Botanical Journal of the Linnean Society, 2017, **185**, 1–26. With 5 figures.

Phylogenetic analyses of plastid DNA suggest a different interpretation of morphological evolution than those used as the basis for previous classifications of Dipterocarpaceae (Malvales)

JACQUELINE HECKENHAUER[1]*, ROSABELLE SAMUEL[1], PETER S. ASHTON[2], BARBARA TURNER[3], MICHAEL H. J. BARFUSS[1], TAE-SOO JANG[1], EVA M. TEMSCH[1], JAMIE MCCANN[1], KAMARIAH ABU SALIM[4], A. M. ACHALA S. ATTANAYAKE[5] and MARK W. CHASE[6,7]

[1]*Department of Botany and Biodiversity Research, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria*
[2]*Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA, USA*
[3]*Department of Integrative Biology and Biodiversity Research, Institute for Botany, University of Natural Resources and Life Science, Gregor-Mendel-Straße 33, 1180 Vienna, Austria*
[4]*Environmental and Life Sciences, Faculty of Science, University of Brunei Darussalam, Tungku Link Road, BE 1410, Brunei Darussalam*
[5]*Royal Botanical Gardens, Peradeniya, Sri Lanka*
[6]*Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond TW9 3DS, UK*
[7]*School of Biological Sciences, The University of Western Australia, Crawley, Western Australia, 6009, Australia*

Received 5 January 2017; revised 30 May 2017; accepted for publication 17 June 2017

Phylogenetic and molecular clock analyses were performed including all genera except one (*Pseudomonotes*) for the three subfamilies of Dipterocarpaceae. We also included representatives of Sarcolaenaceae and Cistaceae with Bixaceae as the ultimate outgroup. Three plastid regions (six markers), partial *rbcL*, *trnK-matK-trnK* (partial *trnK* intron including complete *matK*) and *trnT-trnL-trnF* (partial *trnT*, complete *trnT-trnL* intergenic spacer, complete *trnL*, complete *trnL-trnF* intergenic spacer and partial *trnF*), were analysed. We also investigated additional accessions for genome size and chromosome numbers. Our phylogenetic results differ in three important respects from previous interpretations of morphological characters, as reflected in recent classifications. First, our analyses strongly support assignment of *Pakaraimaea* (subfamily Pakaraimaeoideae) to Cistaceae. Second, the morphological concepts of Dipterocarpeae and Shoreeae in subfamily Dipterocarpoideae are not supported because *Dipterocarpus* is sister to *Dryobalanops* plus tribe Shoreeae. Our analysis revealed four clades: (1) *Dipterocarpus*; (2) *Dryobalanops*, for which tribal assignment has been contentious; (3) genera of Shoreeae; and (4) the remaining genera of Dipterocarpeae. Third, *Shorea* is not monophyletic. Monotoideae are weakly supported as sister to Dipterocarpoideae; Sarcolaenaceae (endemic to Madagascar) are sister to this pair. Divergence in extant Dipterocarpoideae occurred *c.* 55 Mya. Genome sizes for all accessions examined are small (0.3264–0.6724 pg), and the additional chromosome numbers we collected fit into the patterns previously observed for Dipterocarpaceae.

ADDITIONAL KEYWORDS: Cistaceae – chromosome numbers – Dipterocarpoideae – genome size – Monotoideae – Pakaraimaeoideae – Sarcolaenaceae.

INTRODUCTION

Dipterocarpaceae comprise > 500 species and have usually been considered to include three subfamilies (Maury-Lechon & Curtet, 1998), Monotoideae with three genera (30 species), monospecific Pakaraimaeoideae and Dipterocarpoideae (470 species), with nine to 19 genera depending on the author (Table 1). Their distribution is pantropical, with Monotoideae (Gilg, 1925) in Africa, Madagascar and the Colombian Amazon, Pakaraimaeoideae (Maguire &

*Corresponding author. E-mail: jacqueline.heckenhauer@univie.ac.at.

Table 1. Comparative classifications of Dipterocarpaceae according to different authors after Maury-Lechon & Curtet (1998)

Authors	Genera	Section (s.)/subgenus (s.g.)	Subsection (s.s)/subgroup (s.gr.)
Ashton (1964, 1968, 1980, 1982)	*Hopea**	s. *Hopea*	s.s. *Hopea*
			s.s *Pierra*
		s. *Dryobalanoides*	s.s. *Dryobalanoides*
			s.s. *Sphaerocarpae*
	*Neobalanocarpus**	–	
	*Shorea**	s. *Shorea*	s.s. *Shorea*
			s.s. *Barbata*
		s. *Richetioides*	s.s. *Richetioides*
			s.s. *Polyandrae*
		s. *Anthoshorea*	
		s. *Mutica*	s.s. *Mutica*
			s.s. *Auriculatae*
		s. *Ovalis*	–
		s. *Neohopea*	–
		s. *Rubella*	–
		s. *Brachypterae*	s.s. *Brachypterae*
			s.s. *Smithiana*
		s. *Pachycarpae*	–
		s. *Doona*	–
		s. *Pentacme*	–
	*Parashorea**	–	
	*Dryobalanop*s*	–	
	*Dipterocarpus**	–	
	*Anisoptera**	s. *Anisoptera*	–
		s. *Glabrae*	–
	*Upuna**	–	
	*Cotylelobium**	–	
	*Vatica**	s. *Sunaptea*	–
		s. *Vatica*	–
		(s. *Pachynocarpus*, 1964)	
	*Stemonoporus**	–	
	*Vateria**	–	
	*Vateriopsis**	–	
	*Marquesia***	–	
	*Monotes***	–	
	*Pakaraimaea****	–	
Meijer & Wood (1964, 1976), Meijer (1979)	*Hopea*	–	
	Shorea	s.g. *Euchorea* = *Shorea*	–
		s.g. *Richetia*	–
		s.g. *Anthoshorea*	–
		s.g. *Rubroshorea*	s.gr. *Parvifolia*
			s.gr. *Ovalis*
			s.gr. *Pauciflora*
			s.gr. *Smithiana*
			s.gr. *Pinanga*
	Parashorea	–	
	Dryobalanops	–	
	Dipterocarpus	–	
	Anisoptera	s. *Pilosa*	–

Table 1. *Continued*

Authors	Genera	Section (s.)/subgenus (s.g.)	Subsection (s.s)/subgroup (s.gr.)
		s. *Glabrae*	–
	Upuna	–	
	Cotylelobium	–	
	Vatica	s.g. *Synaptea*	–
		s.g. *Isauxis*	–
		s.g. *Pachynocarpus*	–
Maury (1978), Maury-Lechon (1979a, b)	*Hopea*	s. *Hopea*	s.s. *Hopea*
			s.s. *Pierra*
		s. *Dryobalanoides*	s.s. *Dryobalanoides*
			s.s. *Sphaerocarpae*
	Balanocarpus heimii	–	
	Shorea	s. *Shoreae*	–
		s. *Barbatae*	–
	Richetia	s. *Richetioides*	–
		s. *Maximae*	–
	Anthoshorea	–	
	Rubroshorea	s. *Mutica*	s.s. *Mutica*
			s.s. *Auriculatae*
		s. *Ovalis*	–
		s. *Neohopea*	–
		s. *Rubella*	–
		s. *Brachypterae*	s.s. *Brachypterae*
			s.s. *Smithianeae*
		s. *Pachycarpa*	–
	Doona	–	
	Pentacme	–	
	Parashorea	–	
	Dryobalanops	–	
	Dipterocarpus	–	
	Anisoptera	s. *Anisoptera*	–
		s. *Glabrae*	–
	Upuna	–	
	Cotylelobium	–	
	Sunaptea	–	
	Vatica	s. *Vatica*	–
		s. *Pachynocarpus*	–
	Stemonoporus	–	
	Vateria	–	
	Vateriopsis	–	

–, no further classification; *, subfamily Dipterocarpoideae; **, subfamily Monotoideae; ***, subfamily Pakaraimaeoideae.

Ashton, 1977) in the Guianan Highlands of South America and Dipterocarpoideae in the Seychelles, Sri Lanka, India and Southeast Asia to New Guinea. The last have their greatest diversity in Borneo, where they dominate the canopy of lowland forests (Ashton, 1988).

Ashton (2003) defined Dipterocarpaceae by their diversity of epidermal hairs, especially fascicled hair tufts (a malvalean character), spiral or alternate geniculate entire penninerved leaves with paired stipules and mainly paniculate or racemose inflorescences with paired bracteoles. The bisexual actinomorphic scented flowers are pentamerous with an imbricate perianth and have a persistent calyx with the sepals becoming aliform in fruit. The petals have unicellular hairs outside. The stamens are centrifugally arranged with basifixed (Dipterocarpoideae) or versatile (Monotoideae, Pakaraimaeoideae) anthers that are two-celled and generally latrorse. The anthers have (two–) four pollen sacs with more or less prominent connectival appendages. The superior ovary has three (–five) locules, each locule with two (–four) axile anatropous ovules. Ovules

are bitegmatic, with a ventral raphe and a superior micropyle, and only one survives as a viable seed. The indehiscent fruit has a woody pericarp splitting irregularly or along three sutures with persistent sepals. The embryo sac development is of the *Polygonum* type, and endosperm is of the nuclear type. The ripe seeds generally lack endosperm. The cotyledons are generally unequal, one more or less enclosing the other, laminar or fleshy, entire or lobed enclosing the radical. Ashton regarded the presence of many stamens and ovules, the pentaloculate ovary and loculicidally dehiscent pericarp in some taxa to be primitive generalized traits in the family. Dipterocarpaceae are ectotrophic and mycorrhizal (Malloch, Pirozynski & Raven, 1980; Smits, 1994; Tedersoo *et al.*, 2007; Brearley, 2012; Phosri *et al.*, 2012; Sato, Tanabe & Toju, 2015); their seeds lack dormancy.

Although the phylogenetic assignment of Dipterocarpaceae among angiosperms has previously been problematic, Ashton (1982) supported their placement in the order Malvales, a position formally accepted by the Angiosperm Phylogeny Group (APG) (1998, 2003, 2009, 2016). Ashton recognized similarities with Tiliaceae and also cited Sarcolaenaceae as tropical evergreen canopy trees with compatible biogeography. Fascicled hairs, stipules, floral characters and loculicidal capsules of Dipterocarpaceae are shared with many Malvales (Kubitzki & Chase, 2003). More morphological characters are given in Table 2; a review of these and further characters is provided in Maury-Lechon & Curtet (1998). Vestured pits are shared by some Dipterocarpaceae and Cistaceae (Arrington & Kubitzki, 2003). A distinct 'bixoid' chalazal region of the seed coat is shared by Monotoideae and Pakaraimaeoideae with Cistaceae and Bixaceae (including *Cochlospermum* Kunth; Nandi, 1998).

Dayanandan *et al.* (1999) concluded based on molecular evidence that Dipterocarpaceae, including *Monotes* A.DC. (Monotoideae) and *Pakaraimaea* Maguire & P.S.Ashton (Pakaraimaeoideae), form a clade closely related to Sarcolaenaceae, but they did not include enough outgroup genera (e.g. Cistaceae) to make any conclusive assessment of interfamilial relationships, leaving the positions of Monotoideae and Pakaraimaeoideae under discussion. According to the recent APG IV classification (2016), *Pakaraimaea* should be considered a member of an expanded Cistaceae based on the plastid *rbcL* analysis of Ducousso *et al.* (2004), in which *Pakaraimaea* was sister (with 88% bootstrap) to the two genera of Cistaceae included in that study. *Monotes* and *Pseudomonotes* A.C.Londoño, E.Alvarez & Forero (Monotoideae) were moderately supported (88%) as sister to Sarcolaenaceae plus Dipterocarpoideae, with Sarcolaenaceae weakly supported (62%) as sister to Dipterocarpoideae. A recent molecular phylogenetic study of Sarcolaenaceae that included several genera of Cistaceae and Dipterocarpaceae raised questions about the monophyly of Dipterocarpaceae with respect to Sarcolaenaceae (Aubriot *et al.*, 2016). Several other molecular phylogenetic studies have been conducted on Dipterocarpaceae, including use of PCR-RFLP (Tsumura *et al.*, 1996; Indrioko, Gailing & Finkeldey, 2006), RAPD (Rath *et al.*, 1998), AFLPs (Cao *et al.*, 2006), other plastid sequences (Kajita *et al.*, 1998; Kamiya *et al.*, 1998; Dayanandan *et al.*, 1999; Gamage *et al.*, 2003, 2006; Yulita, Bayer & West, 2005; Choong *et al.*, 2008; Tsumura *et al.*, 2011; Yulita, 2013), the nuclear gene *Pgi*C (Kamiya *et al.*, 2005; Choong *et al.*, 2008) and internal transcribed spacer regions (Yulita *et al.*, 2005). These studies have used only one to three plastid or nuclear markers (e.g. Kamiya *et al.*, 1998: *trnL* intron and intergenic spacer between *trnL* and *trnF*; Dayanandan *et al.*, 1999: *rbcL*; Gamage *et al.*, 2003: *trnL-trnF* spacer and *trnL* intron region; Gamage *et al.*, 2006: *matK*, *trnL* intron and *trnL-trnF* intergenic spacer region), only included a limited number of taxa (e.g. Kajita *et al.*, 1998: 17 species; Rath *et al.*, 1998: 12 species; Tsumura *et al.*, 1996: 30 species; Dayanandan *et al.*, 1999: 35 species, Choong *et al.*, 2008: 30 species) or did not include all three subfamilies.

Reconciliation of discordant intuitively constructed morphological classifications and molecular phylogenetics in some cases has presented problems (e.g. sectional classifications in *Leontodon* L., Asteraceae, Samuel *et al.*, 2003; *Diospyros* L., Ebenaceae, Duangjai *et al.*, 2009; *Polystachya* Hook., Orchidaceae, Russell *et al.*, 2010). Molecular phylogenetic studies have paved the way to reclassifications at tribal level in Rubioideae (Bremer & Manen, 2000) and Orchidaceae (Chase *et al.*, 2015). A taxonomic revision of Bromeliaceae subfamily Tillandsioideae was based on molecular phylogenetics of plastid and nuclear markers and new or re-evaluated morphology, which enabled circumscription of monophyletic units using synapomorphic combination of diagnostic morphological characters (Barfuss *et al.*, 2016). In general, traditional classifications have been based on a few characters intuitively selected by a well-informed specialist, and these classifications have typically excluded other generally conflicting characters; these classifications generally cannot be reproduced with a formal cladistic analysis of these data for the same group of organisms. For example, molecular phylogenetic results for the angiosperms (e.g. Chase *et al.*, 1993) appeared to be in conflict with previous 'morphological' systems (e.g. Cronquist, 1981). However, it became clear that when a formal non-molecular cladistic analysis was performed (Nandi, Chase & Endress, 1998), the conflict was not between morphology and molecules, but rather between an intuitive interpretation of a few characters and a formal objective analysis of a broader set of data. In general, such intuitive

Table 2. Distinctive morphological characters of Cistaceae, Sarcolaenaceae and Dipterocarpaceae according to Ashton (2003), Maury-Lechon & Curtet (1998) and Watson & Dallwitz (http://delta-intkey.com/angio/www/cistacea.htm, accessed 14 July 2017)

Character	Cistaceae	Sarcolaenaceae	Dipterocarpoideae	Monotoideae	Pakaraimaeoideae
Inflorescence					
paniculate			+	+	
racemi-paniculate		+	(+)	+	+
cyme	+		(+)		
Perianth pentamerous	x	x	+	+	+
Flower bud sepals					
imbricate	+	+	+		+
valvate			+	+	
Leaves					
alternate	x	+	+	+	+
opposite	x				
Stipules	+	+	x	x	x
One- or two-layered hypodermis		+	x		+
Contorted corolla	x	+	x	+	+
Two-celled anthers generally dehiscing longitudinally	+		x	+	+
Subversatile anthers			+	+	+
Imbricate perianth with unequal persistent sepals					
two smaller sepals: outer	+	+			
two smaller sepals: inner			+	+	+
Mucilage canals and cells in epidermis	+	+	+		+
Fruit					
capsular	+	+	+	+	+
nut		+	+	+	
dehiscent	+	+		+	+
indehiscent		+	+	+	

+, present; x, present and other possibilities; in parentheses, exceptions.

classifications have been re-interpreted in the face of consistent, well-supported, 'conflicting' results of molecular analyses (e.g. the intuitive interpretation of morphological data upon which these classifications have been based is discarded), generally leading to the conclusion that morphological evolution has been more complicated than previously assumed. Our intention in this study was to compare our molecular results with the previous classifications (Ashton, 1964; Meijer & Wood, 1964, 1976; Ashton, 1968, 1980, 1982; Maury, 1978; Maury-Lechon 1979a, b; Meijer, 1979) to determine to what extent they were mutally corroborative. We do not here undertake a formal analysis of morphological data, which is beyond the scope of this study.

