
















RESEARCH ARTICLE

Towards a species-level phylogeny for Neotropical Myrtaceae: Notes on topology and resources for future studies

The Neotropical Myrtaceae Working Group | Vanessa G. Staggemeier¹  |
 Bruno Amorim²  | Mariana Bünger³  | Itayguara R. Costa³  |
 Jair Eustáquio Quintino de Faria⁴ | Jonathan Flickinger⁵  | Augusto Giaretta⁶ |
 Marcelo T. Kubo⁷ | Duane Fernandes Lima⁸  | Leidiana Lima dos Santos⁹ |
 Ana Raquel Lourenço¹⁰ | Eve Lucas¹¹  | Fiorella Fernanda Mazine¹²  |
 José Murillo-A¹³  | Marla Ibrahim Uehbe de Oliveira¹⁴ | Carlos Parra-O¹³ |
 Carolyn E. B. Proença¹⁵  | Marcelo Reginato¹⁶ | Priscila Oliveira Rosa¹⁷  |
 Matheus Fortes Santos¹⁸ | Aline Stadnik^{4,19}  | Amélia Carlos Tuler²⁰  |
 Karinne Sampaio Valdemarin²¹  | Thais Vasconcelos²² 

Correspondence

Thais Vasconcelos, Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA.
 Email: tvasc@umich.edu

Abstract

Premise: Increasingly complete phylogenies underpin studies in systematics, ecology, and evolution. Myrteae (Myrtaceae), with ~2700 species, is a key component of the exceptionally diverse Neotropical flora, but given its complicated taxonomy, automated assembling of molecular supermatrices from public databases often lead to unreliable topologies due to poor species identification.

Methods: Here, we build a taxonomically verified molecular supermatrix of Neotropical Myrteae by assembling 3909 published and 1004 unpublished sequences from two nuclear and seven plastid molecular markers. We infer a time-calibrated phylogenetic tree that covers 712 species of Myrteae (~28% of the total diversity in the clade) and evaluate geographic and taxonomic gaps in sampling.

Results: The tree inferred from the fully concatenated matrix mostly reflects the topology of the plastid data set and there is a moderate to strong incongruence between trees inferred from nuclear and plastid partitions. Large, species-rich genera are still the poorest sampled within the group. Eastern South America is the best-represented area in proportion to its species diversity, while Western Amazon, Mesoamerica, and the Caribbean are the least represented.

Conclusions: We provide a time-calibrated tree that can be more reliably used to address finer-scale eco-evolutionary questions that involve this group in the Neotropics. Gaps to be filled by future studies include improving representation of taxa and areas that remain poorly sampled, investigating causes of conflict between nuclear and plastid partitions, and the role of hybridization and incomplete lineage sorting in relationships that are poorly supported.

KEYWORDS

chronogram, Eugenia, macroecology, macroevolution, Myrcia, Myrteae, Psidium, tropical botany

For affiliations refer to page 15.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *American Journal of Botany* published by Wiley Periodicals LLC on behalf of Botanical Society of America.

The exceptionally diverse Neotropical flora accounts for more species of flowering plants than tropical Africa and Asia combined (Antonelli and Sanmartín, 2011). One of the dominant components in these environments are the myrtles (Myrtaceae), a group of trees and shrubs of outstanding diversity and abundance in both neotropical rainforests and savannas (Mori, 1983; Oliveira-Filho and Fontes, 2000; Ibisch et al., 2002; Zipparro et al., 2005; Mendonça et al., 2008; Joly et al., 2012; Santos et al., 2012; Françoise et al., 2016; Staggemeier et al., 2017). Besides comprising a significant part of the biomass that forms these ecosystems, Myrtaceae species are an important source of floral resources and fleshy-fruits, which are produced throughout all seasons of the year (Staggemeier et al., 2010, 2015a, 2017), to pollinators and frugivores in the region (Gressler et al., 2006; Valadão et al., 2019; Martins et al., 2023). Some species are also a promising source of biologically active compounds with medicinal properties such as antimicrobial, antilarval, antioxidant, anti-inflammatory, and cytotoxic (Stefanello et al., 2011; Cascaes et al., 2015). Widespread exploitation of fruits of economic value, such as guava (*Psidium guajava* L.), pitanga (*Eugenia uniflora* L.), and jaboricaba (*Plinia peruviana* (Poir.) Govaerts), make Myrtaceae one of the most used plant families by rural populations showing their high socio-biodiversity value in Latin America (Souza et al., 2018; Silva et al., 2022). Because several recent studies focused on solving the systematics of the family (Mazine et al., 2014, 2016, 2018; Lucas et al., 2007, 2011, 2018, 2019), Neotropical Myrtaceae have been also elevated from the status of a neglected group to a potential model system for addressing broad questions regarding the dynamics of Neotropical biodiversity (Giaretta et al., 2015; Lucas and Bünge, 2015; Staggemeier et al., 2015b; Vasconcelos et al., 2019b).

In the center of this process is the inference of increasingly robust and complete phylogenetic trees for the tribe Myrteae, the most diverse clade in the family with ~2700 species and ~50 genera (Figure 1; Lucas et al., 2019). Although Myrteae is a pantropical clade, most of its species diversity is found in the New World where about ~2200 species occur. Conversely, all species of Myrtaceae native to the New World are part of tribe Myrteae, with the exception of *Metrosideros stipularis* (Hook. & Arn.) Hook.f. in the tribe Metrosidereae, a species native to Chile. The main morphological traits characterizing Myrteae include opposite leaves with abundant presence of oil glands, brochidodromous venation, inferior ovaries, and fleshy berries (Wilson, 2011). Myrteae first became the subject of molecular systematics studies in the 2000s with the works of Wilson et al. (2005) and Lucas et al. (2005, 2007), and since then, many other studies tackling individual clades and geographic areas (Lucas et al., 2011; Murillo-A et al., 2012; Mazine et al., 2014; Staggemeier et al., 2015b; Bünge et al., 2016; Santos et al., 2016; Wilson et al., 2016; Amorim et al., 2019; Flickinger et al., 2020; Lima et al., 2021; Santos et al., 2021) or reassessing the

tribe's topology in light of a wider species and molecular sampling (e.g., Lucas et al., 2007; Vasconcelos et al., 2017a) have been published (for a complete review see Souza Neto et al., 2022). These studies have provided data and evidence to justify several taxonomic rearrangements (e.g., Proença et al., 2020) and consistent frameworks to shape and test biogeographical, ecological, and evolutionary hypotheses (e.g., Staggemeier et al. 2010, 2015b, 2015a; Bünge et al., 2016; Santos et al., 2017; Amorim et al., 2019; Vasconcelos et al., 2017a, 2018, 2019a).

In spite of these advances, it is not uncommon that larger phylogenetic trees built for the purpose of exploring eco-evolutionary questions (e.g., Seger et al., 2013; Smith and Brown, 2018) contain taxonomic mistakes in this group, for instance by using invalid or outdated names, potentially leading to topologies that differ strongly from more complete and carefully assembled data sets. The reason for this is that automated methods of tree inference (e.g., Bennett et al., 2018; Smith and Walker, 2019) often source taxonomic information from public databases of sequences and names that have not been updated at the same pace as the Myrteae taxonomy. Because of poor identification of samples and outdated taxonomy, it is not uncommon that sequences from different vouchers, which are often different species, are assembled together as part of a single tip in tree building. This issue should also affect the interpretation of all downstream macro-evolutionary and macroecological analyses that use phylogenies, for instance by assigning incorrect geographical ranges and morphological traits to tips in the tree based on the wrong taxon name (see also Bortolus, 2008). A taxonomically verified molecular matrix—including all markers that have been most frequently used to infer phylogenies in the group—would likely solve these issues and would be an important resource for future finer-scale studies. Much of the molecular data required to produce a satisfactory result in this sense have already been produced, but is either scattered throughout several parallel studies or still unpublished. In this sense, coordinating data sharing to build a broad, inclusive, and taxonomically reliable molecular matrix is the most effective way to produce a robust species-level tree for Neotropical Myrtaceae and to make this resource widely available for future studies.

Here we retrace the recent developments in the phylogenetics of the Neotropical clade (*sensu* Vasconcelos et al., 2017b) of tribe Myrteae, a group that is composed all the Neotropical species within the tribe, to combine molecular data generated from studies published in the last 15 years plus ~1000 unpublished sequences. We further reassessed identification of vouchers to increase taxonomic accuracy and explore the support and consistency of topologies inferred from different molecular markers. We also identify taxonomic and geographical gaps in sampling to be targeted for

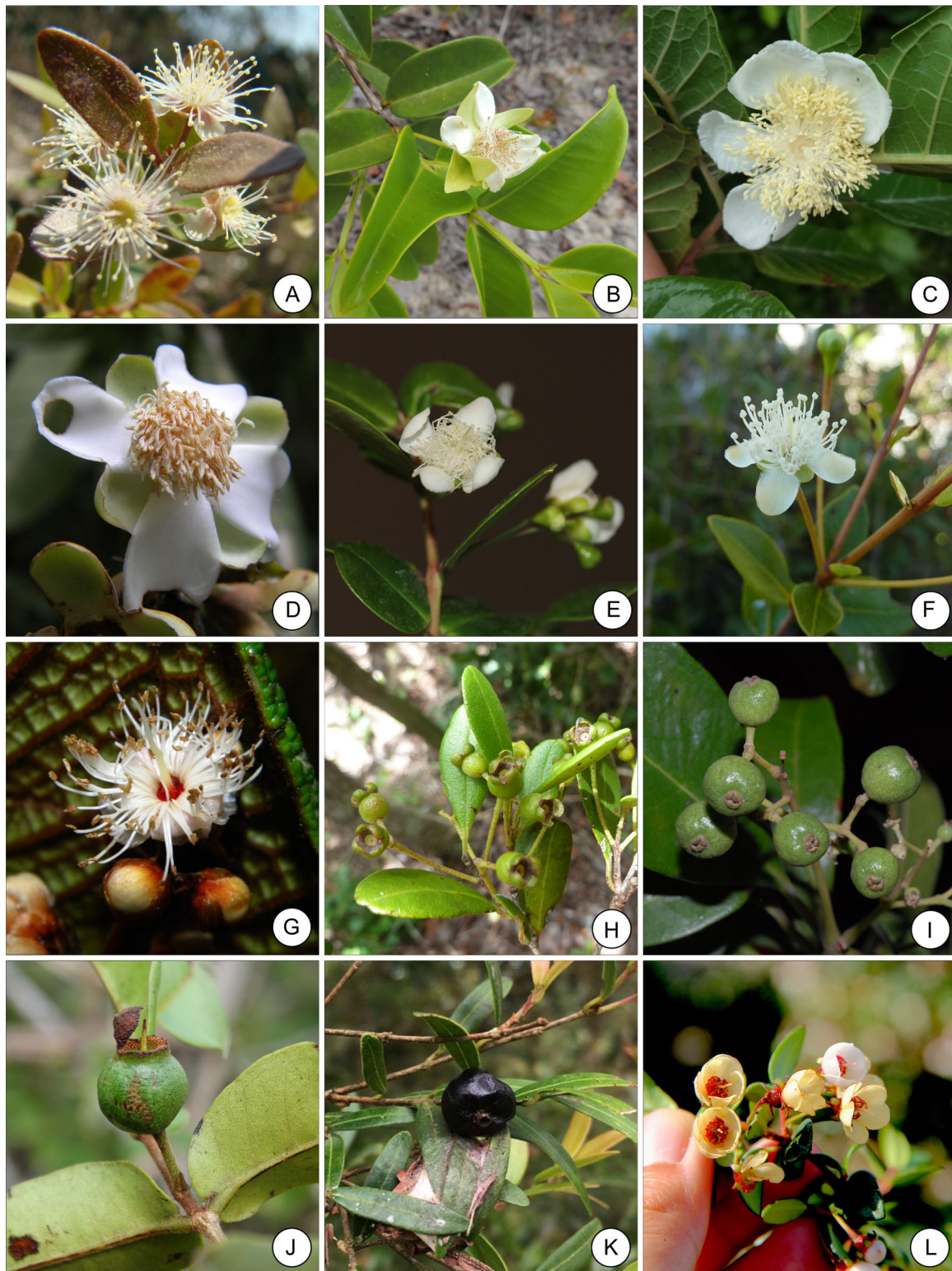


FIGURE 1 Diversity of flowers and fruits in Myrteae (Myrtaceae). (A) *Blepharocalyx salicifolius* (Kunth) O.Berg (Blepharocalycinae), (B) *Calycolpus legrandii* Mattos (Myrtinae), *Campomanesia schlechtendaliana* (O.Berg) Nied. (Pimentinae), *Eugenia quilombola* B.S.Amorim, M.A.D.Souza & Giaretta (Eugeniinae), *Luma apiculata* (DC.) Burret (Luminae), *Mosiera longipes* (O.Berg) Small (Pimentinae), *Myrcia ruschii* B.S.Amorim (Myrciinae), *Myrcianthes fragrans* (Sw.) McVaugh (Eugeniinae), *Pimenta dioica* (L.) Merr. (Pimentinae), *Psidium brownianum* Mart. ex DC. (Pimentinae), *Siphoneugena reitzii* D.Legrand (Pliniinae), *Ugni candollei* (Barnéoud) O.Berg (Ugninae). Photo credits: Aline Stadnik (A), Marla Ibrahim Uehbe de Oliveira (B, C), Bruno Amorim (D, G), Thais Vasconcelos (E, I, L), Jonathan Flickinger (F, H), Amélia Carlos Tuler (J), Vanessa G. Staggemeier (K).

sequencing and suggest guidelines to improve phylogenetic resolution in future studies. Finally, we provide a densely sampled time-calibrated tree that can be more reliably used to explore macroecological and macroevolutionary questions in the Neotropics. These include, but are not limited to, which biogeographical processes were responsible for shaping gradients of diversity observed between modern ecosystems in the region (e.g., Antonelli et al., 2018), how plant and animal interactions evolved over time (e.g., Dellinger et al., 2019; Reginato et al., 2020), and what are the biotic and abiotic factors correlated with spatial distribution of different functional traits (e.g., Vasconcelos et al., 2023). There are three specific aims in this study: (1) to infer a large and, most importantly, taxonomically verified species-level phylogeny of Myrteae; (2) to investigate how to improve species sampling and support for relationships and consistency of topologies inferred from different markers and where to focus resources for future studies; and (3) to provide a time-calibrated phylogenetic tree with trustworthy species identification that can be used for eco-evolutionary studies in the Neotropics.

