RESEARCH ARTICLE

One in, one out: Generic circumscription within subtribe Manilkarinae (Sapotaceae)

Aina Randriarisoa,[1,2] Yamama Naciri,[1,2] Kate Armstrong,[3] Carlos Galan Boluda,[1,2] Stephanie Dafreville,[4] Charles Pouchon[1,2] & Laurent Gautier[1,2]

1 *Conservatoire & Jardin botaniques de Genève, Chemin de l'Impératrice 1, 1292 Chambésy, Geneva, Switzerland*
2 *Laboratoire de Systématique végétale et Biodiversité, Department of Botany and Plant Biology, Université de Genève, Chemin de l'Impératrice 1, 1292 Chambésy, Geneva, Switzerland*
3 *Institute of Systematic Botany, New York Botanical Garden, 2900 Southern Boulevard, Bronx, New York 10458, U.S.A.*
4 *UMR PVBMT, 97410 Saint-Pierre, La Réunion island, France*
Address for correspondence: *Yamama Naciri, yamama.naciri@ville-ge.ch*

DOI https://doi.org/10.1002/tax.12863

Abstract Previous phylogenetic studies have demonstrated that the Manilkarinae are a monophyletic subtribe if *Northia* is excluded. The subtribe consists of four genera: *Faucherea*, *Labourdonnaisia*, *Labramia* and *Manilkara*. However, the same phylogenetic studies also raised taxonomic issues concerning unclear generic delimitations and unresolved relationships. The current study's aims are: to resolve these taxonomic issues using a molecular phylogeny based on hundreds of nuclear markers sequenced from a representative sampling of taxa across the four genera; to find relevant morphological characters allowing the distinction of the clades retrieved with the phylogeny; and finally to understand the evolutionary history of the subtribe by conducting a divergence time estimation and ancestral state reconstructions. Our phylogeny shows a well-resolved backbone with four main lineages: the *Labramia* clade, the main clade of *Manilkara*, a clade in which all species of *Labourdonnaisia* and *Faucherea* are mixed, and a clade of three Pacific *Manilkara* species. The main clade of *Manilkara* is retrieved as sister to *Labramia*, and the *Labourdonnaisia-Faucherea* clade is clearly assessed as sister to the three Pacific *Manilkara* species. As a consequence, *Faucherea* is synonymized with *Labourdonnaisia*, and the three Pacific *Manilkara* are considered to be a separate genus, for which the name *Abebaia* is resurrected. We provide emended descriptions for *Labourdonnaisia* and *Abebaia* as well as the necessary new combinations. The ancestral state reconstruction of flower characters shows that ancestral Manilkarinae were characterized by a hexamerous corolla, well-developed dorsal appendages and staminodes, and a pubescent ovary. These character states have been retained in the main *Manilkara* clade, but surprisingly also in *Abebaia*, which appears as a cryptic genus. The lack of dorsal appendages and the reduction of staminodes observed in *Labourdonnaisia* appeared after the split from *Abebaia*. The increase in corolla merism observed mainly in the Mascarene *Labourdonnaisia*, which was used to separate it from *Faucherea*, appears to be a derived state, which evolved separately in a few species during the radiation of *Labourdonnaisia* on Madagascar and the Mascarenes. The glabrous ovary state observed in *Labramia* also constitutes a derived synapomorphic state in the genus.

Keywords ancestral state reconstruction; integrative taxonomy; Madagascar; Pacific Islands; target gene capture

Supporting Information may be found online in the Supporting Information section at the end of the article.

■ INTRODUCTION

The timber of Sapotaceae is highly valued for its strength, rot-resistance and durability, which make it widely used in construction. Almost all species are characterized by slow growth and a long generation time and can be considered, from a conservation point of view, to be markers of primary forest (L. Gautier, pers. obs.). In Madagascar, Sapotaceae are affected by illegal logging and overexploitation, even in protected areas (Gautier & al., 2022). Therefore, implementing a conservation strategy is a matter of urgency. However, in the absence of robust taxonomic data, uncertainties regarding the delimitation at both generic and specific levels impede conservation action in this family.

Generic circumscription in the Sapotaceae has always been a matter of debate. Although recognizing the family is easy due to the remarkable homogeneity of its morphological vegetative characters, its classification is still highly controversial at lower taxonomic ranks, especially at the generic level (Pennington, 1991; Anderberg & Swenson, 2003). Aubréville's (1964) generic monograph was followed by that of Baehni (1965) and Pennington's (1991). Each of them considered a different classification system leading to a highly unstable number of genera, varying from 53 to 122 depending on the author (Gautier & al.,

Article history: Received: 31 Mar 2022 | returned for (first) revision: 16 May 2022 | (last) revision received: 23 Sep 2022 | accepted: 29 Sep 2022
Associate Editor: Dianxiang Zhang | © 2022 The Authors.
TAXON published by John Wiley & Sons Ltd on behalf of International Association for Plant Taxonomy.

2013; Swenson & al., 2013). Currently, 65–70 genera are accepted (Swenson & al., 2020). Discrepancies in generic delimitation are due to the large number of species that display an "often overlapping distribution of many character states" (Anderberg & Swenson, 2003). According to Pennington (1991), "characters unique to a genus are extremely rare in Sapotaceae, so the use of a single character to define genera causes instability, depending on which character is selected". Consequently, using a dichotomous key is almost impossible and does not enable one to adequately discern between taxa, leading to an "artificial classification" (Pennington, 1991).

Nonetheless, advances in molecular studies have greatly improved our understanding of evolutionary relationships and classification within Sapotaceae, especially from subfamily to generic levels (Anderberg & Swenson, 2003; Swenson & Anderberg, 2005; Gautier & al., 2013; Stride & al., 2014; Borg & al., 2019). The currently accepted classification includes three subfamilies (Swenson & Anderberg, 2005): Chrysophylloideae, Sapotoideae and Sarcospermatoideae. Within the Sapotoideae, four tribes are currently recognized (Gautier & al., 2013): Isonandreae, Sapoteae, Sideroxyleae and Tseboneae, but several genera still remain outside these tribes.

Tribe Sapoteae corresponds to Mimusopeae sensu Pennington (1991) with the exclusion of the subtribe Gluemineae. It consists of genera with a calyx of two whorls of sepals, corolla lobes usually with lateral appendages, and stamens isomerous with the corolla lobes. In Pennington's (1991) system, genera with 4 + 4 sepals constitute subtribe Mimusopinae and genera with 3 + 3 sepals belong to subtribe Manilkarinae (equivalent to the Manilkarées tribe of Aubréville, 1964), which comprises six genera.

A study focusing on the genus *Manilkara*, but including a limited number of specimens belonging to the subtribe Manilkarinae, was conducted using nuclear (ITS) and chloroplast loci (*rpl32-trnL*, *rps16-trnK*, *trnS-trnFM*) separately. The phylogenetic reconstruction using the ITS dataset demonstrated the monophyly of subtribe Manilkarinae (Armstrong, 2010; Armstrong & al., 2014) excluding *Northia* Hook.f., which has been demonstrated to be even outside Sapoteae using ITS and the chloroplast locus *trnH-psbA* (Gautier & al., 2013; Armstrong & al., 2014). Accordingly, it had already been spotted as the only Manilkarinae genus lacking endosperm in the seeds (Pennington, 1991). Regarding the monotypic *Letestua* Lecomte, it has been recovered within *Manilkara* Adans. (Smedmark & al., 2006; Armstrong 2010), despite the fact that Smedmark & al. (2006) casted doubt about the identity of the sampled vouchers.

Out of the four remaining genera, three are endemic to the Western Indian Ocean Islands biodiversity hotspot (Aubréville, 1974; Dafreville, 2013): *Labramia* A.DC. is a Malagasy subendemic genus, with 10 described species, including one in the Comoros (Aubréville, 1974; Labat & al., 1997; Randriarisoa & al., 2020); *Labourdonnaisia* Bojer is known from six species – three in the Mascarenes and three in Madagascar (Aubréville, 1974; Friedmann, 1981); and *Faucherea* Lecomte is endemic to Madagascar with 11 described species (Aubréville, 1974). The notable exception is *Manilkara*, with 78 species distributed widely across the tropics: 30 in Africa, 5 in Madagascar, 13 in the Asia-Pacific and 30 in South and Central America (Armstrong & al., 2014). In the following, we will use the Manilkarinae subtribe as restricted above with only four genera.

The genus *Labramia* is clearly defined by the following combination of morphological characters: a glabrous ovary, the presence of staminodes, a pair of appendages at the base of each corolla lobe, and a corolla tube often as long as the corolla lobes (Aubréville, 1974; Pennington, 1991). *Faucherea* and *Labourdonnaisia* are more problematic: they have very similar morphological characteristics including the absence of dorsal appendages, a pubescent ovary, reduced staminodes and a basiventral seed scar. The only notable distinction between them is the merosity of the corolla and androecium, 6–11 in *Faucherea* versus 10–18 in *Labourdonnaisia* (Lecomte, 1920; Aubréville, 1964). Moreover, the status of *Labourdonnaisia* in Madagascar is considered to be in great need of revision (Friedmann, 1981): only 10 collections of *Labourdonnaisia* were cited in the Flora of Madagascar treatment (Aubréville, 1974), and not even one of them displayed mature fruits. With the recent increase in available collections, numerous new putative species have been found in this group, but in the absence of flowers, it is very hazardous to decide if they belong to *Faucherea* or *Labourdonnaisia*. Moreover, *Faucherea* and *Labourdonnaisia* were recovered as paraphyletic in the recent molecular studies (Armstrong, 2010; Armstrong & al., 2014). A further unexpected taxonomic issue was that *Manilkara*, as traditionally circumscribed, was not monophyletic: two non-sister lineages were detected. The main *Manilkara* lineage, including the type species, was retrieved sister to all other Manilkarinae. A second *Manilkara* lineage, consisting of three Pacific species, was resolved as sister to the *Labourdonnaisia* + *Faucherea* clade. However, support values were low for some key nodes and no morphological differences were found between the two *Manilkara* clades.

As preliminary steps towards a future revision of these genera in Madagascar, the main aims of this study are to resolve the generic circumscription of subtribe Manilkarinae, to discern relevant morphological characters that match this circumscription, and finally to understand the evolutionary history of the subtribe. These main goals have been approached by integrating: (a) a molecular phylogenetic reconstruction based on hundreds of nuclear sequences on a sampling more representative of the Madagascar region endemic genera; (b) an exploration of the phylogenetic tree space; (c) a divergence time estimation; (d) a morphological analysis based on flower characters; and (e) an ancestral state reconstruction (ASR) combining the output of the former analyses. It is however by no means our intention to resolve species delimitation issues within these genera, which will be dealt with in forthcoming papers using an extended sampling.

■ MATERIALS AND METHODS

Specimen identification. — As a first step, all recent Malagasy collections belonging to subtribe Manilkarinae were

compared to specimens hosted in the herbaria G, MO, P, TAN and TEF, and were assigned to one of the four genera. Within each genus, specimens were grouped following a splitter approach, based on morphology, using both reproductive and vegetative characters. These groups were used to build species hypotheses that will be further referred to as "morphospecies". If a type specimen was included, the morphospecies received its species name, if not, an incremental number was assigned (e.g., *Labramia* sp. 1, *Labramia* sp. 2. and so on).

Taxon sampling for molecular study. — All material came either from herbarium specimens (64%) or from silica gel-dried leaf samples (36%) collected in parallel to voucher herbarium specimens in the field. A total of 89 specimens were used for the molecular analyses: 84 specimens in the subtribe Manilkarinae plus 5 outgroups selected from elsewhere in the Sapoteae tribe based on the results of Anderberg & Swenson (2003) and Armstrong (2010): *Baillonella toxisperma* Pierre, *Mimusops* cf. *antorakensis* Aubrév., *Tieghemella heckelii* (A.Chev.) Pierre ex Dubard, *Vitellaria paradoxa* C.F.Gaertn., and *Vitellariopsis cuneata* (Engl.) Aubrév. The ingroup consisted of 44 morphospecies, including 36 recognized species, 1 variety and 7 undescribed morphospecies. Samples of *Faucherea*, *Labourdonnaisia* and *Labramia* were chosen to represent the morphological diversity of the genera as exhaustively as possible. Within *Manilkara*, the main lineages of the genus core were represented, plus the three Pacific species retrieved as a separate lineage: *M. dissecta* Dubard, *M. fasciculata* (Warb.) H.J.Lam & Maas Geest. and *M. udoido* Kaneh (Armstrong, 2010; Armstrong & al., 2014). The type species of each genus, as currently circumscribed, were included (Table 1), with the exception of *Manilkara kauki* (L.) Dubard. However, this species is nested within the *Manilkara* s.str. clade (Armstrong, 2010; Armstrong & al., 2014), which was represented by 10 species in our sampling. All the type species of genera currently synonymized with *Manilkara* were analyzed earlier by Armstrong (2010) and Armstrong & al. (2014). All but one were confirmed to belong to the *Manilkara* s.str. clade, so we did not include them in our sampling. The only exception is *M. fasciculata* (Warb.) H.J.Lam & Maas Gest., type of the genus *Abebaia* Baehni, which is represented in our sampling. Conversely *Faucherea*, *Labourdonnaisia* and *Labramia* do not include, in their current circumscription, any other genus whose type species would have needed to be analyzed.

Each morphospecies is represented by one to four samples, with the following exceptions, to account for possible cryptic taxa: six accessions of the morphologically variable species *Labourdonnaisia madagascariensis* Pierre ex Baill., and eight of *Labramia bojeri* A.DC. because its geographic distribution spreads all along the east coast of Madagascar, from the extreme north to the extreme south (Appendix 1).

As far as allowed by the availability of material, at least one flowering specimen per accepted species was selected in order to be able to link fertile characters with molecular data.

DNA extraction. — Leaves were disrupted using Qiagen TissueLyser II. Genomic DNA was extracted using the CTAB method (Doyle & Doyle, 1987), with some modifications, including a Sorbitol pre-treatment to remove mucilaginous polysaccharides (Souza & al., 2012). A fragment analyzer Qsep100, that uses a capillary gel electrophoresis system, allowed us to estimate the average DNA fragment size of each sample, which was visualized afterwards with the Q-Analyzer v.2.0.0.0 software.

Library construction, target capture enrichment, and sequencing. — The appropriate DNA fragment size for next-generation sequencing (NGS) is expected to range from 200 to 600 bp. Therefore, DNA with longer fragment sizes had to be cut first using a sonicator QSonica machine Q800R3. A genomic library of each specimen was constructed and labeled with dual-indexed primers (Kircher & al., 2012) from the NEXTflex Barcodes kit (BIOO Scientific Austin, Texas U.S.A.). Equimolar quantities from each library were then pooled for the gene capture process.

Target capture was performed following the methodology described in Christe & al. (2021). A total of 792 protein-coding genes were targeted on the 89 specimens, using the probe set developed for Sapotaceae (Christe & al., 2021). Hybridization reactions were performed according to the standard protocol of MyBaits v.5.01, for 24 hours. The targeted sequences were captured using a hybridization step with specific biotinylated oligonucleotide probes complementary to the loci of interest. Hybridized sequences were retained on streptavidin-covered magnetic beads while all non-target DNA was washed away. Sequencing was processed on a HiSeq4000 Illumina machine (100 bp pair-end reads) at iGE3, the genomic platform of the University of Geneva.

Capture data processing. — Quality controls were first performed on the captured data using FastQC v.0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Then, trimmomatic v.0.38 (Bolger & al., 2014) was used to remove the adaptors. The program HybPiper v.1.3 (Johnson & al., 2016) was run to map the reads against the reference sequence, targeting exons sequences of the 792 nuclear genes.

Sequences were concatenated using the AMAS software v.1.0.0 (Borowiec, 2016), in order to calculate the percentage of missing data with BioEdit v.7.2.5 (Hall, 1999). Filtering treatments were processed using trimAl v.1.4 (Capella-Gutiérrez & al., 2009) to remove putative positions with more than 20% of missing data within exons across the specimens retained.

Putative paralogous sequences were identified and removed using the "FilterParalogs.py" python function available within the ORTHOSKIM software v.1.0 (Pouchon & al., 2022). This function is based on a sliding-window approach to detect and remove hyper-variable sites and regions within the alignments files according to a consensus sequence. We applied this function on the gene alignments previously trimmed by trimAl by inferring the consensus at the genus level (-q genus). We also used a 100 nt sliding window (-w 100) with a minimal number of 15 polymorphic sites (-p 15), which represented >15% variable sites.

Phylogenetic reconstruction. — Consensus sequences of the successfully mapped loci were obtained and aligned

using MAFFT v.7 (Katoh & Standley, 2013). We generated one gene tree for each locus with RAxML v.8.2.4 (Stamatakis, 2014). Discordance between gene trees and the species tree can be due to incomplete lineage sorting, especially at species and population levels, or to hybridization among species (Naciri & Linder, 2015). At higher taxonomic levels, ancient incomplete lineage sorting can be a challenge because of ancestral polymorphism (Shi & Yang, 2017). For the phylogenetic reconstruction, and given our taxonomic aims, we opted for ASTRAL-II v.5.7.3 (Mirarab & al., 2014; Mirarab & Warnow, 2015). This methodology infers the species tree using individual gene trees obtained with RAxML. Although ASTRAL is considered to be only a pseudo-coalescent method, many studies argue for its reliability and accuracy when inferring species trees, especially with significant impact of incomplete lineage sorting (Sayyari & Mirarab, 2016; Shi & Yang, 2017; Mirarab, 2019). Furthermore, unlike Bayesian approaches that need considerable computing time, ASTRAL is able to handle large phylogenies based on several hundreds of genes. The analyses were computed on the Baobab cluster of the University of Geneva, and the resulting ASTRAL tree was visualized using Figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Exploration of the phylogenetic tree space. — Phylogenetic landscape of gene trees was examined in order to check for alternative evolutionary histories carried by the different genes. We first computed a pairwise matrix of topological distances between each pair of gene trees inferred with RAxML by using the normalized Robinson-Foulds distance as implemented in "phangorn" R package v.2.8.1 (Schliep, 2011). To do so, each pair of gene trees was rooted using the outgroup taxa and trimmed to the common tip names using "ape" R package v.5.6.1 (Paradis & Schliep, 2019). Next we used a classical multidimensional scaling method with a k-means

Table 1. Ingroup genera with distribution and described species included in this study.