Beside phylogenetic relationships, the age of clades is of interest so that an appropriate geographical interpretation of the evolutionary history of a group can be developed. The three subfamilies occupy different phytogeographical zones along the tropical belt of three continents with Wallace's Line as a major phytogeographical boundary in Southeast Asia (Maury-Lechon & Curtet, 1998). A Gondwanan origin, with subsequent migration to Indomalesia, was proposed by Croizat (1952, 1964) and Ashton (1982). This is supported by the significant decline in the number of species to the east of Wallace's Line. Based on an assumption that high species diversity of Dipterocarpaceae in Southeast Asia is associated with their origin, another hypothesis suggested that Dipterocarpaceae originated on the Eurasian plate with subsequent migration to South Asia, Africa and South America (Merrill, 1923; Prakash, 1972; Meher-Homji, 1979). Both hypotheses involve overland seed dispersal, which was suggested by Ashton (1982) on the basis of the limited seed dispersal capacity of these species, ectomycorrhizal symbiosis, lack of seed dormancy and salt-intolerant seeds. Morley (2000)

inferred the likely migration of Dipterocarpoideae to India/Seychelles directly from Africa, which is consistent with the presence of fossil wood identified as *Dipterocarpus* C.F.Gaertn. in East Africa in the Tertiary (Bancroft, 1935; Ashton & Gunatilleke, 1987). A phylogenetic and ectomycorrhizal study revealed that Sarcolaenaceae (endemic to Madagascar) and Dipteocarpoideae share an ectomycorrhizal common ancestor (Ducousso *et al.*, 2004). Ducousso *et al.* (2004) further suggested that the last common ancestor was located on the India–Madagascar landmass and produced the current Sarcolaenceae in southeastern Madagascar, whereas the Asian Dipterocarpaceae drifted away with the India–Seychelles landmass and then dispersed throughout Asia. Ducousso *et al.* (2004) cited Bossuyt & Milinkovitch (2001), who proposed a similar scenario for amphibians. The separation of Madagascar from the India–Seychelles block occurred 87.6 ± 0.6 Mya.

Chromosome counts are available for seven genera of Dipterocarpoideae (Rice *et al.*, 2015), which indicated the basic chromosome number in Dipterocarpeae is $x = 11$, but $x = 7$ for Shoreeae (Jong & Kaur, 1979). Most species appear to be diploid, but there are a few reports of polyploidy in *Shorea* Roxb. ex C.F.Gaertn. and *Hopea* Roxb. ranging from triploid and near triploid to tetraploid: e.g. *Hopea beccariana* Burck.: $2n = 20–22$ (Ashton, 1982) and *Shorea ovalis* (Korth.) Blume: $2n = 28$ (Kaur *et al.*, 1986). Based on published genome size measurements, most species of Dipterocarpaceae are characterized by small genomes (Ohri & Kumar, 1986; Ng *et al.*, 2016). Recently published genome size values showed a 2.64-fold difference, ranging from 0.267 pg in *Shorea hemsleyana* King ex Foxw. to 0.705 pg in *Shorea ovalis* (Ng *et al.*, 2016).

There have been morphological classifications of Dipterocarpaceae that differ with respect to numbers of genera, sections and subsections (Table 1), and the molecular studies cited above also exhibited some consistent differences in topology from those classifications. This has compromised understanding of the evolution of Dipterocarpaceae. However, the intention of this study is not to reclassify Dipterocarpaceae or to attempt a formal analysis of character evolution, but to obtain information that could help to solve some uncertainities in the current classification of this ecologically and economically important family. We address here the following topics: (1) clarification of the position of subfamilies Pakaraimaeoideae and Monotoideae; (2) phylogenetic placement of *Hopea*, *Parashorea* Kurz and *Shorea* (tribe Shoreeae) and phylogenetic relationships within *Shorea*, which comprises > 190 species; (3) placement of *Dipterocarpus*, which has been placed in Dipterocarpeae with other genera based on morphology, but showed a closer relationship to members of Shoreeae than to other members of Dipterocarpeae in previous molecular studies; (4) an examination of the position of *Dryobalanops* C.F.Gaertn. previously assigned to tribe Shoreeae by Ashton (1979) and placed in an intermediate position between Shoreeae and Dipterocarpeae by Maury-Lechon (1979a); (5) estimation of divergence times of the major clades in Dipterocarpaceae; and (6) investigation of genome size and chromosomal diversity using published as well as newly collected data.

MATERIAL AND METHODS

Plant material

Here, 238 accessions of Dipterocarpoideae representing 143 species were included. Of the 11 sections and eight subsections in the species-rich genus *Shorea* reported by Ashton (1964, 1968, 1980, 1982), nine sections and seven subsections were represented in this study. Samples were mainly collected in Brunei, Sri Lanka and Thailand. Detailed sampling locations can be found in the appendix (Supporting Information, Table S1). The sampling further comprised two accessions of the single species of *Pakaraimaea* and four species of *Marquesia* Gilg and *Monotes* in Monotoideae. This covers all described genera except for *Pseudomonotes* (Monotoideae), a notable increase in generic coverage over previous studies. Even though included in Ducousso *et al.* (2004), *Pseudomonotes*, which paired with *Monotes* (98% bootstrap), was omitted because the *rbcL* sequence used by Ducousso *et al.* (2004) was not available in GenBank. Furthermore, only sequences for which at least sequences of two of the three matrices (1) *rbcL*, (2) *trnK-matK-trnK* and (3) *trnT-trnL-trnF* were available were included in the combined analysis. Additionally, four species belonging to three genera of the closely related families Sarcolaenaceae (three genera) and three genera of Cistaceae (three species) were included. Outgroup sampling included members of Bixaceae, *Bixa orellana* L. and *Cochlospermum vitifolium* Spreng. (Supporting Information, Table S1).

DNA extraction and PCR amplification

For some accessions, sequence data were obtained from previous studies (Kajita *et al.*, 1998; Kamiya *et al.*, 1998; Gamage *et al.*, 2003, 2006: see Supporting Information, Table S1). For new accessions, DNA from the Royal Botanic Gardens, Kew, DNA Bank (apps.kew.org/dnabank/, accessed 14 July 2017) was used or genomic DNA was extracted from *c.* 20 mg of silica gel-dried (Chase & Hills, 1991) material (bark or leaves) using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. To

avoid degradation, material was frozen in liquid nitrogen and then ground to a fine powder using glassbeads. To remove mucilaginous polysaccharides, which are a problem for many members of Malvales due to the mucilaginous epidermal cells, the ground material was initially washed with sorbitol buffer (Russell *et al.*, 2010; Souza *et al.*, 2012) until there was no visible mucilage in the sample.

Three plastid regions (including six markers) were amplified: partial *rbcL*, *trnK-matK-trnK* (partial *trnK* intron including complete *matK*) and *trnT-trnL-trnF* (partial *trnT*, complete *trnT-trnL* intergenic spacer, complete *trnL*, complete *trnL-trnF* intergenic spacer and partial *trnF*), resulting in a *c.* 5.9 kb alignment. PCRs included 7.5 µL 2× Phusion Green HF HS PCR Master Mix with 1.5 mM $MgCl_2$ (Life Technologies, LT, Vienna, Austria), 0.15 µL bovine serum albumin (0.2 g/L), 1.5 µL each primer (3.2 µM), 1 µL template DNA and H_2O up to a final volume of 15 µL. The primers used in this study are provided in Table 3. Thermal cycle conditions were as follows: initial denaturation at 98 °C for 30 s, 35 cycles of denaturation at 98 °C for 10 s, annealing at 63–68 °C (depending on the primers, Table 3) for 30 s and extension at 72 °C for 30 s (*rbcL*) to 1 min (*trnK-matK-trnK, trnT-trnL-trnF*), followed by final extension of 5 min at 72 °C. PCR products were cleaned with 1.5 µL exonuclease I and FastAP thermosensitive alkaline phosphatase mixture (7 U Exo I, 0.7 U FastAP) at 37 °C for 45 min and 85 °C for 15 min. Sequencing reactions were performed with the BigDye Terminator Kit v3.1 (LT) using the same primers that were used for amplification or with internal primers (Table 3) according to the manufacturer's instructions. Sanger sequencing was carried out using a 3730 DNA analyser (LT).

Sequence alignment and phylogenetic analyses

Sequences were assembled and edited using Geneious (version 8.0.5, http://www.geneious.com; accessed 14 July 2017; Kearse *et al.*, 2012). To generate the *trnT-trnL-trnF* alignment, the partial *trnL* intron and the *trnL-trnF* accessions obtained from GenBank were combined in BioEdit v7.0.4 (Hall, 1999). The final alignment was performed online using MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/, accessed 14 July 2017) and inspected manually with BioEdit v7.0.4. Unsequenced regions were coded as missing data in the combined matrix. To infer phylogenetic relationships, maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) analyses were performed. MP analyses were conducted in PAUP version 4.0a149 (Swofford, 2016). For each data set, heuristic searches were conducted using 1000 replicates of random addition sequence, tree-bisection–reconnection (TBR) branch-swapping and 'keeping multiple trees' (MulTrees), but saving only 20 trees per replicate. Clade support was estimated by the bootstrap (Felsenstein, 1985) with 1000 replicates, TBR branch swapping and simple addition sequence. To explore the variability of each marker, four matrices were analysed with MP: (1) *rbcL*, (2) *trnK-matK-trnK*, (3) *trnT-trnL-trnF* and (4) all regions combined. Information about the alignment characteristics and number of variable and potentially parsimony informative sites is presented in Table 4. ML and BI analyses were conducted using the combined data only. An ML rapid bootstrap analysis (1000 replicates) with search for best-scoring ML tree in one run was conducted in RAxML v8.2.0 (Stamatakis, 2014). The best fitting substitution model was determined with jModeltest v2.1.7 (Darriba *et al.*, 2012; Guindon & Gascuel, 2003) using the Akaike information criterion. Evolutionary substitution models for each marker were calculated. The most complex substitution model, general time reversible (GTR+I+GAMMA) model with six substitution types (one for each pair of nucleotides) and gamma-distributed rate variation across sites and a proportion of invariable sites was finally chosen for the analysis. BI was performed using MrBayes v3.2.6 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). A partition scheme was set up by creating character sets for each of the three combined parts of the alignment: (1) *rbcL*, (2) *trnK-matK-trnK* and (3) *trnT-trnL-trnF*. Parameters were unlinked so that each partition has its own parameters. Overall rate variation was allowed to be different across partitions. By changing it to variable, the rates are allowed to vary under a flat Dirichlet prior. Two independent Metropolis-coupled Markov chain Monte Carlo (MCMC) analyses each with 10 million generations, sampling each 1000th generation, were run. The initial 25% of trees obtained from each MCMC run was removed as the burn-in. Each run consisted of three heated and one cold chain. A 50% majority rule consensus tree was calculated using the remaining trees to obtain posterior probabilities for each node. Outgroup taxa were specified to be Bixaceae. Trees were visualized and edited in FigTree v1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/, accessed 14 July 2017).

Molecular clock analysis

To obtain age estimates for the major clades of the groups of interest, a molecular clock analysis was performed in BEAST v2.4.4. (Drummond *et al.*, 2012) with an uncorrelated log-normal relaxed clock excluding the proportion of invariant sites parameter under the TVM+G4 model. This model was obtained by the model test implemented in IQ-TREE software (http://www.iqtree.org/, accessed 14 July 2017) under the

Table 3. Details of primers used in this study

Region	Primer	Sequence (5′–3′)	Usage	T_A (°C)	Reference
rbcL	rbcLa_f	ATGTCACCACAAACAGAGACTAAAGC	PCR and sequencing	63	Levin *et al.* (2003)
	rbcL_724R	TCGCATGTACCTGCAGTAGC	PCR and sequencing		Fay, Swensen & Chase (1997)
trnK-matK-trnK	trnK-799f	CCYTGTTYTRACYRTATYGCACTATGTAT	PCR and sequencing	65	Barfuss *et al.* (2016)
	trnK-2662r	CTCGAACCCGGAACTAGTCGG	PCR and sequencing		Castello *et al.* (2016)
	matK-DipF* (ratio 1:2):		Sequencing		Heckenhauer, Barfuss & Samuel (2016)
	matK-413f-1	TAATTTACRATCAATTCATTCAATATTTCC			
	matK-413f-4	TAATTTMCRATCAATTCATTCCATATTTCC			
	matK-DipR* (ratio 1:1:1):		Sequencing		
	matK-1227r-4	GARGATCCRCTRTRATAATGAGAAAAATTT			Heckenhauer, Barfuss & Samuel (2016)
	matK-1227r-5	GARGATCCRCTRTRATAATGAGAAATATTT			
	matK-1227r-7	GARGATCCGCTATRATAATGATAAATATTT			
trnT-trnL-trnF	a†	CATTACAAATGCGATGCTCT	PCR and sequencing	60	Taberlet *et al.* (1991)
	f†	ATTTGAACTGGTGACACGAG	PCR and sequencing		Taberlet *et al.* (1991)
	a_mod†	CATTACAAATGCGATGCTCTAAC	PCR and sequencing	68	This study
	f_mod†	ATTTGAACTGGTGACACGAGGAT	PCR and sequencing		This study
	c	CGAAATCGGTAGACGCTACG	Sequencing		Taberlet *et al.* (1991)
	h	CCATTGAGTCTCTGCACCTATC	Sequencing		Taberlet *et al.* (2007)

*Primers *matK*-DipF and *matK*-DipR were obtained by multiplexing several degenerate primers in different ratios according to Heckenhauer *et al.* (2016).
†Because of a higher annealing temperature (T_A), predominantly modified primers (a_mod and f_mod) of Taberlet *et al.* (1991) were used for amplification of *trnT-trnL-trnF*.

Bayesian information criterion (BIC). The input file for BEAST was first generated using Beauti implemented in BEAST and edited manually. The dating analysis was based on the study of Ducousso *et al.* (2004), which revealed that the last common ancestor of Sarcolaenaceae and Asian dipterocarps was ectomycorrhizal before the India–Madagascar separation, *c.* 87.6 ± 0.6 Mya.

