



Molecular phylogenetics, phylogenomics, and phylogeography

UCE phylogenomics, biogeography, and classification of long-horned bees (Hymenoptera: Apidae: Eucerini), with insights on using specimens with extremely degraded DNA

Felipe V. Freitas^{1-4,*}, Michael G. Branstetter², Vinícius H. Franceschini-Santos^{1,5}, Achik Dorchin^{6,7}, Karen W. Wright^{8,9}, Margarita M. López-Urbe¹⁰, Terry Griswold², Fernando A. Silveira^{11,†}, Eduardo A.B. Almeida¹

¹Laboratório de Biologia Comparada e Abelhas (LBCA), Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, SP, Brazil, ²U.S. Department of Agriculture, Agricultural Research Service (USDA-ARS), Pollinating Insects Research Unit, Utah State University, Logan, UT, USA, ³Departamento de Ciências Biológicas, Instituto de Biociências, Letras e Ciências Exatas (IBILCE), Universidade Estadual Paulista 'Júlio de Mesquita Filho' (UNESP), São José do Rio Preto, SP, Brazil, ⁴Department of Entomology, Washington State University, Pullman, WA, USA, ⁵Division of Gene Regulation and Onco Institute, Netherlands Cancer Institute, Amsterdam, The Netherlands, ⁶Laboratory of Zoology, Research Institute for Biosciences, University of Mons, Mons, Belgium, ⁷Department of Biology—Invertebrates, Royal Museum for Central Africa, Tervuren, Belgium, ⁸Plant Protection Division, Washington State Department of Agriculture, Yakima, Washington, USA, ⁹Sevilleta Long-Term Ecological Research, University of New Mexico, Albuquerque, NM, USA, ¹⁰Department of Entomology, The Pennsylvania State University, University Park, PA, USA, ¹¹Laboratório de Sistemática de Insetos, Departamento de Zoologia, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil†In memoriam *Corresponding author, e-mail: felipevieirafreitas@gmail.com; f.vieiradefreitas@wsu.edu

Subject Editor: Jessica Gillung

Received on 24 November 2022; revised on 19 April 2023; accepted on 29 May 2023

Long-horned bees (Apidae, Eucerini) are found in different biomes worldwide and include some important crop pollinators. In the Western Hemisphere, Eucerini received extensive taxonomic study during the twentieth century, resulting in several revisions of its genera. In contrast, progress on eucerine phylogenetic research and the genus-level classification has been slow, primarily due to the relatively homogeneous external morphology within the tribe and the rarity of many of its species in collections. Here, we present a comprehensive phylogenetic study of Eucerini based on ultraconserved elements, including 153 species from nearly all genera and subgenera and from all biogeographic regions where they occur. Many of these specimens are from museums and were collected as far back as 1909. We discuss the challenges of working with specimens with highly degraded DNA, present insights into improving phylogenetic results for both species-tree and concatenation approaches, and present a new pipeline for UCE curation (Curation of UltraconserveD Elements—CURE). Our results show the existence of seven main lineages in Eucerini and most of the genera and subgenera to be reciprocally monophyletic. Using a comprehensive and up-to-date phylogenetic framework, we: (1) propose taxonomic changes, including a new subtribal classification and reorganized generic and subgeneric limits; (2) estimate divergence times; and (3) conduct a detailed exploration of historical biogeography of long-horned bees. We find that eucerine lineages expanded their range onto most continents only after their initial diversification in southern South America during the Eocene.

Key words: CURE, ddBD, MCMCtree, museomics, UCE curation

Introduction

Phylogenetic relationships among bees have been investigated using molecular data since the early 1990s (Cameron 1992, 1993, Sheppard and McPherson 1991), and these studies have brought significant advances to the understanding of bee evolution (Danforth et al. 2013, 2019). The recent development of high-throughput sequencing and ultraconserved element (UCE) phylogenomics has catalyzed a revolution in the exploration of phylogenetic relationships and divergence times of bees and closely related groups (Faircloth et al. 2015, Branstetter et al. 2017, Blaimer et al. 2018, 2023, Bossert et al. 2019, Freitas et al. 2021, 2022, Sless et al. 2022). UCE markers have been highly effective for phylogenetic research, lowering costs to produce vast amounts of data and enabling the use of older museum specimens, which often include species from areas difficult to sample or even from extinct species (Blaimer et al. 2015, 2016, McCormack et al. 2016). UCes are frequently highlighted for allowing the obtention of large amounts of data from museum specimens (McCormack et al. 2016, Blaimer et al. 2018, Derkarabetian et al. 2019). However, discussions regarding the quality of the data obtained from these specimens and its effect on phylogenetic results are still scarce. Most discussions regarding the impact of including museum specimens in studies using UCE data are limited to the results of concatenation analyses. The effects of using data from museum specimens on Multispecies Coalescent (MSC) model analyses remain underexplored (Smith et al. 2020, Salter et al. 2022).

Despite recent progress in bee phylogenetics research, some large and relevant clades (especially tribes) of bees remain understudied. One such taxon is the tribe Eucerini (Hymenoptera: Apidae: Eucerinae), colloquially known as long-horned bees (Fig. 1). This group includes more than 750 extant species and occurs in all zoogeographic regions of the world, except for Antarctica and the Australian region (Michener 2007). Eucerine bees are common elements of the bee fauna in many different biomes and regions (e.g., Brazilian Atlantic Forest—Cure et al. 1992, Silveira et al. 1993; Cerrado—Silveria and Campos 1995, Santos et al. 2004; the Mediterranean region—Dafni and O’Toole 1994; the Argentinean Monte—Michelette and Camargo 2000; and warm deserts in the southwestern United States and northern Mexico—Minckley and Radke 2021) and include species that specialize on the pollination of major crops such as pumpkins, squash, gourds [Cucurbitaceae], and sunflowers [Asteraceae] (Hurd et al. 1971, Parker 1981a, 1981b, López-Urbe et al. 2016). Because of their diversity and importance as pollinators, a better understanding of the higher-level classification of this group of bees is necessary.

The current classification of the Eucerini comprises about 30 genera, divided into more than 50 subgenera (Michener 2007, Moure et al. 2012, Ascher and Pickering 2021). In the Western Hemisphere, the species and genus-level taxonomy of the Eucerini is relatively well resolved, thanks to seminal papers that delimited most of the genera (Michener et al. 1955, Moure and Michener

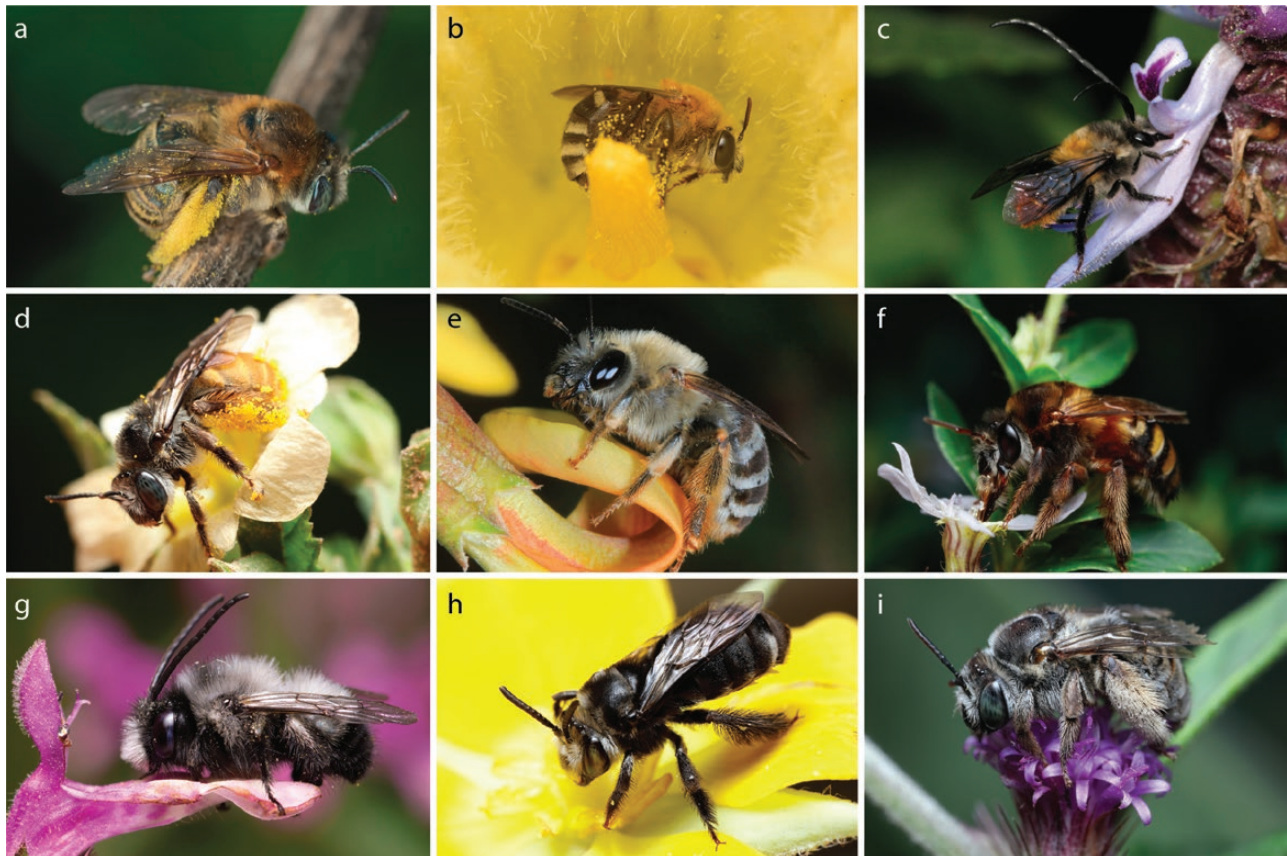


Fig. 1. Representative species of Eucerini: a) *Melissodes (Ecleptica)* sp. ♀ [Melissodina], b) *Xenoglossa (Peponapis) fervens* (Smith, 1879) ♀ [Eucerina], c) *Thygater (Thygater) aethiops* (Smith 1854) ♀ [Thygaterina], d) *Melissoptila* sp. ♀ [Melissodina], e) *Eucera (Synhalonia)* sp. ♀ [Eucerina], f) *Thygater (Thygater) analis* (Lepeletier 1841) ♀ [Thygaterina], g) *Alloscirtetica gayi* (Spinola 1851) ♀ [Alloscirteticina], h) *Pachysvastra leucocephala* (Bertoni and Schrottky 1910) ♀ [Gaesischiina], i) *Gaesischia (Gaesischia) nigra* (Moure 1968) ♀ [Gaesischiina]. Credits of the pictures: a, b, d, f, h, i—Adriana Tiba and Julio Pupin; c—Sergio Jansen-Gonzalez, g—Javier Gross; and e—Shan Gui.

1955, Michener and Moure 1957). Extensive taxonomic treatments of the eucerine genera in this region were conducted by two important bee taxonomists, Danuncia Urban (Martins et al. 2015) and Wallace LaBerge (Rasmussen et al. 2013). These authors have provided descriptions and keys allowing for the identification of most of the species-level diversity (e.g., LaBerge 1956a, 1956b, 1961, 2001, Urban 1967a, 1967b, 1967c, 1968a, 1968b, 1971, 1972, 1973, 1982a, 1982b, 1998). In the Eastern Hemisphere, however, weak morphological boundaries among taxa have led to unstable generic delimitation that hampered full taxonomic treatment (Michener 2007, Dorchin et al. 2018). More recently, efforts have been made to improve eucerine classification with the support of molecular data (Freitas and Silveira 2017, Dorchin et al. 2018, Dorchin et al. 2018, Freitas et al. 2018, 2019). These recent investigations have contributed to eucerine systematics and historical biogeography. However, these recent studies focused only on a few genera/lineages and were limited in their geographical coverage. Consequently, additional work is needed to more thoroughly assess the limits of morphologically complex genera and their phylogenetic relationships.

Here we present a comprehensive phylogenetic analysis of long-horned bees (Apidae: Eucerini) based on UCE phylogenomic data and more thorough taxon sampling. The resulting phylogenetic hypothesis is used to reassess the subtribal classification of Eucerini, resulting in the proposition of four new subtribes and the re-organization of many genera and subgenera. Additionally, we investigate historical biogeography, exploring the following questions: (i) which regions of South America were most relevant for the early diversification of Eucerini; (ii) when did faunal exchanges among continents occur; and (iii) which paleoclimatic and geologic events likely influenced range expansions. Finally, we consider challenges associated with including historical specimens in UCE phylogenomics, particularly for gene-tree species-tree analyses, and present a new pipeline (Curation of UltraconserveRved Elements—CURE) for automating UCE curation that can improve gene-tree estimation and model fit in general.

Material and Methods

Classification and Taxon Sampling

All genera and subgenera of Eucerini recognized in current classifications (Michener 2007, Moure et al. 2012, Dorchin et al. 2018) were sampled, except for two taxa of Neotropical Eucerini: *Lophothygater* and *Trichocerapis* (*Dithygater*). We included as many species as possible to represent genera and subgenera with questionable monophyly, dubious phylogenetic placement, or taxa with extensive geographic range and morphological diversity. As outgroups, we chose representatives of all the five remaining tribes of Eucerinae *sensu* Bossert et al. (2019): two species of Ancyloini, two Ancyloscelidini *sensu* Freitas et al. (2021), four Emphorini, three Exomalopsini, and three species of Tapinotaspidini. In total, we included 164 species of Eucerinae in our dataset. The complete taxon sampling is listed in Supplementary Table S1.

Ultraconserved Element Data Generation

Ultraconserved elements were sequenced by combining target enrichment of UCES with multiplexed, next-generation sequencing (Faircloth et al. 2012, Branstetter et al. 2017). DNA extraction, library preparation, UCE enrichment, sample pooling, and next-generation sequencing followed the same procedures described in Freitas et al. (2021), using the ‘bee-ant-specific hym-v2’ bait set (Branstetter et al. 2017, Grab et al. 2019).

Bioinformatics and Matrix Assembly

The PHYLUCE v1.7.1 toolkit (Faircloth 2015) was used to clean raw reads with Illuminiprocessor (Faircloth 2013, Bolger et al. 2014), assemble reads into contigs with SPAdes (Bankevich et al. 2012), extract UCE contigs, align contigs with MAFFT (Katoh and Standley 2013), and trim alignments with GBLOCKS (Castresana 2000, Talavera and Castresana 2007), using its relaxed parameters (b1: 0.5, b2: 0.85, b3: 8, e b4: 10). We used the default parameters for all these procedures.

Using alignments from the above steps, a concatenated matrix was created, and a preliminary tree was estimated in IQ-TREE v2.1.3 (Minh et al. 2020), with each UCE locus set as a partition and GTR + I + G used as the substitution model for all partitions. Using this matrix as input, we performed additional data trimming with the program SPRUCEUP (Borowiec 2019), which is an effective tool for trimming poorly aligned sequences from individual samples row-by-row instead of column-by-column. By trimming positions in poorly aligned samples, this program removes or reduces artificially long branches that often occur in extremely old and/or low-quality samples. SPRUCEUP was run under the following parameters: distances were calculated using the Jukes–Cantor correction method, with a window size of 20 bp, an overlap size of 15 bp, under a lognormal distribution, a general cutoff value of 0.98, and specific manual cutoffs values for each one of those terminals showing abnormally long branches in the preliminary tree [*Svaetropsis bipunctata* (Friese 1908), **comb. n.:** 0.06; *Tetralonia viator* (Cockerell, 1911), **comb. n.:** 0.06; *Thygater* (*Thygater*) *melanotricha* (Urban, 1967), **comb. n.:** 0.05; *Florilegus* (*Florilegus*) *lanierii* (Guérin-Méneville, 1845), **comb. n.:** 0.09; *Florilegus* (*Florilegus*) *purpurascens* (Cockerell, 1914), **comb. n.:** 0.06; *Dasyhalonia* (*Pachyhalonia*) *cearensis* (Ducke, 1910), **comb. n.:** 0.07; *Tetralonia friesei* (Meade-Waldo, 1914), **comb. n.:** 0.1; *Eucera* (*Cubitalia*) *morio* (Friese, 1922), **comb. n.:** 0.07; *Epimelissodes* (*Brachymelissodes*) *ressoni* (Dalla Torre, 1896), **comb. n.:** 0.09; *Epimelissodes* (*Idiomelissodes*) *duplocincta* (Cockerell, 1905), **comb. n.:** 0.095], no guiding tree was used. Cutoff values were defined after a few attempts evaluating both distance distribution plots and resulting branch lengths in testing trees. Afterward, trimmed matrices were re-split into individual alignments using PHYLUCE and AMAS (Borowiec 2016).