MATERIALS AND METHODS

Taxon sampling

The Neotropical clade of tribe Myrteae, henceforward Neotropical Myrteae, is an informal group that encompasses nine subtribes (sensu Lucas et al., 2019): Blepharocalycinae (three species), Eugeniinae (~1200 species), Luminae (~50 species), Myrciinae (~750 species), Myrtinae (~20 species), Pimentinae (~200 species), Pliniinae (~120 species), and Ugninae (12 species). This clade includes ~2500 species in total (sensu POWO, 2023) and is informally named Neotropical Myrteae because ~2200 of its species are Neotropical. *Eugenia* sect. *Jossinia* (DC.) Nied., a clade of over 200 species within the Eugeniinae, is the largest non-Neotropical radiation within this clade, but it is nested within a clade composed of all Neotropical *Eugenia* L. (Mazine et al., 2016, 2018) and thus is also included as part of the ingroup in our analysis to preserve the monophyly of the group. *Myrtus communis* L. within Myrtinae and *Lophomyrtus* Burret (two species) and *Lenwebbia* N.Snow & Guymer (two species) within Ugninae are the other non-Neotropical representatives of this group, and were also included as part of the ingroup because they are also nested within clades predominantly distributed in the Neotropical region (Vasconcelos et al., 2017b). To reconstruct the most species-inclusive phylogenetic hypothesis for Neotropical Myrteae, DNA sequences were retrieved from two sources: (1) by data-mining GenBank, and (2) by compiling unpublished sequences generated by the coauthors of this study. Each approach is described below.

Data-mining GenBank

We searched and downloaded all nucleotide entries for each of the 52 Myrteae genera recognized by Wilson (2011) in February 2023. Given the problematic nature of Myrteae's taxonomy and to guarantee that all vouchers included in the tree can be traced back to an herbarium for future consultation, several steps of manual curation had to be performed after downloading all sequences available on GenBank. First, the preliminary list of sequences was filtered to remove those not belonging to Myrteae (e.g., sequences from pathogens and parasites that were also occasionally recovered by the search engine). The remaining sequences were classified by subtribe (sensu Lucas et al., 2019) and organized in distinct folders in Geneious v. 11 (<https://www.geneious.com>). Sequences belonging to the subtribe Decaspermia, a non-Neotropical clade sister to all the remaining extant Myrteae, were also removed except for those of 25 species selected to be part of the outgroup in the phylogenetic inference. Thirty-six species belonging to other Myrtaceae tribes were also searched and included in the data set for the same purpose. Next, we filtered these results to exclude all but nine molecular markers that showed the greatest coverage among all species: the nuclear regions internal transcribed spacer (ITS) and external transcribed spacer (ETS) and the plastid regions *matK*, *ndhF*, *psbA-trnH*, *rpl16*, *rpl32-trnL*, *trnL-trnF*, *trnQ-rps16*. The remaining sequences were renamed as “species_voucher” and reorganized in different folders for individual markers. Sequences for which no voucher information was found, or that could not be easily traced back to an herbarium specimen, were also excluded in this step. When there were two sequences for the same molecular region and voucher, we kept the one with the longest length. We kept duplicate entries from the same species as long as they belonged to different vouchers. This filtering process resulted in 3909 sequences recovered from GenBank, in a data set that is restricted to specimens for which the vouchers are known and can be traced back to an herbarium specimen or living collection.

Previously unpublished data—extraction and sequencing protocols

Previously unpublished molecular data contributed by coauthors of this study account for 1004 sequences of the same nine molecular markers targeted in our GenBank search. DNA extraction for most samples was carried out from silica-dried leaf material using the cetyltrimethylammonium bromide (CTAB) extraction protocol (Doyle and Doyle, 1987) and left for precipitation under conditions of -18°C in 100% ethanol followed by a purification by equilibrium centrifugation in cesium chloride (CsCl)-ethidium bromide gradients ($1.55\text{ g}\times\text{mL}^{-1}$). Butanol

extraction followed by dialysis was employed to remove the ethidium bromide and caesium chloride. QIAGEN DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) was also employed for extracting DNA, following the manufacturer's instructions. The target regions were amplified on the GeneAmp PCR System 9700 (Applied Biosystems, Foster City, California, U.S.A.) and Mastercycler nexus (Eppendorf, Hamburg, Germany). The PCR products were purified using QIAGEN QIAquick PCR Purification Kit according to the manufacturer's protocol. Sequencing reactions were carried out with the Taq DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems Inc, Waltham, Massachusetts). Sequences were read on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems Inc, Waltham, Massachusetts). Primers, PCR, and sequencing conditions are detailed in the Supplemental Information, Appendix S1: Tables S1, S2. All previously unpublished sequences have been deposited in GenBank; DNA samples have been deposited in the Royal Botanic Gardens Kew DNA and Tissue Collections.

Taxonomic verification and alignment

In the next step, all species names and vouchers included in the data set were circulated among all coauthors to have their taxonomic identities checked and, when required, to update their identifications prior to tree inference analyses. That also included vouchers marked as “sp.” (unidentified species) from both GenBank and unpublished data, enabling new identifications to be included. This procedure of manual reassessment of all data was time-consuming, but is still the most effective way to address the problem of taxonomic inaccuracy that is widespread in Neotropical Myrteae. We then ran alignments for each molecular region separately using the Muscle algorithm (Edgar, 2004) implemented in Geneious (<https://www.geneious.com>), with default settings. Alignments were visually inspected and adjusted for issues with sequence inversions. Sequences that aligned poorly or that were too short (<200 bp) were excluded from the matrix. At the end of this process, the resultant alignment was considered cleaned and was used in all the subsequent analyses. Voucher information and GenBank numbers for all sequences utilized are available in Appendix S2.

Tree inference, analyses of support, and time calibration

We first inferred a phylogenetic tree including all the molecular data available in our data set, composed of a concatenated matrix of all nine molecular markers and all tips (henceforward “the supermatrix tree”). This tree was used to contrast species positions against previously published trees using smaller data sets and to observe general patterns in support and topology among the

eight subtribes focused on here. To this end, we used a maximum likelihood (ML) algorithm implemented in RAxML-HPC v.8 on XSEDE v.8.2.12 (Stamatakis, 2006) through CIPRES (Miller et al., 2010), setting it to 1000 bootstrap replications. To investigate how consistent the topologies inferred from different markers are, we filtered three subset matrices complete for all molecular markers: (1) a subset of 53 tips for which all nine markers were available (henceforward the “all53” set); (2) a subset of 169 tips representing the Myrciinae subtribe (including the large genus *Myrcia* DC.) for which five markers were sampled (i.e., nuclear ITS and the plastids *ndhF*, *psbA-trnH*, *trnL-trnF*, and *trnQ-rps16*); and (3) a subset of 197 tips for which five markers frequently used in studies of the large genus *Eugenia* were sampled (the nuclear ITS and the plastids *psbA-trnH*, *rpl16*, and *rpl32-trnL* and *trnQ-rps16*). Using these subsets, we ran separate ML analyses for each different marker in each subset. We also ran ML analyses for combining partitions from different organelles (“nuclear” and “plastid”) and one analysis for the full-concatenated matrix in each subset (i.e., “full” analysis). All ML analyses were run using the same software and settings as described for the supermatrix tree. Nodes recovered with bootstrap (BS) values over 75 were considered strongly supported, between 50 and 75 were considered moderately supported, and below 50 were weakly supported in all cases.

To compare topologies and supports of each marker and combined partition against the full analysis, we used three approaches. First, we extracted all bootstrap values in each tree and compared their median, highlighting which marker inferred trees with highest general support. Then, we estimated the similarity between the tree topology inferred by each marker against the full analysis by extracting a matrix of distances based on the covariance among tips, and calculating Mantel statistics between pairs of matrices. The post hoc significance values were estimated following the Pearson method and 1000 permutations, using functions of the R packages ‘ape’ (Paradis et al., 2004) and ‘vegan’ (Oksanen et al., 2007). Values for Mantel correlation ranged from 0 (topologies are completely different) to 1 (topologies are completely similar). Last, we used functions of the R package ‘treespace’ (Jombart et al., 2017) to simultaneously explore resolution and topology inferred from different markers and partitions, using a sample of 100 trees randomly extracted from the set of bootstrap replications of each ML analysis. This analysis is similar to a multidimensional scaling, and visually depicts how different topologies are (i.e., how much they overlap in the multidimensional tree space) and how robust the support from each inference is (i.e., the larger the area, the lower the support).

Finally, we produced a time-calibrated tree using the concatenated supermatrix alignment considering two secondary calibration points as well as three fossil constraints

based on pollen fossil data. Prior distribution and settings for each calibration point followed “Approach B” in Vasconcelos et al. (2017b) (see details in Appendix S1: Table S3). The time calibration was performed in BEAST v.2.7.3 (Bouckaert et al., 2014) through CIPRES. The clock model was selected as “Uncorrelated relaxed clock” and tree prior was selected as “Speciation: birth-death incomplete sampling”. Nucleotide substitution models were set for nuclear and plastid partitions separately. Best substitution models for each partition were selected based on Akaike information criterion (AIC) weight using the modelTest function of the R package ‘phangorn’ (Schliep, 2011), resulting in GTR+I+G model for both nuclear and plastid partitions. Because of computational limitations in running BEAST with data sets of this size, we constrained the topology of the calibrated tree to be the same as that recovered in the RAxML analysis. The Markov Chain Monte Carlo (MCMC) simulation was set for 200,000,000 generations, sampling every 10,000th generation. Convergence was assessed using Tracer v.1.7.2 and accepted when the effective sampling size (ESS) was around or above 200 for all parameters. The resultant time-calibrated tree was further pruned to remove outgroups and to leave only one tip per species, resulting in a coverage of 712 species of Neotropical Myrteae (~28% of the group). The representative voucher for species with more than one entry in the original tree was selected based on the voucher with most molecular data or which best represents the morphological diagnosis of the species, particularly in cases of species that appear nonmonophyletic in the ML tree. All alignments and trees produced are available in Appendix S3–S8. Codes used for the analyses above are available at https://github.com/tncvasconcelos/myrteae_tree.

Identifying taxonomical and geographical gaps

To inspect how balanced the final pruned tree is in terms of taxonomic and geographical coverage, we estimated the sampling proportion of species in each subtribe and for each Level-3 Botanical Country based on the categorization of the World Geographical Scheme for Recording Plant Distributions (Brummitt et al., 2001) using data from POWO (2023). This information was used to discuss which clades and geographical areas are particularly over- or undersampled and where to invest resources for future molecular sampling. For discussion on geographical coverage, we focused on the Neotropics because this is the center of diversity in the group with ~90% of the species and because we noticed some taxonomic inaccuracies with paleotropical species of *Eugenia* in the POWO data set. This means that mapping sampling gaps for these taxa may lead to inaccurate results. A future assessment of this kind for extra-Neotropical taxa is commendable upon revision of their taxonomy.