There are fossils attributed to Dipterocarpaceae (e.g. Maury-Lechon & Curtet, 1998; Dutta *et al.*, 2011; Feng *et al.*, 2013), but they are not clearly assignable to any extant clade of the family, making them unusable as calibration points. Without expanding our analysis to include a much greater set of Malvales, we were unable to use fossils as calibration points. Here, a log-normal distribution with a mean of 87.5 My was used as calibration point to the most recent common ancestor of Sarcolaenaceae and Dipterocarpoideae. The following time of most recent common ancestor settings were log normal prior distribution with a mean of 87.5 and log standard deviation of 0.015 (real space). A log-normal prior with mean of 0.005 and standard deviation of 0.5 was placed on the mean of the log-normal relaxed clock rate. In our analyses, Monotoideae were sister to Dipterocarpoideae (Figs 1, 2), but this is not well supported. Thus, the correct position of Monotoideae remains unclear, and two alternative dating analyses were therefore run. In the first, a constraint consisting of Sarcolaenaceae, Monotoideae and Dipterocarpoideae was defined. In the second analysis, Sarcolaenaceae and Dipterocarpoideae were considered a clade. For each of our two analyses, we ran two separate chains for 300 million generations to achieve

Table 4. Parsimony characteristics and molecular evolutionary model for each locus and combined data set including Bixaceae, Cistaceae, Sarcolaenaceae and Dipterocarpaceae

	rbcL	*trnK-matK-trnK*	*trnT-trnL-trnF*	Combined data
Total number of accessions	192	252	250	254
Length of alignment	697	1908	3306	5911
Number of variable characters (%)	125 (17.9)	765 (40.1)	961 (29.1)	1851 (31.3)
Number of potentially parsimony-informative characters (%)	98 (14.1)	546 (28.6)	648 (19.6)	1292 (21.9)
Tree length of best parsimony tree (steps)	254	1280	1588	3185
Trees saved (parsimony analysis)	11460	20000	3600	14000
Consistency index	0.58	0.76	0.75	0.73
Retention index	0.93	0.95	0.71	0.68
Molecular evolutionary model	TVM+I+G	TVM+G	TVM+G	TVM+G
Number of substitution types (Nst)	6	6	6	6
Rates	Gamma shape	Gamma shape	Gamma shape	Gamma shape
Number of rate categories (Ncat)	4	4	4	4

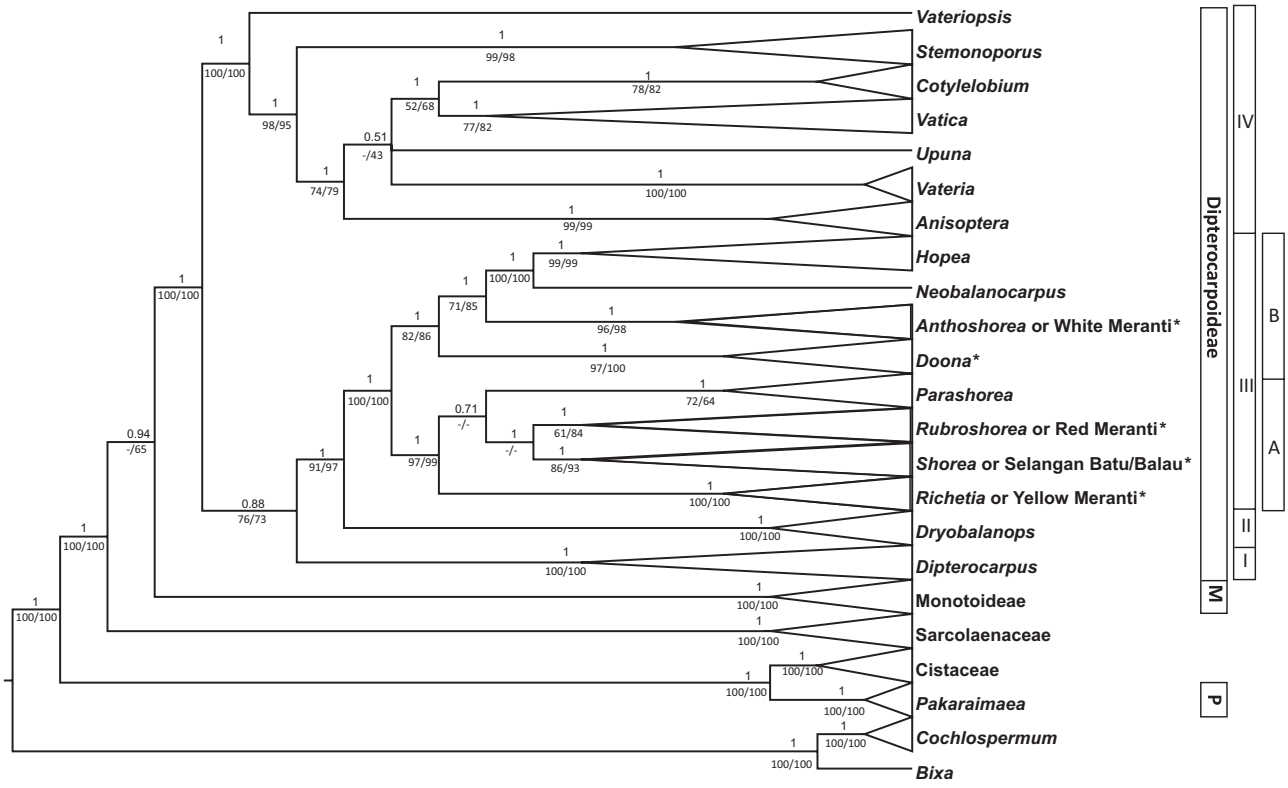
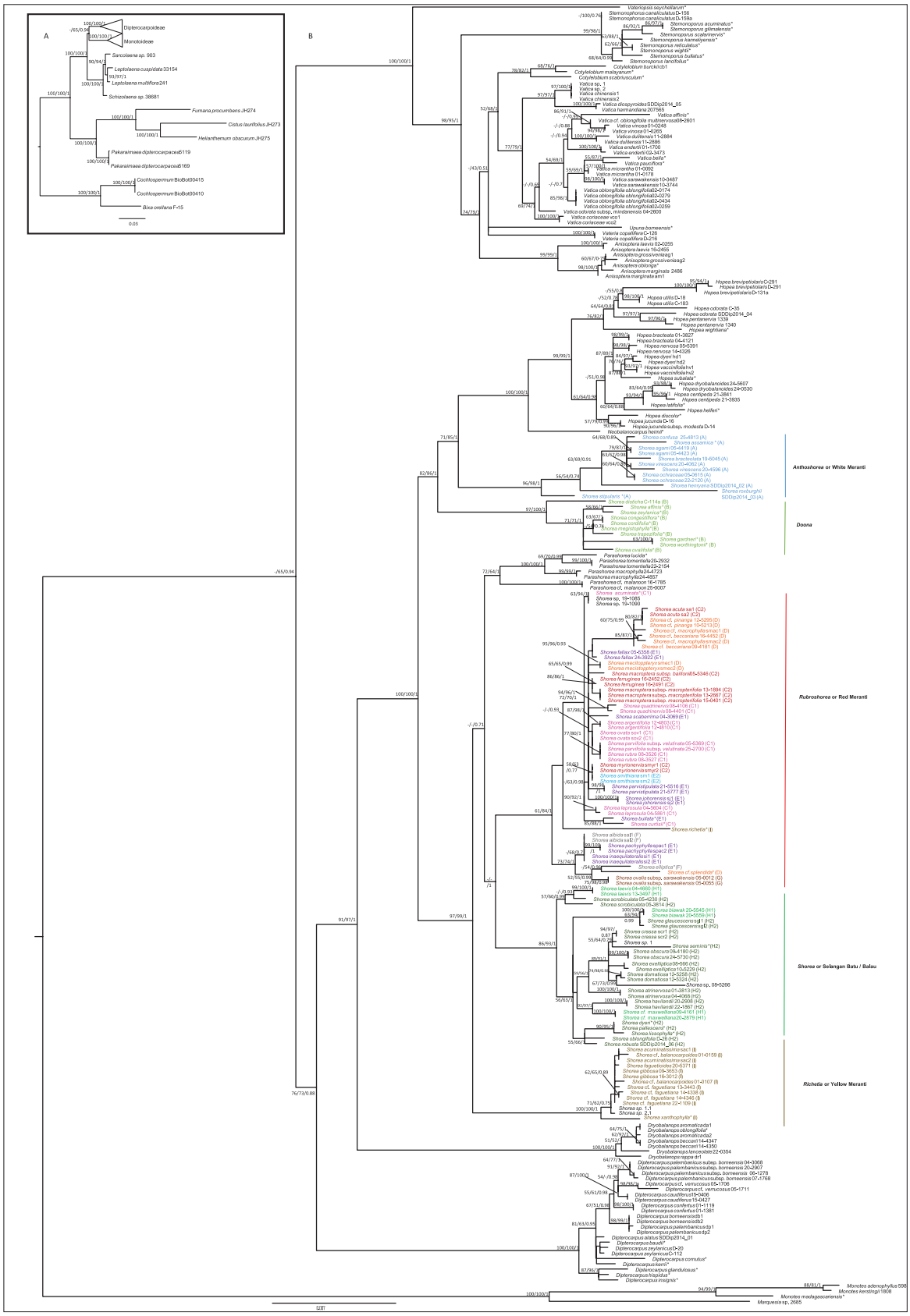


Figure 1. Bayesian 50% majority rule consensus tree from analyses of the combined plastid loci. Taxa are collapsed to major clades. Posterior probabilities ($BI_{PP} \geq 0.7$) are given above the nodes and bootstrap percentages (≥ 50%) from maximum parsimony and maximum likelihood analyses are shown below the nodes in this order. A hyphen indicates bootstrap support < 50%. The current classification of Dipterocarpaceae (Dipterocarpoideae, Monotoideae = M, Pakaraimaeoideae = P) is shown. The four major clades (I, II, III, IV) of Dipterocarpoideae and subclades (A, B) of the tribe Shoreeae are indicated. Different groups of *Shorea* are marked with an asterisk.

a reasonable effective sample size (ESS) of at least 200. Convergence and mixing of each run were assessed with Tracer v1.5.0 (http://tree.bio.ed.ac.uk/software/tracer/, accessed 14 July 2017). Both log and tree files were then trimmed to 250 million generations. The two log files were combined using LogCombiner using 5000



state samples each. For each chain, 60% of generations were discarded as burn-in. We combined the post burn-in trees in TreeAnnotator to construct a maximum clade credibility tree, which was displayed with Figtree v1.4.1. Since we are interested in the ages of the major clades, the maximum clade credibility tree was collapsed. To explore which of our two hypotheses [monophyly of (Sarcolaenaceae+Monotoideae+Dipterocarpoideae) or monophyly of only (Sarcolaenaceae+Dipterocarpoideae)] is better supported, we estimated marginal likelihoods for the two models using the path sampling (PS) method implemented in BEAST. PS analyses were conducted with 112 path steps, each run until the ESS reached 200. Marginal likelihood estimates where then used for calculation of the Bayes factor. We included the whole dataset used in the combined analysis, but the node used for the calibration thus becomes the root of the analysis, arranging the outgroups as sister to Sarcolaenaceae/Dipterocarpoideae. Their age assignments were thus not correctly estimated and are therefore not discussed here.

Chromosome counts and genome size measurements in Dipterocarpoideae

Actively growing root tips were pretreated with 0.002 M 8-hydroxyquinoline for 2.5 h at room temperature and 4 °C for 2.5 h, fixed in 3:1 ethanol/acetic acid and stored at –20 °C until use. Chromosome numbers were initially assessed by standard Feulgen staining of meristematic root cells (Jang *et al.*, 2013). Due to the small size of these chromosomes, additional preparations were also made using enzymatic digestion of cell walls to improve resolution of karyotypes. Preparations were made in a drop of 60% acetic acid with the cover slips off and the material was stained with 2 ng/µL DAPI (4′, 6-diamidino-2-2phenylindole) dissolved in the mounting antifade medium Vectashield (Vector Laboratories, Burlingame, CA, USA). Chromosomes were examined with an AxioImager M2 epifluorescence microscope with a high-resolution microscopy camera (Carl Zeiss, Vienna, Austria), and files were processed using AxioVision 4.8 (Carl Zeiss). At least three well-spread metaphases were analysed for each species.

Genome size was measured with flow cytometry performed on leaf material. Fresh tissue from plants growing in the Hortus Botanicus Vindobonensis (HBV) and recently collected silica-gel dried material from Sri Lanka were used. Together with leaves of the internal standard species, samples were chopped in Otto I buffer (Otto *et al.*, 1981) according to Galbraith *et al.* (1983). Standards were *Solanum pseudocapsicum* L., 1C = 1.30 pg (Temsch, Greilhuber & Krisai, 2010) or *Pisum sativum* L. 'Kleine Rheinländerin', 1C = 4.42 pg (Greilhuber & Ebert, 1994). After filtering of the isolate through a 30-µm nylon mesh, RNA was digested with 15 mg/L RNase A for 30 min at 37 °C. Afterward, DNA was stained in propidium iodide (50 mg/L) complemented with Otto II buffer (Otto *et al.*, 1981). Mean fluorescence intensity of at least 10 000 particles was measured with a CyFlow cytometer (Partec, Münster, Germany) equipped with a green laser (Cobolt Samba, Cobolt AB, Stockholm, Sweden); the 1C-value was calculated according to the formula: (MFIobject/MFIStandard) × 1C-value standard, where MFI is the mean fluorescence intensity of the G1 nuclei population. All measurements were carried out three times.

RESULTS

Sequence and alignment characteristics

There was no length variation in *rbcL* (697 bp), whereas the *trnK-matK-trnK* and *trnT-trnL-trnF* regions were variable among taxa. The aligned sequence length of the partial *trnK* intron region (including complete *matK*) was 1908 bp and that of the *trnT-trnL-trnF* region was 3306 bp. The *trnK-matK-trnK* region was the most informative region with 546 (28.61%) potentially parsimony-informative sites. The number of potentially parsimony-informative sites was 98 (14.06%) and 648 (19.6%) for *rbcL* and *trnT-trnL-trnF*, respectively (Table 4).

Phylogenetic analysis of the plastid loci

All three methods of phylogenetic inference (MP, ML, BI) for the combined data set revealed congruent results for the main clades, but there was some

Figure 2. Best-scoring maximum likelihood tree of a rapid bootstrap analysis with 1000 replicates of the combined data set. Bootstrap values (≥ 50%) obtained from maximum parsimony and maximum likelihood analyses and posterior probabilities ($BI_{PP} \geq 0.7$) obtained from Bayesian interference are given in this order. A hyphen indicates bootstrap support < 50% or $BI_{PP} < 0.7$. The relationships between different (sub-)families used in this study (A) and within Dipterocarpoideae (B) are shown. Sequences obtained from GenBank are indicated with an asterisk (*). Different groups of *Shorea* according to Maury (*Anthoshorea*, *Doona*, *Rubroshorea*, *Shorea*, *Richetia*) are indicated. Sections and subsections according to Ashton are given for each of the *Shorea* accessions: A, section *Anthoshorea*; B, section *Doona*; C, section *Mutica*; C1, section Mutica; C2, section Auriculatae; D, section *Pachycarpae*; E, section *Brachypterae*; E1, subsection *Brachypterae*; E2, subsection *Smithiana*; F, section *Rubella*; G, section *Ovalis;*, H, section *Shorea*; H1, subsection *Barbata*; H2, subsection *Shorea*, I, section *Richetioides*, subsection *Richetioides*.

variation in topologies in the terminal clades. The Bayesian (Fig. 1) and the maximum likelihood (Fig. 2) trees with bootstrap percentages from the MP (BS_{MP}) and ML (BS_{ML}) analyses and posterior probabilities from the BI (PP_{BI}) are shown.

