



Molecular phylogenetics of Sarcolaenaceae (Malvales), Madagascar's largest endemic plant family

XAVIER AUBRIOT^{1,2,*†}, ANAËLLE SOULEBEAU^{1†}, THOMAS HAEVERMANS¹, GEORGE E. SCHATZ³, CORINNE CRUAUD⁴ and PORTER P. LOWRY II^{1,3*}

¹Institut de Systématique, Évolution, et Biodiversité (ISYEB), Unité Mixte de Recherche 7205, Centre National de la Recherche Scientifique/Muséum National d'Histoire Naturelle/École Pratique des Hautes Études/Université Pierre et Marie Curie, Sorbonne Universités, CP 39, 57 rue Cuvier, 75231 Paris CEDEX 05, France

²Department of Life Sciences, Natural History Museum, Cromwell Road, London, SW7 5BD, UK

³Missouri Botanical Garden, P.O. Box 299, St. Louis, MO, 63166-0299, USA

⁴Genoscope, Centre National de Séquençage, CP 5706, 2 rue Gaston Crémieux, 91057 Evry CEDEX, France

Received 26 April 2016; revised 3 July 2016; accepted for publication 27 July 2016

With 72 species belonging to ten genera, Sarcolaenaceae are the largest and most diverse of Madagascar's endemic plant families. Comprising shrubs and trees, with members found in nearly all of this island nation's biogeographic regions, they are characterised by the presence of a distinctive extra-floral involucre that is more or less accrescent, partially or completely covering or enveloping the fruit. We present the first molecular phylogenetic study of Sarcolaenaceae, using broad sampling that encompasses the family's taxonomic and morphological diversity, including 46 species representing all ten genera and sequence data from one nuclear marker (ITS) and three plastid regions (*psaA-ORF170*, *psaA-trnH* and *rbcL*), to reconstruct phylogenetic relationships using Bayesian inference and maximum likelihood. Results confirm the monophyly of Sarcolaenaceae and of eight of the ten genera; the monophyly of *Rhodolaena* remains ambiguous, although the taxa sampled were recovered in two well supported clades that are coherent in terms of morphology and geography. Only a single species of *Eremolaena* was available for study. The phylogenetic backbone of Sarcolaenaceae is not fully resolved, making it difficult to identify potential morphological synapomorphies or ecological preferences between and within genera. In the family, two monophyletic groups were, however, found [*Pentachlaena* + *Eremolaena* + *Perrierodendron* (Clade A) and *Xyloolaena* + *Leptolaena* + *Sarcolaena* + *Mediusella* + *Xerochlamys* (Clade B)] that are consistent with previous results based on morphology. Expanded species sampling and data from additional, more quickly evolving markers will be needed to produce a fully resolved phylogenetic tree for Sarcolaenaceae, which could then serve as a basis for exploring macroevolutionary patterns and processes in this remarkable family and reconstructing its biogeographic history. © 2016 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2016, 182, 729–743

ADDITIONAL KEYWORDS: angiocarpy – generic delimitation.

INTRODUCTION

Sarcolaenaceae Caruel, the largest of Madagascar's endemic plant families, have been the focus of several morphological studies, in particular regarding the peculiar pollen tetrads produced by all genera (Straka, 1963, 1964, 1971; Carlquist, 1964; Straka & Friedrich,

1983; Nilsson & Randrianasolo, 1999). A recent series of taxonomic revisions has clarified species circumscriptions and described new taxa in each of the ten genera now recognised (Lowry *et al.*, 1999, 2000, 2014; Schatz, Lowry & Wolf, 2000, 2001; Lowry, Schatz & Wolf, 2002; Lowry & Rabehevitra, 2006; Hong-Wa, 2009; Rabehevitra & Lowry, 2009; Ramananjanahary *et al.*, 2010; Andriamihajarivo, Lowry & Schatz, 2016). While this emblematic family has attracted the attention of botanists for more than two centuries (Du Petit-Thouars, 1806; Baillon, 1872a,b, 1879, 1884;

*Corresponding authors. E-mails: x.aubriot@gmail.com; Pete.Lowry@mobot.org

†These authors contributed equally to the study.

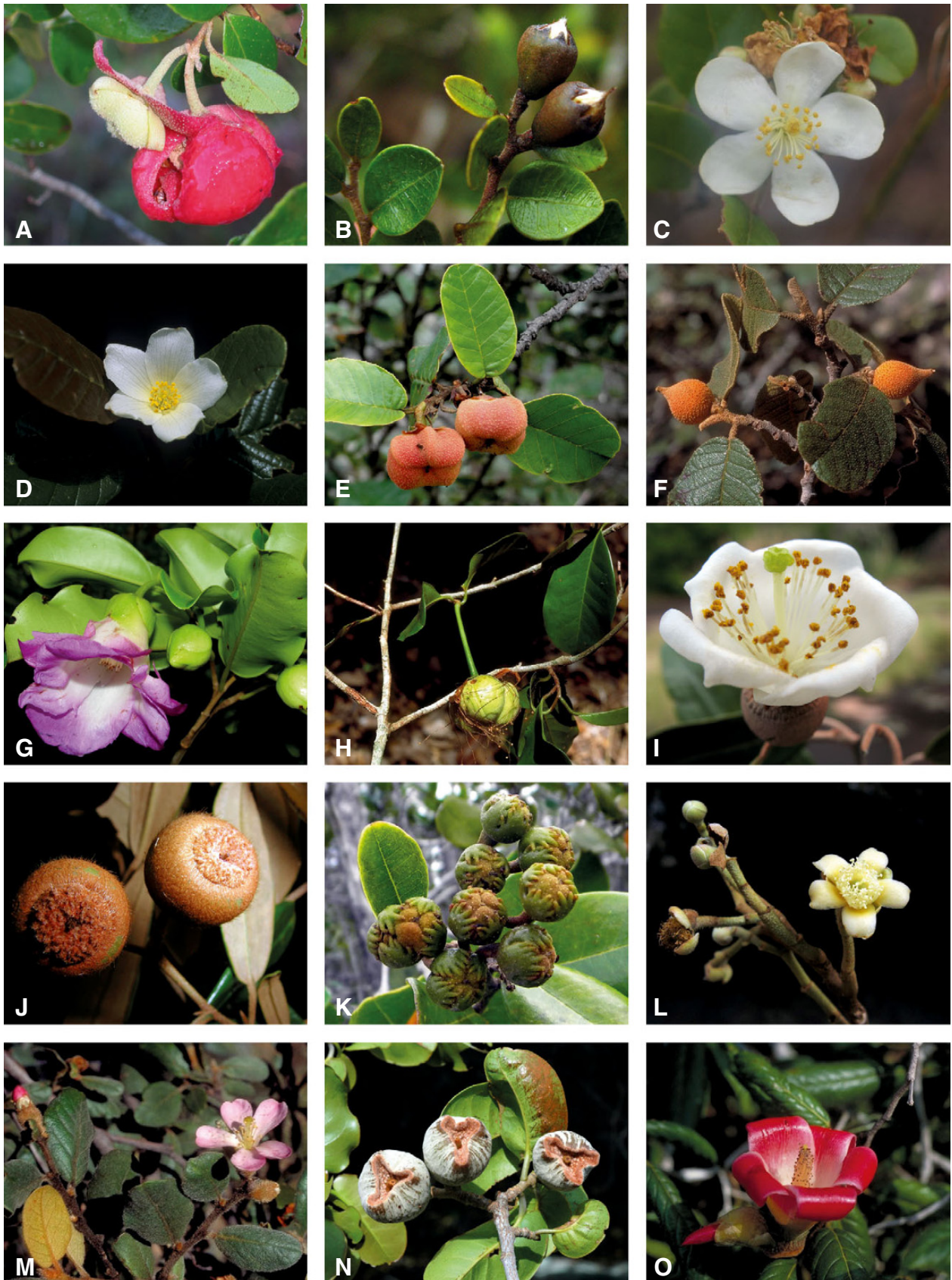
Gerard, 1919; Cavaco, 1952a,b; Dehay, 1957; Straka, 1963, 1971; Carlquist, 1964; Capuron, 1970; Coetzee & Muller, 1984; Goldblatt & Dorr, 1986; Nilsson & Randrianasolo, 1999; Ducouso *et al.*, 2004; Ramananjanahary *et al.*, 2010), no comprehensive molecular phylogenetic study has been conducted to date and little if any research has been done on what may be the family's most distinctive morphological feature, i.e. an extra-floral involucre that expands in many taxa to form a large, fleshy or woody structure, often covering or enveloping the fruit (a syndrome referred to as angiocarpy; Fig. 1A, B, I–K and N). Recent taxonomic work has led to a broad overall consensus on generic limits in Sarcolaenaceae, resolving many long-standing issues (Randrianasolo & Miller, 1999; Schatz *et al.*, 2001; Hong-Wa, 2009), but the monophyly of these genera and the relationships among them remain to be tested using molecular phylogenetic tools.