RESULTS

Phylogenetic analyses and time calibration

The final cleaned supermatrix of nine molecular markers contains a total of 4913 sequences, of which 3909 (80%) were retrieved from GenBank and 1004 (20%) are newly published here. The concatenated alignment has a length of 10,312 base pairs, of which 46% represent missing data; ITS and *psbA-trnH* are the markers most commonly represented throughout the supermatrix. The tree inferred from this supermatrix encompasses 1029 tips, each represented by a single voucher, representing 773 species and 48 genera (including outgroups), and 712 species and all the 34 accepted genera of Neotropical Myrteae (excluding outgroups), ~28% of the species diversity in the clade (Figure 2, Table 1). Of the 1029 vouchers included in the tree, 140 vouchers are represented only by newly generated sequences, whereas 889 had at least some sequence already deposited on GenBank prior to the study. Among the latter, 153 had some type of name alteration, the most common being updates in accepted name (97), followed by redeterminations (54), and typographical corrections in voucher or collector number (2). Among the 78 previously unidentified vouchers (i.e., those marked as “sp.”), 17 could be newly identified to species level, whereas eight previously misidentified vouchers were newly assigned as “sp.”; 54 vouchers remain unidentified after taxonomic verification. Additional details on name alterations can be found in Appendix S2.

Tribe Myrteae is recovered with high support (Figure 2; BS 100) in the ML analysis, as are the recently recircumscribed subtribes Eugeniinae and Myrtinae (BS 97 and 90, respectively). Myrciinae and the Neotropical clade of Myrteae are recovered with moderate support (BS 72 and 62, respectively) and Pliniinae, Pimentinae, and Luminae all appear as low supported clades (BS < 50). Oldest splits in the phylogeny of Eugeniinae and *Eugenia* appear highly supported (Figure 3) whereas supports tended to be lower for relationships among sections of *Myrcia* in the Myrciinae (Figure 4). Based on estimated ages of the crown groups, the oldest of the Neotropical Myrteae subtribes appears to be Eugeniinae (33.21 mya, with a confidence interval (CI) of 30.22–36.12), followed by Pimentinae (33.02 mya, 28.92–36.96), Luminae (29.89 mya, 23.4–35.48), Pliniinae (27.42 mya, 23.49–31.48), Myrciinae (26.95 mya, 23.39–30.14), Myrtinae (26.33 mya, 19.51–32.97), Blepharocalycinae (24.82 mya, 17.59–32.58), and finally Ugninae (24.79 mya, 22.12–29.85) (Table 1). A chronogram depicting mean ages and confidence intervals for all nodes can be found in Appendix S5.

Assessment of taxonomic and geographic coverage

Subtribes with the best taxonomic coverage in our supermatrix data set tend to be the species-poor ones, i.e., those with less than 100 species (Table 1). Luminae presents 68% of its ~50 species sampled, Blepharocalycinae 66% of its three species

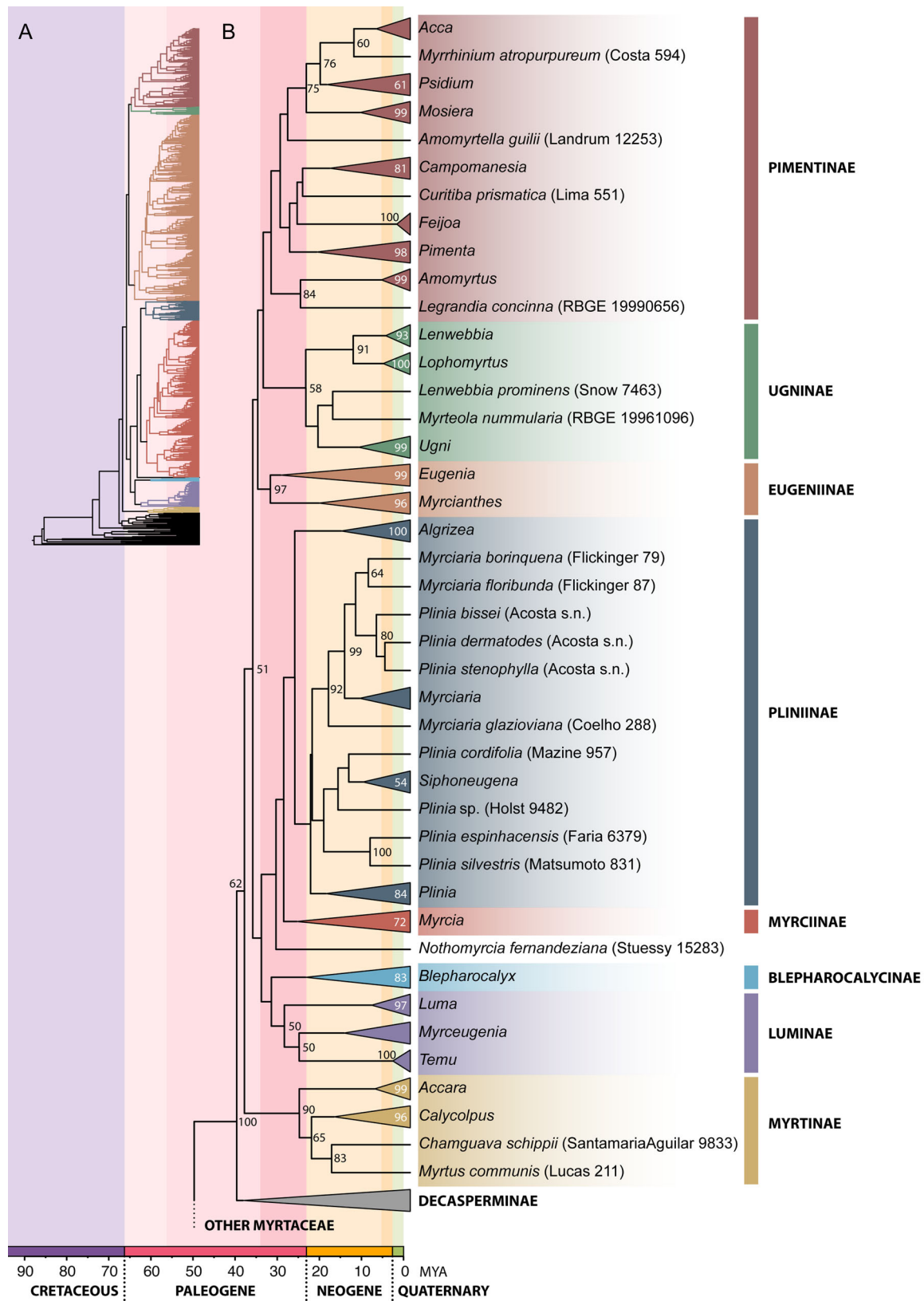


FIGURE 2 Time calibrated maximum likelihood phylogenetic tree inferred from the concatenated supermatrix of nine markers for Myrtaceae: (A) shows the full, uncollapsed topology including 1029 tips; (B) highlights relationships between the genera and unplaced species comprising the eight predominantly Neotropical subtribes of tribe Myrteae (Myrtaceae). Bootstrap supports <50 are not shown.

TABLE 1 Mean, minimum (min) and maximum (max) crown node age for Myrteae, the Neotropical clade, and focal subtribes. Taxonomic coverage indicates the proportion of species within each group sampled in the tree (number of species sampled divided by total number of species estimated to be placed in the clade). *Note that the taxonomic coverage for Myrteae is lower because subtribe Decasperminae is not part of the Neotropical clade and thus was included in the tree only as part of the outgroup.

Group	Age (mean)	Age (min)	Age (max)	Taxonomic coverage
Myrteae	41.19	39.53	43	0.27 (737 out of 2700)*
Neotropical clade	39.38	37.18	41.49	0.28 (712 out of 2500)
Blepharocalycinae	24.84	17.59	32.58	0.67 (2 out of 3)
Eugeniinae	33.21	30.22	36.12	0.16 (192 out of 1200)
Luminae	29.89	23.4	35.48	0.68 (34 out of 50)
Myrciinae	26.95	23.39	30.14	0.27 (203 out of 750)
Myrtinae	26.33	19.51	32.97	0.31 (6 out of 20)
Pimentinae	33.02	28.92	36.96	0.44 (88 out of 200)
Pliniinae	27.42	23.49	31.48	0.21 (25 out of 120)
Ugninae	24.80	22.12	29.85	0.64 (8 out of 12)

sampled (all but *Blepharocalyx myriophyllus* (Casar.) P.O.Morais & Sobral), and Ugninae 64% of its 12 species sampled. The exception is Myrtinae, where only 31% of its ~20 species have been sampled. The largest subtribes tend to have a lower proportion of their species richness covered, with 16% of the ~1200 species of Eugeniinae, 27% of the ~750 species of Myrciinae, and 21% of the ~120 Pliniinae sampled in the tree. Pimentinae, a relatively species rich subtribe of ~200 species, appear relatively well covered, with 44% of its species sampled. In terms of proportional sampling across the distribution of Myrteae in the Neotropics, the Amazon basin, Ecuador, Peru, as well as Mesoamerica and some areas in the Caribbean appear particularly poorly sampled in relation to their species richness (Figure 5A, B).

Contrasting resolution and topology of each marker

Analyses using the three subsets with complete molecular matrices highlight differences in bootstrap support resulting from analyses of each partition and individual molecular marker and topological conflicts between nuclear and plastid partitions (Figure 6). In the Myrteae and Myrciinae data sets, full concatenated matrices including all sampled markers consistently present the highest bootstrap support values, followed by the concatenated plastid partitions (Figure 6A, C). The opposite is observed in the Eugeniinae data set, where the plastid partition presents higher support than the full concatenated one (see Figure 6B). Some of the most recurrently used markers in Myrteae phylogenetics, such as *psbA-trnH*, *ndhF*, *rpl16*, and *trnL-trnF* yielded low

support when analyzed in isolation and were highly incongruent with the full matrix. Surprisingly, the most congruent individual markers in the full analysis are *matK* and *trnQ-rps16*—two regions less frequently sequenced for the whole tribe. The support yielded by these two markers is even higher than the concatenated nuclear partition in this data set (Figure 6A). In general, Mantel correlation values (Table 2) were congruent to the analyses of support, i.e., phylogenies with lower support also had topologies that were more distinct from the full concatenated one, with a few exceptions. Analyses of the phylogenetic landscape shows strong conflict between plastid and nuclear partitions in the full Myrteae data set, with low overlap between partitions in the treespace (Figure 6D). This conflict is also observed in the Myrciinae data set (Figure 6F), but less so in the Eugeniinae data set (Figure 6E) where the plastid partition seems to overlap more closely with the nuclear one. In all data sets, the topology of the full matrix overlaps mainly with the topology inferred from the plastid partition.

DISCUSSION

General comments on topology and time calibration

Our inferred phylogeny represents the first Myrtaceae tree to include all 34 genera of the hyper-diverse Neotropical clade of Myrteae. Most of the resulting suprageneric relationships are congruent with recent phylogenetic analyses performed using different samples within tribe Myrteae (e.g., Vasconcelos et al., 2017b; Maurin et al., 2021), but there are some important differences that must be highlighted when contrasting our results to recently published trees. First and foremost, although our topology recovers all major groups of Myrteae—that is, all subtribes and sections within the large genera *Myrcia* and *Eugenia* (sensu Mazine et al., 2016; Lucas et al., 2018)—support across our tree is generally lower than in previous studies that focused on specific groups within the tribe, potentially due to the proportion of missing data (see discussion below). For practical reasons, our discussion will consider these major groups as clades for now even if support is low and focus on the relationship among and within them. However, given the low support of some relationships at both higher and lower taxonomic levels, it is important to emphasize that our supermatrix tree should not be considered as the final word in terms of phylogenetic relationships within Neotropical Myrtaceae. It is rather a first step towards evaluating sampling gaps and a potential preliminary resource for macroecological and macroevolutionary analyses that need a densely sampled, time calibrated tree.