Using modified scripts from van Dam et al. (2021), we identified the genomic regions [genic (exon or intron) or intergenic] of all UCE loci by mapping loci to the latest *Apis mellifera* Linnaeus reference genome (Amel_Hav3.1). We then merged UCES belonging to the same genomic area with two approaches: by genes and by region, generating two matrices. In the former approach, we merged all UCES in the same gene and treated each genic region (exons and introns) as an independent partition—this matrix is referred to as “Genes”. In the latter, only UCES in the same exon were merged—referred to as “Exons”. In both approaches, we did not combine UCES in intergenic regions.

In addition to the two matrices described above, we created two more matrices in which only UCES with at least 300bp were selected: One containing only UCES with at least 75% of the terminals present, referred to as UCES_75p and another containing only UCES comprising at least 98% (UCES_98p) of the terminals. The UCES_75p supermatrix was analyzed partitioned by UCE locus and by CURE regions. The UCES_98p matrix was used to estimate divergence times only, as described below.

Phylogenetic Analyses

Individual UCE loci were submitted to Sliding-Window Site Characteristics for phylogenetic analyses based on entropy

(SWSC-EN, Tagliacollo and Lanfear 2018) to identify and delimit the core, the right flank, and the left flank regions. This approach has shown good performance in defining partitions for UCE datasets, resulting in improved model fit (Tagliacollo and Lanfear 2018, Branstetter and Longino 2019, Branstetter et al. 2021, Freitas et al. 2021). It was also implemented in our pipeline CURE (see next section) to run in parallel, which speeds up considerably the execution time.

All matrices were used to estimate the species-tree by concatenation using Maximum Likelihood (ML) in IQ-TREE2. We let IQ-TREE search for the best-fit partitioning scheme and for the best substitution model for each partition (-m MFP+MERGE -rclusterf 10) using ModelFinder2 (Kalyaanamoorthy et al. 2017). UFBoot (Minh et al. 2013, Hoang et al. 2018) and SH-like approximate likelihood ratio test (Guindon et al. 2010) scores were calculated with 1,000 replicates each.

To estimate species-trees using the MSC model, which considers potentially divergent gene histories, gene-trees were estimated for all loci in the Genes, Exons, and UCEs_75p matrices. Gene-trees from individual loci (here, the loci are a combination of all UCEs in the same gene) in the Genes dataset were analyzed with each exon and intron defined as one partition. Gene-trees from unmerged UCEs were analyzed partitioned according to SWSC-EN regions following the strategy proposed by Freitas et al. (2021), which showed a significant increase in gene-tree average bootstrap support compared to unpartitioned gene-tree searches. Gene-trees for the Exons dataset were generated with each locus considered a single partition. The gene-trees in each matrix were summarized by applying the MSC in ASTRAL-MP v5.7.1 (Zhang et al. 2018, Yin et al. 2019), using default settings, and calculating branch support as local posterior probabilities.

Curation of UltraconserVed Elements—CURE

Despite not being newly proposed approaches for UCE curation, both (1) identifying and merging UCEs in the same exons and genes and (2) partitioning UCEs according to their core and flanking regions to generate gene-trees, were strategies without fully automated or scalable implementations. To do so, we implemented these procedures in a scalable tool (Curation of UltraconserVed Elements—CURE, available at: <https://github.com/vhfsantos/CURE>). CURE uses GNU-parallel (Tange 2018) to automatically and in parallel run: (a) the scripts from van Dam et al. (2021); (b) the SWSC-EN scripts from Tagliacollo and Lanfear (2018); (c) exporting of all the outputs necessary to run the approach described by Freitas et al. (2021) using the results from the SWSC-EN run; and (d) wrapper scripts for estimating gene-trees in parallel. CURE exports alignments and their respective charsets files ready to be analyzed by software for phylogenetic analyses for all these implementations.

Divergence Time Estimation

To select the most appropriate loci for divergence time estimations, three metrics were calculated for each one of the 737 loci in the UCEs_98p dataset: mean gene-tree bootstrap support; the number of parsimony informative sites (nPIS), both measures of the quality and signal; and root-to-tip distance of gene-trees, with smaller root to tip distance meaning a more clocklike locus. All these metrics were calculated using the script good genes (https://github.com/marekborowiec/good_genes, Borowiec et al. 2015) and AMAS. To reach an alignment with approximately 20k bp, a number that has been previously shown to be adequate for precise divergence time estimation (Freitas et al. 2022), we selected only

those loci ranked in the top 300 for all three metrics at the same time. After this first selection, the worst loci according to root-to-tip distance ($n = 14$) were discarded to have the number of bp desired. The final dataset for divergence time estimation included 27 loci.

The loci selected for divergence time estimation were analyzed as three partitions according to the results of SWSC-EN, one including all the core regions, the second with all the left flanks, and the third one including all the right flanks, attempting to improve precision by combining regions with similar substitution rates (dos Reis and Yang 2019).

Divergence time estimation was conducted in the software MCMCtree 4.9j, part of package PAML (Yang 1997, 2007), using the approximate likelihood method (dos Reis and Yang 2011), under a birth-death diversification model (Kendall 1948, Nee et al. 1994, Yang and Rannala 2006). The substitution model used was HKY, with five gamma categories. A preliminary substitution rate estimation was conducted in baseml to inform the rgene_gamma prior to subsequent MCMCtree analyses. We selected the independent rates (clock = 2 in MCMCtree control file) clock model, which is appropriate for datasets with none or small numbers of calibration points and when the rate variation among branches is high (Brown and Yang 2011, Freitas et al. 2022). The analyses were conducted using a data-driven birth-death (ddBD) diversification model, a recently proposed approach that increases the precision and accuracy of divergence time estimation when no or only a few calibration points are available (Tao et al. 2021). For implementing ddBD priors, a preliminary divergence time analysis was conducted in RelTime (Tamura et al. 2012) implemented in MEGA X (Kumar et al. 2018) without calibrations and using default parameters. The tree resulting from RelTime was used as input to run the script that calculates the Birth and Death prior values. We used the tree from ASTRAL generated using the gene-trees from the Genes dataset, but with the placement of some terminals—those with extremely high levels of missing data (highlighted in Fig. 3)—corrected manually according to their position in the trees generated through concatenated analyses. Only one specimen per species was kept, and all outgroups were removed except those of Ancylini. These manual modifications in the ASTRAL tree were done using Mesquite 3.61 (Maddison and Maddison 2019). We ran all MCMCtree analyses twice for 10^7 Markov chain Monte Carlo (MCMC) generations (sampling every 10k generations), with a burn-in of 50k iterations. Each run's convergence was assessed by calculating effective sample size (ESS) values for all nodes and parameters using the package “coda” (Plummer et al. 2006) in R (R Core Team 2020). Convergence was also assessed by visually inspecting trace plots and posterior distributions of ages for nodes using a modified version of the script made available by dos Reis and Yang (2019). We also confirmed that all ESS values were above 200 for both runs.

The root (Eucerini + Ancylini) was calibrated according to the 95% height posterior density (HPD) of the corresponding node in the results of Freitas et al. (2022), which included a dense sampling of the entire subfamily Eucerinae. We used the bounds 45–70 Mya and a skew-normal distribution, with the following parameters: location-0.45, scale-0.1, and shape-50 adjusted in the R package “MCMCtreeR” (Puttick and Schwartz 2019).

Biogeography

Biogeographical reconstructions were conducted in RevBayes (Höhna et al. 2016). We used a Bayesian implementation of the DEC model (Ree 2005, Ree and Smith 2008, Landis et al. 2018),

following the scripts for unstratified analyses available on the RevBayes web page, although not accounting for geographic distances. We ran two MCMC chains for 10^7 generations. Convergence was inferred when ESS values of all parameters in both runs were, before and after being combined, above 200. These values were obtained through Tracer 1.7 (Rambaut et al. 2018).

To avoid having an excessive number of areas (states) included in a single analysis, which could make the reconstruction computationally intractable, we broke the analysis into two parts. Initially, we investigated the history of Eucerini in South America, where the group is known to have originated (Freitas et al. 2022). Six areas were delimited (five in South America and the rest of the world coded as another area), considering the distribution of species in the present and areas already identified as relevant for other organisms in this region (Rueda et al. 2013, Varela et al. 2019, Freitas et al. 2022). These areas in South America were: (a) Andean—including the Andes Mountains, its western slope, and Patagonia; (b) southern South America; (c) central South America; (d) eastern South America; and (e) Amazonian. Following this first analysis, we investigated how the occupation of other zoogeographical regions occurred. A set of six different areas were delimited (South America, western North America, eastern North America, Palearctic, Afrotropical, and Oriental), considering classical zoogeographic regions of the world (Holt et al. 2013, Rueda et al. 2013) and how eucerine taxa are currently distributed. In both analyses, the maximum number of areas allowed to be occupied at the same time was three.

As our taxon sampling accounts for all geographic variation in each genus and subgenus, we coded only the distributions of the species included in our taxon sampling, considering the distributional data available in online catalogs (Moure et al. 2012, Ascher and Pickering 2021).

Nomenclature

This paper and the nomenclatural acts it contains have been registered in Zoobank (www.zoobank.org), the official register of the International Commission on Zoological Nomenclature. The LSID (Life Science Identifier) number of the publication is: urn:lsid:zoobank.org:pub:31786887-D775-450E-AC10-BC9801ABF846.

Results

Taxon Sampling and Matrix Generation

We captured a total of 2,495 UCEs from 156 specimens belonging to 151 species, of which 80 were sequenced for the first time. These represent 19.7% of the valid species in the tribe Eucerini. After alignment and trimming, we assembled a total of five matrices from these data (Table 2). Eight samples were collected >30 years ago and had extreme levels of missing data (~90%), with values considerably below the mean in all datasets (Table 3). For these eight low-quality samples, most of the sequence data mapped to the core region of UCE loci rather than more variable flanking regions (Table 3). In seven of the eight samples, cores represented $\geq 60\%$ of sites, compared to an average of 1.5% in the remaining samples.

The curation of the UCEs using our CURE pipeline assigned 1,040 UCE loci as exonic, 482 as intronic, 392 as exonic and intronic, 333 as intergenic, and 245 as unassigned to region. Concatenating UCE loci identified as part of the same genes produced alignments considerably longer than independent UCEs (a mean of 907bp for genes and 608bp for independent UCEs). For exons, the mean length of loci was almost the same as UCEs; however, longer alignments were produced (Table 2).

Phylogenetic Results

All phylogenetic analyses recovered similar results, identifying seven main clades of the Eucerini, the relationships among them, and the limits of genera and subgenera (Fig. 2). Results of all of the analyses with concatenated data recovered the same tree topology, except for the result from matrix UCEs_75p. The latter recovered the eight low-quality samples in unexpected positions (Fig. 3), differing from the remaining ML analyses (Supplementary Figs. S1–S5). The four terminals with the most missing data (Table 2) were also recovered in seemingly erroneous positions in most ASTRAL results (Fig. 3 and Supplementary files). The minor inconsistent relationships recovered among different phylogenetic results (Fig. 2) did not affect any of the classificatory decisions presented in the Appendix.

Given the strong congruence among the results obtained using different analyses, we confidently propose new subtribes to organize the diversity of Eucerini that reflect the current understanding about these phylogenetic relationships. This follows a trend started by Roig-Alsina and Michener (1993), who recognized three subtribes in Eucerini: Eucerinodina Michener and Moure, to accommodate *Eucerinodina gayi* (Spinola 1851), and *Canephorulina Michener*, LaBerge and Moure, for *Canephorula apiformis* (Friese 1908), both considered relictual taxa; and Eucerina for the bulk of the Eucerini diversity. Considering our phylogenetic results, we propose four new subtribes that comprise monophyletic groups (Appendix), which represent all major groups previously included in Eucerina *sensu* (Roig-Alsina and Michener 1993): Alloscirteticina **new subtribe**, Gaesischina **new subtribe**, Thygaterina **new subtribe**, and Melissodina **new subtribe**. These four new subtribes, in addition to the existing subtribe names Eucerina Robertson, Canephorulina, and Eucerinodina, represent the seven major lineages of Eucerini recognized in our results. Our newly proposed classification and its associations with our phylogenetic hypothesis are summarized in Table 1 and Fig. 4.

Biogeography and the Timeline of Eucerini Diversification

Eucerini started its diversification in the mid-Eocene (~50–35 Mya HPD), but most clades diversified later in the Miocene (~25–5 Mya HPD). As for the subtribes, all except Eucerinodina and Canephorulina began to diverge from each other around the mid-Miocene (~15 Mya), and most genera diversified only after the first half of the Miocene (15 Mya–present) (Fig. 5).

Considering that there were no incompatibilities between the two sets of areas used in the independent biogeographical reconstructions, and that there were no discordances between the two reconstructions (Supplementary Figs. S8 and S9), the results were manually combined into a summary figure (Fig. 6). The biogeographic reconstruction indicates that the initial diversification of Eucerini occurred in an area comprising southern South America and the Andean region, followed by later range expansions northward (Fig. 6). While Eucerinodina, Canephorulina, and most of the species of Alloscirteticina remain associated with southern South America, the remaining groups diversified and dispersed during the Miocene. It was in the early Miocene (~16 Mya) that the ancestor of the remaining Eucerini reached central South America (Fig. 6). During the mid-Miocene (~14 Mya), the ancestor of Melissodina + Eucerina reached North America, and during the late Miocene (~13 Mya), species of Eucerina expanded their distribution into the Palearctic (Fig. 6). The latter event is implied by the reconstructed areas for the nodes uniting *Xenoglossa* plus *Tetralonia* with *Eucera*. The ancestral area reconstructed for this clade comprised western North America and the Palearctic. *Tetralonia* would have reached the Oriental region during the Miocene/Pliocene transition (~5 Mya) and, from there, the Afrotropic during Pliocene/Pleistocene transition (Fig. 6).