The flora of Madagascar is both rich and highly endemic, with an estimated 13 000–14 000 species of native vascular plants, almost 90% of which occur nowhere else (Phillipson *et al.*, 2006; Callmander *et al.*, 2011). The island's diverse array of natural ecosystems is also highly threatened, qualifying Madagascar as a biodiversity hotspot (Myers *et al.*, 2000; Ganzhorn *et al.*, 2001). Sarcolaenaceae comprise 72 described species, in addition to which several recently discovered novelties remain to be published (Madagascar Catalogue, 2016). The largest genus, *Schizolaena* Thouars, includes 22 species, followed by *Sarcolaena* Thouars, in which eight described species are recognised in addition to c. six more that are unpublished, whereas the smallest genera are *Eremolaena* Baill. (three species) and *Mediusella* (Cavaco) Hutch. (two species) (Madagascar Catalogue, 2016). Members of the family are trees or shrubs with simple alternate leaves and they usually have showy flowers with five or six petals and three to five sepals (Fig. 1C, D, G, I, L, M and O). As mentioned above, they are characterised by an extra-floral involucre, which resembles but is developmentally different from the accrescent calyx that envelopes the fruit in some other genera such as *Corylus* L. (Betulaceae) and *Physalis* L.

(Solanaceae) and in some Malvaceae. The extra-floral involucre of Sarcolaenaceae can be well developed at anthesis, as in *Xyloolaena* Baill. (Fig. 1N), or at later stages of floral development, as in *Leptolaena* Thouars (Fig. 1B). It can form an accrescent cup (e.g. in *Leptolaena*, *Sarcolaena* and *Xyloolaena*; Fig. 1B, J and N), may comprise a compound structure of free parts (as in *Schizolaena*; Fig. 1K) or can form a minute cupule at the base of the fruit (as in *Perrierodendron* Cavaco; Fig. 1F); it can be either fleshy, as in *Eremolaena* and *Schizolaena* (Fig. 1A and K), or woody, as in *Leptolaena* and *Xyloolaena* (Fig. 1B and N). Sarcolaenaceae are restricted to Madagascar, where they occupy all bioclimatic regions (Cornet, 1974), being particularly abundant in the perhumid forests of the east coast and much less common in the bushland of subarid southern and southwestern Madagascar (Soulebeau *et al.*, 2016). Fossil pollen attributed to Sarcolaenaceae has been found in the early Miocene of South Africa (Coetzee & Muller, 1984), indicating that the family once had a larger distribution and that its present restriction to Madagascar is probably the result of extinction elsewhere.

The family was described by Du Petit-Thouars in his 'Histoire des végétaux recueillis dans les îles australes d'Afrique' (Du Petit-Thouars, 1806) as 'Chlaenaceae', based on the Greek word χλαενα (coat), referring to the characteristic involucre. This descriptive name was in use for almost a century, but was rejected in favour of Sarcolaenaceae because it was not based on a legitimate generic name, as required under Article 18 of the Code of Nomenclature (McNeill *et al.*, 2012; see also Turland, 2013). Initially, Du Petit-Thouars (1806) regarded the family as being closely related to Malvaceae or Tiliaceae and Gerard (1919) placed it in Malvales, indicating that it had similarities with Theaceae. Cavaco (1952b) developed the concept of a polyphyletic Sarcolaenaceae, part of which he thought was close to Tiliaceae and the remainder to Theaceae, without clearly stating which genera belonged to which group. Dehay (1957) also saw close affinities with Malvales and, in particular, with Tiliaceae, based on

Figure 1. Diversity of Sarcolaenaceae. A, *Eremolaena rotundifolia* (Birkinshaw 1651, Iabakoho); B, *Leptolaena pauciflora* (Antilahimena *et al.* 7042, Antaniditra forest); C, *Mediusella bernieri* (Hong-Wa 206, Sahafary forest); D, *Pentachlaena latifolia* (Schatz 4135 *et al.*, Ibity massif); E, *Pentachlaena latifolia* (Lowry 7286, Ibity massif); F, *Perrierodendron quartzitorum* (Schatz 3973, Itremo massif); G, *Rhodolaena coriacea* (Birkinshaw & Be 1962, Tsihomanaomby forest); H, *Rhodolaena macrocarpa* (Andrianjafy *et al.* 61, Tsihomanaomby forest); I, *Sarcolaena oblongifolia* (Rakotoarivelo *et al.* 316, Anjoman'Ankona); J, *Sarcolaena* sp. (Razakamalala *et al.* 5575, Ampanola forest); K, *Schizolaena elongata* (Razanatsoa *et al.* 524, Ampanasanarato); L, *Schizolaena manomboensis* (Birkinshaw *et al.* 1738, Analalava forest); M, *Xerochlamys diospyroidea* (Schatz 3960, Itremo massif); N, *Xyloolaena humbertii* (Lowry *et al.* 6968, Ankoba); O, *Xyloolaena perrieri* (Noyes *et al.* 1025, Kirindy forest). Photograph credits: (A, G) C.R. Birkinshaw; (B) P. Antilahimena; (C) F. Ratovoson; (D, E, H, L, N) P. Lowry; (F, M) G.E. Schatz; (I) F. Rajaonary; (J) C. Rakotovao; (K) N.V. Manjato; (O) D.K. Harder. Complete collection information can be found in the Madagascar Catalogue (2016).



petiole characters. Hutchinson (1959, 1973) placed the family in Ochnales, with Ochnaceae, Scytopetalaceae, Dipterocarpaceae, Ancistrocladaceae and Sphaerosepalaceae.

Recent molecular studies have confirmed that Sarcolaenaceae belong in Malvales *sensu* APG IV (2016) and have either placed the family as sister to Dipterocarpaceae (Dayanandan *et al.*, 1999; Morton, Dayanandan & Dissanayake, 1999) or included it in a poorly resolved clade comprising Dipterocarpaceae, Cistaceae and Sarcolaenaceae (Ducousso *et al.*, 2004). In a study focusing primarily on Dipterocarpaceae (including subfamilies Dipterocarpoideae, Monotoideae and Pakaraimoideae), Dayanandan *et al.* (1999) reported that Dipterocarpaceae were sister to the single representative of Sarcolaenaceae used in their analyses (a species of *Sarcolaena*), a finding that was confirmed by Morton *et al.* (1999). Several years later, Ducousso *et al.* (2004) used expanded sampling, including three species of Sarcolaenaceae representing two genera (*Leptolaena* and *Sarcolaena*), to reassess relationships in Dipterocarpaceae. Their results placed both Cistaceae and Sarcolaenaceae in Dipterocarpaceae, with Sarcolaenaceae as sister to Dipterocarpoideae and Pakaraimoideae as sister to Cistaceae, although the use of only one plastid gene (*rbcL*) for inferring phylogenetic relationships and the moderate to low levels of support for these relationships did not preclude the monophyly of Dipterocarpaceae.

The phylogenetic studies conducted to date have only included a limited number of exemplar taxa of Sarcolaenaceae, making it impossible to assess whether the family is monophyletic or to elucidate relationships among its ten genera. In the present study, we report the results of the first molecular phylogenetic analyses of Sarcolaenaceae, using sequence data from one nuclear (ITS) and three plastid (*psaA-ORF170*, *psbA-trnH* and *rbcL*) markers. We sampled 46 of the 72 currently recognised species, including representatives of all ten genera. Our goals are to: (1) test the monophyly of Sarcolaenaceae with respect to those families identified in previous studies as being closely related, in particular Cistaceae and Dipterocarpaceae; (2) test the monophyly of each of the ten currently recognised genera of Sarcolaenaceae; and (3) evaluate the relationships among these genera.

MATERIAL AND METHODS

TAXON SAMPLING

Our sampling accounts for nearly two-thirds of Sarcolaenaceae diversity (46 of 72 accepted species) and is representative of the morphological and

ecogeographic diversity of each of the ten genera. It includes one of three species of *Eremolaena*, two of four species of *Pentachlaena* H.Perrier, three of five species of *Perrierodendron*, three of eight species of *Sarcolaena*, ten of 22 species of *Schizolaena*, two of five species of *Xyloolaena* and all members of *Leptolaena*, *Mediusella*, *Rhodolaena* Thouars and *Xerochlamys* Baker (comprising eight, two, seven and eight species, respectively). For eight of the ten genera of Sarcolaenaceae, the type species of the genus was included in the sampling, the only exceptions being *Eremolaena humblotiana* Baill. and *Xerochlamys pilosa* Baker. All species were sampled using leaf fragments preserved in silica gel that were collected recently by the authors or by the staff of the Madagascar Program of the Missouri Botanical Garden.

To test the monophyly of Sarcolaenaceae, we drew on information from previous molecular phylogenetic studies that have included representatives of Sarcolaenaceae (Dayanandan *et al.*, 1999; Morton *et al.*, 1999; Ducousso *et al.*, 2004), which informed our selection of a set of species from the two most closely related families, Cistaceae and Dipterocarpaceae, as outgroup taxa. This sampling, which made use of sequences available in GenBank, was designed to provide maximum possible taxonomic coverage at the generic level. It includes representatives of five of the nine accepted genera of Cistaceae and nine of 16 recognised genera of Dipterocarpaceae (The Plant List, <http://www.theplantlist.org/>). In Dipterocarpaceae, the outgroup sampling includes representatives of all three subfamilies: *Pakaraima* *dipterocarpacea* Maguire & P.S.Ashton (which belongs to the monotypic subfamily Pakaraimaoidae), one member each of two of the three genera of Monotoideae (*Monotes glaber* Sprague and *Pseudomonotes tropenbosii* A.C.Londoño, E.Alvarez & Forero) and one representative each of six of 13 genera of Dipterocarpoideae [*Anisoptera marginata* Korth., *Cotylelobium lanceolatum* Craib, *Hopea mengarawan* Miq., *Neobalanocarpus heimii* (King) P.S.Ashton, *Parashorea chinensis* Wang Hsie and *Shorea robusta* C.F.Gaertn.]. We used one species of Cistaceae (*Helianthemum scopulicola* L.Sáez, Roselló & Alomar) to root all analyses. In total our sampling thus included 60 species of Sarcolaenaceae, Cistaceae and Dipterocarpaceae. All vouchers are cited in Appendix 1 and complete collection information for the Sarcolaenaceae accessions sampled here can be found in the Madagascar Catalogue (2016, <http://www.tropicos.org/project/mada>). We were unable to trace voucher information for six of the accessions downloaded from GenBank for Cistaceae and Dipterocarpaceae (*Cotylelobium lanceolatum*, *Hopea mengarawan*, *Neobalanocarpus heimii*,

Parashorea chinensis, *Pseudomonotes tropenbosii* and *Shorea robusta*).