Although all subtribes (sensu Lucas et al., 2019) are recovered as clades in our analysis, support varies among them. Specifically, Pliniinae, Pimentinae, and Luminae are all only poorly supported in our supermatrix analysis (BS < 50;

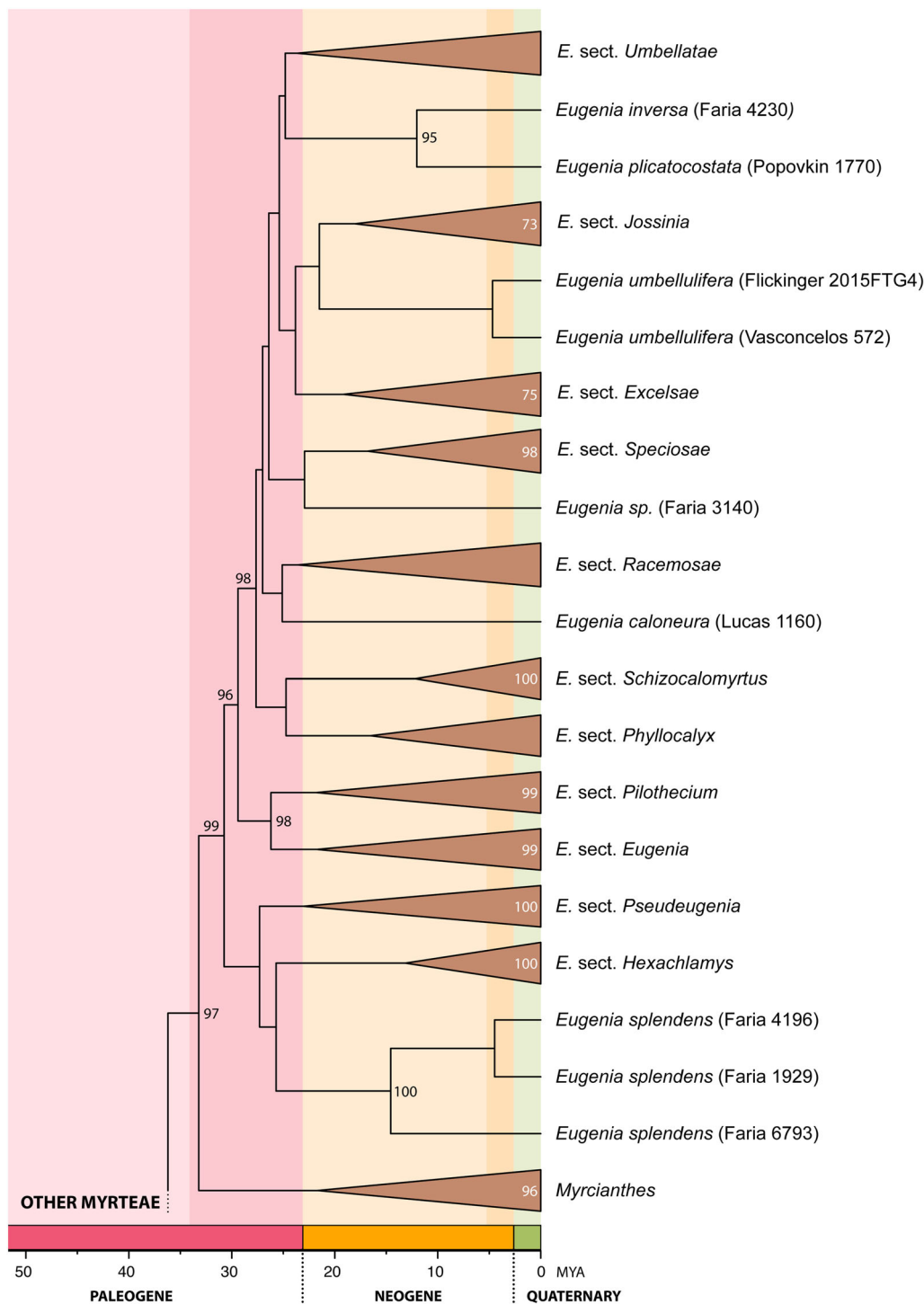


FIGURE 3 Time calibrated maximum likelihood phylogenetic tree inferred from the concatenated supermatrix of nine markers, highlighting relationships between sections of *Eugenia* and genera within subtribe Eugeniinae. Bootstrap supports <50 are not shown.

see all supports in detail for the RAxML tree in Appendix S4). The poor support for Pliniinae might be due to the positioning of *Algrizea* Proença & NicLugh., which has been inconsistently placed as either sister to Myrciinae (e.g., in Maurin et al., 2021) or Pliniinae (e.g., in Vasconcelos et al., 2017b; Stadnik, 2020), depending on the molecular data set used in the inference

(Lucas et al., 2007). Furthermore, the support for Pliniinae as sister to Myrciinae, although recovered, is only poorly supported here (BS < 50). Within Pliniinae, *Plinia* L. emerged as polyphyletic, with species from Central America placed within *Myrciaria* O.Berg in a strongly supported clade (BS 92). *Plinia espinhacensis* Sobral and *P. silvestris* (Vell.) Mazine &

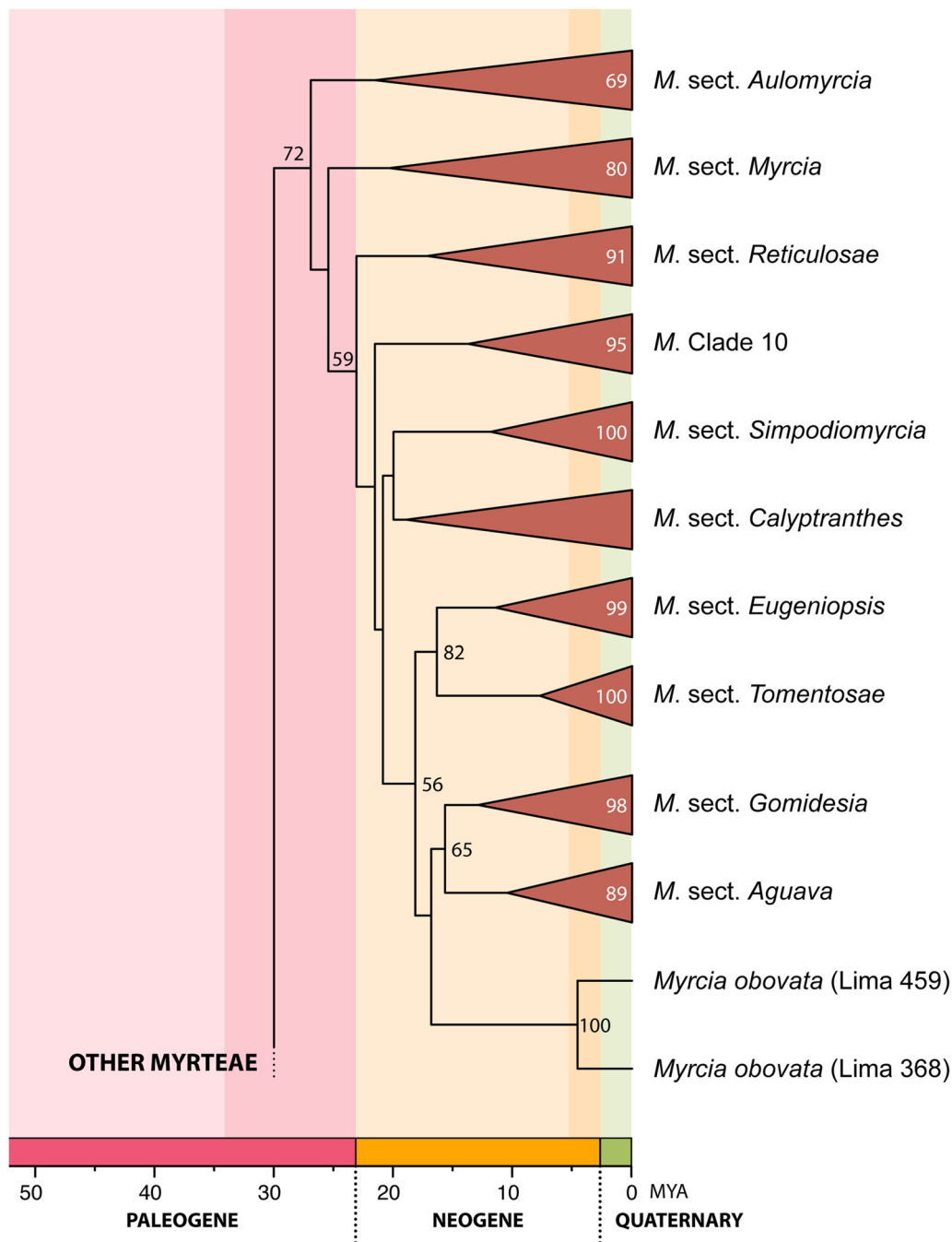


FIGURE 4 Time calibrated maximum likelihood phylogenetic tree inferred from the concatenated supermatrix of nine markers, highlighting relationships between sections of the mega-diverse genus *Myrcia* (subtribe Myrciinae). Bootstrap supports <50 are not shown.

Sobral belong to another well-supported lineage within the genus (BS 100). Both species share well-developed panicle-like inflorescences, an uncommon feature in the genus. Species of *Siphoneugena* O.Berg and *Neomitranthes* D.Legrand are both included in a clade with moderate support (BS 54). These two genera cover a broad morphological variation, mainly in inflorescence architecture and calyx fusion. However, they also share certain vegetative features, such as density and sizes of oil glands and a prominent main vein on the adaxial side of the leaf blade (Proença et al., 1990; Souza, 2009; Stadnik, 2020).

Genera included in Pimentinae have been previously found to be polyphyletic in other tribe-level analyses (e.g., Vasconcelos et al., 2017b, which divided them into informal *Pimenta* and *Psidium* groups) and the group has also been poorly supported in phylogenomic analyses of the order Myrtales (Maurin et al., 2021). In the tree inferred here, relationships within *Psidium* L. were reasonably consistent with those recovered in the recent phylogeny of the genus (Proença et al., 2022), although with lower support. The monophyly of the four sections of *Psidium* was confirmed,

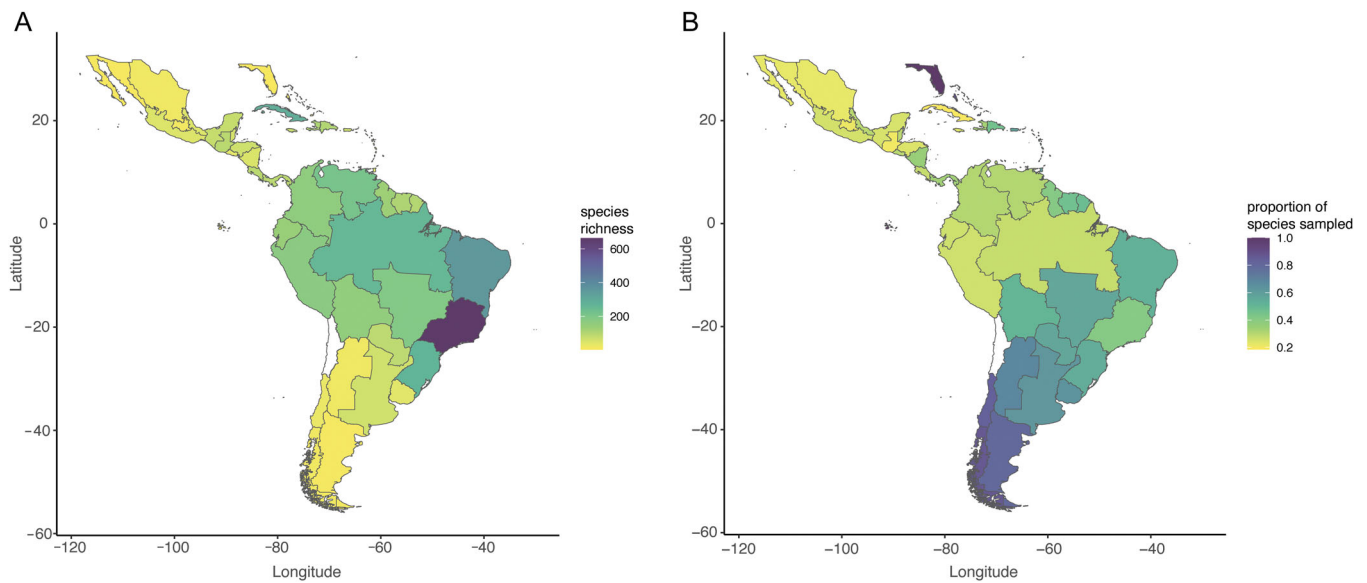


FIGURE 5 Assessment of geographic gap in the phylogenetic sampling: (A) total Myrteae species-richness per level-3 botanical country based on POWO (2023); (B) proportion of species sampled in the phylogeny in each level-3 botanical country. No species occurs in Northern Chile, so the area is shown in white.

although the relative position of *Psidium* sect. *Obversifolia* O.Berg differed from that of their tree. *Pimenta pseudocaryophyllus* (Gomes) Landrum is recovered together with the rest of *Pimenta*, contrary to Maurin et al. (2021). This is an important difference because the placement of *P. pseudocaryophyllus* in the Myrtinae subtribe in the latter study indicated the possible resurrection of the monotypic genus *Pseudocaryophyllus* O.Berg. Pimentinae is found along the whole of the Myrteae geographic distribution in the Neotropics and includes genera with wide variation in embryo (Landrum and Stevenson, 1986), floral (Vasconcelos et al., 2015, 2019b) and fruit traits (Pittarelli et al., 2021), indicating that it might in fact constitute more than one natural lineage. The issue with the monophyly of Pimentinae was discussed when the subtribe was first proposed (i.e., Lucas et al., 2019) and it definitely deserves to be the focus of future studies focused on this clade, especially given the economic importance of members of the group (e.g., allspice, *Pimenta dioica* (L.) Merr., guava, *Psidium guajava*, and the aggressive invasive *Psidium cattleianum* Sabine).

