American Journal of Botany

RESEARCH ARTICLE

INVITED SPECIAL ARTICLE

For the Special Issue: Exploring Angiosperms353: a Universal Toolkit for Flowering Plant Phylogenomics

A nuclear phylogenomic study of the angiosperm order Myrtales, exploring the potential and limitations of the universal Angiosperms353 probe set

Olivier Maurin[1,25], Artemis Anest[2,3], Sidonie Bellot[1], Edward Biffin[4,5], Grace Brewer[1], Tristan Charles-Dominique[6], Robyn S. Cowan[1], Steven Dodsworth[7], Niroshini Epitawalage[1], Berta Gallego[1], Augusto Giaretta[8], Renato Goldenberg[9], Deise J.P. Gonçalves[10], Shirley Graham[11], Peter Hoch[11], Fiorella Mazine[12], Yee Wen Low[1,13,14], Catherine McGinnie[1], Fabián A. Michelangeli[15], Sarah Morris[1], Darin S. Penneys[16], Oscar Alejandro Pérez Escobar[1], Yohan Pillon[17], Lisa Pokorny[1,18], Gustavo Shimizu[19], Vanessa G. Staggemeier[20], Andrew H. Thornhill[4,5], Kyle W. Tomlinson[2,21], Ian M. Turner[13,22], Thais Vasconcelos[23], Peter G. Wilson[24], Alexandre R. Zuntini[1], William J. Baker[1,*], Félix Forest[1,*], and Eve Lucas[1,25,*]

Manuscript received 29 October 2020; revision accepted 29 May 2021.

[1] Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE, UK

[2] Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun, Yunnan 666303, China

[3] University of Chinese Academy of Sciences, Beijing 100049, China

[4] School of Biological Sciences, Faculty of Science, The University of Adelaide, Adelaide, South Australia 5005, Australia

[5] State Herbarium of South Australia, PO Box 1047, Adelaide, South Australia 5001, Australia

[6] Centre National de la Recherche Scientifique (CNRS), Sorbonne University, 4 Place Jussieu, Paris 75005, France

[7] School of Life Sciences, University of Bedfordshire, University Square, Luton LU1 3JU, UK

[8] Faculdade de Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados - UFGD, Dourados, MS, Brazil

[9] Departamento de Botânica, Universidade Federal do Paraná, Curitiba, Paraná, Brazil

[10] Department of Ecology and Evolutionary Biology, University of Michigan, USA

[11] Missouri Botanical Garden, St. Louis, MO 63110, USA

[12] Departamento de Ciências Ambientais, Centro de Ciências e Tecnologias para a Sustentabilidade, Universidade Federal de São Carlos - campus Sorocaba, Sorocaba, SP 18052-780, Brazil

[13] Singapore Botanic Gardens, National Parks Board, 1 Cluny Road, 259569, Singapore

[14] School of Biological Sciences, University of Aberdeen, Aberdeen AB24 3UU, UK

[15] Institute of Systematic Botany, The New York Botanical Garden, Bronx, NY 10458-5126, USA

[16] Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, NC 28403, USA

[17] LSTM, IRD, INRAE, CIRAD, Institut Agro, Univ. Montpellier, Montpellier, France

[18] Centre for Plant Biotechnology and Genomics (CBGP UPM – INIA), Autopista M-40, Km 38, Pozuelo de Alarcón (Madrid) 28223, Spain

[19] Department of Plant Biology, University of Campinas, Campinas, São Paulo 13083-970, Brazil

[20] Departamento de Ecologia, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, RN 59078-970, Brazil

[21] Center of Conservation Biology, Core Botanical Gardens, Chinese Academy of Sciences, Menglun, Mengla, Yunnan 666303, China

[22] Singapore Botanical Liaison Officer, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK

[23] Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA

[24] Royal Botanic Gardens Sydney, Mrs Macquaries Rd, Sydney, NSW 2000, Australia

[25]Authors for correspondence (e-mails: e.lucas@kew.org, o.maurin@kew.org)

*Joint senior authors.

Citation: Maurin, O., A. Anest, S. Bellot, E. Biffin, G. Brewer, T. Charles-Dominique, R. S. Cowan, S. Dodsworth, et al. 2021. A nuclear phylogenomic study of the angiosperm order Myrtales, exploring the potential and limitations of the universal Angiosperms353 probe set. *American Journal of Botany* 108(7): 1087-1111.

doi:10.1002/ajb2.1699

PREMISE: To further advance the understanding of the species-rich, economically and ecologically important angiosperm order Myrtales in the rosid clade, comprising nine families, approximately 400 genera and almost 14,000 species occurring on all continents (except Antarctica), we tested the Angiosperms353 probe kit.

METHODS: We combined high-throughput sequencing and target enrichment with the Angiosperms353 probe kit to evaluate a sample of 485 species across 305 genera (76% of all genera in the order).

RESULTS: Results provide the most comprehensive phylogenetic hypothesis for the order to date. Relationships at all ranks, such as the relationship of the early-diverging families, often reflect previous studies, but gene conflict is evident, and relationships previously found to be uncertain often remain so. Technical considerations for processing HTS data are also discussed.

CONCLUSIONS: High-throughput sequencing and the Angiosperms353 probe kit are powerful tools for phylogenomic analysis, but better understanding of the genetic data available is required to identify genes and gene trees that account for likely incomplete lineage sorting and/or hybridization events.

KEY WORDS Alzateaceae; Combretaceae; Crypteroniaceae; Lythraceae; Melastomataceae; Myrtaceae; Onagraceae; Penaeaceae; phylogenomics; Vochysiaceae.

Targeted sequence capture is a high-throughput sequencing method that uses probes to selectively capture hundreds to thousands of low-copy nuclear orthologs. This approach is desirable as these orthologs are generally expected to have high phylogenetic information content and be sufficiently conserved to be alignable across large phylogenetic distances. Target capture probe sets can be designed for different objectives and can be specific to a narrow taxonomic group (e.g., yams, *Dioscorea* L.; Soto-Gomez et al., 2019) or designed to work universally across larger groups (Johnson et al., 2019), such as entire families Buerki et al., 2021; Clarkson et al., 2021; Pillon et al., 2021; Shah et al., 2021), orders (Antonelli et al., 2021; Zuntini et al., 2021), or even all angiosperms (Baker et al., 2021) as described in this special issue. The cost of developing probe sets remains high, but some off-the-shelf kits already exist and, although the price of implementing such techniques is not negligible, expense-cutting approaches are proliferating (Hale et al., 2020), which, coupled with the significant amount of information generated (in comparison to traditional sequencing techniques), make the use of these techniques cost-efficient.

The order Myrtales Reichenbach is a clade of ca. 399 genera in nine families in the rosid angiosperm clade (APG IV, 2016; POWO, 2020). The nine families in the order are Alzateaceae S.A.Graham, Combretaceae R.Br., Crypteroniaceae A.DC., Lythraceae J.St.-Hil., Melastomataceae Juss., Myrtaceae Juss., Onagraceae Juss., Penaeaceae Sweet ex Guill., Vochysiaceae A.St.-Hil. (Table 1). The latest phylogenetic and biogeographic synthesis of Myrtales (Berger et al., 2016) indicates Combretaceae as sister to a clade composed of all other families, with these forming two further clades, one including Onagraceae + Lythraceae and the other comprising, in root to tip order, Melastomataceae + the CAP clade (Crypteroniaceae-Alzateaceae-Penaeaceae) and Myrtaceae + Vochysiaceae.

Combretaceae are trees, shrubs, lianas, and mangroves (Table 1; POWO 2020). Combretaceae are divided into two subfamilies: Strephonematoideae (genus *Strephonema* Hook.f.) and Combretoideae (9 genera). Combretoideae are divided into two tribes: Laguncularieae Engl. & Diels, comprising four genera of trees, shrubs, and mangroves; and Combreteae DC., which includes the large pantropical tree and shrub genera, *Combretum* Loefl. (also scandent or lianas) and *Terminalia* L. (Table 1). With developments and advances in molecular techniques, several such studies have brought additional understanding to the family (e.g., Tan et al., 2002; Maurin et al., 2010, 2017, 2020; Gere et al., 2013, 2015; Berger et al., 2016). Although the family has been well studied through the years, many aspects remain unclear at all taxonomic levels. Molecular evidence based on a few organellar markers has also raised questions on the monophyly of the Laguncularieae (Maurin et al., 2017) and generic relationships and delimitations within Combretoideae (*Combretum*, *Getonia* Roxb., *Guiera* Adans. ex Juss., *Meiostemon* Exell & Stace). Within the broader Myrtales, there remains uncertainty with regards to the placement of Combretaceae with respect to Lythraceae and Onagraceae (Maurin et al., 2010; Sun et al., 2016; Kriebel et al., 2017).

Lythraceae are small herbs to tall trees with the majority being subshrubs (Table 1; POWO, 2020); the American endemic *Cuphea* P.Browne, with ca. 260 species, is the largest genus. The only monograph of the Lythraceae (Koehne, 1903), recognized 22 genera classified in two tribes, four subtribes, and seven unnamed series. Tribes were based on complete vs. incomplete ovarian septa; however, the degree of septal development varies widely (Tobe et al., 1998; Graham and Graham, 2014). Ten new genera have been described since 1903, and four placed in synonymy (Webb, 1967; Graham, 2010; Graham et al., 2011). Some generic-level taxonomic revisions and phylogenetic studies in the Lythraceae have been published (e.g., Inglis and Cavalcanti, 2018) and other works are ongoing.

Onagraceae are herbs or small shrubs native to open habitats in mountainous and temperate areas, but semi-aquatic plants (e.g., some species of *Ludwigia* L.) and a few trees (e.g., *Xylonagra* Donn. Sm. & Rose) are also represented (Table 1; POWO, 2020). The family is currently classified into two subfamilies, distinguished by the persistence (Ludwigioideae W.L.Wagner & Hoch, genus *Ludwigia*)

TABLE 1. List of Myrtales families. Numbers of genera, species, and geographical distribution.

Family, Author	No. of genera (POWO, 2020)	Nb of species	Distribution	Largest genera (number of accepted species)
Alzateaceae S.A.Graham	1	1	Neotropics	
Combretaceae R.Br.	10	ca. 525	Tropics and subtropics; Africa, Central and South America, southern Asia and northern Australia	*Combretum* Loefl. (ca. 250); *Terminalia* L. (ca. 150)
Crypteroniaceae A.DC.	3	10	South East Asia, Malesia, Sri Lanka	
Lythraceae J.St.-Hil.	28	ca. 625	Old and New World, mostly tropical, but some temperate	*Cuphea* P.Browne (ca. 260)
Melastomataceae Juss.	174	ca. 5700	Very largely tropical, also subtropical, 70% New World	*Blakea* P.Browne (ca. 190); *Medinilla* Gaudich. (ca. 360); *Memecylon* L. (ca. 380) *Miconia* Ruiz & Pav. (ca. 1900); *Microlicia* D.Don: (ca. 170); *Pleroma* D.Don (ca. 160); *Sonerila* Roxburgh (ca. 160)
Myrtaceae Juss.	130	ca. 6079	Worldwide, throughout the tropics and subtropics (mostly tropical-warm temperate), with a paucity of species in Africa and one reaching the Mediterranean	*Eucalyptus* L'Hér. (ca. 750), *Eugenia* P.Micheli ex L. (ca. 1150), *Myrcia* DC. ex Guill. (ca. 750) and *Syzygium* P.Browne ex Gaertn. (ca. 1180)
Onagraceae Juss.	21	ca. 650	Worldwide, from tropical to temperate regions, but particularly diverse in North America	*Epilobium* Dill. ex L. (c. 186); *Fuchsia* Plum. ex L. (c. 108); *Oenothera* L. (c. 155)
Penaeaceae Sweet ex Guill.	3	29	E. and S. Africa, overwhelmingly South African, also St. Helena	
Vochysiaceae A.St.-Hil.	8	ca. 240	Lowland tropical America, from southern Mexico to southern South America, with a large number of species occurring in the Amazon, Cerrado and Brazilian Atlantic Forest apart from *Erismadelphus* Mildbr. and *Korupodendron* Litt and Cheek from W. Africa	

or shedding (Onagroideae Eaton, remaining genera) of sepals in the fruiting stage and by several floral anatomical features (Wagner et al., 2007). Subfamily Onagroideae are further divided into six tribes. Relationships among the seven subunits (including subfamily Ludwigioideae) have been somewhat stable since the first molecular phylogenies that sampled genera from multiple tribes (Conti et al., 1996), and the monophyly of most tribes has been confirmed in subsequent studies focused on the family (Levin et al., 2003, 2004). Generic delimitations have changed considerably over the 2000s, particularly within the species-richest tribes Epilobieae Endlicher and Onagreae Dumortier, and eventually, leading to the re-establishment of tribe Gongylocarpeae J. Donnell Smith & Rose (Levin et al., 2003, 2004; Wagner et al., 2007). Significant nomenclatural rearrangements within Onagreae mainly affected *Oenothera* L. and *Camissonia* Link. The former genus had its circumscription broadened to include several other genera that were found to be nested within it (*Gaura* L., *Calylophus* Spach, and *Stenosiphon* Spach). *Camissonia*, on the other hand was found to be highly paraphyletic, and is now treated as nine genera, including two new names (*Camissoniopsis* W.L.Wagner & Hoch and *Neoholmgrenia* W.L.Wagner & Hoch) (Levin et al., 2003, 2004; Wagner et al., 2007; Wagner and Hoch, 2009). Wagner et al. (2007) provided a detailed synopsis of the family and since then, the suprageneric classification within Onagraceae has remained stable. However, relationships between tribes and genera (especially those within Onagreae) vary slightly, depending on the species sampled and molecular markers used to infer the tree (e.g., Berger et al., 2016; Freyman and Hohna, 2019).

Vochysiaceae are a family of trees (Table 1; POWO, 2020), with several species occurring in the Amazon, Cerrado, and Brazilian Atlantic Forest (Marcano-Berti, 2005; Kawasaki, 2007; Shimizu and Gonçalves, 2017; Flora do Brasil, 2020). The family is subdivided into the tribes Erismeae Dumort. and Vochysieae Dumort. (Litt and Stevenson, 2003; Kawasaki, 2007). The two afrotropical genera *Korupodendron* Litt & Cheek and *Erismadelphus* Mildbr. and the neotropical genus *Erisma* Rudge are included in Erismeae, whereas Vochysieae comprises all remaining neotropical genera. Molecular studies focusing on Vochysiaceae are those of Litt (1999) and Gonçalves et al. (2020). Combined phylogenetic analyses, with plastid *mat*K and morphology (Litt, 1999) and with plastid *ndh*F and ribosomal ITS1 (G. Shimizu et al., unpublished manuscript), support the monophyly of Erismeae, while that of Vochysieae remain uncertain. On the other hand, phylogenies based on all coding plastome genes (Gonçalves et al., 2019, 2020) do not support the monophyly of Erismeae, while Vochysieae appear monophyletic. Further studies are still needed for the remaining nonmonotypic genera, especially *Qualea* Aubl. For example, *Ruizterania* (Stafleu) Marc.-Berti, previously treated as *Qualea* section *Trichanthera* Stafleu, may be indeed embedded in *Qualea*, but more extensive species sampling is required before taxonomic decisions can be made.

Myrtaceae are a family of woody species of trees, treelets, shrubs, and very occasionally lianas (*Metrosideros* Banks ex Gaertn.) (Table 1; Wilson, 2005, 2010; Heywood et al., 2007; POWO, 2020). Myrtaceae comprise two subfamilies, Heteropyxidoideae, with two monogeneric tribes, and Myrtoideae, comprising the remaining 15 tribes (Wilson et al., 2005). Myrtaceae have four large genera (>500 spp.; Frodin, 2004): *Eucalyptus* L'Hér., *Eugenia* L., *Myrcia* DC. ex Guill., and *Syzygium* P.Browne ex Gaertn. The species of these genera are often morphologically extremely homogeneous and thus difficult to manage in herbaria or in the biomes in which they occur (e.g., Lucas and Bünger, 2015; Cámara-Leret et al., 2020). Three of the four large Myrtaceae genera, *Eugenia*, *Myrcia*, and *Syzygium* have a few large seeds and fleshy fruits and fall into two exclusively fleshy-fruited tribes Myrteae DC. (ca. 50 genera) and Syzygieae Peter

G.Wilson (one genus) that comprise about half of the species of the family. With few exceptions, the remaining tribes of Myrtaceae have dry fruits with many, wind-dispersed seeds.

Molecular analyses have consistently recovered a clade comprising the families Alzateaceae, Crypteroniaceae, and Penaeaceae (CAP clade; van Beusekom-Osinga and van Beusekom, 1975; Conti et al., 1994, 1997, 2002; Renner et al., 2001; Schönenberger and Conti, 2003; Sytsma et al., 2004), as sister to the Melastomataceae. These three families of shrubs and small trees (Table 1; POWO, 2020) have affinity to Melastomataceae based on morphological and anatomical analyses (Johnson and Briggs, 1984; Tobe and Raven, 1983, 1984a–c, 1987), but relationships within remain unclear. Asian Crypteroniaceae have been recovered as sister to a clade that contains Alzateaceae, Rhynchocalycaceae L.A.S.Johnson & B.G.Briggs, Oliniaceae Harv. & Sond. and Penaeaceae (Schönenberger and Conti, 2003; Rutschmann et al., 2004; Sytsma et al., 2004; Rutschmann et al., 2007). While prior to 2002 the latter three families were all recognized, APG III (2009) chose to recognize an expanded Penaeaeceae that includes monotypic Rhynchocalycaceae and monogeneric Oliniaceae.

Melastomataceae are a predominantly tropical family of mostly shrubs or small trees, but they can also be herbs, root climbers, epiphytes, hemi-epiphytes, and large trees (Table 1; POWO, 2020). The monophyly of the Melastomataceae has rarely been challenged from a morphological standpoint and has been corroborated by molecular data (Berger et al., 2016; Clausing and Renner, 2001). The main issue has been whether to recognize the Memecylaceae as a separate family or as a subfamily, the Olisbeoideae. The last complete monograph of the family was that of Cogniaux (1891). Based on morphological and anatomical data, Renner (1993) reorganized the tribal taxonomy, but many of the groups proposed by Cogniaux (1891) and Renner (1993) have been shown to be nonmonophyletic, while new clades, not initially suggested by morphology, have been recovered with molecular data. In consequence, several new tribes have been proposed in the last decade (Clausing and Renner, 2001; Michelangeli et al., 2004, 2011, 2013; Penneys et al., 2010, 2020; Goldenberg et al., 2012; Penneys and Judd, 2013; Rocha et al., 2016, 2018; Bacci et al., 2019; Bochorny et al., 2019). While much progress has been made on the phylogeny of New World groups (previous references) and African Olisbeoideae (Stone, 2006, 2014) and Melastomateae Bartl. (Veranso-Libalah et al., 2017, 2018, 2020), much work is still needed on the Asian and Madagascan Melastomataceae. Most of these phylogenetic analyses have been based on analyses of seven or fewer nuclear and plastid loci. Two phylogenetic trees are available based on complete plastomes, but, in one of them, the taxonomic sampling is limited (Reginato et al., 2016), and in the other, sampling is restricted to just one tribe (Zhou et al., 2019). In consequence, even though the composition of most clades is now coming into focus, relationships among tribes are still weakly supported. One large clade within the family that has been consistently recovered includes the tribes Melastomateae, Marcetieae M.J.R.Rocha, P.J.F.Guim. & Michelang., Microlicieae Naudin and Rhexieae DC. (Clausing and Renner, 2001; Michelangeli et al., 2013; Rocha et al., 2016), mostly composed of plants with capsular fruits and stamens with a sterile section at the base of the anther dubbed the "pedoconnective clade" (Michelangeli et al., 2013).

