Biogeography and early diversification of Tapinotaspidini oil-bees support presence of Paleocene savannas in South America

Antonio J.C. Aguiar[a,*], Gabriel A.R. Melo[b], Thais N.C. Vasconcelos[c], Rodrigo B. Gonçalves[b], Lilian Giugliano[d], Aline C. Martins[a,e]

[a] *Departamento de Zoologia, Universidade de Brasilia, 70910-900 Brasilia, Distrito Federal, Brazil*
[b] *Department de Zoologia, Universidade Federal do Paraná, PR 19020, 81531-980 Curitiba, Parana, Brazil*
[c] *Departmento de Botanica, Universidade de Sao Paulo, Sao Paulo, SP, Brazil*
[d] *Departmento de Genética e Morfologia, Universidade de Brasilia, 70910-900 Brasilia, Distrito Federal, Brazil*
[e] *Departmento de Ecologia, Universidade de Brasilia, 70910-900 Brasília, Distrito Federal, Brazil*

ABSTRACT

Worldwide distributed tropical savannas were established only in the Miocene, with climatic cooling and rise of C4 grasses. However, there is evidence for an earlier presence of savanna-like vegetation in southern parts of South America. Here we investigated the biogeographic history of a clade of solitary bees which have endemic groups in areas covered by savannas and other types of open vegetation as well as forested areas. We hypothesized that these bees originated in savanna-like biomes and that shifts to forested areas and floral host shifts increased species diversification along their evolutionary history. We reconstructed a comprehensive phylogeny for Tapinotaspidini bees based majorly on original DNA sequences. We then used macroevolutionary tools to estimate ancestral range area and reconstructed ancestral habitat (open versus forested) and host plant association to analyze the effects of shifts in vegetation type and flower hosts on their diversification. Tapinotaspidini bees originated in the Paleocene and in a savanna-type, Cerrado-like, which is reinforced by reconstruction of open vegetation as the most probable ancestral area, thus bringing additional evidence to a much earlier origin of this vegetation type in South America. Shifts to forested areas occurred at least three times in a period of 30 Ma and were responsible for slight increases in diversification rates. Malpighiaceae is the ancestral floral host; host broadening occurred only in the Miocene and at least in three occasions. Host shifts, i.e. from Malpighiaceae to other oil families, occurred in the Eocene and Miocene. Both host broadening and host shifts did not significantly alter diversification rates, however exploitation of other oil sources were important in occupying new habitats. The link between biomes and host plant shifts and changes in diversification rate brings us additional insights into the evolution of bees and associated flora in South America.

1. Introduction

South America is one of the most diverse areas for plants and animals in the world (Antonelli and Sanmartín, 2011). Within this continent, diversity is unevenly distributed and both tropical humid forests and savannas comprise particularly species-rich areas. Present-day distribution of South American biomes shows large portions of tropical rainforests divided by a diagonally organized continuum of open vegetation biomes, including the Caatinga, the Brazilian Savannas (Cerrado) and the Chaco, and others in the Andes and in isolated patches of seasonally dry tropical forests (Werneck, 2011). The Andean Mountain uplift – a readjustment process in plate tectonics that began in the early Paleogene – formed a barrier to rainfalls over an extensive area in the

continent, which, associated to climatic cooling and drying during the Oligocene/Miocene, influenced the expansion of open vegetation biomes (Ortiz-Jaureguizar and Cladera, 2006; Simon et al., 2009). The dry diagonal of open vegetation represents an ensemble of different biomes sharing a warm climate with seasonal drought (Pennington et al., 2006), and may have not been fully established before the late Pliocene (Simon et al. 2009, Pennington and Hughes, 2014).

Unlike most living beings that abound in tropical rainforests (Willig et al., 2003), bees are predominantly abundant and diverse in xeric and warm temperate regions of the world (Michener, 1979). Although an extensive treatment of the bees biogeography anchored in phylogenetic evidence is still lacking, their current distribution pattern associated to xeric areas indicates a probable origin associated to this kind of climate

(Michener, 1979). The association of bees with dry areas is directly connected to their nesting habits. Most bees nest in the ground, usually using the soil as substrate to dig their brood cells and provisioning them with a mixture of pollen and nectar to feed the growing larvae (Michener, 2007). Ground-nesting bees face a significant challenge: soil humidity facilitating infestation by mold and other enemies, particularly in tropical humid areas, can spoil the larval food and compromise their survival and reproductive success. Ancestral bees possibly evolved a suite of nesting adaptations, such as use of foreign material, to overcome constraints brought by the exploitation of pollen as food, thus benefiting from the role as pollen feeders (Litman et al. 2011)