Phylogenetic relationships in Dipterocarpaceae

Our main aim in this study was the clarification of the position of the three subfamilies of Dipterocarpaceae relative to Sarcolaenaceae and Cistaceae. Besides *Bixa* and *Cochlospermum* (Bixaceae), which were used as the outgroup and arranged as a clade sister to all other taxa, our analyses revealed four groups (Figs 1, 2): (1) Cistaceae including *Pakaraimaea* (the sole member of Pakaraimaeoideae; Fig. 1: P; BS_{MP} 100, BS_{ML} 100, PP_{BI} 1.00; this order will be used throughout; a hyphen indicates support < 50; Figs 1, 2), (2) Sarcolaenaceae (100, 100, 1.00), which were strongly supported (100, 100, 1.00) as sister to the clade containing Dipterocarpoideae (100, 100, 1.00) plus Monotoideae (100, 100, 1.00), (3) Monotoideae (consisting of *Monotes* and *Marquesia*, Fig. 1: M) and (4) all taxa belonging to the Asian subfamily Dipterocarpoideae (Fig. 1). The sister relationship between Monotoideae and Dipterocarpoideae was only weakly supported (-, 65, 0.94).

Phylogenetic relationships in subfamily Dipterocarpoideae

Dipterocarpoideae were divided in four clades (Fig. 1: I, II, III, IV), which are almost in accordance to the tribal division *sensu* Ashton except that *Dipterocarpus* (Fig. 1, clade I, 100, 100, 1.00) was weakly supported (76, 73, 0.88) as sister to clades II and III and thus separated from the remaining genera of Dipterocarpeae (clade IV). The sister relationship of *Dryobalanops* (Fig. 1, clade II, 100, 100, 1) to tribe Shoreeae (Fig. 1, clade III) was strongly supported (91, 97, 1.00). This third major clade (Fig. 1, clade III, 100, 100, 1.00) can be further divided into two main subclades (designated as A and B in Fig. 1). Subclade A (97, 99, 1.00) consisted of *Parashorea* (71, 64, 1.00), *Rubroshorea* (ined., 61, 84, 1.00), *Richetia* F.Heim or the yellow meranti group (100, 100, 1.00) and *Shorea* or selangang batu/balau group (86, 93, 1.00). Subclade B (82, 86, 1.00) contained three groups with the following taxa: (1) *Hopea* and *Neobalanocarpus* P.S.Ashton (100, 100, 1.00); (2) *Anthoshorea* Pierre or white meranti wood group (96, 98, 1.00); and (3) *Doona* Thwaites (97, 100, 1.00). It is notable that *Shorea richetia* Symington (obtained from GenBank), which has been assigned to *Richetia*, clustered with *Rubroshorea* (Fig. 2). This is possibly due to a missidentification. Species of *Anisoptera* Korth., *Cotylelobium* Pierre, *Stemonoporus* Thwaites, *Upuna* Symington, *Vateria* L., *Vateriopsis* F.Heim and *Vatica* L. formed a fourth major clade (Fig. 1, IV; 100, 100, 1.00). Monophyly of *Anisoptera* and *Stemonoporus* was strongly supported (99, 99, 1.00 and 99, 98, 1.00, respectively). In *Anisoptera*, *A. laevis* Ridl. was sister to the other three species, *A. grossivenia* Slooten, *A. marginata* Korth. and *A. oblonga* Dyer (Fig. 2). Species of *Vatica* and *Cotylelobium* each formed sister clades with weak to moderate support (77, 79, 1.00 and 78, 82, 1.00, respectively). *Vateriopsis seychellarum* F.Heim was sister to the other genera in that clade (100, 100, 1.00). Positions of *Upuna* and *Vateria* in this major clade were not well supported (Figs 1, 2).

Molecular dating analysis

The Bayes factor tests using the marginal likelihoods from the BEAST analyses found a clear preference (Bayes factor: 5.6) for the model with monophyletic constraint consisting of only Sarcolaenaceae and Dipterocarpoideae (marginal likelihood estimate: –29 614) over the model with the monophyletic constraint consisting of Monotoideae, Sarcolaenaceae and Dipterocarpoideae (marginal likelihood estimate: –29 619.6). Therefore, results from the analysis using the first model are presented. The age estimates obtained for the major clades showed a wide range (e.g. age estimate for Dipterocarpoideae: 39.3–71.6 Mya). The median crown age estimate for Dipterocarpoideae was 54.9 Mya. Further age estimates for the major clades can be found in Figure 5, but because of the way BEAST works (and our decision not to use a fossil to set the age of the deeper nodes because we judged none of them to be specific enough to be of use in our study) the divergences for the outgroup taxa are not relevant and will not be discussed.

Chromosomes and genome sizes in Dipterocarpoideae

The chromosome numbers determined in this study are given in Table 5 with those from earlier reports on Dipterocarpaceae. Chromosome numbers for five species (*Dipterocarpus zeylanicus* Thwaites: $2n = 22$; *Shorea megistophylla* P.S.Ashton: $2n = 14$; *Hopea jucunda* Thwaites: $2n = 21$; *Shorea oblongifolia* Thwaites: $2n = 14$; and *Vatica endertii* Slooten: $2n = 22$) are reported here for the first time (Fig. 3). Most of the newly counted species were diploid (Fig. 3A–B, D–F), but our chromosome counts of *Hopea jucunda* reveal triploidy (Fig. 3C). Karyotypes were similar and symmetrical for all with small metacentric, submetacentric or subtelocetric chromosomes in all analysed species in Dipterocarpaceae, which makes

Table 5. Chromosome numbers for Dipterocarpaceae

Taxon	Chromosome number	Putative ploidy	Reference(s)
Anisoptera costata Korth.	$2n = 20$	$2x$	Tixier (1953)
Anisoptera laevis Ridl.	$2n = 22$	$2x$	Jong & Lethbridge (1967)
Anisoptera scaphula Pierre	$2n = 20$	$2x$	Tixier (1960)
Anisoptera thurifera Blume	$2n = 22$	$2x$	Oginuma *et al.* (1998)
Dipterocarpus alatus Roxb. & G.Don	$2n = 20$	$2x$	Tixier (1953)
	$2n = 22$	$2x$	Roy & Jha (1965)
Dipterocarpus baudii Korth.	$2n = 22$	$2x$	Jong & Lethbridge (1967)
Dipterocarpus costatus C.F.Gaertn.	$2n = 20$	$2x$	Tixier (1960)
Dipterocarpus intricatus Dyer	$2n = 20$	$2x$	Tixier (1953)
Dipterocarpus kunstleri King	$2n = 20$	$2x$	Pancho (1971)
Dipterocarpus oblongifolius Blume	$2n = 22$	$2x$	Kaur *et al.* (1986)
Dipterocarpus sarawakensis Slooten	$2n = 22$	$2x$	Jong & Lethbridge (1967)
Dipterocarpus tuberculatus Roxb.	$2n = 20$ and 30	$2x$ and $3x$	Tixier (1960)
Dipterocarpus turbinatus C.F.Gaertn.	$2n = 20$	$2x$	Tixier (1960)
Dipterocarpus validus Blume	$2n = 20$	$2x$	Pancho (1971)
Dipterocarpus zeylanicus Thwaites	$2n = 22$	$2x$	*(PDA: D-20)
Dryobalanops oblongifolia Dyer	$2n = 14$	$2x$	Jong & Lethbridge (1967), Kaur *et al.* (1986)
Dryobalanops sumatrensis (J.F.Gmel.) Kosterm.	$2n = 14$	$2x$	Jong & Lethbridge (1967), Kaur *et al.* (1986)
Hopea beccariana Burck	$2n = 20, 21, 22$	$2x$, $3x$	Ashton (1982)
Hopea jucunda Thwaites	$2n = 21$	$3x$	*(PDA: D-16)
Hopea latifolia Symington	$2n = 21$	$3x$	Jong & Kaur (1979)
Hopea odorata Roxb.	$2n = 20–22$	$3x$	Kaur *et al.* (1986)
	$n = 7$	–	Sarkar *et al.* (1982)
	$2n = 14$	$2x$	Jong & Lethbridge (1967), Roy & Jha (1965)
	$2n = 20$	$2x$	Tixier (1960)
Hopea subalata Symington	$2n = 21$	$3x$	Kaur *et al.* (1986)
	$2n = 21$	$3x$	Jong & Kaur (1979)
Neobalanocarpus heimii (King) P.S.Ashton	$2n = 14$	$2x$	Jong & Lethbridge (1967)
Shorea acuminata Dyer	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea agami P.S.Ashton	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea argentifolia Symington	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea gardneri (Thwaites) P.S.Ashton	$2n = 14$	$2x$	Jong & Kaur (1979)
Shorea leprosula Miq.	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea macrophylla (de Vriese) P.S.Ashton	$2n = 14$	$2x$	Jong & Kaur (1979), Kaur *et al.* (1986)
Shorea macroptera Dyer	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea megistophylla P.S.Ashton	$2n = 14$	$2x$	*(PDA: D-24)
Shorea multiflora (Burck) Symington	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea oblongifolia Thwaites	$2n = 14$	$2x$	*(PDA: D-26)
Shorea ovalis (Korth.) Blume subsp. *ovalis*	$2n = 28$	$4x$	Kaur *et al.* (1986)
Shorea ovalis (Korth.) Blume subsp. *sericea* (Dyer) P.S.Ashton	$2n = 21, 27, 28$	$3x$ and $4x$	Jong & Kaur (1979)
Shorea pauciflora King	$2n = 14$	$2x$	Kaur *et al.* (1986)
Shorea pinanga Scheff.	$2n = 14$	$2x$	Jong & Kaur (1979), Kaur *et al.* (1986)
Shorea platyclados Slooten ex Endert	$2n = 14$	$2x$	Kaur *et al.* (1986)

Table 5. *Continued*

Taxon	Chromosome number	Putative ploidy	Reference(s)
Shorea resinosa Foxw.	$2n = 21$	$3x$	Jong & Kaur (1979), Kaur *et al.* (1986)
Shorea robusta C.F.Gaertn.	$2n = 14$	$2x$	Roy & Jha (1965), Pal *et al.* (1993)
Shorea roxburghii G.Don	$2n = 14$	$2x$	Roy & Jha (1965), *(S. Duangjai_Dip2014_03)
Shorea splendida (de Vriese) P.S.Ashton	$2n = 14$	$2x$	Jong & Kaur (1979), Kaur *et al.* (1986)
Shorea stenoptera Burck	$2n = 14$	$2x$	Jong & Kaur (1979), Kaur *et al.* (1986)
Shorea trapezifolia (Thwaites) P.S.Ashton	$2n = 14$	$2x$	Jong & Kaur (1979)
Vateria indica L.	$n = 10$	–	Mehra (1976)
Vatica endertii Slooten	$2n = 22$	$2x$	*(UBDH: UBD-CTFS: 01-1700)
Vatica odorata (Griff.) Symington	$2n = 22$	$2x$	Roy & Jha (1965)

Previously published chromosome counts and its references were obtained from http://ccdb.tau.ac.il/ (Rice *et al.*, 2015, accessed 14 July 2017). Counts from the present study are indicated with an asterisk (*). Herbarium voucher of mother plant is given in parentheses. –, not indicated.

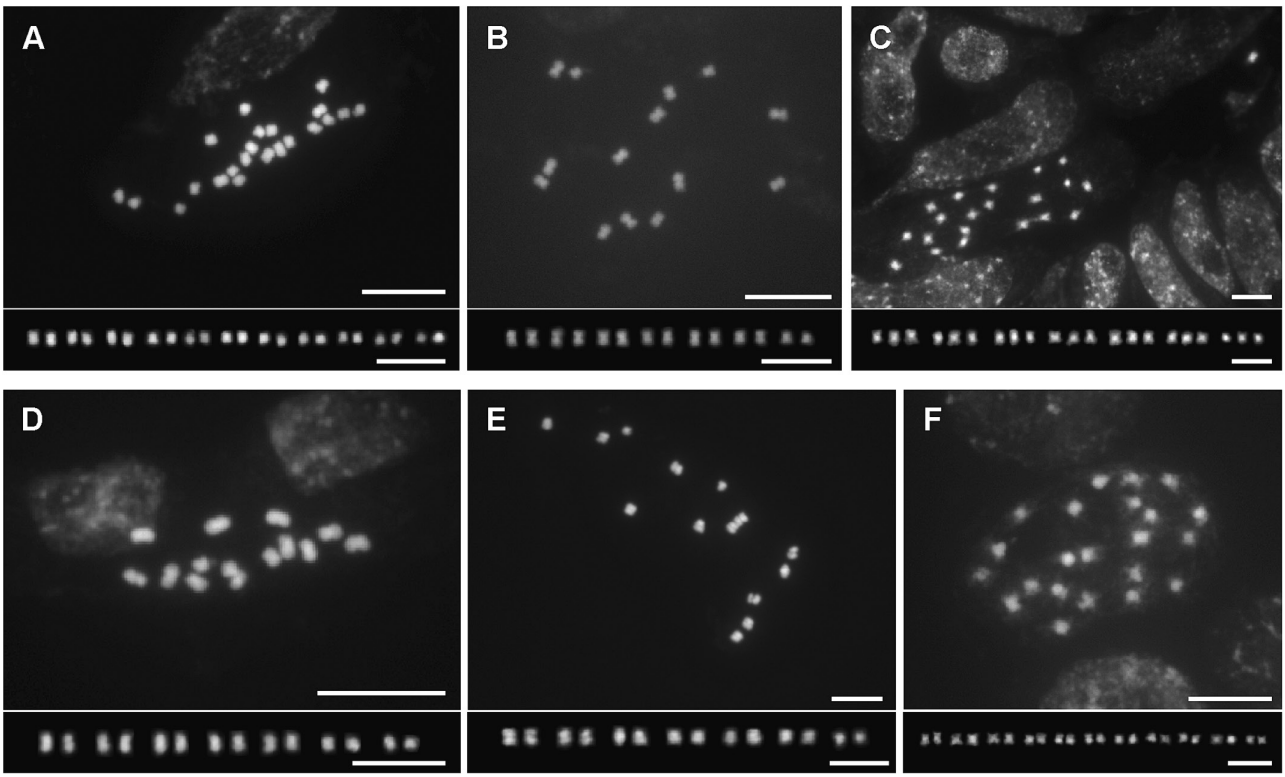


Figure 3. Mitotic chromosomes of some species of Dipterocarpaceae. A, *Dipterocarpus zeylanicus* ($2n = 2x = 22$). B, *Shorea megistophylla* ($2n = 2x = 14$). C, *Hopea jucunda* ($2n = 2x = 21$). D, *Shorea oblongifolia* ($2n = 2x = 14$). E, *Shorea roxburghii* ($2n = 2x = 14$). F, *Vatica endertii* ($2n = 2x = 22$). Scale bars = 5 μm.

identification of individual chromosome pairs difficult (Fig. 3). Similar to a recent study of genome sizes in Dipterocarpaceae (Ng *et al.*, 2016), our measurements of genome size showed differences among and within genera (Table 6, Fig. 4) and range from 1C = 0.3264 pg in *Shorea roxburghii* G.Don to 0.6724 pg in *Vatica*

Table 6. Genome size measurements in Dipterocarpaceae

Taxon	Taxon-ID	C-value	SD
Dipterocarpus hispidus Thwaites	D-7a	0.3287	0.0017
Dipterocarpus zeylanicus Thwaites	D-20	0.3933	0.0178
Hopea brevipetiolaris (Thwaites) P.S.Ashton	D-291	0.3931	0.0016
Hopea brevipetiolaris (Thwaites) P.S.Ashton	C-291	0.3955	0.0028
Hopea brevipetiolaris (Thwaites) P.S.Ashton	D-131a	0.3899	0.0039
Hopea jucunda Thwaites	D-16	0.5949	0.0035
Hopea jucunda subsp. *modesta* (A.DC.) Kosterm.	D-14	0.3277	0.0001
Hopea odorata Roxb.	C-35	0.4216	0.002
Hopea odorata Roxb.	C-36	0.6051	0.0042
Hopea odorata Roxb.	C-37	0.6094	0.0036
Hopea utilis (Bedd.) Bole	D-18	0.3663	0.0013
Shorea lissophylla Thwaites	D-27	0.3586	0.0047
Shorea oblongifolia Thwaites	D-26	0.3669	0.0024
Shorea roxburghii G.Don	S.Duangjai_Dip2014_03	0.3264	0.0033
Shorea zeylanica (Thwaites) P.S.Ashton	C-160	0.4537	0.0012
Stemonoporus canaliculatus Thwaites	D-156	0.4005	0.0011
Vateria copallifera (Retz.) Alston	D-216	0.4581	0.0011
Vatica endertii P.S.Ashton	01-1700	0.5505	0.0006
Vatica diospyroides Symington	S.Duangjai_Dip2014_05	0.6724	0.0079
Vatica sp.	D-198a	0.448	0.0012