Table 1. Summary of the new classification of Eucerini, with data on the number of species known per genus or subgenus, the number of species sampled, and occurrence of each subgenus according to the four main zoogeographic domains: Afrotropical (AF), Nearctic (NA), Neotropical (NT), and Palearctic (PA)

EUCERINI (Robertson 1904)	Spp. total (Sampled)	Occurrence
SUBTRIBE ALLOSCIRTETICINA FREITAS AND SILVEIRA, NEW SUBTRIBE		
Genus <i>Alloscirtetica</i> (Holmberg 1909)	39 (7)	NT
Genus <i>Dasyscirtetica</i> (Michener et al. 1955)	4 (2)	NT
Genus <i>Megascirtetica</i> (Moure and Michener 1955)	1 (1)	NT
SUBTRIBE CANEPHORULINA MICHENER, LABERGE AND MOURE		
Genus <i>Canephorula</i> (Jørgensen 1909)	1 (1)	NT
SUBTRIBE EUCERINA ROBERTSON		
Genus <i>Eucera</i> (Scopoli 1770)		
Subgenus <i>Eucera</i> s. str.	123 (11)	PA
Subgenus <i>E. (Cubitalia)</i> (Friese 1911)	8 (3)	PA
Subgenus <i>E. (Synhalonia)</i> (Patton 1879)	101 (4)	NA, PA, OR
Genus <i>Protohalonia</i> (Dorchin 2018)	3 (1)	NA
Genus <i>Simanthedon</i> (Zavortink 1975)	1 (1)	NA
Genus <i>Tetralonia</i> (Spinola 1838)	100 (6)	AF, PA, OR
Genus <i>Xenoglossa</i> (Smith 1854)		
Subgenus <i>Xenoglossa</i> s. str.	7 (2)	NA, NT
Subgenus <i>X. (Cemolobus)</i> (Robertson 1891)	1 (1)	NA
Subgenus <i>X. (Peponapis)</i> (Robertson 1902)	15 (3)	NA, NT
Subgenus <i>X. (Syntrichalonia)</i> (LaBerge 1957)	2 (2)	NA, NT
Subgenus <i>X. (Xenoglossodes)</i> (Ashmead 1899)	3 (3)*	NA, NT
<i>Xenoglossa, incertae sedis</i> *	38(5)	NA, NT
SUBTRIBE EUCERINODINA MICHENER AND MOURE		
Genus <i>Eucerinoda</i> (Michener and Moure 1957)	1 (1)	NT
SUBTRIBE GAESISCHINA FREITAS AND SILVEIRA, NEW SUBTRIBE		
Genus <i>Dasyhalonia</i> (Michener et al. 1955)		
Subgenus <i>Dasyhalonia</i> s. str.	2 (2)	NT
Subgenus <i>D. (Pachyhalonia)</i> (Moure and Michener 1955)	3 (3)	NT
Genus <i>Florilegus</i> (Robertson 1900)		
Subgenus <i>Florilegus</i> s. str.	5 (4)	NA, NT
Subgenus <i>F. (Euflorilegus)</i> (Ogloblin 1955)	5 (3)	NT
Subgenus <i>F. (Floriraptor)</i> (Moure and Michener 1955)	1 (1)	NT
Genus <i>Gaesischia</i> (Michener et al. 1955)		
Subgenus <i>Gaesischia</i> s. str.	18 (8)	NT
Subgenus <i>G. (Gaesischiana)</i> (Michener et al. 1955)	1 (1)	NA, NT
Subgenus <i>G. (Gaesischiopsis)</i> (Michener et al. 1955)	7 (5)	NT
Subgenus <i>G. (Gaesischioides)</i> (Freitas 2022)	1 (1)	NT
<i>Gaesischia, incertae sedis (G. caracicola, G. cipoana, G. horizontina)</i>	3 (0)	NT
Genus <i>Gaesochira</i> (Moure and Michener 1955)	1 (1)	NT
Genus <i>Hamatothrix</i> (Urban 1989b)	1 (1)	NT
Genus <i>Micromyichapis</i> (Moure and Michener 1955)	1 (1)	NT
Genus <i>Pachysvastra</i> (Moure and Michener 1955)	2 (1)	NT
Genus <i>Platysvastra</i> (Moure 1967)	1 (1)	NT
Genus <i>Santiago</i> (Urban 1989b)	2 (1)	NT
Genus <i>Savannychapis</i> (Freitas 2022)	1 (1)	NT
Genus <i>Svastrina</i> (Moure and Michener 1955)	1 (1)	NT

Table 1. Continued

EUCERINI (Robertson 1904)	Spp. total (Sampled)	Occurrence
SUBTRIBE MELISSODINA FREITAS AND SILVEIRA, NEW SUBTRIBE		
Genus <i>Agapanthinus</i> (LaBerge 1957)	1 (1)	NA
Genus <i>Epimelissodes</i> (Ashmead 1899)		
Subgenus <i>Epimelissodes</i> s. str.	15 (7)	NA, NT
Subgenus <i>E. (Anthedonia)</i> (Michener 1942)	2 (2)	NA
Subgenus <i>E. (Brachymelissodes)</i> (LaBerge 1956a)	2 (1)	NA
Subgenus <i>E. (Idiomelissodes)</i> (LaBerge 1956a)	1 (1)	NA
Genus <i>Martinapis</i> (Cockerell 1929)	2 (2)	NA
Genus <i>Melissodes</i> (Latreille 1829)		
Subgenus <i>Melissodes</i> s. str.	24 (3)	NA, NT
Subgenus <i>M (Apomelissodes)</i> (LaBerge 1956a)	10 (7)	NA
Subgenus <i>M (Callimelissodes)</i> (LaBerge 1961)	14 (3)	NA
Subgenus <i>M (Eclectica)</i> (Holmberg 1884)	10 (4)	NT
Subgenus <i>M (Eumelissodes)</i> (LaBerge 1956a)	71 (9)	NA, NT
<i>Melissodes, incertae sedis (M. cajannensis)</i> (Lepeletier 1841), <i>M. morosa</i> (Cresson 1878)	2 (0)	NT
Genus <i>Melissoptila</i> (Holmberg 1884)	54 (6)	NT
Genus <i>Mirnapis</i> (Urban 1997)	2 (1)	NT
Genus <i>Svastra</i> (Holmberg 1884)	3 (3)	NT
Genus <i>Svastrides</i> (Michener et al. 1955)	5 (3)	NT
Genus <i>Svastropsis</i> (Moure and Michener 1955)	1 (1)	NT
SUBTRIBE THYGATERINA FREITAS AND SILVEIRA, NEW SUBTRIBE		
Genus <i>Lophothygater</i> (Moure and Michener 1955)	3 (0)	NT
Genus <i>Thygater</i> (Holmberg 1884)		
Subgenus <i>Thygater</i> s. str.	21 (3)	NT
Subgenus <i>T. (Nectarodiaeta)</i> (Holmberg 1903)	9 (2)	NT
Genus <i>Trichocerapis</i> (Cockerell 1904)		
Subgenus <i>Trichocerapis</i> s. str.	4 (1)	NT
Subgenus <i>T. (Dithygater)</i> (Moure and Michener 1955)	1 (0)	NT
Total	761 (150)—19.7% of spp.	

*As the limits of *Xenoglossa* (*Xenoglossodes*) are uncertain, we are considering this subgenus to only include the three species recovered in the clade of its type species. The remaining species of ‘*Tetraloniella*’ from the Western Hemisphere are kept as ‘*Xenoglossa incertae sedis*’ given the polyphyletic nature of the grade comprising these species.

Discussion

Phylogenetics and Higher-level Classification of Eucerini

We gathered the most comprehensive sample of eucerine diversity to date for phylogenetic analysis, including representatives of nearly all previously recognized genera and subgenera. Many of these taxa are herein analyzed in a phylogenetic framework for the first time. This study represents a significant advancement in our understanding of the evolutionary history of a group of bees that historically has been considered morphologically homogeneous and confusing with regard to phylogenetic relationships (Michener 2007).

Our results show *Eucerinoda gayi* and *Canephorula apiformis* as the earliest diverging lineages of Eucerini, with the former placed as the sister group of all remaining Eucerini (Fig. 2). These results provide additional support for the current understanding of the initial divergence of Eucerini based on morphological (Moure

and Michener 1955, Michener and Moure 1957, Roig-Alsina and Michener 1993, Michener 2007), and molecular evidence (Praz and Packer 2014, Dorchin et al. 2018, Bossert et al. 2019, Freitas et al. 2019, 2021).

Alloscirteticina was placed as the sister group of the majority of Eucerini (Fig. 2), and this subtribe comprises three genera: *Alloscirtetica* (~40 spp.), *Dasyscirtetica* stat. n. (4 species), and *Megascirtetica* (1 species). Progress in resolving phylogenetic relationships among remaining eucerine genera in previous analyses was largely obscured by morphological homogeneity and the extreme rarity of some taxa. These obstacles were overcome here by using UCE phylogenomics and next-generation sequencing, which allowed us to obtain large amounts of previously inaccessible molecular data from museum specimens. Using this approach, we have confidently shown that Gaesischiina is the sister group of the large clade encompassing Eucerina, Melissodina, and Thygaterina. The divergence between Gaesischiina and closely related subtribes

Table 2. Basic information for each one of the matrices assembled and analyzed here. All these values refer to the loci after alignment and trimming by GBLOCKS and SPRUCEUP

Dataset	Nbr. loci	Mean loci bp	Min-max loci bp	Total bp	Missing data*
All UCEs	2495	608	126-1,716	1,518,342	14.44%
UCEs_75p	2009	659	154-1,716	1,324,063	18.8%
Genes	1302	907	212-11,741	1,181,407	24.9%
Exons	1411	609	139-2,100	859,710	25.4%
UCEs_98p	738	718	311-1451	529,378	14.4%
UCEs_98p_20k	27	765	605-944	20,665	16.5%

*These values are the amount of missing data in the concatenated alignment after aligning each locus individually.

Table 3. UCE statistics for low-quality samples that were sometimes recovered in spurious positions in the phylogenetic analyses. UCE regions are those identified using SWSC-EN. Statistics for all samples are provided for comparison

Taxon	Coll. year (age)	Loci (cores + both flanks)			Cores			Left flanks			Right flanks		
		Nbr loci	Mean bp	% missing	Nbr loci	Mean bp	Mean length	Nbr loci	Mean bp	Mean length	Nbr loci	Mean bp	Mean length
<i>Svastropsis bipunctata</i>	?	388	179	96.79	336	60,082	191.50	130	10,500	31.00	131	9,998	26.00
<i>Tetralonia viator</i>	1909 (112)	838	180	93.90	732	131,902	197.00	320	26,871	33.00	319	24,553	32.00
<i>Florilegus purpurascens</i>	1973 (48)	915	170	93.27	810	137,769	193.00	360	28,367	28.50	361	27,989	30.00
<i>Thygater melanotricha</i>	1967 (54)	957	170	93.51	799	135,829	184.00	319	27,822	36.00	336	26,241	26.00
<i>Eucera morio</i>	1997 (24)	1,333	186	91.68	1,199	222,684	206.00	629	60,965	50.00	607	54,465	45.00
<i>Dasyhalonia cearensis</i>	2001 (20)	1,396	186	91.48	1,244	231,368	201.00	690	69,073	55.00	682	66,935	57.00
<i>Florilegus lanierii</i>	1964 (57)	1,669	185	88.56	1,519	281,766	208.00	882	183,863	43.00	857	182,417	45.00
<i>Epimelissodes cressoni</i>	1979 (42)	1,983	224	86.44	1,923	431,048	228.00	1,463	77,771	82.00	1,497	75,285	83.00
All samples mean	15 years	2,031	229	25.37*	2,040	470,816	229.10	1,883	331,389	168.97	1,883	330,073	168.02

*This value is the mean of mean missing data for each terminal included.

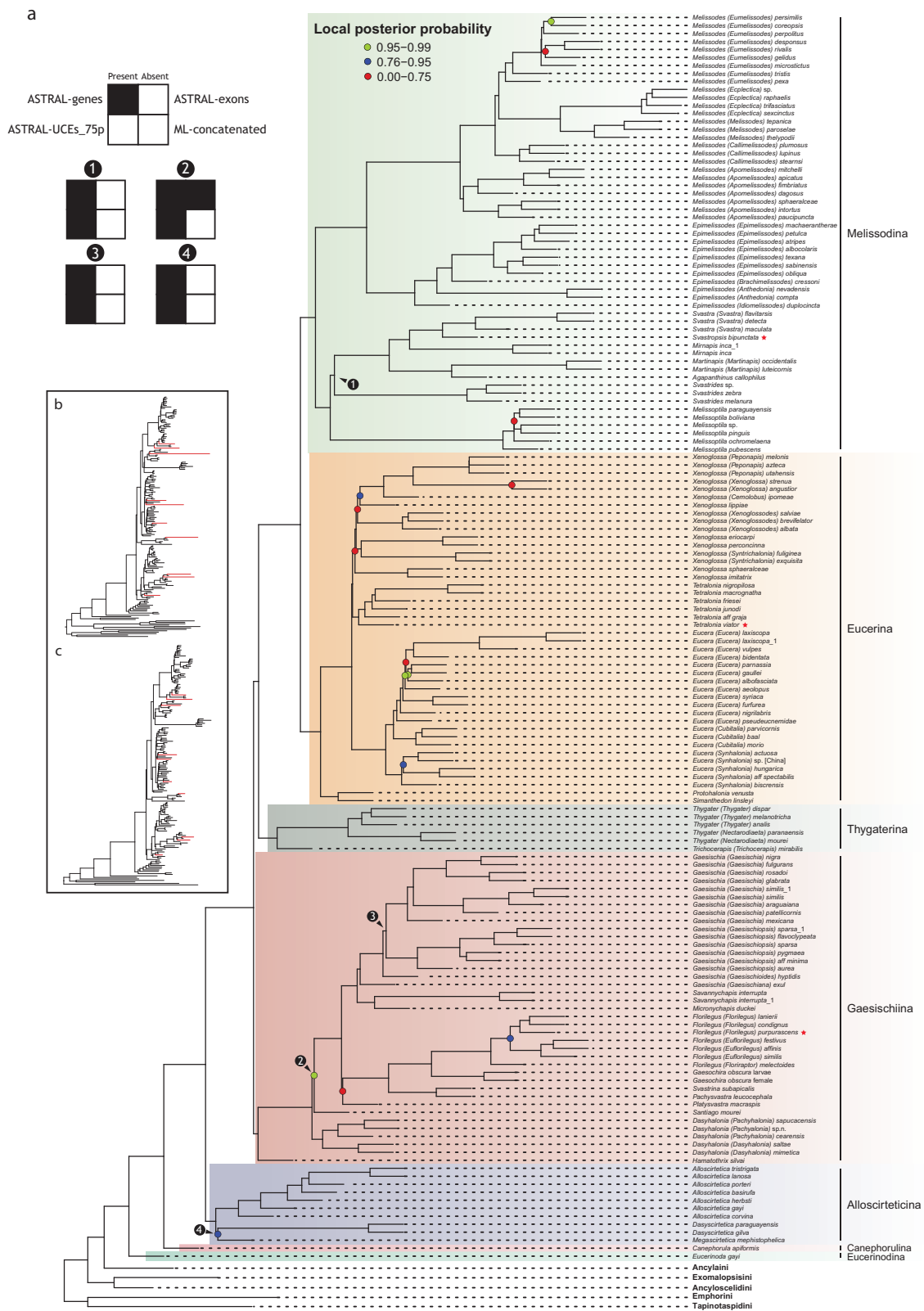
occurred during the mid-Miocene, probably in association with the expansion of the dry diagonal of South America (see next section). The diversification of Thygaterina was also probably associated with the formation of the dry diagonal of South America. Eucerina and Melissodina, placed as sister clades, are predominantly distributed in the Northern Hemisphere (see next section). Melissodina is exclusive to the Western Hemisphere, and Eucerina, in contrast, is more widely distributed. In addition to the Northern Hemisphere (North America and the Palearctic), Eucerina is also found in the Afrotropics and the Oriental region (Michener 2007, Dorchin et al. 2018). Most of the generic diversity of Eucerina is concentrated in the Western Hemisphere, but most of its species-level richness occurs in the Eastern Hemisphere (Dorchin et al. 2018).

Biogeography of Eucerini

Overall, bee species richness is higher in open habitats in mid-latitudes, particularly in xeric environments (e.g., Michener 1979, Orr et al. 2021). Eucerinae originated and initially diversified in such environments in southern South America during the Paleocene (Praz and Packer 2014, Freitas et al. 2022). Our results show that the early diversification of the long-horned bees (Eucerini) occurred in an area including southern South America and the Andean region (Fig. 5). This early history of Eucerini took place before the complete formation of the Andes, which gradually became a major geographic barrier for the biotic connection east and west of this mountain chain. Our results indicate that Eucerini initially radiated in parallel with the uplift of the Andes and other processes that contributed to

aridification in the eastern portion of South America and the “dry diagonal” (Zanella 2011, Dunn et al. 2015, Masa-Iranzo et al. 2021). This pattern of diversification associated with the initial formation and expansion of the set of open vegetation habitats—so-called the Dry Diagonal of South America—is also recovered by other studies of Neotropical bee lineages including Centridini [Apidae] (Martins and Melo 2015); Emphorini [Apidae] (Wilson et al. 2014, Freitas et al. 2022); Tapinotaspini [Apidae] (Aguar et al. 2020, Freitas et al. 2022); and Neopasiphaeinae [Colletidae] (Almeida et al. 2012, 2019). This consistent pattern of several bee groups originating in open vegetation areas during the Paleogene and currently inhabiting these habitats suggests that there is likely niche conservatism. The existence of savanna-like environments in southern South America during the Paleocene has been proposed by previous studies on bee historical biogeography (Aguar et al. 2020, Freitas et al. 2022).