DNA EXTRACTION, AMPLIFICATION, SEQUENCING AND DATA ASSEMBLY

Samples were ground at room temperature using a Mixer Mill MM 300 with tungsten carbide beads (Qiagen Inc., Valencia, CA, USA). Total genomic DNA was extracted using the DNeasy 96 Plant kit (Qiagen Inc.) according to the manufacturer's protocol.

We selected DNA regions for amplification that would be useful for making phylogenetic inferences at the family and generic levels and that enabled us to maximise outgroup coverage using previously published sequence data. One nuclear region, the internal transcribed spacer (ITS), and one plastid gene, RuBisCO large subunit (*rbcL*), were particularly well represented in GenBank for the two outgroup families (Dipterocarpaceae and Cistaceae) and were thus chosen as core markers. We also added two variable plastid intergenic spacers, *psaA-ORF170* and *psbA-trnH*, in an attempt to increase both resolution and node support of the phylogenetic backbone.

The entire ITS region (ITS1 + 5.8S + ITS2) was amplified using two primers, ITS4 and ITS5 (White *et al.*, 1990), with the following cycling parameters: 2-min initial denaturation at 94 °C, 30 cycles of 1-min denaturation at 94 °C, 1-min annealing at 52 °C and 3-min extension at 72 °C followed by a final extension of 7 min at 72 °C. The three plastid regions (*psaA-ORF170*, *psbA-trnH* and *rbcL*) were amplified following standard procedure described by Shaw *et al.* (2007). All PCR reactions were performed in a total volume of 25 µL containing: 5 µL Taq&Go (Qiogene, Irvine, CA, USA) 5× mastermix, DMSO 3–5% final volume (for ITS only), 1 µL 10 µM forward and reverse primer, 1–3 µL template DNA of unknown concentration, and up to 25 µL water. Purification and sequencing of PCR products were performed by the National Centre for Sequencing (Génoscope) in Évry, France, using the same primers as for amplification, except for ITS, for which primers ITS 5 and ITS 2 were used for ITS1, and primers ITS 3b and ITS 4 for ITS2 (White *et al.*, 1990).

The newly generated sequences were assembled and edited in Geneious v.7.1.9 (Biomatters Ltd., Auckland, New Zealand) and, with previously published sequences obtained from GenBank, were automatically aligned with MAFFT v.7.221 (Katoh & Standley, 2013) using the L-INS-i algorithm. All newly generated sequences, with their accompanying voucher information, are archived in GenBank (Appendix 1).

PHYLOGENETIC ANALYSES

Maximum likelihood (ML) and Bayesian inference (BI) analyses were performed on each single-marker matrix and on the combined marker dataset (see Table 1 for details of datasets). For each of the four DNA regions, the best fitting nucleotide substitution model was selected with MrModeltest v.2.3 (Nylander, 2004) using the Akaike information criterion (AIC). The ITS and *rbcL* regions followed the GTR + I + G substitution model; the two intergenic spacers, *psaA-ORF170* and *psbA-trnH*, followed the GTR + I model (Table 1). Gaps were treated as missing data. The BI and ML analyses were performed via the CIPRES Science Gateway (Miller, Pfeiffer & Schwartz, 2010) using MrBayes 3.2.6 (Huelsenbeck & Ronquist, 2001) and RaxML-HPC v.8.1.24 (Stamatakis, 2014), respectively. The BI analyses, performed using the selected substitution model, constituted two independent parallel runs of four Markov chains each, implemented for 10 million generations and sampled every 1000 generations. Adequate mixing of the Markov chains and convergence of the two runs were confirmed with Tracer v1.6 (Rambaut *et al.*, 2014). After removing a 10% burn-in, the remaining trees were used to generate a 50% Bayesian majority-rule consensus tree. For the ML analyses, we applied for each single-marker matrix a GTR + GAMMA rate substitution model to fit them to those implemented in the BI analyses, and we used a rapid bootstrap algorithm with 1000 replicates. After visual inspection for congruence between the topologies obtained for each of the four markers in the ML and BI analyses, all subsequent analyses used the combined dataset.

Our attempts to assemble a sampling of Cistaceae and Dipterocarpaceae that is representative of the generic diversity in these two families were hampered by the need to rely on the somewhat limited number of sequences available in GenBank and the fact that four of the species we used were represented by a single DNA region (*rbcL*). Moreover, obtaining successful amplicons from DNA extracts of Sarcolaenaceae proved to be particularly challenging and, despite numerous attempts to resolve this issue, we were unable to generate sequences for all four markers from some accessions. Specifically, we obtained data for all four regions from 25 of the 46 new accessions and from three regions in 11 accessions; six accessions yielded sequences for just two markers and four accessions yielded sequences for only one region (see Appendix 1 for detailed information). Sequences that were not available were coded as missing data ('N'). To evaluate the possible effects of missing data on our phylogenetic hypotheses, we assembled a smaller matrix (comprising 39 species) that excluded the

Table 1. Characteristics of the DNA regions used for the separate and combined Bayesian and ML analyses

	<i>psaA-ORF170</i>	<i>psbA-trnH</i>	<i>rbcL</i>	ITS	Combined
Number of accessions (/60)	34	39	50	54	60
Aligned sequence length (bp)	998	274	1402	792	3466
Number of variable characters (bp)	83	61	191	389	724
Number of potentially parsimony informative characters (bp)	49	52	100	299	500
Model selected	GTR + I	GTR + I	GTR + I + G	GTR + I + G	–

accessions for which we lacked more than one DNA region. These combined matrices were divided into four partitions, corresponding to the four DNA regions, to which the best-fitted substitution models were applied in BI (Table 1). ML and BI analyses were run using the combined dataset under the same conditions detailed above for the single-marker matrices. In analysing the results, nodes with bootstrap support (BS) $\geq 90\%$ and posterior probabilities (PP) ≥ 0.99 were considered to be strongly supported, those with BS $> 80\%$ and PP > 0.9 as moderately supported and all others as weakly supported.

RESULTS

Visual comparison of the 50% majority-rule trees obtained from the single markers showed no strongly supported topological conflicts between the markers or between the methods used for phylogenetic reconstruction (trees not shown). The topologies only differed in the degree of resolution, with the ITS and *psaA-ORF170* consensus trees having higher proportions of supported nodes than those obtained with the *psbA-trnH* and *rbcL* matrices.

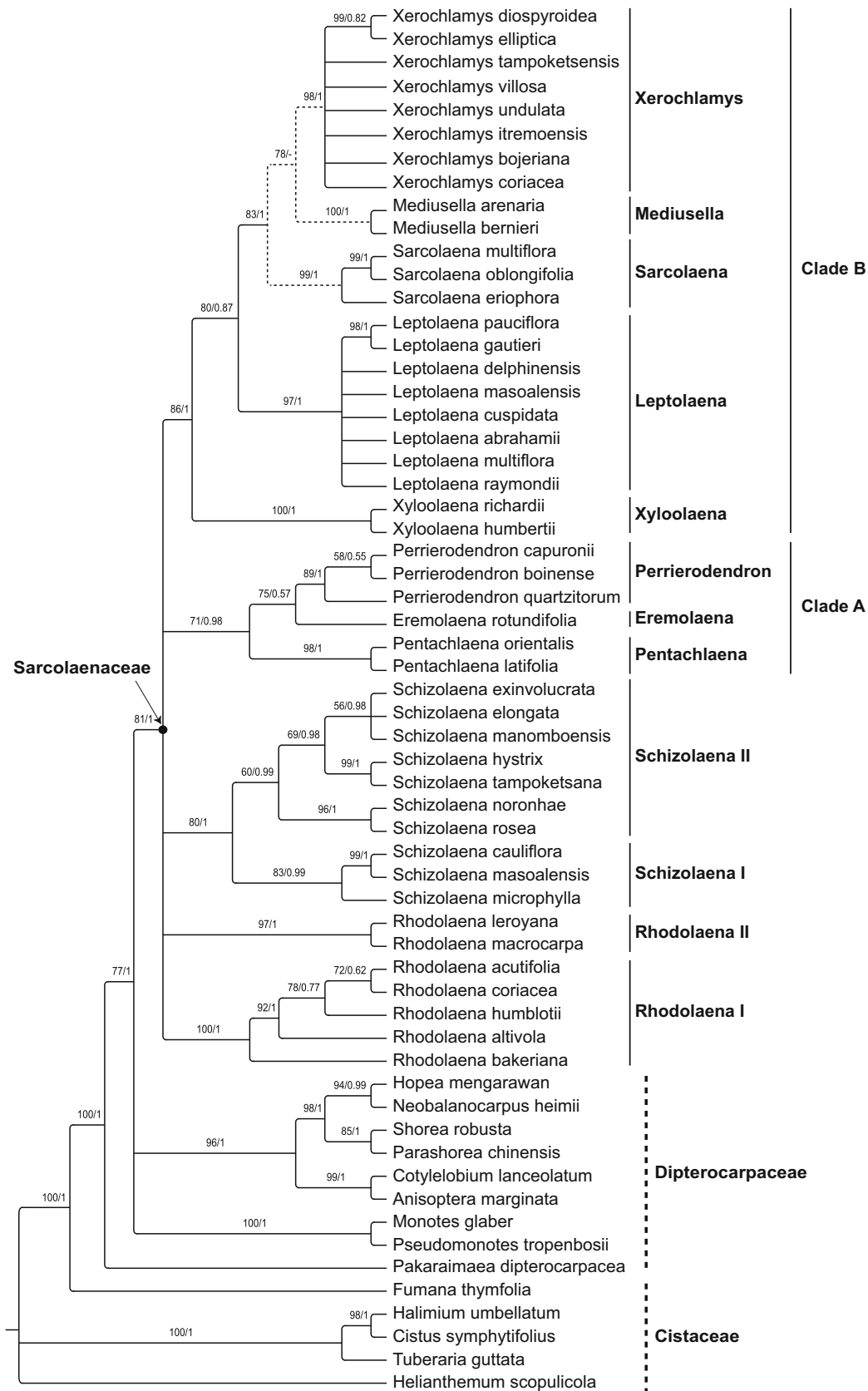
Trees obtained with the more complete dataset (i.e. the 39 accessions that yielded sequences for three or four markers; trees not shown) were similar to those derived from the full, 60-accession dataset (Fig. 2), although they were less well resolved. As the focus of this study is on testing the monophyly of the genera of Sarcolaenaceae and inferring relationships among them, the larger dataset was more suitable and will be used for the remainder of this paper. The complete combined dataset (60