Genus, with number of accepted species	Distribution	Species included in this study
Faucherea Lecomte 11 species	Madagascar	*Faucherea ambrensis* Aubrév. *Faucherea glutinosa* Aubrév. ***Faucherea hexandra* (Lecomte) Lecomte** *Faucherea laciniata* Lecomte *Faucherea manongarivensis* Aubrév. *Faucherea parvifolia* Lecomte *Faucherea tampoloensis* Aubrév. *Faucherea thouvenotii* Lecomte *Faucherea urschii* Capuron ex Aubrév.
Labourdonnaisia Bojer 6 species	Mascarenes Islands and Madagascar	***Labourdonnaisia calophylloides* Bojer** *Labourdonnaisia glauca* Bojer *Labourdonnaisia lecomtei* Aubrév. *Labourdonnaisia madagascariensis* Pierre ex Baill. *Labourdonnaisia revoluta* Bojer
Labramia A.DC. 10 species	Madagascar and Comoros	*Labramia ankaranaensis* Aubrév. *Labramia ankaranaensis* var. *antsingensis* Aubrév. *Labramia boivinii* (Pierre) Aubrév. ***Labramia bojeri* A.DC.** *Labramia capuronii* Aubrév. *Labramia costata* (M.M.Hartog ex Baill.) Aubrév. *Labramia louvelii* Aubrév. *Labramia mayottensis* Labat, Pignal & O.Pascal *Labramia platanoides* Capuron ex Aubrév. *Labramia sambiranensis* Aubrév.
Manilkara Adans. s.str. 78 species	Pantropical (Africa, America, Asia, Pacific Asia)	*Manilkara bidentata* (A.DC.) A.Chev. *Manilkara boivinii* Aubrév. *Manilkara cuneifolia* (Baker) Dubard *Manilkara dissecta* (L.f.) Dubard *Manilkara fasciculata* (Warb.) H.J.Lam & Maas Geest. **(type species of the genus *Abebaia*)** *Manilkara hexandra* (Roxb.) Kuntze *Manilkara lacera* (Baker) Dubard *Manilkara longifolia* (A.DC.) Dubard *Manilkara multinervis* (Baker) Dubard *Manilkara obovata* (Sabine & G.Don) J.H.Hemsl. *Manilkara sansibarensis* (Engl.) Dubard *Manilkara udoido* Kaneh. *Manilkara zapota* (L.) P.Royen

Type species of each genus is given in bold. The type species of *Manilkara*, *M. kauki* Adans., is not represented in the sampling.

clustering approach to identify gene families based on the distance matrix (MacQueen, 1967).

To visualize the topological distances among trees, we first computed the principal coordinates of the distances with the "cmdscale" R function of the package stats v.3.6.3. Then, we applied the "kmeans" function on these coordinates to partition them into *k* predefined groups. The optimal *k* was identified by using the silhouette method of the "fviz_nbclust" function from the factoextra R package v.1.0.7 (Kassambara & Mundt, 2020), with *k* ranging from 1 to 10. This method chooses automatically the *k* value that maximizes the average silhouette over the range of *k* (Kaufman & Rousseeuw, 1990).

In addition, we also performed a hierarchical clustering approach by using the "heatmap" R function of the package stats v.3.6.3 directly on the topological distance matrix. This allows to visualize the matrix by simultaneously re-ordering the columns/rows, i.e., gene trees, according to a dendrogram based on the similarity between them, while coloring the cells according to their values in the distance matrix. This approach depicts how gene trees vary across the phylogenetic space.

To summarize the topology for each cluster, we inferred a respective species tree from all gene trees assigned to each of the clusters using ASTRAL-II v.5.7.3. A comparative graphic representation was drawn using the "tanglegram" R function of dendextend v.1.15.1 (Galili, 2015) to check for topological differences among species trees.

Divergence time estimation. — We used BEAST (Suchard & al., 2018) to estimate divergence times. Due to computing time limitations when running on a large dataset, we had to select a subsample of genes. We randomly selected 20 genes with a percentage of variable sites higher than the median value. The input file was prepared with BEAUTI2 (Bouckaert & al., 2014). Primary calibration points from available fossils data were used: a *Tetracolporpollenites* pollen from 37.2 to 48.6 million years ago (mya) from England (Harley, 1991) to constrain the crown of the tribe Sapoteae (offset: 42.9, mean 0.095) and a series of fossil leaves from a putative *Manilkara* sp. from 23 to 33.9 mya from Ethiopia (Jacobs & al., 2005) to constrain the *Manilkara* s.str. crown node (offset: 28.0, mean 0.1; Armstrong & al., 2014). Site models were kept unlinked, while we linked the clock and tree models. All substitution models were set as GTR with a gamma category of 4, estimating neither the substitution rate nor the proportion of invariant sites, nor fixing the mean substitution rate, but estimating gamma shapes and rate frequencies between nucleotides, except for CT rates. A relaxed clock with lognormal distribution was used to allow rate heterogeneity between lineages. Tree prior was set as Yule model with a gamma distribution. The priors gammaShape, nucleotide rates, and ucldStdev were set with a gamma distribution, while proportionInvariant was kept as uniform. A log normal prior was used to constrain calibration points. In addition to node calibration using fossils, we constrained the main clades found as monophyletic and strongly supported in the ASTRAL analysis.

Two independent runs with the same priors and parameters were executed in BEAST v.2.5.2 (Bouckaert & al., 2014) with 600 million MCMC iterations each, sampling every 10,000th iteration. BEAST runs were computed on the Baobab cluster of the University of Geneva. Each log file was visualized on Tracer v.1.6 (Rambaut & al., 2014) to ensure that the log-likelihood values of the sample points reached an equilibrium with effective sample size (ESS) values above 200. Then the two runs were combined using LogCombiner (program included in BEAST package), discarding the first 25% trees as burn-in, to ensure that all parameters converged even with different starting points.

Morphological analyses. — In order to assess to which extent the results of the molecular phylogeny could be corroborated by morphology, we conducted a factorial analysis of mixed data on a selection of qualitative and quantitative characters and a linear discriminant analysis on the latter. Our character selection was based mainly on flower characters used by Aubréville (1964) and Pennington (1991) for delimiting genera in Sapotaceae. Further characters more specific to generic delimitation in Manilkarinae, were added based on our own observations (Table 2). As all Manilkarinae share a double trimerous calyx, calyx characters were not included. An attempt was also made to investigate fruit characters, but due to the scarcity of fruiting specimens, this was unfortunately not possible.

Character scoring was primarily performed on herbarium specimen measurements, but also partly on morphological data from the literature. Sampling consisted of 56 terminals (suppl. Table S1), including the following 51 specimens:

- 27 flowering specimens that were included in the molecular sampling;
- 3 additional flowering specimens that were chosen to represent species for which no recent flowering material was available when the molecular sampling was conducted;
- 21 specimens that were scored previously in the context of a recent taxonomic revision of Asian-Pacific *Manilkara* (Armstrong, 2013).

The five remaining terminals consisted of five clearly circumscribed species, for which character states were extracted from the literature (Aubréville, 1936; Friedmann, 1981; Pennington, 1991): *Labourdonnaisia glauca* Bojer, *L. revoluta* Bojer, *Manilkara bidentata* (A.DC.) A.Chev., *M. lacera* (Baker) Dubard and *M. zapota* (L.) P.Royen.

The morphological dataset includes 10 quantitative variables (8 continuous, 2 integer) and 3 qualitative (Table 2). In order to avoid overweighting flower size, 7 of the 8 continuous variables are ratios, the only one chosen to represent the size of the flower being the total length of the corolla (Cor.length).

The character "merosity" here refers to the number of internal whorls of the flower: corolla including dorsal appendages, stamens, and staminodes, but excluding the number of ovary locules. The importance of merism within Sapotaceae has been discussed by Kümpers & al. (2016) and has been used by Aubréville (1964, 1974) to separate *Labourdonnaisia* and *Faucherea*. This character has been coded as an integer variable with five possible states (6, 7, 8, 12, 18). The number of ovary locules was included as a separate variable, as it is

rarely related to the merosity of other elements of the flower (Kümpers & al., 2016). It has also been coded as an integer variable with five possible states (6, 8, 9, 10, 12). Both dorsal appendages and staminodes can be absent. If present, their apex can be variously dissected. Dorsal appendages have been coded as absent, entire, or laciniate (three possible states). Staminodes have been coded as absent, entire, toothed or laciniate (four possible states). Both have been treated as qualitative variables. Lack of ovary pubescence is used by Pennington (1991) as a distinctive character for *Labramia*, given that among all Manilkarinae, *Labramia* is the only genus that consistently has a glabrous ovary. This trait was coded as a binary qualitative variable (pubescent or glabrous).

In order to consider both quantitative and qualitative data in the same analysis, we used a factorial analysis for mixed data (FAMD – Hill & Smith, 1976; Escofier, 1979; Pagès, 2004). It performs a multivariate analysis with the same approach as a principal component analysis (PCA). Quantitative traits are treated as in a normal PCA, while qualitative ones are transformed into quantitative traits by dividing each character state by its respective frequency (Pagès, 2013) as in a multiple correspondence analysis (MCA).

To highlight which of the quantitative characters could best define the main clades, a linear discriminant analysis (LDA) was conducted with the "lda" function of the MASS library (Venables & Ripley, 2002). Groups were define a priori according to the main clades retrieved from the molecular phylogenetic reconstruction.

Ancestral state reconstruction. — The tree topology from the dated Bayesian analysis was reduced to 47 terminals

Table 2. Characters and character states used for the morphological analyses and the ancestral state reconstructions (ASR).

Characters	Abbreviations	Character states	Variable type (morphological analyses)	Character type (ASR)
Dorsal appendage dissection	App.Integr	Absent	Qualitative	Discrete
		Entire		
		Laciniate		
Staminode apex dissection	Stam.Integr	Absent		
		Entire		
		Toothed		
		Laciniate		
Ovary pubescence	Ov.Pub	Pubescent		
		Glabrous		
Merosity of corolla	Merosity	6	Quantitative integer	
		7		
		8		
		12		
		18		
Number of ovary locules	Ov.locules	6		
		8		
		9		
		10		
		12		
Total length of corolla (lobe + tube)	Cor.length	Millimeters	Quantitative continuous	Continuous
Corolla tube length to corolla lobe length ratio	Tube_Cor	Ratio		
Dorsal appendage length to corolla lobe length ratio	App_Cor	Ratio		
Staminode length to corolla lobe length ratio	Stam_Cor	Ratio		
Style length to corolla lobe length ratio	Style_Cor	Ratio		Not used for ancestral state reconstruction
Staminode width to length ratio	Stam.l_L	Ratio		
Corolla lobe width to length ratio	Cor.l_L	Ratio		
Dorsal appendage width to length ratio	App.l_L	Ratio		

after removing the sterile specimens (42 Manilkarinae and 5 outgroups) and used as the basis for ancestral state reconstruction (ASR). For the ingroup, the data are those of the morphological analysis. For the five outgroups, which were not analyzed in the morphological study, character states were scored from literature.

Matrices were generated on a set of targeted morphological traits. Character state changes were mapped on the pruned tree. Reconstruction of the ancestral states has been conducted using two different approaches, depending on whether the character is discrete (including qualitative and integer characters) or continuous.

For discrete characters, we used MBASR, a toolkit which performs ASR relying on the MrBayes mechanism (Heritage, 2021). MBASR provides a statistical estimate for discrete character states at ancestral nodes based on continuous-time Markov modelling against a tree's topology (our reduced dated Bayesian tree) and branch lengths (time and substitution rate). Discrete characters include the three qualitative variables that were used in the morphological analysis (ovary pubescence, dissection levels of dorsal appendages and that of staminodes) and the two integer variables (merosity and number of ovary locules). Ovary pubescence presence/absence has been coded as a binary character; dissection levels of dorsal appendages and that of staminode apices as ordered discrete characters; merosity and number of ovary locules as unordered discrete characters (Table 2).

For continuous characters, we used phytools v.0.7-80 on R, a widely used tool for comparative biology that includes methods for mapping trait evolution on trees, in order to reconstruct the ancestral states at internal nodes (Revell, 2012). The function "fastAnc" of phytools was used to estimate the most probable states at ancestral nodes for continuous characters, based on a maximum likelihood approach. Continuous characters include the dorsal appendages length to corolla lobe length ratio, the staminode length to corolla lobe length ratio, the corolla tube length to corolla lobe length ratio and the total length of the corolla. As the outgroup character states were scored from the literature, we were not able to retrieve some character values, especially those related to the width of the different flower organs. Thus, the characters Cor.l_L, App.l_L and Stam.l_L were not used for the ASR.

■ RESULTS

Target capture efficiency. — As expected, most herbarium specimens contained highly fragmented DNA. However, the average sizes were compatible with NGS and the HiSeq Illumina technology. Silica gel-dried samples from recent field-trips yielded, on average, longer fragment sizes due to their better preservation (Chase & Hills, 1991). Out of the 794-exon probe set developed by Christe & al. (2021), we retrieved 787 protein-coding genes from HybPiper, for which paralog sequences were removed. The dataset contained 89 specimens with less than 35% of missing data. The proportion of variable sites per gene varied between 0.074% and 0.689%, with an average value of 0.228% ± 0.003. The proportion of parsimonious sites per gene ranged between 0.029% and 0.422%, with an average value of 0.102% ± 0.001. Alignments are available in the BioSample database (https://www.ncbi.nlm.nih.gov/biosample) under the reference PRJNA849733.

Phylogenetic reconstruction. — The phylogenetic reconstruction of Manilkarinae shows a well-resolved backbone with two main clades (Fig. 1). The first one (PP 0.99) contains *Labramia* and *Manilkara* s.str. as sister clades. In the second one (PP 1.00), the three Pacific *Manilkara* species recovered in a separate clade by Armstrong & al. (2014) are sister to a clade containing all species of *Labourdonnaisia* and *Faucherea*.

Labramia is recovered as monophyletic with strong support (PP 1.00). Branch lengths throughout the crown of the clade are short, giving rise to a radiation-like topology (Glor, 2010) with moderate support for species relationships. In turn, and pending extension of the sampling, seven described species are well supported (PP 1.00) within the genus: *Labramia capuronii* Aubrév., *L. costata* (M.M.Hartog ex Baill.) Aubrév., *L. louvelii* Aubrév., *L. mayottensis* Labat & al., *L. platanoides* Capuron ex Aubrév., *L. sambiranensis* Aubrév. and *L. ankaranaensis* Aubrév. The latter species includes a well-supported clade with the variety *antsingensis* (PP 1.00). *Labramia bojeri* A.DC. is recovered as a moderately supported clade (PP 0.90) and contains two distinct strongly supported lineages: the first is comprised of the majority of the *L. bojeri* specimens (PP 1.00), while the second only contains specimens of *L. bojeri* from the southeastern part of Madagascar (PP 1.00). *Labramia boivinii* (Pierre) Aubrév. is represented by a single specimen. Five potential new species stand out as isolated and monophyletic (PP 1.00) in this phylogenetic tree: *Labramia* sp. 1, *L.* sp. 3, *L.* sp. 4, *L.* sp. 6 and *L.* sp. 8.

Manilkara is resolved as polyphyletic due to the three Pacific species *M. dissecta*, *M. fasciculata* and *M. udoido*, which are placed on a strongly supported branch outside the main *Manilkara* clade (PP 1.00). This main *Manilkara* clade, excluding the former three species, is strongly supported (PP 1.00) and will be referred to hereafter as *Manilkara* s.str. clade, whereas the clade with the three remaining species will be referred to as the *Abebaia* clade, as it contains *Manilkara fasciculata*, the type species of the genus *Abebaia* Baehni. Within *Manilkara* s.str., a first split is observed between a clade of American *Manilkara* species (*M. bidentata* (A.DC.) A.Chev., *M. longifolia* (A.DC.) Dubard, *M. zapota* (L.) P.Royen; PP 1.00) and a clade (PP 1.00) grouping a species from Southeast Asia, *M. hexandra* (Roxb.) Kuntze) and all African species (Continental Africa: *M. cuneifolia* (Baker) Dubard, *M. lacera* (Baker) Dubard, *M. multinervis* (Baker) Dubard, *M. obovata* (Sabine & G.Don) J.H.Hemsl., *M. sansibarensis* (Engl.) Dubard; Madagascar: *M. boivinii* Aubrév.). The African lineage itself is monophyletic (PP 1.00).

Neither *Labourdonnaisia* nor *Faucherea* are resolved as monophyletic, since species assigned to the two genera render them reciprocally paraphyletic. However, the lineage containing the two genera forms a well-supported clade (PP 1.00)

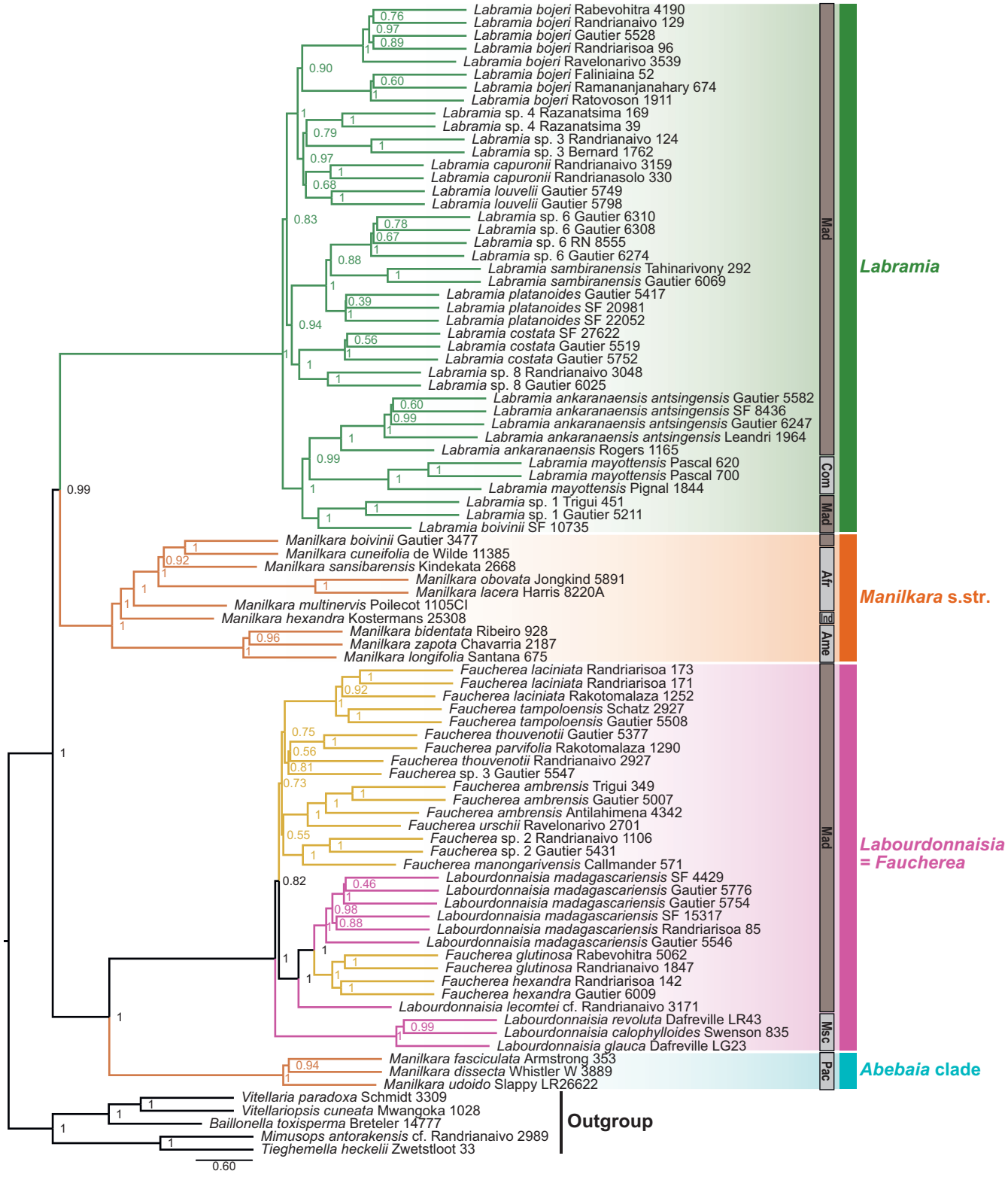


Fig. 1. Species tree reconstruction inferred from ASTRAL-II using 89 specimens and 787 individual gene trees obtained using RAxML. The node labels represent ASTRAL support values. Note that ASTRAL only calculates internal branch length and that tip lines are artificially fixed with the same length for all the specimens. Tip labels include the species names and the collector codes. Branch colors represent the traditional classification: *Labramia* (dark green), *Manilkara* (orange), *Faucherea* (yellow) and *Labourdonnaisia* (pink). The revised four major genetic clades are highlighted by a colored bar as follows: *Labramia* (dark green), *Manilkara* s.str. (orange), *Faucherea* and *Labourdonnaisia* (pink), and the *Abebaia* clade (blue). The main regions are indicated as follows: Afr: Africa; Ame: Americas; Com: Comoros; Ind: Indonesia; Mad: Madagascar; Msc: Mascarenes; Pac: Pacific Asia. RN: Réserves Naturelles; SF: Service Forestier.