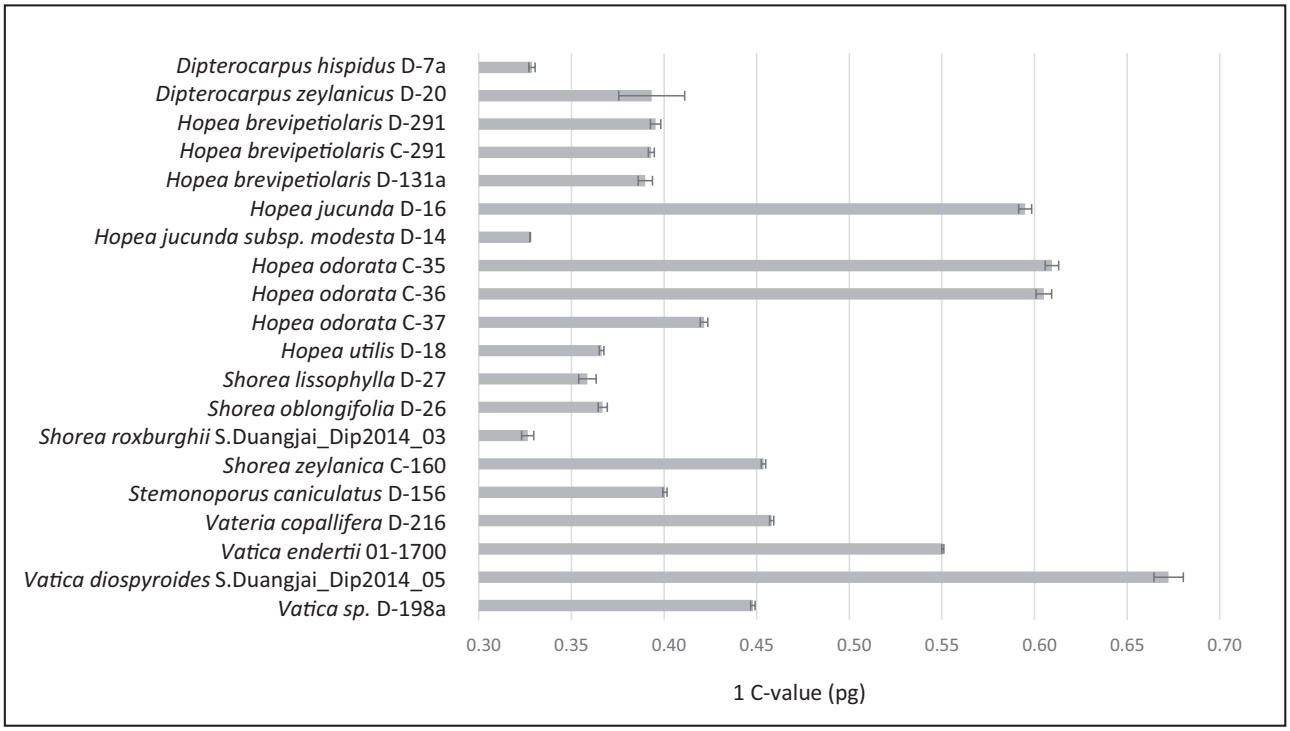


Figure 4. Genome size in several species of Dipterocarpaceae with standard deviation based on three measurements of each individual.

diospyroides Symington. Although most species show uniform genome size, intraspecific variation was detected in *Hopea odorata* Roxb. (1C = 0.4216, 0.6051 and 0.6094 pg).

DISCUSSION

This study provides a comprehensive molecular phylogenetic tree of the ecologically and economically important family Dipterocarpaceae including all three subfamilies, Cistaceae and Sarcolaenaceae, the largest Madagascan endemic family. Taxonomic issues among the three subfamilies, especially in Asian Dipterocarpoideae, could be refined. Molecular phylogenetic analyses have assigned Dipterocarpaceae to Malvales (APG IV, 2016), and recent genetic studies have shown that at least Dipterocarpoideae share a unique common ancestor with Sarcolaenaceae, a family of trees endemic to Madagascar (Dayanandan *et al.*, 1999; Ducousso *et al.*, 2004). The close relationship between Dipterocarpaceae and Sarcolaenaceae has been emphasized by Maguire & Ashton (1977) and Ashton (1982) on morphological evidence and was supported by anatomical features (Capuron, 1970; de Zeeuw, 1977). In addition, results of numerous molecular studies employing plastid and nuclear genes have indicated Cistaceae to be the closest relatives of Dipterocarpaceae in the broadly circumscribed order Malvales (APG, 1998; Savolainen *et al.*, 2000; Soltis *et al.*, 2000). This is supported by the similarity in the structure of the chalazal region of the mature seed (Nandi, 1998) and strongly suggests a common ancestry of at least Monotoideae, Pakaraimaeoideae, Sarcolaenaceae and possibly Dipterocarpoideae, Bixaceae and Cistaceae. All three subfamilies of Dipterocarpaceae (Högberg, 1982; Alexander & Högberg, 1986; Högberg & Piearce, 1986; Lee, 1990; Moyersoen, 2006), Sarcolaenaceae (Ducousso *et al.*, 2004) and Cistaceae (Smith & Read, 1997) are ectomycorrhizal. Our results differed from the widely used subfamily concept for Dipterocarpaceae based on morphological and anatomical evidence consisting of three subfamilies, Dipterocarpoideae, Monotoideae and Pakaraimaeoideae. *Pakaraimaea* is more closely related to Cistaceae, but their exact relationship could not be determined from our limited sampling of the latter. The close relationship between Cistaceae and *Pakariamaea* has been already suggested by Alverson *et al.* (1998), Kubitzki & Chase (2003), Ducousso *et al.* (2004) and Horn, Wurdack & Dorr (2016). *Pakaraimaea* was recently included in Cistaceae in APG (2016). This was an unexpected result from an ecological point of view. Cistaceae are also woody with a few herbaceous members (Proctor, 1978). Dipterocarpaceae (including *Pakaraimaea*) and Sarcolaenaceae are exclusively tropical, but Cistaceae are distributed primarily in the temperate areas of Europe, principally in the Mediterranean Basin and, to a much more limited extent, in North and South America (http://www.mobot.org/mobot/research/apweb/orders/malvalesweb.htm#Cistaceae, accessed 14 July 2017). *Pakaraimaea* are relatively small trees (Maury-Lechon & Curtet, 1998), recalling *Stemonoporus* in architecture. Leaf venation of *Pakaraimaea* shows similarities to those of *Cotylelobium* and *Anisoptera* (Ashton, 2003). On the other hand, there are features not shared by *Pakaraimaea* and Asian Dipterocarpaceae, which supports removing *Pakaraimaea* from Dipterocarpaceae. Contrary to the thick-walled intricately structured pericarp wall of Dipterocarpaceae, the thin fruit pericarp of *Pakaraimaea* has a simple structure. The five-celled fruit dehisces loculidally. There is continuing growth of the cotyledons following germination, and albumen occurs in the ripe embryo, all as in *Monotes*. *Pakaraimaea* petals are shorter than the sepals, and the anthers appear versatile as in Monotoideae (Maury-Lechon & Curtet, 1998). Wood rays are biseriate (Maury-Lechon & Curtet, 1998). The ovary of Dipterocarpoideae and Monotoideae is three-celled, each bearing two seeds (four in *Monotes*). The five-celled ovary of *Pakaraimaea*, each cell bearing two (rarely four) ovules per loculus (Maguire & Ashton, 1977), is unique in Dipterocarpeaceae but typically malvalean and could therefore be primitive within the family. Locules with two to > 30 ovules and two to many have been observed in Sarcolaenaceae (Bayer, 2003) and Cistaceae (Arrington & Kubitzki, 2003), respectively. Ripe fruits of Dipterocarpoideae are one-seeded nuts, generally woody, sometimes corky (Ashton, 2003), and *Pakaraimaea* fruits contain at most one fertile seed although other aborted seeds persist. In *Pakaraimaea* and Monotoideae pollen is tricolporate, with well-developed endexine and a distinct foot layer, whereas in Dipterocarpoideae pollen grains are tricolpate and lack endexine. Anthers are basifixed in Dipterocarpoideae and basi-versatile in Monotoideae and Pakaraimaeoideae. In *Pakaraimaea* and Monotoideae, wood, leaves and ovary are devoid of resin (Maury-Lechon & Curtet, 1998), whereas Dipterocarpoideae are distinguished by the universal presence of intercellular resin canals. Our analyses showed that Monotoideae are probably sister to the Asian dipterocarps, but this was not well supported (Figs 1, 2). The position of Monotoideae needs to be further investigated with a broader taxon sampling and more data. A more detailed analysis is required to obtain further insights into the relationships of Sarcolaenaceae–Dipterocarpaceae–Cistaceae, which could be combined in an expanded family concept as discussed in APG (2009, 2016). Cistaceae is the

oldest name of these three, but conservation of Dipterocarpaceae may be considered as an option to preserve the name of this economically important group of forest trees, if these are to be combined in a single family.

With respect to the large clade of Asian Dipterocarpoideae, we discuss the four clades obtained in our molecular analysis (Fig. 1) and some morphological features. The concept of two tribes, Dipterocarpeae and Shoreeae, was not supported in our analyses. Our results separated *Dipterocarpus* from the remaining genera of Dipterocarpeae (Fig. 1: clades I and IV) and it was weakly supported (76, 73, 0.88) as sister to *Dryobalanops* (Fig. 1: clade II) and Shoreeae (Fig. 1: clade III). This has also been observed in earlier molecular studies (e.g. Kajita *et al.*, 1998; Yulita *et al.*, 2005; Gamage *et al.*, 2006). In the study of Indrioko *et al.* (2006), depending on the outgroup, *Dipterocarpus* was sister either to the remaining Dipterocarpeae (bootstrap support: 80%) or to Shoreeae (bootstrap support: 83%). We acknowledge that the weak support obtained from our analysis limits our ability to interpret this relationship. *Dipterocarpus* could perhaps be sister to other Dipterocarpeae or the latter form a separate tribe. *Dipterocarpus* makes trees that are columnar but hardly buttressed with untidy globose crowns and prominently lenticellate orange–brown massively flaky bark, which at once makes these recognizable as distinct from other large forest dipterocarps. They have the chromosome number, $2n$ = 20–22, as in other Dipterocarpeae (e.g. Tixier, 1953; Tixier, 1960; Table 5) but differ from other Dipterocarpoideae further in their dispersed resin canals in the wood (Meijer, 1979; Ashton, 1982). Other typical characters are large leaf buds, amplexicaul bud scales and stipules furnished with diverse species-defining indumenta, plicate venation resulting in corrugation of their coriaceous leaves, thickly geniculate and often long petioles with often complex rings of vascular bundles and resin canals, large flowers bearing a tubular calyx united at base into a smooth, angled, tuberculate or flanged tube enclosing but free from the ovary, two aliform, valvate sepals, and 15–40 stamens that are larger than in all other dipterocarp taxa and have elongate orange anthers and stout tapering connectival appendages. First-branching Dipterocarpoideae exhibit relatively large orange anthers, whereas those are reduced in size and white in most derived clades. Dipterocarpaceae are pollinated by pollenivores (Thysanoptera, Ashton, Givnish & Appanah, 1988; Kondo *et al.*, 2016; multiple species of Coleoptera, Appanah & Chan, 1981; Momose *et al.*, 1998; Nagamitsu, Harrison & Inoue, 1999; Sakai *et al.*, 1999; flies, Khatua, Chakrabarti & Mallick, 1998; and bees, Khatua, Chakrabarti & Mallick, 1998; Momose, Nagamitsu & Inoue, 1996; see also Corlett, 2004). *Shorea acuminata* Dyer is also pollinated by a species of *Geocoris*, a major predator of thrips (Kondo *et al.*, 2016). *Dipterocarpus* is mainly pollinated by nectarivorous Lepidoptera (Ghazoul, 1997; Harrison *et al.*, 2005; Ashton, 2014), but also by Hymenoptera (Apis; Harrison *et al.*, 2005) and, to a small extent, by Coleoptera (Harrison *et al.*, 2005) and birds (Ghazoul, 1997).

Furthermore, our results revealed a sister relationship of *Dryobalanops* (II) to Shoreeae (III) (91, 97, 1.00; Figs 1, 2), which is in agreement with earlier molecular studies (Tsumura *et al.*, 1996; Kajita *et al.*, 1998, Kamiya *et al.*, 1998, Gamage *et al.*, 2003, 2006; Yulita, 2013). However, in the study of Indrioko *et al.* (2006), depending on the outgroup selection, *Dryobalanops* clustered with either Dipterocarpeae or Shoreeae. This ambiguity over the placement of *Dryobalanops* with either Shoreeae or Dipterocarpeae is reflected in its morphology and chromosome number. It shares wood anatomical characters (fibres with bordered pits, scattered resin canals and solitary vessels) with Dipterocarpeae, whereas its chromosome number, n = 7, and a thickened fruit sepal base are similar to those of Shoreeae (Gottwald & Parameswaran, 1966; Ashton, 1982). Moreover, being subvalvate, the sepals in fruit are intermediate between these tribes (Dipterocarpeae, valvate; Shoreeae, imbricate). Besides the strong bootstrap support, it is not clear from morphological characters if *Dryobalanops* could be included in the tribe Shoreeae or kept as an independent tribe.

Regarding the third clade (Fig. 1, clade III), our analyses clearly showed that *Hopea*, *Parashorea*, *Neobalanocarpus* and paraphyletic *Shorea* (tribe Shoreeae) should probably not be separated into distinct genera without additional evidence. This also has been reported in earlier molecular analyses (e.g. Yulita *et al.*, 2005; Gamage *et al.*, 2006). Pollen morphology of *Shorea*, *Hopea*, *Parashorea* and *Neobalanocarpus* is fairly uniform (Talip, 2008) and there are no obvious morphological characters to separate these four genera. *Anthoshorea* and *Doona* (endemic to Sri Lanka) form distinct groups, sister to *Hopea* and *Neobalanocarpus* (Fig. 1: clade III, subclade B), an observation also reported by Gamage *et al.* (2006). For species-rich *Shorea*, 11 sections have been proposed by Ashton (1982), based on the independent characters of androecium and bark morphology proposed by Symington (1943) and amplified by Whitmore (1963). Sections and subsections for each species of *Shorea* included in this study are given in Figure 2. However, our molecular analyses could not clearly separate these sections, but five groups of *Shorea* were observed. These groups were also recovered by Gamage *et al.* (2006) and correspond to the classification of Maury (1978; Table 1; Fig. 1, clade III; Fig. 2). According to Maury (1978), *Shorea* consists of six genera, *Anthoshorea*, *Rubroshorea*,

Richetia, *Shorea*, *Doona* and *Pentacme* A.DC. (the last was not included in our study), generic limits which correlate with the field characters of bark and wood anatomy proposed by Symington (1943; *Anthoshorea* = white meranti, *Rubroshorea* = red meranti, *Richetia* = yellow meranti and *Shorea* = selangang batu/balau). *Rubroshorea* is held together solely by the red colour of their wood, a character also found in some other *Shorea* spp. All species for which characteristics have been observed (two thereby excepted) are unambiguously attributable to the five sections in *Rubroshorea* (*Brachypterae* F.Heim, *Mutica* P.S.Ashton, *Ovalis* Symington ex P.S.Ashton, *Pachycarpae* P.S.Ashton and *Rubella* P.S.Ashton) recognized by Ashton. To evaluate whether the classifications proposed by Ashton (1964, 1968, 1980, 1982), Maury (1978) and Maury-Lechon (1979a, b) can be supported, *Pentacme* needs to be included in subsequent analyses. Furthermore, in our results, it is obvious that *Shorea* should include *Hopea*, *Parashorea* and *Neobalanocarpus*. Here, *Parashorea* clustered with *Rubroshorea* (red meranti) and *Shorea* (selangan batu/balau). A close relationship between *Shorea* and *Parashorea* was also confirmed in earlier molecular studies (Tsumura *et al.*, 1996; Kajita *et al.*, 1998; Kamiya *et al.*, 2005; Gamage *et al.*, 2003, 2006; Indrioko *et al.*, 2006). In an AFLP analysis, *Parashorea* clustered with *Hopea*, which could be explained by interspecific hybridization or ancestral polymorphisms as suggested by Cao *et al.* (2006). *Neobalanocarpus* was sister to *Hopea* (Figs 1, 2), which contradicts the nuclear *PgiC* analysis of Kamiya *et al.* (2005), in which *Neobalanocarpus* is nested in the white meranti group of *Shorea*. This could indicate hybridization between a species of white meranti and one of *Hopea*, as suggested by Kamiya *et al.* (2005). Evidence for hybridization comes from irregular meiosis (*Neobalanocarpus*) and existence of morphologically intermediate individuals between other species in Shoreeae (Ashton, 2003).

