mean age estimate of the origin of the group (node d) is 57.7 Ma and the lower limit of the 95% HPD is 76.6 Ma. An aspect of the results that contradicts rafting on the Indian subcontinent is the fact that the Indian plate collided with the Asian plate in the middle Eocene (50–39 Ma), and it is estimated that the radiation in Southeast Asia did not begin until 23.5 Ma, at the end of the Oligocene (node h, 95% HPD 30.9–16.5 Ma). Moreover, long-distance dispersal between Africa and Asia seems less likely because Temnopteryx, Pentaloncha and the two unresolved basal clades within clade j (Fig. 2) all occur along the west coast of Africa, from Sierra Leone in the north to Congo in the south. This would imply that dispersal had taken place from the west coast of Africa to Southeast Asia, which seems less probable than dispersal from the east coast of Africa.

CONCLUSIONS

Based on the results presented in this study we conclude that boreotropical migration provides a somewhat better explanation of the disjunct pantropical distribution patterns of extant Urophylleae than long-distance dispersal, and that the inconclusive result is due to the uncertainty in estimated divergence times. This study also shows clearly that results from biogeographical analyses based on single phylogenetic trees, such as an ML or consensus tree, can be misleading, and that it may be

very important to take the uncertainty in age estimates into account. Confidence intervals of estimated divergence times obtained in molecular dating analyses are usually large, and age estimates of individual nodes may be very important for the conclusions of historical biogeography. The problem is probably more pronounced when no calibration point is present within the group under study, which often is the case in published studies. Methods that account for the uncertainty in topology, branch lengths and estimated divergence times are not commonly used in biogeographical inference today but should definitely be preferred in order to avoid unwarranted conclusions.

ACKNOWLEDGEMENTS

REFERENCES

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

Andersson, L. (1996) Circumscription of the tribe Isertieae (Rubiaceae). *Opera Botanica Belgica*, **7,** 139–164.

Andersson, L. & Antonelli, A. (2005) Phylogeny of the tribe Cinchoneae (Rubiaceae), its position in Cinchonoideae, and description of a new genus, *Ciliosemina*. *Taxon*, **54,** 17–28.

Andersson, L. & Rova, J.H.E. (1999) The *rps*16 intron and the phylogeny of the Rubioideae (Rubiaceae). *Plant Systematics and Evolution*, **214,** 161–186.

Antonelli, A., Nylander, J.A.A., Persson, C. & Sanmartín, I. (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences USA*, **106,** 1–6.

Backlund, M., Bremer, B. & Thulin, M. (2007) Paraphyly of Paederieae, recognition of Putorieae and expansion of *Plocama* (Rubiaceae–Rubioideae). *Taxon*, **56,** 315–328.

Baldwin, B.G. & Markos, S. (1998) Phylogenetic utility of the external transcribed spacer (ETS) of 18S–26S rDNA: congruence of ETS and ITS trees in *Calycadenia* (Compositeae). *Molecular Phylogenetics and Evolution*, **10,** 449–463.

Bremekamp, C.E.B. (1952) The African species of *Oldenlandia* L. sensu Hiern et K. Schumann. *Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, afd. Natuurkunde_Ser. 2*, **18,** 1–297.

Bremekamp, C.E.B. (1966) Remarks on the position, the delimitation and the subdivision of the Rubiaceae. *Acta Botanica Neerlandica*, **15,** 1–33.

Bremer, B. (1996) Phylogenetic studies within Rubiaceae and relationships with other families based on molecular data. *Opera Botanica Belgica*, **7,** 33–50.

Bremer, B. & Eriksson, T. (2009) Timetree of Rubiaceae – phylogeny and dating the family, subfamilies and tribes. *International Journal of Plant Sciences*, **170,** 766–793.

Bremer, B. & Thulin, M. (1998) Collapse of Isertieae, re-establishment of Mussaendeae, and a new genus of Sabiceeae (Rubiaceae); phylogenetic relationships based on *rbc*L data. *Plant Systematics and Evolution*, **211,** 71–92.

Crane, P.R., Friis, E.M. & Pedersen, K.R. (1995) The origin and diversification of angiosperms. *Nature*, **374,** 27–33.

Davis, C.C., Fritsch, P.W., Bell, C.D. & Mathews, S. (2004) High-latitude Tertiary migrations of an exclusively tropical clade: evidence from Malpighiaceae. *International Journal of Plant Sciences*, **165,** S107–S121.

Doyle, J.J. & Doyle, J.L. (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12,** 13–15.

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

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

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

Govaerts, R. (2009) *World checklist of Rubiaceae*. The Board of Trustees of the Royal Botanic Gardens, Kew. Available at: http://www.kew.org/wcsp/ (accessed 31 August 2009).