accessions) yielded a total of 3466 base pairs (bp) of which 500 bp were potentially parsimony informative. We observed no strongly supported conflicts between the ML bootstrap 50% majority-rule consensus tree (Fig. 2) and the 50% Bayesian majority-rule consensus tree (not shown). The only topological incongruence lies in the relative branching of *Mediusella*, *Sarcolaena* and *Xerochlamys*. *Xerochlamys* is sister to *Mediusella* in the ML topology (Fig. 2), whereas it is grouped with *Sarcolaena* in the BI consensus tree, although these relationships are only poorly supported (BS = 78% for ML and PP = 0.82 for BI) and the incongruence points toward a lack of phylogenetic resolution among these three genera. The topology resulting from analysis of the combined dataset is consistent with those obtained for the individual markers while significantly improving overall phylogenetic resolution. It shows that: (1) Sarcolaenaceae species form a monophyletic group; (2) all the genera for which we have more than one representative (i.e. excluding *Eremolaena*) are monophyletic, with the exception of *Rhodolaena*; and (3) the phylogenetic relationships among the ten genera of Sarcolaenaceae are poorly resolved but several monophyletic lineages are recovered.

All species of Sarcolaenaceae included in our sampling are nested in a moderately supported monophyletic group (Fig. 2; BS = 81%; PP = 1). Sarcolaenaceae are part of a polytomy that includes all of the species of Dipterocarpaceae sampled, with the exception of *Pakaraimaea dipterocarpacea* (BS = 77%; PP = 1).

Relationships among the genera of Sarcolaenaceae are poorly resolved, with a polytomy that includes: (1) a clade grouping five species of *Rhodolaena*

Figure 2. Fifty per cent majority-rule tree from the maximum likelihood partitioned analysis of the combined dataset including internal transcribed spacer (ITS), two plastid intergenic spacers (*psaA-ORF170* and *psbA-trnH*) and one plastid gene (*rbcL*). Numbers above each branch are bootstrap values $>50\%$ followed by posterior probabilities from the Bayesian analysis. Broken lines represent relationships that are not present in the 50% majority-rule tree from the Bayesian tree of the combined dataset. Clades discussed in the text are labeled; solid lines indicate Sarcolaenaceae clades and broken lines show the two families of outgroup taxa (Cistaceae and Dipterocarpaceae).



(referred to as '*Rhodolaena* I'; Fig. 2; BS = 100%; PP = 1); (2) a smaller clade comprising two other *Rhodolaena* species, *R. leroyana* G.E.Schatz, Lowry & A.-E.Wolf and *R. macrocarpa* G.E.Schatz, Lowry & A.-E.Wolf ('*Rhodolaena* II'; BS = 97%; PP = 1); (3) a clade that includes all species of *Schizolaena* (BS = 80%; PP = 1), which is itself divided into two moderately supported clades, '*Schizolaena* I' (BS = 83%; PP = 0.99) and '*Schizolaena* II' (BS = 60%; PP = 0.99). The polytomy at the base of the family also includes two additional clades, hereafter referred to as 'Clade A' and 'Clade B', in which the phylogenetic relationships among the eight remaining genera of Sarcolaenaceae are resolved. Clade A (BS = 71%; PP = 0.98) includes all of the sampled species of three genera, *Eremolaena*, *Pentachlaena* and *Perrierodendron*, and in this clade *Pentachlaena* and *Perrierodendron* are supported as monophyletic (BS = 98%; PP = 1, and BS = 89%; PP = 1, respectively); the monophyly of *Eremolaena* could not be inferred as only a single sample was available. In Clade A, *Perrierodendron* is sister to *Eremolaena*, although this relationship is only weakly supported (BS = 75%; PP = 0.57). All representatives of the five remaining genera of Sarcolaenaceae belong to Clade B (BS = 86%; PP = 1), viz. *Leptolaena*, *Mediusella*, *Sarcolaena*, *Xerochlamys* and *Xyloolaena*. At the base of this clade, *Xyloolaena* is monophyletic (BS = 100%; PP = 1) and is sister to the other genera, which form a moderately supported group (BS = 80%; PP = 0.87). The monophyly of *Leptolaena* is well supported (BS = 97%; PP = 1) and this genus is sister to a clade that includes the three remaining genera (BS = 83%; PP = 1). Each of these three genera is likewise monophyletic, with *Sarcolaena* (BS = 99%; PP = 1) sister to a clade (supported only in the ML tree; BS = 78%) comprising *Mediusella* (BS = 100%; PP = 1) and *Xerochlamys* (BS = 98%; PP = 1).

DISCUSSION

This study represents the first attempt to reconstruct the phylogeny of Sarcolaenaceae based on molecular sequence data. Three previous studies of relationships within the family relied on morphological analyses and used limited sampling (Straka, 1963, 1971; Haevermans, 1999). Each of them confirmed the monophyly of the genera sampled and yielded results that are consistent with our findings. Straka (1971) distinguished two hierarchical 'groupings' originating from a single 'Urtyp' ('archetype' in Swedish), one in which he regarded *Pentachlaena* as sister to *Perrierodendron* + *Eremolaena* and another comprising *Rhodolaena*, *Xyloolaena* and a group containing

Xerochlamys + *Sarcolaena* and *Mediusella* + *Leptolaena*. His analysis did not, however, include any members of *Schizolaena*, the largest genus in the family (with 22 species). Haevermans (1999) conducted a cladistic analysis based on morphological characters, which resulted in a tree that strongly supported *Schizolaena* as sister to the other genera of Sarcolaenaceae, which formed two main clades. In one of these clades, he found weak support for a sister relationship between *Rhodolaena* and a subclade comprising *Pentachlaena* and *Perrierodendron* + *Eremolaena*. The other main clade placed *Xyloolaena* sister to two subclades, *Mediusella* + *Leptolaena* and *Xerochlamys* + *Sarcolaena*. The results of these analyses based on morphological data, although broadly compatible with our findings, are not totally congruent with them. In particular, our results did not recover *Schizolaena* as sister to the other members of the family and they did not place *Rhodolaena* as sister either to a *Pentachlaena* + *Perrierodendron* + *Eremolaena* clade or a clade comprising *Xyloolaena*, *Mediusella*, *Leptolaena*, *Xerochlamys* and *Sarcolaena* (Fig. 2).

Our molecular analyses confirm the monophyly of Sarcolaenaceae. This is consistent with morphology and indicates that several characters represent synapomorphies for the family, in particular a distinct type of pollen organised in tetrads and the presence of an extra-floral involucre, on which the family was originally defined (Du Petit-Thouars, 1806). Our results also support the monophyly of each of the genera of Sarcolaenaceae (excluding *Eremolaena*, for which a single sample was available), with the notable exception of *Rhodolaena*, which was found to comprise two distinct, well supported clades. Although basal affinities within Sarcolaenaceae remain unclear, the relationships among the genera included in Clade A and in Clade B are resolved (see Fig. 2). In the following paragraphs we discuss each genus in turn. For each genus, we cite key morphological and ecogeographical features; more detailed descriptions and a recent key to the genera are provided in Ramananjanahary *et al.* (2010) and can be accessed through the Madagascar Catalogue (2016, <http://www.tropicos.org/project/mada>).