The fourth clade comprised *Anisoptera*, *Cotylelobium*, *Stemonoporus*, *Upuna*, *Vateria*, *Vateriopsis* and *Vatica* (Fig. 1, clade IV). *Anisoptera laevis* was highly divergent from the other three species, *A. grossivenia*, *A. marginata* and *A. oblonga* (Fig. 2) to which it was sister. This fits the classifications of Ashton (1964, 1968, 1980, 1982), Maury (1978) and Maury-Lechon (1979a, b), in which *Anisoptera* is divided into two sections, *Glabrae* (ined.), to which *A. laevis* is assigned, and *Anisoptera* Korth. containing the other three species included in this study. This is well supported by the morphological features of the flower buds, number of stamens, style and stigma (Ashton, 2003). Monophyly of *Stemonoporus*, which is endemic to Sri Lanka, was strongly supported (99, 100, 1.00) in all analyses, consistent with previous molecular studies (Dayanandan *et al.*, 1999; Gamage *et al.*, 2003, 2006) and its distinctive morphological features, including peculiar anthers with apical dehiscence, leaf traces that separate from the central vascular cylinder well before the node and the absence of wing-like sepals (Ashton, 1982). *Cotylelobium* was weakly (MP, ML) to highly supported (BI) as sister to *Vatica* (Figs 1, 2). Similar results have occurred in previous studies (Kajita *et al.*, 1998; Kamiya *et al.*, 1998; Gamage *et al.*, 2006). In our results, the positions of *Vateria* and *Upuna* remained unresolved or weakly supported, although a sister relationship of *Upuna* to *Anisoptera* has been suggested by one of the co-authors (P. S. Ashton, pers. comm.). *Vateriopsis seychellarum*, which is endemic to the Seychelles, has unique anatomical features: many stamens, implying a primitive condition (Ashton, 1982), with anthers of a type attracting bees, although no native bees currently survive on the islands. It is sister to the remaining genera in clade IV (Figs 1, 2). *Stemonoporus* and *Vateria*, the endemic genera of Gondwanan peninsular India, also have wingless fruits.

Besides clarification of phylogenetic relationships in Dipterocarpaceae and allied families, one of the aims in this study was to obtain estimation of divergence times of Dipterocarpaceae and infer ages of major clades in Dipterocarpoideae. The biogeography and origin of Cistaceae, Sarcolaenaceae and Dipterocarpaceae have been widely discussed. The age of crown-group Cistaceae is *c.* (18.5–) 14.2 (–10.2) Mya (Guzmán & Vargas, 2009). Diversification in Sarcolaenaceae possibly began only 4.5 Mya (http://www.mobot.org/mobot/research/apweb/orders/malvalesweb.htm#Sarcolaenacea, accessed 14 July 2017). Wikström, Savolainen & Chase (2001) estimated the origin of Dipterocarpaceae as 14–28 Mya, but these dates are based on an analysis that included only one dipterocarp. Such limited sampling was stated by those authors to underestimate ages in terminal clades. On the other hand, based on the ectomycorrhizal status of *Pakaraimaea*, Moyersoen (2006) suggested that Dipterocarpaceae occurred on Gondwana *c.* 135 Mya. Fossil resin and pollen grains from the early Eocene of western India (Dutta *et al.*, 2009, 2011; Rust *et al.*, 2010) suggested an origin or early occurrence of Dipterocarpaceae in India and later dispersal to Southeast Asia–Malesia and southern China after contact of the two was established *c.* 50 Mya (Feng *et al.*, 2013; Shukla, Mehrotra & Guleria, 2013). In addition, Tertiary fossils of East Africa have been attributed to *Dipterocarpus* (Bancroft, 1935). Thus, *Dipterocarpus* is the only genus of Dipterocarpoideae known from Africa (Bancroft, 1935; Ashton & Gunatilleke, 1987). Although it would seem unlikely that the diversity of other dipterocarpoid genera could all have originated on the Indian Noah's Ark, some may have later dispersed there and gone extinct in Africa. During the late Oligocene and early Miocene

(20–23 Mya), Dipterocarpoideae occurred in the monsoon forests of the Sunda region and were therefore already distributed across Southeast Asia at the time of widespread expansion of evergreen rainforest in the later part of the early Miocene and probably have become a major part of the Southeast Asian rainforest only since then (Morley, 2000). The irregular flowering pattern followed by the distinctive masting behaviour of Dipterocarpaceae, which depends on sudden cool spells resulting from El Niño oscillations, supports their origin in a seasonal climate (Ashton, 1988). For molecular clock analyses, fossils are often used as calibration points for defined clades. Several dipterocarp fossils are reported in the literature (e.g. Dutta *et al.*, 2011; Feng *et al.*, 2013). However, placing fossils in the correct position on the phylogenetic tree is crucial for correct interpretation (Forest, 2009), and we faced several problems in assigning the described fossils to clades in our trees. For example, winged fruits and associated leaves of *Shorea* are reported from the late Eocene of South China by Feng *et al.* (2013) and are described as *Shorea maomingensis* Feng, Kodrul & Jin. According to Feng *et al.* (2013), this fossil can be attributed to *Shorea ovalis* (Korth.)Bl. subsp. *sericea* (Dyer) P.S.Ashton. Feng *et al.* (2013) suggested that the fossil leaves show the greatest similarities to *Shorea*. According to P. S. Ashton (pers. comm.), these leaves differ from Dipterocarpaceae in the nature of their reticulate tertiary venation, whereas the fruit is not that of the tetraploid *S. ovalis*, but almost certainly a species with subauriculate sepal bases in *Shorea* section *Anthoshoreae.* Another problem of using fossils for calibration is that they represent minimum ages. Coetzee & Muller (1984) reported intricate pollen tetrads of extant taxa of Sarcolaenaceae from South Africa in the Miocene, but these probably do not represent the oldest occurrence of this family (Nilsson, Coetzee & Grafström, 1996). It may have been an ancient endemic African taxon that migrated to Madagascar where it became restricted (Raven & Axelrod, 1974). As the Sarcolaneaceae pollen fossils are young (Miocene), we decided not to use them as a calibration point. To avoid the problem of incorrect placement, we did not use any fossils and instead applied the time of separation of Madagascar from the India–Seychelles block (87.6 ± 0.6 Mya) as a calibration point for Sarcolaenaceae plus Dipterocarpoideae. Potentially due to differences between phylogenetic models or implementation in BEAST used for the age estimation, the topology of the dated maximum clade credibility tree (Fig. 5) differed slightly from the trees obtained in our other analyses (Figs 1, 2). However, differences in topologies were not well supported in either result (see posterior probabilities; Fig. 5). Our dating study gives a general time frame for the major clades in Dipterocarpoideae and shows that they had already diverged into the extant genera by the end of the Miocene. Our median crown age estimate for Dipterocarpoideae was 54.9 Mya. The emergence of Dipterocarpeae was dated to 47.7 Mya (crown age). The dating analysis revealed 43.3 Mya as the median age of Shoreeae and *Dryobalanops*. Monotypic *Vateriopsis* is endemic to the Seychelles and has wingless fruits and seeds that are inviable in salt water, implying early separation from other Dipterocarpoideae (63 Mya; Ashton, 2014). Our results here indicate 34.9 Mya as the median age of *Vateriopsis*. The separation of the Seychelles from India began *c.* 63.4 Mya (Collier *et al.*, 2008). Therefore, our age estimates imply that *Vateriopsis* reached its current position not by continental drift, but rather by long-distance dispersal. Our dating estimate of 15.4 Mya for *Vateria* corresponds to the occurrence of the fossil *Vaterioxylon* in northern India in the Miocene (Maury-Lechon & Curtet, 1998) and suggest parallel evolution of *Vateria* and *Upuna*. An expanded analysis including a much larger set of Malvales would permit the use of multiple calibrations points and would be suitable to obtain further insights into the ages of clades in the larger set of taxa included here.

Earlier reports of chromosome numbers for Dipterocarpoideae indicated a high level of uniformity in the species and genera with *Anisoptera*, *Dipterocarpus*, *Upuna* and *Vatica* having $x = 11$ and *Dryobalanops*, *Hopea*, *Neobalanocarpus*, *Parashorea* and *Shorea* having $x = 7$ as the basic chromosome numbers. Some species of the last exhibit a chromosome number of 20, 21 and 22, assuming that $x = 11$ might have been derived from $x = 7$ through hybridization and polyploidization (Bawa, 1998). Our additional chromosome counts confirm those of earlier studies and demonstrate further evidence of polyploidy in Dipterocarpaceae (*Hopea jucunda*: $2n = 21$), which has been reported in *Shorea* [e.g. *S. ovalis* (Korth.) Blume with $2n = 28$ and *S. resinosa* Foxw. with $2n = 21$; Kaur *et al.*, 1986)] and *Hopea* (e.g. *H. odorata*: $2n = 20–22$ and *H. subalata* Symington: $2n = 21$, Kaur *et al.*, 1986). Furthermore, intraspecific variation in chromosome numbers has been observed (e.g. *H. odorata*: $2n = 14, 20, 21, 22$; Jong & Lethbridge, 1967; Kaur *et al.*, 1986). However, variation in chromosome numbers has to be interpreted with caution due to the often small sample size (Bawa, 1998). For example, the form of intraspecific variation in *H. odorata* is dysploid or polyploid, but it remains unclear if it is in the form of occasional dysploid individuals or polyploid populations (Ashton, 1982). Sampling of several individuals in the same population and of the same species from different populations would be helpful in evaluating variation and its significance (Bawa, 1998). Genome

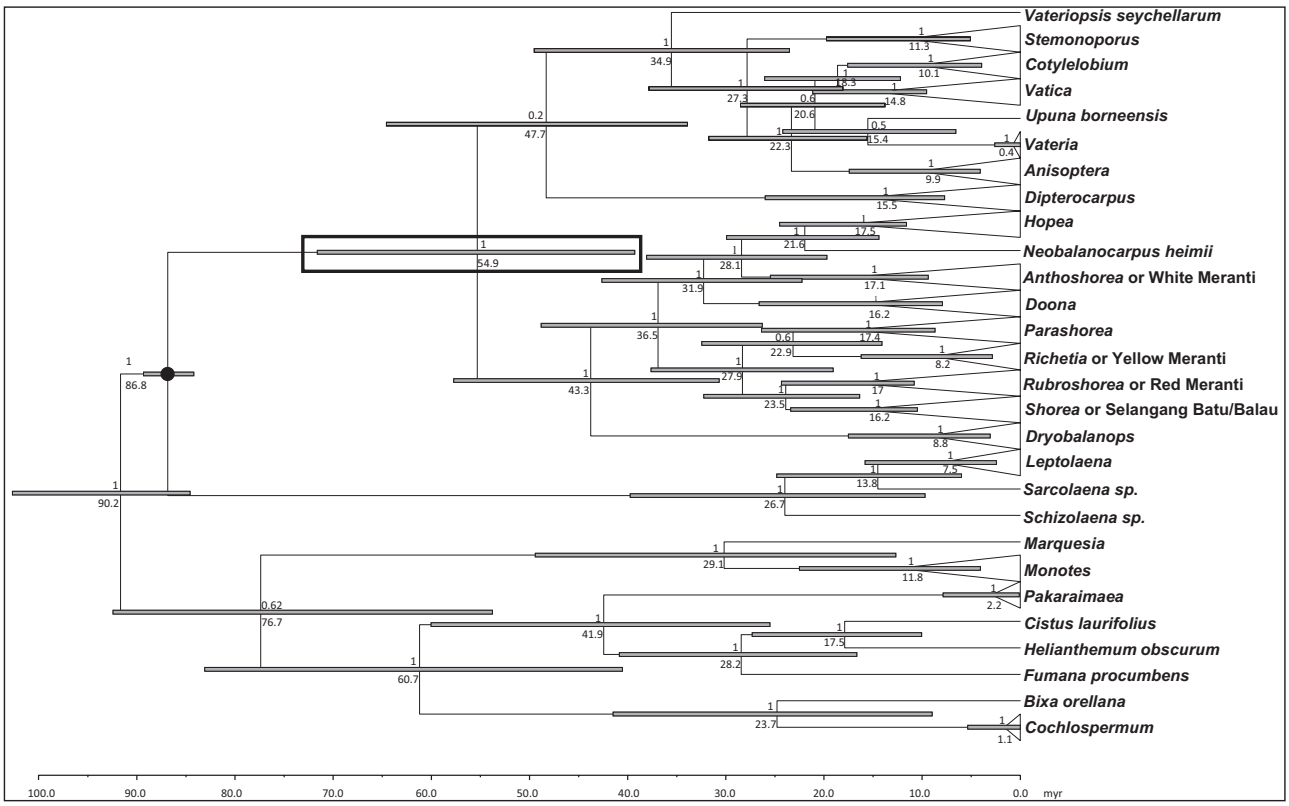


Figure 5. Dated maximum clade credibility tree obtained from BEAST analysis. Taxa are collapsed to major clades. The node that was calibrated is marked with a black dot. Grey node bars represent the 95% highest posterior density interval. Posterior probabilities are given above each node and the mean age estimates are shown below each node. The node in the black box shows the age estimates for Dipterocarpoideae. Geological time scale is given in millions of years.

size of Dipterocarpaceae was first reported from a diploid *Shorea robusta* C.F.Gaertn. (2C = 1.15 pg; Ohri & Kumar, 1986), and insights into evolution of genome size in Dipterocarpaceae were recently reported by Ng *et al.* (2016). Genome sizes of 20 individuals representing 15 species in six genera were obtained in this study (Table 6, Fig. 4) and ranged from 1C = 0.3264 pg in *Shorea roxburghii* to 0.6724 pg in *Vatica diospyroides*. Genome size variation was observed between and within genera, corresponding well to the results of Ng *et al.* (2016; Table 6). Moreover, genome size variation was observed within species, e.g. in *Hopea odorata* (1C = 0.4216, 0.6051 and 0.6094 pg). Dipterocarpaceae have relative small genome sizes, corresponding to previous observations of small genome sizes in woody angiosperms that are hypothesized to rarely experience polyploidization (Ohri, 2005; Chen *et al.*, 2014). Compared to closely related families (Bennett & Leitch, 2012), genome size in dipterocarps was smaller than those in Cistaceae (median 1C = 2.53 and 0.88–4.50 pg, respectively), but larger than those in Bixaceae (1C = 0.20 pg). Although negative correlations have been observed between genome size and species richness (e.g. Vinogradov, 2004; Knight *et al.*, 2005), in their study Ng *et al.* (2016) argued that excluding any correlation between the high species diversity of Dipterocarpoideae and their small genome size is premature, and further studies are needed.