Gradstein, M.N. & Ogg, J.G. (2004) Geologic time scale 2004 – why, how, and where next! *Lethaia*, **37,** 175–181.

Graham, A. & Jarzen, D.M. (1969) Studies in neotropical paleobotany. I. The Oligocene communities of Puerto Rico. *Annals of the Missouri Botanical Garden*, **56,** 308–357.

Green, P. (1999) *PHRAP, version 0.990329*. Available at: http://www.phrap.org.

Green, P. & Ewing, B. (2002) *PHRED, version 0.020425c*. Available at: http://www.phrap.org.

Hallé, N. (1966) Famille des Rubiacées (1re partie). *Flore du Gabon* (ed. by A. Aubréville), pp. 1–278. Muséum National d'Histoire Naturelle, Laboratoire de Phanérogamie, Paris.

Hasegawa, M., Kishino, H. & Yano, T. (1985) Dating the human–ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22,** 160–174.

Huelsenbeck, J.P. & Ronquist, F. (2000) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17,** 754–755.

Ingle, N.R. (2003) Seed dispersal by wind, birds, and bats between Philippine montane rainforest and successional vegetation. *Oecologia*, **134,** 251–261.

Lo Presti, R.M. & Oberprieler, C. (2009) Evolutionary history, biogeography and eco-climatological differentiation of the genus *Anthemis* L. (Compositae, Anthemideae) in the circum-Mediterranean area. *Journal of Biogeography*, **36,** 1313–1332.

McDill, J., Repplinger, M., Simpson, B.B. & Kadereit, J.W. (2009) The phylogeny of *Linum* and Linaceae subfamily Linoideae, with implications for their systematics, biogeography, and evolution of heterostyly. *Systematic Botany*, **34**, 386–405.

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

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

Muellner, A.N., Savolainen, V., Samuel, R. & Chase, M.W. (2006) The mahogany family 'out-of-Africa': divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. *Molecular Phylogenetics and Evolution*, **40**, 236–250.

Negrón-Ortiz, V. & Watson, L.E. (2002) Molecular phylogeny and biogeography of *Erithalis* (Rubiaceae), an endemic of the Caribbean Basin. *Plant Systematics and Evolution*, **234**, 71–83.

Nepokroeff, M., Bremer, B. & Sytsma, K.J. (1999) Reorganization of the genus *Psychotria* and tribe Psychotrieae (Rubiaceae) inferred from ITS and *rbcL* sequence data. *Systematic Botany*, **24**, 5–27.

Novotny, V., Basset, Y., Miller, S.E., Weiblen, G.D., Bremer, B., Cizek, L. & Drozd, P. (2002) Low host specificity of herbivorous insects in a tropical forest. *Nature*, **416**, 841–844.

Nylander, J.A.A. (2004) *MrAIC.pl, version 1.4*. Available at: http://www.abc.se/~nylander/.

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

Oxelman, B. (1996) RAPD patterns, nrDNA ITS sequences and morphological patterns in *Silene* section Sedoideae (Caryophyllaceae). *Plant Systematics and Evolution*, **201**, 93–116.

Oxelman, B., Lidén, M. & Berglund, D. (1997) Chloroplast *rps16* intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution*, **184**, 393–410.

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.

Piesschaert, F., Andersson, L., Jansen, S., Dessein, S., Robbrecht, E. & Smets, E. (2000) Searching for the taxonomic position of the African genus *Colletoecema* (Rubiaceae): morphology and anatomy compared to an *rps16*-intron analysis of the Rubioideae. *Canadian Journal of Botany*, **78**, 288–304.

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

Popp, M. & Oxelman, B. (2001) Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **20**, 474–481.

Rambaut, A. (1996) *Se-Al: Sequence Alignment Editor, version 2.0a11*. Available at: http://evolve.zoo.ox.ac.uk/.

Rambaut, A. & Drummond, A.J. (2003) *Tracer, version 1.2*. Available at: http://evolve.zoo.ox.ac.uk/.

Razafimandimbison, S.G. & Bremer, B. (2002) Phylogeny and classification of Naucleeae s.l. (Rubiaceae) inferred from molecular (ITS, *rbcL*, and *trnT-F*) and morphological data. *American Journal of Botany*, **89**, 1027–1041.

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.

Renner, S.S. (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, **165**(Suppl.)**,** 23–33.

Renner, S.S. (2005) Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science*, **10**, 550–558.

Renner, S.S., Clausing, G. & Meyer, K. (2001) Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany*, **88**, 1290–1300.

Richardson, J.E., Chatrou, L.W., Mols, J.B., Erkens, R.H.J. & Pirie, M.D. (2004) Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 1495–1508.