Here, we present a novel phylogenomic study of the Myrtales with near-complete (ca. 76%) generic sampling, based on target capture data generated with the universal Angiosperms353 nuclear probe set. We examine the recovery of plastome data from off-target reads. We apply the phylogenomic framework established to address the following questions: (1) Can exon data from the 353 nuclear target loci improve resolution of relationships among families (i.e., the backbone)? (2) Can recalcitrant relationships among genera be resolved? (3) Can the addition of intron data captured from flanking regions shed new light on interspecies relationships within highly diverse genera such as *Combretum* (Combretaceae), *Miconia* Ruiz & Pav. and *Tibouchina* Aubl. (Melastomataceae), and *Eucalyptus*, *Eugenia*, *Myrcia*, and *Syzygium* (Myrtaceae)?

MATERIALS AND METHODS

Sampling

One specimen was selected per genus from a list of genera standardized according to the online World Checklist of Selected Plant Families (Govaerts et al., 2020) and the Plants of the World Online (POWO, 2020), supplemented with recommendations by taxonomic specialists. In total, 485 species in 305 genera were sampled, representing 76% of all genera in the order (Appendix S1). Two outgroup species from families Francoaceae A.Juss. and Geraniaceae Juss. were included (Appendix S1). For species-rich genera (Table 1), additional representatives were included based on previously published classifications. We included a particularly comprehensive sample of *Combretum* subgenus *Combretum* Loefl. section *Ciliatipetala* Engler & Diels to evaluate the potential of the Angiosperms353 probe set in discriminating closely related species. Selection of this last subsample was based on current taxonomic understanding and tissue availability.

Selection of plant tissue was based on the best available quality, considering that an ideal sample extraction should contain 100–200 ng of DNA with fragment sizes ≥ 350 bp as recommended in the library preparation protocol (New England BioLabs, Ipswich, MA, USA). Material was selected from, in order of preference, (1) leaf tissue dried in silica gel from field collections or the living collection at the Royal Botanic Gardens, Kew (RBGK), which generally gave optimal results; (2) DNA aliquots from the DNA and Tissue Bank at RBGK; and (3) plant tissue material selected from herbarium specimens, which yielded DNA of variable quality.

DNA extraction, library preparation, hybridization, and sequencing

DNA extractions were performed using a modified CTAB protocol (Doyle and Doyle, 1987) and purified using Mag-Bind TotalPure NGS magnetic beads (Omega Bio-tek, Norcross, GA, USA). Purified DNA extracts were run in a 1.5× agarose gel to assess the average fragment size and then quantified using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA extracts with fragment sizes >350 bp were sonicated using a M220 Focused-ultrasonicator with microTUBEs AFA Fiber Pre-Slit Snap-Cap (Covaris, Woburn, MA, USA), following the manufacturer's protocol for ~350-bp insert sizes. Dual-indexed libraries for Illumina sequencing were prepared using the DNA NEBNext UltraTM II Library Prep Kit at half the recommended volume, with Dual Index Primers Set 1, NEBNext Multiplex Oligos for Illumina (New England BioLabs). Quality of libraries was evaluated on a 4200 TapeStation System using High Sensitivity D1000 ScreenTape (Agilent Technologies, Santa Clara, CA, USA).

Subsequently, libraries were quantified using the Qubit fluorometer. Equimolar pools comprising 20–25 DNA libraries for a total of 1 μg of DNA were hybridized using the myBaits Expert Predesigned Panel (Arbor Biosciences, Ann Arbor, MI, USA) Angiosperms353 v1 (Catalog #308196; Johnson et al., 2019) following the manufacturer's protocol with v4 chemistry (http://www.arborbiosci.com/mybaits-manual). Hybridizations were performed at 65°C for 28–32 h in a Hybex Microsample Incubator (SciGene, Sunnyvale, CA, USA), using an equal volume of red Chill-out Liquid Wax (Bio-Rad, Hercules, CA, USA) to prevent evaporation. Enriched products were amplified with KAPA HiFi (2×) HotStart ReadyMix PCR Kit (Roche, Basel, Switzerland) for 10 cycles. PCR products were then cleaned using the QIAquick PCR purification kit (Qiagen, Germantown, MD, USA). Products were quantified with the Qubit fluorometer and in some cases re-amplified a second time for 3–8 cycles. Final products were run on a 4200 TapeStation System using High Sensitivity D1000 ScreenTape (Agilent Technologies) to assess quality and average fragment size. Several library pools were multiplexed and sequencing was performed at RBGK on an Illumina MiSeq (Illumina, San Diego, CA, USA) with v3 reagent chemistry (2 × 300-bp paired-end reads) or on an on Illumina HiSeq (2 × 150-bp paired-end reads) at Genewiz (Takeley, UK) or at Macrogen (Geumcheon, Republic of Korea).

Sequence data processing

Sequencing output reads (FASTQ files) were trimmed using Trimmomatic (Bolger et al., 2014) to remove base pairs at the beginning of the reads as long as their Phred quality score remained below 30 and to trim the end of the reads once encountering a 4-bp window with quality below 30 (Trimmomatic settings LEADING: 30; TRAILING: 30; and SLIDING WINDOW:4:30). Trimmed reads shorter than 36 bp were then removed to decrease the risk of short reads that might not be uniquely positioned over sequences (MINLEN: 36).

Nuclear recovery—Paired reads and combined unpaired reads were used to recover target sequences using HybPiper v1.3.1 (Johnson et al., 2016) using the target file available at https://github.com/mossmatters/Angiosperms353 containing de-gapped medoid sequences (Johnson et al., 2019). HybPiper was run using the BLASTx option (Camacho et al., 2009) since it has been found to produce longer contig sequences (Murphy et al., 2020). Reads were mapped to each of these target sequences (with the abovementioned BLASTx algorithm), were then assembled de novo using SPAdes (Bankevich et al., 2012), and coding sequences (hereafter called exons) were extracted using exonerate (Slater and Birney, 2005). Noncoding sequences (i.e., introns and untranslated regions [UTRs]) flanking the exons were recovered and combined to the exon sequences into so-called supercontigs using the script intronerate.py available with HybPiper. Sequences from all taxa were combined for each exon and supercontig, and the resulting single locus matrices were each aligned separately using MAFFT v7 (mafft-7.419-gcc_fc6.x86; Katoh and Standley, 2013), with accuracy-oriented methods (--localpair; --maxiterate 1000) and the option to generate reverse complement sequences to align them with the remaining sequences based on 6-mer counting (--adjustdirectionaccurately). Single-locus alignments were subsequently trimmed using phyutility (https://github.com/blackrim/phyutility) to remove nucleotide sites missing in at least 80% of the taxa (-clean 0.8).

Plastome recovery—To evaluate the success of Hyb-Seq (Weitemier et al., 2014) in our analyses, we performed a plastome recovery on the entire sample using the HybPiper pipeline with a target file of coding sequences (genes, rRNA, tRNA, etc.) and noncoding sequences (intergenic regions; available at Zenodo [https://doi.org/10.5281/zenodo.4268317]) generated from plastomes available on GenBank for *Eugenia uniflora* O.Berg (Myrtaceae; NC_027744), *Lagerstroemia subcostata* var. *fauriei* (Koehne) Hatus. ex Yahara (Lythraceae; NC_02980), *Oenothera argillicola* Mack. (Onagraceae; NC_010358), *Penaea sarcocolla* L. (Penaeaceae; MK726025), *Tibouchina urvilleana* (DC.) Cogn (Melastomataceae; MK726030), and *Vochysia acuminata* Bong (Vochysiaceae; MK726031). This step allowed evaluation of the overall recovery of the plastome across the study samples, and between families, type of material used (herbarium versus silica dried), sample DNA concentration, library length (in bp), sequencing success (read number), and enrichment success (number of reads on target nuclear exons + flanking introns). The comparison was made by fitting linear models including additive effects of some or all these predictors on plastome length recovered (in bp), trying all possible combinations of predictors. The function regsubsets of the R package leaps (scripts from Thomas Lumley based on Fortran code by Alan Miller [2020]) was used to select the best model among models with a given number of predictors, and the model with the smallest corrected AIC (estimated using the function AICc of the R package AICcmodavg; Mazerolle, 2019) was then chosen among these best models. Analyses were performed in R Studio (R Core Team, 2020; RStudio Team, 2020).

Recovery statistics—Recovery statistics and heatmaps were generated using scripts from HybPiper v1.3 (Johnson et al., 2016). Additional statistics were generated using HybPiper_stats_general.sh (scripts available at Zenodo, https://doi.org/10.5281/zenodo.4268317), which allowed evaluation of exon and (flanking partial) intron coverage. For this, we combined all reads found by HybPiper to map the reference target files, then mapped them against the recovered sample gene sequences using the BWA algorithm (Li and Durbin, 2010). Resulting SAM files were parsed using a custom python script to keep only reads mapping with less than three mismatches and a score >30 to produce conservative coverage estimates. The filtered SAM files were analyzed with SAMtools (Li et al., 2009) mpileup function to produce coverage information per base pair, and outputs were parsed with custom Python scripts to calculate intron and exon average coverage for each gene of each sample. Intron–exon boundaries were obtained from the genomic feature format (GFF) annotation files produced by HybPiper.

Phylogenomic inference workflow

Tree reconstruction—Species trees were first inferred by analyzing matrices made of multiple single locus alignments concatenated using FASconCAT-G v1.0 (Kück and Meusemann, 2010). Analyses were conducted on two data matrices: one matrix made of all the exons, and one matrix made of all the supercontigs (i.e., exons and their flanking regions). For each matrix, a species tree was generated using RAxML (Stamatakis, 2014). This approach is hereafter referred to as the maximum likelihood (ML) approach.

Species trees were also inferred using a multispecies-coalescent approach, based on the analysis of single-locus trees (hereafter, gene trees). Gene trees were generated from each trimmed locus

alignment using IQ-TREE v2.0 (Minh et al., 2020), with ultrafast bootstrap (1000 replicates, UFBoot2; Chernomor et al., 2016) and model selection (-m MFP). Branches with support values below 10% (Mirarab, 2019 [Preprint]) were collapsed in each gene tree using Newick Utilities v1.6 (Junier and Zdobnov, 2010). A first coalescent analysis (Analysis V1) was performed with this set of gene trees using ASTRAL-III (Mirarab and Warnow, 2015) with extensive branch annotations (-t 2 flag). These annotations allowed recovery of both normalized quartet score (QS) values and local posterior probabilities (LPP). The set of gene trees used in Analysis V1 was evaluated using TreeShrink (Mai and Mirarab, 2018) to identify outlier taxa that increased the diameter of each gene tree (i.e., the maximum distance between any two tips of the tree) by more than 20%, using centroid re-rooting (-b 20 -c). Each locus alignment was then cleaned of the outlier taxa, realigned, trimmed, and analyzed using IQ-TREE as described above. A second coalescent analysis was conducted on this new set of gene trees, using the same method as for analysis V1 (Analysis V2). Finally, a third set of gene trees was made by excluding genes with <25% of the taxonomic sample from the second set of gene trees. A third coalescent analysis was performed on this new set of gene trees, using the same method as for analyses V1 and V2 (Analysis V3). Analyses V1, V2 and V3 were performed separately on exons and on supercontigs. The number of parsimony informative sites (PIS) was calculated using AMAS (Borowiec, 2016), before and after trimming, for each alignment. R (R Core Team, 2020) packages ape (Paradis and Schliep, 2019), ggimage (Yu, 2019), ggtree (Yu et al., 2017), and treeio (Wang et al., 2020) were used to plot the trees.

Polytomy test—A quartet-based polytomy test was conducted using ASTRAL-III as described by Sayyari and Mirarab (2018). This test was used on the exons-only ASTRAL analysis to evaluate gene tree discordance in the data set and identify potential hard polytomies.

Model violation and tree landscape assessment—To explore the influence of potential incorrect nucleotide substitution models that might impact the accuracy of phylogenetic reconstructions, model assumption violations were tested. The test was implemented in IQ-TREE v2.0 (Naser-Khdour et al., 2019). Tests of two less-constraining conditions than the defaults parameters rejected 38 and 52 exons from the data set with p-value of 0.0001 and 0.001, respectively. The default parameters (p-value: 0.05) excluded 106 loci. The three resulting data sets with the remaining loci (Analyses V4, V5, and V6, respectively) were used to assess the impact of model violation on the topology of Myrtales. The phylogenetic landscape was investigated using the R package treespace (Jombart et al., 2017). The statistical distribution of the tree topologies was calculated using principal component analysis with the Kendall–Colijn (KC) distance for rooted trees (Kendall and Colijn, 2016). The trees were rooted using the function reroot of the package phytools v.0.6-99 (Revell, 2012) on *Hypseocharis bilobata* Killip (Geraniaceae).

RESULTS

We used the results to address performance of the Angiosperms353 kit in Myrtales, on one hand, and inferred relationships, on the other. Inferred relationships are presented based on the species tree obtained from the multispecies-coalescent ASTRAL analysis V3 from exon-only gene trees (see Materials and Methods), with quartet score (QS; for the main, and the first and second alternative topologies) and LPP values displayed. This ASTRAL tree is then compared to the ML analysis of concatenated exons, with bootstrap support (BS) values displayed at the nodes (Fig. 1). Topologies resulting from other analyses or with alternative support metrics are presented in Fig. 2 (QS), Fig. 3A–C (QS and LPP), and Appendix S2 (BS) and Appendix S3 (QS, LPP). Appendix S4 supplements these values, for each branch, with the ASTRAL polytomy test results. The ASTRAL topologies derived from exons (coding regions) vs. supercontig (exons+flanking introns) datasets are also compared (Fig. 4). References to numbered clades throughout refer to those shown in Figs. 2 and 3A–C. All alignments, and Newick tree files can be found at Zenodo (https://doi.org/10.5281/zenodo.4268317).

Angiosperms353 performance in Myrtales

Myrtales sequences amounted to ~228 GB of data, resulting in ~1.4 billion reads. When mapping reads on the target nuclear genes, the median number of reads per sample reached 1,898,354 with a median of reads on target of 3.5%. At the order level, the mean number of loci retrieved was 295, representing 83.6% of loci with sequences that extend >50% of the target length and with a median read depth of 28. The median of reads on target varied between families from 2.1% in Combretaceae to 9.9% in Onagraceae, though the number of loci retrieved did not correlate with reads on target. The lowest number of loci retrieved was 260 in the CAP clade, and the maximum was 329 for Onagraceae. Statistics are summarized in Appendix S5 and presented fully in Appendix S6. Recovery for the nuclear and plastid genes is respectively illustrated in Appendices S7 and S8. Enrichment success for the order is presented in Appendix S9 and detailed for each family in Fig. 5, providing recovery statistics for the exons and introns.

The best model predicting plastome length recovery was the one including all the predictors (difference in AICc with second best model = 51, $p < 0.001$, adjusted $R^2 = 0.2506$). In this model, log-transformed DNA library length, number of reads sequenced, and number of reads on nuclear exon targets and flanking introns all had positive significant effects on log-transformed plastome length recovery ($p < 0.05$, < 0.001, and < 0.001, respectively), while the effect of other predictors was not significant ($p > 0.05$). When mapping reads on the target plastome genes, the median of reads on target was of 1%. At the order level, the mean number of loci retrieved was 29, representing 12% of loci with sequences that extend >50% of the target length. The median of reads on target varied between families from 0.9% in Lythraceae and Vochysiaceae to 1.8% in Onagraceae. The lowest number of loci retrieved was 15 in the Combretaceae, and the maximum was 59 for the CAP clade. Statistics are summarized in Appendix S5 and presented fully in Appendix S6. Due to low and unequal recovery across the sample used here, the plastome data were not used to produce a phylogenetic tree.

Final nuclear gene alignments reached respectively 119,990 for the exons and 162,909 base pairs for the supercontigs. Missing data was respectively 7.45% and 8.17% and the percentage of parsimony informative sites (PIS) was 0.67% (79,036 bp) and 0.77% (125,384 bp), respectively. Complete statistics for combined exons and supercontigs alignments as well as for each single locus alignment are provided in Appendix S10. For the molecular model violation testing, 38, 53, and 106 loci (with p-values of 0.0001, 0.001, and 0.05, respectively), were excluded from the V3 gene data set. Axis 1 of

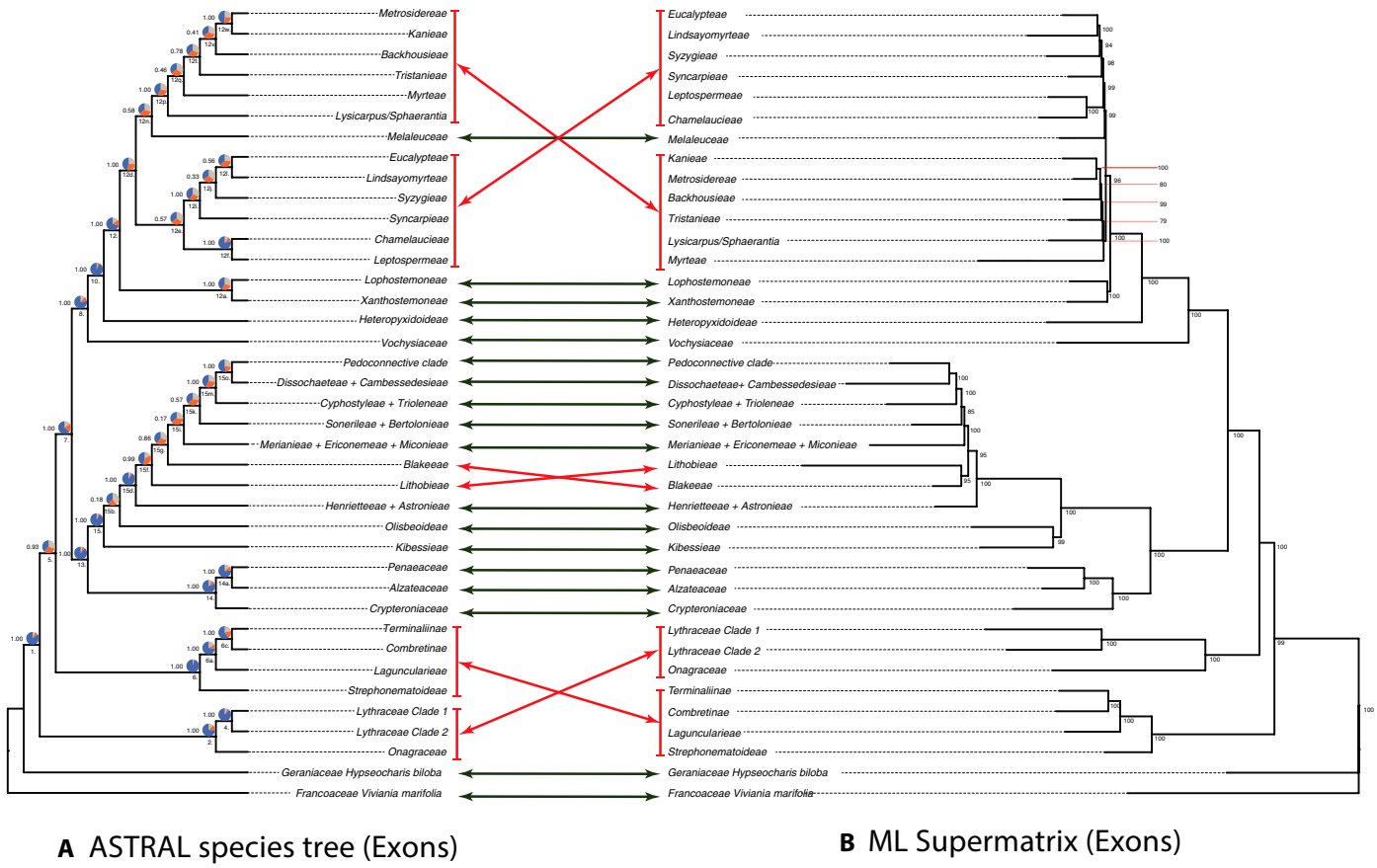

FIGURE 1. Exon coalescent ASTRAL (A) versus exon RAxML supermatrix (B) topology comparison in major families and subclades. Annotations in (part A): Pies charts above branches display quartet score (QS) values for each node (blue = species tree topology QS; orange = first alternative topology QS; gray = second alternative topology QS). Local posterior probability values are presented next to the pies chart. Values next to branches in 1B are bootstrap support percentages (BP).

the PCA using the KC distance (Appendix S11) explained 95% of the variation in the tree topology, while axis 2 explained 4.1%. The topology without modification (V1) clustered with the topologies obtained after excluding the genes that did not pass the model violation tests (V5 and V6). This result indicates that model violation has little impact on topology and supports the argument that the V3 data set is appropriate for downstream analysis and taxonomic discussion.