RHODOLAENA

Species of *Rhodolaena* are arborescent, have fruit surrounded by a fleshy involucre and occur in humid forest areas in eastern Madagascar (Schatz *et al.*, 2000; Madagascar Catalogue, 2016). We were able to include all seven members of the genus in our analyses, the results of which did not support its monophyly but rather recovered two separate, well supported clades, *Rhodolaena* I and *Rhodolaena* II

(Fig. 2). *Rhodolaena* II contains *R. leroyana* and *R. macrocarpa*, two species in which the fleshy involucre completely covers the fruits at maturity (Fig. 1H) and the petals are yellow. This last character state contrasts with the distinctive pinkish-magenta corolla of the members of *Rhodolaena* I, which gave the genus its name (Fig. 1G). The two species of *Rhodolaena* II are found in low elevation humid forests and are known from limited areas north-west of Toamasina and Sambava, respectively (Madagascar Catalogue, 2016). The five remaining species of *Rhodolaena* differ not only in flower colour, but also in having an involucre that does not always expand over the mature fruits (Schatz *et al.*, 2000). They also occur in eastern Madagascar, but their geographical and elevational ranges are larger (Madagascar Catalogue, 2016).

Based on several shared morphological features (including paired flowers, a fleshy involucre and septicidally dehiscent fruits), *Rhodolaena* was suggested to be closely related to *Schizolaena* (see Lowry *et al.*, 1999; Schatz *et al.*, 2000). Our results, in which the affinities of *Rhodolaena* I and *Rhodolaena* II remain unresolved, neither support nor refute this notion, but they do raise the possibility that *Rhodolaena* may not be monophyletic. If this proves to be the case, flower colour could represent a reliable character for recognising them as separate genera, but a final decision must await further analyses using additional, rapidly evolving markers, perhaps supplemented with even more comprehensive sampling in the other genera of the family.

SCHIZOLAENA

Schizolaena (22 species) is the only genus of Sarcolaenaceae to have three sepals rather than five (Fig. 1L), and it also possesses bracts on the inflorescence axes in addition to the involucre, confirming that these represent distinct structures (Lowry *et al.*, 1999). The involucres are accrescent over the fruit, fleshy and sometimes deeply divided, with dentate to lacinate margins (Fig. 1K). As in *Rhodolaena*, species of *Schizolaena* are mostly found in humid forests of eastern Madagascar, although a few members of the genus are restricted to the subhumid central plateau (*S. isaloensis* Rabeh. & Lowry, *S. microphylla* H.Perrier and *S. tampoketsana* Lowry, G.E.Schatz, J.-F.Leroy & A.-E.Wolf), one species is found in the north-east in the humid Sambirano region [*S. parviflora* (F.Gérard) H.Perrier] and another occurs in northern dry forests (*S. viscosa* F.Gérard) (Madagascar Catalogue, 2016). Our sampling includes nearly half of the diversity in the genus (ten of 22 species), which formed a strongly supported clade (Fig. 2) in which two moderately supported subclades can be

distinguished, *Schizolaena* I, comprising three species (*S. cauliflora* Thouars, *S. masoalensis* Lowry, G.E.Schatz, J.-F.Leroy & A.-E.Wolf and *S. microphylla*) and *Schizolaena* II, which includes all the remaining taxa sampled (*S. elongata* Thouars, *S. exinvolucrata* Baker, *S. hystrix* Capuron, *S. manomboensis* Lowry, G.E.Schatz, J.-F.Leroy & A.-E.Wolf, *S. noronhae* G.E.Schatz & Lowry, *S. rosea* Thouars and *S. tampoketsana*). Based on the current sampling, we have not identified any morphological synapomorphies that might characterise these two clades or any clear ecological preferences that could differentiate them. As with *Rhodolaena*, the broader affinities of *Schizolaena* remain unresolved (Fig. 2).

CLADE A:

PENTACHLAENA + EREMOLAENA + PERRIERODENDRON

This clade includes three relatively small genera, *Eremolaena* (three species), *Pentachlaena* (four species) and *Perrierodendron* (five species) (Madagascar Catalogue, 2016). We were able to sample one species of *Eremolaena* [*E. rotundifolia* (F.Gérard) Danguy], two species of *Pentachlaena* (*P. latifolia* H.Perrier and *P. orientalis* Capuron) and three species of *Perrierodendron* [*P. boinense* (H.Perrier) Cavaco, *P. capuronii* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz and *P. quartzitarum* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz]. In our analyses, all of the taxa in Clade A form a monophyletic group, in which both *Pentachlaena* and *Perrierodendron* are supported as monophyletic and *Eremolaena* (the monophyly of which cannot be tested given our sample size) is sister to *Perrierodendron*, albeit with poor support (Fig. 2). These findings are consistent with the results based on morphological data alone reported by Straka (1963, 1971) and Haevermans (1999).

Although these three genera form a clade, they are clearly distinct from one another: species of *Eremolaena* have tricarpellate ovaries (vs. bicarpellate in *Perrierodendron* and pentacarpellate in *Pentachlaena*); and *Perrierodendron* has indehiscent fruits (Fig. 1F) whereas those of *Eremolaena* and *Pentachlaena* are dehiscent (Fig. 1A and E). They also have somewhat different bioclimatic preferences, with members of *Eremolaena* and *Pentachlaena* occurring mostly in subhumid areas in the north-eastern (*E. darainensis* Nusb. & Lowry) and central (*P. latifolia*) parts of Madagascar and humid areas in the east (*E. humblotiana*, *E. rotundifolia*, *P. betamponensis* Lowry, Haev., Labat & G.E.Schatz and *P. orientalis*), whereas *Perrierodendron* exhibits a broader range, with three of the five species occurring in the dry areas in the centre and west (*P. boinense*, *P. occidentalis* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz).

and the north-east (*P. rodoense* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz), with one each in subhumid (*P. quartzitorum*) and humid areas (*P. capuronii*) (Madagascar Catalogue, 2016). The relationships between the dehiscent-fruited genera, *Eremolaena* and *Pentachlaena*, and *Perrierodendron*, whose fruits are indehiscent, are not strongly supported. Increased sampling in Clade A might provide better resolution and thereby help develop an improved understanding of the evolution of this and other characters that differ among these three genera.

CLADE B: *XYLOOLAENA* + *LEPTOLAENA* +
SARCOLAENA + *MEDIUSELLA* + *XEROCHLAMYS*

The five genera included in Clade B are rather heterogeneous in terms of species richness. We were able to sample all members of *Leptolaena* (eight species), *Mediusella* (two species) and *Xerochlamys* (eight species) and three of the eight species of *Sarcolaena* (*S. eriophora* Thouars, *S. oblongifolia* F.Gérard and *S. multiflora* Thouars) and two of the five species of *Xyloolaena* [*X. humberitii* Cavaco and *X. richardii* (Baill.) Baill.]. All of the species belonging to each genus formed a well supported clade and, although our sampling is not exhaustive for *Sarcolaena* and *Xyloolaena*, the results confirm the monophyletic nature of the currently accepted generic delimitations based on morphology.

In Clade B, *Xyloolaena*, the first-branching lineage (Fig. 2), is readily distinguishable from the other genera by its woody involucre (hence the generic name), which is already well developed at anthesis (Fig. 1N). Also, the ovaries of *Xyloolaena* have numerous multiseriate ovules, whereas the other genera exhibit no more than five ovules in two series per ovary (Capuron, 1970). The species of *Xyloolaena* occur in areas with dry and subhumid bioclimates in western Madagascar (*X. perrieri* F.Gérard) to humid areas in the north-east (*X. sambiranensis* Lowry & G.E.Schatz and *X. speciosa* Lowry & G.E.Schatz), east (*X. richardii*) and south-east (*X. humberitii* Cavaco) (Madagascar Catalogue, 2016).

Leptolaena is sister to the three remaining genera in Clade A (Fig. 2). The monophyly of this genus has been the subject of a long-standing debate. Cavaco (1952a; see also Cavaco, 1952b), in his treatment of *Sarcolaenaceae* for the *Flore de Madagascar*, defined *Leptolaena* broadly to include the previously recognized genera *Mediusella* and *Xerochlamys*, which were treated as subgenera, along with '*Euleptolaena*' (a superfluous name for subgenus *Leptolaena*) (Cavaco, 1951), including its type species, *L. multiflora* Thouars. Schatz *et al.* (2001) adopted a narrow circumscription of *Leptolaena* based on morphology, excluding the species assigned by Cavaco to

subgenera *Mediusella* and *Xerochlamys*, which was subsequently also adopted by Hong-Wa (2009). Our results support this interpretation, with *Mediusella* and *Xerochlamys* being more closely related to *Sarcolaena* than to *Leptolaena* (Fig. 2). *Leptolaena*, as currently circumscribed, can be distinguished from the four other genera of Clade B by having smaller flowers with fewer stamens (six to 12 in *Leptolaena* vs. 20 or more in *Mediusella*, *Sarcolaena* and *Xerochlamys*) and its involucre is woody at maturity (Fig. 1B), a feature found in only one other genus of Clade B, namely *Mediusella*. Most members of the genus occur in areas with a humid bioclimate, including sites at low elevation (e.g. *L. delphinensis* G.E.Schatz & Lowry, *L. raymondii* G.E.Schatz & Lowry, *L. masoalensis* G.E.Schatz & Lowry and *L. multiflora*) and more upland localities (*L. abrahamii* G.E.Schatz & Lowry and *L. gautieri* G.E.Schatz & Lowry), but one species (*L. cuspidata* Baker) is found in dry and subhumid areas in the north-west and another (*L. pauciflora* Baker) ranges widely from subhumid and humid to montane areas (Madagascar Catalogue, 2016). Although our sampling of *Leptolaena* is exhaustive, the topology of this part of our phylogenetic tree is almost totally unresolved (Fig. 2), suggesting that additional markers will be needed to elucidate relationships among the species and examine character evolution in the genus.