CONCLUSIONS

Several molecular and many morphological studies on Dipterocarpaceae have been conducted in the past. Here, we present the first molecular phylogenetic study including all three subfamilies of Dipterocarpaceae and closely related families. In our study, there are conflicts between molecular results and the distribution of some of the intuitively selected morphological characters that in the past have been the basis of previous classifications. Broad and critical observations on well-defined morphological characters are important for classical taxonomy, but ultimately such decisions should be taken on the bases of all data, not just a set of intuitively selected characters that are thought to be more reliable than others. For example, Ashton's

circumscription of *Shorea* was based on a single character, the number of long versus short fruit sepals. However, many *Shorea* spp. only have short subequal fruit sepals. This concept is further complicated by the fact that *Parashorea* also has unequal fruit sepals, which could be interpreted as three long and two short as in *Shorea*. Our molecular results were not supported well enough to resolve the 11 sections in *Shorea* proposed by Ashton on the basis of morphological characters. We therefore assume that next-generation techniques, such as restriction-site associated sequencing (RADseq), which allows sampling of genome-wide single nucleotide polymorphisms, could give better resolution at the species level in *Shorea* and be able to detect instances of hybridization, which have been suggested in some previous studies (e.g. AFLP; Cao *et al.*, 2006). To conclude, our study strengthens the phylogenetic hypotheses for the larger clade to which Dipterocarpaceae are related (*Pakaraimaea* + Cistaceae) (Sarcolaenaceae + Monotoideae + Dipterocarpoideae). Nevertheless, there are still some relationships between (Sarcolaenaceae + Monotoideae + Dipterocarpoideae) that still need to be clarified. This paper clearly demonstrates that morphological and molecular evidence are both important, although there are still some discrepancies between them that need to be better addressed in future research.

ACKNOWLEDGEMENTS

The authors wish to thank all the people and institutions who provided the material needed for this study, especially the Smithsonian Institution (NMNH Biorepository) for providing the silica-gel dried specimens of *Pakaraimaea* collected by Kenneth Wurdack, Eric Feltz for *Monotes* from the Missouri Botanical Garden (USA), Sutee Duangjai of the Faculty of Forestry, Kasetsart University, the Royal Botanic Gardens, Kew (UK) and the Royal Botanical Gardens, Peradeniya (Sri Lanka), for DNA/seeds/leaf material, without which we could not have done this study. We greatly appreciate and thank Dayanandan Selvadurai of the University of Concordia for providing one of the *Pakaraimaea* sequences and Ovidiu Paun for helping with analyses. Fieldwork was done in a research plot in Kuala Belalong Field Study Centre (KBFSC) with collaboration of University of Brunei Darussalam (UBD). The plot is part of a global network of large-scale demographic tree plots, established by UBD in collaboration with the Centre for Tropical Forest Science (CTFS) of the Smithsonian Tropical Research Institute, USA. Field assistants Fiona Willinathy, Anak Amdani, Sawai Anak Amba and Teddy Chua of the KBFS, Brunei, are acknowledged for their support during our fieldwork. CTFS is acknowledged for all information on the plots. We thank both Brunei Heart of Borneo Secretariat and the Forest Department of Sri Lanka for granting permission to export material for research purposes. Verena Klejna and Elfriede Grasserbauer helped in the laboratory with DNA extraction and sequencing. We thank the two reviewers for their helpful comments and suggestions. This work was supported by the Austrian Science Fund FWF (grant P26548-B22 to R.S.).

REFERENCES

Alexander IJ, Högberg P. 1986. Ectomycorrhizas of tropical angiospermous trees. *New Phytologist* **102:** 541–549.

Alverson WS, Karol KG, Baum DA, Chase MW, Swensen SM, McCourt R, Sytsma KJ. 1998. Circumscription of the Malvales and relationships to other Rosidae: evidence from *rbcL* sequence data. *American Journal of Botany* **85:** 876–887.

Appanah S, Chan HT. 1981. Thrips: the pollinators of some dipterocarps. *Malaysian Forester* **44:** 234–252.

APG. 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* **85:** 531–553.

APG II. 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* **141:** 399–436.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161:** 105–121.

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181:** 1–20.

Arrington JM, Kubitzki K. 2003. Cistaceae. In: Kubitzki K, Bayer C, eds. *The families and genera of flowering plants, Vol. 5*. New York: Springer, 62–70.

Ashton PS. 1964. *A manual of the dipterocarp trees of Brunei State.* Oxford: Oxford University Press.

Ashton PS. 1968. *A manual of the dipterocarp trees of Brunei State and of Sarawak - supplement*. Hong Kong: Borneo Literature Bureau for Sarawak, Forest Department.

Ashton 1979. Final discussion. In: Maury-Lechon G, ed. *Dipterocarpaceae: taxonomie-phylogénie-ecologie.* Paris: Memoires du Museum National d'Histoire Naturelle, serie B, Botanique 26, Editions du Museum, 159.

Ashton PS. 1980. Dipterocarpaceae. In: Dassanayake MD, Fosberg FR, eds. *A revised handbook to the flora of Ceylon I.* Washington: Smithsonian Institution, 364–423.

Ashton PS. 1982. Dipterocarpaceae. In: Van Steenis CGGJ, ed. *Flora Malesiana, series 1, Spermatophyta, Vol. 9*. The Hague: Nijhoff, 237–552.

Ashton PS, Gunatilleke CVS. 1987. New light on the plant geography of Ceylon I. Historical plant geography. *Journal of Biogeography* **14:** 249–285.

Ashton PS. 1988. Dipterocarp biology as a window to the understanding of tropical forest structure. *Annual Review of Ecology and Systematics* **19:** 347–370.

Ashton PS, Givnish TJ, Appanah S. 1988. Staggered flowering in the Dipterocarpaceae: new insights into floral induction and the evolution of mast fruiting in the aseasonal tropics. *American Naturalist* **132:** 44–66.

Ashton PS. 2003. Dipterocarpaceae. In: Kubitzki K, Bayer C, eds. *The families and genera of vascular plants, Vol. 5*. New York: Springer, 182–197.

Ashton PS. 2014. *On the forests of tropical Asia, lest the memory fade.* Richmond: Kew Publishing.

Aubriot X, Soulebeau A, Haevermans T, Schatz GE, Cruaud C, Lowry PP. 2016. Molecular phylogenetics of Sarcolaenaceae (Malvales), Madagascar's largest endemic plant family. *Botanical Journal of the Linnean Society* **182:** 729–743.

Bancroft H. 1935. Some fossil dicotyledonous wood from Mount Elgon, East Africa.I. *American Journal of Botany* **22:** 164–183.

Barfuss MHJ, Till W, Leme EMC, Pinzón JP, Manzanares JM, Halbritter H, Samuel R, Brown GK. 2016. Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa* **279:** 1–97.

Bawa KS. 1998. Conservation of genetic resources in the Dipterocarpaceae. In: Appanah S, Turnbull JM, eds. *A review of dipterocarps, taxonomy, ecology and silviculture.* Bogor: Center for Forest Research Institute, 45–56.

Bayer C. 2003. Sarcolaenaceae. In: K. Kubitzki, C. Bayer, eds. *The families and genera of vascular plants, Vol. 5*. New York: Springer, 345–352.

Bennett MD, Leitch IJ. 2012. *Angiosperm DNA C-values database.* Release 8.0, December 2012. Available at: http://data.kew.org/cvalues/ accessed 14 July 2017.

Bossuyt F, Milinkovitch MC. 2001. Amphibians as indicators of early Tertiary 'out-of-India' dispersal of vertebrates. *Science* **292:** 93–95.

Brearley FQ. 2012. Ectomycorrhizal associations of the Dipterocarpaceae. *Biotropica* **44:** 637–648.

Bremer B, Manen JF. 2000. Phylogeny and classification of the subfamily Rubioideae (Rubiaceae). *Plant Systematics and Evolution* **225:** 43–72.

Cao CP, Gailing O, Siregar I, Indrioko S, Finkeldey R. 2006. Genetic variation at AFLPs for the Dipterocarpaceae and its relation to molecular phylogenies and taxonomic subdivisions. *Journal of Plant Research* **119:** 553–558.

Capuron R. 1970. Observations sur les Sarcolaenaceés. *Adansonia II* **10:** 247–265.

Castello LV, Barfuss MHJ, Till W, Galetto L, Chiapella JO. 2016. Disentangling the *Tillandsia capillaris* complex: phylogenetic relationships and taxon boundaries in Andean populations. *Botanical Journal of the Linnean Society* **181:** 391–414.

Chase MW, Hills HG. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* **40:** 215–220.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu YL, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedrén M, Gaut BS, Jansen RK, Kim KJ, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang QY, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Graham SW, Barrett SCH, Dayanandan S, Albert VA. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* **80:** 528–580.

Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, Van den Berg C, Schuiteman A. 2015. An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society* **177:** 151–174.

Chen SC, Cannon CH, Kua CS, Liu JJ, Galbraith DW. 2014. Genome size variation in the Fagaceae and its implications for trees. *Tree Genetics & Genomes* **10:** 977–988.

Choong CY, Wickneswari R, Norwati M, Abbott RJ. 2008. Phylogeny of *Hopea* (Dipterocarpaceae) inferred from chloroplast DNA and nuclear PgiC sequences. *Molecular Phylogenetics and Evolution* **48:** 1238–1243.

Coetzee JA, Muller J. 1984. The phytogeographic significance of some extinct Gondwana pollen types from the Tertiary of the southwestern Cape (South Africa). *Annals of the Missouri Botanical Garden* **71:** 1088–1099.

Collier JS, Sansom V, Ishizuka O, Taylor RN, Minshull TA, Whitmarsh RB. 2008. Age of Seychelles–India breakup. *Earth and Planetary Science Letters* **272:** 264–277.

Corlett RT. 2004. Flower visitors and pollination in the Oriental (Indomalayan) Region. *Biological Reviews* **79:** 497–532.

Croizat L. 1952. *Manual of phytogeography.* The Hague: published by the author.

Croizat L. 1964. Thoughts on high systematics, phylogeny and floral morphology with a note on the origin of Angiospermae. *Candollea* **19:** 17–96.

Cronquist A. 1981. *An integrated system of classification of flowering plants.* New York: Columbia University Press.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Dayanandan S, Ashton PS, Williams SM, Primack RB. 1999. Phylogeny of the tropical tree family Dipterocarpaceae based on nucleotide sequences of the chloroplast *rbcL* gene. *American Journal of Botany* **86:** 1182–1190.

De Zeeuw C. 1977. Stem anatomy. *Taxon* **26:** 368–380.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Ducousso M, Béna G, Bourgeois C, Buyck B, Eyssartier G, Vincelette M, Rabevohitra R, Randrihasipara L, Dreyfus B, Prin Y. 2004. The last common ancestor of Sarcolaenaceae and Asian dipterocarp trees was ectomycorrhizal before the India–Madagascar separation, about 88 million years ago. *Molecular Ecology* **13:** 231–236.

Duangjai S, Samuel R, Munzinger J, Forest F, Wallnöfer B, Barfuss MH, Fischer G, Chase MW. 2009. A multilocus plastid phylogenetic analysis of the pantropical genus Diospyros (Ebenaceae), with an emphasis on the radiation

and biogeographic origins of the New Caledonian endemic species. *Molecular Phylogenetics and Evolution* **52:** 602–620.

Dutta S, Mallick M, Bertram N, Greenwood PF, Mathews RP. 2009. Terpenoid composition and class of Tertiary resins from India. *International Journal of Coal Geology* **80:** 44–50.

Dutta S, Tripathi SM, Mallick M, Mathews RP, Greenwood PF, Rao MR, Summons RE. 2011. Eocene out-of-India dispersal of Asian dipterocarps. *Review of Palaeobotany and Palynology* **166:** 63–68.

Fay MF, Swensen SM, Chase MW. 1997. Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae). *Kew Bulletin* **52:** 111–120.

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

Feng X, Tang B, Kodrul TM, Jin J. 2013. Winged fruits and associated leaves of *Shorea* (Dipterocarpaceae) from the Late Eocene of South China and their phytogeographic and paleoclimatic implications. *American Journal of Botany* **100:** 574–581.

Forest F. 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. *Annals of Botany* **104:** 789–794.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220:** 1049–1051.

Gamage DT, de Silva MP, Yoshida A, Szmidt AE, Yamazaki T. 2003. Molecular phylogeny of Sri Lankan Dipterocarpaceae in relation to other Asian Dipterocarpaceae based on chloroplast DNA sequences. *Tropics* **13:** 79–87.

Gamage DT, de Silva MP, Inomata N, Yamazaki T, Szmidt AE. 2006. Comprehensive molecular phylogeny of the sub-family Dipterocarpoideae (Dipterocarpaceae) based on chloroplast DNA sequences. *Genes & Genetic Systems* **81:** 1–12.

Ghazoul J. 1997. The pollination and breeding system of *Dipterocarpus obtusifolius* (Dipterocarpaceae) in dry deciduous forests of Thailand. *Journal of Natural History* **31:** 901–916.

Gilg E. 1925. Dipterocarpaceae. In: Engler A, Prantl K, eds. *Die natürlichen Pflanzenfamilien*, *2nd edn.* Leipzig: Engelmann, 237–269.

Gottwald H, Parameswaran N. 1966. Das sekundäre Xylem der Familie Dipterocarpaceae, anatomische Untersuchungen zur Taxonomie und Phylogenie. *Botanische Jahrbücher* **5:** 410–508.

Greilhuber J, Ebert I. 1994. Genome size variation in *Pisum sativum*. *Genome* **37:** 645–655.

Guindon S, Gascuel O. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* **52:** 696–704.

Guzmán B, Vargas P. 2009. Historical biogeography and character evolution of Cistaceae (Malvales) based on analysis of plastid *rbcL* and *trnL-trnF* sequences. *Organisms Diversity & Evolution* **9:** 83–99.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Harrison RD, Nagamitsu T, Momose K, Inoue T. 2005. Flowering phenology and pollination of *Dipterocarpus* (Dipterocarpaceae) in Borneo. *Malayan Nature Journal* **57:** 67–80.

Heckenhauer J, Barfuss MHJ, Samuel R. 2016. Universal multiplexable *matK* primers for DNA barcoding of angiosperms. *Applications in Plant Sciences* **4:** 1500137.

Högberg P. 1982. Mycorrhizal associations in some woodland and forest trees and shrubs in Tanzania. *New Phytologist* **92:** 407–415.

Högberg P, Piearce GD. 1986. Mycorrhizas in Zambian trees in relation to host taxonomy, vegetation type and successional patterns. *Journal of Ecology* **74:** 775–785.

Horn JW, Wurdack KJ, Dorr LJ. 2016. Phylogeny and diversification of Malvales. Botany (abstract). Available at: http://2016.botanyconference.org/engine/search/index.php?func=detail&aid=919 accessed 14 July 2017.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Indrioko S, Gailing O, Finkeldey R. 2006. Molecular phylogeny of Dipterocarpaceae in Indonesia based on chloroplast DNA. *Plant Systematics and Evolution* **261:** 99–115.

Jang TS, Emadzade K, Parker J, Temsch EM, Leitch AR, Speta F, Weiss-Schneeweiss H. 2013. Chromosomal diversification and karyotype evolution of diploids in the cytologically diverse genus *Prospero* (Hyacinthaceae). *BMC Evolutionary Biology* **13:** 136.