Inferred relationships

Support values—Phylogenetic results presented follow the phylogenomic reconstructions presented in Figs. 2 and 3 (summarized in Appendix S4). Quartet score values are interpreted as follows: if Q1 ≥ 0.75, congruence is high; if 75 > Q1 ≥ 50, congruence is moderate; and if Q1 < 50, congruence is low. When interpreting LPP values: if LPP = 1.0, support is full; if 1.0 < LPP > 0.7, support is high; if 0.5 < LPP < 0.7, support is moderate; and if LPP < 0.5, support is low. When interpreting BP values: if BS = 100, support is strong (full); if BS ≥ 95, support is moderate; and if BS < 95, support is weak (low). The polytomy test of Sayyari and Mirarab (2018) evaluates the null hypothesis that a given branch is a polytomy. The test's ability to reject the null hypothesis is influenced by the length of a given branch and the number of genes available; the shorter the branch, the more genes are required for the null hypothesis to be rejected. This pattern explains why in some instances the null hypothesis is maintained on short branches although they are supported with high LPP in the ASTRAL analysis.

Myrtales backbone—Results for relationships of the Myrtales backbone can be seen in Figs. 2, 3A–C (ASTRAL), Appendix S2 (ML). Highly congruent gene trees corroborate the well-supported monophyly of Myrtales (clade 1). Relationships between families Combretaceae, Lythraceae, and Onagraceae conflict between analytical approaches, probably due to the high incongruence between gene trees, as highlighted by the quartet score values in the ASTRAL tree. The ASTRAL topology suggests the Lythraceae-Onagraceae (clade 2) is sister to the remaining families within the order, whereas in the ML analysis Combretaceae are strongly supported as sister to all other Myrtales. The sister relationship between Lythraceae and Onagraceae is moderately congruent in the ASTRAL (QS) and is fully supported in both the ASTRAL (LPP) and ML (BS) trees. Combretaceae are well supported as monophyletic in all analyses (ASTRAL, ML; clade 6). Relationships between these families and the remaining families within the order are moderately to strongly supported in all analyses, although the QS indicates that just 63% of informative genes support this topology (clade 7). A summary comparing key topologies resulting

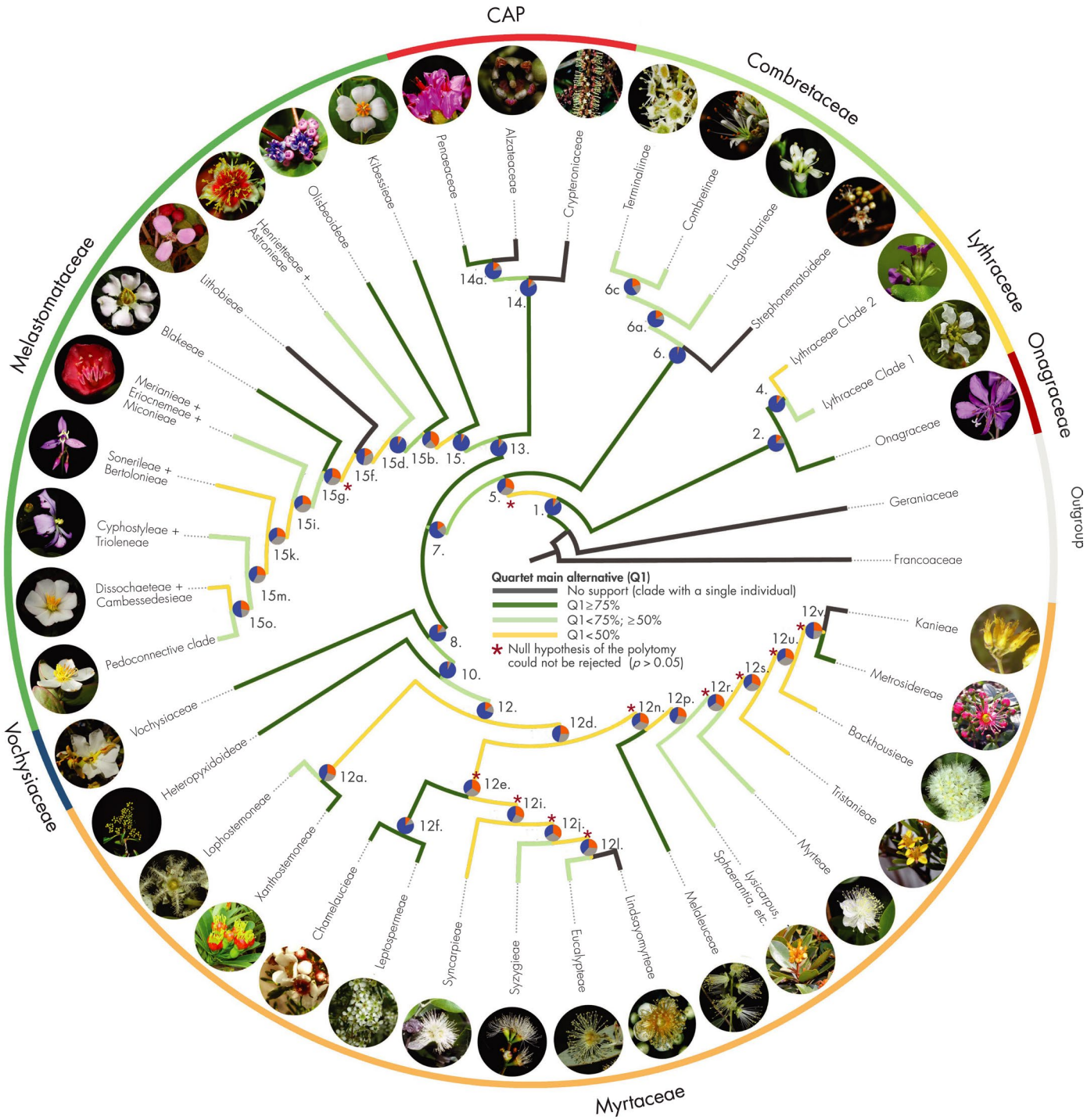

FIGURE 2. Coalescent astral tree summary showing relationships in key Myrtales families and subclades, using exons, inferred in ASTRAL. Pies charts on branches display quartet score (QS) values as described for Fig. 1A. Branches and clade color code follow the quartet main topology (Q1) support value, with dark green ≥0.75; light green ≥50, <75; and yellow <50. Red asterisks next to pie charts indicate the null hypothesis of a polytomy (equal support for species tree, first alternative, and second alternative topology) could not be rejected with the polytomy test ($p > 0.05$). Numbers on branches relate to clades reported in Appendix S4. *Photo credits:* Counter clockwise from Onagraceae: OFS, WMC, ST, DH, OM (×3), JD, PM, OM, DP (×3), FM, DP, FM, DP, FM (×3), GS, OM, OFS, YP, PW, FF, CSIRO, EL, OFS, CSIRO, RBGK, TV, FF (×2), DD, EL, WT. Commonwealth Scientific and Industrial Research Organisation (CSIRO); D. Dixon (DD); J. Dransfield (JD); F. Forest (FF); O. Fragman-Sapir (OFS); D. Harris (DH); P. Maas (PM); O. Maurin (OM); W. McCleland (WMC); E. Lucas (EL); F. Michelangeli (FM); D. Penneys (DP); Y. Pillon (YP); Royal Botanic Gardens, Kew (RBGK); G. Shimizu (GS); W. Takeuchi (WT); S. Turner (ST); T. Vasconcelos (TV) and P. Wilson (PW).

from previous studies to results from the different analyses presented here is available in Fig. 6A.

A well-supported sister relationship (clade 8) is reported for Vochysiaceae and Myrtaceae (clades 9 and 10, respectively), each of them strongly supported as monophyletic. In Myrtaceae, Heteropyxidoideae, including the monotypic genera *Psiloxylon* Thouars ex Tul. and *Heteropyxis* Harv., are highly supported (clade 11) and sister to a clade comprising all the remaining Myrtaceae genera (subfamily Myrtoideae), also with high support (clade 12). Finally, the clade comprising the Crypteroniaceae, Alzateaceae, and Penaeaceae (CAP) clade and Melastomataceae, is well supported, with each family also receiving high support as monophyletic (clades 13, 14, and 15).

Some improvement in support values was observed in the supercontig tree (Fig. 4; Appendix S3) vs. the exon only tree. Within the two largest families, Melastomataceae and Myrtaceae, the inclusion of introns seems to provide a more robust backbone at the infrafamilial level (with both LPP and QS values increasing); however, relationships between closely related subclades tend not to improve and gene incongruence remains strong, as highlighted by the QS.

Onagraceae—Results for relationships of the Onagraceae can be seen in Figs. 2, 3A (ASTRAL), and Appendix S2 (ML). All analyses support *Ludwigia octovalvis* (Jacq.) P.H.Raven (representing the monogeneric subfamily Ludwigioideae as sister to Onagroideae (represented by all remaining accessions, clade 3a). Sixteen of the 21 genera in Onagroideae are here sampled, encompassing all tribes except for monogeneric Hauyeae Raim. The monophyly of the subfamily is fully supported in all

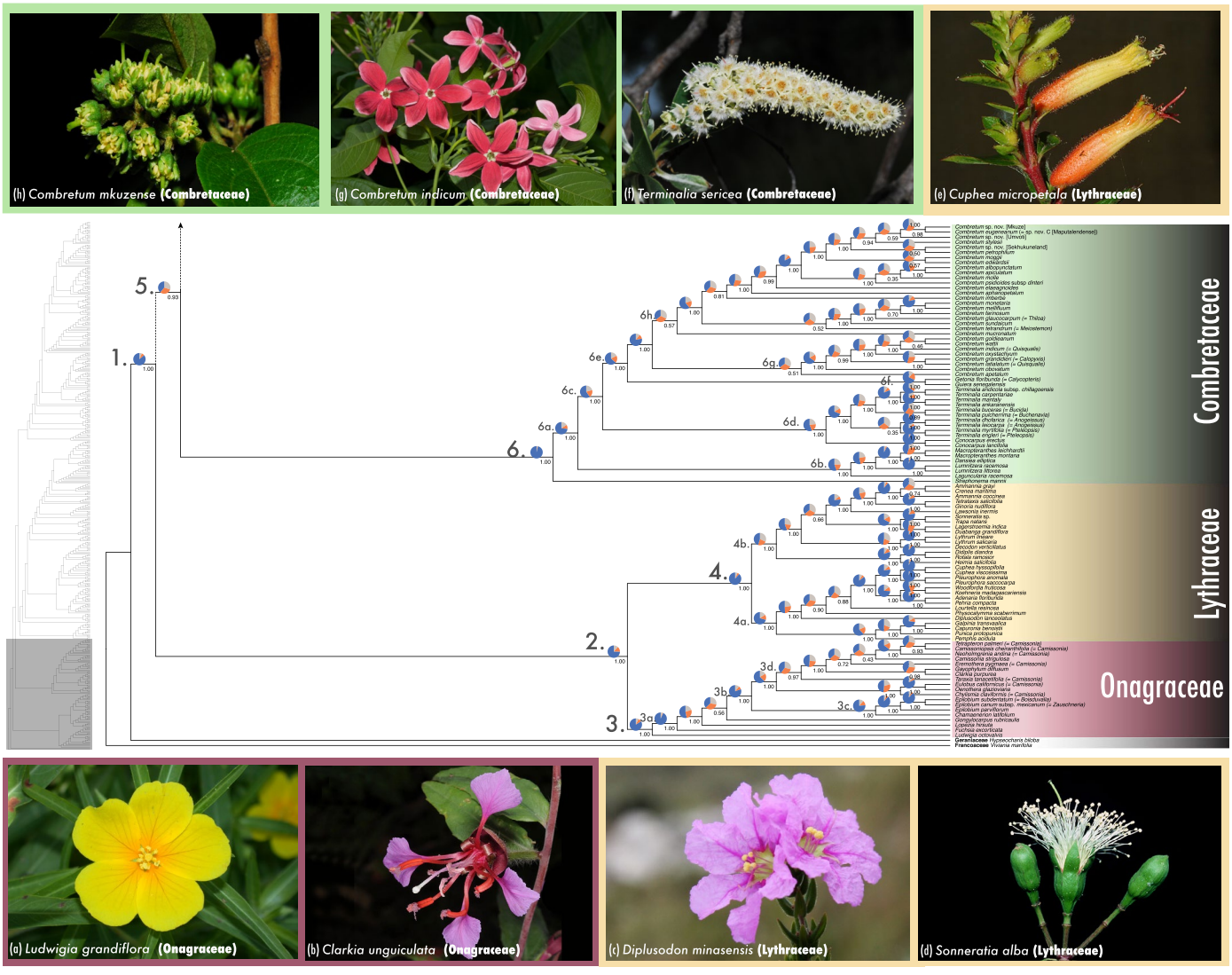

FIGURE 3A.

FIGURE 3. Coalescent tree using exons showing relationships in Myrtales inferred using ASTRAL. (A) Combretaceae, Lythraceae, Onagraceae; (B) Myrtaceae and Vochysiaceae; (C) Alzateaceae, Crypteroniaceae, Melastomataceae, Penaeaceae. Annotations above as pies chart above branches are as described for Fig. 1A and local posterior probability values are presented below branches. Numbers on branches relate to clades reported in Appendix S4 (Parts A, B, and C of this figure appear on separate pages). *Photo credits:* from bottom and counterclockwise (initials as in Fig. 2 legend); (A): P. Hoch.; OFS; W. Milliken; D. Goldman; OFS; OM (x3). (B): GS (×2); OFS (x2); EL; OM; CSIRO; EL. (C): PM; OM; FM; DP; FM; DP; FM; DP; FM.

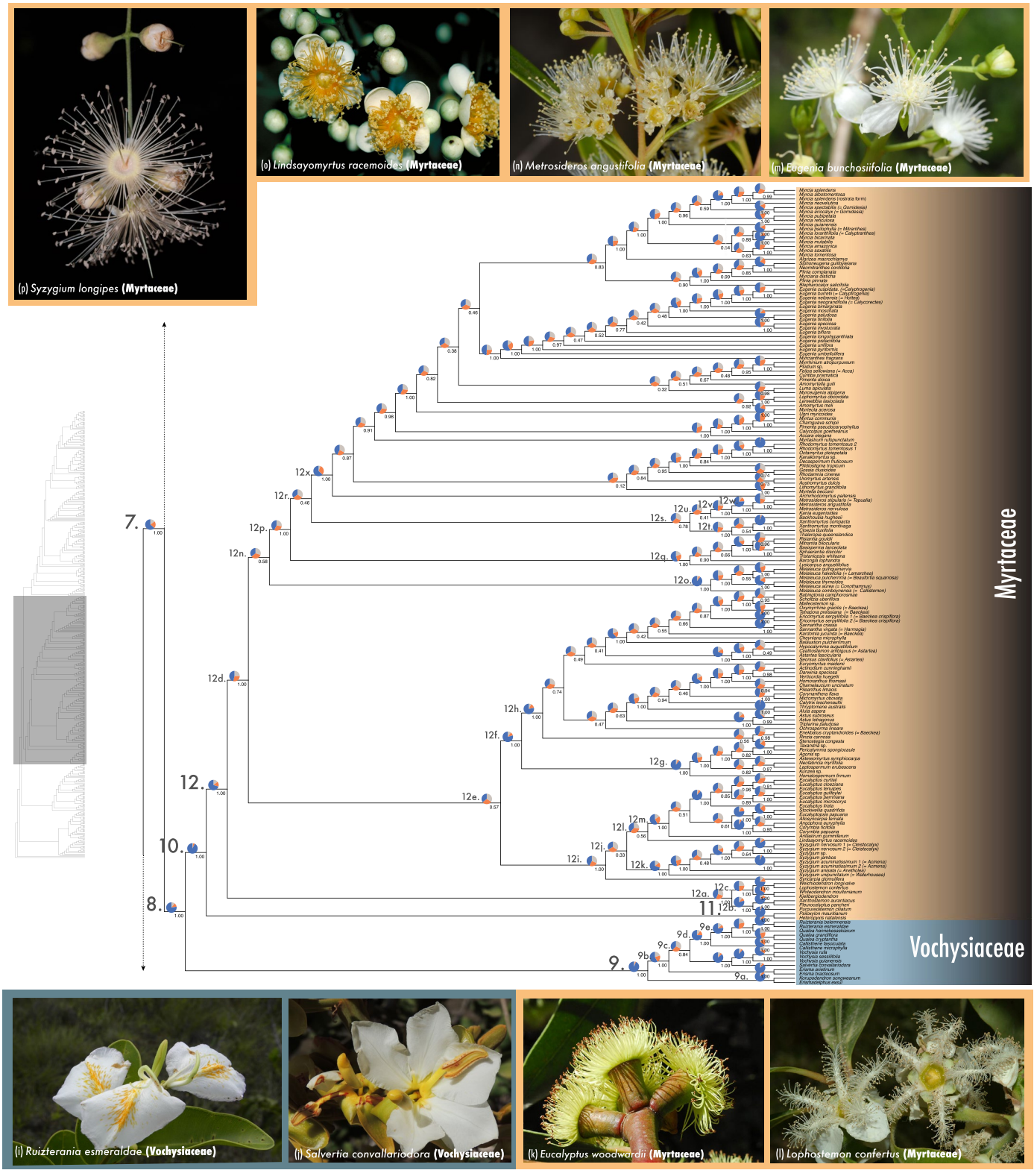

FIGURE 3B.

analyses. The position of tribe Circaeeae Dumort. (represented by *Fuchsia excorticata* L.f.), as sister to the remaining tribes in Onagroideae, is fully supported in both ASTRAL and ML analyses. In root to tip order, tribes Lopezieae Spach (represented by *Lopezia hirsuta* Jacq.), Gongylocarpeae (*Gongylocarpus* Schltdl. & Cham.), Epilobieae, and Onagreae are recovered as a grade in all analyses, but the position of Gongylocarpeae as sister to the fully supported clade formed by tribes Epilobieae and Onagreae