The next branching genus, *Sarcolaena*, is represented in our sampling by just three species (*S. eriophora*, *S. oblongifolia* and *S. multiflora*) from among the 14 currently recognized species, six of which remain to be described (Madagascar Catalogue, 2016), but they form a well supported clade that is sister to *Mediusella* + *Xerochlamys* or only *Xerochlamys* (Fig. 2). Members of *Sarcolaena* share several morphological characters, including the presence of prefoliation marks on leaves (absent, however, in *S. isaloensis* Randrian. & J.S.Mill.), fruits with two seeds per locule and a fleshy involucre that completely envelopes the fruit at maturity (Fig. 1J; Randriana-solo & Miller, 1994, 1999). They occur in a wide range of bioclimates, from dry and subhumid areas in the south-west (*S. isaloensis*) and centre of Madagascar (*S. oblongifolia*) to humid areas in the north-west (*S. codonochlamys* Baker) and east (e.g. *S. eriophora* and *S. multiflora*) (Madagascar Catalogue, 2016).

Both species of *Mediusella* were included in our sample. They formed a well supported clade (Fig. 2), confirming the monophyly of this genus, which was recently resurrected by Hong-Wa (2009). Several characters distinguish *Mediusella* from its close relatives, including its free stipules and inflorescence bracts, woody involucre and fruit with a thick pericarp not dissociating into Malpighian hairs. Species of *Mediusella* also have large, white flowers with numerous

stamens (Fig. 1C). Both species occur in areas with a dry bioclimate; *M. arenaria* (F.Gérard) Hong-Wa extends from the north-western Madagascar to the far north, whereas *M. bernieri* (Baill.) Hutch. is restricted to the far north (Madagascar Catalogue, 2016).

Our results also supported the monophyly of *Xerochlamys*, based on comprehensive sampling of all eight species (Fig. 2). As in *Sarcolaena*, members of the genus possess a fleshy involucre, a fruit pericarp that completely dissociates into Malpighian hairs (Cavaco, 1952a,b; Schatz *et al.*, 2001) and inflorescence bracts and stipules that are united, the latter forming a single hood-like structure (Fig. 1M; Capuron, 1970; Schatz *et al.*, 2001). These shared character states would be consistent with the sister relationship suggested by the topology of our BI tree, but given that the ML tree instead places *Xerochlamys* as sister to *Mediusella*, we are not able to determine whether they represent synapomorphies or are the result of convergent evolution. All species of *Xerochlamys* occur in areas with a subhumid bioclimate, although some extend into subarid to dry areas (*X. tampoketsensis* F.Gérard, *X. villosa* F.Gérard and *X. undulata* Hong-Wa) (Madagascar Catalogue, 2016). As with *Leptolaena*, additional molecular data will be required to clarify the phylogenetic relationships among the species of *Xerochlamys*.

CONCLUSIONS AND PERSPECTIVES

This first molecular phylogenetic reconstruction of Sarcolaenaceae, the largest plant family endemic to Madagascar, provides insights into the internal relationships among its genera, including those that traditionally comprised *Leptolaena s.l.* Several characteristics of the family make it an ideal group for exploring patterns and processes of evolution in the Malagasy flora. In particular, Sarcolaenaceae are well diversified both in terms of genera (ten) and species (72 described and nearly 80 recognized), far more than any of the other families endemic to Madagascar (the second largest being Sphaerosepalaceae, with two genera and 20 species; Madagascar Catalogue, 2016), and they possess a distinctive extra-floral involucre (angiocarpy) unknown elsewhere among the flowering plants. Moreover, Sarcolaenaceae, with Dipterocarpaceae, have evolved a characteristic symbiotic relationship with ectomycorrhizal fungi (Ducousso *et al.*, 2004, 2008), also found in another endemic Malagasy family, Asteropeiaceae (one genus, eight species) and in *Uapaca* Baill. (Phyllanthaceae), represented in Madagascar by 12 species (Madagascar Catalogue, 2016). Species of Sarcolaenaceae occur in a wide range of ecogeographic conditions, with populations found in many

parks and reserves in Madagascar (Soulebeau *et al.*, 2016). With sampling that covers most of the diversity in the family, we have established a framework within which analyses of morphological and ecological diversification can now be undertaken. The recent study of Soulebeau *et al.* (2016) also showed how phylogenetic information can be used to evaluate the effectiveness of Madagascar's network of protected areas, pointing towards the potential use of phylogenetic results from other plant groups to inform conservation planning.

Our study includes outgroup taxa representing several genera of Cistaceae and Dipterocarpaceae, the two families putatively most closely related to Sarcolaenaceae. While we did not seek to assess the monophyly of Cistaceae and Dipterocarpaceae or the relationships between them and Sarcolaenaceae, our results place the latter in a large clade that includes all sampled Dipterocarpaceae except one (*Pakaraimaea dipterocarpacea*). This raises questions about the monophyly of Dipterocarpaceae with respect to Sarcolaenaceae and highlights the need for further analyses using broader, fully representative sampling from Cistaceae and Dipterocarpaceae with data from an expanded set of molecular markers in an attempt to clarify relationships among these three families and ensure that their circumscriptions correspond to monophyletic groups. Such a study would benefit from the significantly broadened sampling of Sarcolaenaceae assembled for the present study and the results presented here.

To understand the historical biogeography of Sarcolaenaceae, including the implications of its occurrence in South Africa during the Miocene (Nilsson, Coetzee & Grafstrom, 1996; Buerki *et al.*, 2013) and to explore patterns and tempos of diversification in the family, a more fully resolved phylogenetic tree will be needed. This will require both the use of additional rapidly evolving DNA markers (Shaw *et al.*, 2007) and expanded sampling, especially in *Eremolaena*, *Perrierodendron*, *Sarcolaena* and *Schizolaena*. The present molecular phylogenetic study represents an important initial step toward achieving this goal, opening the way for an array of macroevolutionary studies on this fascinating and emblematic family.

ACKNOWLEDGEMENTS

The authors thank the staff of the Missouri Botanical Garden's (MBG) Madagascar Program for providing DNA samples, which were supplemented by material from the MBG DNA Bank. Fieldwork by MBG botanists was conducted under collaborative agreements with the Parc Botanique et Zoologique de Tsimbazaza and the Direction de la Recherche Forestière et

Piscicole, FOFIFA, Antananarivo, Madagascar. We thank Madagascar National Parks and the 'Direction des Eaux et Forêts' for permission to collect and the Botany-Entomology-Mycology (BoEM) molecular laboratory ['Muséum National d'Histoire Naturelle' (MNHN) Paris] for supporting the molecular work. Laboratory work was funded by the 'Agence Nationale de la Recherche' (ANR) project entitled EVORANGE (ANR-09-PEXT-011). Sequencing was performed under agreement No. 2005/67 between the Genoscope and MNHN under the project entitled 'Macrophylogy of life' directed by Guillaume Lecointre.