Some general comments are also required regarding the topologies of the super diverse Eugeniinae and Myrciinae (Figures 3 and 4). Because of the size and sparse nature of the molecular matrix, relationships among sections within the large *Eugenia* and *Myrcia* also presented generally lower support in the full data set than those recovered by recently published studies focused on fewer species and using more complete molecular data sets. The relationships among *Myrcia* sections, for instance, received higher support in the supermatrix analysis of Amorim et al. (2019). In our analysis, *Myrcia* sect. *Aulomyrcia* (O.Berg) Griseb. appears as a clade with moderate support (BS 69) sister to the rest of *Myrcia*, followed by *Myrcia* sect. *Myrcia* (BS 80) as sister to

the remaining sections. *Myrcia* sect. *Aulomyrcia* has not been consistently recovered as a clade in previous analyses (e.g., Santos et al., 2017) and the reason might be linked to partition conflict (see below). Both *M.* sect. *Aulomyrcia* and *M.* sect. *Myrcia* are the most geographically widespread sections of *Myrcia*, occurring across the whole Neotropical region, whereas the diversity of species in other sections is mainly concentrated in the Cerrado and Atlantic Forest domains of Eastern Brazil (except perhaps for *M.* sect. *Calypttranthes* (Sw.) A.R.Lourenço & E.Lucas, which is also diverse in the Caribbean region). A better understanding of the relationships in these earliest splits and across the backbone of *Myrcia* is necessary for untangling the biogeographical history of the group in the Neotropics. Among other sections, *M.* sect. *Calypttranthes* is recovered with low support (BS < 50) and placed as sister to *M.* sect. *Sympodiomyrcia* M.F.Santos & E.Lucas (BS 100) as in previous studies (e.g., Santos et al., 2016, 2017). Although this relationship is also poorly supported, these two sections share some morphological characteristics, such as sympodial branching, cataphylls, and deciduous calyx lobes (Santos et al., 2016). *Myrcia* “Clade 10” (BS 95) (Amorim et al., 2019; Lima et al., 2021) is recovered with high support. *Myrcia* sect. *Gomidesia* (O.Berg) B.S.Amorim & E.Lucas (BS 98) is recovered with high support, and so is *Myrcia* sect. *Aguava* (Raf.) D.F.Lima & E.Lucas (BS 89), as long as *M. obovata* (O.Berg) Nied. is not included as part of the latter. These two sections are recovered as a clade, as supported by Lucas et al. (2011) and Amorim et al. (2019). *Myrcia obovata* appears as sister to this clade in our tree, contradicting previous molecular analyses and morphological characteristics that place this species within *M.* sect. *Aguava* (Lima et al., 2021). *Myrcia* sect. *Eugeniopsis*

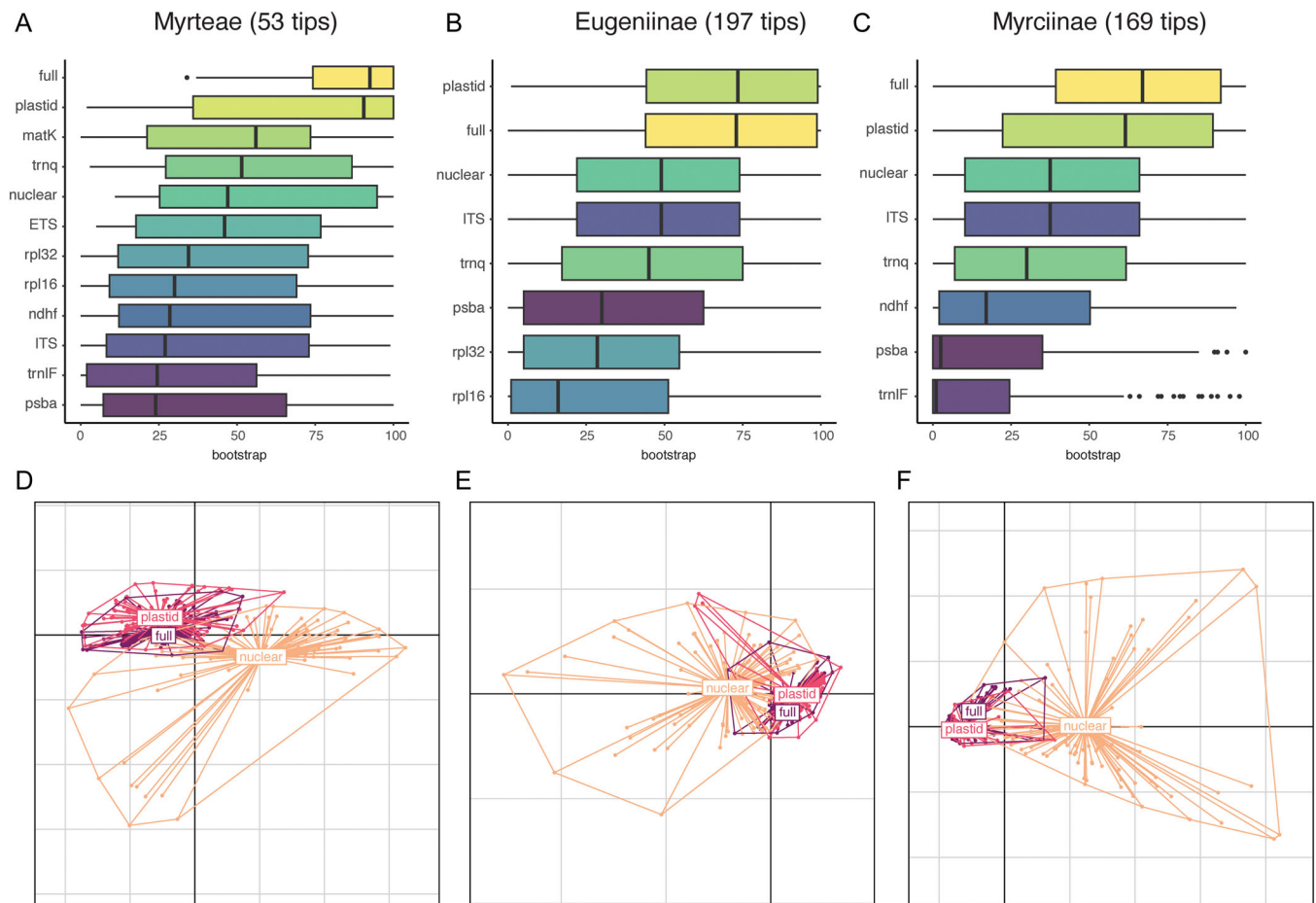


FIGURE 6 Distribution of bootstrap supports per molecular marker and partition subset and treespaces comparing nuclear and plastid partitions with full data sets. Columns represent subsets for (A, D) Myrteae (53 tips with complete molecular data sets for nine markers); (B, E) Eugeniinae (197 tips with complete molecular data set for five markers); (C, F) Myrciinae (169 tips with complete molecular data set for five markers).

TABLE 2 Mantel correlation values for comparisons between the full concatenated topology and topologies resulting from each individual partition and marker. Values range from 0 (topologies are completely different) to 1 (topologies are completely similar). All values were significant in the post hoc comparison.

	nuclear	ETS	ITS	plastid	matK	ndhF	psbA-trnH	rpl16	rpl32-trnL	trnL-trnF	trnQ-rps16
all53	0.98	0.97	0.97	1	0.96	0.95	0.91	0.88	0.93	0.91	0.94
Eugeniinae	0.85		0.85	0.99			0.63	0.79	0.85		0.83
Myrciinae	0.82		0.82	0.99		0.94	0.75			0.58	0.95

(O.Berg) M.F.Santos & E.Lucas (BS 99) is recovered as sister to *M. sect. Tomentosae* E.Lucas & D.F.Lima (BS 100) with high support (BS 82), consistent with previous studies (Santos et al., 2016, 2017; Amorim et al., 2019). Relationships within each section are mostly consistent with previous studies focused on specific groups (e.g., Wilson et al., 2016; Santos et al., 2017; Amorim et al., 2019; Lima et al., 2021; Santos et al., 2021), with the exception of *M. sect. Aulomyrcia*, as stated above.

The reconstructed phylogenetic hypothesis of Eugeniinae mostly confirms the general topology proposed by

Mazine et al. (2014, 2018), Vasconcelos et al. (2017b), and Giaretta et al. (2022). *Eugenia* sect. *Jossinia* (BS 73), the largest non-neotropical clade in our sample, is geographically structured. The relationships in this clade suggest a westward route of round-the-world colonization, with a clade composed of southeast Asia and New Caledonia species sister to species distributed in Mauritius, India, Madagascar, and Africa (Appendix S4). Disjunct geographical patterns in this and other Myrteae groups are likely a result of long-distance dispersal events or dispersal through Antarctica before the Miocene, when the ice sheet that

covers that continent was formed (Vasconcelos et al., 2017b; Estrella et al., 2019). Because these relationships are poorly supported, further studies that improve the resolution of the relationships within this group are necessary to confirm potential biogeographical routes. Finally, several species represented by more than one voucher appear polyphyletic, including the widespread *Blepharocalyx salicifolius* (Kunth) O.Berg, *Myrcia amazonica* DC., *Myrcia multiflora* (Lam.) DC., *Myrcia splendens* (Sw.) DC., and *Myrcia guianensis* (Aubl.) DC. These species should be the target of future studies that aim to better understand species limitations and improve diagnostic characters within Myrteae.

In terms of time estimates, divergence ages among subtribes are generally in accordance with previous inferences for this group that used a similar fossil data set for time calibration (e.g., Thornhill et al., 2015; Vasconcelos et al., 2017a). It is interesting to note that, when confidence intervals are considered, all subtribes have very similar age, and have likely originated in the Oligocene (33–23 mya). It is also worth noting that our calibration approach used only pollen fossil data, because their systematic placement was recently thoroughly reviewed (Thornhill et al., 2012). However, a promising way forward in improving our current understanding of time estimates is by reviewing the macrofossil record for Myrteae as well. Several macrofossils assigned to Myrtaceae have been recovered from the Cretaceous of southern South America, with some being placed within tribe Myrteae (e.g., Ragonese, 1980; Poole et al., 2001; Tosolini et al., 2021); including them in future calibration analyses may push back the crown age of Myrteae (e.g., Murillo-A et al., 2016). Although that means that time divergences may change once more fossil data and better time divergence models are incorporated in the analyses, we believe the current time calibrated, taxonomically verified tree (available as Appendices S5–S8) can be preliminarily used as a resource for novel studies on ecology and evolution until further data is available. Other ways forward in improving the tree are discussed below.

Lack of phylogenetic resolution and incongruence between partitions

The low bootstrap support in our supermatrix tree likely results from using a sparse and patchy molecular matrix, which contains several gaps that occasionally reduce branch support. Evidence for this is that the mean bootstrap support increases when we analyze subsets with complete molecular sampling for Myrteae, Myrciinae, and Eugeniinae (i.e., Figure 6A, C, E). Analysis of closely related lineages in a subset is likely to incorporate less genealogical discordance and avoid genuine phylogenetic signal being blurred by stochastic processes such as gene duplication and loss, hybridization, and incomplete lineage sorting (see Smith et al., 2015; Molloy and Warnow, 2018). These analyses will also have considerably fewer missing data, which have been shown to lead to spurious relationships and low support in

supermatrices analyses under likelihood inferences (Simmons, 2012). In *Eugenia*, a more data rich matrix obtained from targeted sequencing suggested that conflicting signals are involved in the low support recovered along the backbone (Giarretta et al., 2022), which indicates that incomplete lineage sorting may play a role in Myrteae phylogenetics. Rapid radiation and recent divergence have been implicated in phylogenetic incongruence (Cai et al., 2021; Meleshko et al., 2021) and further studies will be able to determine if this is also the case in Myrteae.

Another potential reason for the lower resolution in the concatenated supermatrix is the conflict between plastid and nuclear partitions. When comparing topologies resulting from analyses of different markers, we observe that topologies inferred from nuclear and plastid data sets are strongly incongruent in the full data set (Figure 6B, D, F). Diverging topologies between plastid and nuclear data sets can be a result of evolutionary rates in nuclear partitions being five times faster than chloroplasts in angiosperms (Drouin et al., 2008), leading nuclear partitions to accumulate more molecular changes over time. Phylogenetic analyses based on genes with diverse evolutionary rates can yield incongruent signals. Fast-evolving genes are more susceptible to saturation, potentially leading to multiple substitutions at a given site that may erode phylogenetic signal (Philippe et al., 2000; Zhong et al., 2011). This is especially true if distantly related taxa share convergent mutations. The increased likelihood of multiple mutations in fast-evolving genes may result in homoplasy, introducing conflicting signals in phylogenetic analyses. Incongruences are also possible if genes experiencing strong selection pressures evolve more slowly, creating DNA evolution rate disparities and potentially resulting in an incongruent phylogenetic signal (Philippe and Roure, 2011; Kapli et al., 2020).