The presence of *Eucerinoda* on the western flank of the Andes (Chile) and *Canephorula* on the eastern side (Argentina) constitutes important evidence of the early diversification of Eucerini before the complete formation of the southern portion of the mountain chain during the early Eocene (~55 Mya, Boschman 2021, Pérez-Escobar et al. 2022). The first species-rich lineage to originate (Alloscirteticina, ~10 Mya) retained the original distribution of Eucerini, i.e., southern South America + Andean region. The diversification of Alloscirteticina was likely strongly affected by the final uplift of the south-central/central Andes *sensu* Boschman (2021), which represents most of the Chilean territory northern to Patagonia and southern Peru (between latitudes of 34–14°S). Despite the



Downloaded from https://academic.oup.com/isd/article/7/4/3/7222693 by ESA Member Access user on 13 July 2023

Fig. 2. Phylogenetic relationships among 151 species of Eucerini: a) tree resulting from ASTRAL analyses, using the matrix “Genes”, with branch lengths shown as coalescent units. Node supports of 1 were omitted. Terminal taxa marked with a red star were manually moved to the phylogenetic placement recovered in the concatenated results. This was done because missing data created a bias in gene-tree estimation and led to poor summary tree results for these taxa (see details in the text). The Navajo rugs next to some nodes represent variation in relationships among main taxa (genera or subgenera) in different analyses (Supplementary Figs. S1–S7); different results within subgenera were not considered. b) Preliminary tree topology generated through a ML analysis of all (2,495) UCE loci concatenated and partitioned by UCE locus. Long branches highlighted in red are those of specimens with extremely degraded DNA. These terminals were the ones with specific cutoffs in SPRUCEUP analyses. c) Tree generated through a ML analysis of all UCE loci concatenated and partitioned by locus after trimming the alignment with SPRUCEUP. The branches highlighted in red are the same as those highlighted in b).

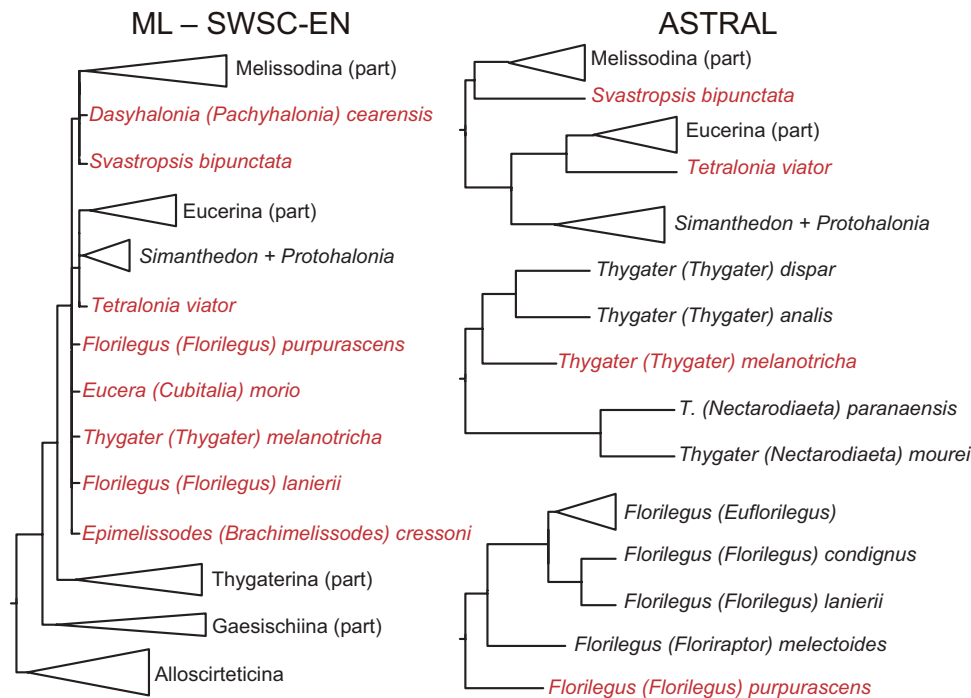


Fig. 3. Phylogenetic placement of terminal species of Eucerini with highly degraded DNA in ML analysis using partitions defined by SWSC-EN and ASTRAL analysis.

estimated origin of Alloscirteticina on the eastern flank of the Andes, species of this subtribe are distributed on both sides of the mountain chain. This distribution suggests that the final uplift of this portion of Andes either promoted vicariant events in *Alloscirteticina* or that dispersals may have occurred westward through the Andes.

The initial diversification of long-horned bees began between the Oligocene and the mid-Miocene (25–15 Mya) when Eucerinae diversified more rapidly according to lineage through time plots (Freitas et al. 2022). During this same timeframe, connections between areas of open vegetation in South America and North America formed, likely contributing to range expansions in different Eucerinae lineages to North America (Wilson et al. 2014, Freitas et al. 2022). The same signal was recovered here, as evidenced by the Eucerina + Melissodina clade. A few basal nodes of Melissodina suggest the occupation of an area that includes southeastern South America and western North America (Fig. 6). This supports the hypothesis that connections between open vegetation areas in both continents existed during the Miocene. These connections allowed groups of Eucerini (as well as other tribes of Eucerinae) to reach North America before the complete closing of the Isthmus of Panama, likely due to suitable environmental conditions (Freitas et al. 2022). Additional biogeographic studies of bees have shown comparable events of range expansion to North and Central America during the same time frame in Calliopsini and Protandrenini (Bossert et al. 2021, Ramos et al. 2022), and Centridini (Martins and Melo 2016). Others have shown range expansion in the same time frame but in the opposite direction for groups originating in northern latitudes (e.g., Bombini; Hines 2008; Santos Júnior et al. 2022). The congruent results from these studies using different organisms reinforce the understanding that connections existed between the areas where these bees occur in the present in both South and Central/North America. Most species of long-horned bees and other groups mentioned as having possibly crossed to or from North and South America are ground nesters, and considering the challenges to nidify in the ground of humid forested

habitats (the food mass can liquefy, drowning the eggs and larvae; the same food mass can be attacked by fungi—Michener 2007), we infer that corridors of open dry habitats may have existed allowing these bees to cross the area occupied by forests in the present between the areas of open vegetation.

After this initial radiation, which was probably affected by a new short period of planet-warming (mid-Miocene Climatic Optimum) and the re-establishment of forested environments, this connection seems to have disappeared. However, aridification and the spread of habitats dominated by open vegetation apparently continued in both South America (Dunn et al. 2015, Azevedo et al. 2020) and North America (Strömberg 2005) over this period, certainly associated with the constant planet cooling (Zachos et al. 2008). This probably contributed to the diversification of long-horned bees in both continents, particularly Gaesischiina and Thygaterina in South America, and Melissodina and Eucerina in North America (Fig. 5).

Our results indicate that the ancestor of Eucerina minus *Protohalonia* + *Simanthedon* occupied North America and the Palearctic concurrently during the late Miocene (~10 Mya). Dispersal to the Palearctic probably occurred through the Bering land bridge. Periods of high mean annual temperatures occurred in Beringia during the late Miocene (~13–8 Mya), with summer temperatures reaching 14°C (Wolfe 1994). This is a well-accepted bridge connecting the Holarctic region, with a plethora of different organisms reconstructed as having used Bering as a bridge to cross both eastward and westward (Sanmartín et al. 2001, Jiang et al. 2019). Several studies have indicated that this route may have been used by different bee groups (e.g., Hines 2008, Trunz et al. 2016, Branstetter et al. 2021) including Eucerina (Praz and Packer 2014, Dorchin et al. 2018).

The connection between North America and the Palearctic was broken in the late Miocene (~8 Mya) due to a drastic drop in mean annual temperature in Beringia (Wolfe 1994). This allowed for the isolation and diversification of the ancestors of three clades of

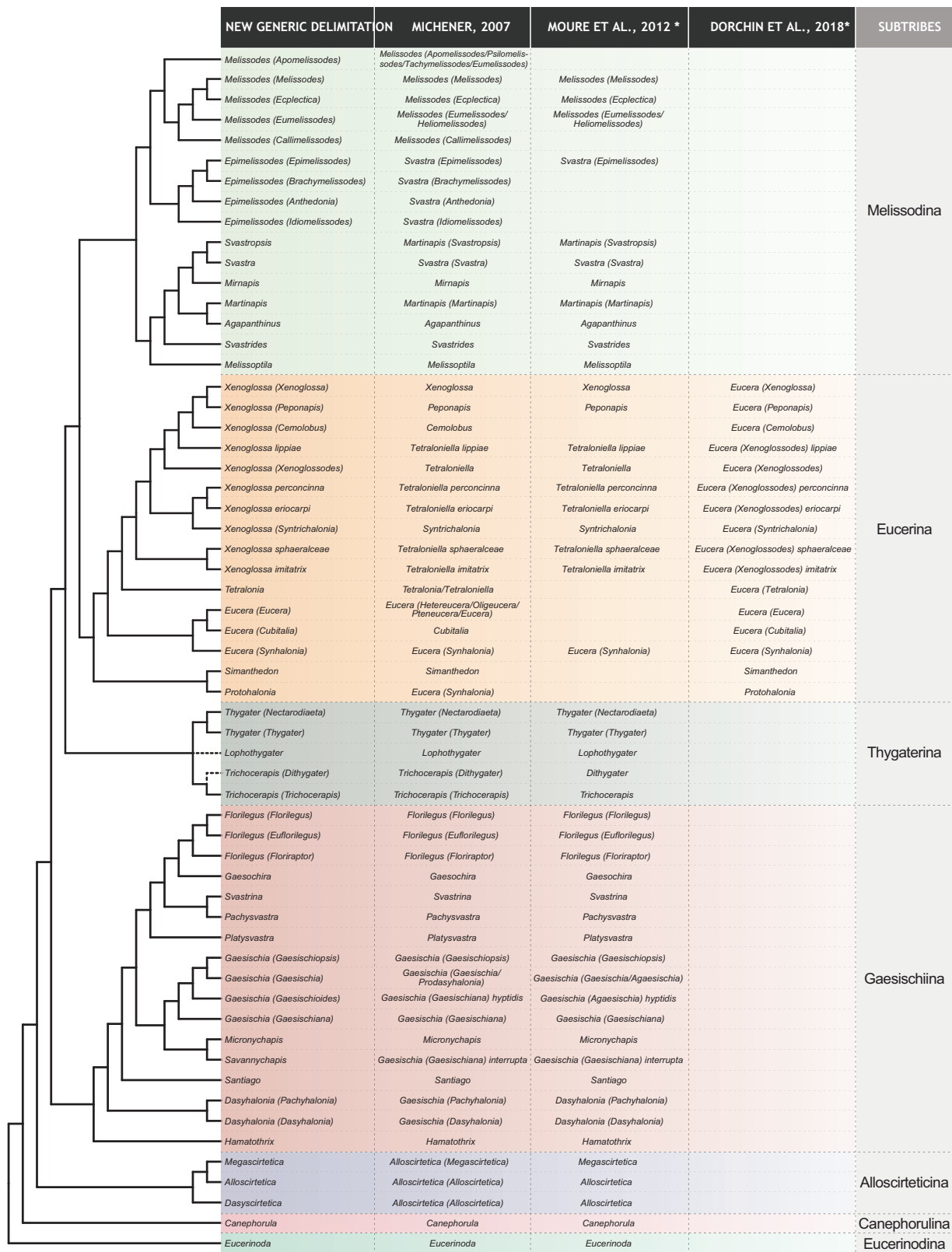


Fig. 4. Summary of the new classificatory proposal for Eucerini compared to previous classifications (* Moure et al. [2012] only apply to Neotropical taxa; Dorchin et al. [2018] only apply to Eucerina). Dotted lines refer to taxa not included here but included in the tree based on previous phylogenetic results or knowledge about their morphology (Moure and Michener 1955, Urban 1967a, Michener 2007, Freitas et al. 2019). The revised subtribal classification of Eucerini proposed in this study is shown on the right column.

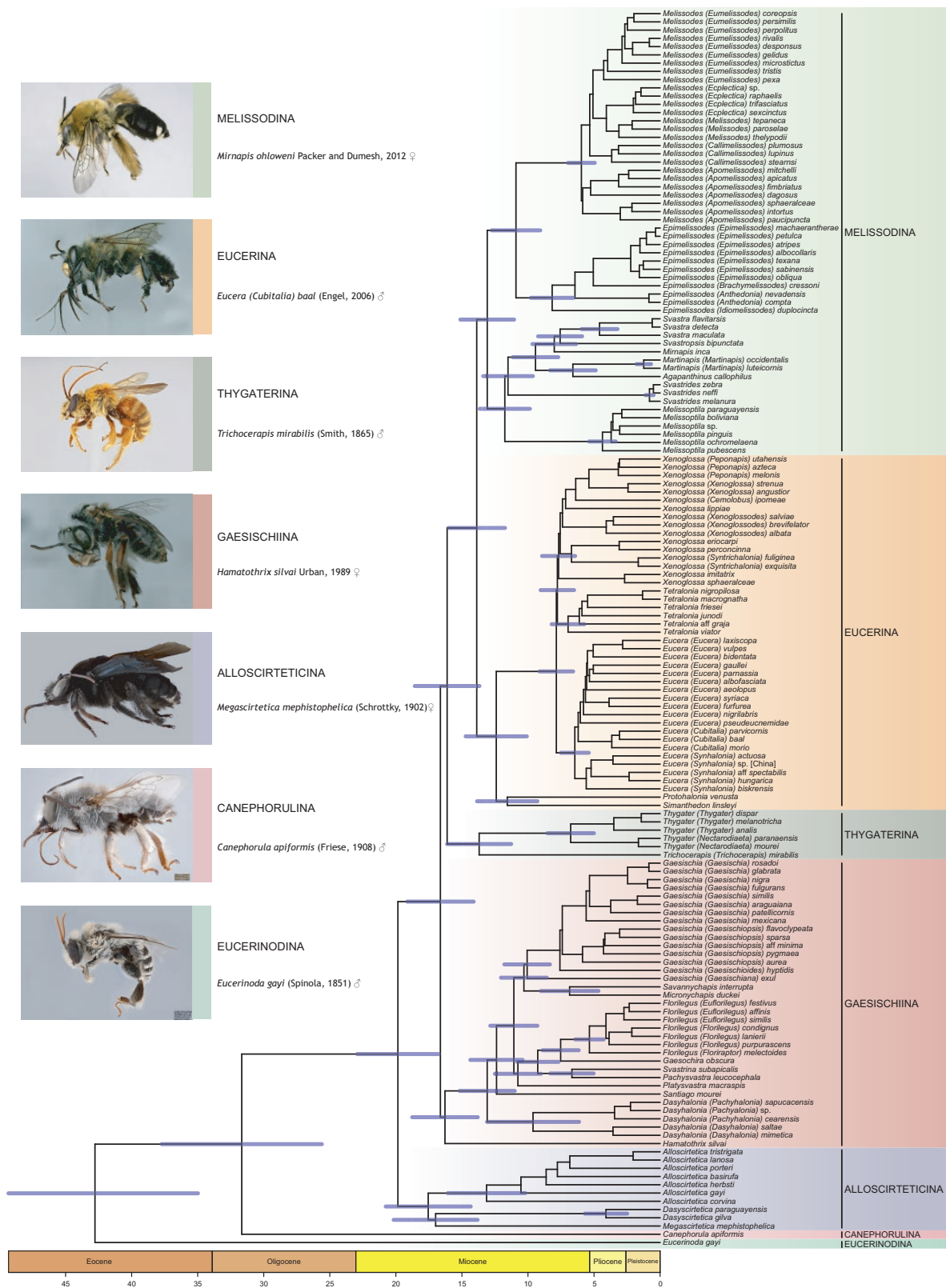


Fig. 5. Divergence time estimation of Eucerini. The chronogram is the result of MCMCtree analysis conducted using a data-driven Birth and Death (ddBD) prior. Pictures on the left represent diversity in each of the subtribes and are organized in the same order as the lineages appear in the tree. All pictures are from Packer lab (<https://www.yorku.ca/bugsrus/>), except *Eucera (Cubitalia) baal* and *Hamatothrix silvai*, which are from authors.