REFERENCES

- Andriamihajarivo T, Lowry PP II, Schatz GE. 2016.** Endemic families of Madagascar. XIV. A new restricted range species of *Pentachlaena* (Sarcocaulaceae) from central Madagascar. *Candollea* **71**: 167–172.
- 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.
- Baillon H. 1872a.** Sur le fruit d'une nouvelle Chlaenacée. *Adansonia* **10**: 234–239.
- Baillon H. 1872b.** *Histoire des Plantes, vol. 4*. Paris: Hachette & Co.
- Baillon H. 1879.** *Dictionnaire de botanique, vol. 2*. Paris: Hachette & Co.
- Baillon H. 1884.** Les *Xylolaena* et la valeur de la famille des Chlénacées. *Bulletin Mensuel de la Société Linnéenne de Paris* **1**: 410–414.
- Buerki S, Devey DS, Callmander MW, Phillipson PB, Forest F. 2013.** Spatio-temporal history of the endemic genera of Madagascar. *Botanical Journal of the Linnean Society* **171**: 304–329.
- Callmander MW, Phillipson PB, Schatz GE, Andriambololona S, Rabarimanarivo M, Rakotonirina N, Raharimampionona J, Chatelain C, Gautier L, Lowry PP II. 2011.** The endemic and non-endemic vascular flora of Madagascar updated. *Plant Ecology and Evolution* **144**: 121–125.
- Capuron R. 1970.** Observations sur les Sarcocaulaceae. *Adansonia, série 2* **10**: 247–265.
- Carlquist S. 1964.** Pollen morphology and evolution of Sarcocaulaceae (Chlaenaceae). *Brittonia* **16**: 231–254.
- Cavaco A. 1951.** Remarques sur les genres *Leptolaena* et *Xerochlamys* (Chlaenacées). Un nouveau genre de Chlaenacées. *Adansonia, série 2* **10**: 247–265.
- Cavaco A. 1952a.** Chlénacées. *Flore de Madagascar* **126**: 1–37.
- Cavaco A. 1952b.** Recherches sur les Chlénacées, famille endémique de Madagascar. *Mémoires de l'Institut Scientifique de Madagascar, série B, Biologie Végétale* **4**: 59–92.
- 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.
- Cornet A. 1974.** *Essai de cartographie bioclimatique à Madagascar. Notice explicative N° 55*. Paris: ORSTOM.
- 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.
- Dehay C. 1957.** Anatomie comparée de la feuille des Chlaenacées. *Mémoires de l'Institut Scientifique de Madagascar, série B, Biologie Végétale* **8**: 145–203.
- Du Petit-Thouars LMA. 1806.** *Histoire des végétaux recueillis dans les Isles australes d'Afrique*. Paris: Levrault, Schoell & Co.
- Ducouso M, Bena G, Bourgeois C, Buyck B, Eyssartier G, Vincelette M, Rabevohitra R, Randrihasipara L, Dreyfus B, Prin Y. 2004.** The last common ancestor of Sarcocaulaceae and Asian dipterocarp trees was ectomycorrhizal before the India-Madagascar separation, about 88 million years ago. *Molecular Ecology* **13**: 231–236.
- Ducouso M, Ramanankierana H, Duponnois R, Rabevohitra R, Randrihasipara L, Vincelette M, Dreyfus B, Prin Y. 2008.** Mycorrhizal status of native trees and shrubs from eastern Madagascar littoral forests with special emphasis on one new ectomycorrhizal endemic family, the Asteropeiaceae. *New Phytologist* **178**: 233–238.
- Ganzhorn JU, Lowry PP II, Schatz GE, Sommer S. 2001.** The biodiversity of Madagascar: one of the world's hottest hotspots on its way out. *Oryx* **35**: 346–348.
- Gerard F. 1919.** Etude systématique, morphologique et anatomique des Chlaenacées. *Annales de l'Institut Botanique-Géologique Colonial de Marseille, sér 3* **3**: 1–135.
- Goldblatt P, Dorr LJ. 1986.** Chromosome number in Sarcocaulaceae. *Annals of the Missouri Botanical Garden* **73**: 828–829.
- Haevermans T. 1999.** *Phylogénie des genres de Sarcocaulaceae, une famille endémique de Madagascar*. Paris: Mémoire de DEA, (in French), Muséum National d'Histoire Naturelle.
- Hong-Wa C. 2009.** Endemic families of Madagascar. XII. Resurrection and taxonomic revision of the genera *Mediusella* (Cavaco) Hutchinson and *Xerochlamys* Baker (Sarcocaulaceae). *Adansonia, série 3* **31**: 311–339.
- Huelsenbeck JP, Ronquist F. 2001.** MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.
- Hutchinson J. 1959.** *The families of flowering plants. Vol I. Dicotyledons*, 2nd ed. Oxford: Clarendon Press.
- Hutchinson J. 1973.** *The families of flowering plants arranged according to a new system based on their probable phylogeny*. Oxford: Clarendon Press.
- Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.
- Lowry PP II, Haevermans T, Labat JN, Schatz GE, Leroy JF, Wolf AE. 2000.** Endemic families of Madagascar. V. A synoptic revision of *Eremolaena*, *Pentachlaena*

- and *Perrierodendron* (Sarcolaenaceae). *Adansonia, série 3* **22**:11–31.
- Lowry PP II, Nusbaumer L, Randrianasolo A, Schatz GE, Hong-Wa C. 2014.** Endemic families of Madagascar. XIII. New, restricted range species of *Eremolaena* Baill. and *Schizolaena* Thouars (Sarcolaenaceae). *Candollea* **69**: 183–193.
- Lowry PP II, Rabehevitra D. 2006.** Endemic families of Madagascar. IX. A new littoral forest species of *Schizolaena* (Sarcolaenaceae). *Adansonia, série 3* **28**:149–153.
- Lowry PP II, Schatz GE, Leroy JF, Wolf AE. 1999.** Endemic families of Madagascar. III. A synoptic revision of *Schizolaena* (Sarcolaenaceae). *Adansonia, série 3* **21**:183–212.
- Lowry PP II, Schatz GE, Wolf AE. 2002.** Endemic families of Madagascar. VIII. A synoptic revision of *Xyloolaena* Baill. (Sarcolaenaceae). *Adansonia, série 3* **24**:7–19.
- Madagascar Catalogue. 2016.** *Catalogue of the vascular plants of Madagascar*. St. Louis and Antananarivo: Missouri Botanical Garden. <http://www.tropicos.org/project/mada> (Accessed 26 February 2016).
- McNeill J, Barrie FR, Buck WR, Demoulin V, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Marchal K, Prado J, van Prud'homme Reine WF, Smith GF, Wiersema JH, Turland NJ. 2012.** *International Code of Nomenclature for algae, fungi, and plants (Melbourne Code): adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011*. Regnum Vegetabile 154. Oberreifenberg: Koelz Scientific Books.
- Miller MA, Pfeiffer W, Schwartz T. 2010.** Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environment Workshop 14 November 2010, New Orleans*, 1–8.
- Morton CM, Dayanandan S, Dissanayake D. 1999.** Phylogeny and biosystematics of *Pseudomonotes* (Dipterocarpaceae) based on molecular and morphological data. *Plant Systematics and Evolution* **216**: 197–205.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000.** Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858.
- Nilsson S, Coetzee J, Grafstrom E. 1996.** On the origin of the Sarcolaenaceae with reference to pollen morphological evidence. *Grana* **35**: 321–334.
- Nilsson S, Randrianasolo A. 1999.** Morphology and functional aspects of pollen in the Sarcolaenaceae. *Palaeoecology of Africa* **26**: 191–200.
- Nylander JAA. 2004.** *MrModeltest v2*. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. <https://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html>
- Phillipson PB, Schatz GE, Lowry PP II, Labat JN. 2006.** A catalogue of the vascular plants of Madagascar. In: Ghazanfar SA, Beentje HJ, eds. *Taxonomy and ecology of African plants, their conservation and sustainable use: Proceedings of the 17th AETFAT Congress, Addis Ababa, Ethiopia*. Kew: Kew Publishing, 613–627.
- Rabehevitra D, Lowry PP II. 2009.** Endemic families of Madagascar. XI. A new critically endangered species of *Schizolaena* (Sarcolaenaceae) from Tapia woodlands in south-central Madagascar. *Adansonia, série 3* **31**:149–155.
- Ramananjanahary RH, Fraiser CL, Lowry PP II, Rajaonary FA, Schatz GE. 2010.** *Madagascar's endemic plant families: species guide*. Antananarivo: Missouri Botanical Garden Madagascar Research and Conservation Program.
- Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014.** Tracer v1.6. Program distributed by the author. <http://beast.bio.ed.ac.uk/Tracer>
- Randrianasolo A, Miller JS. 1994.** *Sarcolaena isaloensis*, a new species of Sarcolaenaceae from Isalo, south-central Madagascar. *Novon* **4**: 290–292.
- Randrianasolo A, Miller JS. 1999.** Revision of the genus *Sarcolaena* (Sarcolaenaceae). *Annals of the Missouri Botanical Garden* **86**: 702–722.
- Schatz GE, Lowry PP II, Wolf AE. 2000.** Endemic families of Madagascar. VI. A synoptic revision of *Rhodolaena* (Sarcolaenaceae). *Adansonia, série 3* **22**:239–252.
- Schatz GE, Lowry PP II, Wolf AE. 2001.** Endemic families of Madagascar. VII. A synoptic revision of *Leptolaena* Thouars *sensu stricto* (Sarcolaenaceae). *Adansonia, série 3* **23**:171–189.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007.** Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94**: 275–288.
- Soulebeau A, Pellens R, Lowry PP II, Aubriot X, Evans M, Haevermans T. 2016.** Conservation of phylogenetic diversity in Madagascar's largest endemic plant family, Sarcolaenaceae. In: Pellens R, Grandcolas P, eds. *Biodiversity conservation and phylogenetic systematics: preserving our evolutionary heritage in an extinction crisis*. New York: Springer, 355–374.
- Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.
- Straka H. 1963.** Betrachtungen zur Phylogenie der Sarcolaenaceae (Chlaenaceae). *Berichte der Deutschen Botanischen Gesellschaft* **76**: 55–62.
- Straka H. 1964.** Palynologia madagassica et mascarenica. Fam. 126, Sarcolaenaceae (Chlaenaceae). *Pollen et Spores* **6**: 289–301.
- Straka H. 1971.** Über das System der madagassischen Sarcolaenaceae. *Berichte der Deutschen Botanischen Gesellschaft* **84**: 731–735.
- Straka H, Friedrich B. 1983.** Familie 121–127, Microscopie électronique à balayage et addenda. *Pollen et Spores* **25**: 63–70.
- The Plant List. 2016.** <http://www.theplantlist.org/> (Accessed 26 February 2016).

Turland N. 2013. The code decoded. A user's guide to the International Code of Nomenclature for algae, fungi, and plants. Königstein: Koeltz Scientific Books.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand D, Sninsky J, White TJ, eds. *PCR protocols: a guide to methods and applications*. San Diego: Academic Press, 315–322.

Appendix 1

Summary of species, country, voucher (collector and number), herbarium acronym where deposited and GenBank accession numbers for taxa used in this study, provided in the following order: ITS, *psaA-ORF170*, *psbA-trnH* and *rbcL*. For four of the out-group species, *Anisoptera marginata*, *Cotylelobium lanceolatum*, *Hopea mengarawan* and *Neobalanocarpus heimii*, the data available in GenBank for ITS1 and ITS2 were archived separately; both were used in the analyses and the two corresponding accession numbers are indicated. Dashed lines indicate that the region was not sampled for this accession. Newly generated sequences are indicated with an asterisk following the accession number.