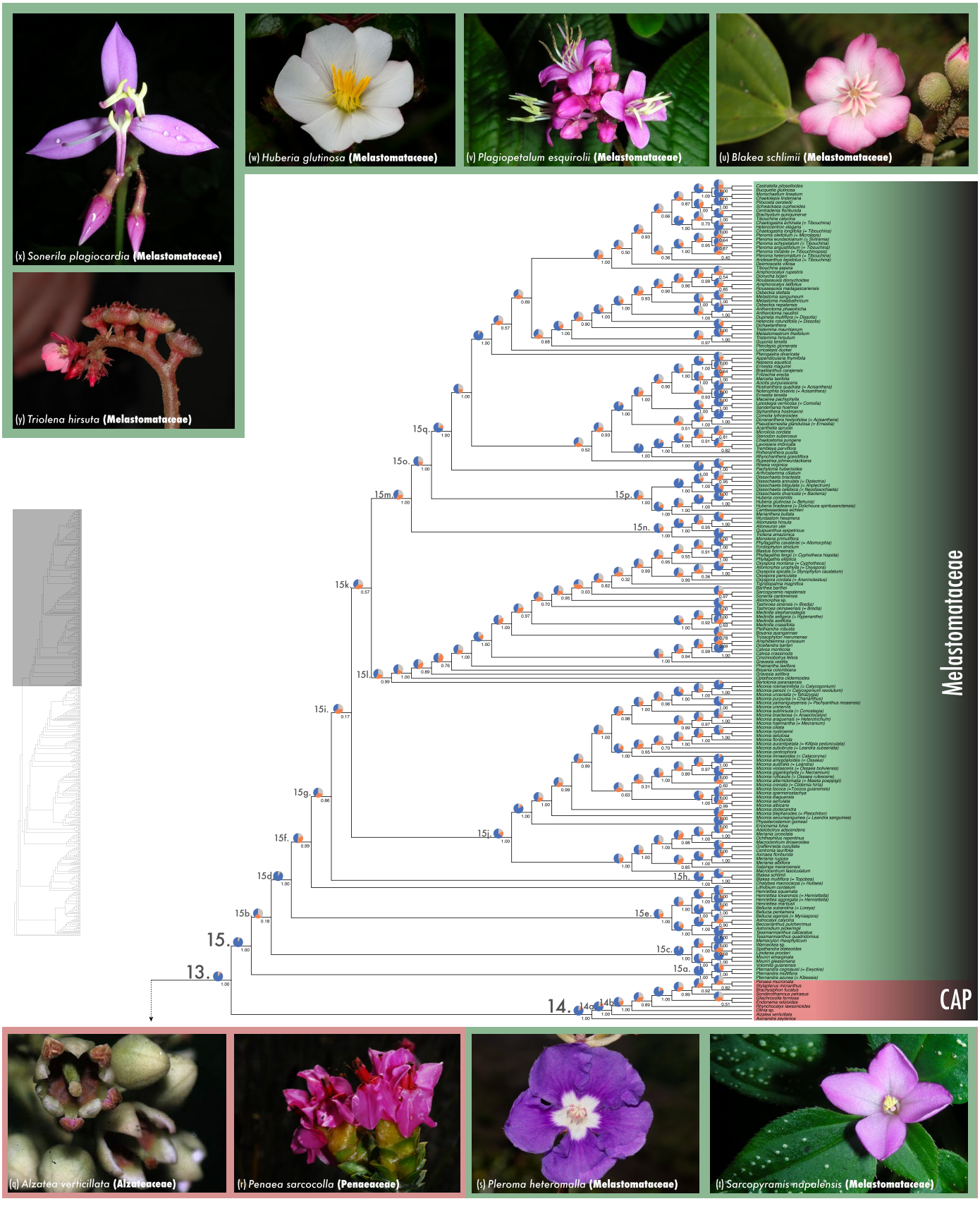
(x) Sonerila plagiocardia (Melastomataceae)
(w) Huberia glutinosa (Melastomataceae)
(v) Plagiopetalum esquirolii (Melastomataceae)
(u) Blakea schlimii (Melastomataceae)
(y) Triolena hirsuta (Melastomataceae)
Melastomataceae
CAP
15.
13.
14.
15a.
15b.
15c.
15d.
15e.
15f.
15g.
15h.
15i.
15j.
15k.
15l.
15m.
15n.
15o.
15p.
15q.
14a.
14b.
(q) Alzatea verticillata (Alzateaceae)
(r) Penaea sarcocolla (Penaeaceae)
(s) Pleroma heteromalla (Melastomataceae)
(t) Sarcopyramis napalensis (Melastomataceae)

FIGURE 3C.

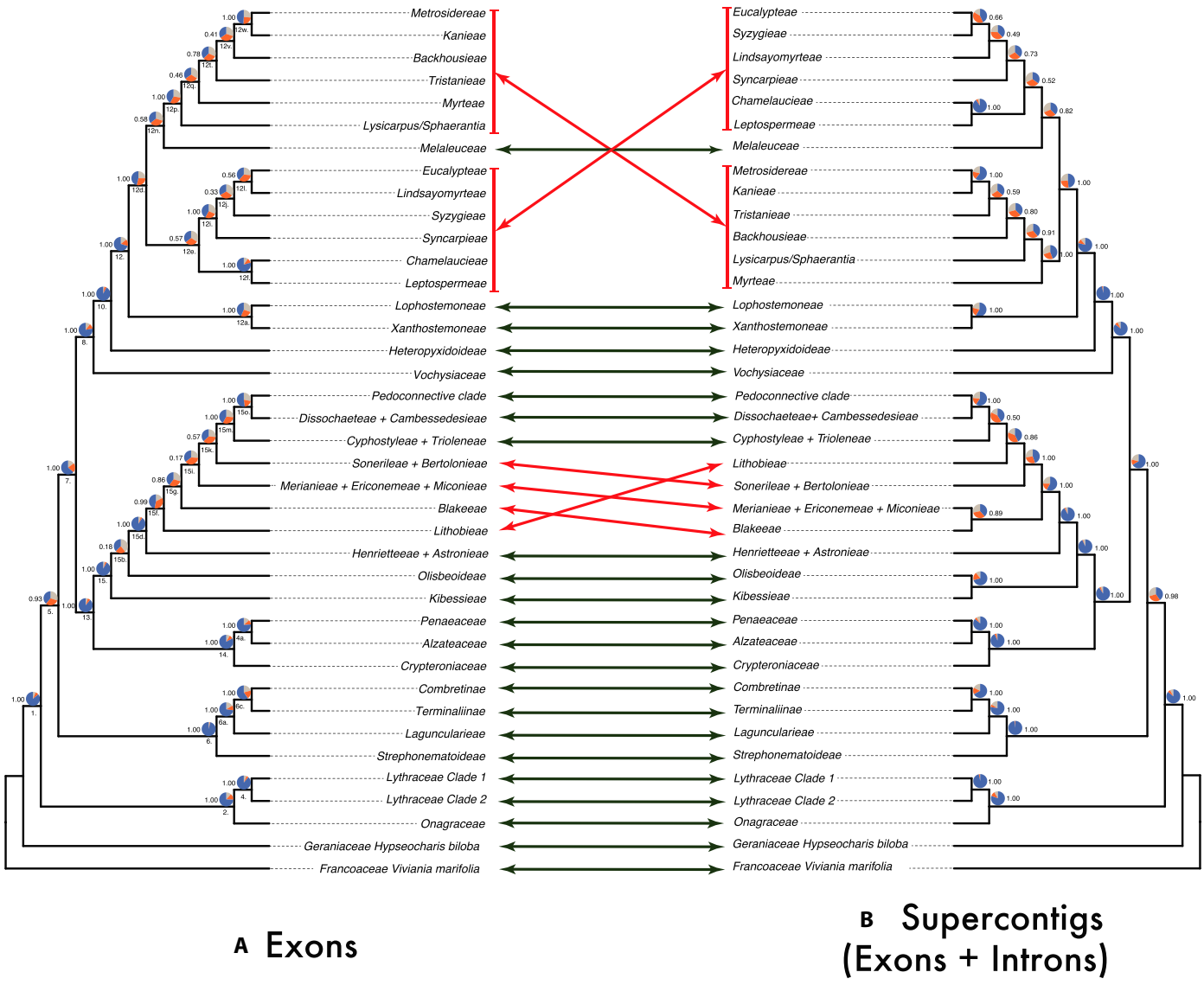

FIGURE 4. Exons coalescent ASTRAL (A) versus coalescent supercontigs ASTRAL (B) topology comparison in major families and subclades. Annotations above and below branches are respectively as described for Figs. 1A and 3A–C.

(clade 3b) received weak support in both ML (74 BS) and ASTRAL (LPP 0.56) trees; all other relationships are fully supported. Tribes Epilobieae (*Epilobium* L. and *Chamaenerion* Ség., clade 3c) and Onagreae (*Camissoniopsis*, *Tetrapteron* (Munz) W.L.Wagner & Hoch, *Neoholmgrenia*, *Camissonia*, *Eremothera* (P.H.Raven) W.L.Wagner & Hoch, *Clarkia* Pursh., *Gayophytum* A.Juss., *Taraxia* (Torr. & A.Gray) Nutt. ex Raim., *Eulobus* Nutt. ex Torr. & A.Gray, *Oenothera*, and *Chylismia* (Torr. & A.Gray) Nutt. ex Raim.; clade 3d) are both recovered as monophyletic with full support. All of the intergeneric relationships within Onagreae are fully supported by the ML analyses, but not so in the ASTRAL tree, which shows low support in parts of the backbone (e.g., the position of *Eremothera* as sister to *Camissonia* s.s. and that of a clade formed by the genera *Neoholmgrenia*, *Camissoniopsis*, and *Tetrapteron*).

Lythraceae—Results for relationships of the Lythraceae can be seen in Figs. 2, 3A (ASTRAL), and Appendix S2 (ML). In Lythraceae, all genera except *Lafoensia* Vand. are here represented. An initial dichotomy at the base of the Lythraceae produces two consistent, major lineages (clades 4a and 4b) that receive strong ML and ASTRAL support except clade 4b that receives weak QS. Clade 4a consists of two subclades: subclade 1 (LPP 1, strong QS) includes *Capuronia* Lourteig, *Galpinia* N.E.Br., *Pemphis* J.R.Forst. & G.Forst., and *Punica* L. Subclade 2 (LPP 1, weak QS) includes *Adenaria* Kunth, *Cuphea*, *Diplusodon* Pohl, *Koehneria* S.A.Graham, Tobe & Baas, *Lourtella* S.A.Graham, *Pehria* Sprague, *Physocalymma* Pohl, *Pleurophora* D.Don, and *Woodfordia* Salisb. Clade 4b comprises four subclades: subclade 1 (LPP 1, strong QS) comprises *Heimia* Link, *Didiplis* Raf., and *Rotala* L., sister to the rest; subclade 2 (LPP 1, moderate QS) includes *Lythrum*

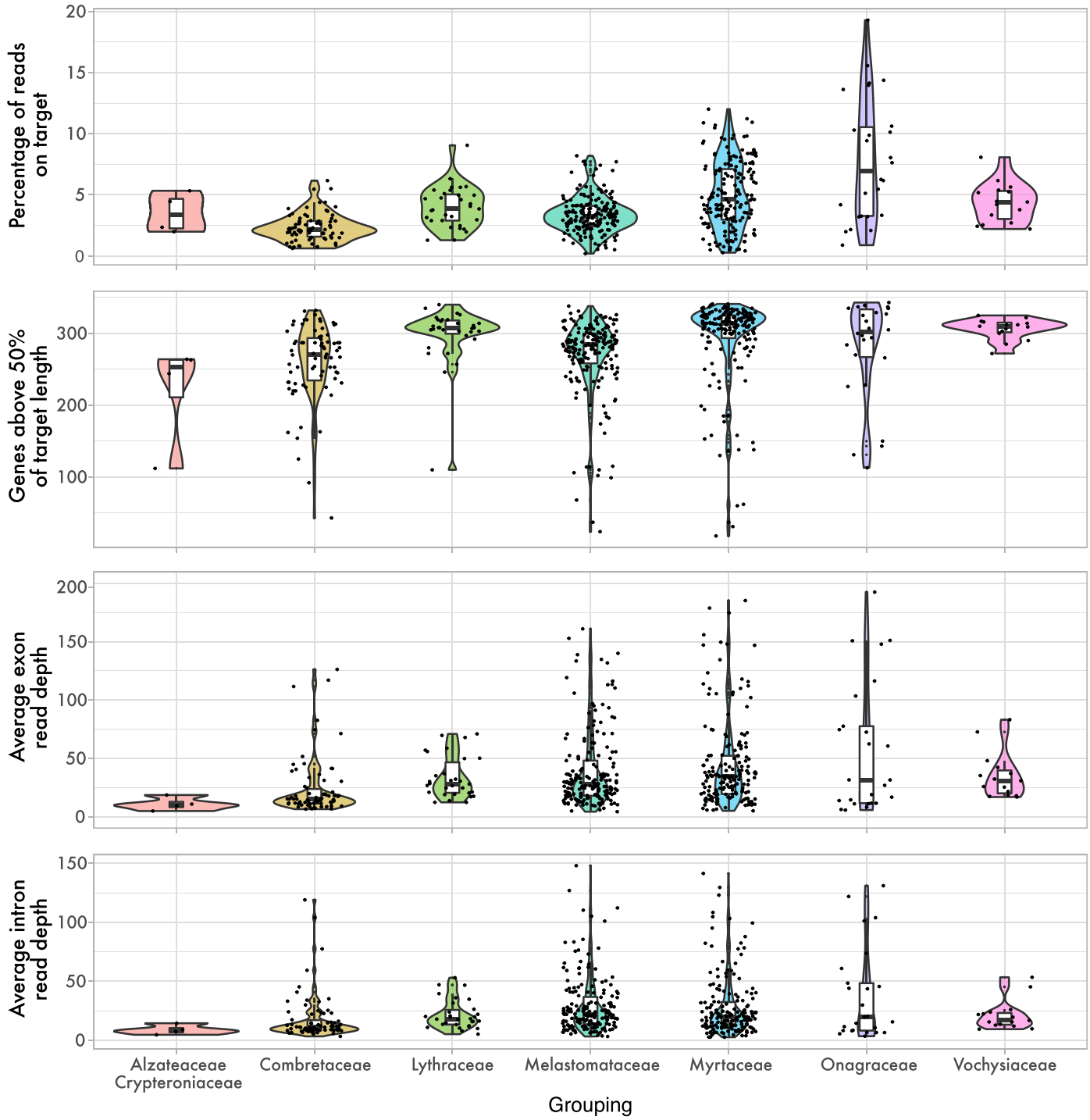

FIGURE 5. Visual recovery statistics per family.

L. and *Decodon* J.F.Gmel., sister to subclade 3 (LPP 1, strong QS), which comprises *Sonneratia* L.f., *Trapa* L., *Lagerstroemia* L., and *Duabanga* Buch.-Ham.; and subclade 4 (LPP 1, strong QS) including *Lawsonia* L., *Ginoria* Jacq., *Tetrataxis* Hook.f., *Ammannia* L., and *Crenea* Aubl.

Combretaceae—Results for relationships of the Combretaceae can be seen in Figs. 2, 3A (ASTRAL), and Appendix S2 (ML). Within Combretaceae (clade 6), the monotypic subfamily Strephonematoideae is well supported as sister to subfamily Combretoideae (clade 6a) within which tribe Laguncularieae is sister to Combreteae, with moderate QS but high LPP and BS support (clade 6b). Tribe Combreteae is strongly supported as monophyletic (clade 6c) including subtribes Terminaliinae and Combretinae and receives strong LPP and BS but with moderate QS (clade 6d and 6e). In Terminaliinae, *Conocarpus* L. is sister to all other *Terminalia*

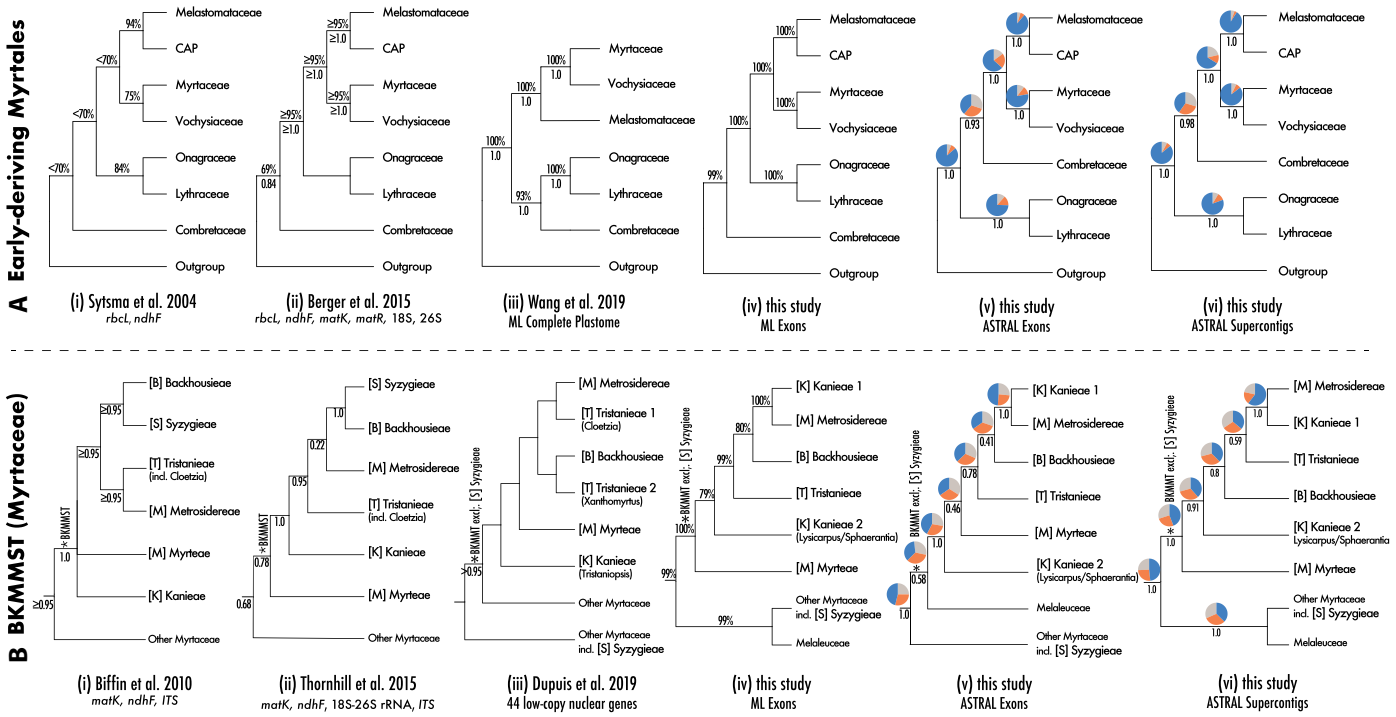

FIGURE 6. Topological contrast of two controversial clades of Myrtales. (A) Relationships between Combretaceae, Lythraceae and Onagraceae. CAP = Crypteroniaceae, Alzateaceae, Penaeaceae. (B) The BKMMT clade (Myrtaceae) and the position of Syzygieae. Support values (when available) are reported from this study or from the relevant publications. Supports above branches are bootstrap support or quartet score values for each node (as described in Figs. 1A and 3A–C). Local posterior probability values are presented below branches. BKMMST = Backhousieae, Kanieae, Metrosidereae, Myrteae, Syzygieae, Tristanieae in Myrtaceae (BKMMT excludes Syzygieae).

s.l. In Combretinae, the monotypic genera *Getonia* Roxb. and *Guiera* Adans. ex Juss. are well supported as closely related (clade 6f) and sister to the remainder of the subtribe. Within subtribe Combretinae, the relationships between *Combretum* subgenera are unclear. The monospecific subgenera *Apetalanthum* Exell & Stace and *Cacoucia* (Aubl.) Exell & Stace group with low QS and LPP with *Combretum mucronatum* Schumach. & Thonn., a member of the subgenus *Cacoucia* section *Mucronata* Engl. & Diels, which is sister to subgenus *Combretum*, again with weak support. Subgenus *Apetalanthum* receives high BP support as sister to subgenus *Combretum*, with *C. mucronatum* being sister to Combretinae (excluding *Getonia* and *Guiera*). Within *Combretum*, the inclusion of partial flanking introns provides no additional resolution for the relationships between the three subgenera *Apetalanthum*, *Cacoucia*, and *Combretum* (Fig. 3A vs. Appendix S3). Relationships between closely related species in *Combretum* section *Ciliatipetala* are not better resolved with the inclusion of introns.