Another possible cause of incongruence could be hybridization resulting in reticulate evolution. In *Angophora*, *Corymbia*, and *Eucalyptus* (Myrtaceae, tribe Eucalypteae), incongruence between chloroplast (cp) DNA trees and ITS trees has been largely attributed to hybridization and introgression resulting in shared chloroplasts across taxa, giving rise to geographically structured “chloroplast clades” (Schuster et al., 2018). In *Hieracium* (Asteraceae), incongruence between ITS and cpDNA phylogenetic trees was attributed to ancient intergeneric hybridization events that resulted in chloroplast capture (Fehrer et al., 2007; Schilling, 2011); i.e., the chloroplasts of the introgressed hybrid are identical to those of the maternal parent taxon. Hybridization occurs naturally and artificially in *Eucalyptus* (Griffin et al., 1988) and in *Psidium* (Landrum et al., 1995; Costa et al., 2012). The fact that the chromosome count $2n = 22$ is dominant in Myrtaceae (Rye, 1979) and in tribe Myrteae (Costa and Forni-Martins, 2006 and references therein) is significant in this context. A low variability in chromosome number potentially removes one of the major barriers to hybridization in angiosperms, the imperfect pairing of gametes when the number of chromosomes differs between parent taxa (Stace et al., 2015).

The incongruence observed between nuclear and full data sets, as well as the overlap between plastid and full topologies in our tree-space analysis may be also due to the plastid partition being more data-rich (i.e., longer) than the nuclear one. These analyses show that our full concatenated topology is probably being driven mostly by the plastid partition, where most of the data comes from. That could explain some small inconsistencies between this and the Maurin et al. (2021) genus-level topology using the Angiosperms-353 probe set, which is a nuclear-only data set (Johnson et al., 2019). Examples of these inconsistencies include the placement of *Algrizea* and *Pimenta pseudocaryophyllus* as discussed above and support for Ugniinae as a clade (recovered in ours but not in Maurin et al., 2021). However, given that most higher-level relationships are congruent between our supermatrix and Angiosperms-353 topologies (Maurin et al., 2021; Giaretta et al., 2022), and that Angiosperms-353 topologies are inferred with significantly larger amounts of molecular data and generally result in higher node support, we believe that using the Angiosperms-353 in a broader Myrteae sample is a promising way forward to improve our current understanding of relationships in the group. This method will also open the opportunity of contrasting complementary data sets to best assess deep and shallow phylogenetic levels, as well as perform relevant genomic analyses to recognize if processes other than genuine conflict signals are taking place. For instance, Giaretta et al. (2022) recovered several infrageneric level topologies of *Eugenia* contrasting different data sets (exons, introns, and off-target plastid) and reconstructions methods based on the Angiosperms-353. Even though their general topology agrees with the previous phylogenetic reconstructions (e.g., Mazin et al., 2014, 2018), some inconsistencies were better understood, such as the relationships between groups within *E. sect. Umbellatae* O.Berg.

Tackling sampling gaps in future phylogenetic studies of Neotropical Myrteae

Another important way forward to a better understanding of relationships in the tribe is to increase species sampling, particularly in clades and areas that are currently under-sampled. For instance, we show that the largest subtribe in Myrteae, Eugeniinae, remains with less than a fifth of its diversity with trustworthy sequences available for integrative phylogenetic analyses. Sampling is also uneven within *Eugenia*, which makes up most of Eugeniinae with over 1200 species: the gigantic *Eugenia* sect. *Umbellatae* (~500 species) has only ~12% of its species richness sampled in our phylogeny. The second largest section, *Eugenia* sect. *Jossinia* is also poorly represented, with only 37 out of the estimated 200 species sampled in our tree. There is a bias in geographical collecting effort as well. Some areas of southern and eastern South America are better represented than the Amazon basin and Mesoamerica. Furthermore, although eastern Brazil appears relatively well sampled (Figure 5B), it is

likely that many of the unsampled species in the tree occur in this region, given the exceptional species richness of the group in these areas (Figure 5A). Increased sampling from these areas will also help shed light on species delimitation of widespread species. Some *Myrcia* species (e.g., *Myrcia amazonica*, *M. guianensis*, *M. multiflora*, and *Myrcia splendens*) are not monophyletic on the full analysis and in the Myrciinae data set. These species are widely distributed across several ecosystems in the Caribbean, Mesoamerica, and South America, and most of them exhibit high morphological variability. Further analyses are needed for understanding if such constitute separate lineages or not, and if so if they could be split into two or more species, according to clades obtained in the phylogeny. Given some reported difficulties in extracting sequences from herbarium material even using target sequencing approaches (e.g., Brewer et al., 2019), more field collections will probably be required for such endeavors, especially focusing on groups that lack sampling and in poorly collected areas. New collaborative efforts with botanists in underrepresented regions will represent a key step forward towards this goal.

CONCLUSIONS

Here we assembled a densely sampled molecular supermatrix for the Neotropical clade of tribe Myrteae, Myrtaceae. We showed that reassessing the identification of vouchers used in other studies improved identification accuracy of currently available data, which is of extreme importance to diminish biases in analyses using taxonomically complex groups such as this. The careful taxonomic validation of data sets of this size can be time-consuming, but it still is the best approach for assembling trustworthy phylogenies of taxonomically complicated groups such as Myrteae. Working toward an extended specimen network could facilitate taxonomic updates of other specimen-derived data and research products (Lendemer et al., 2020).

We also provide ~1000 new sequences, representing one fifth of the currently available molecular information in the group, which can now be reliably used for other future ecological and evolutionary inferences in this important tropical group. This process resulted in a broadly sampled time calibrated tree that is expert-verified by taxonomists and available for use by ecologists and evolutionary biologists to test hypotheses in their fields, avoiding data of dubious identifications. Gaps for future studies are highlighted; addressing these will involve resolving the incongruence between plastid and nuclear partitions, including a broader sample of species from mega-diverse groups and poorly sampled areas as well as overcoming noise from stochastic processes such as hybridization, introgression or incomplete lineage sorting to better accommodate gene tree heterogeneity. These future endeavors may provide resolution for some persistent taxonomic issues highlighted in the discussion. Results presented here guide future studies to clades that are most urgently in need of sampling. Given the difficulty in

accessing suitable samples for sequencing, compiling large phylogenies of tropical groups is a challenge. However, these phylogenies represent a window for greater understanding of highly diverse, taxonomically complex groups from tropical ecosystems where these lineages are often most species rich.

AUTHOR CONTRIBUTIONS

V.G.S. and T.V. conceived the idea and led the study. V.G.S., B.A., M.B., I.C., J.E.Q.F., A.G., D.F.L., L.L., A.R.L., F.F.M., J.M.A., M.I.U.O., C.P.O., C.E.B.P., M.F.S., A.S., A.C.T., K.S.V., and T.V. provided new sequences. V.G.S., B.A., M.B., I.C., J.E.Q.F., J.F., A.G.O., D.F.L., L.L., A.R.L., F.F.M., M.I.U.O., C.E.B.P., P.O.R., M.F.S., A.S., A.C.T., K.S.V., and T.V. confirmed voucher identification. V.G.S., D.F.L., C.E.B.P., P.O.R., and T.V. organized the data set. V.G.S., M.R., and T.V. performed the analyses. V.G.S., B.A., M.B., A.G.O., D.F.L., F.F.M., M.F.S., A.S., A.C.T., and T.V. wrote the first draft of the manuscript. All authors revised and approved the final version of the manuscript.

AFFILIATIONS

¹Departamento de Ecologia, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, RN 59092-970, Brazil

²Programa de Pós-Graduação em Biotecnologia e Recursos Naturais da Amazônia, Universidade do Estado do Amazonas, Manaus, AM, Brazil

³Programa de Pós-Graduação em Sistemática, Uso e Conservação da Biodiversidade, Department de Biologia, Universidade Federal do Ceará, Fortaleza, CE 60355-636, Brazil

⁴Instituto Interamericano de Cooperação para a Agricultura - IICA - SHIS QI 5, Chácara 16, Lago Sul, Brasília, DF 71600-530, Brazil

⁵Lady Bird Johnson Wildflower Center, The University of Texas at Austin, 4801 La Crosse Ave., Austin, TX 78739, USA

⁶Universidade Federal da Grande Dourados, Faculdade de Ciências Biológicas e Ambientais, Unidade II, Dourados, MS 79804-970, Brazil

⁷Departamento de Botânica, Laboratório de Sistemática Vegetal, Instituto de Biociências, Universidade de São Paulo, São Paulo, São Paulo 05508-900, Brazil

⁸Programa de Pós-Graduação em Biologia de Fungos, Algas e Plantas, Centro de Ciências Biológicas, Universidade Federal de Santa Catarina, Florianópolis, SC 88040-900, Brazil

⁹Instituto Federal de Roraima, campus Amajari, Amajari/Roraima 69343-000, Brazil

¹⁰Independent researcher

¹¹Herbarium, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK

¹²Universidade Federal de São Carlos, Campus Sorocaba, Rodovia João Leme dos Santos (SP-264), km 110, Sorocaba, SP 18052-780, Brazil

¹³Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Carrera 30 No. 45-03, Bogotá, Colombia

¹⁴Departamento de Biologia, Universidade Federal de Sergipe, Av. Marcelo Déda Chagas, s/n, Bairro Jardim Rosa Elze, São Cristóvão, SE 49107-230, Brazil

¹⁵Departamento de Botânica, Universidade de Brasília, Brasília, DF 70910-900, Brazil

¹⁶Departamento de Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 90650-001, Brazil

¹⁷Jardim Botânico de Brasília, Diretoria de Vegetação e Flora, Área Especial SMDB Estação Ecológica Jardim Botânico de Brasília, Brasília, DF 71.680-001, Brazil

¹⁸Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Alameda da Universidade s/n, Anchieta, São Bernardo do Campo, SP 09606-045, Brazil

¹⁹Programa de Pós-Graduação em Botânica, Universidade Estadual de Feira de Santana, Departamento de Ciências Biológicas, Av. Transnordestina s/n, Feira de Santana, BA 44036-900, Brazil

²⁰Centro de Estudos da Biodiversidade, Universidade Federal de Roraima, Campus Paricarana, Av. Cap. Ene Garcez, 2413, Boa Vista, RR 69304-000, Brazil

²¹Departamento de Ciências Biológicas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP 13418-260, Brazil

²²Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

ACKNOWLEDGMENTS

V.G.S. was funded by FAPESP (16/02312-8), CNPq (473468/2010-7 and 236687/2012-3), CAPES Ph.D. fellowship, Bentham-Moxon Trust, and Serrapilheira (R-2111-39858). B.S.A. was funded by CAPES (88882.315044/2019-01). D.F.L. was funded by CAPES (88887.371829/2019-00). C.E.B.P., F.F.M., and M.R. were funded by CNPq (314593/2020-9, 305105/2022-1, and 302490/2022-1 respectively). I.R.C., V.G.S., and C.E.B.P. also received funding from CNPq through the project “Estudos filogenéticos e macroecológicos voltados à conservação do gênero *Psidium* L. (Myrtaceae Juss. – 2011–2015, REFLORA Program) and I.R.C. through the Kew Latin American Research Fellowship (KLARF). Authors acknowledge staff from herbaria ASE, HUEFS, K, SPF, UB, and collection permits from SISBIO (21120-1 to M.I.U.O., 41156-6 to D.F.L.); Laszlo Csiba, Dion Devey, Jim Clarkson, Laura Martinez, J. Floriano Pastore, and Maria José G. de Andrade provided laboratory support; Rob Smissen, Marcos Sobral, John Freudenstein, and two anonymous reviewers provided comments that greatly improved the manuscript.

DATA AVAILABILITY STATEMENT

All raw data used in this paper are available as part of the Supporting Information online. The code used to perform analyses can be found at https://github.com/tncvasconcelos/myrteae_tree.