sister to the *Abebaia* clade (PP 1.00). The Mascarene *Labourdonnaisia* species (*L. calophylloides* Bojer, *L. glauca* Bojer, *L. revoluta* Bojer) form a well-defined monophyletic group with strong support (PP 1.00). However, the relationship between the Mascarene lineage and the Malagasy species is not resolved due to moderate support in the crown Malagasy lineage (PP 0.82). Not all species are retrieved confidently and the obtained topology does not allow for the resolution of all the relationships among species. Six described species show moderate to strong support within the genus: *Faucherea ambrensis* Aubrév. (PP 1.00), *F. glutinosa* Aubrév. (PP 1.00), *F. hexandra* (Lecomte) Lecomte (PP 1.00), *F. tampoloensis* Aubrév. (PP 1. 00) and *Labourdonnaisia madagascariensis* (PP 1.00) and to a lesser extent *F. laciniata* Lecomte (PP 0.92). *Faucherea manongarivensis* Aubrév., *F. urschii* Capuron ex Aubrév., and *Labourdonnaisia lecomtei* Aubrév. are represented by single specimens. The remaining taxa were not recovered confidently as species: an unsupported clade (PP 0.81) contains two specimens of *Faucherea thouvenotii* Lecomte together with a specimen of *F. parvifolia* Lecomte and an undescribed morphospecies (*Faucherea* sp. 3). The single specimen of *F. parvifolia* analyzed is sister to a *F. thouvenotii* specimen (PP 1.00), making *F. thouvenotii* not monophyletic. Two undescribed morphospecies of *Faucherea* were included on this phylogenetic tree, the first one, *Faucherea* sp. 2 stands out as isolated and monophyletic (PP 1.00) while the second one, *Faucherea* sp. 3 is represented by a single specimen.

Exploration of the phylogenetic tree space. — The k-means analysis based on the Robinson-Foulds distance matrix computed on gene tree topologies shows that the number of clusters that best adjust to our dataset is $k = 2$ (Fig. 2A). The two clusters are unequal in size: cluster 1 constitutes 66% of the total number of loci (517 loci), while cluster 2 represents 34% (270 loci). A quite similar result is obtained with the hierarchical clustering approach (suppl. Fig. S1). Interestingly, the heatmap shows that gene trees from cluster 1 display close pairwise relationships while cluster 2 was mostly composed by gene trees which are equally dissimilar to those from cluster 1, and not similar to each other within cluster 2.

According to the Wilcoxon test, genes of cluster 1 display, on average, significantly more informative sites than that of cluster 2 (p-value = 2.15e-71; 137 and 51, respectively), longer gene sizes (p-value = 1.17e-70; 1334 and 533 bp, respectively), and less missing data (p-value = 1.09e-24; 2.90% and 5.53%, respectively; Fig. 2B).

The ASTRAL trees inferred from the two clusters show discordances in topology and branch support (Fig. 3). The ASTRAL tree of cluster 1 is in agreement with that of Fig. 1. Branch support values are generally higher in cluster 1 compared to cluster 2, while most nodes in cluster 2 have no significant support, except for the four major clades retrieved in the species tree (Fig. 1) that are monophyletic and well-supported in both clusters (PP 1.00). Notable discrepancies concerning branch support are seen for the relationship of *Manilkara* s.str. with the other genera. As in the species tree, *Manilkara* s.str. and *Labramia* are each monophyletic (PP 1.00) and both genera form a monophyletic clade in cluster 1 (PP 0.99), while there is uncertainty for their common monophyly in cluster 2 (PP 0.54). However, the tanglegram (Fig. 3) clearly demonstrates that each clade is monophyletic and that discordance in topologies is only recorded at the inter-species relationship level. A notable incongruence is found for the placement of the Mascarenes *Labourdonnaisia*. In cluster 1, the Mascarenes clade (PP 1.00) is found outside a clade containing all Malagasy species (PP 0.87), while it is retrieved within a Malagasy clade in cluster 2 with high support (PP 0.97).

Divergence time estimations. — The 20 genes used for BEAST analysis are all comprised in cluster 1 (Fig. 2A). In the combined BEAST log file built on two runs of 600 million generations each, convergence was reached, and all parameters showed high ESS values ($209 \leq ESS \leq 55{,}162$). The Manilkarinae subtribe evolved during the late Eocene at about 38 mya (highest posterior density [HPD] 42–33 mya; Table 3 and Fig. 4). *Manilkara* s.str. diverged from the genus *Labramia* during the early Oligocene (ca. 34 mya; HPD 38–30 mya), its crown age being about 29 mya (HPD 32–28 mya), while that of *Labramia* is much younger, around 11 mya (HPD 13–9 mya), during the late Miocene. The only non-Malagasy species, *Labramia mayottensis*, described from the Comoros archipelago, has a crown age of 5 mya (HPD 7–3 mya). The *Labourdonnaisia-Faucherea* clade diverged from the *Abebaia* clade 30 mya (HPD 36–24 mya) during the middle Oligocene, both clades evolving since the middle Miocene: the crown of *Labourdonnaisia-Faucherea* is slightly older (ca. 17 mya; HPD 21–14 mya) than that of the *Abebaia* clade (14 mya; HPD 20–9 mya). The Mascarene *Labourdonnaisia* species crown age is about 6 mya (HPD 9–3 mya).

Morphological analyses. — The factorial analysis for mixed data (FAMD) shows that the proportions of variance explained by the first four axes are 29.4%, 16.4%, 11.4% and 9.5%, respectively, which together explain 66.7% of the variance (Fig. 5C). All three qualitative variables are well-represented on the first FAMD axes (suppl. Fig. S2) with the staminodes apex dissection (Stam.Integr) represented on Dim-1 to Dim-4, while the dorsal appendages dissection (App.Integr) and the ovary pubescence (Ov.Pub) are represented only on Dim-1 and Dim-2, respectively. The quantitative variables that have a significant influence on defining the factorial axes are listed in declining order for each of the four first axes:

Dim-1: dorsal appendage length to corolla lobe length ratio (App_Cor), width to length ratio of staminode (Stam.l_L), staminode length to corolla lobe length ratio (Stam_Cor), number of corolla lobes, staminodes and stamens (Merosity), corolla lobe width to length ratio (Cor.l_L), and dorsal appendage width to length ratio (App.l_L).

Dim-2: corolla tube length to corolla lobe length ratio (Tube_Cor), style length to corolla lobe length ratio (Style_Cor), and number of ovary locules (Ov.locules).

Dim-3: total length of the corolla (Cor.length), corolla lobe width to length ratio (Cor.l_L), merosity, and dorsal appendage width to length ratio (App.l_L).

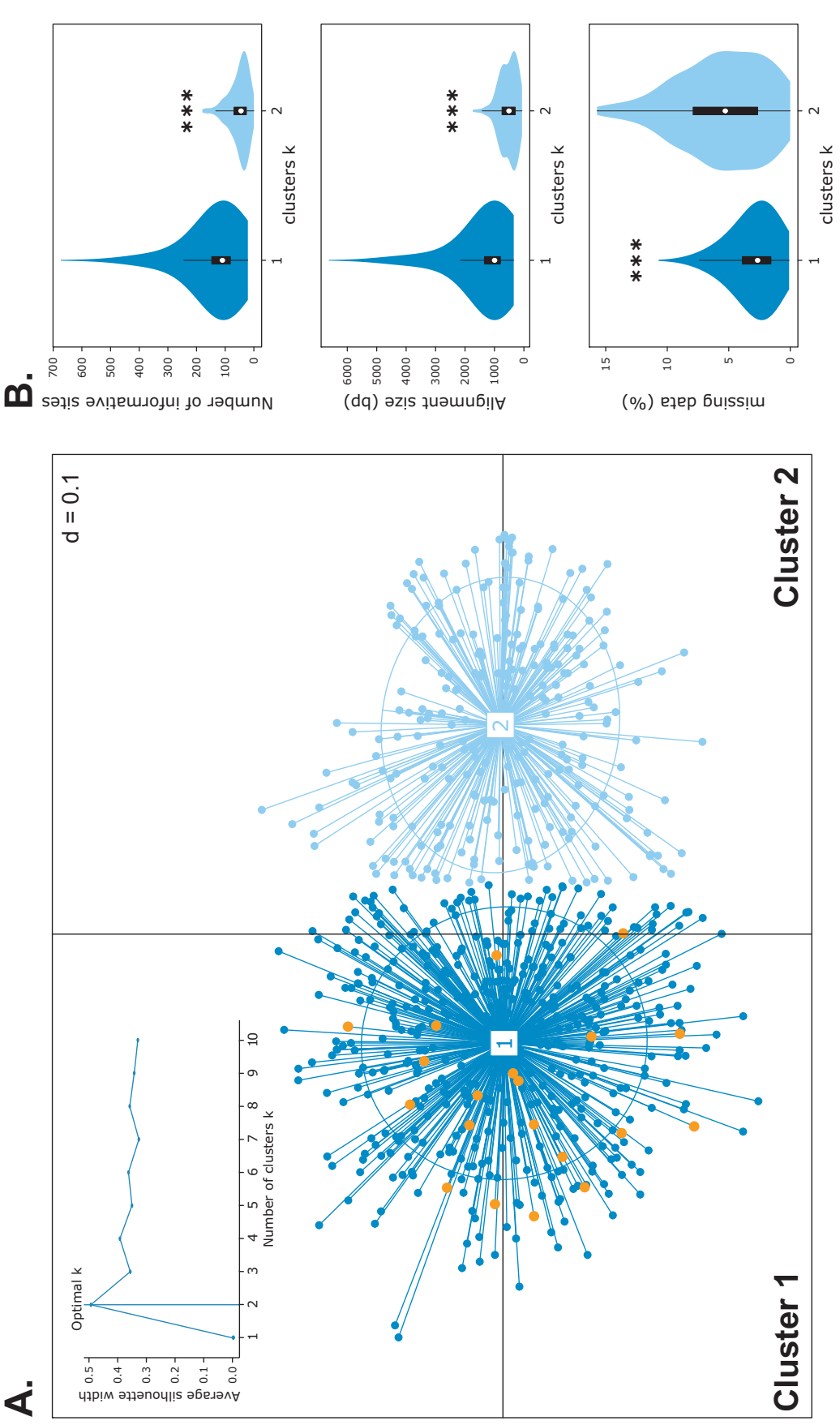


Fig. 2. A, Multidimensional scaling analysis of gene tree topologies, pairwise distances and k-means analysis (upper left corner) highlighting two clusters: cluster 1 (dark blue) and cluster 2 (light blue). The orange points within cluster 1 represent the 20 genes used for the phylogenetic dating analysis of Fig. 4; **B**, Violin plots of gene characteristics.

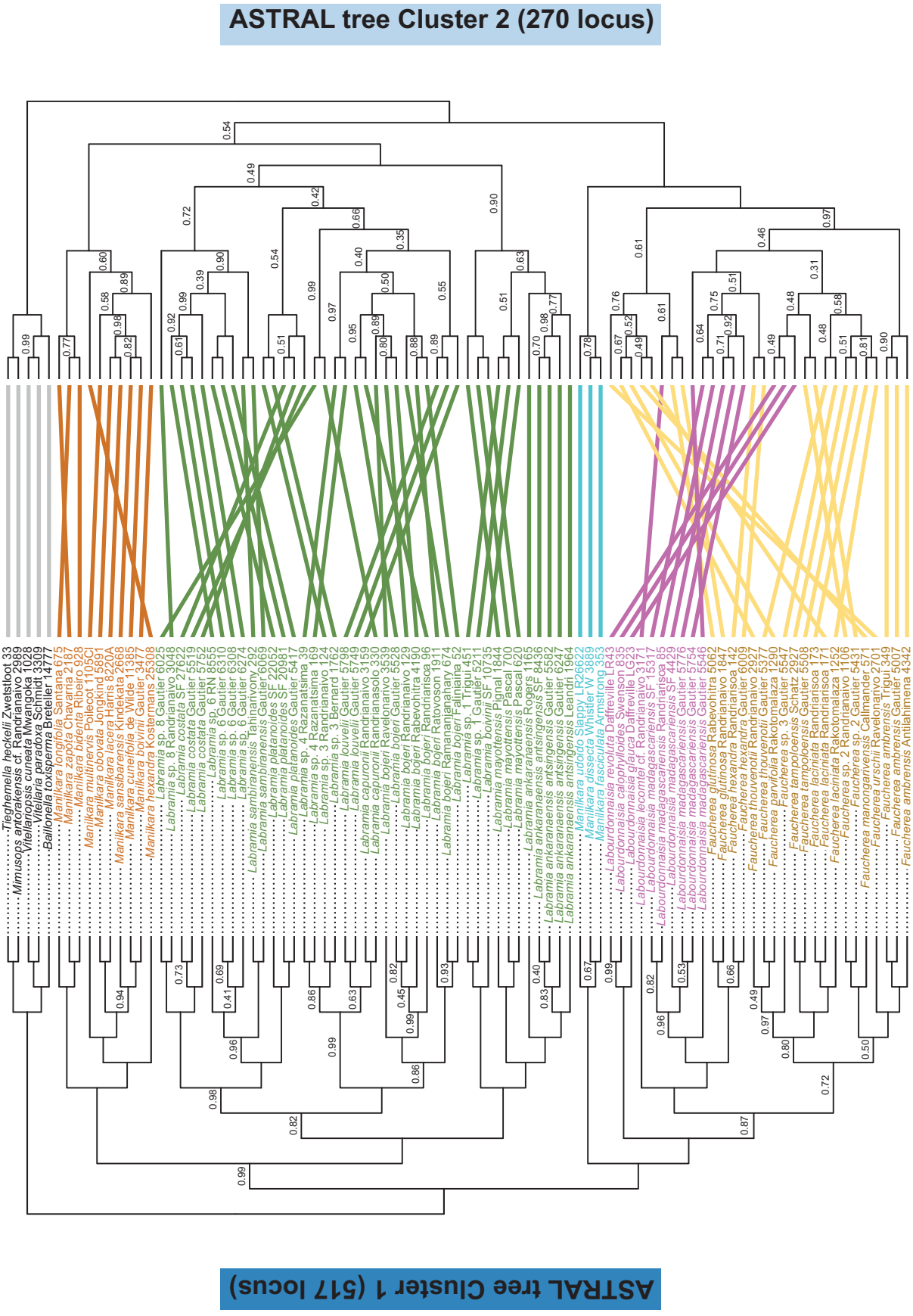


Fig. 3. Exploration of the phylogenetic tree space. Tanglegram comparison of the two ASTRAL trees inferred from cluster 1 (left) and cluster 2 (right). Tanglegram colors: *Labramia* (dark green), *Manilkara* s.str. (orange), the *Abebaia* clade (blue), *Faucherea* (yellow), *Labourdonnaisia* (pink) and the outgroup (grey). — Only ASTRAL supports lower than 1.00 are represented on the branches. RN: Réserves Naturelles; SF: Service Forestier.

Dim-4: number of ovary locules (Ov.locules) and number of corolla lobes, staminodes and stamens (Merosity).

Sample distribution on the first plane (Dim-1 × Dim-2) represents 45.8% of the total variance (Fig. 5A). *Labourdonnaisia* and *Faucherea* are grouped together on the right side of the FAMD, with an important overlap in their confidence interval. The Malagasy *Labourdonnaisia* samples (two *L. madagascariensis* – No. 15 and No. 16 and two *L.* sp. 1 – No. 13 and No. 14) are clustered within the *Faucherea* group, whereas the Mascarenes *Labourdonnaisia* samples are distinct (*L. calophylloides* – No. 11, *L. glauca* – No. 12, *L. revoluta* – No. 17). *Labramia* and *Manilkara* s.str. are distributed on the left side of the FAMD, and partially overlap due to the following five samples: *L. ankaranaensis* var. *antsingensis* – No. 32 and *L. bojeri* – No. 19; *M. hoshinoi* – No. 54; *M. sansibarensis* – No. 49 and *M. zapota* – No. 50. *Labramia* samples have an upper-left distribution trend, while *Manilkara* s.str. show a lower-left distribution, with some outliers for both genera. The specimens belonging to the *Abebaia* clade are distributed in between *Manilkara* s.str. and *Labourdonnaisia-Faucherea*. Six of them are clustered within the *Manilkara* s.str. confidence interval (three *M. dissecta* – No. 36, No. 41 and No. 43; and three *M. fasciculata* – No. 34, No. 38, and No. 39). Five further samples stand outside *Manilkara* s.str. and tend to be closer to *Labourdonnaisia-Faucherea* (two samples of *M. dissecta* – No. 35, No. 37; one *M. fasciculata* – No. 40; two *M. udoido* – No. 33 and No. 42). A total of 40.8% of the total variance is represented in Dim-1 × Dim-3 projection (Fig. 5B). Compared to the Dim-1 × Dim-2 representation, this new perspective shows that: (i) the isolation of Mascarene *Labourdonnaisia* samples from Malagasy *Labourdonnaisia* and *Faucherea* is even more pronounced; (ii) *Labramia* and *Manilkara* s.str. completely overlap; (iii) the specimens belonging to the *Abebaia* clade are still embedded in *Manilkara* s.str. The Dim-1 × Dim-4 projection does not provide further information.

Table 3. Divergence time estimations.

Node	Mean age BEAST (mya)	95% HPD BEAST (mya)
Crown of Manilkarinae subtribe	37.65	33.13–42.37
Labramia A.DC.–*Manilkara* Adans. s.str. divergence	33.86	30.34–37.97
Labourdonnaisia Bojer–*Abebaia* Baehni divergence	30.33	24.24–36.42
Crown of *Labramia*	10.72	8.6–12.83
Crown of *Manilkara* s.str.	29.38	28.08–31.57
Crown of *Labourdonnaisia*	17.29	13.85–20.75
Crown of *Abebaia* clade	13.96	8.54–19.91
Crown of *Labramia mayottensis* Labat & al.	4.63	2.82–6.6
Crown of Mascarene *Labourdonnaisia*	6.00	3.38–8.98

mya: million years ago; HPD: highest posterior density.