Robbrecht, E. (1988) Tropical woody Rubiaceae. *Opera Botanica Belgica*, **1**, 1–271.

Ronquist, F. (2001) *DIVA, version 1.2*. Available at: http://www.ebc.uu.se/systzoo/research/diva/diva.html.

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

Rova, J.H.E., Delprete, P.G., Andersson, L. & Albert, V.A. (2002) A *trnL-F* cpDNA sequence study of the Condamineeae-Rondeletieae-Sipaneeae complex with implications on the phylogeny of the Rubiaceae. *American Journal of Botany*, **89**, 145–159.

Rydin, C., Razafimandimbison, S.G. & Bremer, B. (2008) Rare and enigmatic genera (*Dunnia*, *Schizocolea*, *Colletoecema*), sisters to species-rich clades: phylogeny and aspects of conservation biology in the coffee family. *Molecular Phylogenetics and Evolution*, **48**, 74–83.

Rydin, C., Kainulainen, K., Razafimandimbison, S.G., Smedmark, J.E.E. & Bremer, B. (2009) Deep divergences in the coffee family and the systematic position of *Acranthera*. *Plant Systematics and Evolution*, **278**, 101–123.

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

Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, **19**, 101–109.

Sanderson, M.J. & Doyle, J.A. (2001) Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *American Journal of Botany*, **88**, 1499–1516.

Smedmark, J.E.E., Rydin, C., Razafimandimbison, S.G., Khan, S.A., Liede-Schumann, S. & Bremer, B. (2008) A phylogeny of Urophylleae (Rubiaceae) based on *rps16* intron data. *Taxon*, **57**, 1–9.

Sonké, B., Dessein, S., Taedooumg, H., Groeninckx, I. & Robbrecht, E. (2008) A new species of *Colletoecema* (Rubiaceae) from southern Cameroon with a discussion of relationships among basal Rubioideae. *Blumea*, **53**, 533–547.

Staden, R., Beal, K.F. & Bonfield, J.K. (1998) The Staden Package. *Bioinformatics: methods and protocols* (ed. by S. Misener and S.A. Krawetz), pp. 115–130. Computer Methods in Molecular Biology, Vol. 132. The Humana Press Inc., Totowa, NJ.

Swofford, D.L. (2002) *PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b10*. Sinauer, Sunderland, MA.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991) Universal primers for amplification of 3 noncoding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Tavaré, S. (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Some mathematical questions in biology – DNA sequence analysis* (ed. by R.M. Miura), pp. 57–86. American Mathematical Society, Providence, RI.

Thiers, B. (2009) *Index Herbariorum: a global directory of public herbaria and associated staff*. New York Botanical Garden's Virtual Herbarium. Available at: http://sweetgum.nybg.org/ih/ (accessed 31 August 2009).

Thorne, J.L. & Kishino, H. (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology*, **51**, 689–702.

Thorne, R.F. (1973) Floristic relationships between tropical Africa and tropical America. *Tropical forest ecosystems in Africa and South America: a comparative review* (ed. by B.J. Meggers and W.D. Duckworth), pp. 27–47. Smithsonian Institution, Washington, DC.

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

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

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

Verdcourt, B. (1958) Remarks on the classification of the Rubiaceae. *Bulletin du Jardin Botanique de l'État à Bruxelles*, **28**, 209–290.

Weeks, A.D., Caly, 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.

White, T.J., Bruns, T., Lee, S. & Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* (ed. by M.A. Innis, D.H. Gelfand, J.J. Sninsky and T.J. White), pp. 315–322. Academic Press, London.

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

Wolfe, J.A. (1978) The paleobotanical interpretation of Tertiary climates in the Northern Hemisphere. *American Scientist*, **66**, 694–703.

Wolfe, J.A. (1980) Tertiary climates and floristic relationships at high latitudes in the northern hemisphere. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **30**, 313–323.

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

Zerega, N.J.C., Clement, W.L., Datwyler, S.H. & Weiblen, G.D. (2005) Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution*, **37**, 402–416.

BIOSKETCHES

Jenny Smedmark is a post-doctoral researcher interested in flowering plant phylogeny, especially reticulate evolution. She is currently focusing on the phylogeny and historical biogeography of the two clades Urophylleae and Lasiantheae in the coffee family.

Torsten Eriksson is a researcher and bioinformatics officer at the Bergius Foundation. His current research interests are uncovering allopolyploidy in the Rosaceae, focusing on low-copy gene phylogenies, and divergence time studies of major clades such as the Rubiaceae.

Birgitta Bremer is a professor at the Royal Swedish Academy of Sciences and the director of the Bergius Botanic Garden in Stockholm. She is interested in flowering plant phylogeny with a particular focus on the asterids and the coffee family (Rubiaceae).

Editor: Malte Ebach