ORCID

Vanessa G. Staggemeier  <http://orcid.org/0000-0003-4911-9574>

Bruno Amorim  <https://orcid.org/0000-0002-8109-9254>


Mariana Bünger  <http://orcid.org/0000-0003-4241-2399>

Itayguara R. Costa  <http://orcid.org/0000-0003-0135-6706>

Jonathan Flickinger  <http://orcid.org/0000-0001-7693-131X>

Duane Fernandes Lima  <http://orcid.org/0000-0003-4240-9077>

Eve Lucas  <http://orcid.org/0000-0002-7603-435X>

Fiorella Fernanda Mazine  <http://orcid.org/0000-0002-2604-6088>


José Murillo-A  <http://orcid.org/0000-0002-3703-4250>

Carolyn E. B. Proença  <http://orcid.org/0000-0002-8924-2692>

Priscila Oliveira Rosa  <http://orcid.org/0000-0002-3821-5520>

Aline Stadnik  <http://orcid.org/0000-0002-2370-5869>

Amélia Carlos Tuler  <https://orcid.org/0000-0002-1512-5493>

Karinne Sampaio Valdemarin  <http://orcid.org/0000-0002-9564-1163>

Thais Vasconcelos  <http://orcid.org/0000-0001-9991-7924>

REFERENCES

- Amorim, B. S., T. N. C. Vasconcelos, G. Souza, M. Alves, A. Antonelli, and E. J. Lucas. 2019. Advanced understanding of phylogenetic relationships, morphological evolution and biogeographic history of the mega-diverse plant genus *Myrcia* and its relatives (Myrtaceae: Myrteae). *Molecular Phylogenetics and Evolution* 138: 65–88.
- Antonelli, A., and I. Sanmartín. 2011. Why are there so many plant species in the Neotropics? *Taxon* 60: 403–414.
- Antonelli, A., A. Zizka, F. A. Carvalho, R. Scharn, C. D. Bacon, D. Silvestro, and F. L. Condamine. 2018. Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences of the United States of America* 115: 6034–6039.
- Bennett, D. J., H. Hettling, D. Silvestro, A. Zizka, C. D. Bacon, S. Faurby, R. A. Vos, and A. Antonelli. 2018. Phylotax: An automated pipeline for retrieving orthologous DNA sequences from GenBank in R. *Life* 8: 1–11.
- Bortolus, A. 2008. Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *Ambio*, 37, 114–118.
- Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, et al. 2014. BEAST 2: A software platform for bayesian evolutionary analysis. *PLoS Computational Biology* 10: e1003537.
- Brewer, G. E., J. J. Clarkson, O. Maurin, A. R. Zuntini, V. Barber, S. Bellot, N. Biggs, et al. 2019. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science* 10: 1–14.
- Brummitt, R. K., F. Pando, and S. Hollis. 2001. Plant taxonomic database standards No. 2. World geographical scheme for recording plant distributions. 2nd ed. Pittsburgh, Pennsylvania.
- Bünger, M. O., F. F. Mazine-Capelo, F. Forest, M. L. Bueno, J. R. Stehmann, and E. J. Lucas. 2016. The evolutionary history of *Eugenia* sect. *Phyllocalyx* (Myrtaceae) corroborates historically stable areas in the southern Atlantic forests. *Annals of Botany* 118: 1209–1223.
- Cai, L., Z. Xi, E. M. Lemmon, A. R. Lemmon, A. Mast, C. E. Buddenhagen, L. Liu, and C. C. Davis. 2021. The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology* 70: 491–507.
- Cascaes, M. M., G. M. S. P. Guilhon, E. H. A. Andrade, M. G. B. Zoghbi, and L. S. Santos. 2015. Constituents and pharmacological activities of *Myrcia* (Myrtaceae): A review of an aromatic and medicinal group of plants. *International Journal of Molecular Sciences* 16: 23881–23904.
- Costa, I. R., and E. R. Forni-Martins. 2006. Chromosome studies in Brazilian species of *Campomanesia* Ruiz & Pavon and *Psidium* L. (Myrtaceae Juss.). *Caryologia* 59: 7–13.
- Costa, S. R., C. A. F. Santos, and J. M. C. Castro. 2012. Assessing *Psidium guajava* × *P. guineense* hybrids tolerance to *Meloidogyne enterolobii*. *Acta Horticulturae* 959: 59–66.
- Dellinger, A. S., M. Chartier, D. Fernández-Fernández, D. S. Penneys, M. Alvear, F. Almeda, F. A. Michelangeli, et al. 2019. Beyond buzz-pollination-departures from an adaptive plateau lead to new pollination syndromes. *New Phytologist* 221: 1136–1149.
- Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin of the Botanical Society of America* 19: 11–15.
- Drouin, G., H. Daoud, and J. Xia. 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Molecular Phylogenetics and Evolution* 49: 827–831.
- Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- Estrella, M. de la, S. Buerki, T. N. C. 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.
- Fehrer, J., B. Gemeinholzer, J. Chrtek, and S. Bräutigam. 2007. Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in *Pilosella* hawkweeds (Hieracium, Cichorieae, Asteraceae). *Molecular Phylogenetics and Evolution* 42: 347–361.
- Flickinger, J. A., B. Jestrow, R. Oviedo-Prieto, E. Santiago-Valentín, J. Sustache-Sustache, F. Jiménez-Rodríguez, K. Campbell, et al. 2019. A phylogenetic survey of Myrtaceae in the Greater Antilles. Botany 2019, Tucson, Arizona, USA.
- Flickinger, J. A., B. Jestrow, R. Oviedo Prieto, E. Santiago-Valentín, J. A. Sustache, F. Jiménez-Rodríguez, K. 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.
- Françoso, R. D., R. Haidar, and R. B. Machado. 2016. Tree species of South America central Savanna: endemism, marginal areas and the relationship with other biomes. *Acta Botanica Brasílica* 30: 1–9.
- Giaretta, A., L. F. T. Menezes, and A. L. Peixoto. 2015. Diversity of Myrtaceae in the southeastern Atlantic forest of Brazil as a tool for conservation. *Revista Brasileira de Botânica* 38: 175–185.
- Giaretta, A., B. Murphy, O. Maurin, F. F. Mazine, P. Sano, and E. J. Lucas. 2022. Phylogenetic relationships within the hyper-diverse genus *Eugenia* (Myrtaceae: Myrteae) based on target enrichment sequencing. *Frontiers in Plant Science* 12: 1–20.
- Gressler, E., M. A. Pizo, and L. P. C. Morellato. 2006. Polinização e dispersão de sementes em Myrtaceae do Brasil. *Revista Brasileira de Botânica* 29: 509–530.
- Griffin, A. R., I. P. Burgess, and L. Wolf. 1988. Patterns of natural and manipulated hybridization in the genus *Eucalyptus* L'Hérit.—A review. *Australian Journal of Botany*, 36: 41–66.
- Ibisch, P.L., G. Mérida (eds.). 2002. Estado de conocimiento y conservación. In Biodiversidad: La Riqueza de Bolivia. Ministerio de Desarrollo sostenible, 47–88. Editorial Fundación Amigos de la Naturaleza, Santa Cruz de la Sierra, Bolivia.
- 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.
- Joly, C. A., M. A. Assis, L. C. Bernacci, J. Y. Tamashiro, M. C. R. de Campos, J. A. M. A. Gomes, M. S. Lacerda, et al. 2012. Florística e fitossociologia em parcelas permanentes da Mata Atlântica do sudeste do Brasil ao longo de um gradiente altitudinal. *Biota Neotropica* 12: 125–145.
- 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.
- Kapli, P., Z. Yang, and M. J. Telford. 2020. Phylogenetic tree building in the genomic age. *Nature Reviews Genetics* 21: 428–444.
- Landrum, L. R., W. D. Clark, W. P. Sharp, and J. Bredecke. 1995. Hybridization between *Psidium guajava* and *P. guineense* (Myrtaceae). *Economic Botany* 49: 153–161.
- Landrum, L. R., and D. Stevenson. 1986. Variability of embryos in subtribe Myrtinae (Myrtaceae). *Systematic Botany* 11: 155–162.
- Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. Levan, et al. 2020. The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. *BioScience* 70: 23–30.
- Lima, D. F., R. Goldenberg, F. Forest, R. S. Cowan, and E. J. Lucas. 2021. Phylogeny and biogeography of *Myrcia* sect. *Aguava* (Myrtaceae, Myrteae) based on phylogenomic and Sanger data provide evidence for a Cerrado origin and geographically structured clades. *Molecular Phylogenetics and Evolution* 157: 107043.

- 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: 1–12.
- Lucas, E. J., S. R. Belsham, E. M. Nic-Lughadha, D. A. Orlovich, C. M. Sakuragui, M. W. Chase, and P. G. Wilson. 2005. Phylogenetic patterns in the fleshy-fruited Myrtaceae? preliminary molecular evidence. *Plant Systematics and Evolution* 251: 35–51.
- Lucas, E. J., and M. O. Büniger. 2015. Myrtaceae in the Atlantic forest: their role as a 'model' group. *Biodiversity and Conservation* 24: 2165–2180.
- Lucas, E. J., S. A. Harris, F. F. Mazine-Capelo, S. R. Belsham, E. M. Nic-Lughadha, A. Telford, P. E. Gasson, and M. W. Chase. 2007. Suprageneric phylogenetics of Myrteae, the generically richest tribe in Myrtaceae (Myrtales). *Taxon* 56: 1105–1128.
- Lucas, E. J., B. K. Holst, M. E. G. Sobral, F. F. Mazine, E. Nic Lughadha, C. E. B. Proença, I. R. Costa, and T. N. C. Vasconcelos. 2019. A new subtribal classification of tribe Myrteae (Myrtaceae). *Systematic Botany* 44: 560–569.
- Lucas, E. J., K. Matsumoto, S. A. Harris, E. M. Nic-Lughadha, B. Benardini, and M. W. Chase. 2011. Phylogenetics, morphology, and evolution of the large genus *Myrcia* s.l. (Myrtaceae). *International Journal of Plant Sciences* 172: 915–934.
- Martins, A. C., C. E. B. Proença, T. N. C. Vasconcelos, A. J. C. Aguiar, H. C. Farinasso, A. T. F. de Lima, J. E. Q. Faria, et al. 2023. Contrasting patterns of foraging behavior in Neotropical stingless bees using pollen and honey metabarcoding. *Scientific Reports* 13: e-14474.
- Maurin, O., A. Anest, S. Bellot, E. Biffin, G. Brewer, T. Charles-Dominique, R. S. Cowan, 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: 1087–1111.
- Mazine, F. F., M. O. Büniger, J. E. Q. Faria, E. J. Lucas, and V. C. Souza. 2016. Sections in *Eugenia* (Myrteae, Myrtaceae): nomenclatural notes and a key. *Phytotaxa* 289: 225–236.
- Mazine, F. F., J. E. Q. Faria, A. Giaretta, T. N. C. 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.
- Mazine, F. F., V. C. Souza, M. E. G. Sobral, F. Forest, and E. J. Lucas. 2014. A preliminary phylogenetic analysis of *Eugenia* (Myrtaceae: Myrteae), with a focus on Neotropical species. *Kew Bulletin* 69: 1–14.
- Meleshko, O., M. D. Martin, T. S. Korneliusson, C. Schröck, P. Lamkowski, J. Schmutz, A. Healey, et al. 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus. *Molecular Biology and Evolution* 38: 2750–2766.
- Mendonça, R. C., J. M. Felfili, B. M. T. Walter, M. C. Silva Júnior, A. V. Rezende, T. S. Filgueiras, P. E. Nogueira and C. W. Fagg. 2008. Flora vascular do Bioma cerrado checklist com 12.356 espécies. In S.M. Sano and S.P. Almeida [eds]. *Cerrado Ecologia e Flora*, 423–1270. Embrapa Cerrados, Brasília, DF.
- Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop, GCE 2010, 1–8. New Orleans, Louisiana.
- Molloy, E. K., and T. Warnow. 2018. To include or not to include: The impact of gene filtering on species tree estimation methods. *Systematic Biology* 67: 285–303.
- Mori, S. A., B. M. Boom, A. M. Carvalino, and T. S. Santos. 1983. Ecological importance of Myrtaceae in an eastern Brazilian wet forest. *Biotropica* 15: 68–70.
- Murillo-A, J., P. E. Ruiz, L. R. Landrum, T. F. Stuessy and M. H. Barfuss. 2012. Phylogenetic relationships in myrceugenia (Myrtaceae) based on plastid and nuclear DNA sequences. *Molecular Phylogenetics and Evolution* 62: 764–776.
- Murillo-A, J., T. F. Stuessy, and E. Ruiz. 2013. Phylogenetic relationships among *Myrceugenia*, *Blepharocalyx*, and *Luma* (Myrtaceae) based on paired-sites models and the secondary structures of ITS and ETS sequences. *Plant Systematics and Evolution* 299: 713–729.
- Murillo-A, J. C., T. F. Stuessy, and E. Ruiz. 2016. Explaining disjunct distributions in the flora of southern South America: Evolutionary history and biogeography of *Myrceugenia* (Myrtaceae). *Journal of Biogeography* 43: 979–990.
- Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, and M. H. H. Stevens. 2007. The vegan package. *Community Ecology Package* 10: 719.
- Oliveira-Filho, A. T., and M. A. L. Fontes. 2000. Patterns of floristic differentiation among Atlantic forests in southeastern Brazil and the influence of climate. *Biotropica* 32: 793–810.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.
- Philippe, H. 2000. Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proceedings of the Royal Society B: Biological Sciences* 267, 1213–1221.
- Philippe, H., and B. Roue. 2011. Difficult phylogenetic questions: more data, maybe; better methods, certainly. *BMC Biology* 9: 91.
- Pittarelli, B. F. S., K. S. M. Mourão, and M. Thadeo. 2021. Pericarp development in *Campomanesia* Ruiz & Pav. (Myrtaceae) species and systematic implications for the genus. *Flora* 282: 151885.
- Pittarelli, B. F. d. S., K. S. M. Mourão, and M. Thadeo. 2021. Pericarp development in *Campomanesia* Ruiz & Pav. (Myrtaceae) species and systematic implications for the genus. *Flora: Morphology, Distribution, Functional Ecology of Plants* 282.
- Poole, I., R. J. Hunt, and D. J. Cantrill. 2001. A fossil wood flora from King George Island: Ecological implications for an Antarctic eocene vegetation. *Annals of Botany* 88: 33–54.
- POWO. 2023. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Website: <http://www.plantsoftheworldonline.org/> [accessed 1 September 2023].
- Proença, C. 1990. A revision of *Siphoneugena* Berg. *Edinburgh Journal of Botany* 47: 239–271.
- Proença, C. E. B., J. E. Q. Faria, A. Giaretta, E. J. Lucas, V. G. Staggemeier, A. C. Tuler, and T. N. C. Vasconcelos. 2020. Nomenclatural and taxonomic changes in tribe Myrteae (Myrtaceae) spurred by molecular phylogenies. *Heringeriana* 14: 49–61.
- Proença, C. E. B., A. C. Tuler, E. J. Lucas, T. N. D. C. Vasconcelos, J. E. Q. De Faria, V. G. Staggemeier, P. S. De-Carvalho, et al. 2022. Diversity, phylogeny and evolution of the rapidly evolving genus *Psidium* L. (Myrtaceae, Myrteae). *Annals of Botany* 129: 367–388.
- Ragonese, A. M. 1980. Leños fosiles de dicotiledoneas del paleoceno de Patagonia argentina. *Ameghiniana* 17: 297–311.
- Reginato, M., T. N. 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.
- Rye, B. L. 1979. Chromosome number variation in the Myrtaceae and its taxonomic implications. *Australian Journal of Botany* 27: 547–573.
- Santos, L. L. dos, F. Forest, D. F. Lima, M. F. Sales, T. N. Vasconcelos, V. G. Staggemeier, and E. Lucas. 2021. Phylogenetic and biogeographic analysis in *Myrcia* sect. *Myrcia* (*Myrcia* s.l., Myrtaceae) with focus on highly polyphyletic *Myrcia splendens*. *International Journal of Plant Sciences* 182: 778–792.
- Santos, M. F., E. J. Lucas, P. T. Sano, S. Buerki, V. G. Staggemeier, and F. Forest. 2017. Biogeographical patterns of *Myrcia* s.l. (Myrtaceae) and their correlation with geological and climatic history in the Neotropics. *Molecular Phylogenetics and Evolution* 108: 34–48.
- Santos, M. F., P. T. Sano, F. Forest, and E. J. Lucas. 2016. Phylogeny, morphology and circumscription of *Myrcia* sect. *Sympodiomyrcia* (*Myrcia* s.l., Myrtaceae). *Taxon* 65: 759–774.
- Santos, T. R. R., J. R. R. Pinto, E. Lenza, and H. A. Mews. 2012. The tree-shrub vegetation in rocky outcrop cerrado areas in Goiás State, Brazil. *Brazilian Journal of Botany* 35: 281–294.
- Schilling, E. E. 2011. Hybrid genera in *Liatrinae* (Asteraceae: Eupatorieae). *Molecular Phylogenetics and Evolution* 59: 158–167.
- Schliep, K. P. 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics* 27: 592–593.