Vochysiaceae—Results for relationships of the Vochysiaceae can be seen in Figs. 2, 3B (ASTRAL), and Appendix S2 (ML). Vochysiaceae are well recovered as monophyletic (clade 9) with the genera appearing in two highly supported clades. Clade 9a formed by a *Korupodendron* Litt & Cheek-*Erismadelphus* Mildbr. clade sister to a clade comprising the rest of the family. The remaining genera appear in three highly supported clades (9b, 9c, and 9d) with the following relationships (*Erisma* Rudge (*Vochysia* Aubl., *Salvertia* A.St.-Hil.)) and (*Callisthene* Mart. (*Qualea* Aubl., *Ruizterania* Marc.-Berti)). Still the relationship between these three clades is not well resolved, due to low support (LPP 0.84, BS 60) for Vochysieae ([*Vochysia* + *Salvertia*] and [*Callisthene* + *Qualea* (including *Ruizterania*)] and for the sister relationship between *Erisma* and *Vochysia* + *Salvertia*. *Salvertia* is sister to *Vochysia*, and *Callisthene* is sister to a clade in which *Ruizterania* is embedded in *Qualea*. All nodes within each of these three clades receive maximum support (LPP 1, BS 100).

Myrtaceae—Results for relationships of the Myrtaceae can be seen in Figs. 2, 3B (ASTRAL), and Appendix S2 (ML). *Kania* Schltr. is separated from the rest of former-tribe Kanieae Engl. (clade 12q) and placed sister to Metrosidereae (Benth.) Peter G.Wilson (clade 12v). All tribes except Xanthostemoneae Peter G.Wilson, Lophostemoneae Peter G.Wilson, Melaleuceae Burnett, and presumably its associate (Thornhill et al., 2015), Osbornieae Peter G.Wilson, are divided into the well-supported "core" clade (clade 12d) comprising two clades: (1) the poorly supported LESSLC (clade 12e) which includes Lindsayomyrteae Peter G.Wilson, Eucalypteae (Benth.) Peter G.Wilson, Syncarpieae Peter G.Wilson, Syzygieae, Leptospermeae DC., and Chamelaucieae DC.; and (2) the moderately supported BKMMT (clade 12p), which comprises Backhousieae (Nied.) Peter G.Wilson, Kanieae, Myrteae, Metrosidereae and Tristanieae). Fleshy fruited Syzygieae are placed within the LESLC clade, sister to Eucalypteae (weakly supported clade 12j) plus monogeneric Lindsayomyrteae (clade 12l). Syncarpieae is sister to (Syzygieae (Eucalypteae, Lindsayomyrteae)) (clade 12i). Melaleuceae is placed with low support, within the "core" clade and sister to BKMMT (clade 12n). Fig. 6B provides a summary comparing key topologies

of the relationship between the BKMMST tribes from previous studies to results from the different analyses presented here. Tribes Xanthostemoneae and Lophostemoneae remain sister to each other (clade 12a) and sister to the "core" clade. The ASTRAL and ML trees of Myrteae are congruent, except subtribe Pliniinae E.Lucas & T.Vasc. + *Blepharocalyx salicifolius* (Kunth) O.Berg (a previously unreported, relatively weak relationship; LPP 0.9, weak QS), which is sister to subtribe Myrciinae or Eugeninae in the ASTRAL and ML trees, respectively (LPP 0.83, weak QS). Subtribe Ugninae is not monophyletic as currently circumscribed, with *Ugni* Turcz. and *Myrteola* O.Berg emerging sister to all Myrteae (LPP 1, high QS), except for Myrtinae and Decasperminae, while *Lophomyrtus* Burret and *Lenwebbia* N.Snow & Guymer form a clade with *Amomyrtus* (Burret) D.Legrand & Kausel (LPP 0.92, weak QS). Pimentinae are returned as expected, except that *Pimenta pseudocaryophyllus* (Gomes) Landrum emerges instead within subtribe Myrtinae (LPP 1, high QS). Previously unsequenced *Amomyrtella* Kausel is sister to Pimentinae (LPP 0.51, weak QS), both sister to subtribe Luminae (LPP 0.32, weak QS). The KARPO clade (Vasconcelos et al., 2019) formed by *Kanakomyrtus* N.Snow, *Rhodomyrtus* (DC.) Rchb., *Pilidiostigma* Burret and *Octamyrtus* Diels here also includes *Decaspermum* J.R.Forst. & G.Forst. (LPP 1, high QS), but not *Archirhodomyrtus* (Nied.) Burret. Remaining Decasperminae genera emerge as subsequent sister clades either singly or in pairs, including *Myrtella* and *Lithomyrtus* F.Muell., emerging as sisters (LPP 1, high QS). New Caledonian *Myrtastrum* Burret emerges as sister to all neotropical Myrtaceae (LPP 0.91; weak QS). Within Chamelaucieae, subtribe Rinziinae (Rye et al., 2020) emerges sister to the rest of the tribe (LPP 0.56; weak QS), and *Astus* Trudgen & Rye and *Triplarina* Raf. are not placed within the subtribe.

Examining relationships in the species-rich clades in the exon–intron analysis (Fig. 4; Appendix S3), in *Myrcia*, *M. saxatilis* (Amshoff) McVaugh, and *M. amazonica* DC. of section *Aulomyrcia* (O.Berg) Griseb. inferred in a clade sister to *M. tomentosa* (Aubl.) DC., of section *Tomentosae* E.Lucas & D.F.Lima. While support for this relationship is moderate in the exon-only tree (LPP 0.63), it is high in the supercontig tree (LPP 0.94). The relationship of these two sections (clade *Aulomyrcia-Tomentosae*) with respect to the highly supported clade composed of sections *Calyptranthes* (Sw.) A.R.Lourenço & E.Lucas (represented by *M. loranthifolia* (DC.) G.Burton & E.Lucas and *M. psilophylla* Flickinger) and *Sympodiomyrcia* M.F.Santos & E.Lucas (represented by *M. mutabilis* (O.Berg) N.Silveira and *M. bicarinata* (O.Berg) D.Legrand) shows low congruence and moderate support in both trees. Within *Eugenia*, both exon and supercontig topologies return *Eugenia umbellulifera* (Kunth) Krug & Urb. (*Pseudanamomis* Kausel) sister (LPP 1, high QS) to the rest of *Eugenia* O.Berg, with high congruence and full support. The exon only analysis places *Eugenia* section *Jossinia* sister to section *Umbellatae* O.Berg with low congruence and weak support (LPP 0.48); in the supercontig analysis support is moderate (LPP 0.65). The supercontig analysis places *Eugenia* section *Racemosae* O.Berg [*E. biflora* (L.) DC.] sister to a clade formed by *E. speciosa* Cambess. (*Eugenia* section *Speciosae* Bünger & Mazine) and *E. involucrata* DC. [*Eugenia* section *Phyllocalyx* (O.Berg) Nied.], with low congruence and moderate support (LPP 0.52). The exon only analysis, instead, places *Eugenia biflora* sister to the clade comprising *Eugenia* sections *Umbellateae*, *Jossinia*, *Speciosae*, and *Phyllocalyx* with low congruence and high support (LPP 0.77). Within Eucalypteae and Syzygieae, species arrangement is almost identical in the exon and supercontig topologies, with weakly or moderately supported nodes sharing similar QS values but LPP almost exclusively slightly higher in the latter analysis. *Eucalyptus* is inferred as sister to a clade formed by *Corymbia* K.D.Hill & L.A.S.Johnson /*Angophora* Cav. + the "mesicalypts" (sensu Thornhill et al., 2019; *Allosyncarpia* S.T.Blake, *Eucalyptopsis* C.T.White, and *Stockwellia* D.J.Carr, S.G.M.Carr & B.Hyland) with low congruence in both trees and moderate (LPP 0.51) to high (LPP 0.88) support in the exon and supercontig trees, respectively). *Arillastrum* Pancher ex Baill. (= "newcalypt" sensu; Thornhill et al., 2019) is fully supported in all trees as sister to the clade comprising all the former genera. Unexpectedly in Syzygieae, *Syzygium unipunctatum* (B.Hyland) Craven & Biffin is placed sister to the rest of *Syzygium* s.s. with equally strong support (LPP 1; high QS) in both exon-only and supercontig analyses.

CAP-Melastomataceae clade—Results for relationships of the CAP-Melastomataceae clade can be seen in Figs. 2, 3C (ASTRAL), and Appendix S2 (ML). In both the ASTRAL and concatenated ML trees, Crypteroniaceae are recovered as sister to a clade where monotypic Alzateaceae is sister to Penaeaceae (the CAP clade; clade 14). Within Penaeaceae, the only difference between the ASTRAL and ML tree is that in the former, *Glischrocolla* A.DC. is recovered as sister to *Endonema* A.Juss., while in the latter, *Glischrocolla* is recovered as sister to the remaining Penaeaceae genera.

The ASTRAL tree shows the CAP clade as sister to the Melastomataceae (clade 13), where in turn the Kibessieae Krasser (clade 15a) is sister to the rest of the family. Olisbeoideae (clade 15c) is the next diverging clade, followed by a clade formed by the Henrietteeae Penneys, Michelangeli, Judd & Almeda and Astronieae Triana (clade 15e). Next (clade 15f), the monotypic *Lithobium* Bong. is resolved in an isolated position, followed by the Blakeeae Benth. & Hook.f. (clade 15h). Progressing along the backbone of the tree (clade 15i), from the root, the next clade is formed by the Merianieae Triana as sister to a subclade containing the Eriocnemeae Penneys & Almeda as sister to the Miconieae DC. (clade 15j). The monotypic *Catocoryne* Hook.f., never sampled before, sits also within the Miconieae, and the monotypic *Ochthephilus* Wurdack is resolved within the Merianieae. The next clade contains Bertolonieae Triana as sister to the Sonerileae Triana (clade 15i). Within the predominantly Old World Sonerileae, a neotropical grade (*Opisthocentra* Hook.f. and *Boyania colombiana* Humberto Mend.) is followed by a paleotropical one (African *Gravesia* Naudin) with neotropical *Phainantha* Gleason intercalated between the latter two. Other neotropical species are more deeply nested (*Tryssophyton* Wurdack, *Boyania ayangannae* Wurdack) in the Sonerileae. The concatenated ML topology is similar, but *Phainantha* is sister to the two species of *Gravesia* Naudin, and some internal relationships within the Sonerileae differ.

Continuing along the backbone of the tree (clade 15m), the next four diverging lineages in order are the Cyphostyleae Gleason + Trioleneae Bacci, Michelang. & R.Goldenb. (clade 15n), Cambessedesieae Bochorny, Almeda, Michelang. & R.Goldenb. + Dissochaeteae Triana (clade 15p), Rhexieae, and then the di-specific *Rupestrea* R.Goldenb., Almeda & Michelang. is resolved as sister to the Marcetieae + Microlicieae and this larger clade as sister to the Melastomateae. Within the Melastomateae neotropical *Pterogastra* Naudin is sister to neotropical *Loricalepis* Brade + the remaining of the tribe composed of two subclades. One contains neotropical *Pterolepis* (DC.) Miq., as sister to a monophyletic Old World Melastomateae, while the second clade contains all other neotropical

genera. However, support for the backbone of the Melastomateae is weak. There are a few but important differences in the ML vs. the ASTRAL tree; in the ML tree Kibessieae and Olisbeoideae form a clade sister to the remaining of the family (instead as sequentially diverging clades). Additionally, the *Lithobium* is resolved as sister to the Blakeeae with high support and not in an isolated position. Within the Pedoconnective clade, *Rupestrea* is recovered as sister to the Melastomateae, and *Loricalepis* is recovered as sister to the rest of the tribe with high support. *Pterogastra* + *Pterolepis* are recovered forming a clade sister to the Old World Melastomateae (albeit with moderate support), with this latter clade then sister to the remaining New World Melastomateae.

The addition of exons to the CAP clade and Melastomataceae analysis recovers all the same major clades with the same generic composition and some nodes along the backbone have increased LPP support (Fig. 4; Appendix S3). The two notable differences are the position of *Lithobium*, which is no longer recovered isolated between Henrietteeae and Blakeeae (exon tree) in the supercontig tree, but as sister to the Bertolonieae plus Sonerileae, and Blakeeae, recovered as sister to the Merianieae-Ericonemeae-Miconieae clade. Within the Miconieae or Melastomateae the inclusion of introns does not provide a substantial increase in LPP support values, and in fact, some relationships along the internal backbone of these two clades are instead more-weakly supported.

DISCUSSION

Myrtales backbone

Results presented here, based on hundreds of low-copy nuclear genes, still do not fully clarify the backbone topology of Myrtales. Only the ML analysis supports previous results by Berger et al. (2016), suggesting that Combretaceae are sister to the rest of the order, with a Lythraceae-Onagraceae clade sister to all other families. More recently Wang et al. (2021) and Zhang et al. (2021) produced topologies based on data from the complete plastome; results differ again, with Combretaceae sister to a Lythraceae-Onagraceae clade, forming one of two major clades in the order (with all remaining families in the other clade). Our ASTRAL topologies, from either exons or supercontigs, suggest that the clade formed by the latter two families is sister to the rest of the order with high LPP, but with similarly low congruence. Indeed, ASTRAL QS values indicate clear gene tree conflict among the nodes closest to the root of the order, indicating a complex evolutionary history at this node, also supported by the ASTRAL polytomy test that does not allow for rejection of the null hypothesis, i.e., the root node being a polytomy. Relationships between these Myrtales families (Combretaceae, Lythraceae, and Onagraceae) remain ambiguous based on the results of this study, indicating incomplete lineage sorting (ILS) and potential ancient hybridisation. For the remainder of the families, the backbone is congruent with previous studies, but with higher support values.

Onagraceae

The infra-familial topology recovered here and described as (Ludwigioideae (Circaeeae (Lopezieae (Gongylocarpeae (Epilobeae, Onagreae))))) is similar to previous tribe-level relationships based on ITS and few plastid markers (Levin et al., 2003, 2004; Berger et al., 2016; Freyman and Hohna, 2019). Interestingly, the position of Gongylocarpeae as sister to a clade comprising tribes Epilobieae and Onagreae, a well-supported relationship in previous analyses using few markers (Levin et al., 2004) and by some morphological characters (e.g., lack of stipules, Raven 1964), is only weakly (74 BS) to moderately (0.56 LPP) supported here. This result is noteworthy as it confirms that *Gongylocarpum*, only recently segregated from Onagreae (Levin et al., 2004; Wagner et al., 2007), is indeed distinct from the latter.

Tribes Epilobieae and Onagreae are both confirmed to be monophyletic with full support, corroborating all previous phylogenetic studies in the family (Baum et al., 1994, Levin et al., 2003, 2004; Berger et al., 2016). A close relationship between the genera that form these two tribes has been suggested (Wagner et al., 2007) ever since *Oenothera* L., *Gaura* L. (today a synonym of *Oenothera*) (Onagreae), and *Epilobium* (Epilobieae) were described by Linnaeus (1753). The backbone of Onagreae and relationships among its genera have been mostly poorly supported in previous studies (Levin et al., 2004). Although low support values are still present in ASTRAL trees, the ML analyses suggest for the first time a well-supported relationship for the genera within this tribe. A close relationship between *Oenothera*, *Eulobus* Nutt., and *Chylismia* Nutt. had been previously recovered (e.g., Levin et al., 2004; Freyman and Hohna, 2019); however, in all those cases, *Oenothera* was placed as sister to *Chylismia*, whereas results presented here suggest a fully supported relationship between *Oenothera* and *Eulobus*. Similarly, *Tetrapteron* (Munz) W.L.Wagner & Hoch, *Camissoniopsis* W.L.Wagner & Hoch, *Neoholmgrenia* W.L.Wagner & Hoch, *Camissonia*, and *Eremothera* are inferred as belonging to the same clade in previous analyses using ITS and a few plastid markers (e.g., Levin et al., 2004), but in a slightly different configuration from the one we infer. The position of *Taraxia* as sister to a clade including all genera of Onagreae except for *Oenothera*, *Eulobus*, and *Chylismia* is another novelty of this study and provides an alternative scenario for the diversification of the tribe, previously recovered as a polytomy (Wagner et al., 2007). Increased species sampling, comparison with plastid data, and a better understanding of the evolutionary mechanisms behind the rapid diversification of tribes Onagreae and Epilobieae (e.g., Hollister et al., 2019) are necessary to further resolve relationships within Onagraceae.

Lythraceae

In the analysis of Graham et al. (2005), the arrangement into two subclades (here 4a and 4b) can be detected, although different topologies presented in the former analysis show that support is generally low and genera flip between clades. Berger et al. (2016), using a much-reduced sample, return the same arrangement as presented here, but with only moderate support for clade 4a. Relationships in clades 4a and 4b are fully or nearly fully congruent in the topologies returned by the ASTRAL and ML analyses presented here. The exclusion of *Heimia* and *Rotala* from clade 4a in this study is incongruent with the combined molecular analysis of Graham et al. (2005); there is also incongruence in relationships between genera. In the analysis presented here, *Lythrum* is strongly supported as forming part of clade 4b, incongruent with the study of Graham et al. (2005), who found *Lythrum* sister to the remainder of Lythraceae. Clade 4b also associates *Decodon* and *Lythrum* with *Ammannia*, *Lawsonia*, *Sonneratia*, and *Duabanga* as found in previous studies (Graham et al., 2005; Berger et al., 2016). Lythraceae nodes are universally

strongly supported; the only two nodes that do not receive LPP 1 unite *Ammannia* with *Crenea* and *Lawsonia*, *Ginoria*, *Tetrataxis*, *Ammannia*, and *Crenea* with *Lagerstroemia*, *Trapa*, and *Sonneratia*. The primarily small herb clade of *Ammannia/Crenea* encompasses a morphologically diverse group of taxa. *Trapa* may be responsible for some of this incongruence as it is highly autapomorphic and unique in the family in a multitude of ways, both micro- and macromorphologically (Kadono and Schneider, 1986).

Results from the ML analysis (Appendix S2) show several genera on long branches, some represented today by a single, or few species, such as *Lawsonia*, *Trapa* (maybe three species or dozens depending on taxonomic judgments), and *Pemphis*. *Diplusodon*, on a long branch, is exceptionally species rich (see also Inglis and Cavalcanti, 2018). In contrast, *Cuphea*, the largest genus of the family, appears to be the result of a more recent radiation. These long-branch clades may be an area of focus in future studies as long branch attraction is known to interfere with phylogenetic inference. The two clades and eight subclade structure is not supported by comparative morphology but reflects to a certain extent previous phylogenetic studies in Lythraceae (Graham et al., 2005; Gu et al., 2019) and confirms that current classifications of the family and of *Cuphea* are highly incongruent with evolutionary relationships.