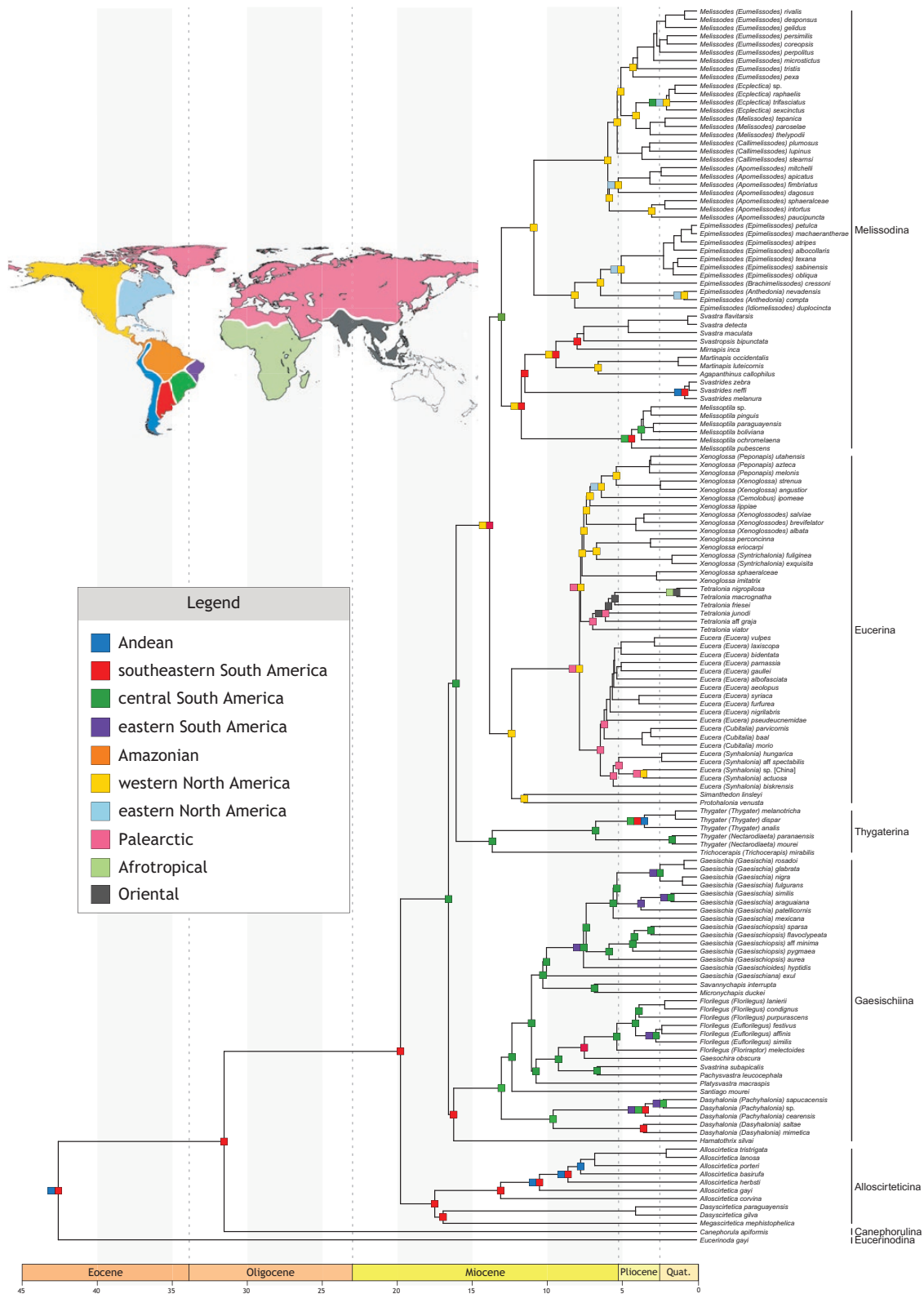


Fig. 6. Summary of two biogeographical reconstructions conducted in RevBayes under the DEC model using 10 areas that represent the global distribution of Eucerini. The biogeographic analyses were conducted on the MCMCtree chronogram. The first reconstruction was conducted using a detailed subdivision of South America into five areas and all other regions coded as a single area (Supplementary Fig. S8), the second analysis treated South America as a single area and included a more refined subdivision of the Nearctic and the Eastern Hemisphere (Supplementary Fig. S9); refer to the raw results (Supplementary Figs. S8 and S9) for the distributions of each terminal and detailed analytical output of the DEC reconstructions.

Eucerina: (1) *Xenoglossa* in North America, and (2) *Tetralonia* + (3) *Eucera* in the Palearctic (Fig. 5). This continuous cooling of the planet during the Miocene is also associated with the probable extinction of other groups of bees adapted to subtropical or paratropical climates in higher latitudes, like the stingless bees (Meliponini). This is a group with a Pan-tropical distribution in the present, but with fossils in Baltic amber, which shows that the group distribution extended much norther during periods when the planet was warmer (Engel and Michener 2013, Melo 2020).

Tetralonia subsequently reached the Oriental region during the late Miocene (Fig. 5). This happened when the planet was likely warmer and wetter, which probably led to deserts (predominant in areas between the Palearctic and Oriental regions) being largely replaced by shrublands and savannas (Pound et al. 2011). *Tetralonia* reached the Afrotropics in the Pliocene (Fig. 5), probably taking advantage of the cycles of humid and arid periods that made the Sahara Desert contract and expand since the Miocene-Pliocene transition (Swezey 2009). This same pattern of range expansion to the Afrotropics was found also in some clades of andrenid bees (Bossert et al. 2021, Ramos et al. 2022).

The last major clade of Eucerini to expand its range and later become isolated is *Eucera* (*Synhalonia*). This subgenus is the most species-rich in the entire tribe (Ascher and Pickering 2021) and is composed of Palearctic and Nearctic clades (Fig. 5; Dorchin et al. 2018, Ascher and Pickering 2021). Despite our restricted taxon sampling of the Nearctic clade of Eucerina, previous phylogenetic evidence using a denser taxon sampling supports its monophyly (Dorchin et al. 2018). The divergence between the Nearctic and Palearctic clades of Eucerina in the Pliocene is also consistent with our findings. This movement back to the Nearctic region probably occurred through Beringia—the same route used by their ancestors to reach the Palearctic—during the early Pliocene, the last period when mean temperatures were considerably warmer than today (Ravelo et al. 2004, Stroynowski et al. 2015). The drop in mean temperatures throughout the Pliocene and especially during the Pleistocene (<3 Mya) (Raymo et al. 1996, Ravelo et al. 2004, Stroynowski et al. 2015) would have isolated the populations on either side of Beringia.

UCEs and Missing Data

We found that most of the UCE data sequenced from highly degraded specimens mapped to the core regions of UCEs (as shown in Table 3). This result is expected because the core region is the area specifically targeted by the enrichment baits (Faircloth et al. 2012). Thus, due to the high levels of DNA fragmentation in low-quality samples, there are not many fragments that include both core and flanking regions, impeding sequencing and assembly of loci beyond the core. This observation means that in UCE studies, specimens with highly degraded DNA may be more challenging to place confidently in a phylogeny due to a lack of variable sites. This is particularly true in the context of species-tree summary methods, which rely on individual gene-trees as input. Including these old specimens in phylogenomic datasets must therefore be considered with caution. We recommend that the use of these suboptimal samples should be limited and considered only in the absence of more recent samples, especially when the samples are key to answering pressing and important phylogenetic questions.

UCEs are effective markers to accurately place taxa with high levels of missing data in a phylogenetic context (Blaimer et al. 2015, 2016, McCormack et al. 2016, Derkarabetian et al. 2019). This is usually true for both concatenation and species-tree summary

methods. However, in extreme cases, in which samples have ~90% missing data or more, samples can be misplaced in some analyses. In these cases, it is important to examine results from both species-tree summary methods and concatenation, as we did here, in order to identify problematic samples and correct erroneous results.

Species-tree summary methods sometimes perform better in phylogenetic reconstruction than concatenation because of their capacity to deal with incomplete lineage sorting (Kubatko and Degnan 2007, Jiang et al. 2020, but see Springer and Gatesy 2016). However, sometimes these summary approaches are ineffective at correctly placing taxa with high levels of missing data (Moyle et al. 2016, Supplementary Figs. S1–S3), likely due to a lack of phylogenetic signal and poor gene-tree resolution. In these cases, concatenation can outperform summary methods because the complete alignment is used (Moyle et al. 2016, Derkarabetian et al. 2019), maximizing the amount of phylogenetic signal available to place low-quality taxa.

Combining UCEs into genes using CURE produces better gene-trees (van Dam et al. 2021) by increasing the number of variable sites for the highly degraded samples, which usually only have the core region of their UCEs available. In addition to the well-known benefits of having longer alignments (e.g., Adams and Castoe 2019, van Dam et al. 2021), species-tree summary methods may also benefit from merging UCEs found in the same gene by not over-representing single genes, as discussed previously by van Dam et al. (2021).

Museomics and Limits of Phylogenomic Research Based on Highly Degraded DNA Samples

Despite the high interest shown for the use of museum specimens for UCE phylogenomics and how it can be beneficial for phylogenetic studies (e.g., Moyle et al. 2016, Blaimer et al. 2018, Derkarabetian et al. 2019), discussions about the limitations and ways of addressing these limitations are still scarce in the literature (Smith et al. 2020, Salter et al. 2022). Samples with extremely degraded DNA are often recovered in seemingly erroneous or odd positions in species-tree summary methods. They tend to be recovered as diverging from the base of the clades where they belong. Furthermore, as the length and number of loci decrease, the deeper into the tree the samples get placed (e.g., Moyle et al. 2016). This observation may be one of the reasons why some museomics studies do not present species-tree analyses. Doing so would have revealed contrasting results regarding the position of these samples as compared to concatenation (e.g., Blaimer et al. 2018, Derkarabetian et al. 2019).

In our study, most terminals with highly degraded DNA (those listed in Table 3) continued to be recovered in spurious positions by species-tree summary methods, even after CURE treatment. Concatenation analyses produced better results, because even with the flanks in each UCE locus being short, they provided a considerable amount of variation when combined together, allowing for better placement of the taxa. For example, the most degraded sample in our dataset (*Svastropsis bipunctata* comb. n.) had 336 loci, and of these, at least 130 of them had more variable flanking sites present. Together, these loci represent more than 20k bp (Table 3), which is a considerable amount of molecular data, at least when compared to traditional Sanger sequencing studies. It is also worth noting that although core regions tend to be highly conserved, they often do have some phylogenetically informative sites. This efficiency of concatenation methods to deal with considerable amounts of missing data has been widely discussed and is confirmed here (e.g., Wiens 2003, Driskell et al. 2004, Philippe et al. 2004). Overall, our results

suggest that combining species-tree summary methods and concatenation methods can be an effective approach to benefit from both kinds of analysis, especially when datasets include specimens with extreme levels of missing data.

New approaches to obtaining data from UCEs, like sequencing low-coverage genomes to isolate the ultraconserved regions, have shown promising results. This may be an excellent alternative to maximize the recovery of data from flanking regions of UCEs, allowing a better assessment of the position of these taxa in species-tree summary methods or even isolating other kinds of loci for phylogenetic inference (Zhang et al. 2019, Ribeiro et al. 2021, Orr et al. 2022). It could also be the case that designing UCE probe sets that enrich more of the flanking regions of loci may improve data quality in older samples. To our knowledge, this approach has not yet been tested.

Supplementary Data

Supplementary data are available at *Insect Systematics and Diversity* online.

Acknowledgments

This article is dedicated to Professor Fernando A. Silveira, mentor of the junior and senior authors. Fernando had an active role in the development of this research, especially in the proposal of a revised subtribal classification of Eucerini, but sadly passed away months before the submission of the manuscript. We are grateful to Antonio J. C. Aguiar (Universidade de Brasília, Brazil), Christophe Praz (Université de Neuchâtel, Switzerland), Jerome G. Rozen Jr. (American Museum of Natural History, USA), Carlos Roberto F. Brandão (Museu de Zoologia da Universidade de São Paulo, Brazil), Laurence Packer (York University, Canada), Maximilian Schwarz (Ansfelden, Austria), and Seán G. Brady (US National Museum of Natural History—Smithsonian Institution, USA) for generously providing specimens or tissue samples, essential for our broad taxonomic sampling. We are thankful to Gabriel A.R. Melo for kindly discussing some key aspects of the taxonomy of Neotropical Eucerini, and to three anonymous reviewers and the associate editor of *Insect Systematics and Diversity*, Jessica Gillung, for their valuable suggestions that helped improve the manuscript. We are grateful to Adriana Tiba, Julio Pupin, Laurence Packer, Sergio Jansen-Gonzalez, Javier Gross, and Shan Gui for allowing the use of their beautiful photographs of eucerine bee. This study was supported by USDA-ARS, and NSF (Project No. 2080-21000-019-00-D, and grant DEB-2127744, respectively) to M.G.B.; by grants #2018/09666-5, #2019/09215-6, and #2021/02196-6, São Paulo Research Foundation (FAPESP) to E.A.B.A. and F.V.F.; by the Brazilian National Council for Scientific and Technological Development (CNPq grants #310111/2019-6 and #422019/2018-6) to E.A.B.A. and F.V.F.; and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)—Brazil (Finance Code 001). USDA is an equal opportunity provider and employer.

Author Contributions

Felipe Freitas (Conceptualization-Lead, Data curation-Lead, Formal analysis-Lead, Funding acquisition-Lead, Investigation-Lead, Methodology-Lead, Project administration-Lead, Software-Equal, Writing – original draft-Lead, Writing – review & editing-Lead), Michael Branstetter (Data curation-Equal, Formal analysis-Equal, Funding acquisition-Lead, Investigation-Equal, Methodology-Equal, Resources-Equal, Supervision-Equal, Validation-Lead, Writing – review & editing-Equal), Vinicius Franceschini-Santos (Formal analysis-Equal, Investigation-Equal, Methodology-Equal, Software-Equal, Writing – review & editing-Equal), Achik Dorchin (Investigation-Equal, Resources-Equal, Validation-Equal, Writing – review & editing-Equal), Karen Wright (Investigation-Equal, Resources-Equal, Validation-Equal, Writing – review & editing-Equal), Margarita Lopez-Urbe (Investigation-Equal, Validation-Equal, Writing – review & editing-Equal), Terry Griswold (Investigation-Equal, Resources-Equal, Validation-Equal, Writing – review & editing-Equal), Fernando Silveira (Conceptualization-Equal, Investigation-Equal,

Supervision-Equal), Eduardo Almeida (Conceptualization-Lead, Funding acquisition-Lead, Investigation-Equal, Methodology-Equal, Resources-Lead, Supervision-Lead, Writing – original draft-Equal, Writing – review & editing-Equal)

Data Availability

All contigs, data sets, and complementary files and results are available at Zenodo ([10.5281/zenodo.7820143](https://doi.org/10.5281/zenodo.7820143)). All raw Illumina reads for Eucerinae have been deposited at the NCBI Sequence Read Archive under BioProject PRJNA632049. The samples originally sequenced for this paper are those under the BioSample SAMN30656088—SAMN30656168.