Jong K, Lethbridge A. 1967. Cytological studies in the Dipterocarpaceae, I. Chromosome numbers of certain Malaysian genera. *Notes from the Royal Botanical Garden, Edinburgh* **27:** 175–184.

Jong K, Kaur A. 1979. Cytotaxonomic view of Dipterocarpaceae with some comments on polyploidy with apomixis. In: Maury-Lechon G, ed. *Dipterocarpaceae: taxonomie-phylogénie-ecologie.* Paris: Memoires du Museum National d'Histoire Naturelle, serie B, Botanique 26, Editions du Museum, 41–49.

Khatua AK, Chakrabarti S, Mallick N. 1998. Abundance, activity and diversity of insects associated with flower of sal (*Shorea robusta*) in Midnapore, (Arabari) West Bengal, India. *The Indian Forester* **124:** 62–74.

Kajita T, Kamiya K, Nakamura K, Tachida H, Wickneswari R, Tsumura Y, Yoshimaru H, Yamazaki T. 1998. Molecular phylogeny of Dipetrocarpaceae in Southeast Asia based on nucleotide sequences of *matK*, *trnL* intron, and *trnL-trnF* intergenic spacer region in chloroplast DNA. *Molecular Phylogenetics and Evolution* **10:** 202–209.

Kamiya K, Harada K, Ogino K, Kajita T, Yamazaki T, Lee HS, Ashtion PS. 1998. Molecular phylogeny of dipterocarp species using nucleotide sequences of two non-coding regions in chloroplast DNA. *Tropics* **7:** 195–207.

Kamiya K, Harada K, Tachida H, Ashton PS. 2005. Phylogeny of *PgiC* gene in *Shorea* and its closely related genera (Dipterocarpaceae), the dominant trees in Southeast Asian tropical rain forests. *American Journal of Botany* **92:** 775–788.

Kaur A, Jong K, Sands VE, Soepadmo E. 1986. Cytoembryology of some Malaysian dipterocarps, with some evidence of apomixis. *Botanical Journal of the Linnean Society* **92:** 75–88.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647–1649.

Knight CA, Molinari NA, Petrov DA. 2005. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* **95:** 177–190.

Kondo T, Nishimura S, Tani N, Ng KKS, Lee SL, Muhammad N, Okuda T, Tsumura Y, Isagi Y. 2016. Complex pollination of a tropical Asian rainforest canopy tree by flower-feeding thrips and thrips-feeding predators. *American Journal of Botany* **103:** 1912–1920.

Kubitzki K, Chase MW. 2003. Introduction to Malvales. In: Kubitzki K, Bayer C, eds. *The families and genera of flowering plants, Vol. 5*. New York: Springer, 12–16.

Lee SS. 1990. The mycorrhizal association of the Dipterocarpaceae in the tropical rain forests of Malaysia. *Ambio* **19:** 383–385.

Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA, Sytsma KJ. 2003. Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. *American Journal of Botany* **90:** 107–115.

Maguire BPC, Ashton PS. 1977. Pakaraimoideae, Dipterocarpaceae of the Western Hemisphere II. Systematic, geographic and phyletic considerations. *Taxon* **26:** 341–385.

Malloch DW, Pirozynski KA, Raven PH. 1980. Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (a review). *Proccedings of the National Acadamy of Sciences of the United States of America* **77:** 2113–2118.

Maury G. 1978. *Diptérocarpacées: du fruit à la plantule.* [3 vols: IA, IB, II]. D. Phil. Thesis, University Toulouse.

Maury-Lechon G. 1979a. Conséquences taxonomiques de l'étude des caractères des fruits/germinations, embryons et plantules des Diptérocarpacées. In: Maury-Lechon G, ed. *Dipterocarpaceae: taxonomie-phylogénie-ecologie.* Paris: Memoires du Museum National d'Histoire Naturelle, serie B, Botanique, Editions du Museum, 81–106.

Maury-Lechon G. 1979b. Interprétation phylogénique des caractères des pollens, fruits germinations et plantules des Diptérocarpacées. In: Maury-Lechon G, ed. *Dipterocarpaceae: taxonomie-phylogénie-ecologie.* Paris: Memoires du Museum National d'Histoire Naturelle, serie B, Botanique, Editions du Museum, 139–144.

Maury-Lechon G, Curtet L. 1998. Biogeography and evolutionary systematics of family Dipterocarpaceae. In: Appanah S, Turnvull JM, eds. *A review of dipterocarps, taxonomy, ecology and silviculture.* Bogor: Center for Forest Research Institute, 5–44.

Meher-Homji VM. 1979. Distribution of Dipterocarpaceae: some phytogeographic considerations on India. *Phytocoenologia* **6:** 85–93.

Mehra PN. 1976. *Cytology of Himalayan hardwoods*. Calcutta: Sree Saraswaty Press.

Meijer W, Wood GHS. 1964. Dipterocarps of Sabah (North Borneo). In: Forest Department Sandakan, ed. *Sabah Forest Record* **5:** 1–344.

Meijer W, Wood GHS. 1976. *Keys to dipterocarps on Sabah.* Bogor: Biotrop.

Meijer W. 1979. Taxomonic studies in the genus *Dipterocarpus.* In: Maury-Lechon G. (ed.) *Diptérocarpacées: taxonomie-phylogénie-ecologie.* Paris: Mémoires du Muséum National d'Histoire Naturelle: First International Round Table on Dipterocarpaceae, Série B, Botanique 26, Editions du Muséum, 50–56.

Merrill ED. 1923. Distribution of the Dipterocarpaceae. *Philippine Journal of Science* **23:** 1–32.

Momose K, Nagamitsu T, Inoue T. 1996. The reproductive ecology of an emergent dipterocarp in a lowland rain forest in Sarawak. *Plant Species Biology* **11:** 189–198.

Momose K, Yumoto T, Nagamitsu T, Kato M, Nagamasu H, Sakai S, Harrison RD, Itioka T, Hamid AA, Inoue T. 1998. Pollination biology in a lowland dipterocarp forest in Sarawak, Malaysia. I. Characteristics of the plant-pollinator community in a lowland dipterocarp forest. *American Journal of Botany* **85:** 1477–1501.

Morley RJ. 2000. Distributions of palms, oaks and dipterocarps. In: Morley RJ (ed.) *Origin and evolution of tropical rain forests.* Chichester: Wiley, 275–278.

Moyersoen B. 2006. *Pakaraimaea dipterocarpacea* is ectomycorrhizal, indicating an ancient Gondwanaland origin for the ectomycorrhizal habit in Dipterocarpaceae. *New Phytologist* **172:** 753–762.

Nagamitsu T, Harrison RD, Inoue T. 1999. Beetle pollination of *Vatica parvifolia* (Dipterocarpaceae) in Sarawak, Malaysia. *Gardens' Bulletin Singapore* **51:** 43–54.

Nandi OI. 1998. Ovule and seed anatomy of Cistaceae and related Malvanae. *Plant Systematics and Evolution* **209:** 239–264.

Nandi OI, Chase MW, Endress PK. 1998. A combined cladistics analysis of angiosperms using *rbcL* and non-molecular data sets. *Annals of the Missouri Botanical Garden* **85:** 137–214.

Ng CH, Lee SL, Tnah LH, Ng KKS, Lee CT, Madon M. 2016. Genome size variation and evolution in Dipterocarpaceae. *Plant Ecology & Diversity* **9:** 437–446.

Nilsson S, Coetzee J, Grafström E. 1996. On the origin of the Sarcolaenaceae with reference to pollen morphological evidence. *Grana* **35:** 321–334.

Ohri D, Kumar A. 1986. Nuclear DNA amounts in some tropical hardwoods. *Caryologia* **39:** 303–307.

Ohri D. 2005. Climate and growth form: the consequences for genome size in plants. *Plant Biology* **7:** 449–458.

Oginuma K, Kiaptranis R, Damas K, Tobe H. 1998. A cytological study of some plants from Papua New Guinea. *Acta Phytotaxonomica et Geobotanica* **49:** 105–114.

Otto F, Oldiges H, Göhde W, Jain VK. 1981. Flow cytometric measurement of nuclear DNA content variations as a potential *in vivo* mutagenicity test. *Cytometry* **2:**189–191.

Pal JK, Mandal S, Bhattacharya GN 1993. Cytopalynological studies in *Shorea robusta* Gaertn. f. *Science and Culture* **59:** 55–57.

Pancho JV. 1971. In IOPB chromosome number reports XXXIV. *Taxon* **20:** 785–797.

Phosri C, Põlme S, Taylor AFS, Kõljalg U, Suwannasai N, Tedersoo L. 2012. Diversity and community composition of ectomycorrhizal fungi in a dry deciduous dipterocarp forest in Thailand. *Biodiversity and Conservation* **21:** 2287–2298.

Prakash U. 1972. Palaeoenvironmental analysis of Indian tertiary floras. *Geophytology* **2:** 178–205.

Proctor MCF. 1978. Cistaceae. In Heywood VH, ed. *Flowering plants of the world*. Oxford: Oxford University Press, 108–100.

Rath P, Rajaseger G, Goh CJ, Kumar PP. 1998. Phylogenetic analysis of dipterocarps using random amplified polymorphic DNA markers. *Annals of Botany* **82:** 61–65.

Raven PH, Axelrod DI. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* **61:** 539–673.

Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O, Mayrose I. 2015. The chromosome counts database (CCDB) – a community resource of plant chromosome numbers. *New Phytologist* **206:** 19–26.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:**1572–1574.

Roy RP, Jha RP. 1965. Cytological studies in Dipterocarpaceae I. *The Journal of the Indian Botanical Society* **44:** 387–397.

Russell A, Samuel R, Rupp B, Barfuss MHJ, Šafran M, Besendorfer V, Chase MW. 2010. Phylogenetics and cytology of a pantropical orchid genus *Polystachya* (Polystachyinae, Vandeae, Orchidaceae): evidence from plastid DNA sequence data. *Taxon* **59:** 389–404.

Rust J, Singh H, Rana RS, McCann T, Singh L, Anderson K, Sarkar N, Nascimbene PC, Stebner F, Thomas JC, Solórzano Kraemer M, Wiliams CJ, Engel MS, Sahni A, Grimaldi D. 2010. Biogeographic and evolutionary implications of a diverse paleobiota in amber from the early Eocene of India. *Proceedings of the National Academy of Sciences USA* **107:** 18360–18365.

Sakai S, Momose K, Yumoto T, Kato M, Inoue T. 1999. Beetle pollination of *Shorea parvifolia* (section *Mutica*, Dipterocarpaceae) in a general flowering period in Sarawak, Malaysia. *American Journal of Botany* **86:** 62–69.

Samuel R, Stuessy TF, Tremetsberger K, Baeza CM, Siljak-Yakovlev S. 2003. Phylogentic relationships among species of *Hypochaeris* (Asteraceae, Lactuceae) based on ITS, plastid *trnL* intron, *trnl-F* spacer and *matK* sequences. *American Journal of Botany* **90:** 496–507.

Sarkar AK, Datta N, Chatterjee U, Hazra D. 1982. IOPB chromosome number reports LXXVI. *Taxon* **31:** 576–579.

Sato H, Tanabe AS, Toju H. 2015. Contrasting diversity and host association of ectomycorrhizal basidiomycetes versus root-associated ascomycetes in a dipterocarp rainforest. *PLoS One* **10:** e0125550.

Savolainen V, Fay MF, Albach DC, Backlund A, Van der Bank M, Cameron KM, Johnson SA, Lledó MD, Pintaud JC, Powell M, Sheahan MC, Soltis DE, Soltis PS, Weston P, Whitten WM, Wurdack KJ, Chase MW. 2000. Phylogeny of the eudicots: a nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bulletin* **55:** 257–309.

Shukla A, Mehrotra RC, Guleria JS. 2013. Emergence and extinction of Dipterocarpaceae in western India with reference to climate change: fossil wood evidence. *Journal of Earth System Science* **122:** 1373–1386.

Smith SE, Read DJ. 1997. *Mycorrhizal symbiosis.* London: Academic Press.

Smits WTM. 1994. *Dipterocarpaceae: mycorrhizae and regeneration.* Wageningen: Tropenbos.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swenson SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, *atpB* sequences. *Botanical Journal of the Linnean Society* **133:** 381–461.

Souza HAV, Muller LAC, Brandão RL, Lovato MB. 2012. Isolation of high quality and polysaccharide-free DNA from leaves of *Dimorphandra mollis* (Leguminosae), a tree from the Brazilian cerrado. *Genetics and Molecular Research* **11:** 756–764.

Stamatakis A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.

Swofford DL. 2016. *PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4.0a149*. Sunderland: Sinauer Associates.

Symington CF. 1943. *Foresters' manual of dipterocarps.* Kuala Lumpur: Syonan-Hakubutukan.

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

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E. 2007. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* **35:** e14.

Talip N. 2008. Systematic significance of pollen morphology of *Shorea*, *Hopea*, *Parashorea* and *Neobalanocarpus* (Dipterocarpaceae) in Malaysia. *Sains Malaysiana* **37:**169–176.

Temsch EM, Greilhuber J, Krisai R. 2010. Genome size in liverworts. *Preslia* **82:** 63–80.

Tedersoo L, Suvi T, Beaver K, Kõljag U. 2007. Ectomycorrhizal fungi of the Seychelles: diversity patterns and host shifts from the native *Vateriopsis seychellarum* (Dipterocarpaceae) and *Intsia bijuga* (Caesalpiniaceae) to the introduced *Eucalyptus robusta* (Myrtaceae), but not *Pinus caribea* (Pinaceae). *New Phytologist* **175:** 321–333.

Tixier P. 1953. Données cytologiques sur quelques Guttiferales du Viet-Nam. *Revue Cytolologique Biologique Végetal* **14:**1–12.

Tixier P. 1960. Données cytologiques sur quelques Guttiferales au Laos. *Revue Cytolologique Biologique Végetal* **22:** 65–70.

Tsumura Y, Kawahara T, Wickneswari R, Yoshimura K. 1996. Molecular phylogeny of Dipterocarpaceae in Southeast Asia using RFLP of PCR-amplified chloroplast genes. *Theoretical and Applied Genetics* **93:** 22–29.

Tsumura Y, Kado T, Yoshida K, Abe H, Ohtani M, Taguchi Y, Fukue Y, Tani N, Ueno S, Yoshimura K, Kamiya K,

Harada K, Takeuchi Y, Diway B, Finkeldey R, Na'iem M, Indrioko S, Ng KK, Muhammad N, Lee SL. 2011. Molecular database for classifying *Shorea* species (Dipterocarpaceae) and techniques for checking the legitimacy of timber and wood products. *Journal of Plant Research* **124:** 35–48.

Vinogradov AE. 2004. Genome size and extinction risk in vertebrates. *Proceedings of the Royal Society* B: Biological Sciences **271:** 1701–1705.

Whitmore TC, 1963. Studies in systematic bark morphology. III. Bark taxonomy in Dipterocarpaceae. *Gardens' Bulletin Singapore* **19:** 321–371.

Wikström N, Savolainen V, Chase MW. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London B* **268:** 2211–2220.

Yulita KS, Bayer RJ, West JG. 2005. Molecular phylogenetic study of *Hopea* and *Shorea* (Dipterocarpaceae): evidence from the *trnL–trnF* and internal transcribed spacer regions. *Plant Species Biology* **20:** 167–182.

Yulita KS. 2013. Secondary structures of chloroplast *trnL* intron in Dipterocarpaceae and its implications for the phylogenetic reconstruction. *Hayati Journal of Biosciences* **20:** 31–39.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Specimens used in this study. The collection number, herbarium voucher, and location is given. GenBank accession numbers of species used for phylogenetic analysis is stated.