- Schuster, T. M., S. D. Setaro, J. F. G. Tibbits, E. L. Batty, R. M. Fowler, T. G. B. McLay, S. Wilcox, et al. 2018. Chloroplast variation is incongruent with classification of the Australian bloodwood eucalypts (genus *Corymbia*, family Myrtaceae). *PLoS One* 13: 1–28.
- Seger, G. D. S., L. D. S. Duarte, V. J. Debastiani, A. Kindel, and J. A. Jarenkow. 2013. Discriminating the effects of phylogenetic hypothesis, tree resolution and clade age estimates on phylogenetic signal measurements. *Plant Biology* 15: 858–867.
- Silva, A. C. da C., D. M. de Oliveira, and L. J. Gomes. 2022. What does the list of Brazilian sociobiodiversity species of food value show us? *Rodriguesia* 73: e00472021.
- Simmons, M.P. 2012. Radical instability and spurious branch support by likelihood when applied to matrices with non-random distributions of missing data. *Molecular Phylogenetics and Evolution* 62: 472–484.
- Smith, S. A., and J. W. Brown. 2018. Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany* 105: 302–314.
- Smith, S. A., M. J. Moore, J. W. Brown, and Y. Yang. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evolutionary Biology* 15: 1–15.
- Smith, S. A., and J. F. Walker. 2019. PyPHLAWD: A python tool for phylogenetic dataset construction. *Methods in Ecology and Evolution* 10: 104–108.
- Souza, M. D. C., 2009. Estudos taxonômicos em Myrtaceae no Brasil: Revisão de *Neomitranthes* Kausel ex D. Legrand e contribuição ao conhecimento da diversidade e conservação de *Plinia* L. no Domínio Atlântico (Doctoral dissertation, Ph.D. Thesis. Jardim Botânico do Rio de Janeiro. Rio de Janeiro, Brazil).
- de Souza, R. G., M. L. Dan, M. A. Dias-Guimarães, L. A. O. P. Guimarães, and J. M. A. Braga. 2018. Fruits of the Brazilian Atlantic forest: Allaying biodiversity conservation and food security. *Anais da Academia Brasileira de Ciências* 90: 3583–3595.
- de Souza Neto, J. D., E. K. dos Santos, E. Lucas, N. M. Vetö, O. Barrientos-Díaz, V. G. Staggemeier, T. Vasconcelos, and A. C. Turchetto-Zolet. 2022. Advances and perspectives on the evolutionary history and diversification of Neotropical Myrteae (Myrtaceae). *Botanical Journal of the Linnean Society* 199: 173–195.
- Stace, C., C. D. Preston, D. A. Pearman, and R. J. Abbott. 2015. Hybrid Flora of the British Isles. Botanical Society of Britain and Ireland, Bristol, United Kingdom.
- Stadnik, A. M. S. 2020. Estudos Sistemáticos na subtribo Pliniinae E.Lucas & T.Vasc. (Myrtaceae). PhD thesis, Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, Brazil.
- Staggemeier, V. G., E. Cazetta, and L. P. C. Morellato. 2017. Hyperdominance in fruit production in the Brazilian Atlantic rain forest: The functional role of plants in sustaining frugivores. *Biotropica* 49: 71–82.
- Staggemeier, V. G., J. A. F. Diniz-Filho, F. Forest, and E. J. Lucas. 2015b. Phylogenetic analysis in *Myrcia* section *Aulomyrcia* and inferences on plant diversity in the Atlantic rainforest. *Annals of Botany* 115: 747–761.
- Staggemeier, V. G., J. A. F. Diniz-Filho, and L. P. C. Morellato. 2010. The shared influence of phylogeny and ecology on the reproductive patterns of Myrteae (Myrtaceae). *Journal of Ecology* 98: 1409–1421.
- Staggemeier, V. G., J. A. F. Diniz-Filho, V. B. Zipparro, E. Gressler, E. R. Castro, F. F. Mazine-Capelo, I. R. Costa, et al. 2015a. Clade-specific responses regulate phenological patterns in Neotropical Myrtaceae. *Perspectives in Plant Ecology, Evolution and Systematics* 17: 476–490.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- Stefanello, M. É. A., A. C. R. F. Pascoal, and M. J. Salvador. 2011. Essential oils from neotropical Myrtaceae: Chemical diversity and biological properties. *Chemistry and Biodiversity* 8: 73–94.
- Thornhill, A. H., S. Y. W. Ho, C. K. Kühn, and M. D. Crisp. 2015. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution* 93: 29–43.
- Thornhill, A. H., L. W. Popple, R. J. Carter, S. Y. W. Ho, and M. D. Crisp. 2012. Are pollen fossils useful for calibrating relaxed molecular clock dating of phylogenies? A comparative study using Myrtaceae. *Molecular Phylogenetics and Evolution* 63: 15–27.
- Tosolini, A. M. P., D. J. Cantrill, and J. E. Francis. 2021. Paleocene high-latitude leaf flora of Antarctica Part 1: entire-margined angiosperms. *Review of Palaeobotany and Palynology* 285: 104317.
- Valadão, H., C. E. B. Proença, M. P. Kuhlmann, S. A. Harris, and R. Tidon. 2019. Fruit-breeding drosophilids (Diptera) in the Neotropics: playing the field and specialising in generalism? *Ecological Entomology* 44: een.12769.
- Vasconcelos, T., J. D. Boyko, and J. M. Beaulieu. 2023. Linking mode of seed dispersal and climatic niche evolution in flowering plants. *Journal of Biogeography* 50: 43–56.
- Vasconcelos, T. N. C., M. Chartier, G. Prenner, A. C. Martins, J. Schönenberger, A. Wingler, and E. Lucas. 2019a. Floral uniformity through evolutionary time in a species-rich tree lineage. *New Phytologist* 221: 1597–1608.
- Vasconcelos, T. N. C., E. J. Lucas, J. E. Q. Faria, and G. Prenner. 2018. Floral heterochrony promotes flexibility of reproductive strategies in the morphologically homogeneous genus *Eugenia* (Myrtaceae). *Annals of Botany* 121: 161–174.
- Vasconcelos, T. N. C., G. Prenner, M. O. Bünger, P. S. Carvalho, A. Wingler, and E. J. Lucas. 2015. Systematic and evolutionary implications of stamen position in Myrteae (Myrtaceae). *Botanical Journal of the Linnean Society* 179: 388–402.
- Vasconcelos, T. N. C., G. Prenner, and E. J. Lucas. 2019b. A Systematic Overview of the Floral Diversity in Myrteae (Myrtaceae). *Systematic Botany* 44: 570–591.
- Vasconcelos, T. N. C., G. Prenner, M. F. Santos, A. Wingler, and E. J. Lucas. 2017a. Links between parallel evolution and systematic complexity in angiosperms—A case study of floral development in *Myrcia* s.l. (Myrtaceae). *Perspectives in Plant Ecology, Evolution and Systematics* 24: 11–24.
- Vasconcelos, T. N. C., C. E. B. Proença, B. Ahmad, D. S. Aguilar, R. Aguilar, B. S. Amorim, K. Campbell, et al. 2017b. Myrteae phylogeny, calibration, biogeography and diversification patterns: Increased understanding in the most species rich tribe of Myrtaceae. *Molecular Phylogenetics and Evolution* 109: 113–137.
- Wilson, C. E., F. Forest, D. S. Devey, and E. J. Lucas. 2016. Phylogenetic relationships in *Calyptanthes* (Myrtaceae) with particular emphasis on its monophyly relative to *Myrcia* s. l. *Systematic Botany* 41: 378–386.
- Wilson, P. G. 2011. Myrtaceae. In K. Kubitzki [ed.], *The Families and Genera of Vascular Plants*. Vol. 10. Flowering Plant Eudicots: Sapindales, Cucurbitales, Myrtaceae, 212–217. Springer, Heidelberg, Germany.
- Wilson, P. G., M. M. O'Brien, M. M. Heslewood, and C. J. Quinn. 2005. Relationships within Myrtaceae sensu lato based on a matK phylogeny. *Plant Systematics and Evolution* 251: 3–19.
- Zhong, B., O. Deusch, V. V. Goremykin, D. Penny, P. J. Biggs, R. A. Atherton, S. V. Nikiforova, and P. J. Lockhart. 2011. Systematic error in seed plant phylogenomics. *Genome Biology and Evolution* 3: 1340–1348.
- Zipparro, V. B., F. A. G. Guilherme, R. J. Almeida-Scabbia, and L. P. C. Morellato. 2005. Levantamento florístico de Floresta Atlântica no sul do estado de São Paulo, Parque Estadual Intervales, base Saibadela. *Biota Neotropica* 5: 147–170.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Primers, polymerase chain reaction (PCR), and fossil information.

Appendix S2. Vouchers and GenBank accession numbers for all tips sampled in the supermatrix tree.

Appendix S3. Concatenated alignment used to generate the supermatrix tree.

Appendix S4. Maximum likelihood supermatrix tree depicting bootstrap supports.

Appendix S5. Time-calibrated supermatrix tree depicting ages and confidence intervals for all nodes.

Appendix S6. Time-calibrated supermatrix tree in Newick format.

Appendix S7. Time-calibrated supermatrix tree in Newick format with only one tip for each species kept.

Appendix S8. Time-calibrated supermatrix tree in Newick format with only one tip for each species kept and outgroups pruned.

How to cite this article: The Neotropical Myrtaceae Working Group, V. G. Staggemeier, B. Amorim, M. Bünger, I. R. Costa, J. E. Q. de Faria, J. Flickinger, A. Giaretta, M. T. Kubo, D. F. Lima, L. L. dos Santos, A. R. Lourenço, E. Lucas, F. F. Mazine, J. Murillo-A, M. I. U. de Oliveira, C. Parra-O, C. E. B. Proença, M. Reginato, P. O. Rosa, M. F. Santos, A. Stadnik, A. C. Tuler, K. S. Valdemarin, and T. Vasconcelos. 2024. Towards a species-level phylogeny for Neotropical Myrtaceae: Notes on topology and resources for future studies. *American Journal of Botany* 111(5): e16330. <https://doi.org/10.1002/ajb2.16330>