References

- Adams RH, Castoe TA. Statistical binning leads to profound model violation due to gene tree error incurred by trying to avoid gene tree error. *Mol Phylogenet Evol.* 2019;134:164–171. <https://doi.org/10.1016/j.ympev.2019.02.012>
- Aguiar AJC, Melo GAR, Vasconcelos TNC, Gonçalves RB, Giugliano L, Martins AC. Biogeography and early diversification of Tapinotaspini oil-bees support presence of Palaeocene savannas in South America. *Mol Phylogenet Evol.* 2020;143:106692. <https://doi.org/10.1016/j.ympev.2019.106692>
- Almeida EAB, Packer L, Melo GAR, Danforth BN, Cardinal SC, Quintero FB, Pie MR. The diversification of neopasiphaeine bees during the Cenozoic (Hymenoptera: Colletidae). *Zool Scr.* 2019;48:226–242.
- Almeida EAB, Pie MR, Brady SG, Danforth BN. Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. *J Biogeogr.* 2012;39:526–544.
- Ascher JS, Pickering J. Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila); 2021.
- Ashmead WH. Classification of the bees or the superfamily Apoidea. *Trans Am Entomol Soc.* 1899;26:49–100.
- Azevedo JAR, Collevatti RG, Jaramillo CA, Strömberg CAE, Guedes TB, Matos-Maraví P, Bacon CD, Carillo JD, Faurby S, Antonelli A. On the young savannas in the land of ancient forests. In: Rull V, Carnaval AC, editors. Neotropical diversification: patterns and processes. Cham: Springer International Publishing; 2020. p. 271–298.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 2012;19:455–477.
- Blaimer BB, Brady SG, Schultz TR, Lloyd MW, Fisher BL, Ward PS. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evol Biol.* 2015;15(1):1–14.
- Blaimer BB, Lloyd MW, Guillory WX, Brady SG. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS One.* 2016;11(8):e0161531. <https://doi.org/10.1371/journal.pone.0161531>
- Blaimer BB, Mawdsley JR, Brady SG. Multiple origins of sexual dichromatism and aposematism within large carpenter bees. *Evolution.* 2018;72:1874–1889.
- Blaimer BB, Santos BF, Cruaud A, Gates MW, Kula RR, Mikó I, Rasplus J-Y, Smith DR, Talamas EJ, Brady SG, et al. Key innovations and the diversification of Hymenoptera. *Nat Commun.* 2023;14(1):1212. <https://doi.org/10.1038/s41467-023-36868-4>
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–2120.
- Borowiec ML. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ.* 2016;2016:e1660.
- Borowiec ML. Spruceup: fast and flexible identification, visualization, and removal of outliers from large multiple sequence alignments. *J Open Source Softw.* 2019;4:1635.
- Borowiec ML, Lee EK, Chiu JC, et al. Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. *BMC Genomics.* 2015;16:987. <https://doi.org/10.1186/s12864-015-2146-4>

- Boschman LM. Andean mountain building since the late Cretaceous: a paleoelevation reconstruction. *Earth-Sci Rev.* 2021;220:103640. <https://doi.org/10.1016/j.earscirev.2021.103640>
- Bossert S, Murray EA, Almeida EAB, Brady SG, Blaimer BB, Danforth BN. Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. *Mol Phylogenet Evol.* 2019;130:121–131. <https://doi.org/10.1016/j.ympev.2018.10.012>
- Bossert S, Wood TJ, Patiny S, Michez D, Almeida EAB, Minckley RL, Packer L, Neff JL, Copeland RS, Straka J, et al. Phylogeny, biogeography and diversification of the mining bee family Andrenidae. *Syst Entomol.* 2021;47:283–302.
- Bossert S, Wood TJ, Patiny S, Michez D, Almeida EAB, Minckley RL, Packer L, Neff JL, Copeland RS, Straka J, Pauly A, Griswold T, Brady SG, Danforth BN, Murray EA. Phylogeny, biogeography and diversification of the mining bee family Andrenidae. *Syst. Entomol.* 2021;47:283–302.
- Branstetter MG, Danforth BN, Pitts JP, Faircloth BC, Ward PS, Buffington ML, Gates MW, Kula RR, Brady SG. Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. *Curr Biol.* 2017;27(7):1019–1025. <https://doi.org/10.1016/j.cub.2017.03.027>
- Branstetter MG, Longino JT. Ultra-conserved element phylogenomics of new world *Ponera* (Hymenoptera: Formicidae) illuminates the origin and phylogeographic history of the endemic exotic ant *Ponera* *exotica*. *Insect Syst Diversity.* 2019;3(2):1–13.
- Branstetter MG, Longino JT, Ward PS, Faircloth BC. Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. *Methods Ecol Evol.* 2017;8(6):768–776. <https://doi.org/10.1111/2041-210x.12742>
- Branstetter MG, Müller A, Griswold TL, Orr MC, Zhu C. Ultraconserved element phylogenomics and biogeography of the agriculturally important mason bee subgenus *Osmia* (*Osmia*). *Syst Entomol.* 2021;46(2):453–472. <https://doi.org/10.1111/syen.12470>
- Brown RP, Yang Z. Rate variation and estimation of divergence times using strict and relaxed clocks. *BMC Evol Biol.* 2011;11(1):2711–2712. <https://doi.org/10.1186/1471-2148-11-271>
- Cameron SA. A new tribal phylogeny of the Apidae inferred from mitochondrial DNA sequences. *Diversity in the Genus Apis*. Boulder, CO: Westview Press; 1992. p. 71–87.
- Cameron SA. Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. *Proc Natl Acad Sci U S A.* 1993;90(18):8687–8691. <https://doi.org/10.1073/pnas.90.18.8687>
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 2000;17:540–552.
- Cockerell TDA. New genera of bees. *Entomol News.* 1904;15:292.
- Cockerell TDA. New name for a genus of bees. *Entomologist.* 1929;62:19.
- Cresson ET. Descriptions of new species of North American bees. *Proc Acad Nat Sci Phil.* 1878;30:181–221.
- Cure JR, Thiengo M, Silveira FA, Rocha LB. Levantamento da fauna de abelhas silvestres na ‘zona da mata’ de Minas Gerais: III. Mata secundária na região de Viçosa (Hymenoptera, Apoidea). *Rev Bras Zool.* 1992;9(3–4):223–239. <https://doi.org/10.1590/s0101-81751992000200008>
- Dafni A, O’Toole C. Pollination syndromes in the Mediterranean: generalizations and peculiarities. In: Arianoutsou M, Groves RH, editors. *Plant-animal interactions in mediterranean-type ecosystems*. Dordrecht: Springer; 1994. p. 125–135.
- van Dam MH, Henderson JB, Esposito L, Trautwein M. Genomic characterization and curation of UCEs improves species tree reconstruction. *Syst Biol.* 2021;70(2):307–321. <https://doi.org/10.1093/sysbio/syaa063>
- Danforth BN, Cardinal S, Praz C, Almeida EAB, Michez D. The impact of molecular data on our understanding of bee phylogeny and evolution. *Annu Rev Entomol.* 2013;58:57–78.
- Danforth BN, Minckley RL, Neff JL. *The solitary bees: biology, evolution, conservation*. Princeton: Princeton University Press; 2019.
- Derkarabetian S, Benavides LR, Giribet G. Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. *Mol Ecol Resour.* 2019;19(6):1531–1544. <https://doi.org/10.1111/1755-0998.13072>
- Dorichin A, Danforth BN, Griswold T. A new genus of eucerine bees endemic to southwestern North America revealed in phylogenetic analyses of the *Eucera* complex (Hymenoptera: Apidae: Eucerini). *Arthropod Syst Phylo.* 2018;76:215–234.
- Dorichin A, López-Urbe MM, Praz CJ, Griswold T, Danforth BN. Phylogeny, new generic-level classification, and historical biogeography of the *Eucera* complex (Hymenoptera: Apidae). *Mol Phylogenet Evol.* 2018;119:81–92. <https://doi.org/10.1016/j.ympev.2017.10.007>
- Driskell AC, Ané C, Burleigh JG, McMahon MM, O’Meara BC, Sanderson MJ. Prospects for building the tree of life from large sequence databases. *Science.* 2004;306:1172–1174.
- Dunn RE, Strömberg CAE, Madden RH, Kohn MJ, Carlini AA. Linked canopy, climate, and faunal change in the Cenozoic of Patagonia. *Science.* 2015;347:258–261.
- Engel MS, Michener CD. Geological history of the stingless bees (Apidae: Meliponini). In: Vit P, Roubik DW, editors. *Stingless bees process honey and pollen in the cerumen pots*. Mérida: Facultad de Farmacia y Bioanálisis, Universidad de Los Andes; 2013.
- Faircloth BC. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics.* 2016;32:786–788.
- Faircloth BC. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. *GitHub*; 2013.
- Faircloth BC, Branstetter MG, White ND, Brady SG. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Mol Ecol Resour.* 2015;15(3):489–501. <https://doi.org/10.1111/1755-0998.12328>
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst Biol.* 2012;61(5):717–726. <https://doi.org/10.1093/sysbio/sys004>
- Freitas FV. New genus and subgenus of South American long-horned bees (Apidae, Eucerini). *Zootaxa.* 2022;5196(4):595–600. <https://doi.org/10.11646/zootaxa.5196.4.8>
- Freitas FV, Branstetter MG, Casali DM, Aguiar AJC, Griswold T, Almeida EAB. Phylogenomic dating and Bayesian biogeography illuminate an antitropical pattern for eucerine bees. *J Biogeogr.* 2022;49(6):1034–1047. <https://doi.org/10.1111/jbi.14359>
- Freitas FV, Branstetter MG, Griswold T, Almeida EAB. Partitioned gene-tree analyses and gene-based topology testing help resolve incongruence in a Phylogenomic study of host-specialist bees (Apidae: Eucerinae). *Mol Biol Evol.* 2021;38(3):1090–1100. <https://doi.org/10.1093/molbev/msaa277>
- Freitas FV, Santos Júnior JE, Santos FR, Silveira FA. Species delimitation and sex associations in the bee genus *Thygater*, with the aid of molecular data, and the description of a new species. *Apidologie.* 2018;49(4):4841–4496. <https://doi.org/10.1007/s13592-018-0576-0>
- Freitas FV, Santos Júnior JE, Santos FR, Silveira FA. A phylogenetic study of the *Thygater-Trichocera* group and new scopes for the subgenera of *Thygater* Holmberg (Hymenoptera, Apidae). *Syst Entomol.* 2019;44:728–744.
- Freitas FV, Silveira FA. Synopsis of the bee genus *Thygater* Holmberg 1884 (Hymenoptera, Apidae) in the Brazilian state of Minas Gerais, with the description of a new species and a key to all Brazilian species. *Zootaxa.* 2017;4238(1):1–29. <https://doi.org/10.11646/zootaxa.4238.1.1>
- Friese H. Die Apidae (Blumenwespen) von Argentina nach den Reisenergebnissen der Herren A. C. Jensen-Haarup und P. Jörgensen in den Jahren 1904–1907. *Flora Fauna.* 1908;10:1–94
- Friese H. Neue Bienen-Arten der palaearktischen Region (Hym.). *Archiv für Naturgeschichte.* 1911;77:135–143.
- Grab H, Branstetter MG, Amon N, Urban-Mead KR, Park MG, Gibbs J, Blitzer EJ, Poveda K, Loeb G, Danforth BN. Agriculturally dominated landscapes reduce bee phylogenetic diversity and pollination services. *Science.* 2019;363:282–284.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010;59(3):307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hines HM. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). *Syst Biol.* 2008;57(1):58–75. <https://doi.org/10.1080/10635150801898912>

- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 2018;35(2):518–522. <https://doi.org/10.1093/molbev/msx281>
- Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst Biol.* 2016;65(4):726–736. <https://doi.org/10.1093/sysbio/syw021>
- Holmberg EL. Viajes al Tandil y a la Tinta, 2nd Parte, Zoología, Insectos. I. Himenópteros-Hymenoptera. *Act Acad Nac Cienc Córdoba.* 1884;5:117–136.
- Holmberg EL. Delectus Hymenopterologicus Argentinus. Hymenopterorum Argentinorum et quorundam exoticorum observationes synonymicas, addendas, novorumque generum specierumque descriptiones continens. *Anales Mus Nac Hist Nat Buenos Aires.* 1903;2:377–468.
- Holmberg EL. *Alloscirtetica*, Holmberg., n. nom. *Apuntes Hist Nat.* 1909;1:77.
- Holt BG, Lessard J-P, Borregaard MK, Fritz SA, Araújo MB, Dimitrov D, Fabre P-H, Graham CH, Graves GR, Jönsson KA, et al. An update of Wallace's zoogeographic regions of the World. *Science.* 2013;339:74–78.
- Hurd PD, Linsley EG, Whitaker TW. Squash and gourd bees (*Peponapis*, *Xenoglossa*) and the origin of the cultivated cucurbita. *Evolution.* 1971;25:218–234.
- Jiang D, Klaus S, Zhang YP, Hillis DM, Li JT. Asymmetric biotic interchange across the Bering land bridge between Eurasia and North America. *Natl Sci Rev.* 2019;6(4):739–745. <https://doi.org/10.1093/nsr/nwz035>
- Jiang X, Edwards SV, Liu L, Faircloth B. The multispecies coalescent model outperforms concatenation across diverse Phylogenomic data sets. *Syst Biol.* 2020;69:795–812.
- Jørgensen P. Beobachtungen über Blumenbesuch, Biologie, Verbreitung usw. der Bienen von Mendoza. (Hym.). *Dtsch Entomol Z.* 1909;1909(2):21153–21227. <https://doi.org/10.1002/mmnd.48019090206>
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14(6):587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 2013;30:772–780.
- Kendall DG. On the generalized 'birth-and-death' process. *Ann Math Stat.* 1948;19(1):1–15. <https://doi.org/10.1214/aoms/1177730285>
- Kubatko LS, Degnan JH. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol.* 2007;56(1):17–24. <https://doi.org/10.1080/10635150601146041>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35(6):1547–1549. <https://doi.org/10.1093/molbev/msy096>
- LaBerge WE. Notes on the genus *Gaesischia* Michener, LaBerge and Moure, with descriptions of a new species and subgenus from Mexico. *Pan-Pacific Entomol.* 1958;34:195–201.
- LaBerge WE. A revision of the bees of the genus *Melissodes* in North and Central America. Part I. *Univ Kansas Sci Bull.* 1956a;38:911–1194.
- LaBerge WE. A revision of the bees of the genus *Melissodes* in North and Central America. Part II. *Univ Kansas Sci Bull.* 38:533–578. 1956.
- LaBerge WE. The genera of bees of the tribe Eucerini in North and Central America (Hymenoptera, Apoidea). *Am Mus Novit.* 1957;1837:1–44.
- LaBerge WE. A revision of the bees of the genus *Melissodes* in North and Central America. Part III. *Univ Kansas Sci Bull.* 1961;42:283–663.
- LaBerge WE. Revision of the bees of the genus *Tetraloniella* in the New World (Hymenoptera: Apidae). *Ill Nat Hist Surv Bull.* 2001;36(1–3):67–162. <https://doi.org/10.21900/j.inhs.v36.125>
- Landis MJ, Freyman WA, Baldwin BG. Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution.* 2018;72(11):2343–2359. <https://doi.org/10.1111/evo.13594>
- Latreille PA. Les insectes. In: Cuvier GCLD, editor. Règne animal. Paris: J.-B. Baillière; 1829. p. 556.
- Lepeletier ALM. Histoire naturelle des insectes. In: Hyménoptères. 2nd ed. Paris: Librairie Encyclopédique de Roret; 1841.
- López-Uribe MM, Cane JH, Minckley RL, Danforth BN. Crop domestication facilitated rapid geographical expansion of a specialist pollinator, the squash bee *Peponapis pruinosa*. *Proc Biol Sci.* 2016;283(1833):20160443. <https://doi.org/10.1098/rspb.2016.0443>
- Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis; 2019.
- Martins AC, Melo GAR. The New World oil-collecting bees *Centris* and *Epicharis* (Hymenoptera, Apidae): molecular phylogeny and biogeographic history. *Zool. Scr.* 2016;45:22–33.
- Martins AC, Melo GAR. The New World oil-collecting bees *Centris* and *Epicharis* (Hymenoptera, Apidae): molecular phylogeny and biogeographic history. *Zool Scr.* 2016;45(1):22–33.
- Martins AC, Santos LM, Dalmolin A. Danuncia Urban: a life devoted to entomology. In: Aguiar AJC., Gonçalves RB, Ramos KS, editors. *Ensaio Sobre as Abelhas Da Região Neotropical: Homenagem Aos 80 Anos de Danuncia Urban.* Curitiba: Editora UFPR; 2015. p. 11–48.
- Masa-Iranzo I, Sanmartín I, Caruzo MBR, Riina R. Skipping the dry diagonal: spatio-temporal evolution of *Croton* section *Cleodora* (Euphorbiaceae) in the Neotropics. *Bot J Linn Soc.* 2021;197:61–84.
- McCormack JE, Tsai WLE, Faircloth BC. Sequence capture of ultraconserved elements from bird museum specimens. *Mol Ecol Resour.* 2016;16:1189–1203.
- Melo GAR. Stingless bees (Meliponini). In: Starr C, editors. *Encyclopedia of social insects.* Cham: Springer; 2020. p. 1–18.
- Michelette ERF, Camargo JMF. Bee-plant community in a xeric ecosystem in Argentina. *Rev Bras Zool.* 2000;17:651–665.
- Michener CD. Taxonomic observations on bees with descriptions of new genera and species (Hymenoptera; Apoidea). *J N Y Entomol Soc.* 1942;50:273–282.
- Michener CD. Biogeography of the bees. *Ann Mo Bot Gard.* 1979;66(3):277–347. <https://doi.org/10.2307/2398833>
- Michener CD. The bees of the World. 2nd ed. Baltimore: The Johns Hopkins University Press; 2007.
- Michener CD, LaBerge WE, Moure JS. Some American Eucerini bees. *Dusenía.* 1955;6:213–228.
- Michener CD, Moure JS. Generic positions of some South American eucerine bees (Hymenoptera, Apoidea). *Dusenía.* 1956;7:277–290.
- Michener CD, Moure JS. A study of the classification of the more primitive non-parasitic Anthophorine bees (Hymenoptera, Apoidea). *Bull Am Mus Nat Hist.* 1957;112:395–452.
- Minckley RL, Radke WR. Extreme species density of bees (Apiformes, Hymenoptera) in the warm deserts of North America. *J Hymenopt Res.* 2021;82:317–345. <https://doi.org/10.3897/jhr.82.60895>
- Minh BQ, Nguyen MAT, Von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 2013;30(5):1188–1195. <https://doi.org/10.1093/molbev/mst024>
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37:1530–1534.
- Moure JS. A new genus and two new species of eucerine bees from South America (Hymenoptera, Apoidea). *Dusenía.* 1967;8:147–152.
- Moure JS, Michener CD. A contribution toward the classification of neotropical Eucerini (Hymenoptera, Apoidea). *Dusenía.* 1955;6:239–331.
- Moure JS, Urban D, Melo GAR. Catalogue of Bees (Hymenoptera, Apoidea) in the Neotropical Region—online version; 2012. <http://www.moure.cria.org.br/catalogue>.
- Moyle RG, Oliveros CH, Andersen MJ, Hosner PA, Benz BW, Manthey JD, Travers SL, Brown RM, Faircloth BC. Tectonic collision and uplift of Wallace triggered the global songbird radiation. *Nat Commun.* 2016;7(1):1–7.
- Nee S, May RM, Harvey PH. The reconstructed evolutionary process. *Philos Trans R Soc Lond B Biol Sci.* 1994;344(1309):305–311. <https://doi.org/10.1098/rstb.1994.0068>
- Orr MC, Branstetter MG, Straka J, Yuan F, Leijes R, Zhang D, Zhou Q, Zhu C-D. Phylogenomic interrogation revives an overlooked hypothesis for the early evolution of the bee family Apidae (Hymenoptera: Apoidea), with a focus on the subfamily Anthophorinae. *Insect Syst Divers.* 2022;27:1019–1025.
- Orr MC, Hughes AC, Chesters D, Pickering J, Zhu C-D, Ascher JS. Global patterns and drivers of bee distribution. *Curr Biol.* 2021;31:1–8.