Combretaceae

Results in Combretaceae, which include all 10 currently accepted genera (POWO, 2020; Maurin et al, 2020), as well as formerly recognized genera (*Anogeissus* (DC.) Wall. ex Guill. & Perr., *Buchenavia* Eichler, *Meiostemon* Exell & Stace, *Thiloa* Eichler, *Pteleopsis* Engl.), confirm previous findings and relationships within the family (Tan et al., 2002; Maurin et al., 2010, 2017). The monophyly of the family is supported with the monogeneric subfamily Strephonematoideae sister to all other genera in Combretoideae. However, results presented here increase resolution beyond that previously observed and confirm the monophyly of tribe Laguncularieae, a clade including mangrove taxa, as sister to the Combreteae. Inclusion of all mangrove genera in the Laguncularieae clade supports the relatively recent biogeographic split between eastern and western mangrove habitats as observed in other mangrove taxa (Plaziat et al., 2001). As a result, mangrove habitats have been separated by the Eurafrican gap (Plaziat et al., 2001), with *Laguncularia* distributed from the New World tropics to tropical West Africa and *Lumnitzera* from tropical East Africa to Australia and the Pacific and finally, the northwestern Australian clade comprising *Dansiea* and *Macropteranthes*. These results suggest that further morphological and molecular investigations are required to assess generic delimitation between northwestern Australian taxa more accurately. Within tribe Combreteae, the inclusion in this study of several previously accepted genera confirms the monophyly of both subtribes Combretinae and Terminaliinae with the inclusion of *Meiostemon*, *Quisqualis* L., and *Thiloa* in the first and *Anogeissus*, *Buchenavia*, and *Pteleopsis* in the last. The more densely sampled Combretaceae, beyond a single representative per currently accepted genus, permits further discussion regarding the relationships within the family. In Terminaliineae, *Conocarpus* is confirmed as sister to the rest of the subtribe, while within the rest of the subtribe our sampling remains too limited to distinguish any geographical or morphological clades. In Combretiinae, inclusion for the first time in a phylogenetic analysis of the monotypic subgenus *Apetalanthum* from Myanmar, confirms the position of this subgenus and relationships between the two additional subgenera, *Cacoucia* and *Combretum*. *Combretum apetalum* Wall. ex Kurz has long been considered, based on morphological characters, as a sister lineage to subgenera *Combretum* and *Cacoucia*. In both topologies (Fig. 3A, exons; Appendix S3, supercontigs), *C. apetalum* is highly supported as sister to subgenus *Cacoucia*, while *C. mucronatum* appears sister to subgenus *Combretum*, with very weak support.

Both exon and supercontig analyses show poorly resolved relationships within the three subgenera, and only a more thorough investigation or a denser sampling will allow resolution of these specific relationships. At a shallower taxonomic level, the potential for exons and introns to resolve relationships within a clade of closely related taxa, such as subgenus *Combretum* section *Ciliatipetala*, is low. This section is predominantly southern African and contains numerous, morphologically homogeneous species. Field observations have recently revealed new species (Maurin et al., 2011; Boon et al., 2020) and it is suspected that more undescribed species remain. Overall, the topology within this clade is similar in both analyses, and support values are low overall, with quartet scores highlighting gene tree incongruence.

Vochysiaceae

In the present study, Erismeae was recovered as nonmonophyletic, with a clade formed by *Korupodendron* + *Erismadelphus* appearing as sister to a clade comprising the rest of the family. The monophyly of Vochysieae remains uncertain, based on incongruent topologies retrieved with the two methods of phylogenetic inference used in this study (ML and ASTRAL). In the ML analysis, Vochysieae are monophyletic, but the support is low; in the ASTRAL analysis the clade *Vochysia* + *Salvertia* is sister to *Erisma*, with moderate support. *Salvertia* was retrieved as sister to *Vochysia*, corroborating some previous studies (Sytsma et al., 2004; Gonçalves et al., 2019, 2020). The placement of *Salvertia* has always been intriguing in the history of the family. While most species of Vochysiaceae have flowers with one or three petals and are pollinated by bees, *Salvertia* bears five-petaled, completely white flowers, associated with hawkmoth pollination (Oliveira et al., 2004). *Callisthene* appears sister to a clade formed by *Qualea*, with *Ruizterania* embedded in it, corroborating a placement highly supported in Gonçalves et al., 2020). Further studies using more molecular markers and increasing the number of species of *Erisma*, *Callisthene*, *Qualea* (including species assigned to *Ruizterania*), and *Vochysia* are necessary before taxonomic decisions are reached. Ultimately, the present study is the first to use a large number of nuclear markers for the genera of Vochysiaceae. The relationships recovered in the ML analysis appear congruent with plastome phylogenetic studies, despite relatively low support for the monophyly of Vochysieae. Our ASTRAL topology does not correspond to the plastome phylogeny of Gonçalves et al. (2020), since we retrieve *Erisma* as sister to *Vochysia* + *Salvertia* instead.

Myrtaceae

The division of Kanieae has been recorded elsewhere and a new tribe will house the non-*Kania* genera (P. G. Wilson et al., unpublished manuscript). Comparison of previously published topologies of the relationships of Backhousieae, Kanieae, Metrosidereae, Myrteae, Syzygieae, and Tristanieae (BKMMST, Biffin et al., 2010; Thornhill et al., 2015; Dupuis et al., 2019) to results presented here

indicates that uncertainty remains with regards to relationship among these tribes. However, support values from our analysis indicate to some degree that Syzygieae are not part of the BKMMST clade, as noted by Dupuis et al. (2019). This arrangement suggests that the two most species-rich, fleshy-fruits tribes are more distantly related than previously believed.

The sister relationship of *Lindsayomyrtus racemoides* (Greves) Craven to Eucalypteae is notable as the former is a rainforest species with few morphological characters in common with the latter. *Lindsayomyrtus racemoides* has large seeds, fruits apparently not wind dispersed and has previously been considered most closely related to *Xanthostemon* F.Muell. or *Eugenia* (White, 1942; Craven, 1990; Zich et al., 2018). The *Eucalyptus* arrangement differs from the results of Thornhill et al. (2019) where the "mesicalypts" and "newcalypt" clades are interchangeably sister to a clade of *Eucalyptus* sister to *Corymbia*/*Angophora*, more similar to the result from the ML analysis presented here. *Corymbia* is not monophyletic, as previously reported by Schuster et al. (2018); Thornhill et al. (2019), with the closer relationship of *Arillastrum* to Australian *Corymbia ficifolia* (F.Muell.) K.D.Hill & L.A.S.Johnson in our ASTRAL analysis more closely resembling the results of those studies.

Within tribe Myrteae, *Algrizea* Proença & NicLugh. was included in the Pliniinae by Lucas et al. (2019) following studies that place it as sister to the remainder of Pliniinae (e.g., Vasconcelos et al., 2017). However, the position of *Algrizea* as sister to *Myrcia* (as reported by Proença et al., 2006) indicates the need for taxonomic adjustment, as does the association of *Lophomyrtus*, *Lenwebbia*, and *Amomyrtus*. The position of *Amomyrtella* suggests it should be included in subtribe Pimentinae. In *Eugenia*, the absence of representatives of *Eugenia* sections *Hexachlamys* (O.Berg) Mazine and *Pseudeugenia* Mazine & Faria may influence the resulting topology. The sister relationship of *Eugenia umbellulifera* (Kunth) Krug & Urb. (*Pseudanamomis umbellulifera* (Kunth) Kausel), to the rest of the genus suggests the former could be a lineage endemic to the Caribbean and its morphological distinction from *Eugenia* (Flickinger et al., 2020) may have systematic significance. Previously *Pseudanamomis* was embedded in *Eugenia*, but its position was labile (Vasconcelos et al., 2017; Flickinger et al., 2020). The novel proximity recovered of *Eugenia* sections *Phyllocalyx* and *Speciosae* is also of note, as species of these sections were previously included in genera *Phyllocalyx* O.Berg and *Stenocalyx* O.Berg, morphologically similar in having enlarged, showy sepals. *Eugenia* section *Speciosae* Bünger & Mazine also has an unstable relationship, of note as species of these sections were previously included in genus *Phyllocalyx* O.Berg. The apparent polyphyly of *Pimenta* reflects historical uncertainty in Myrtinae and Pimentinae (Lucas et al., 2019; Vasconcelos et al., 2019) with small but significant differences apparent in nuclear only analyses (e.g., Salywon, 2003), vs. those including plastid markers (e.g., Vasconcelos et al., 2017), which recovered a monophyletic *Pimenta*, including *Pimenta pseudocaryophyllus* within Pimentinae. It is of note that several taxa formerly assigned to subtribe Myrtinae sensu Berg (1859), with subtropical or temperate distributions, have unstable relationships. These taxa share physiological characteristics, such as frost-resisting scalariform plates within vessel elements that may be indicative of the existence of Myrteae some 80 Ma over land that now forms Antarctica (de la Estrella et al., 2019). Regarding the exclusion of *Archirhodomyrtus* from Decaspermineae E. Lucas & T. Vasc., it is not clear whether *A. paitensis* shares the thickened pseudo-lamina between seeds typical of the other KARPO genera (Vasconcelos et al., 2019). The previously unrecorded position of *Syncarpia glomulifera* (Sm.) Nied. sister to Syzygieae, Lindsayomyrteae, and Eucalypteae is of note as ovaries fuse during fructification and fruits are capsules, similar to other dry-fruited genera such as *Lophostemon* Schott, with which it has been compared (Wilson and Waterhouse, 1982; Bean, 1995). The sister relationship of a non-capsular fruited *Lindsayomyrtus* B.Hyland & Steenis with Eucalypteae and capsular fruited *Syncarpia* with the remaining genera of Clade 12i, suggests more switches between dry and fleshy fruits in Myrtaceae than previously anticipated.

Analysis of the supercontig topology again indicates agreement with recent infra-generic classifications. In *Myrcia* (see Lucas et al., 2018), sectional relationships have yet to receive strong support, but the proximity of *Myrcia* section *Aulomyrcia* to *Myrcia* section *Tomentosae*, reported for the first time, has morphological congruence as species of these sections share inflorescences with a tendency for asymmetry. The addition of flanking intron data has an unpredictable effect at the level of section in *Eugenia*, increasing support for the position of *Eugenia* section *Jossinia* as sister to section *Umbellatae*, but not for relationships between sections *Racemosae*, *Speciosae*, and *Phyllocalyx* (contrasted with Van der Merwe et al. [2005]; Mazine et al. [2014, 2018]). Within Eucalypteae and Syzygieae, somewhat elevated support levels in the supercontig analysis suggest this extra data should be included in future analyses of species-rich groups enriched with Angiosperms353. However, incongruence remains, as evidenced by the strongly supported relationship of *Syzygium unipunctatum* as sister to the rest of *Syzygium* s.s. in both exon-only and supercontig analyses, that differs considerably from previous studies (Biffin et al., 2010; Craven and Biffin, 2010) where it grouped, also with strong support, with *S. acuminatissimum* (Blume) DC. in *S.* section *Acmena* (DC.) Craven & Biffin. *Syzygium*-focused phylogenomic work in progress by Y. W. Low et al., (unpublished manuscript), using 292 accession based on 1227 BUSCO genes (Simão et al., 2015), instead strongly supports those previously inferred arrangements.

CAP-Melastomataceae clade

Within the CAP clade the position of *Olinia* Thunb. and *Rhynchocalyx* Oliv., as successively sister to all other Penaeaceae, is consistent with previous analyses. Moreover, these phylogenetic relationships are consistent with either taxonomic alternative of recognizing an expanded Penaeaceae or three separate families (Penaeaceae, Oliniaceae, and Rhychocalycaceae).

Within Melastomataceae, the generic composition of all major clades and tribes is consistent with that of previous analyses (Clausing and Renner, 2001; Penneys et al., 2010; Michelangeli et al., 2011, 2013, 2014; Goldenberg et al., 2012, 2015; Penneys and Judd, 2013; Rocha et al., 2016; Veranso-Libalah et al., 2017; Bacci et al., 2019; Bochorny et al., 2019; Wurdack and Michelangeli, 2019; Zhou et al., 2019). However, there are some important differences in the relationships of major clades or tribes, mostly for nodes with low to moderate support along the backbone of the tree. A notable result is the placement of the Kibessieae as sister to the remaining of the family (ASTRAL) or as sister to subfamily Olisbeoideae (ML), both results conflicting with earlier phylogenetic results that recovered the Olisbeoideae as sister to the remainder of the family (Clausing and Renner, 2001; Wurdack and Michelangeli, 2019). It should be noted that these previous analyses were based only on plastid data, and that inclusion of nuclear data in more recent analyses has also recovered the relationship of Kibessieae as sister to Olisbeoideae (Bacci et al., 2019). Those tribe share the presence of included phloem, the anthers

opening by slits (present also in scattered species across the family), antipetalous ovary locules (when the number of locules and petals coincide), and strong differential growth of the inferior ovary (Morley, 1976; Maxwell, 1981; van Vliet, 1981; van Vliet et al., 1981). However, Kibessieae also shares with the rest of Melastomatoideae the more typical acrodromous venation (present in some Olisbeoideae), the absence of anther glands, and numerous ovules per locule. Placentation, with ovules attached to the ovary wall between locules, appears intermediate between Olisbeoideae and the rest of Melastomatoideae (Morley, 1976; Clausing et al., 2000). Another finding of note is the placement of *Tessmannianthus* Markgr,, a neotropical genus, as sister to the remaining Astronieae, an otherwise paleotropical clade, as previously suggested based on morphology (Mancera, 2017).

The large-pedoconnective clade, including *Rupestrea*, Rhexieae, Microlicieae, Marcetieae, and Melastomateae has been recovered in several analyses (Michelangeli et al., 2013; Goldenberg et al., 2015; Wurdack and Michelangeli, 2019), but relationships between those tribes vary from one study to the next, and even within individual studies, depending on taxon and loci sampled and analytical approach taken. Previous analyses had suggested a close relationship between Sonerileae and Dissochaeteae (Clausing and Renner, 2001); however, increased taxon sampling shows that they form two distinct and non-sister lineages as recovered here.

Ochthephilus, a monotypic genus from the Guiana Shield, is recovered here within the Merianieae, but elsewhere forms a clade with *Eriocnema* and *Physeterostemon* R.Goldenb. & Amorim (Penneys et al., 2020). Within the Miconieae, *Pleiochiton* Naudin ex A.Gray, and *Leandra sanguinea* Gleason are recovered as sister to the remaining members of the tribe, but in other analyses have been found closer to the tips and not in the same clade (Goldenberg et al., 2008; Martin et al., 2008; Reginato and Michelangeli, 2016). Consistent with previous analyses, *Macrocentrum* Hook.f., *Meriania* Sw., and *Phyllagathis* Blume as currently defined are not monophyletic, and further work is needed with these groups.

A taxonomic consequence of this work is to consider whether subfamilies within the Melastomataceae should be recognized. Because the tribe Kibessieae, currently in the subfamily Melastomatoideae, is resolved either as sister to the remainder of the family or as sister to subfamily Olisbeoideae, the current classification system does not reflect these phylogenetic results. Each of the main clades could be recognized as tribes, in which case the rank of subfamily could be dispensed with and the Olisbeoideae would be recognized at the tribal level, as Mouririeae Richard, or Kibbessieae could be recognized as a subfamily, for which there is already the name Kibbesioideae Naudin.

CONCLUSIONS

The Angiosperms353 probe set has proved efficient in evaluating relationships in Myrtales at various taxonomic levels. Analytical methods building species trees under the multispecies coalescent process require new ways to evaluate support, since gene tree congruence/incongruence needs to be accounted for. Such methodologies require support to be evaluated beyond traditional bootstrapping, with quartet support allowing to interpret gene tree congruence/incongruence. Recovery of low-copy nuclear loci was satisfactory, with an average of 288 loci retrieved across the data set. McLay et al., 2021) demonstrate that recovery might be further improved by adapting the target file to the taxonomic group. However, recovery of the plastome via genome skimming was less efficient as on average, less than a quarter of the plastome was recovered, with no clear pattern emerging. Nonetheless, combining enriched and unenriched libraries might improve recovery, but needs to be undertaken with care to avoid recovery of plastid reads impacting nuclear recovery, the main goal of the Angiosperms353 kit.

With regards to recovery of the 353 low-copy nuclear loci and in relation to the wide taxonomic sampling employed, the inclusion of just exons resolves the backbone of the order to some extent. Inclusion of flanking partial introns, in addition to exons, contributes to resolution of relationships of closely related species but, generally, only marginally improving statistical support at historically unstable nodes. The most extreme topological conflict inferred occurs between the species tree reconstructions resulting from the summary coalescent method ASTRAL vs. those produced by concatenating genes, regardless of whether flanking intron data are included or not. Overall, the ASTRAL topology is closer than the ML topology to what might be expected based on morphological and/or previous DNA-based studies (e.g., Wilson et al., 2005 in Myrtaceae; Bacci et al., 2019 and Reginato et al., 2020 in Melastomataceae). This outcome reflects current thinking that coalescence analysis is the most appropriate approach for building phylogenomic species trees in the presence of gene tree conflict due to ILS (Liu et al., 2019). However, incongruent gene trees, due to either gene tree error or hybridization (Degnan and Rosenberg, 2006; Linkem et al., 2016; Li et al., 2019), can result in erroneous species trees. Anomalous gene trees resulting from short branches or long-branch attraction exist in the "anomaly zone" of Degnan and Rosenberg (2006; see also Susko and Roger, 2021). At deeper nodes, uncertainty may be linked to ancestral whole-genome duplications or horizontal gene flow resulting, e.g., in chloroplast capture, as identified in the early evolutionary history of the Asteridae (Stull et al., 2020), and may result in statistically significant conflicts between phylogenetic trees based on organellar vs. nuclear DNA sequence data. Results presented here support these considerations, with uncertain nodes often involving short branches at every rank as visible on the ML tree. Examples of two such key fluctuating relationships have been discussed in relation to Fig. 6. Discrepancies between results obtained with nuclear and plastid sequence data have been explored in detail in other groups (e.g., in Compositae; Stull et al., 2020).

In the quest for the ultimate phylogenetic reconstruction with unequivocal support at all nodes, it is tempting to believe that the answers lie in building trees with more and more genetic data. However, using more genes appears to raise more issues, some of which were common with few-gene topologies resulting from Sanger sequencing, and potentially some new problems too. A new form of uncertainty is associated with the coalescence approach, that is most appropriate for evaluating high numbers of genes and highlighting ILS and gene tree conflict that may result from ancient hybridisation or gene duplication. Future studies should focus effort on a better understanding of the genetic data available, at least to provide evidence that strongly supported but controversial relationships are not artefacts of the inference and/or the gene regions sampled. It is important to recognize that incongruence can appear as noise, but also that this noise is an integral part of the evolutionary process. The point at which taxonomic inference can be reliably made is a subject for future refinement of the analytical process. Smith et al. (2020) noted this lack of "panacea" and suggest a "data-centric middle-way" between concatenation and coalescent approaches is desirable. This "data-centric middle way" would help identify subsets of data that can be combined and would allow the

exploration of conflict between different subsets when it arises. This approach could be useful in identifying relationships that need more attention, but also nodes that will be difficult to resolve despite best efforts (hard polytomies). Also, rather than treating data as "good" or "bad", this type of approach has the potential to help us better understand the evolutionary processes underlying the tree of life.

The present study represents only a fraction of the information available. The taxonomic results described are the most remarkable, but exploration of every relationship is not possible in a single work. Additional studies, using the Angiosperms353 probe set and phylogenomic methods relying on the multispecies coalescent process, with increased taxonomic sampling within the individual families of Myrtales, will surely follow. These studies will continue to clarify relationships and increase statistical support, particularly at some of the weaker nodes along the backbone. There are some cases of very short branches that may receive high LPP support but low QS values and for which the ASTRAL polytomy test of Sayyari and Mirarab (2018) cannot reject the null hypothesis that the node is a polytomy. In these cases, such as the nodes closest to the Myrtales root, it is possible that additional genetic data will help resolve these nodes with strong support and reject the null hypothesis of the polytomy test.

The order Myrtales exhibits a wide variety of habits (including herbaceous herbs, lianas, trees, and mangroves), floral forms, and fruit types (berry, capsule, drupe and samara) and exhibits high species diversification in several fleshy-fruited and dry-capsular clades (e.g., within Melastomataceae and Myrtaceae) and will make an excellent model for study of adaptation of form over past and future evolutionary time frames.

ACKNOWLEDGMENTS

We gratefully acknowledge assistance from colleagues at RBGKew: Dion Devey (supporting laboratory work); Imalka Kahandawala and Rosie Woods (DNA Bank); Paul Bailey, Kevin Leempoel, and Eduardo Toledo, (bioinformatics). We are grateful to the editors and anonymous reviewers who have patiently helped improve the content and format of this manuscript.