An a priori group assignment based on the phylogenetic results was used for the linear discriminant analysis (LDA). The four groups are the *Labourdonnaisia-Faucherea* clade, *Labramia*, *Manilkara* s.str. and the *Abebaia* clade. The three discriminant axes account for 84.5%, 9.9% and 5.6% of the total variation, respectively (Fig. 6). Four variables are found to be the most relevant to discriminate the groups and are listed in declining order as follows: corolla lobe width to length ratio (Cor.l_L), dorsal appendage length to corolla lobe length ratio (App_Cor), staminode width to length ratio (Stam.l_L) and number of ovary locules (Ov.locules). For this model, the correctness rate is 83.3%. On the two LD1 × LD2 and LD1 × LD3 projections, *Labourdonnaisia-Faucherea* is clearly differentiated from the rest of the groups due to its coordinates on the LD1 axis (Fig. 6A and 6B, respectively). The barycenters of *Labramia*, *Manilkara* s.str. and the *Abebaia* clade are close to each other. *Labramia* is differentiated from *Manilkara* s.str. and the *Abebaia* clade along the LD2 axis, on which the latter two display overlapping distributions. The third axis (LD3) shows that *Labramia*'s barycenter lays between the *Manilkara* s.str. and *Abebaia* clade barycenters. According to the leave-one-out validation analysis (Fig. 6C), the classification error rate is 18.3%. The *Labourdonnaisia-Faucherea* group is well-delimited, and all a priori assigned samples are correctly classified. One individual of the 15 *Labramia* samples is wrongly attributed to *Manilkara* s.str. as well as two that are retrieved in the *Abebaia* clade. Five individuals out of the 17 *Manilkara* s.str. are wrongly classified: three are assigned to *Labramia*, one to the *Labourdonnaisia-Faucherea* group and one to the *Abebaia* clade. Three members of the *Abebaia* clade out of 11 are wrongly classified as *Manilkara* s.str.

Ancestral state reconstruction for discrete characters. — The most probable state for ovary pubescence in the Manilkarinae node (53) is a pubescent ovary as indicated by the pie charts representing the marginal likelihood at ancestral nodes (Fig. 7A and suppl. Table S2). The Most Recent Common Ancestor (MRCA) of *Labramia* and *Manilkara* s.str. (node 72) is also recovered as having a pubescent ovary. This state is retained in *Manilkara* s.str., whereas an evolution toward a glabrous ovary is observed in the *Labramia* clade (node 80). The ovary of the MRCA of *Labourdonnaisia-Faucherea* and the *Abebaia* clade (node 54) is reconstructed as pubescent, as observed in all its extant members.

The ancestral state reconstruction for the number of ovary locules at the Manilkarinae node is not resolved with confidence. An ovary with eight locules is reconstructed as the most probable state; however, the four other states cannot be excluded (node 53; Fig. 7B). The ancestral node of *Labramia* is also reconstructed with eight locules (node 80), while that of *Manilkara* s.str. remains uncertain (node 73). The ancestor of the *Abebaia* clade displays six locules (node 55) as the one of *Labourdonnaisia-Faucherea* (node 57) with exceptions in the Mascarenes *Labourdonnaisia* (node 61).

For dorsal appendage dissection, the character state at the Manilkarinae ancestral node could not be estimated with

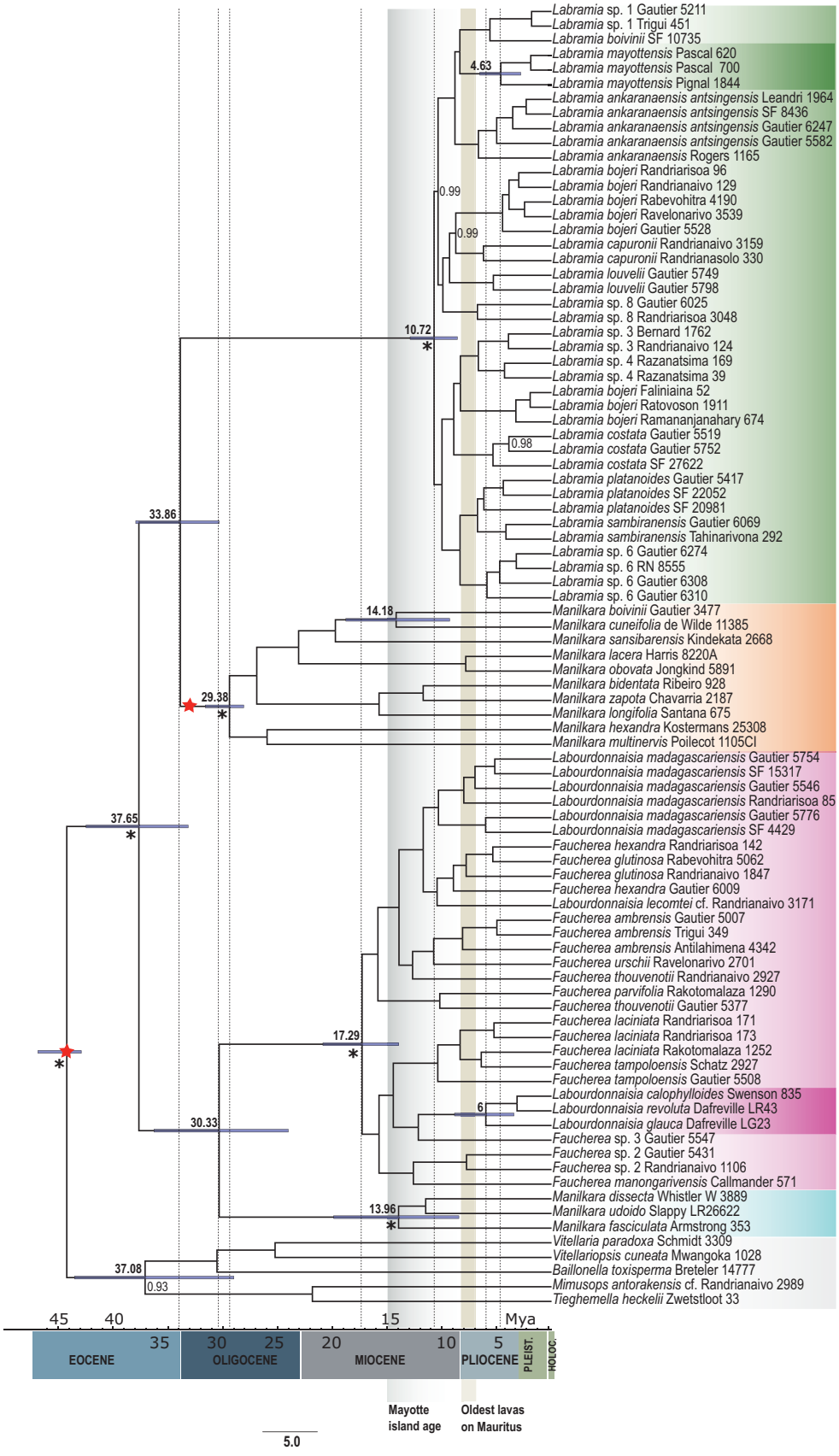


Fig. 4. Maximum clade credibility and divergence time estimations from BEAST reconstruction using 89 specimens and 20 genes. Background colors represent the four Manilkarinae clades. Node labels are given as the mean of node age estimates for the main clades, including their 95% HPD and posterior probability (PP). The latter are only shown when PP < 1. Epoch and ages in million years ago are represented at the bottom. Primary calibration points from fossils data are indicated with a red star, whereas clades constrained as monophyletic are labeled with an asterisk. RN: Réserves Naturelles; SF: Service Forestier.

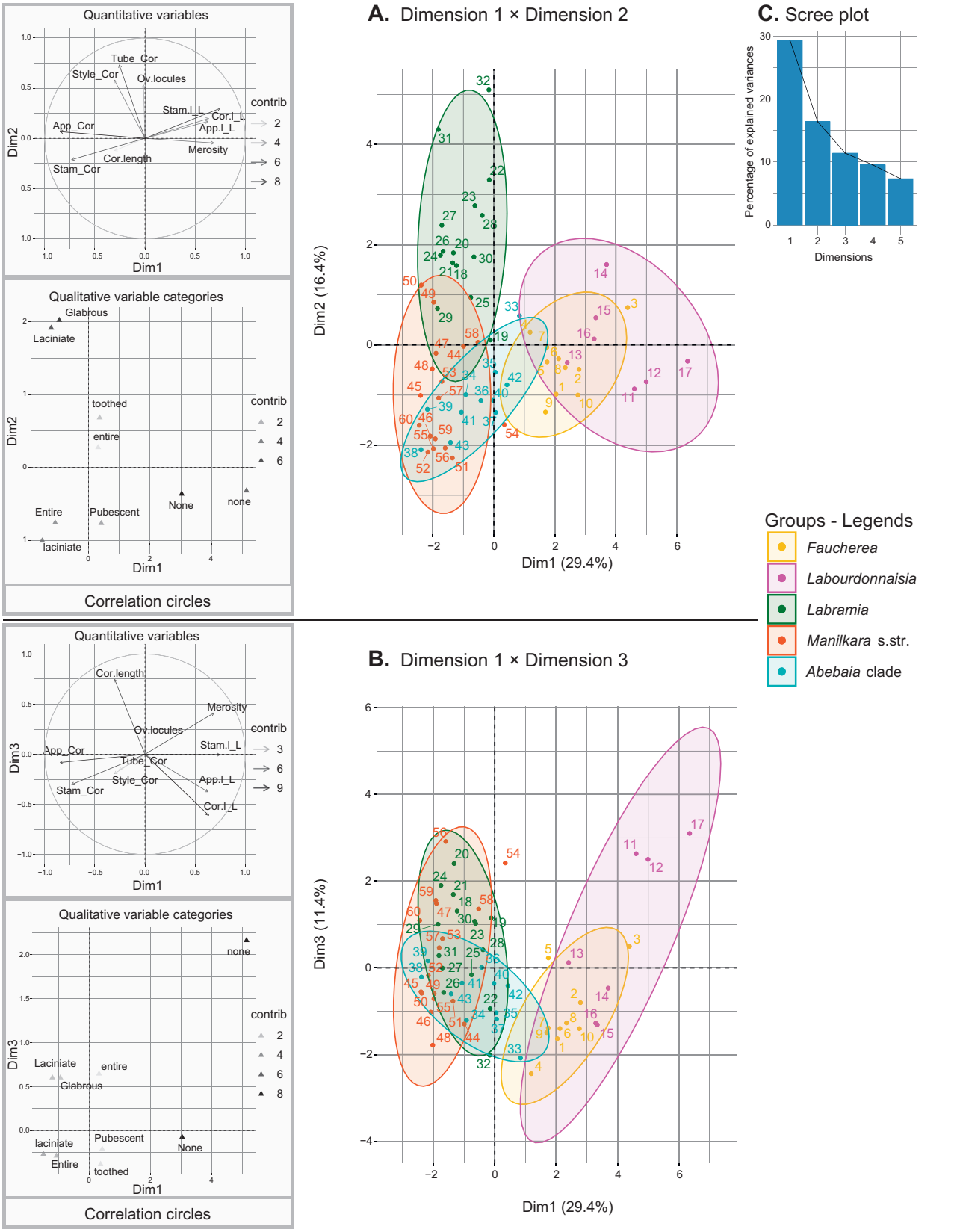
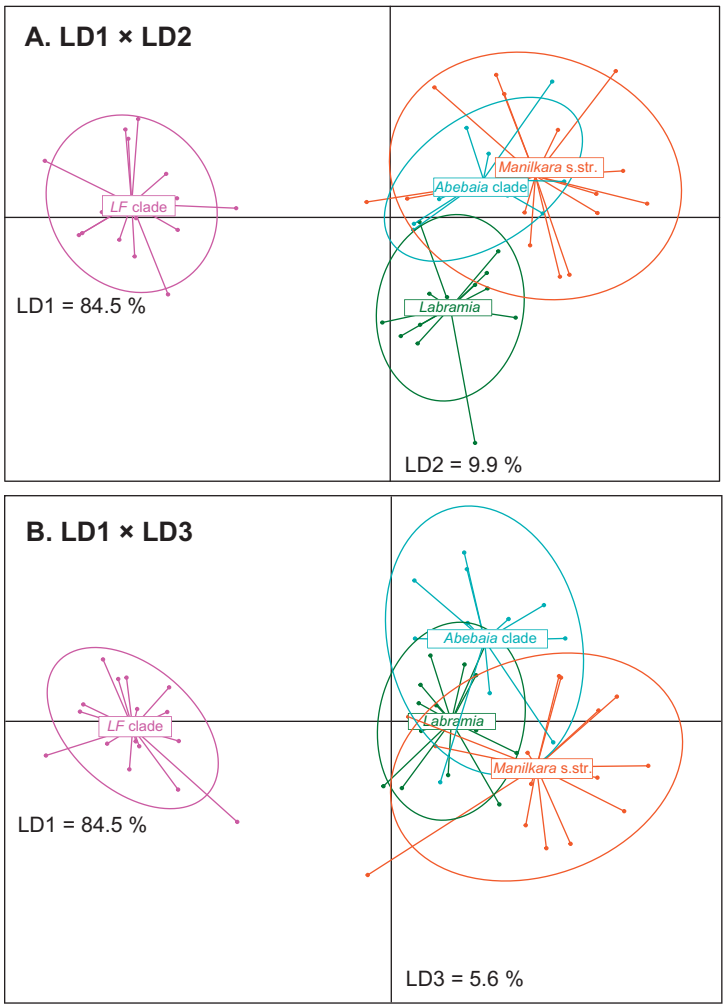


Fig. 5. Factorial analysis of mixed data (FAMD). The dataset includes 10 quantitative and 3 qualitative variables on five groups: *Faucherea, Labourdonnaisia*, *Labramia*, *Manilkara* s.str. and the *Abebaia* clade. **A**, FAMD on Dimension 1 × Dimension 2 with the correlation circles of quantitative and qualitative variables; **B**, FAMD on Dimension 1 × Dimension 3 with the correlation circles of quantitative and qualitative variables; **C**, Variance proportions explained by the axes. contrib: contribution of variables to the axes.

certainty; there is a conflict between a laciniate and an entire apex (node 53; Fig. 8A). The *Labramia* clade most probably had an ancestor with laciniate dorsal appendages (node 80), while that of *Manilkara* s.str. is equivocally resolved (node 73). The *Abebaia* clade is characterized by a MRCA with an entire apex (node 55). A loss of dorsal appendages has probably occurred in the ancestral *Labourdonnaisia-Faucherea* clade (node 57), and this state has been retained by the majority of its descendants.

The reconstruction of the ancestral state for the staminode apex dissection at the Manilkarinae node is not resolved with confidence, but either laciniate or toothed apex are the most probable states (node 53; Fig. 8B). The *Labramia* clade most probably had an ancestor with a toothed apex (node 80), while that of *Manilkara* s.str. is likely to have been laciniate (node 73). The ancestor of the *Labourdonnaisia-Faucherea* clade had either toothed or entire staminodes, with a higher probability for a toothed apex (node 57). It is worth mentioning, however, that the absence of staminodes in all extant Mascarene *Labourdonnaisia* makes it very likely that their ancestor had already lost them (node 61). The ancestor of the *Abebaia* clade is likely to have had a laciniate or toothed apex (node 55).

Regarding the merosity of corolla and androecium, the character state reconstruction shows that an hexamerous corolla is the most probable ancestral state for subtribe Manilkarinae (node 53; Fig. 9). All the ancestors of the main clades have retained this character state and it was further conserved throughout *Manilkara* s.str., *Labramia* and the *Abebaia* clade. In the *Labourdonnaisia-Faucherea* clade, two distinct patterns are observed: (i) the probability of having eight corolla lobes instead of six increases starting from node 66, and becomes very high for the ancestor of both *Labourdonnaisia madagascariensis* (node 71); (ii) the ancestor of all Mascarene *Labourdonnaisia* is retrieved with a high probability of having 12 lobes (node 61). Furthermore, a heptamerous corolla has been observed in our sample of *Faucherea manongarivensis*, but the probabilities of having such a state in its ancestors are very low.



Prior groups

- *Labourdonnaisia-Faucherea* clade
- *Labramia*
- *Manilkara* s.str.
- *Abebaia* clade

C. Confusion table with cross validation

Posterior / Prior	*LF* clade	*Labramia*	*Manilkara* s.str.	*Abebaia* clade
LF clade	17	0	0	0
Labramia	0	12	1	2
Manilkara s.str.	1	3	12	1
Abebaia clade	0	0	3	8

Fig. 6. Linear discriminant analysis (LDA) using quantitative variables only, with groups defined a priori according to the phylogenetic clades. **A**, LD1 × LD2 projection; **B**, LD1 × LD3 projection; **C**, Confusion table with cross-validation. — Colors correspond to prior groups. LD: linear discriminant axis.

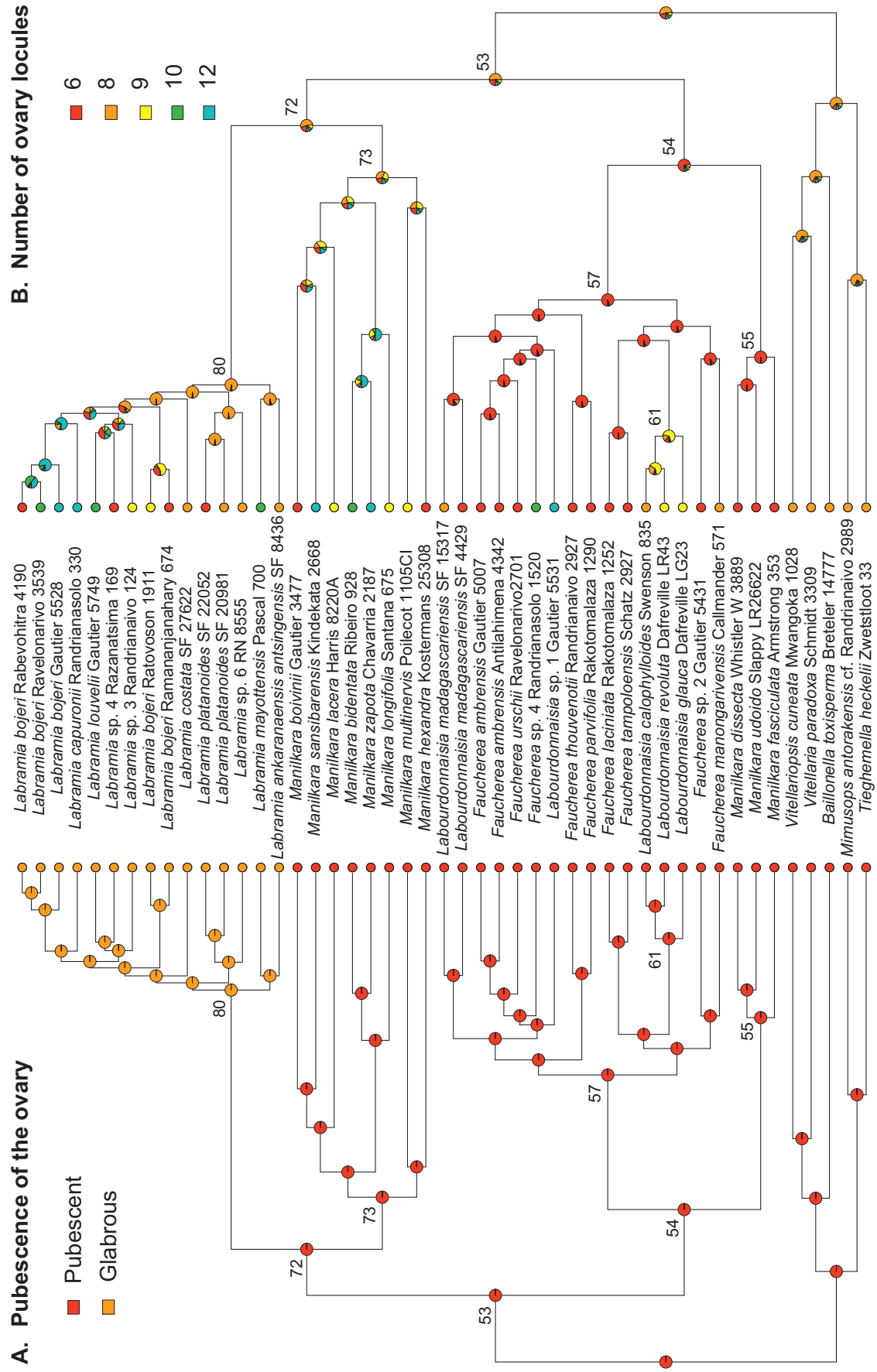


Fig. 7. Ancestral state reconstruction: **A**, Pubescence of the ovary; **B**, Number of ovary locules. Pie charts represent the marginal likelihood at ancestral nodes. The character matrix and support values for the inferred states are available in suppl. Table S2. Morphological traits were mapped on the dated tree topology pruned to 47 taxa. Node labels correspond to node number. — RN: Réserves Naturelles; SF: Service Forestier.