- Parker FD. How efficient are bees in pollinating sunflowers? *J Kans Entomol Soc.* 1981b;54:61–67.
- Parker FD. Notes on the biology of a common sunflower bee, *Melissodes (Eumelissodes) agilis* Cresson. *J N Y Entomol Soc.* 1981a;89:43–52.
- Patton WH. Generic arrangement of the bees allied to *Melissodes* and *Anthophora*, Vol. 5. Washington, D.C.: U S Geol Geogr Surv Territories; 1879. p. 471–479.
- Pérez-Escobar OA, Zizka A, Bermúdez MA, Meseguer AS, Condamine FL, Hoorn C, Hooghiemstra H, Pu Y, Bogarín D, Boschman LM, *et al.* The Andes through time: evolution and distribution of Andean floras. *Trends Plant Sci.* 27–2022;27:364–378.
- Philippe H, Snell EA, Baptiste E, Lopez P, Holland PWH, Casane D. Phylogenomics of eukaryotes: impact of missing data on large alignments. *Mol Biol Evol.* 2004;21(9):1740–1752. <https://doi.org/10.1093/molbev/msh182>
- Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis for MCMC. *R News.* 2006;6:7–11.
- Pound MJ, Haywood AM, Salzmann U, Riding JB, Lunt DJ, Hunter SJ. A Tortonian (late Miocene, 11.61–7.25 Ma) global vegetation reconstruction. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2011;300(1–4):29–45. <https://doi.org/10.1016/j.palaeo.2010.11.029>
- Praz CJ, Packer L. Phylogenetic position of the bee genera *Ancyla* and *Tarsalia* (Hymenoptera: Apidae): a remarkable base compositional bias and an early Paleogene geodispersal from North America to the Old World. *Mol Phylogenet Evol.* 2014;81:258–270. <https://doi.org/10.1016/j.ympev.2014.09.003>
- Puttick MN, Schwartz R. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics.* 2019;35:5321–5322.
- R Core Team. R: a language and environment for statistical computing; 2020.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol.* 2018;67(5):901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ramos KS, Martins AC, Melo GAR. Evolution of andrenine bees reveals a long and complex history of faunal interchanges through the Americas during the Mesozoic and Cenozoic. *Mol Phylogenet Evol.* 2022;172:107484. <https://doi.org/10.1016/j.ympev.2022.107484>
- Rasmussen C, Gonzalez V, Engel MS, Michener CD. In memoriam: Wallace Edmund LaBerge (1927–2013). *J Melittol.* 2013;17:1–22.
- Ravelo AC, Andreasen DH, Lyle M, Lyle AO, Wara MW. Regional climate shifts caused by gradual global cooling in the Pliocene epoch. *Nature.* 2004;429:263–267.
- Raymo ME, Grant B, Horowitz M, Rau GH. Mid-Pliocene warmth: stronger greenhouse and stronger conveyor. *Mar Micropaleontol.* 1996;27:313–326.
- Ree RH. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution.* 2005;59(2):257–265.
- Ree RH, Smith SA. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and Cladogenesis. *Syst Biol.* 2008;57(1):4–14. <https://doi.org/10.1080/10635150701883881>
- dos Reis M, Yang Z. Approximate likelihood calculation on a phylogeny for Bayesian Estimation of Divergence Times. *Mol Biol Evol.* 2011;28(7):2161–2172. <https://doi.org/10.1093/molbev/msr045>
- dos Reis M, Yang Z. Bayesian molecular clock dating using genome-scale datasets. In: Anisimova M, editor. *Evolutionary genomics: statistical and computational methods.* New York: Springer; 2019. p. 309–330.
- Ribeiro PG, Torres Jiménez MF, Andermann T, Antonelli A, Bacon CD, Matos-Maraví P. A bioinformatic platform to integrate target capture and whole genome sequences of various read depths for phylogenomics. *Mol Ecol.* 2021;30:6021–6035.
- Robertson C. Descriptions of new species of North American bees. *Trans Am Entomol Soc.* 1891;18:49–66.
- Robertson C. Some new or little-known bees—II. *Can Entomol.* 1902;34(2):48–49. <https://doi.org/10.4039/ent3448-2>
- Robertson C. Synopsis of Anthophila. *Can Entomol.* 1904;36(2):37–43. <https://doi.org/10.4039/ent3637-2>
- Roig-Alsina AH, Michener CD. Studies of the phylogeny and classification of long-tongued bees. *Univ Kansas Sci Bull.* 1993;55:123–173.
- Rueda M, Rodríguez M, Hawkins BA. Identifying global zoogeographical regions: lessons from Wallace. *J Biogeogr.* 2013;40(12):2215–2225. <https://doi.org/10.1111/jbi.12214>
- Salter JF, Hosner PA, Tsai WLE, McCormack JE, Braun EL, Kimball RT, Brumfield RT, Faircloth BC. Historical specimens and the limits of subspecies phylogenomics in the New World quails (Odontophoridae). *Mol Phylogenet Evol.* 2022;175:107559. <https://doi.org/10.1016/j.ympev.2022.107559>
- Sanmartín I, Enghoff H, Ronquist F. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol J Linn Soc.* 2001;73:345–390.
- Santos Júnior JE, Williams PH, Dias CAR, Silveira FA, Faux P, Coimbra RTF, Campos DP, Santos FR. Biogeography and Diversification of Bumblebees (Hymenoptera: Apidae), with Emphasis on Neotropical Species. *Diversity (Basel).* 2022;14:238.
- dos Santos FM, de Carvalho CAL, Silva RF. Diversidade de abelhas (Hymenoptera: Apoidea) em uma área de transição Cerrado-Amazônia. *Acta Amazonica.* 2004;34:319–328.
- Scopoli GA. *Annus historico-naturalis*, Vol. 4. Lipsiae: C. G. Hilscher; 1770. p. 1–152.
- Sheppard WS, McPherson BA. Ribosomal DNA diversity in Apidae. *Diversity in the Genus Apis.* Boulder: Westview Press; 1991. p. 89–102.
- Silveira FA, Rocha LB, Cure JR, Oliveira MJDO. Abelhas silvestres da zona da mata de Minas Gerais. II. Diversidade, abundância e fontes de alimento em uma pastagem abandonada de ponte nova. *Rev Bras Entomol.* 1993;37:595–610.
- Silveira FA, Campos MJO. A melissofauna de Corumbataí e Paraopeba e uma análise da biogeografia das abelhas do cerrado brasileiro. *Revta bras. ent.* 1995;39:371–401.
- Sless TJJ, Branstetter MG, Gillung JP, Krichilsky EA, Tobin KB, Straka J, Rozen JG, Freitas FV, Martins AC, Bossert S, *et al.* Phylogenetic relationships and the evolution of host preferences in the largest clade of brood parasitic bees (Apidae: Nomadinae). *Mol Phylogenet Evol.* 2022;166:107326. <https://doi.org/10.1016/j.ympev.2021.107326>
- Smith F. Catalogue of hymenopterous insects in the collection of the British museum. Part II Apidae. London: Printed by order of the Trustees; 1854.
- Smith BT, Mauck WM, Benz BW, Andersen MJ. Uneven Missing Data Skew Phylogenomic Relationships within the Lories and Lorikeets. *Genome Biol. Evol.* 2020;12:1131–1147.
- Spinola M. Compte rendu des Hymenopteres recueillis par M. Fischer pendant son voyage en Egypte. *Ann Soc Entomol Fr.* 1838;7:437–546.
- Spinola M. Hymenopteros. In: Gay C, editors. *Historia Fisica y Politica de Chile.* Zoologia, Vol. 6. Paris: Maulde & Renon; 1851. p. 153–569.
- Springer MS, Gatesy J. The gene tree delusion. *Mol Phylogenet Evol.* 2016;94(Pt A):1–33. <https://doi.org/10.1016/j.ympev.2015.07.018>
- Strömberg CAE. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proc Natl Acad Sci U S A.* 2005;102:11980–11984.
- Stroynowski Z, Ravelo AC, Andreasen D. A Pliocene to recent history of the Bering Sea at Site U1340A, IODP Expedition 323. *Paleoceanography.* 2015;30(12):1641–1656. <https://doi.org/10.1002/2015pa002866>
- Swezey CS. Cenozoic stratigraphy of the Sahara, Northern Africa. *J Afr Earth Sci.* 2009;53(3):89–121. <https://doi.org/10.1016/j.jafrearsci.2008.08.001>
- Tagliacollo VA, Lanfear R. Estimating improved partitioning schemes for ultraconserved elements. *Mol Biol Evol.* 2018;35(7):1798–1811. <https://doi.org/10.1093/molbev/msy069>
- Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 2007;56:564–577.
- Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipowski A, Kumar S. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci U S A.* 2012;109(47):19333–19338. <https://doi.org/10.1073/pnas.1213199109>
- Tange O. GNU Parallel 2018; 2018.
- Tao Q, Barba-Montoya J, Kumar S. Data-driven speciation tree prior for better species divergence times in calibration-poor molecular phylogenies. *Bioinformatics.* 2021;37(Suppl_1):i102–i110. <https://doi.org/10.1093/bioinformatics/btab307>

- Trunz V, Packer L, Vieu J, Arrigo N, Praz CJ. Comprehensive phylogeny, biogeography and new classification of the diverse bee tribe Megachilini: can we use DNA barcodes in phylogenies of large genera? *Mol Phylogenet Evol.* 2016;103:245–259. <https://doi.org/10.1016/j.ympev.2016.07.004>
- Urban D. As espécies do gênero *Dasyhalonia* Michener LaBerge & Moure, 1955 (Hymenoptera—Apoidea). *Rev Bras Biol.* 1967b;27:247–266.
- Urban D. As espécies do gênero *Thygater* Holmberg, 1984 (Hymenoptera, Apidae). *Bol Univ Fed Paraná.* 1967c;12:177–309.
- Urban D. O gênero *Lophothygater* Moure & Michener, 1955 (Hymenoptera, Apoidea). *Dusenía.* 1967a;8:135–145.
- Urban D. As espécies de *Gaesischia* Michener, LaBerge e Moure, 1955. *Bol Univ Fed Paraná.* 1968a;3:79–129.
- Urban D. As espécies do gênero *Melissoptila* Holmberg, 1884 (Hymenoptera—Apoidea). *Rev Bras Entomol.* 1968b;13:1–94.
- Urban D. As espécies de *Alloscirtetica* Holmberg, 1909. *Bol Univ Paraná.* 1971;16:307–369.
- Urban D. As espécies de *Svastrides* Michener, LaBerge e Moure, 1955. *Rev Bras Biol.* 1972;32:485–498.
- Urban D. As espécies sulamericanas do gênero *Melissodes* (Latreille, 1829) (Hymenoptera, Apoidea). *Rev Bras Biol.* 1973;33:201–220.
- Urban D. Sobre o gênero *Alloscirtetica* Holmberg, 1909 (Hymenoptera, Apoidea). *Dusenía.* 1982;13:65–80.
- Urban D. Dois gêneros novos de Eucerinae neotropicais (Hymenoptera, Apoidea). *Rev Bras Zool.* 1989b;6(1):117–124. <https://doi.org/10.1590/s0101-81751989000100013>
- Urban D. Espécies novas e notas sobre o gênero *Gaesischia* Michener, LaBerge, Moure, 1955 (Hymenoptera, Apoidea). *Rev Bras Entomol.* 1989a;33:75–102.
- Urban D. *Mirnapis inca*, gênero e espécies novos de Eucerinae da América do sul (Hymenoptera, Apoidea, Anthophoridae). *Rev Bras Zool.* 1997;14(3):565–569. <https://doi.org/10.1590/s0101-81751997000300006>
- Urban D. Espécies novas de *Melissoptila* Holmberg da América do Sul e notas taxonômicas (Hymenoptera, Anthophoridae). *Rev Bras Zool.* 1998;15(1):01–46. <https://doi.org/10.1590/s0101-81751998000100001>
- Urban D, Moure JS, Melo GAR. Eucerini Latreille, 1802. In: Urban D, Moure JS, Melo GAR, editors. *Catalogue of bees (Hymenoptera, Apoidea) in the Neotropical region—Online version.* Curitiba; 2012.
- Varela L, Tambusso PS, McDonald HG, Fariña RA. Phylogeny, Macroevolutionary trends and historical biogeography of sloths: insights from a Bayesian morphological clock analysis. *Syst Biol.* 2019;68(2):204–218. <https://doi.org/10.1093/sysbio/syy058>
- Wiens JJ. Missing data, incomplete taxa, and phylogenetic accuracy. *Syst Biol.* 2003;52(4):528–538. <https://doi.org/10.1080/10635150390218330>
- Wilson JS, Carril OM, Sipes SD. Revisiting the Great American Biotic Interchange through analyses of amphitropical bees. *Ecography.* 2014;37:791–796.
- Wolfe JA. An analysis of Neogene climates in Beringia. *Palaeogeogr Palaeoclimatol Palaeoecol.* 1994;108(3-4):207–216. [https://doi.org/10.1016/0031-0182\(94\)90234-8](https://doi.org/10.1016/0031-0182(94)90234-8)
- Wright KW, Miller KB, Song H. A molecular phylogeny of the long-horned bees in the genus *Melissodes* Latreille (Hymenoptera: Apidae: Eucerinae). *Insect Syst Evol.* 2020;52(4):428–443.
- Yang Z. Paml: a program package for phylogenetic analysis by maximum likelihood. *Bioinformatics.* 1997;13(5):555–556. <https://doi.org/10.1093/bioinformatics/13.5.555>
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24(8):1586–1591. <https://doi.org/10.1093/molbev/msm088>
- Yang Z, Rannala B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol Biol Evol.* 2006;23(1):212–226. <https://doi.org/10.1093/molbev/msj024>
- Yin J, Zhang C, Mirarab S, Schwartz R. ASTRAL-MP: scaling ASTRAL to very large datasets using randomization and parallelization. *Bioinformatics.* 2019;35:3961–3969.
- Zachos J, Dickens G, Zeebe R. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature.* 2008;451:279–283.
- Zanella FCV. Evolução da biota da diagonal de formações abertas secas da América do Sul. In: Carvalho CJB., Almeida EAB, editors. *Biogeografia Da América Do Sul: Padrões e Processos.* São Paulo: ROCA; 2011. p. 198–220.
- Zavortink TJ. A new genus and species of Eucerine bee from North America (Hymenoptera: Anthophoridae). *Proc Calif Acad Sci.* 1975;40:231–242.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinf.* 2018;19 (S6):153.
- Zhang F, Ding Y, Zhu CD, Zhou X, Orr MC, Scheu S, Luan YX. Phylogenomics from low-coverage whole-genome sequencing. *Methods Ecol Evol.* 2019;10:507–517.