This work was funded by grants from the Calleva Foundation and the Sackler Trust to the Plant and Fungal Trees of Life Project (PAFTOL) at the Royal Botanic Gardens, Kew, with additional funding from the Garfield Weston Foundation. A.A. received support from the 2018 CAS-TWAS President's Fellowship for International Doctoral Students. E.L. and A.G. received support from the Bentham-Moxon Trust providing funds for sequencing *Eugenia* species and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - proc. 155480/2018-9). Support from the National Science Foundation is acknowledged (DEB-0515665, DEB-0818399, DEB-1146409, DEB-1343612, DEB-1754667). We also thank the Assemblée de la Province Nord, Assemblée de la Province Sud (New Caledonia) for relevant collection permits. F.M. received support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PriInt) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - proc. 302309/2018-7). K.W.T. received support by a joint Natural Science Foundation of China-Yunnan Government research grant (U1502264).

AUTHOR CONTRIBUTIONS

O.M., E.L., and F.A.M. conceived the project and coordinated the study with contributions from S.B., W.J.B., S.D., F.F., A.G., O.P., and A.Z. O.M., D.G., and A.G. led the analyses. O.M. and E.L. led the sampling of specimens. O.M., E.L., L.P., S.B., and S.D. wrote the Methods section. O.M., A.Z., and B.G. led production of the figures. Text related to particular families was led by S.G. for Lythraceae; T.V. and P.H. for Onagraceae; O.M. and I.T. for Combretaceae, D.G. and G.S. for Vochysiaceae; E.L., P.W., T.V., A.G., F.F.M., Y.W.L., Y.P., V.S., A.H.T., and E.B. for Myrtaceae; F.A.M., R.G., and D.S.P. for Melastomataceae. All authors contributed to general topics in the introduction and discussion and edited the final manuscript. Laboratory work was carried out by G.B., N.E., C.M.G., and L.P. W.J.B. and F.F. acquired funding and supervised the PAFTOL project.

DATA AVAILABILITY

All Sequences files (fastq) generated for this study are deposited in the European Nucleotide Archive (ENA accession PRJEB35285; Appendix S1). All alignments, Newick tree files generated and scripts used in this research are available at Zenodo (https://doi.org/10.5281/zenodo.4268317).

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. Voucher specimens and ENA accessions; collection information. Samples highlighted in gray were sequenced multiple times, and sequence files have been merged.

APPENDIX S2. Concatenation tree showing relationships in Myrtales inferred using RAxML. Bootstrap values are presented below branches.

APPENDIX S3. Coalescent tree using supercontigs (exons + introns) showing relationships in Myrtales inferred using ASTRAL. Annotations above and below branches are respectively as described for Figs. 1A and 3.

APPENDIX S4. Key nodes and clades with ASTRAL quartet score (QS) values and pie charts, local posterior probability (LPP), polytomy test, and bootstrap (BP) support reported from the maximum likelihood (ML) RAxML concatenation analysis. Clades and nodes numbers correspond to Figs. 2, 3A, 3B. Pie charts above branches are as described for Fig. 1A. Color codes of each clade follow the quartet main topology (Q1) support value, with dark green ≥ 0.75; light green ≥50, <75; and yellow <50. For the polytomy test, branches with $p > 0.05$ are marked with a red asterisk.

APPENDIX S5. Summary of gene recovery statistics in Myrtales and for each family.

APPENDIX S6. Gene recovery statistics for all samples: recovery for Angiosperms353 and for the plastome. *Percentage based on target file from Johnson et al. (2016); **Gene = CDS + NCS; *** from the average length of the six plastomes used as the target file.

APPENDIX S7. Heatmap for recovery of the 353 low-copy nuclear genes. Each row shows a sample, and each column is a gene. Cell shadings correspond to the length of the gene recovered for each sample relative to the length of the reference.

APPENDIX S8. Heatmap for the recovery of the plastome.

APPENDIX S9. Number of nuclear genes recovered per sample. (A) Number of genes recovered per sample characterised by percentage of target length. (B) Whole number of genes per category. Percentage target length: green = >75%; yellow = 50% to 75%; red= 25% to 50%; blue = <25%.

APPENDIX S10. Alignment statistics. For all alignments, (1) exons and (2) supercontigs. Averages for (1) and (2) are provided in the first two rows.

APPENDIX S11. Principal coordinate analysis of Kendall–Colijn distance showing ordination of rooted topologies under different approaches. Six approaches were tested: V1, iteration before Treeshrink (standard); V2 ,iteration after TreeShrink; V3, as (V2), but excluding genes with less than a quarter of the all samples; V4, as (V3) with *p*-values of 0.0001; V5, as (V3) *p*-values of 0.001; V6, as (V3) *p*-values of 0.05.

LITERATURE CITED

Antonelli, A., J. J. Clarkson, K. Kainulainen, O. Maurin, G. E. Brewer, A. P. Davis, N. Epitawalage, et al. 2021. Settling a family feud: a high-level phylogenomic framework for the Gentianales based on 353 nuclear genes and partial plastomes. *American Journal of Botany* 108: 1143–1165.

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

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

Bacci, L. F., F. A. Michelangeli, and R. Goldenberg. 2019. Revisiting the classification of Melastomataceae: implications for habit and fruit evolution. *Botanical Journal of the Linnean Society* 190: 1–24.

Baker, W. J., P. Bailey, B. Barber, A. Barker, S. Bellot, D. Bishop, L. R. Botigué, et al. 2021. A comprehensive phylogenomic platform for exploring the angiosperm tree of life. *Systematic Biology* syab035. https://doi.org/10.1093/sysbio/syab035.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.

Baum, D. A., K. J. Sytsma, and P. C. Hoch. 1994. The phylogeny of Epilobium (Onagraceae) based on nuclear ribosomal DNA sequences. *Systematic Botany* 19: 363–388.

Bean, A. R. 1995. A revision of *Syncarpia* Ten. (Myrtaceae). *Austrobaileya* 4: 337–344.

Berg, O. 1859. Myrtaceae. *In* C. F. P. von Martius [ed.], Flora brasiliensis, vol. 14, 1–655. R. Oldenbourg, Munich, Germany.

Berger, B. A., R. Kriebel, D. Spalink, and K. J. Sytsma. 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Molecular Phylogenetics and Evolution* 95: 116–136.

Biffin, E., E. J. Lucas, L. A. Craven, I. Ribeiro da Costa, M. G. Harrington, and M. D. Crisp. 2010. Evolution of exceptional species richness among lineages of fleshy-fruited Myrtaceae. *Annals of Botany* 106: 79–93.

Bochorny, T., F. A. Michelangeli, F. Almeda, and R. Goldenberg. 2019. Phylogenetics, morphology and circumscription of Cambessedesieae: a new Neotropical tribe of Melastomataceae. *Botanical Journal of the Linnean Society* 190: 281–302.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Boon, G. C., M. Jordaan, and A. E. van Wick. 2020. A new species of *Combretum* sect. *Ciliatipetala* (Combretaceae) from South Africa. *Phytotaxa* 434: 1–12.

Borowiec, M. L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4: e1660.

Buerki, S., M. W. Callmander, P. Acevedo-Rodriguez, P. P. Lowry, J. Munzinger, P. Bailey, O. Maurin, et al. 2021. An updated infra-familial classification of Sapindaceae based on targeted enrichment data. *American Journal of Botany* 108: 1234–1251.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.

Cámara-Leret, R., D. G. Frodin, F. Adema, C. Anderson, M. S. Appelhans, G. Argent, S. A. Guerrero, et al. 2020. New Guinea has the world's richest island flora. *Nature* 584: 579–583.

Chernomor, O., A. von Haeseler, and B. Q. Minh. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* 65: 997–1008.

Clarkson, J. J., A. R. Zuntini, O. Maurin, S. R. Downie, G. M. Plunkett, A. A. Nicolas, J. F. Smith, et al. 2021. A higher-level nuclear phylogenomic study of the carrot family (Apiaceae). *American Journal of Botany* 108: 1252–1269.

Clausing, G., K. Meyer, and S. S. Renner. 2000. Correlations among fruit traits and evolution of different fruits within Melastomataceae. *Botanical Journal of the Linnean Society* 133: 203–226.

Clausing, G., and S. S. Renner. 2001. Molecular phylogenetics of Melastomataceae and Memecylaceae: implications for character evolution. *American Journal of Botany* 88: 486–498.

Cogniaux, C. A. 1891. Mélastomacées [Melastomaceae]. *In* A. De Candolle and C. De Candolle [eds.], Monographiae phanerogamarum, vol. 7. G. Masson, Paris, France.

Conti, E., T. Eriksson, J. Schönenberger, K. J. Sytsma, and D. A. Baum. 2002. Early tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. *Evolution* 56: 1931–1942.

Conti, E., A. Litt, and K. J. Sytsma. 1996. Circumscription of Myrtales and their relationships to other rosids: evidence From *rbcL* sequence data. *American Journal of Botany* 83: 221–233.

Conti, E., A. Litt, P. G. Wilson, S. A. Graham, B. G. Briggs, L. A. S. Johnson, and K. J. Sytsma. 1997. Interfamilial relationships in Myrtales: molecular phylogeny and patterns of morphological evolution. *Systematic Botany* 22: 629–647.

Conti, E., K. J. Sytsma, A. Litt, and S. A. Graham. 1994. Phylogenetic relationships of controversial families of Myrtales: evidence from *rbcL* sequence data. *American Journal of Botany* 81(Supplement 6): 149–150.

Craven, L. A. 1990. One new species each in *Acmena* and *Eucalyptopsis* and a new name in *Lindsayomyrtus* (all Myrtaceae). *Australian Systematic Botany* 3: 727–732.

Craven, L. A., and E. Biffin. 2010. An infrageneric classification of *Syzygium* (Myrtaceae). *Blumea* 55: 94–99.

Degnan, J. H., and N. A. Rosenberg. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genetics* 2: e68.

de la Estrella, M., S. Buerki, T. Vasconcelos, E. J. Lucas, and F. Forest. 2019. The role of Antarctica in biogeographical reconstruction: a point of view. *International Journal of Plant Sciences* 180: 63–71.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 19: 11–15.

Dupuis, J. R., Y. Pillon, T. Sakishima, C. E. Gemmill, S. Chamala, W. B. Barbazuk, S. M. Geib, and E. A. Stacy. 2019. Targeted amplicon sequencing of 40 nuclear genes supports a single introduction and rapid radiation of Hawaiian *Metrosideros* (Myrtaceae). *Plant Systematics and Evolution* 305: 961–974.

Flickinger, J. A., B. Jestrow, R. Oviedo Prieto, E. Santiago-Valentín, J. Sustache-Sustache, F. Jiménez-Rodríguez, K. C. S. E. Campbell, and J. Francisco-Ortega. 2020. A phylogenetic survey of Myrtaceae in the Greater Antilles with nomenclatural changes for some endemic species. *Taxon* 69: 448–480.

Flora do Brasil, 2020. Vochysiaceae. *In* Flora do Brasil 2020. Jardim Botânico do Rio de Janeiro. Website: http://floradobrasil.jbrj.gov.br/reflora/listaBrasil/ConsultaPublicaUC/ConsultaPublicaUC.do#CondicaoTaxonCP [accessed 24 January 2020].

Freyman, W. A., and S. Höhna. 2019. Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. *Systematic Biology* 68: 505–519.

Frodin, D. 2004. History and concepts of big plant genera. *Taxon* 53: 753–776.

Gere, J., K. Yessoufou, B. H. Daru, L. Mankga, O. Maurin, and M. van der Bank. 2013. Incorporating *trnH-psbA* to the core DNA barcodes improves significantly species discrimination within southern African Combretaceae. *ZooKeys* 365: 129–147.

Gere, J., K. Yessoufou, B. H. Daru, O. Maurin, and M. van der Bank. 2015. African continent a likely origin of family Combretaceae (Myrtales). A biogeographical view. *Annual Research and Review in Biology* 8: 1–20.

Goldenberg, R., F. Almeda, K. Sosa, R. C. Ribeiro, and F. A. Michelangeli. 2015. *Rupestrea*: A new brazilian genus of Melastomataceae, with anomalous seeds and dry indehiscent fruits. *Systematic Botany* 40: 561–571.

Goldenberg, R., C. N. de Fraga, A. P. Fontana, A. N. Nicolas, and F. A. Michelangeli. 2012. Taxonomy and phylogeny of *Merianthera* (Melastomataceae). *Taxon* 61: 1040–1056.

Goldenberg, R., D. S. Penneys, F. Almeda, W. S. Judd, and F. A. Michelangeli. 2008. Phylogeny of *Miconia* (Melastomataceae): initial insights into broad patterns of diversification in a megadiverse neotropical genus. *International Journal of Plant Sciences* 169: 963–979.

Gonçalves, D. J. P., G. H. Shimizu, E. M. Ortiz, R. K. Jansen, and B. B. Simpson. 2020. Historical biogeography of Vochysiaceae reveals an unexpected perspective of plant evolution in the Neotropics. *American Journal of Botany* 107: 1004–1020.

Gonçalves, D. J. P., B. B. Simpson, E. M. Ortiz, G. H. Shimizu, and R. K. Jansen. 2019. Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Molecular Phylogenetics and Evolution* 138: 219–232.

Govaerts, R., M. Sobral, P. Ashton, F. Barrie, B. K. Holst, L. R. Landrum, K. Matsumoto, et al. 2020. World checklist of Myrtaceae [online]. Facilitated by the Royal Botanic Gardens, Kew. Website www.kew.org/wcsp [accessed 1 October 2020].

Graham, S. A. 2010. Revision of the Caribbean genus *Ginoria* (Lythraceae), including *Haitia* from Hispaniola I. *Annals of the Missouri Botanical Garden* 97: 34–90.

Graham, S. A., M. Diazgranados, and J. Barber. 2011. Relationships among the confounding genera *Ammannia*, *Hionanthera*, *Nesaea* and *Rotala* (Lythraceae). *Botanical Journal of the Linnean Society* 166: 1–19.

Graham, S. A., and A. Graham. 2014. Ovary, fruit, and seed morphology of the Lythraceae. *International Journal of Plant Sciences* 175: 202–240.

Graham, S. A., J. Hall, K. Sytsma, and S. H. Shi. 2005. Phylogenetic analysis of the Lythraceae based on four gene regions and morphology. *International Journal of Plant Sciences* 166: 995–1017.

Gu, C., L. Ma, Z. Wu, K. Chen, and Y. Wang. 2019. Comparative analyses of chloroplast genomes from 22 Lythraceae species: inferences for phylogenetic relationships and genome evolution within Myrtales. *BMC Plant Biology* 19: 281.

Hale, H., E. M. Gardner, J. Viruel, L. Pokorny, and M. G. Johnson. 2020. Strategies for reducing per-sample costs in target capture sequencing for phylogenomics and population genomics in plants. *Applications in Plant Sciences* 8: e11337.

Heywood, V. H., R. K. Brummitt, A. Culham, and O. Seberg. 2007. Flowering plant families of the world. Firefly Books, Richmond Hill, ON, Canada.

Hollister, J. D., S. Greiner, M. T. Johnson, and S. I. Wright. 2019. Hybridization and a loss of sex shape genome-wide diversity and the origin of species in the evening primroses (*Oenothera*, Onagraceae). *New Phytologist* 224: 1372–1380.

Inglis, P. W., and T. B. Cavalcanti. 2018. A molecular phylogeny of the genus *Diplusodon* (Lythraceae), endemic to the campos rupestres and cerrados of South America. *Taxon* 67: 66–82.

Johnson, L. A. S., and B. G. Briggs. 1984. Myrtales and Myrtaceae— A phylogenetic analysis. *Annals of the Missouri Botanical Garden* 71: 700–756.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. Zerega, and N. J. Wickett. 2016. HybPiper: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications Plant Science* 4: apps.1600016.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigué, R. S. Cowan, A. Devault, W. L. Eiserhardt, et al. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using *k*-medoids clustering. *Systematic Biology* 68: 594–606.

Jombart, T., M. Kendall, J. Almagro-Garcia, and C. Colijn. 2017. Treespace: Statistical exploration of landscapes of phylogenetic trees. *Molecular Ecology Resources* 17: 1385–1392.

Junier, T., and E. M. Zdobnov. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26: 1669–1670.

Kadono, Y., and E. L. Schneider. 1986. Floral biology of *Trapa natans* var. *japonica*. *Botanical Magazine, Tokyo* 99: 435–439.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Kawasaki, M. L. 2007. Vochysiaceae. *In* K. Kubitzki [ed.], The families and genera of vascular plants, vol. IX, 480–487. Springer, Berlin, Germany.

Kendall, M., and C. Colijn. 2016. Mapping phylogenetic trees to reveal distinct patterns of evolution. *Molecular Biology and Evolution* 33: 2735–2743.

Koehne, E. 1903. Lythraceae. *In* A. Engler [ed.], Das Pflanzenreich, vol. IV, 1–326. Wilhelm Engelmann, Leipzig, Germany.

Kriebel, R., M. Khabbazian, and K. J. Sytsma. 2017. A continuous morphological approach to study the evolution of pollen in a phylogenetic context: an example with the order Myrtales. *PLoS One* 12: e0187228.

Kück, P., and K. Meusemann. 2010. FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution* 56: 1115–1118.

Levin, R. A., W. L. Wagner, P. C. Hoch, W. J. Hahn, A. Rodriguez, D. A. Baum, L. Katinas, et al. 2004. Paraphyly in tribe Onagreae: insights into phylogenetic relationships of Onagraceae based on nuclear and chloroplast sequence data. *Systematic Botany* 29: 147–164.

Levin, R. A., W. L. Wagner, P. C. Hoch, M. Nepokroeff, J. C. Pires, E. A. Zimmer, and K. J. Sytsma. 2003. Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. *American Journal of Botany* 90: 107–115.

Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589–595.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.

Li, H.-T., T.-S. Yi, L.-M. Gao, P.-F. Ma, T. Zhang, J.-H. Yang, M. A. Gitzendanner, et al. 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* 5: 461–470.

Linnaeus, C. 1753. Species plantarum, ed. 2, vol. 1. L. Salvius, Stockholm, Sweden.

Linkem, C. W., V. N. Minin, and A. D. Leaché. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Systematic Biology* 65: 465–477.

Litt, A. 1999. Floral morphology and phylogeny of Vochysiaceae. Ph.D. dissertation. City University of New York, NY, NY, USA.

Litt, A., and D. Stevenson. 2003. Floral development and morphology of Vochysiaceae. I. The structure of the gynoecium. *American Journal of Botany* 90: 1533–1547.

Liu, L., C. Anderson, D. Pearl, and S. V. Edwards. 2019. Modern phylogenomics: building phylogenetic trees using the multispecies coalescent model. *In* M. Anisimova [ed.], Evolutionary genomics, 2nd ed., 211–239. Humana, NY, NY, USA.

Lucas, E. J., B. S. Amorim, D. F. Lima, A. R. Lima-Lourenço, E. M. Nic Lughadha, C. E. B. Proença, P. O. Rosa, et al. 2018. A new infra-generic classification of the species-rich Neotropical genus *Myrcia* s.l. *Kew Bulletin* 73: 9.

Lucas, E. J., and M. O. Bünger. 2015. Myrtaceae in the Atlantic forest: their role as a 'model' group. *Biodiversity and Conservation* 24: 2165–2180.

Lucas, E. J., B. Holst, M. Sobral, F. Mazine, E. M. Nic Lughadha, C. Proença, I. da Costa, and T. Vasconcelos. 2019. A new subtribal classification of tribe Myrteae (Myrtaceae). *Systematic Botany* 44: 560–569.

Mai, U., and S. Mirarab. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* 19: 272.

Mancera, J. P. 2017. Morphological phylogenetic analysis of the Astronieae (Melastomataceae). Ph.D. dissertation, San Francisco State University, San Francisco, CA, USA.