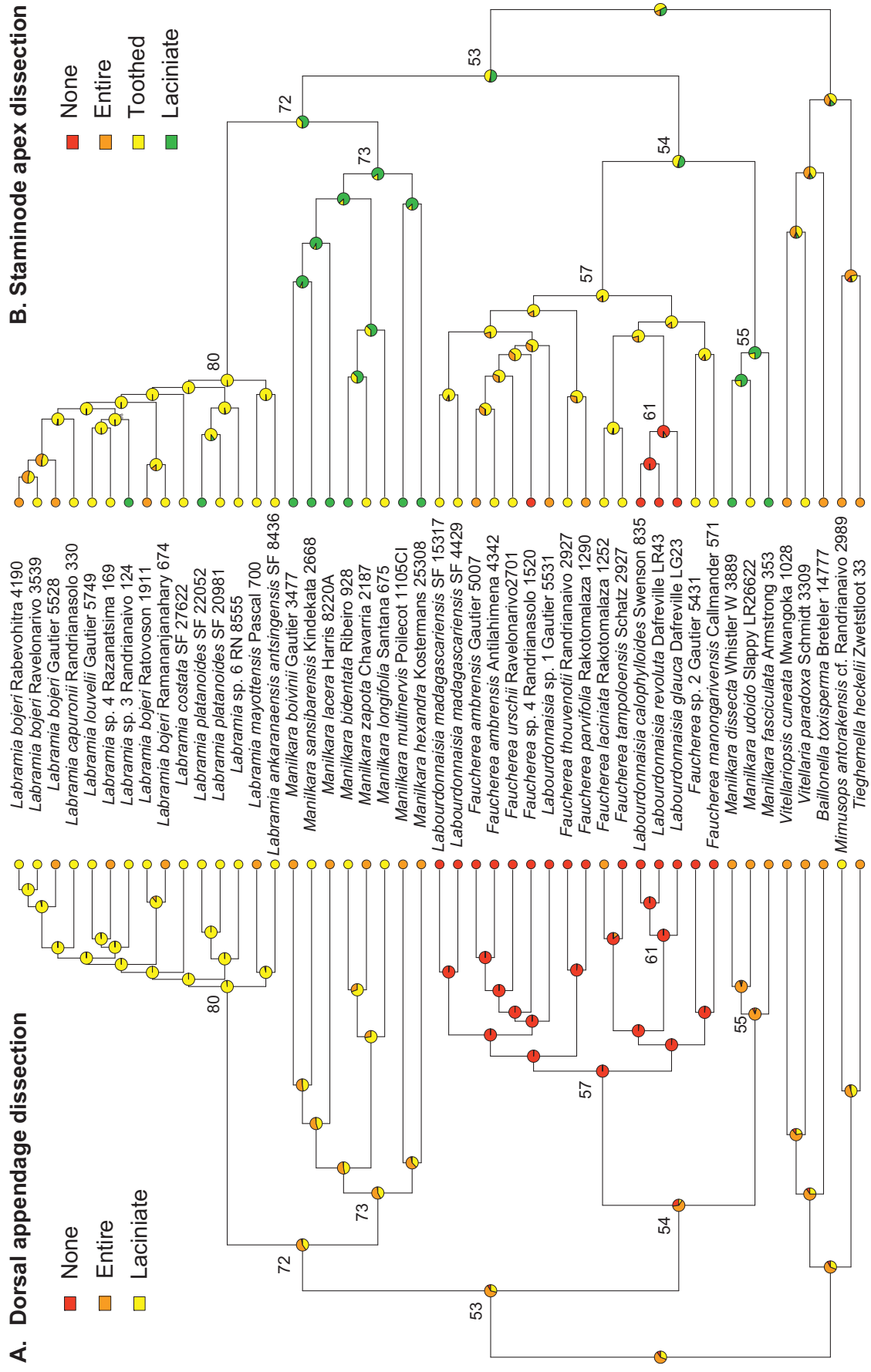


Fig. 8. Ancestral state reconstruction: **A**, Dorsal appendage dissection; **B**, Staminode apex dissection. Pie charts represent the marginal likelihood at ancestral nodes. The character matrix and support values for the inferred states are available in suppl. Table S2. Morphological traits were mapped on the dated tree topology pruned to 47 taxa. Node labels correspond to node number. — RN: Réserves Naturelles; SF: Service Forestier.

Ancestral state reconstruction for continuous characters. — The MRCA of subtribe Manilkarinae had most likely developed dorsal appendages slightly shorter than the corolla lobes (node 53; App_Cor = 0.77 ± 0.01; Fig. 10A and suppl. Table S3). This state is retained in *Labramia* (node 80) and in the *Abebaia* clade (node 55). The *Manilkara* s.str. ancestor displayed larger dorsal appendages, which are almost as long as the corolla lobes (node 73; App_Cor = 0.88 ± 0.009). The *Labourdonnaisia-Faucherea* clade is characterized by the loss of dorsal appendages with only a few taxa that have kept them in vestigial condition (node 57).

Well-developed staminodes, half the length of the corolla lobes, are retrieved at the Manilkarinae ancestral node and maintained in the MRCA of *Manilkara* s.str. (node 53; Stam_Cor ≥ 0.48 ± 0.005; Fig. 10B and suppl. Table S3). Two independent reductions in length are observed within Manilkarinae: one at the ancestral node of *Labramia* (node 80; Stam_Cor = 0.27 ± 0.001), the other at the MRCA of *Labourdonnaisia-Faucherea* and the *Abebaia* clade (node 54; Stam_Cor = 0.29 ± 0.007). The most important reduction in size occurs within the *Labourdonnaisia-Faucherea* clade (MRCA at node 57; Stam_Cor = 0.19 ± 0.003), with a total loss of staminodes in the Mascarene *Labourdonnaisia* species node.

The ratio of the corolla tube to corolla lobe length in the MRCA of subtribe Manilkarinae is 0.38 ± 0.02 (node 53, Fig. 11A and suppl. Table S3). The *Manilkara* s.str. ancestral node has the same value (node 73; 0.38 ± 0.02) but its descendants are very variable. The MRCA of *Faucherea/Labourdonnaisia* and the *Abebaia* clade is characterized by a slight reduction of the corolla tube (node 54; 0.31 ± 0.02),

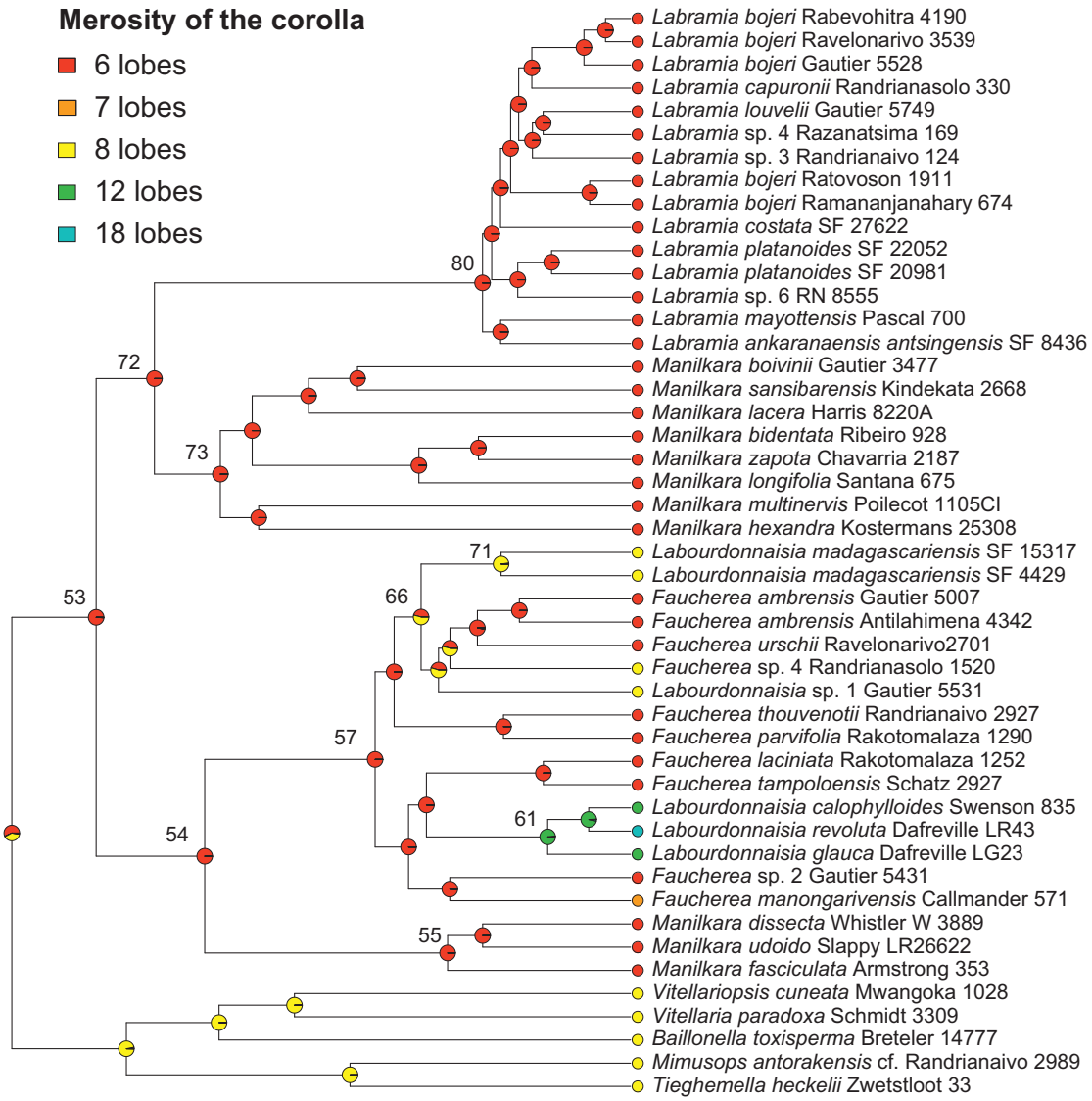


Fig. 9. Ancestral state reconstruction of the merosity of the corolla. Pie charts represent the marginal likelihood at ancestral nodes. The character matrix and support values for the inferred states are available in suppl. Table S2. Morphological traits were mapped on the dated tree topology pruned to 47 taxa. Node labels correspond to node number. — RN: Réserves Naturelles; SF: Service Forestier.

with occasional very low values within the extant species of the clade. On the contrary, a significant increase in relative tube size appears at the *Labramia* ancestral node (node 80; Tube_Cor = 0.58 ± 0.05) and the trend persists within the *Labramia* clade with node ratios from 0.41 to 0.66. In a few extant *Labramia* species, the corolla tube is even as long as the corolla lobes.

The character state reconstruction for the corolla total length (as a measure of flower size) at the Manilkarinae ancestral node was not resolved with certainty, showing very high variance and large confidence interval at 95% (node 53; Cor. length = 6.49 ± 3.52 mm; Fig. 11B and suppl. Table S3). Internal variation in corolla length within each major lineage is high. Some minor trends are noticed according to the character states of the terminals: The *Abebaia* clade was inferred to have the smallest corollas, whereas the *Labramia* clade contains taxa with the largest. However, none of the ancestors of the main lineages have been estimated with certainty according to the large variances of their ancestral node reconstruction. The ancestral nodes of both the *Abebaia* clade and that of *Labourdonnaisia-Faucherea* show a slight decrease in size (respectively node 55, Cor.length = 5.09 ± 2.55 and node 57, Cor.length = 4.98 ± 1.55). The *Manilkara* s.str. ancestor was retrieved with a slight increase (node 73; Cor.length = 6.09 ± 2.6 mm). The *Labramia* ancestor is characterized by the highest Cor.length value (node 80; Cor.length = 7.34 ± 1 mm).

■ DISCUSSION

A robust resolution of phylogenetic relationships within the Manilkarinae subtribe. — Previous phylogenetic reconstructions of the subtribe Manilkarinae using nuclear (ITS) and chloroplast (*rpl32-trnL*, *rps16-trnK*, *trnS-trnFM*) markers separately demonstrated its monophyly (Armstrong, 2010; Armstrong & al., 2014). The study was, however, primarily focused on *Manilkara*, and showed that this genus was monophyletic only if three Pacific species (*M. dissecta*, *M. fasciculata*, *M. udoido*) are excluded. It, however, left some unresolved questions about the relationships among genera in the subtribe. Using 787 nuclear genes with an extended taxon sampling, including the vast majority of the described species in *Faucherea*, *Labourdonnaisia* and *Labramia*, plus several undescribed morphospecies, we were able to resolve the backbone of Manilkarinae with high support. Our phylogeny demonstrates the existence of four major lineages within the Manilkarinae: *Labramia*, *Manilkara* s.str., *Labourdonnaisia-Faucherea*, and the *Abebaia* clade. For the first time, their relationships are well-resolved: the main clade of *Manilkara* is now retrieved as sister to *Labramia* and the strongly supported *Labourdonnaisia-Faucherea* clade is clearly assessed as sister to the *Abebaia* clade. The monophyly of these four major clades is strongly supported in all analyses, as shown in the ASTRAL species trees inferred from all genes (Fig. 1) and from gene clusters 1 and 2 (Fig. 3).

We are able to confirm that the species assigned to *Labourdonnaisia* and *Faucherea* are intermingled in a single clade and that the genus *Labramia* is monophyletic. Within both the *Labramia* and the *Labourdonnaisia-Faucherea* clades, the relationships among species are not supported for most taxa. Both clades have a radiation-like topology (Glor, 2010) which is known to result in unresolved polytomies (Naciri & Linder, 2020). Recent studies suggest that this type of rapid radiation might be common in Malagasy Sapotaceae (Boluda & al., 2021, 2022; Christe & al., 2021) as well as in other Malagasy tree genera (e.g., *Canarium* – Federman & al., 2018; *Dalbergia* – Crameri, 2020). Weak support for species relationships was expected, considering that the current sampling was designed to resolve generic issues and not interspecific relationships within each genus, which will require further sampling and more analyses.

Incongruence in gene tree histories within genera. — Conflicting genealogies are one of the main challenges faced by phylogeneticists, especially for those working at the species level, with high throughput genetic data and a large number of genes (Jeffroy & al., 2006). Obtaining different topologies due to incomplete lineage sorting, selection, ancient hybridization, paralogous genes (Naciri & Linder, 2015) or artefactual issues due to data quality or bioinformatics biases, is expected. Gene tree clustering analyses constitutes an interesting approach to exploring topological space, enabling the teasing apart of genes with non-concordant histories. In our case, two main topologies were found: one supporting the species tree (Fig. 1) and comprising the majority of the genes (66% – cluster 1 of Fig. 2A) and the second supporting an alternative tree (34% – cluster 2 of Fig. 2A).

Both clusters agree that the four Manilkarinae major clades are monophyletic, and the backbones of the two phylogenies demonstrate a similar pattern ((*Manilkara* s.str. + *Labramia*) + (*Labourdonnaisia* + *Abebaia* clade)), although not supported for the alternative cluster with 34% of the total number of loci (cluster 2, PP 0.54). Discordance in topologies is mainly found at inter-species relationship level, although almost never supported with genes of cluster 2.

We demonstrate in Fig. 2B that significant differences on gene evolutionary properties exist between cluster 1 and cluster 2. Indeed, cluster 2 displays genes with less informative sites on average, as well as shorter alignments and more missing data, which can explain the lack of support in the species tree reconstruction of cluster 2, but also partly, the discordance in species tree topologies.

However, as demonstrated by the heatmap analyses (suppl. Fig. S1), whereas gene trees of cluster 1 mainly share the same topologies, gene trees from cluster 2 show a more variable range of topologies. This implies that within cluster 2, there are potentially sub-clusters of genes sharing different histories, which could be explained not only by gene properties but also by other biological reasons such as hybridization, incomplete lineage sorting or/and selection. However, disentangling the different hypotheses would require further investigations similar to those conducted by Pollard & al. (2006) on *Drosophila*.