Notwithstanding, some groups of bees were able to conquest and succeed in tropical humid areas, for example the single pantropical bee group, the Meliponini or stingless bees (Camargo and Vit, 2013) and many groups of oil-collecting bees (e.g. *Centris*, *Epicharis*) (Martins and Melo, 2016; Rocha-filho et al., 2017). These groups overcome the ground-nesting challenge using different strategies, for example: nesting in pre-existing cavities in wood and other materials (Buschini, 2006), excavating galleries in dead wood (Gerling et al., 1989) or in walls of termite nests (Carrijo et al., 2012), occupying hollows in trunks (Camargo and Pedro, 2003) or constructing themselves aerial nests (Camargo and Pedro, 2003). Alternatively some of these bees make use of waterproofing linings, such as wax, plant resins, secreted cellophane-like substances, and floral oils, when building their nests (Michener, 2007). Use of oils as waterproofing material has been argued as one of the key elements that allowed the successful occupation of tropical forests by pollen basket bees, i.e. honeybees, stingless bees and orchid bees bees and their relatives *Centris* and *Epicharis* (Martins et al., 2014).

The tribe Tapinotaspidini (Apidae, Apinae) includes about 180 species of solitary bees distributed throughout the Neotropical region (Aguiar, 2012) and whose females and males collect floral oils for different purposes (Buchmann, 1987; Cappellari et al., 2011; Rozen et al., 2006). In several recent papers, Aguiar (Aguiar and Melo, 2007, 2011) revised the genera of tapinotaspidine bees, however some genera remains relatively unresolved, like *Arhysoceble* and *Chalepogenus* with numerous new species to describe. This tribe possesses representatives restricted to humid forested areas (*e.g.* the Amazon and the Brazilian Atlantic Forest) as well as groups found only in open and seasonally dry areas (the dry diagonal of open areas). In this sense they are a particularly interesting model to understand the evolutionary context behind shifts in habitat preferences between humid and dry areas and how historical biogeography is linked to key aspects of their biology.

Tapinotaspidine bees can vary from specialists to generalists in their oil-collecting behavior. Some species collect oil mostly from one angiosperm family, while others use multiple families as oil source, but are not necessarily pollen specialists on the same species. Among the seven Neotropical plant families that produce floral oils (Renner and Schaefer, 2010), Malpighiaceae and Iridaceae are the main source for Tapinotaspidini (Cocucci and Vogel, 2000; Sigrist and Sazima, 2004; Vogel, 1974), but also Plantaginaceae, Krameriaceae, Calceolariaceae and Orchidaceae can be visited by them (Carneiro et al., 2015; Martins and Alves-dos Santos, 2013; Sérsic, 2004). Malpighiaceae are most common in forested and in relatively wet open habitats, like the Cerrado savannas, not reaching either extremely dry or cold environments (Anderson, 1979). In austral latitudes and in xeric environments other oil plant families (e.g. Krameriaceae, Iridaceae) offer this resource to oil collecting bees that also abound in these regions (Martins et al., 2015). Interestingly, abundance and diversity of these oil-offering plant families changes between open or forested vegetation types. In other words, shifts among vegetation types were most likely accompanied by changes in host plant preferences by Tapinotaspidini bees. Malpighiaceae is arguably the most ancient oil-producing plant family, originated about 60 Mya (Renner and Schaefer, 2010; Xi et al., 2012), thus shifts probably occurred from Malpighiaceae to other, younger oil plant families whose origin took place 40 Mya or less. Both changes in habitat and in host preferences are known to affect diversification rates

in insects (e.g. Condamine et al., 2018) as these new evolutionary circumstances can affect survival and reproduction rates, all of which are important drivers in changes of macroevolutionary patterns. Combining ecological traits – such as mutualistic interactions between bees and plants – with spatial evolution and macroevolutionary dynamics remains elusive, but might reveal important insights to understand how biodiversity changes across time and space (Antonelli et al., 2018a).

We here reconstruct a robust phylogenetic inference to explore the biogeographical scenario that could explain the current distribution of tapinotaspidine bees in forested and open vegetation areas, the diversification dynamics linked to shifts between areas and among floral hosts from the ancient Malpighiaceae to other oil plant clades. The first paper to propose a phylogeny for the tribe based on morphological data was presented by Roig-Alsina (1997), although this is the first comprehensive treatment based on molecular dataset. In an evolutionary context we aimed to determine: (1) the most probable vegetation type occupied by the Tapinotaspidini ancestor (2) when and under which global conditions occurred the exchanges between forests and open areas and if those exchanges were accompanied by shifts in diversification rates (3) and if host plant shifts or host broadening increased diversification rates in the group.