Anisoptera marginata Korth. Chase 2486 (K), –, –, Y15144. *A. marginata* Korth., Indonesia, KY 851 (BO), AY026638/AY026695, –, –, -. *Cistus symphytifolius* Lam., Spain, Guzman 143BGA04 (MA), DQ092940, –, –, FJ225864. *Cotylelobium lanceolatum* Craib No voucher –, –, AB452429, -. *C. lanceolatum* Craib, Indonesia, KY 871 (BO), AY026639/AY026696, –, –, -. *Eremolaena rotundifolia* (F. Gérard) Danguy, Madagascar, Andrianjafy et al. 33 (MO), KX588809*, –, KX588731*, -. *Fumana thymifolia* (L.) Spach ex Webb, Portugal, Guzman 53BGA04 (MA), DQ092926, –, –, FJ225850. *Halimium umbellatum* Spach, Spain, Guzman 71BGA04 (MA), DQ092930, –, –, FJ225857. *Helianthemum scopulicola* L.Sáez, Rosselló & Alomar, Cultivated, Guzman 67BGA04 (MA), DQ092928, –, –, FJ225852. *Hopea mengarawan* Miq. No voucher –, –, AB452465, -. *H. mengarawan* Miq., Singapore, Lum BT_0070230681 –, –, –, KJ594748. *H. mengarawan* Miq., Indonesia, KY 817 (BO), AY026650/AY026708, –, –, -. *Leptolaena abrahamii* G.E.Schatz & Lowry, Madagascar, Andrianjafy et al. 61 (MO), KX588821*, KX588860*, KX588742*, KX588781*. *L. cuspidata* Baker, Madagascar, Andrianjafy et al. 68 (MO), KX588801*, KX588842*, –, KX588762*. *L. delphinensis* G.E.Schatz & Lowry, Madagascar, Andrianjafy et al. 85 (MO), KX588823*, KX588862*, KX588744*, KX588783*. *L. gautieri* G.E.Schatz & Lowry, Madagascar, Andrianjafy et al. 86 (MO), KX588822*, KX588861*, KX588743*, KX588782*. *L. masoalensis* G.E.Schatz & Lowry, Madagascar, Andrianjafy et al. 89 (MO),

KX588803*, KX588843*, –, KX588764*. *L. multiflora* Thouars, Madagascar, Antilahimena & Aridy 1094 (MO), KX588824*, KX588863*, –, KX588784*. *L. pauciflora* Baker, Madagascar, Birkinshaw et al. 823 (MO), KX588818*, KX588857*, KX588739*, KX588778*. *L. raymondii* G.E.Schatz & Lowry, Madagascar, Hong-Wa 54 (MO), KX588802*, –, –, KX588763*. *Mediusella arenaria* (F.Gérard) Hong-Wa, Madagascar, Hong-Wa 78 (MO), KX588804*, KX588844*, KX588725*, KX588765*. *M. bernieri* (Baill.) Hutch., Madagascar, Hong-Wa 97 (MO), KX588805*, KX588845*, KX588726*, KX588766*. *Monotes glaber* Sprague, Zimbabwe, OM2130 (NH), –, –, –, JX572785. *Neobalanocarpus heimii* (King) P.S.Ashton No voucher –, –, AB452515, -. *N. heimii* (King) P.S.Ashton, Malaysia, 28645 (KEP), AY026657/AY026713, –, –, -. *Pakaraimaea dipterocarpacea* Maguire & P.S.Ashton, Venezuela, Ashton BM03M22 (LG), –, –, –, DQ406587. *Parashorea chinensis* Wang Hsie No voucher KR532475, –, KR534161, KR534161. *Pentachlaena latifolia* H.Perrier, Madagascar, Hong-Wa 251 (MO), KX588831*, KX588869*, KX588751*, KX588792*. *P. orientalis* Capuron, Madagascar, Hong-Wa 268 (MO), KX588826*, –, KX588746*, KX588786*. *Perrierodendron boinense* (H.Perrier) Cavaco, Madagascar, Hong-Wa 330 (MO), KX588835*, –, –, -. *P. capuronii* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz, Madagascar, Hong-Wa 380 (MO), KX588832*, –, KX588752*, -. *P. quartzitorum* J.-F.Leroy, Lowry, Haev., Labat & G.E.Schatz, Madagascar, Lowry II et al. 5170 (MO), KX588819*, KX588858*, KX588740*, KX588779*. *Pseudomonotes tropenbosii* A.C.Londoño, E.Alvarez & Forero No voucher –, –, –, AF030238. *Rhodolaena acutifolia* Baker, Madagascar, Lowry II et al. 5194 (MO), KX588828*, KX588866*, KX588750*, KX588789*. *R. altivola* Thouars, Madagascar, Lowry II et al. 6016 (MO), KX588796*, KX588838*, –, KX588758*. *R. bakeriana* Baill., Madagascar, Malcomber et al. 2843 (MO), KX588814*, KX588853*, KX588735*, KX588774*. *R. coriacea* G.E.Schatz, Lowry & A.-E.Wolf, Madagascar, McPherson & Rabenantoandro 18395 (MO), KX588798*, KX588840*, –, KX588760*. *R. humblotii* Baill., Madagascar, Miller et al. 8745 (MO), KX588795*, KX588837*, KX588723*, KX588757*. *R. leroyana* G.E.Schatz, Lowry & A.-E.Wolf, Madagascar, Miller et al. 8812 (MO), KX588797*, KX588839*, KX588724*, KX588759*. *R. macrocarpa* G.E.Schatz, Lowry & A.-E.Wolf, Madagascar, Schatz 3640 (MO), KX588794*, –, KX588722*, KX588756*. *Sarcolaena eriophora* Thouars, Madagascar, Schatz 3947 (MO), KX588811*, KX588850*, KX588733*, KX588771*. *S. multiflora* Thouars, Madagascar, Schatz 3959 (MO), KX588827*,

- KX588865*, KX588748*, KX588787*. *S. oblongifolia* F.Gérard, Madagascar, *Schatz* 3960 (MO), KX588812*, KX588851*, KX588734*, KX588772*.
- Schizolaena cauliflora* Thouars, Madagascar, *Schatz* 3964 (MO), KX588813*, KX588852*, –, KX588773*. *S. elongata* Thouars, Madagascar, *Schatz* 3973 (MO), KX588825*, KX588864*, KX588745*, KX588785*. *S. exinvolucrata* Baker, Madagascar, *Schatz* 3982 (MO), –, –, KX588747*, –.
- S. hystrix* Capuron, Madagascar, *Schatz* 3991 (MO), KX588799*, –, –, *S. manomboensis* Lowry, G.E. Schatz, J.-F.Leroy & A.-E.Wolf, Madagascar, *Schatz et al.* 4185 (MO), KX588793*, KX588836*, KX588721*, KX588755*.
- S. masoalensis* Lowry, G.E.Schatz, J.-F.Leroy & A.-E.Wolf, Madagascar, *Rabenantoandro et al.* 272 (MO), KX588800*, KX588841*, –, KX588761*.
- S. microphylla* H.Perrier, Madagascar, *Rabenantoandro et al.* 563 (MO), KX588810*, KX588849*, KX588732*, KX588770*.
- S. noronhae* G.E.Schatz & Lowry, Madagascar, *Rabevohitra et al.* 3821 (MO), KX588830*, KX588868*, –, KX588791*.
- S. rosea* Thouars, Madagascar, *Rakotondrajaona* 143 (MO), KX588829*, KX588867*, –, KX588790*.
- S. tampoketsana* Lowry, G.E.Schatz, J.-F.Leroy & A.-E.Wolf, Madagascar, *Rakotonirina et al.* 302 (MO), KX588807*, KX588847*, KX588729*, KX588768*.
- Shorea robusta* C.F.Gaertn. No voucher KM514673, –, JX856942, JX856762.
- Tuberaria guttata* (L.) Fourr., Portugal, *Guzman* 44BGA04 (MA), DQ092929, –, –, FJ225853.
- Xerochlamys bojeriana* (Baill.) F.Gérard, Madagascar, *Randrianasolo* 223 (MO), KX588820*, KX588859*, KX588741*, KX588780*.
- X. coriacea* Hong-Wa, Madagascar, *Randrianasolo* 241 (MO), KX588834*, –, KX588754*, –.
- X. diospyroidea* (Baill.) F.Gérard, Madagascar, *Ratovoson* 307 (MO), KX588817*, KX588856*, KX588738*, KX588777*.
- X. elliptica* F.Gérard, Madagascar, *Razafindrakoto et al.* 46 (MO), KX588806*, KX588846*, KX588727*, KX588767*.
- X. itremoensis* Hong-Wa, G.E.Schatz & Lowry, Madagascar, *Razafindrakoto et al.* 70 (MO), KX588816*, KX588855*, KX588737*, KX588776*.
- X. tampoketsensis* F.Gérard, Madagascar, *Razafitsalama* 1 (MO), KX588815*, KX588854*, KX588736*, KX588775*.
- X. undulata* Hong-Wa, Madagascar, *Razafitsalama* 63 (MO), –, –, KX588749*, KX588788*.
- X. villosa* F.Gérard, Madagascar, *Lowry II et al.* 6968 (MO), –, –, KX588728*, –.
- Xyloolaena humbertii* Cavaco, Madagascar, *Lowry II et al.* 6966 (MO), KX588833*, –, KX588753*, –.
- X. richardii* (Baill.) Baill., Madagascar, *Razafitsalama* 31 (MO), KX588808*, KX588848*, KX588730*, KX588769*.