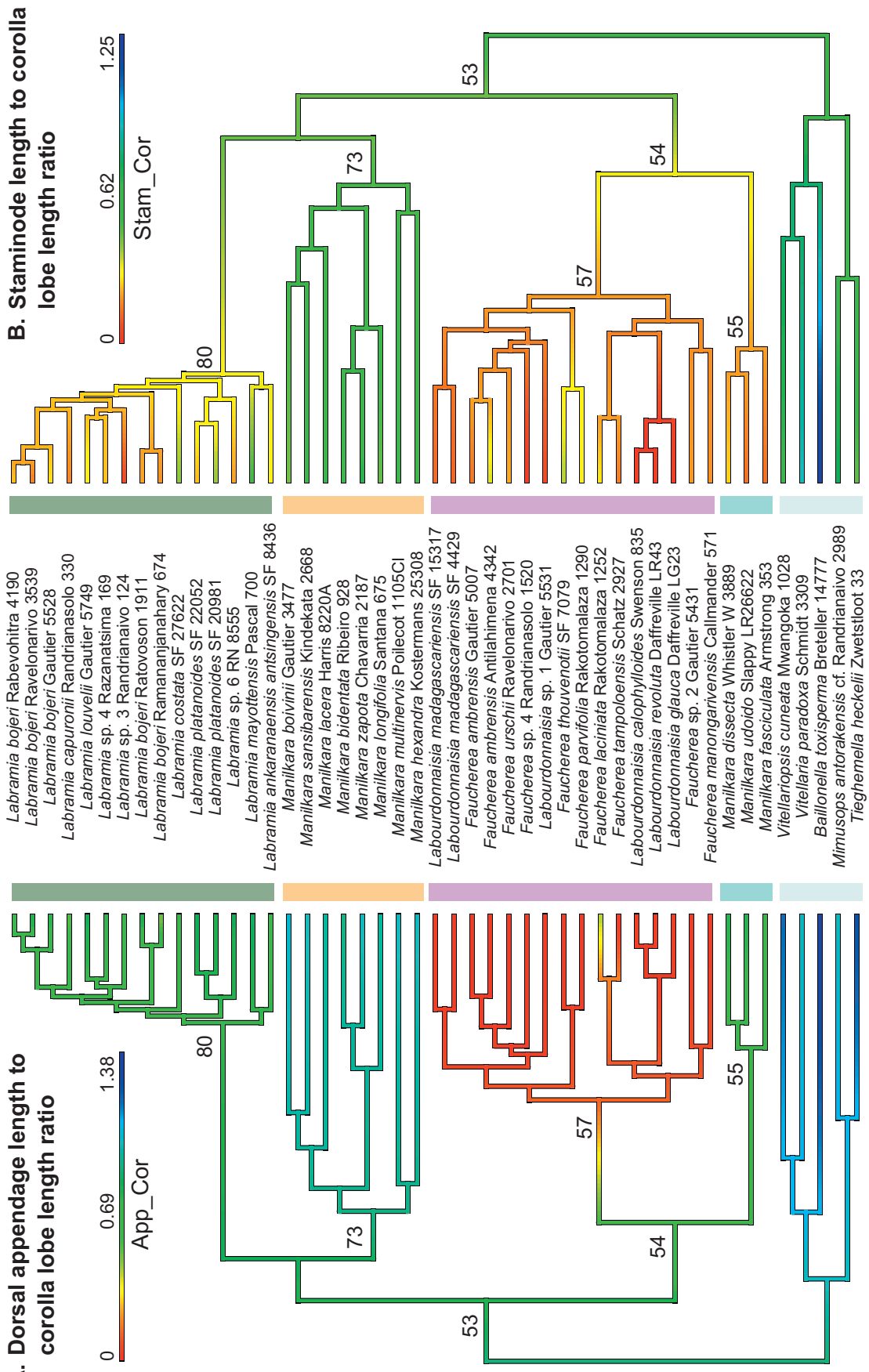


Fig. 10. Ancestral state reconstruction: **A**, Dorsal appendage length to corolla lobe length ratio; **B**, Staminode length to corolla lobe length ratio. Color gradients along the branches represent interpolation of state changes. Probabilities of ancestral node states, their computed variances and 95% confidence intervals are presented in suppl. Table S3. Morphological traits were mapped on the dated tree topology pruned to 47 taxa. — RN: Réserves Naturelles; SF: Service Forestier.

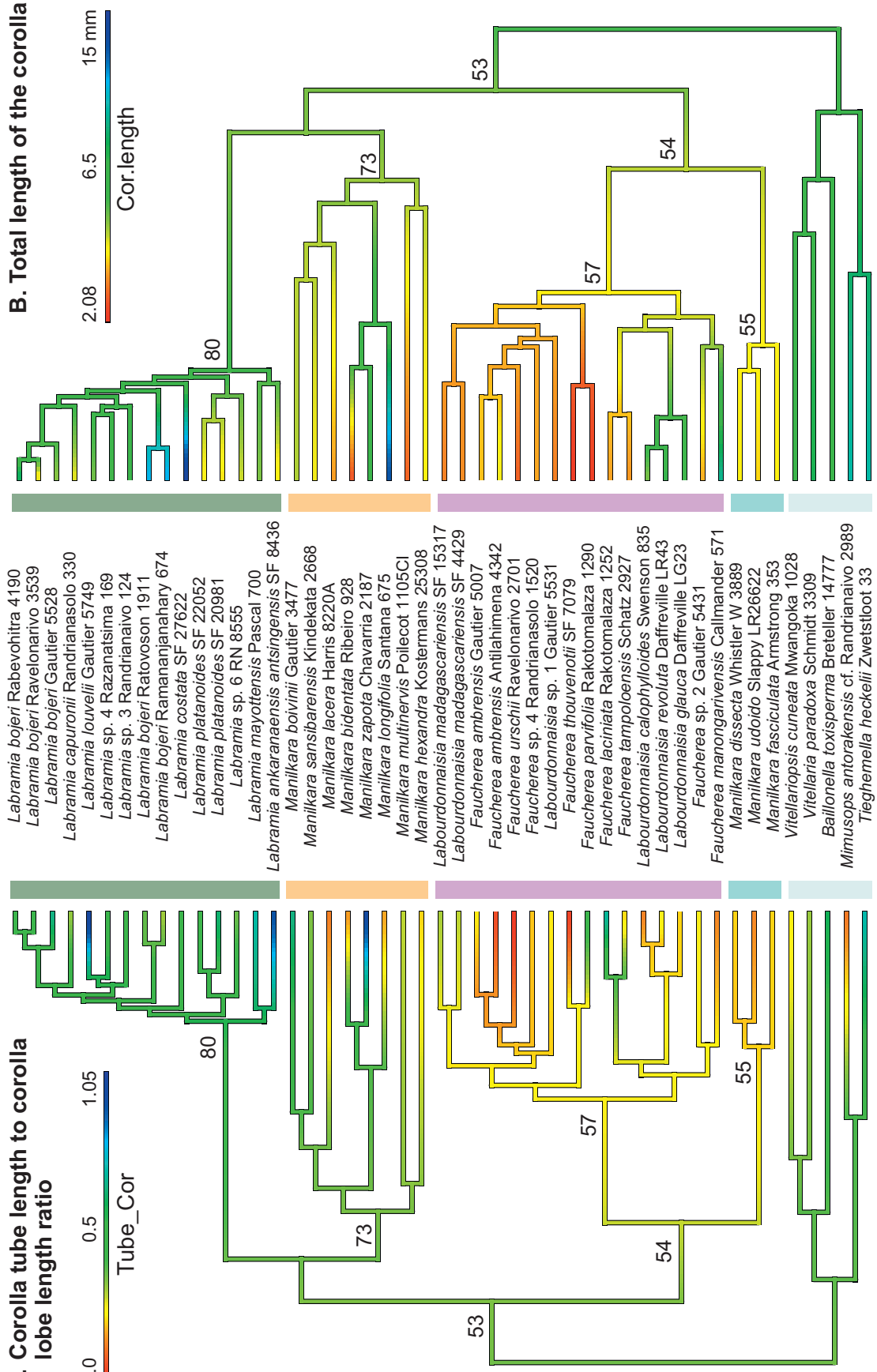


Fig. 11. Ancestral state reconstruction: **A**, Corolla tube length to corolla lobe length ratio; **B**, Total length of the corolla. Color gradients along the branches represent interpolation of state changes. Probabilities of ancestral node states, their computed variances and 95% confidence intervals are presented in suppl. Table S3. Morphological traits were mapped on the dated tree topology pruned to 47 taxa. — RN: Réserves Naturelles; SF: Service Forestier.

They are, however, outside the scope of this study as our main goal is to resolve the generic circumscription, which was confirmed by the vast majority of the genes used.

On the other hand, our results demonstrate the robustness of the ASTRAL methodology for phylogenetic reconstruction when used on NGS data and a high number of loci (787 genes), compared to a concatenation method, which would potentially dilute the phylogenetic signal (e.g., Gatesy & Baker, 2005) when sequences are too short with a low number of informative sites for each locus.

Divergence time estimations. — With the same calibration points used by Armstrong & al. (2014), it is not surprising that despite an extended sampling divergence time estimates are mostly congruent, with overlapping confidence intervals for the crown age of both the subtribe Manilkarinae and the genus *Manilkara* s.str. Thus, assumptions on geological and/or environmental events that could match with these two estimations follow those of Armstrong & al. (2014). We show here that *Manilkara* s.str. probably constitutes the oldest genus in the subtribe. Our data support the biogeographic assumption of a relatively recent (Oligocene/Miocene) onset of the majority of the extant Malagasy endemic or subendemic genera (Buerki & al., 2013), most probably through long-distance dispersal from Africa (Yoder & Nowak, 2006; Buerki & al., 2013), possibly facilitated by predominant eastward marine currents during that period (Federman & al., 2015). Indeed, all dates we obtained postdate the separation of Madagascar from Africa that occurred around 86 mya, excluding a pattern of vicariance for our groups, as has been outlined elsewhere (Yoder & Nowak, 2006; Agnarsson & Kuntner, 2012; Buerki & al., 2013).

In order to preserve as much as possible the nomenclatural stability, and given that our four main lineages diverged from each other as soon as in the early Oligocene, we propose to consider them at the generic rank. The alternative would be to consider an extended *Manilkara* genus, encompassing the four clades, but this would conflict with very clear morphological synapomorphies for both *Labramia* and the *Labourdonnaisia-Faucherea* clade. Such synapomorphies are however missing for the *Abebaia* clade as discussed below.

***Labramia*: a monophyletic and morphologically well-delimited genus. —** The monophyly of *Labramia* is confirmed here by the molecular phylogeny and further supported by the morphological analyses. The flower characters that we retrieved as the best ones to distinguish *Labramia* from the other genera are in concordance with those used traditionally (Aubréville, 1964, 1974; Pennington, 1991): the length of the corolla tube and the glabrous ovary. Across the entire subtribe Manilkarinae, *Labramia* is the only genus in which a long corolla tube, with respect to the total corolla length, is systematically observed. This character state is sometimes also present in *Manilkara* s.str. (*M. zapota*) and is reconstructed as the ancestral state of the MRCA of the two genera. It has apparently been conserved in *Labramia*. Such evidence demonstrates the stability of this character within the clade. The glabrous ovary constitutes a specific synapomorphic feature of the genus that was derived from its ancestor more than 11 mya (HPD 13–8 mya). Having such a well-developed corolla tube and a glabrous ovary might be related to ecological factors, such as pollinators, but no data are presently available to confirm this hypothesis with confidence.

An additional important character traditionally used in the delimitation of *Labramia* is seed scar shape and position, being generally ventral and narrow, compared to a basiventral oblong scar in the rest of the subtribe. Although not statistically analyzed here because our approach was mainly based on flowering material, it should be noted that this character state has been observed on all fruiting herbarium specimens examined for a coming revision of the genus.

The ancestors of *Labramia* diverged from the pantropical *Manilkara* s.str. clade in the early Oligocene, around 34 mya (HPD 38–30 mya). However, the extant species (all endemic to the Western Indian Ocean islands) radiated during the late Miocene around 11 mya (HPD 13–9 mya). This long ~20-million-year interval between stem and crown age, during which long-distance dispersal from Africa to Madagascar probably occurred, can be interpreted in different ways, implying either extinctions in Madagascar or in Africa, or delayed radiation of the genus on Madagascar. In the absence of *Labramia* fossil records, further interpretation can only remain speculative.

A relatively recent colonization event from Madagascar to the Comoros archipelago (5 mya; HPD 7–3) seems to be at the origin of the only known non-endemic Malagasy species *Labramia mayottensis*. This estimation is compatible with the rise of the island of Mayotte estimated to have occurred around 15–7 mya (Hajash & Armstrong, 1972; Emerick & Duncan, 1982; Nougier & al., 1986; Debeuf, 2004; Thébaud & al., 2009).

***Labourdonnaisia-Faucherea*: a single genus. —** Both molecular and morphological evidence retrieved the species belonging to *Labourdonnaisia* and *Faucherea* as being intermingled in a highly supported monophyletic clade, suggesting that the traditional circumscription should be revised. The Mascarene *Labourdonnaisia* species are also found to be monophyletic, albeit their position regarding the Malagasy species inferred from the ASTRAL phylogeny remains uncertain. Whether the two groups of species are sister clades still needs to be assessed. If confirmed, it would imply that the Mascarene *Labourdonnaisia* and the Malagasy taxa could then be considered as distinct taxonomic entities. However, the former topology is inconsistent with the dated Bayesian tree reconstruction, in which the Mascarene clade falls within the Malagasy species as a strongly supported clade (PP 1.00). Such incongruence between the two analyses deserves further discussion. Both methods are coalescent-based models although the ASTRAL model is pseudo-coalescent. However, the ASTRAL dataset was large, with 787 genes, while under current computing limitations, BEAST reconstruction was only based on 20 genes selected to represent the average number of parsimony-informative sites. We might assume that the genes that are informative for the separation of Mascarene *Labourdonnaisia* from the Malagasy lineage are absent from the 20 selected for the dating analysis. Nevertheless, the fact that

this relationship was obtained with no support in the ASTRAL analysis suggests that the mutual monophyly of these two lineages is questionable. Indeed, the exclusion of the Mascarene species is only weakly supported in cluster 1 reconstruction (PP 0.87; Fig. 3), and the alternative topology inferred on cluster 2 genes (Fig. 2C) nests the Mascarene clade within some Malagasy species (PP 0.97). Beyond the lack of support in the ASTRAL species tree, the two scenarios resolved by the two gene clusters demonstrate the lack of robustness of this relationship. Therefore, the decision to consider the Mascarene *Labourdonnaisia* as sister to the Malagasy clade cannot be made solely on the basis of the current genetic results.

Corolla merism is the only floral morphological character used in traditional classification to distinguish the genus *Faucherea* (6–11 lobes) from *Labourdonnaisia* (10–18 lobes), with overlapping ranges. In the light of our phylogeny, the threshold value to separate the two genera appears even more artificial. In fact, the increase in the merism of internal flower whorls could be hypothesized to have evolved through two independent changes within the *Labourdonnaisia-Faucherea* clade: the first one, observed in the Mascarene clade, can be interpreted as a duplication/triplication leading to 12 and 18 corolla lobes. The second type of change, observed in the Malagasy species, is more likely interpreted as either a gradual increase from 6 to 8 (and up to 12) corolla lobes or, alternatively, as a duplication of corolla lobes followed by a loss. Kümpers & al. (2016) have discussed the importance of merism for Sapotaceae classification. Regarding calyx morphology, the change from five sepals in a single whorl to six or eight sepals in two whorls is demonstrated as a key innovation with high taxonomic significance. A change in calyx merism implies a total reorganisation of its structure due to space limitation in the meristem development. On the contrary, changes in corolla and androecium merism are frequently observed between (and sometimes even within) individuals of the same species (e.g., *Capurodendron madagascariensis* and *C. oblongifolium* – Boluda & al., 2022) and are interpreted by Kümpers & al. (2016) as having a much lower taxonomic significance. Indeed, it should be seen as a homoplasic character (Anderberg & Swenson, 2003; Swenson & Anderberg, 2005). This is why the merosity of the corolla and androecium should not be given too much significance at the generic level within the Manilkarinae.

The only other character suggested to separate the two genera is the seed scar that was presented as hollow in *Labourdonnaisia* (Aubréville, 1974). However, it should be emphasized that this character state had only been observed on Mascarene specimens, the only three Malagasy *Labourdonnaisia* specimens cited at that time in the Flora being fruitless. Numerous collections of *L. madagascariensis* have accumulated since, showing that seeds do not always conform to this observation. Furthermore, it appears from our field observations, that a hollow seed scar might be a consequence of seed immaturity before desiccation. This character therefore appears irrelevant to separate the Malagasy *Labourdonnaisia* from *Faucherea*, although further studies might show that it might still be relevant to distinguish the Mascarene *Labourdonnaisia*.

We therefore conclude that the genera *Faucherea* and *Labourdonnaisia*, as traditionally circumscribed, should be synonymized into a single genus, with the name *Labourdonnaisia* having priority. According to the morphological analysis and the ASR, the dorsal appendages loss and staminodes reduction are the most suitable character states for delimiting the enlarged circumscription of *Labourdonnaisia*. Moreover, the fact that these character states are reconstructed as the ancestral states of the clade (17 mya, HPD 21–14 mya), clearly argues (i) for the relative stability of those characters states and (ii) their synapomorphy, consequently reflecting the evolutionary history of the clade. Such evidence helps avoiding discrepancies in the classification, by arbitrarily giving more/less importance to a character in the delimitation of taxonomic groups. The Mascarene *Labourdonnaisia* have likely evolved recently from a Malagasy origin, presumably from a single colonization event dated around 6 mya (HPD 9–3 mya). This crown age estimate is congruent with the biogeographical history of the Mascarene archipelago since the oldest lavas found on Mauritius are reported to be 7.8 million years old (Thébaud & al., 2009). Furthermore, these species have developed a multiplication of corolla merism. This character state seems however irrelevant to deserve a generic separation as already discussed above. To consider them as distinct at the subgeneric level remains debatable depending on further molecular analyses.

We provide below an amended description of the genus *Labourdonnaisia* that accounts for the inclusion of *Faucherea*.

Segregation of three Pacific *Manilkara* species at generic level. — No matter which genes were used for the phylogenetic reconstruction (ASTRAL species tree – Fig. 1; cluster 1 and cluster 2 – Fig. 3; and BEAST phylogeny – Fig. 4), molecular findings clearly show that *Manilkara* is polyphyletic: *Manilkara* s.str. and an alternate lineage (the *Abebaia* clade) more closely related to the *Labourdonnaisia-Faucherea* clade, as already suggested by Armstrong & al. (2014). This relationship is robust despite the fact that the morphological distinction between the two *Manilkara* lineages is not obvious using either flower morphology (Figs. 5, 6), or complementary morphological observations on other characters (vegetative parts and, as far as the sampling allowed for it, fruits). Our multivariate morphological analysis demonstrated important similarities between *Manilkara* s.str. and the *Abebaia* clade, explaining why these three species have been described in *Manilkara*.

Manilkara s.str. is a widely distributed genus, with a greater diversity of species compared to the other, smaller Manilkarinae genera. It also displays a broader range of morphological character variation. This can partially explain the overlapping space of the *Manilkara* s.str. and *Abebaia* clades in the morphological analysis. Indeed, for most of the traits that have been studied here, the full range of variation in character states is present in *Manilkara* s.str. For instance, considering the character "corolla tube length to corolla lobe length ratio", some species within *Manilkara* s.str. cover the range displayed by the genus *Labramia* and that of the *Labourdonnaisia-Faucherea* clade, which respectively show the most

extreme values. A similar pattern is observed for the character "flower length". It is, thus, not surprising that finding the morphological characters separating the *Abebaia* clade from *Manilkara* s.str. remains an unresolved issue. The broad range of morphological variation in *Manilkara* s.str. does not allow a clear-cut distinction between the two clades. However, Armstrong (2013) pointed out "the tendency for the leaves to have striate venation, where the tertiary veins are nearly indistinguishable from the secondary veins" in *M. dissecta*, *M. fasciculata* and *M. udoido*. Nevertheless, she also admitted that this character is not consistently present in one of the species: *M. dissecta* (L.f.) Dubard. Our morphological analysis supports similar patterns in other floral parts, such as the staminodes, which have a tendency to be broad compared to the generally long and narrow in *Manilkara* s.str.