Appendix

All taxonomic decisions considering the phylogenetic results presented in the main text are detailed and formalized in this appendix. Lists of synonyms are provided by Michener (2007) for the taxa distributed worldwide and by Urban et al. (2012) for Neotropical groups. New synonyms are indicated when needed.

Taxonomy

EUCERINI ROBERTSON

ALLOSCIRTETICINA Freitas and Silveira **NEW SUBTRIBE**

(urn:lsid:zoobank.org:act:713B57ED-B32B-4E85-B243-28C33AC00BD3)

Diagnosis: Six, infrequently five, maxillary palpomeres; male sixth and seventh terga without gradular or lateral teeth, male pygidial plate triangular and vestigial; males seventh sterna apical lobes are large, delicate, hairy, and elaborate; between these lobes, there are unpaired median projections between them which are rather large, delicate, hairy, and elaborate.

Included genera: *Alloscirtetica* Holmberg. *Dasyrcirtetica* Michener, LaBerge and Moure **stat. n.**, *Megascirtetica* Moure and Michener.

Comments: *Alloscirteticina* is mainly a southern South American clade, with a few species occurring in central or northern South America. It is composed of the following genera: *Alloscirtetica*, *Dasyrcirtetica*, and *Megascirtetica*. Several subgenera have been proposed in the classification of *Alloscirtetica* (Michener et al. 1955, Moure and Michener 1955, Michener and Moure 1956). Still, in the last comprehensive revision of the genus, Urban (1982) suppressed all of them. Later, Michener (2007) included *Megascirtetica* as a subgenus of *Alloscirtetica*, but it was kept as a genus in the Moure catalog of Neotropical bees (Urban et al. 2012). Considering the distinctiveness of these three groups and that *Megascirtetica* was placed either as the sister of *Alloscirtetica* (Fig. 2) or of *Dasyrcirtetica* and both sister to *Alloscirtetica* (Figs. 2 and Supplementary Figs. S1–S7), these three taxa are treated as genera.

Dasyrcirtetica Michener, LaBerge and Moure, **stat. n.**

(urn:lsid:zoobank.org:act:3904C642-83AF-4AEB-A11A-98AC40B73A35)

Diagnosis: Pilosity predominantly yellow or pale; labrum always yellow; clypeus entirely yellow or at least with a yellow apical band; maxillary palpi always with six palpomeres. Males: seventh tergum not completely round; lateral margins of vestigial pygidial plate never meet each other or the margin of the tergum apically.

Comment: *Dasyrcirtetica* was firstly described as a subgenus of *Alloscirtetica* (Michener et al. 1955). However, its species included here were recovered either as the sister group of *Megascirtetica* plus *Alloscirtetica* (Fig. 2) or as the sister of *Megascirtetica* (Supplementary Figs. S1–S4). Thus, recognizing *Dasyrcirtetica* as a genus is necessary to prevent the recognition of a paraphyletic

Alloscirtetica. Species included in *Dasyscirtetica* here are those initially recognized by Urban (1971) in the subgenus: *Dasyscirtetica alvarengai* (Urban 1971) **comb. n.**; *Dasyscirtetica arrhenica* (Vachal 1904) **comb. n.**; *Dasyscirtetica gilva* (Holmberg 1884) **comb. n.**; and *Dasyscirtetica paraguayensis* (Friese 1909) **comb. n.** They can be identified using the keys presented by Urban (1971, 1982) for the species of *Alloscirtetica*. In the 1971 key, all of them are separated in the couplets 3 of the key to males and 15 of the key to females, while in the 1982 key, the couplets are 27 for males and 18 for females. Apparently, species in this genus are restricted to the eastern slopes of the Andes.

Megascirtetica Moure and Michener

Comment: *Megascirtetica* includes only one easily recognizable species, *M. mephistophelica* (Schrottky). It was already treated as a subgenus of *Alloscirtetica* (Michener 2007) or as a genus (Urban et al. 2012) but given its distinctiveness and the possible sister-group relationship to *Dasyscirtetica*, we are keeping it as a genus.

EUCERINA ROBERTSON

Diagnosis: Four to six maxillary palpomerers (very rarely three); long to short tongue and protuberant or non-protuberant clypeus; tegulae always rounded, with the anterolateral margin convex; wings with two or three submarginal cells; male S7 with the posterior lobe of the lateral process usually with an elevated, heavily sclerotized carina; this carina is usually linearly transverse, but sometimes having the basal or apical portions reduced or modified.

Included genera: *Eucera* Scopoli, *Protohalonia* Dorchin, *Simanthedon* Zavortink, *Tetralonia* Spinola, *Xenoglossa* Smith.

Comments: Eucerina comprehends the lineage referred to as the “Eucera complex” (Michener 2007, Dorchin et al. 2018). This is a primarily Holarctic group, and it is the most species-rich of the Eucerini subtribes. The limits of the lineage here are treated as Eucerina and the tree topology recovered for this subtribe is congruent with the findings of Dorchin et al. (2018). Four main clades are recognized: (i) *Simanthedon* plus *Protohalonia*, as the sister-group of the remaining lineages; (ii) a large, primarily Eastern Hemisphere genus, *Eucera*, with its three subgenera: *Eucera* s. str., *E. (Cubitalia)* and the Holarctic *E. (Synhalonia)*; (iii) *Tetralonia*, including the species traditionally recognized as the genus *Tetralonia* plus the Eastern Hemisphere species placed in the genus ‘*Tetraloniella*’ by Michener (2007); and (iv) the Western Hemisphere clade, the sister-group of *Tetralonia*.

The Western Hemisphere clade was strongly supported in all results and all its lineages are accommodated into a single genus, for which the oldest available name is *Xenoglossa* (Smith 1854). As defined here, *Xenoglossa* includes the following taxa: (i) the squash bees, (*Xenoglossa* s. str., *Peponapis*, and *Cemolobus*) plus *Syntrichalonia* and *Xenoglossodes* (each of them considered here as a subgenus of *Xenoglossa* s. l.); and (ii) clades with no subgeneric assignment. *X. (Xenoglossodes)* and the species not assigned to subgenera correspond to the Western Hemisphere’s species of ‘*Tetraloniella*’ as delimited by Michener (2007), recovered as a polyphyletic group by Dorchin et al. (2018) and confirmed here to represent many unrelated lineages. A broader taxonomic sampling of the ‘*Tetraloniella*’ from the Western Hemisphere should be employed in a future phylogenetic study so that its internal lineages can be better delimited, and their classification revised.

Eucera Scopoli

Included subgenera: *E. (Cubitalia)* Friese; *Eucera* s. str.; *E. (Synhalonia)* Patton.

Comment: Here, the genus *Eucera* is considered in a reduced scope, including only the primarily Palearctic subgenera. Considering the vast diversity in the genus and our limited sampling it is possible that the limits between *Eucera* s. str. and *E. (Cubitalia)* change when more species are included. Previous results by Dorchin et al. (2018), recovered *E. (Cubitalia)* nested among the species here considered in *Eucera* s. str., already suggesting the difficult delimitation of these taxa.

Tetralonia Spinola

Comment: *Tetralonia*, as delimited here, includes all species placed under this genus name by Michener (2007) plus the Eastern Hemisphere species of ‘*Tetraloniella*’, the latter proved to be a junior synonym of the former.

Xenoglossa Smith

Included subgenera: *Xenoglossa* s. str.; *X. (Cemolobus)* Robertson; *X. (Peponapis)* Robertson; *X. (Syntrichalonia)* LaBerge; *X. (Xenoglossodes)* Ashmead.

Comment: All subgenera included in *Xenoglossa* are kept with the same limits defined by Dorchin et al. (2018). The only exception is *X. (Xenoglossodes)* in which only three species are included at this point: *X. (Xenoglossodes) albata* **comb. n.** (Cresson 1872) its type species, *X. (Xenoglossodes) salviae* **comb. n.** (LaBerge 1989), and *X. (Xenoglossodes) brevifellator* **comb. n.** (LaBerge 1957). *X. (Xenoglossodes)* is a historically complicated group and a specific study of the species previously included in ‘*Tetraloniella*’ is needed to clarify if any of them is part of the subgenus and if new subgenera will be needed to accommodate any group of those species.

GAESISCHIINA Freitas and Silveira NEW SUBTRIBE

(urn:lsid:zoobank.org:act:DDD10F6D-7794-417F-92FB-223995741824)

Diagnosis: Three to five maxillary palpomerers; seventh sternum of males with large inner plates; gonostyli usually long and tapering; arolia present or not (minute in *Micronychapis*); with or without gradular spines or teeth on the last terga; males seventh sterna with large inner plates, generally complex. This is a considerably morphological-homogeneous group of genera, primarily from tropical South America, with a few groups restricted to temperate South America and others reaching the southern United States.

Included genera: *Dasyhalonia* Michener, LaBerge and Moure, *Florilegus* Robertson, *Gaesischia* Moure and Michener, *Gaesochira* Moure and Michener, *Hamatothrix* Urban, *Micronychapis* Moure and Michener, *Pachysvastra* Moure and Michener, *Savannychapis* Freitas, *Platysvastra* Moure, *Santiago* Urban, *Svastrina* Moure and Michener.

Comments: Most of the genera and subgenera recognized in classifications were recovered as reciprocally monophyletic. The exceptions are mentioned below.

Interesting results are the sister relationship of *Svastrina* and *Pachysvastra*, two genera with no arolia, which suggests a single loss of this feature in the tribe. The only other genus in the tribe without arolia is *Canephorula*, but it is not closely related to *Gaesischiina*.

Gaesischia Michener, LaBerge and Moure

Included subgenera: *Gaesischia* s. str., *G. (Gaesichiana)* Michener, LaBerge and Moure, *G. (Gaesichiopsis)* Michener, LaBerge and Moure, *G. (Gaesichioides)*.

Comments: *Gaesischia* is a genus with a complicated taxonomic history (Michener et al. 1955, Moure and Michener 1955, LaBerge 1958, Urban 1968a, 1989a). Thanks to our broad taxon sampling, we were able to show that *G. hyptidis* is a distinct lineage in relation to *Gaesischia* s. str., and *G. (Gaesischiopsis)*, which was the main argument for its positioning in a new subgenus (*G. (Gaesischioides)*) by Freitas (2022), and that *G. patellicornis* is part of *Gaesischia* s. tr.

MELISSODINA Freitas and Silveira **NEW SUBTRIBE**

(urn:lsid:zoobank.org:act:2194C10D-8E8F-471F-9F88-F707162AD2B5)

Diagnosis: Three to five maxillary palpomerers (rarely two in *Melissoptila*); arolia always present; tegula rounded or tapering anteriorly with a concave to straight edge; males with or without gradular teeth or spines on the last terga; male S7 variable, generally complex, but small and short in *Melissodes* and *Epimelissodes* **stat. n.**, which are the most species-rich genera in the subtribe.

Included genera: *Agapanthinus* LaBerge, *Epimelissodes* Ashmead **stat. n.**, *Martinapis* Cockerell, *Melissodes* Latreille, *Melissoptila* Holmberg, *Mirnapis* Urban, *Svastra* Holmberg, *Svastrides* Michener, LaBerge and Moure, *Svastropsis* Moure and Michener **stat. n.**

Comments: Melissodina is composed of two main clades: (i) *Melissodes* plus *Epimelissodes*, both primarily Nearctic, and (ii) the remaining genera. This second lineage shows an antitropical distribution, with *Melissoptila*, *Mirnapis*, *Svastra*, *Svastrides*, and *Svastropsis* being South American, and *Agapanthinus* and *Martinapis*, Nearctic. It has *Melissoptila* as its first diverging lineage, followed by *Svastrides*, and a small Nearctic clade with *Agapanthinus* (monotypic) sister to *Martinapis*, now composed only of two Nearctic species. The next lineage in this clade includes *Mirnapis*, *Svastropsis* (previously considered a subgenus of *Martinapis*), sister to *Svastra* (now including only its five South American species). The Nearctic clade has *Epimelissodes* (including all primarily Nearctic groups previously considered as subgenera of *Svastra*—*E. (Anthedonia)* **comb. n.**, *E. (Brachymelissodes)* **comb. n.**, *E. (Idiomelissodes)* **comb. n.**, *Epimelissodes* s. str.); sister to *Melissodes*, now including five subgenera: *M. (Apomelissodes)*; *M. (Callimelissodes)*; *M. (Ecleptica)*; *M. (Eumelissodes)*; and *Melissodes* s. str.

Epimelissodes Ashmead **stat. n.**

(urn:lsid:zoobank.org:act:592EF296-57DE-481B-B7B8-89C9FA75C10E)

Included subgenera: *E. (Anthedonia)* Michener **comb. n.**, *E. (Brachymelissodes)* LaBerge **comb. n.**, *Epimelissodes* s. str., *E. (Idiomelissodes)* LaBerge **comb. n.**

Comment: As *Svastra* was recovered as an independent lineage, *Epimelissodes* was elevated to generic status and now accommodates all the subgenera previously positioned in *Svastra*. *Epimelissodes* is a Nearctic genus with a few species occurring in the Neotropics.

Martinapis Cockerell

Comment: Since *Svastropsis*, previously considered a subgenus of *Martinapis*, was recovered as sister to *Svastra*, *Martinapis*, now includes only the two Nearctic species originally assigned to its nominal subgenus.

Melissodes Latreille

Included subgenera: *M. (Apomelissodes)* LaBerge, *M. (Callimelissodes)* LaBerge, *M. (Ecleptica)* Holmberg, *M. (Eumelissodes)* LaBerge, *Melissodes* s. str.

M. (Eumelissodes) LaBerge

= *Melissodes (Heliomelissodes)* LaBerge 1956b **syn. n.**; Type species: *Melissodes desponsus* (Smith 1854), by original designation.

M. (Apomelissodes) LaBerge

= *Melissodes (Psilomelissodes)* LaBerge 1956b **syn. n.**; Type species: *Melissodes intortus* Cresson, 1872 *Melissodes (Tachymelissodes)* LaBerge 1956b **syn. n.**; Type species: *Melissodes dagosus* (Cockerell 1909)

Comment: *M. (Apomelissodes)* is redefined to accommodate, in addition to the species it originally contained, all the species previously included in *M. (Psilomelissodes)*, *M. (Tachymelissodes)*, and *Melissodes paucipuncta* LaBerge, previously included in *M. (Eumelissodes)*. This scope for this subgenus is also in agreement with the results of Wright et al. (2020).

Svastropsis Moure and Michener **stat. n.**

(urn:lsid:zoobank.org:act:8D714498-4A32-4C54-932D-3DE934E118A9)

Comment: *Svastropsis* includes only *S. bipunctata* (Friese) **stat. n.** a rare species from southern South America. Previously considered a subgenus of *Martinapis*, it was recovered as an independent lineage requiring its elevation to generic status.

Svastra Holmberg

Comment: *Svastra* now only includes the South American species originally classified into *Svastra* s. str. The remaining subgenera of *Svastra* in previous classifications (e.g., Michener 2007) are herein transferred to the genus *Epimelissodes*.

THYGATERINA Freitas and Silveira **NEW SUBTRIBE**

(urn:lsid:zoobank.org:act:2B78BAE6-17A1-4064-91AD-D79BBC849896)

Diagnosis: Three to five maxillary palpomerers; always protuberant clypeus; paraocular carina present only on upper paraocular area; males' antennae extremely long, the longest among Eucerini; wings frequently with three submarginal cells (two in *Trichocerapis (Dithygater)*); females gradulus of S2 weakly biconvex; male's medium plate of S7 divided into two lobes.

Included genera: *Lophothygater* Moure and Michener, *Thygater* Holmberg, *Trichocerapis* Cockerell.

Comments: The recognition of Thygaterina formalizes a well known clade of Eucerini, previously referred to as the *Thygater-Trichocerapis* group (e.g., Michener 2007, Freitas et al. 2019). It was first recognized by Moure and Michener (1955) as the most remarkable of the groups of Eucerini. Our results confirm and reinforce the findings of Freitas et al. (2019) regarding the subgenera of *Thygater* and allow the inclusion of *T. melanotricha* in *Thygater* s. str. This supports the suggestions of Freitas et al. (2019) that *T. (Nectarodiaeta)* is a group restricted to subtropical-temperate regions of South America, while *Thygater* s. str. is a widespread clade throughout the Neotropical region. Another implication of our findings is the positioning of a species with four maxillary palpomerers in *Thygater* s. str., indicating that the number of maxillary palpomerers cannot be considered as an unambiguous diagnostic character to distinguish these subgenera.