Marcano-Berti, L. 2005. Vochysiaceae. *In* P. E. Berry, B. K. Holst, and K. Yatskievych [eds.], Flora of the Venezuelan Guayana, vol. 9, 500–524. Rutaceae - Zygophyllaceae. Missouri Botanical Garden Press, St. Louis, MO, USA.

Martin, C. V., D. Little, R. Goldenberg, and F. A. Michelangeli. 2008. A phylogenetic evaluation of *Leandra* (Miconieae, Melastomataceae): a polyphyletic genus where the seeds tell the story, not the petals. *Cladistics* 24: 317–327.

Maurin, O., M. W. Chase, M. Jordaan, and M. van der Bank. 2010. Combretaceae inferred from nuclear and plastid DNA sequence data: implications for generic classification. *Botanical Journal of the Linnean Society* 162: 453–476.

Maurin, O., G. Gere, M. van der Bank, and J. S. Boatwright. 2017. The inclusion of *Anogeissus*, *Buchenavia* and *Pteleopsis* in *Terminalia* (Combretaceae: Terminaliinae). *Botanical Journal of the Linnean Society* 184: 312–325.

Maurin, O., I. M. Turner, J. S. Boatwright, and M. J. M. Christenhusz. 2020. New combinations in Combretaceae subtribe Combretinae from Africa and Asia. *Phytotaxa* 451: 231–237.

Maurin, O., A. E. Van Wyk, M. Jordaan, and M. van der Bank. 2011. A new species of *Combretum* section *Ciliatipetala* (Combretaceae) from southern Africa, with a key to the regional members of the section. *South African Journal of Botany* 77: 105–111.

Maxwell, J. F. 1981. A revision of the genus *Pternandra* (Melastomataceae). *Garden Bulletin Singapore* 34: 1–90.

Mazerolle, M. J. 2019. AICcmodavg: model selection and multimodel inference based on (Q)AIC(c). R package version 2.2-2. Website: https://cran.rproject.org/package=AICcmodavg/.

Mazine, F. F., V. Castro Souza, M. Sobral, F. Forest, and E. Lucas. 2014. A preliminary phylogenetic analysis of *Eugenia* (Myrtaceae: Myrteae), with a focus on Neotropical species. *Kew Bulletin* 69: 9497.

Mazine, F. F., J. E. Q. Faria, A. Giaretta, T. Vasconcelos, F. Forest, and E. J. Lucas. 2018. Phylogeny and biogeography of the hyper-diverse genus *Eugenia* (Myrtaceae: Myrteae), with emphasis on *E.* sect. *Umbellatae*, the most unmanageable clade. *Taxon* 67: 752–769.

McLay, T. G. B., J. L. Birch, B. F. Gunn, W. Ning, J. A. Tate, L. Nauheimer, E. M. Joyce, et al. 2021. New targets acquired: improving locus recovery from the Angiosperms353 probe set. *Applications in Plant Sciences* 9: e11420.

Michelangeli, F. A., P. J. F. Guimaraes, D. S. Penneys, F. Almeda, and R. Kriebel. 2013. Phylogenetic relationships and distribution of New World Melastomeae (Melastomataceae). *Botanical Journal of the Linnean Society* 171: 38–60.

Michelangeli, F. A., A. Nicolas, M. E. Morales-Puentes, and H. David. 2011. Phylogenetic relationships of *Allomaieta*, *Alloneuron*, *Cyphostyla*, and *Wurdastom* (Melastomataceae) and the resurrection of the Tribe Cyphostyleae. *International Journal of Plant Sciences* 172: 1165–1178.

Michelangeli, F. A., D. S. Penneys, J. Giza, D. Soltis, M. H. Hils, and J. D. Skean Jr. 2004. A preliminary phylogeny of the tribe Miconieae (Melastomataceae) based on nrITS sequence data and its implications on inflorescence position. *Taxon* 53: 279–290.

Michelangeli, F. A., C. Ulloa Ulloa, and K. Sosa. 2014. *Quipuanthus*, a new genus and species of Melastomataceae from the foothills of the Andes in Ecuador and Peru. *Systematic Botany* 39: 533–540.

Miller, A. 2020. Leaps: regression subset selection. R package version 3.1. Website: https://CRAN.R-project.org/package=leaps.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534.

Mirarab, S. 2019. Species tree estimation using ASTRAL: practical considerations. *arXiv*. Website: https://arxiv.org/abs/1904.03826v2 [Preprint].

Mirarab, S., and T. Warnow. 2015. ASTRAL-II: Coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31: 44–52.

Morley, T. 1976. Memecyleae (Melastomataceae). *Flora Neotropica* 15: 1–295.

Murphy, B., F. Forest, T. Barraclough, J. Rosindell, S. Bellot, R. Cowan, and M. Golos, et al. 2020. A phylogenomic analysis of *Nepenthes* (Nepenthaceae). *Molecular Phylogenetics and Evolution* 144: 106668.

Naser-Khdour, S., B. Q. Minh, W. Zhang, E. A. Stone, and R. Lanfear. 2019. The prevalence and impact of model violations in phylogenetic analysis. *Genome Biology and Evolution* 11: 3341–3352.

Oliveira, P. E., P. E. Gibbs, and A. A. Barbosa. 2004. Moth pollination of woody species in the Cerrados of Central Brazil: a case of so much owed to so few? *Plant Systematics and Evolution* 245: 41–54.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Penneys, D. S., F. Almeda, F. A. Michelangeli, R. Goldenberg, A. B. Martins, and P. W. Fritsch. 2020. Lithobieae and Eriocnemeae: two new neotropical tribes of Melastomataceae. *Phytotaxa* 453: 157–178.

Penneys, D. S., and W. S. Judd. 2013. Combined molecular and morphological phylogenetic analyses of the Blakeeae (Melastomataceae). *International Journal of Plant Sciences* 174: 802–817.

Penneys, D. S., F. A. Michelangeli, W. S. Judd, and F. Almeda. 2010. Henriettieae: a new Neotropical tribe of berry-fruited Melastomataceae. *Systematic Botany* 35: 783–800.

Pillon, Y., H. C. F. Hopkins, O. Maurin, N. Epitawalage, J. Bradford, Z. S. Rogers, W. J. Baker, and F. Forest. 2021. Phylogenomics and biogeography of Cunoniaceae (Oxalidales) with complete generic sampling and taxonomic realignments. *American Journal of Botany* 108: 1181–1200.

Plaziat, J.-C., C. Cavagnetto, J.-C. Koeniguer, and F. Baltzer. 2001. History and biogeography of the mangrove ecosystem, based on a critical, reassessment of the paleontological. *Wetlands Ecology and Management* 9: 161–179.

POWO. 2020. Plants of the world online. Facilitated by the Royal Botanic Gardens, Kew. Website: http://www.plantsoftheworldonline.org/ [accessed 15 October 2020].

Proença, C. E., E. M. Nic Lughadha, E. J. Lucas, and E. Woodgyer. 2006. *Algrizea* (Myrteae, Myrtaceae): a new genus from the Highlands of Brazil. *Systematic Botany* 31: 320–326.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: https://www.R-project.org/.

RStudio Team. 2020. RStudio: integrated development for R. RStudio, PBC, Boston, MA, USA. Website: http://www.rstudio.com/.

Raven, P. H. 1964. The generic subdivision of Onagraceae, tribe Onagreae. *Brittonia* 276–288.

Reginato, M., and F. A. Michelangeli. 2016. Untangling the phylogeny of *Leandra* sensu str. (Melastomataceae, Miconieae). *Molecular Phylogenetics and Evolution* 96: 17–32.

Reginato, M., K. M. Neubig, L. C. Majure, and F. A. Michelangeli. 2016. The first complete plastid genomes of Melastomataceae are highly structurally conserved. *PeerJ* 4: e2715.

Reginato, M., T. N. C. Vasconcelos, R. Kriebel, and A. O. Simoes. 2020. Is dispersal mode a driver of diversification and geographical distribution in the tropical plant family Melastomataceae. *Molecular Phylogenetics and Evolution* 148: 106815.

Renner, S. S. 1993. Phylogeny and classification of the Melastomataceae and Memecylaceae. *Nordic Journal of Botany* 13: 519–540.

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

Revell, L. J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Rocha, M. J. R., J. A. N. Batista, P. J. F. Guimaraes, and F. A. Michelangeli. 2016. Phylogenetic relationships in the *Marcetia* alliance (Melastomeae, Melastomataceae) and implications for generic circumscription. *Botanical Journal of the Linnean Society* 181: 585–609.

Rocha, M. J. R., P. J. Guimarães, F. A. Michelangeli, and J. A. Nogueira Batista. 2018. Taxonomy of Marcetieae: A new Neotropical tribe of Melastomataceae. *International Journal of Plant Sciences* 179: 50–74.

Rutschmann, F., T. Eriksson, K. Abu Salim, and E. Conti. 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. *Systematic Biology* 56: 591–608.

Rutschmann, F., T. Eriksson, J. Schönenberger, and E. Conti. 2004. Did Crypteroniaceae really disperse out of India? Molecular dating evidence from *rbcL*, *ndhF*, and *rpl16* intron sequences. *International Journal of Plant Sciences* 165: S69–S83.

Rye, B. L., P. G. Wilson, M. M. Heslewood, A. J. Perkins, and K. R. Thiele. 2020. A new subtribal classification of Myrtaceae tribe Chamelaucieae. *Australian Systematic Botany* 33: 191–206.

Salywon, A. 2003. A monograph of *Mosiera* (Myrtaceae). Ph.D. dissertation, Department of Plant Biology, Arizona State University, Tempe, AZ, U.S.A.

Sayyari, E., and S. Mirarab. 2018. Testing for polytomies in phylogenetic species trees using quartet frequencies. *Genes* 9: 132.

Schönenberger, J., and E. Conti. 2003. Molecular phylogeny and floral evolution of Penaeaceae, Oliniaceae, Rhynchocalycaceae, and Alzateaceae (Myrtales). *American Journal of Botany* 90: 293–309.

Schuster, T. M., S. D. Setaro, J. F. G. Tibbits, E. L. Batty, R. M. Fowler, T. G. B. McLay, et al. 2018. Chloroplast variation is incongruent with classification of the Australian bloodwood eucalypts (genus *Corymbia*, family Myrtaceae). *PLoS ONE* 13: e0195034.

Shah, T., J. V. Schneider, G. Zizka, O. Maurin, W. Baker, F. Forest, G. E. Brewer, et al. 2021. Joining forces in Ochnaceae phylogenomics: a tale of two targeted sequencing probe kits. *American Journal of Botany* 108: 1201–1216.

Shimizu, G. H., and D. J. P. Gonçalves. 2017. Flora das cangas da Serra dos Carajás, Pará, Brasil: Vochysiaceae. *Rodriguésia* 68: 1159–1164.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210–3212.

Slater, G. S., and E. Birney. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31.

Smith, S. A., N. Walker-Hale, J. F. Walker, and J. W. Brown. 2020. Phylogenetic conflicts, combinability, and deep phylogenomics in plants. *Systematic Biology* 69: 579–592.

Soto Gomez, M., L. Pokorny, M. B. Kantar, F. Forest, I. L. Leitch, B. Gravendeel, P. Wilkin, et al. 2019. A customized nuclear target enrichment approach for developing a phylogenomic baseline for *Dioscorea* yams (Dioscoreaceae). *Applications in Plant Sciences* 7: e11254.

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

Stone, R. 2006. Phylogeny of major lineages in Melastomataceae, subfamily Olisbeoideae: utility of nuclear glyceraldehyde 3-phosphate dehydrogenase (*GapC*) gene sequences. *Systematic Botany* 31: 107–121.

Stone, R. D. 2014. The species-rich, paleotropical genus *Memecylon* (Melastomataceae): molecular phylogenetics and revised infrageneric classification of the African species. *Taxon* 63: 539–561.

Stull, G. W., P. S. Soltis, D. E. Soltis, M. A. Gitzendanner, and S. A. Smith. 2020. Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany* 107: 790–805.

Sun, M., R. Naaem, J.-X. Su, Z.-Y. Cao, J. G. Burleigh, P. S. Soltis, D. E. Soltis, and Z.-D. Chen. 2016. Phylogeny of the Rosidae: A dense taxon sampling analysis. *Journal of Systematics and Evolution* 54: 363–391.

Susko, E., and A. J. Roger. 2021. Long branch attraction biases in phylogenetics. *Systematic Biology* 70: 838–843.

Sytsma, K. J., A. Litt, M. L. Zjhra, J. C. Pires, M. Nepokroeff, E. Conti, J. Walker, and P. G. Wilson. 2004. Clades, clocks, and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the Southern Hemisphere. *International Journal of Plant Sciences* 165: S85–S105.

Tan, F., S. Shi, Y. Zhong, X. Gong, and Y. Wang. 2002. Phylogenetic relationships of Combretoideae (Combretaceae) inferred from plastid, nuclear gene and spacer sequences. *Journal of Plant Research* 115: 475–481.

Thornhill, A. H., M. Crisp, C. Külheim, L. Carsten, N. Kristy, L. Nelson, D. Yeates, and J. Miller. 2019. A dated molecular perspective of eucalypt taxonomy, evolution and diversification. *Australian Systematic Botany* 32: 29–48.

Thornhill, A. H., S. Y. W. Ho, C. Külheim, and M. D. Crisp. 2015. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution* 93: 29–43.

Tobe, H., S. A. Graham, and P. H. Raven. 1998. Floral morphology and evolution in Lythraceae sensu lato. *In* S. J. Owens, and P. J. Rudall [eds.], Reproductive biology in systematics conservation and economic botany, 329–344. Royal Botanic Gardens, Kew, UK.

Tobe, H., and P. H. Raven. 1983. An embryological analysis of Myrtales: its definition and characteristics. *Annals of the Missouri Botanical Garden* 70: 71–94.

Tobe, H., and P. H. Raven. 1984a. The embryology and relationships of *Alzatea* Ruiz & Pav. (Alzateaceae, Myrtales). *Annals of the Missouri Botanical Garden* 71: 844–852.

Tobe, H., and P. H. Raven. 1984b. The embryology and relationships of Penaeaceae (Myrtales). *Plant Systematics and Evolution* 146: 181–195.

Tobe, H., and P. H. Raven. 1984c. The embryology and relationships of *Rhynchocalyx* Oliv. (Rhynchocalycaceae). *Annals of the Missouri Botanical Garden* 7: 836–843.

Tobe, H., and P. H. Raven. 1987. The embryology and relationships of *Dactylocladus* (Crypteroniaceae) and a discussion of the family. *Botanical Gazette* 148: 103–111.

Van Beusekom-Osinga, R. J., and C. F. van Beusekom. 1975. Delimitation and subdivision of the Crypteroniaceae (Myrtales). *Blumea* 22: 255–266.

van der Merwe, M. M., A. E. van Wyk, and A. M. Botha. 2005. Molecular phylogenetic analysis of *Eugenia* L. (Myrtaceae), with emphasis on southern African taxa. *Plant Systematics and Evolution* 251: 21–34.

van Vliet, G. J. C. M. 1981. Wood anatomy of the palaeotropical Melastomataceae. *Blumea* 27: 395–462.

van Vliet, G. J. C. M., J. Keok-Noorman, and B. J. H. ter Welle. 1981. Wood anatomy, classification and phylogeny of the Melastomataceae. *Blumea* 27: 463–473.

Vasconcelos, T. N., C. G. Prenner, and E. J. Lucas. 2019. A systematic overview of the floral diversity in Myrteae (Myrtaceae). *Systematic Botany* 44: 570–591.

Vasconcelos, T. N., C. E. Proença, B. Ahmad, D. S. Aguilar, R. Aguilar, B. S. Amorim, K. Campbell, et al. 2017. Myrteae phylogeny, calibration, biogeography and diversification patterns: increased understanding in the most species rich tribe of Myrtaceae. *Molecular Phylogenetics and Evolution* 109: 113–137.

Veranso-Libalah, M. C., G. Kadereit, R. D. Stone, and T. L. P. Couvreur. 2018. Multiple shifts to open habitats in Melastomateae (Melastomataceae) congruent with the increase of African Neogene climatic aridity. *Journal of Biogeography* 45: 1420–1431.

Veranso-Libalah, M. C., R. D. Stone, A. G. N. Fongod, T. L. P. Couvreur, and G. Kadereit. 2017. Phylogeny and systematics of African Melastomateae (Melastomataceae). *Taxon* 66: 584–614.

Veranso-Libalah, M. C., R. D. Stone, and G. Kadereit. 2020. Towards a complete phylogeny of African Melastomateae: Systematics of *Dissotis* and allies (Melastomataceae). *Taxon* 69: 946–991.

Wagner, W. L., and P. C. Hoch. 2009. Nomenclatural corrections in Onagraceae. *Novon: A Journal for Botanical Nomenclature* 19: 130–132.

Wagner, W. L., P. C. Hoch, and P. H. Raven. 2007. Revised classification of the Onagraceae. *Systematic Botany Monographs* 83: 1–240.

Wang, L. G., T. T. Y. Lam, S. Xu, Z. Dai, L. Zhou, T. Feng, P. Guo, C. W. Dunn, et al. 2020. treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. *Molecular Biology and Evolution* 37: 599–603.

Wang, X. Q., W. W. Song, and J. J. Xiao. 2021. Phylogeny of Myrtales and related groups based on chloroplast genome. *Guihaia Plants* 41: 68–80.

Webb, D. A. 1967. Generic limits in European Lythraceae. *Feddes Repertorium* 74: 1–13.

Weitemier, K., S. Straub, R. C. Cronn, and M. Fishbein. 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences* 2: 1400042.

White, C. T. 1942. Contributions to the Queensland Flora 7. *Proceedings of the Royal Society of Queensland* 53: 219.

Wilson, P. G. 2010. Myrtaceae. *In* K. Kubitzki [ed.], The families and genera of vascular plants, vol. X, Flowering plants. Eudicots, 212–271. Springer, Berlin, Germany.

Wilson, P. G., M. M. O'Brien, M. M. Heslewood, and C. J. Quinn. 2005. Relationships within Myrtaceae sensu lato based on a *mat*K phylogeny. *Plant Systematics and Evolution* 251: 3–19.

Wilson, P. G., and J. T. Waterhouse. 1982. A review of the genus *Tristania* R. Br. (Myrtaceae): a heterogeneous assemblage of five genera. *Australian Journal of Botany* 30: 413–446.

Wurdack, K. J., and F. A. Michelangeli. 2019. Systematics and relationships of *Tryssophyton* (Melastomataceae), with a second species from the Pakaraima Mountains of Guyana. *Phytokeys* 136: 1–21.

Yu, G. 2019. ggimage: Use Image in 'ggplot2'. R package version 0.2.1. Website: https://github.com/GuangchuangYu/ggimage.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T. Y. Lam. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8: 28–36.

Zhang, X. F., J. B. Landis, H. X. Wang, Z. X. Zhu, and H. F. Wang. 2021. Comparative analysis of chloroplast genome structure and molecular dating in Myrtales. *BMC Plant Biology* 21: 1–19.

Zhou, Q., C.-W. Lin, W. L. Ng, J. Dai, T. Denda, R. Zhou, and Y. Liu. 2019. Analyses of plastome sequences improve phylogenetic resolution and provide new insight into the evolutionary history of Asian Sonerileae/Dissochaeteae. *Frontiers in Plant Science* 10: 1477.

Zich, F. A., B. P. M. Hyland, T. Whiffin, and R. A. Kerrigan. 2018. Australian tropical rainforest plants, ed. 7. Website: http://www.anbg.gov.au/cpbr/cd-keys/RFK7/key/RFK7/Media/Html/index_rfk.htm/ [accessed 1 October 2020].

Zuntini, A. R., L. P. Frankel, L. Pokorny, F. Forest, and W. J. Baker. 2021. A comprehensive phylogenomic study of the monocot order Commelinales with a new classification of Commelinaceae. *American Journal of Botany* 108: 1066–1086.