The three species of the *Abebaia* clade are in sympatry with other Pacific *Manilkara*, except *M. udoido*, which, in addition to being endemic to Palau, is the only Manilkarinae recorded on the archipelago (Armstrong, 2013).

These results highlight that morphology does not necessarily reflect the evolutionary history, as revealed by a molecular approach. The relationships between lineages can be obscured by the complexity of evolutionary processes such as morphological convergence between distinct lineages or, alternatively, conservation of ancestral traits. The latter could be a putative consequence of large population sizes which would have limited the effect of genetic drift or the consequence of unknown selective effects. Thus, understanding the evolutionary processes behind this molecular and morphological incongruence would deserve more effort.

The *Abebaia* clade diverged from the *Labourdonnaisia-Faucherea* clade at around 30 mya (HPD 36–24 mya). The MRCA of the two lineages has been reconstructed as being of Malagasy origin (Armstrong & al., 2014). This implies that the three members of the *Abebaia* clade have a different origin than the core Pacific *Manilkara*, which were derived from Africa (Armstrong & al., 2014).

Given this evidence, the three species of the *Abebaia* clade should be considered a separate genus, morphologically distinct from its sister group the genus *Labourdonnaisia* but still very similar to *Manilkara* s.str. It would be tempting to qualify this genus as "cryptic", but much more work, including field observations, is needed.

As mentioned above, one of the species, *Manilkara fasciculata* (Warb.) H.J.Lam & Maas Gest., is the type of *Abebaia* Baehni (1964), and consequently this name has to be resurrected to accommodate the three species. It should be noted that in the description of the genus, *Abebaia fasciculata* was excluded from *Manilkara* on the basis of the number of staminodes (Baehni, 1964). Baehni (1965) stated that "*Abebaia* is a Manilkarinae with a variable number of staminodes". This definition is, however, untenable according to the specimens of *A. fasciculata* currently available, as well as when considering the other two species now included. We, therefore, provide below an emended description of the genus, based on the character states of its three species.

While describing *Abebaia*, Baehni (1964) suggested that *Manilkara vitiensis* (H.J.Lam & Olden) B.Meeuse could possibly represent a second species in the genus. Armstrong & al. (2014) demonstrated that this was not the case, the species being included in the *Manilkara* s.str. clade.

It should finally be noted that three rare species of Asian-Pacific *Manilkara* were not available for molecular analyses and that they could possibly belong to *Abebaia*. Two of them, *Manilkara celebica* H.J.Lam and *M. samoensis* H.J.Lam, are distributed within the area of the genus *Abebaia*, while the third one, *M. roxburghiana* (Wight) Dubard would extend its distribution as far as southern peninsular India. Including them in future phylogenetic analyses could help clarifying this issue.

■ TAXONOMIC IMPLICATIONS

As a consequence of the results exposed above, we provide below a brief taxonomic overview of the genus *Abebaia* including an emended generic description, citation of the three currently included species (two being here newly combined in the genus) with typification and distribution. A complete taxonomic treatment of these three species, including full synonymy can be found in Armstrong (2013) where they are treated under *Manilkara*. Regarding the other genera, they are currently under investigation, and a complete taxonomic treatment, including full synonymy and typification, will be published later. We will here only provide an emended description of the genus *Labourdonnaisia* and the necessary combinations to accommodate the currently accepted species of *Faucherea* that do not yet have a name in *Labourdonnaisia*.

Abebaia Baehni in Arch. Sci. 17: 78. 1964 – Type: *A. fasciculata* (Warb.) Baehni (≡ *Mimusops fasciculata* Warb.).

Emended description. – Trees; leaves coriaceous, generally glabrous above and below; leaf blades with striate venation, i.e., the secondary veins generally indistinct from the tertiaries; flowers with a biseriate trimerous calyx. Corolla tube short, less than 1/2 the corolla lobes length, 6 lobes, each with a pair of appendages 0.5–0.8 times as long as the lobes, often entire; stamens isomerous, opposite to corolla lobes; staminodes isomerous, alternate with stamens, broad (breadth/length ratio 0.5–1.6), 0.1–0.6 times as long as the corolla lobes; ovary pubescent or pilose, with 6–8 locules. Fruit single- or multi-seeded, seed scar basiventral or basal.

Diversity and distribution. – Three Pacific species, possibly more, distributed in Indonesia (Borneo, Sulawesi, Moluccas, Papua), Philippines, Palau, New Caledonia, Vanuatu, Samoa, Tonga, Fiji.

Abebaia dissecta (L.f.) Randriarisoa & K.Armstr., **comb. nov.** ≡ *Achras dissecta* L.f., Suppl. Pl.: 210. 1782 – Holotype: Tonga-Tabu [Tonga], *J.G.A. Forster s.n.* (LINN-HS No. 618.1!)

Distribution. – New Caledonia, Vanuatu, Samoa, Tonga, Fiji.

Note. – The LINN-HS specimen was cited as lectotype by Armstrong (2013). However, as explained in Nicolson & Fosberg (2004), it is the only specimen that the author had access to. No lectotypification is therefore needed and this specimen should be considered the holotype. Information on further original material can be found in Nicolson & Fosberg (2004).

Abebaia fasciculata (Warb.) Baehni in Arch. Sci. 17: 78. 1964 ≡ *Mimusops fasciculata* Warb. in Bot. Jahrb. Syst. 13: 401. 1891 – Lectotype (designated by Armstrong in Edinburgh J. Bot. 70: 18. 2013): Indonesia, West Papua, 1888, *O. Warburg 21361* (E barcode E00570193!).

Distribution. – Indonesia (Borneo, Sulawesi, Moluccas, Papua), Philippines.

Abebaia udoido (Kaneh.) Randriarisoa & K.Armstr., **comb. nov.** ≡ *Manilkara udoido* Kaneh., Fl. Micron.: 304. 1933 – Lectotype (designated by Armstrong in Edinburgh J. Bot. 70: 46. 2013): Palau, Aimeliik, Aug 1932, *R. Kanehira 1925* (FU [no barcode attributed, n.v.]; isolectotypes: K barcode K000229505 [image!], NY barcode 00273539 [image!], P barcode P00526499 [image!], US barcode 00113377 [image!]).

Distribution. – Palau (endemic).

Labourdonnaisia Bojer in Mém. Soc. Phys. Genève 9: 295. 1842 – Type (designated by Baehni in Boissiera 11: 147. 1965): *L. sarcophleia* Bojer (= *L. calophylloides* Bojer).
= *Faucherea* Lecomte in Bull. Mus. Hist. Nat. (Paris) 26: 245. 1920, **syn. nov.** – Lectotype (designated by Aubréville in Adansonia, ser. 2, 11: 280. 1971): *F. hexandra* (Lecomte) Lecomte (≡ *Labourdonnaisia hexandra* Lecomte).

Emended description. – Trees; leaves coriaceous; leaf blades with striate venation, i.e., the secondaries generally indistinct from the tertiaries; flowers with a biseriate trimerous calyx. Corolla tube generally very short [tube/lobe ratio 0.05–0.3 (0.7)], lobes 6–18, appendages absent, rarely vestigial; stamens isomerous, opposite with respect to corolla lobes; staminodes absent or vestigial and in that case isomerous and alternate with stamens, 0.1–0.3 times as long as the corolla lobes; ovary pubescent or pilose, with 6–10 locules. Fruit generally single-seeded, seed scar ovate, basiventral.

Diversity and distribution. – Seventeen described species from Madagascar and the Mascarenes, probably several more.

Labourdonnaisia ambrensis (R.Capuron ex Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea ambrensis* R.Capuron ex Aubrév. in Adansonia, ser. 2, 11: 288. 1971.

Labourdonnaisia glutinosa (Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea glutinosa* Aubrév. in Adansonia, ser. 2, 11: 287. 1971.

Labourdonnaisia laciniata (Lecomte) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea laciniata* Lecomte in Bull. Mus. Hist. Nat. (Paris) 26: 251. 1920.

Labourdonnaisia longepedicellata (Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea longepedicellata* Aubrév. in Adansonia, ser. 2, 11: 282. 1971.

Labourdonnaisia manongarivensis (Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea manongarivensis* Aubrév. in Adansonia, ser. 2, 11: 283. 1971.

Labourdonnaisia parvifolia (Lecomte) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea parvifolia* Lecomte in Bull. Mus. Hist. Nat. (Paris) 26: 251. 1920.

Labourdonnaisia sambiranensis (Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea sambiranensis* Aubrév. in Adansonia, ser. 2, 11: 285. 1971.

Labourdonnaisia tampoloensis (Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea tampoloensis* Aubrév. in Adansonia, ser. 2, 11: 285. 1971.

Labourdonnaisia thouvenotii (Lecomte) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea thouvenotii* Lecomte in Bull. Mus. Hist. Nat. (Paris) 26: 248. 1920.

Labourdonnaisia urschii (R.Capuron ex Aubrév.) L.Gaut. & Randriarisoa, **comb. nov.** ≡ *Faucherea urschii* R.Capuron ex Aubrév. in Adansonia, ser. 2, 11: 285. 1971.

■ AUTHOR CONTRIBUTIONS

AR, YN and LG designed the research. AR, CGB, YN and LG collected samples in Madagascar. SD provided some Mascarenes samples for molecular analyses and KA provided some Asian *Manilkara* measurements for morphological analysis. AR conducted the laboratory work. AR and CP analyzed the data. All authors interpreted the results. AR prepared the manuscript under the supervision of YN and LG, and all authors contributed to the final submission. — AR, https://orcid.org/0000-0003-3990-507X; YN, https://orcid.org/0000-0001-6784-8565; KA, https://orcid.org/0000-0001-9850-5717; CGB, https://orcid.org/0000-0001-7922-8718; CP, https://orcid.org/0000-0001-7766-3732; LG, https://orcid.org/0000-0003-4157-3713

■ ACKNOWLEDGMENTS

AR was funded by the Swiss Government Excellence Ph.D. scholarship from September 2018 to August 2021 and by the University of Geneva from September to December 2021. This work is part of a project led by LG on the Malagasy Sapotaceae, financially supported by the Franklinia foundation (grant No. 2019-20) and formerly supported by a grant attributed to YN and LG by the Swiss National Foundation (grant no. 31003A_166349/1) and by two grants attributed to YN by the Fondation Ernst et Lucie Schmidheiny. We would like to express our gratitude to Prof. Roman Ulm and Martine Mir for financial support and fund management,

respectively. We thank the curators of the herbaria G, MO, P, TAN and TEF for allowing us to study their specimens and to perform limited destructive sampling. We also thank people at the iGE3 platform for their help with the sequencing process (https://ige3.genomics.unige.ch). Some computations were performed at the University of Geneva on the Baobab/Yggdrasil cluster. We are grateful to Camille Christe for her precious advice on the bioinformatics analyses, to Regine Niba for her help in the lab, to Romain Dewaele for his help with the figures and to Ulf Swenson for fruitful discussions at the inception of this project. We are grateful to the Malagasy government for providing us with a research permit to collect samples. We would also like to thank Richard Randrianaivo from the MBG, Patrick Ranirison and Jacquis Tahinarivony from the University of Antananarivo (DBEV) for facilitating the fieldtrips in Madagascar; as well as the Malagasy people on site for their help collecting Sapotaceae specimens. SD would like to thank the National Parks and Conservation Service (NPCS) of Mauritius for permission to sample in the national park.

■ LITERATURE CITED

Agnarsson, I. & Kuntner, M. 2012. The generation of a biodiversity hotspot: Biogeography and phylogeography of the Western Indian Ocean islands. Pp. 33–82 in: Anamthawat-Jonsson, K. (ed.), *Current topics in phylogenetics and phylogeography of terrestrial and aquatic systems*. [s.l.]: Tech Publishers. https://doi.org/10.5772/1947

Anderberg, A.A. & Swenson, U. 2003. Evolutionary lineages in Sapotaceae (Ericales): A cladistic analysis based on *ndhF* sequence data. *Int. J. Pl. Sci.* 164: 763–773. https://doi.org/10.1086/376818

Armstrong, K.E. 2010. *Systematics and biogeography of the pantropical genus Manilkara Adans. (Sapotaceae)*. Ph.D. thesis. University of Edinburgh, Edinburgh, U.K.

Armstrong, K.E. 2013. A revision of the Pacific species of *Manilkara* (Sapotaceae). *Edinburgh J. Bot.* 70: 7–56. https://doi.org/10.1017/S0960428612000327

Armstrong, K.E., Stone, G.N., Nicholls, J.A., Valderrama, E., Anderberg, A.A., Smedmark, J., Gautier, L., Naciri, Y., Milne, R. & Richardson, J.E. 2014. Patterns of diversification amongst tropical regions compared: A case study in Sapotaceae. *Frontiers Genet.* 5: 362. https://doi.org/10.3389/fgene.2014.00362

Aubréville, A. 1936. *La flore forestière de la Côte d'Ivoire*, 1st ed., 3 vols. Paris: La Rose.

Aubréville, A. 1964. Les Sapotacées: Taxonomie et phytogéographie. *Adansonia, Mém.* 1: 1–157.

Aubréville, A. 1974. *Flore de Madagascar et des Comores*, 164 fam., *Sapotaceae*. Paris: Museum National d'Histoire Naturelle. https://doi.org/10.5962/bhl.title.6600

Baehni, C. 1964. Genres nouveaux de Sapotacées. *Arch. Sci.* 17: 78.

Baehni, C. 1965. Mémoires sur les Sapotacées. III. Inventaire des genres. *Boissiera* 11: 1–262.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boluda, C.G., Christe, C., Randriarisoa, A., Gautier, L. & Naciri, Y. 2021. Species delimitation and conservation in taxonomically challenging lineages: The case of two clades of *Capurodendron* (Sapotaceae) in Madagascar. *Plants (Switzerland)* 10: 1702. https://doi.org/10.3390/plants10081702

Boluda, C.G., Christe, C., Gautier, L. & Naciri, Y. 2022. A 638-gene phylogeny supports the recognition of twice as many species in the Malagasy endemic genus *Capurodendron* (Sapotaceae). *Taxon* 71: 360–395. https://doi.org/10.1002/tax.12676

Borg, D., Richardson, J.E., Harris, D.J., Gautier, L., Hughes, M. & Mackinder, B. 2019. Phylogeny of two African genera of Sapotaceae – *Englerophytum* and *Synsepalum*. *Edinburgh J. Bot.* 76: 231–267. https://doi.org/10.1017/S0960428619000040

Borowiec, M.L. 2016. AMAS: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4: e1660. https://doi.org/10.7717/peerj.1660

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C., Xie, D., Suchard, M., Rambaut, A. & Drummond, A.J. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computat. Biol.* 10(4): e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Buerki, S., Devey, D., Callmander, M., Phillipson, P. & Forest, F. 2013. Spatio-temporal history of the endemic genera of Madagascar. *Bot. J. Linn. Soc.* 171: 304–329. https://doi.org/10.1111/boj.12008

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T. 2009. TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chase, M.W. & Hills, H.H. 1991. Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon* 40: 215–220. https://doi.org/10.2307/1222975

Christe, C., Boluda, C.G., Koubínová, D., Gautier, L. & Naciri, Y. 2021. New genetic markers for Sapotaceae phylogenomics: More than 600 nuclear genes applicable from family to population levels. *Molec. Phylogen. Evol.* 160: 107–123. https://doi.org/10.1016/j.ympev.2021.107123

Crameri, S. 2020. *Phylogenomics, species discovery and integrative taxonomy in Dalbergia (Fabaceae) precious woods from Madagascar*. Ph.D. thesis. ETH Zurich, Switzerland.

Dafreville, S. 2013. *Diversité et structuration génétiques des Sapotacées endémiques de l'archipel des Mascareignes à différentes échelles spatiales et temporelles*. Ph.D. thesis. Université de la Réunion, Saint-Denis, Réunion.

Debeuf, D. 2004. *Étude de l'évolution volcano-structurale et magmatique de Mayotte, Archipel des Comores, Océan Indien: Approches structurale, pétrographique, géochimique et géochronologique*. Ph.D. thesis. Université de la Réunion, Saint-Denis, Réunion.

Doyle, J.J. & Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leave tissue. *Phytochem. Bull. Bot. Soc. Amer.* 19: 11–15.

Emerick, C.M. & Duncan, R.A. 1982. Age progressive volcanism in the Comoros Archipelago, western Indian Ocean and implications for Somali plate tectonics. *Earth Planet. Sci. Lett.* 60: 415–428. https://doi.org/10.1016/0012-821X(82)90077-2

Escofier, B. 1979. Traitement simultané de variables qualitatives et quantitatives en analyse factorielle. *Cah. Analyse Données* 4: 137–146. http://www.numdam.org/item/CAD_1979__4_2_137_0/

Federman, S., Dornburg, A., Downie, A., Richard, A.F., Daly, D.C. & Donoghue, M.J. 2015. The biogeographic origin of a radiation of trees in Madagascar: Implications for the assembly of a tropical forest biome. *B. M. C. Evol. Biol.* 15: 216. https://doi.org/10.1186/s12862-015-0483-1

Federman, S., Donoghue, M.J., Daly, D.C. & Eaton, D.A.R. 2018. Reconciling species diversity in a tropical plant clade (*Canarium*, Burseraceae). *PLoS ONE* 13(6): e0198882. https://doi.org/10.1371/journal.pone.0198882

Friedmann, F. 1981. *Flore des Mascareignes*, 116, *Sapotacées*. Mauritius: The Sugar Industry Research Imstitute.

Galili, T. 2015. dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* 31: 3718–3720. https://doi.org/10.1093/bioinformatics/btv428

Gatesy, J. & Baker, R.H. 2005. Hidden likelihood support in genomic data: Can forty-five wrongs make a right? *Syst. Biol.* 54: 483–492. https://doi.org/10.1080/10635150590945368

Gautier, L., Naciri, Y., Anderberg, A.A., Smedmark, J.E.E., Randrianaivo, R. & Swenson, U. 2013. A new species, genus and tribe of Sapotaceae, endemic to Madagascar. *Taxon* 62: 972–983. https://doi.org/10.12705/625.17

Gautier, L., Randrianaivo, R., Boluda, G.C., Randriarisoa, A. & Naciri, Y. 2022. Sapotaceae. Pp. 726–739 in: Goodman, S.M.

(ed.), *The new natural history of Madagascar*. Princeton: Princeton University Press.

Glor, R.E. 2010. Phylogenetic insights on adaptive radiation. *Annual Rev. Ecol. Evol. Syst.* 41: 251–270. https://doi.org/10.1146/annurev.ecolsys.39.110707.173447

Hajash, A. & Armstrong, R.L. 1972. Paleomagnetic and radiometric evidence for the age of the Comores islands, west central Indian Ocean. *Earth Planet. Sci. Lett.* 16: 231–236. https://doi.org/10.1016/0012-821X(72)90195-1

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis. Program for Windows 95/98/NT. *Nucl. Acids. Symp.* 41: 95–98.

Harley, M.M. 1991. The pollen morphology of the Sapotaceae. *Kew Bull.* 46: 379–491. https://doi.org/10.2307/4110538

Heritage, S. 2021. MBASR: Workflow-simplified ancestral state reconstruction of discrete traits with MrBayes in the R environment. *BioRxiv*. https://doi.org/10.1101/2021.01.10.426107

Hill, M.O. & Smith, A.J.E. 1976. Principal component analysis of taxonomic data with multistate discrete characters. *Taxon* 25: 249–255. https://doi.org/10.2307/1219449

Jacobs, B.F., Tabor, N., Feseha, M., Pan, A., Kappelman, J., Rasmussen, T., Sanders, W., Wiemann, M., Crabaugh, J. & Massini, J.L.G. 2005. Oligocene terrestrial strata of north western Ethiopia: A preliminary report on paleoenvironments and paleontology. *Palaeontol. Electronica* 8.1.25A.

Jeffroy, O., Brinkmann, H., Delsuc, F. & Philippe H. 2006. Phylogenomics: The beginning of incongruence? *Trends Genet.* 22: 225–231. https://doi.org/10.1016/j.tig.2006.02.003

Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J.C. & Wickett, N.J. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applic. Pl. Sci.* 4: 1600016. https://doi.org/10.3732/apps.1600016

Kassambara, A. & Mundt, F. 2020. factoextra: Extract and visualize the results of multivariate data analyses, R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

Katoh, K. & Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molec. Biol. Evol.* 30: 772–780. https://doi.org/10.1093/molbev/mst010

Kaufman, L. & Rousseeuw, P.J. 1990. *Finding groups in data: An introduction to cluster analysis*. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9780470316801

Kircher, M., Sawyer, S. & Meyer, M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucl. Acids Res.* 40(1): e3. https://doi.org/10.1093/nar/gkr771

Kümpers, B.M.C., Anderberg, A.A., Richardson, J-E., Wilkie, P. & Ronse De Craene, L.P. 2016. The significance of meristic changes in the flowers of Sapotaceae. *Bot. J. Linn. Soc.* 180: 161–192. https://doi.org/10.1111/boj.12363

Labat, J.-N., Pignal, M. & Pascal, O. 1997. Une nouvelle espèce de *Labramia* (Sapotaceae) de l'Ile de Mayotte dans l'Archipel des Comores. *Adansonia*, ser. 3: 19: 213–216.

Lecomte, H. 1920. Genre nouveau de la famille des Sapotacées. *Bull. Mus. Hist. Nat. (Paris)* 26: 245–253.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Pp. 281–297 in: Le Cam, L.M. & Neyman, J. (eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, *Statistics*. Berkeley & Los Angeles: University of California Press.

Mirarab, S. 2019. Species tree estimation using ASTRAL: Practical considerations. *arXiv* 1904.03826v2. https://doi.org/10.48550/arXiv.1904.03826

Mirarab, S. & Warnow, T. 2015. ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31: i44–i52. https://doi.org/10.1093/bioinformatics/btv234

Mirarab, S., Reaz, R., Bayzid, Md.S., Zimmermann, T., Swenson, M.S. & Warnow, T. 2014. ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics* 30: i541–i548. https://doi.org/10.1093/bioinformatics/btu462

Naciri, Y. & Linder, H.P. 2015. Species delimitation and relationships: The dance of the seven veils. *Taxon* 64: 3–16. https://doi.org/10.12705/641.24

Naciri, Y. & Linder, H.P. 2020. The genetics of evolutionary radiations. *Biol. Rev. (Cambridge)* 95: 1055–1072. https://doi.org/10.1111/brv.12598

Nicolson, D.H. & Fosberg, F.R. 2004. *The Forsters and the botany of the Second Cook Expedition (1772–1775)*, 2nd ed. Regnum Vegetabile 139. Ruggell: Gantner.

Nougier, J., Cantagrel, J.M. & Karche, J.P. 1986. The Comores archipelago in the western Indian Ocean: Volcanology, geochronology and geodynamic setting. *J. African Earth Sci.* 5: 135–145. https://doi.org/10.1016/0899-5362(86)90003-5

Pagès, J. 2004. Analyse factorielle de données mixtes. *Rev. Statist. Appl.* 4: 93–111.

Pagès, J. 2013. *Analyse factorielle multiple avec R*. Paris: EDP sciences.

Paradis, E. & Schliep, K. 2019. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pennington, T.D. 1991. *The genera of Sapotaceae*. Richmond: Royal Botanic Gardens, Kew.

Pollard, D.A., Iyer, V.N., Moses, A.M. & Eisen, M.B. 2006. Widespread discordance of gene trees with species tree in *Drosophila*: Evidence for incomplete lineage sorting. *PLoS Genet.* 2: e173. https://doi.org/10.1371/journal.pgen.0020173

Pouchon, C., Boyer, F., Roquet, C., Denoeud, F., Chave, J., Coissac, E., Alsos, I.G., The PhyloAlps Consortium, The PhyloNorway Consortium & Lavergne, S. 2022. ORTHOSKIM: In silico sequence capture from genomic and transcriptomic libraries for phylogenomic and barcoding applications. *Molec. Ecol. Resources* 22: 2018–2037. https://doi.org/10.1111/1755-0998.13584

Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A.J. 2014. Tracer, version 1.6. http://beast.community/tracer.html

Randriarisoa, A., Naciri, Y. & Gautier, L. 2020. *Labramia ambondrombeensis* (Sapotaceae), a Critically Endangered new species from Madagascar. *Candollea* 75: 83–87. https://doi.org/10.15553/c2020v751a8

Revell, L.J. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Meth. Ecol. Evol.* 3: 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Sayyari, E. & Mirarab, S. 2016. Fast coalescent-based computation of local branch support from quartet frequencies. *Molec. Biol. Evol.* 33: 1654–1668. https://doi.org/10.1093/molbev/msw079

Schliep, K. 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics* 27: 592–593. https://doi.org/10.1093/bioinformatics/btq706

Shi, C.M. & Yang, Z. 2017. Coalescent-based analyses of genomic sequence data provide a robust resolution of phylogenetic relationships among major groups of gibbons. *Molec. Biol. Evol.* 35: 159–179. https://doi.org/10.1093/molbev/msx277

Smedmark, J.E.E., Swenson, U. & Anderberg, A.A. 2006. Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Molec. Phylogen. Evol.* 39: 706–721. https://doi.org/10.1016/j.ympev.2006.01.018

Souza, H.A.V., Muller, L.A.C., Brandão, R.L. & Lovato, M.B. 2012. Isolation of high quality and polysaccharide-free DNA from leaves of *Dimorphandra mollis* (Leguminosae), a tree from the Brazilian Cerrado. *Genet. Molec. Res.* 11: 756–764. https://doi.org/10.4238/2012.March.22.6

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stride, G., Nylinder, S. & Swenson, U. 2014. Revisiting the biogeography of *Sideroxylon* (Sapotaceae) and an evaluation of the taxonomic status of *Argania* and *Spiniluma*. *Austral. Syst. Bot.* 27: 104–118. https://doi.org/10.1071/SB14010

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J. & Rambaut, A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4: vey016. https://doi.org/10.1093/ve/vey016

Swenson, U. & Anderberg, A.A. 2005. Phylogeny, character evolution, and classification of Sapotaceae (Ericales). *Cladistics* 21: 101–130. https://doi.org/10.1111/j.1096-0031.2005.00056.x

Swenson, U., Nylinder, S. & Munzinger, J. 2013. Towards a natural classification of Sapotaceae subfamily Chrysophylloideae in Oceania and Southeast Asia based on nuclear sequence data. *Taxon* 62: 746–770. https://doi.org/10.12705/624.11

Swenson, U., Lowry, P.P., Cronholm, B. & Nylinder, S. 2020. Resolving the relationships of the enigmatic Sapotaceae genera *Beauvisagea* and *Boerlagella*, and the position of *Planchonella suboppositifolia*. *Taxon* 69: 998–1015. https://doi.org/10.1002/tax.12313

Thébaud, C., Warren, B.H., Cheke, A. & Strasberg, D. 2009. Mascarene Islands, biology. In: Gillepsie, R.G. & Clague, D.A. (eds.), *The Encyclopedia of islands*. Berkeley: University of California Press.

Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*, 4th ed. New York: Springer. https://doi.org/10.1007/978-0-387-21706-2

Yoder, A.D. & Nowak, M.D. 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Rev. Ecol. Evol. Syst.* 37: 405–431. https://doi.org/10.1146/annurev.ecolsys.37.091305.110239

Appendix 1. Voucher information.

Original identification in bold face, followed by country of origin, collection year, collector and number in *italics*, herbarium code, and BioSample number.

Faucherea ambrensis Capuron ex Aubrév., Madagascar, 2005, *Antilahimena 4342* (G), SAMN29129799. ***Faucherea ambrensis***, Madagascar, 2006, *Gautier 5007* (G), SAMN29129833. ***Faucherea ambrensis***, Madagascar, 2008, *Trigui 349* (G), SAMN29129798. ***Faucherea glutinosa*** Aubrév., Madagascar, 2004, *Rabevohitra 5062* (G), SAMN29129793. ***Faucherea glutinosa***, Madagascar, 2011, *Randrianaivo 1847* (G), SAMN29129863. ***Faucherea hexandra*** (Lecomte) Lecomte, Madagascar, 2013, *Gautier 6009* (G), SAMN29129804. ***Faucherea hexandra***, Madagascar, 2018, *Randriarisoa 142* (G), SAMN29129810. ***Faucherea laciniata*** Lecomte, Madagascar, 1997, *Rakotomalaza 1252* (G), SAMN29129834. ***Faucherea laciniata***, Madagascar, 2018, *Randriarisoa 171* (G), SAMN29129808. ***Faucherea laciniata***, Madagascar, 2018, *Randriarisoa 173* (G), SAMN29129809. ***Faucherea manongarivensis*** Aubrév., Madagascar, 2006, *Callmander 571* (G), SAMN29129803. ***Faucherea parvifolia*** Lecomte, Madagascar, 1997, *Rakotomalaza 1290* (G), SAMN29129835. ***Faucherea tampoloensis*** Aubrév., Madagascar, 1990, *Schatz 2927* (G), SAMN29129823. ***Faucherea tampoloensis***, Madagascar, 2010, *Gautier 5508* (G), SAMN29129801. ***Faucherea thouvenotii*** Lecomte, Madagascar, 2009, *Gautier 5377* (G), SAMN29129802. ***Faucherea thouvenotii***, Madagascar, 2018, *Randrianaivo 2997* (G), SAMN29129865. ***Faucherea urschii*** Aubrév., Madagascar, 2008, *Ravelonarivo 2701* (G), SAMN29129864. ***Faucherea* sp. 2**, Madagascar, 2004, *Randrianaivo 1106* (G), SAMN29129806. ***Faucherea* sp. 2**, Madagascar, 2010, *Gautier 5431* (G), SAMN29129807. ***Faucherea* sp. 3**, Madagascar, 2010, *Gautier 5547* (G), SAMN29129805. ***Labourdonnaisia calophylloides*** Bojer, Réunion, 2009, *Swenson 835* (S), SAMN29129797. ***Labourdonnaisia glauca*** Bojer, Mauritius, 2010, *Dafreville LG23* (MAU), SAMN29129816. ***Labourdonnaisia*** cf. ***lecomtei*** Aubrév., Madagascar, 2018, *Randrianaivo 3171* (G), SAMN29129812. ***Labourdonnaisia madagascariensis*** Pierre ex Baill., Madagascar, 1956, *Service Forestier 15317* (G), SAMN29129836. ***Labourdonnaisia madagascariensis***, Madagascar, 2010, *Gautier 5546* (G), SAMN29129796. ***Labourdonnaisia madagascariensis***, Madagascar, 2011, *Gautier 5754* (G), SAMN29129795. ***Labourdonnaisia madagascariensis***, Madagascar, 1952, *Service Forestier 4429* (G), SAMN29129794. ***Labourdonnaisia madagascariensis***, Madagascar, 2011, *Gautier 5776* (G), SAMN29129792. ***Labourdonnaisia revoluta*** Bojer, Mauritius, 2010, *Dafreville LR43* (MAU), SAMN29129817. ***Labourdonnaisia madagascariensis***, Madagascar, 2018, *Randriarisoa 85* (G), SAMN29129811. ***Labramia ankaranaensis*** Aubrév., Madagascar, 2006, *Rogers 1165* (G), SAMN29129843. ***Labramia ankaranaensis*** var. ***antsingensis*** Aubrév., Madagascar, 1952, *Leandri 1964* (G), SAMN29129858. ***Labramia ankaranaensis*** var. ***antsingensis***, Madagascar, 1953, *Service Forestier 8436* (P), SAMN29129844. ***Labramia ankaranaensis*** var. ***antsingensis***, Madagascar, 2011, *Gautier 5582* (G), SAMN29129777. ***Labramia ankaranaensis*** var. ***antsingensis***, Madagascar, 2016, *Gautier 6247* (G), SAMN29129819. ***Labramia boivinii*** (Pierre) Aubrév., Madagascar, 1954, *Service Forestier 10735* (P), SAMN29129857. ***Labramia bojeri*** A.DC., Madagascar, 1997, *Randrianaivo 129* (G), SAMN29129856. ***Labramia bojeri***, Madagascar, 2000, *Faliniaina 52* (G), SAMN29129851. ***Labramia bojeri***, Madagascar, 2002, *Rabevohitra 4190* (G), SAMN29129848. ***Labramia bojeri***, Madagascar, 2004, *Ravelonarivo 3539* (G), SAMN29129785. ***Labramia bojeri***, Madagascar, 2010, *Gautier 5528* (G), SAMN29129790. ***Labramia bojeri***, Madagascar, 2012, *Ramanjanahary 674* (G), SAMN29129849. ***Labramia bojeri***, Madagascar, 2012, *Ratovoson 1911* (G), SAMN29129850. ***Labramia bojeri***, Madagascar, 2018, *Randriarisoa 96* (G), SAMN29129860. ***Labramia capuronii*** Aubrév., Madagascar, 2002, *Randrianasolo 330* (G), SAMN29129862. ***Labramia capuronii***, Madagascar, 2018, *Randrianaivo 3159* (G), SAMN29129847. ***Labramia costata*** (M.M.Hartog ex Baill.) Aubrév., Madagascar, 1967, *Service Forestier 27622* (G), SAMN29129846. ***Labramia costata***, Madagascar, 2010, *Gautier 5519* (G), SAMN29129845. ***Labramia costata***, Madagascar, 2011, *Gautier 5752* (G), SAMN29129778. ***Labramia louvelii*** Aubrév., Madagascar, 2011, *Gautier 5749* (G), SAMN29129789. ***Labramia louvelii***, Madagascar, 2011, *Gautier 5798* (G), SAMN29129779. ***Labramia mayottensis*** Labat, Pignal & O.Pascal, Comoros, 1996, *Pascal 620* (G), SAMN29129781. ***Labramia mayottensis***, Comoros, 1996, *Pascal 700* (G), SAMN29129837. ***Labramia mayottensis***, Comoros, 2001, *Pignal 1844* (G), SAMN29129782. ***Labramia platanoides*** Capuron ex Aubrév., Madagascar, 1962, *Service Forestier 20981* (P), SAMN29129840. ***Labramia platanoides***, Madagascar, 1962, *Service Forestier 22052* (G), SAMN29129784. ***Labramia platanoides***, Madagascar, 2010, *Gautier 5417* (G), SAMN29129838. ***Labramia sambiranensis*** Aubrév., Madagascar, 2009, *Tahinarivony 292* (G), SAMN29129854. ***Labramia sambiranensis***, Madagascar, 2013, *Gautier 6069* (G), SAMN29129783. ***Labramia* sp. 1**, Madagascar, 2007, *Gautier 5211* (G), SAMN29129788. ***Labramia* sp. 1**, Madagascar, 2008, *Trigui 451* (G), SAMN29129839. ***Labramia* sp. 3**, Madagascar, 1997, *Randrianaivo 124* (G), SAMN29129787. ***Labramia* sp. 3**, Madagascar, 2010, *Bernard 1762* (G), SAMN29129786. ***Labramia* sp. 4**, Madagascar, 2005, *Razanatsima 39* (G), SAMN29129853. ***Labramia* sp. 4**, Madagascar, 2006, *Razanatsima 169* (G), SAMN29129852. ***Labramia* sp. 6**, Madagascar, 1956, *RN 8555* (G), SAMN29129841. ***Labramia* sp. 6**, Madagascar, 2016, *Gautier 6274* (G), SAMN29129855. ***Labramia* sp. 6**, Madagascar, 2016, *Gautier 6308* (G), SAMN29129859. ***Labramia* sp. 6**, Madagascar, 2016, *Gautier 6310* (G), SAMN29129842. ***Labramia* sp. 8**, Madagascar, 2013, *Gautier 6025* (G), SAMN29129780. ***Labramia* sp. 8**, Madagascar, 2018, *Randrianaivo 3048* (G), SAMN29129861. ***Manilkara bidentata*** (A.DC.) A.Chev., Brazil, 1993, *Ribeiro 928* (G), SAMN29129830. ***Manilkara boivinii*** Aubrév., Madagascar, 1999, *Gautier 3477* (G), SAMN29129818. ***Manilkara cuneifolia*** (Baker) Dubard, Gabon, 1994, *de Wilde 11385* (G), SAMN29129821. ***Manilkara dissecta*** (L.f.) Dubard, Samoa, 1977, *Whistler W 3889* (BISH), SAMN29129815. ***Manilkara fasciculata*** (Warb.) H.J.Lam & Maas Geest., Indonesia, 2008, *Armstrong 353* (E), SAMN29129814. ***Manilkara hexandra*** (Roxb.) Dubard, Sri Lanka, 1974, *Kostermans 25308* (G), SAMN29129822. ***Manilkara lacera*** (Baker) Dubard, Gabon, 2005, *Harris 8220A* (G), SAMN29129800. ***Manilkara longifolia*** (A.DC.) Dubard, Brazil, 1998, *Sant'ana 675* (G), SAMN29129828. ***Manilkara multinervis*** (Baker) Dubard, Côte d'Ivoire, 1986, *Poilecot 1105CI* (G), SAMN29129832. ***Manilkara obovata*** (Sabine & G.Don) J.H.Hemsl., Gabon, 2003, *Jongkind 5891* (G), SAMN29129831. ***Manilkara sansibarensis*** (Engl.) Dubard, NA, 2005, *Kindekata 2668* (G), SAMN29129820. ***Manilkara udoido*** Kaneh., Palau, 1996, *Slappy LR26622* (BISH), SAMN29129813. ***Manilkara zapota*** (L.) P.Royen, Costa Rica, 2001, *Chavarria 2187* (G), SAMN29129829. — OUTGROUPS: ***Baillonella toxisperma*** Pierre, Gabon, 1999, *Breteler 14777* (G), SAMN29129826. ***Mimusops*** cf. ***antorakensis*** Aubrév., Madagascar, 2018, *Randrianaivo 2989* (G), SAMN29129791. ***Tieghemella heckelii*** (A.Chev.) Pierre ex Dubard, Côte d'Ivoire, 1980, *Zwetstloot 33* (G), SAMN29129824. ***Vitellaria paradoxa*** C.F.Gaertn., Ghana, 1999, *Schmidt 3309* (G), SAMN29129825. ***Vitellariopsis cuneata*** (Engl.) Aubrév., Tanzania, 1999, *Mwangoka* 1028 (G), SAMN29129827.