2. Material and methods

2.1. Taxon sampling

Tapinotaspidini bees occur broadly in the Neotropical region from northern Argentina to central Mexico and the Caribbean. Some taxa are particularly diverse in forested areas (e.g. *Paratetrapedia* and close relatives), whereas most genera inhabit only open vegetation areas. We tried to include representatives of different vegetation types, covering a broad geographic extension in the Neotropical biogeographic regions, including Andes and Central America. We broadly sampled bees from the Tapinotaspidini tribe, including all genera, most of them represented by at least two species, totalizing 52 out of approximately 177 (29%) accepted species on the tribe. For only three species we used GenBank sequences, therefore we newly produced 218 sequences. We sampled the following taxa (in parentheses, sampled species, total number of species): *Arhysoceble* (4, 11), *Caenonomada* (2, 4), *Chalepogenus* (10, 21) *Lanthanomelissa* (2, 5), *Lophopedia* (4, 20), *Monoeca* (3, 15), *Nasutopedia* (2, 7), *Paratetrapedia* (6, 33), *Tapinotaspis* (2, 3), *Tapinotaspoides* (2, 7), *Trigonopedia* (2, 11), *Tropidopedia* (5, 20), *Urbanapis* (2, 5), *Xanthopedia* (6, 15). The following three species from our ingroup already had sequences available on GenBank: *Lophopedia* cf. *nigrispinis* (deposited as *Paratetrapedia* sp.), *Chalepogenus caeruleus* and *Chalepogenus herbsti* (Cardinal et al. 2010). Table S1 and S2 included all GenBank accession numbers.

To better accommodate fossil calibrations and to reinforce the monophyly of the tribe Tapinotaspidini, we made a large and representative outgroup sampling (32 species) (Cardinal et al. 2010). The choice of outgroup terminals was based on broader phylogenetic hypotheses for Apinae bees (Cardinal et al., 2010) and included: *Apine line*: Apini (2 species), Bombini (2); *Eucerine line*: Ancylaini (1), Emphorini (5), Eucerini (4), Exomalopsini (2); *Xylocopine line*: Allodapini (3), Ceratinini (5) and Xylocopini (3); and Anthophorini (5). Under the alternative classification system of Bossert et al. (2019) the mentioned groups are given subfamily status: Anthophorinae (Anthophorini), Xylocopinae (Xylocopine line), Eucerinae (Eucerine line, including Tapinotaspidini) and Apinae (Apine line).

2.2. Molecular sampling, alignment and phylogenetic inference

Dry pinned and alcohol preserved specimens were used for DNA extraction with commercial extraction kits. Tissue was taken from the thoracic musculature when extracted from fresh specimens preserved in alcohol or through washing the entire dry specimen in extraction

buffer. Five genes were selected due to the informative potential observed in previous studies of bee phylogeny (Almeida and Danforth, 2009; Danforth et al., 2013): elongation factor-1α, F2 copy; long wavelength rhodopsin; wingless; 28S (D2-D3 expansion region) and cytochrome *c* oxidase unit one (COI). The amplification was performed using published and newly designed primers (Table S3). PCR products were purified and sequenced by Macrogen Inc., South Korea. Sequences were trimmed and assembled using the software Geneious v. 8.1 (Biomatters, 2015). Alignments were performed in MAFFT v.7 (Katoh and Standley, 2013) also in Geneious using the 1PAM/k = 2, for the nucleotide scoring matrix; 1.53 for gap opening penalty and 0 offset value (default); and omitting gappy regions. Different alignment strategies were chosen depending on the gene characteristics and performed on the MAFFT online server (28S) or on the Geneious R8: Q-INS-I for 28S ribosomal gene, considering secondary structure; E-INS-I, recommended for sequences with multiple conserved domains and long gaps, for the protein coding genes that contain introns, i.e. EL-1alpha, LW-Rhodopsin and Wingless; and G-INS-I, recommended for sequences with global homology, for CO1. The resulting alignments were then checked and corrected manually for obvious errors. Intron and exon boundaries were determined through comparisons with GenBank sequences of closely related bee groups. We excluded introns of opsin for outgroup sequences due to the great variability of that region but maintained for ingroup sequences to increase the amount of data.

Each gene matrix was submitted to individual gene tree searches to check for well-supported incongruences among datasets. As no well-supported incongruences were found among gene trees, all genes were concatenated in one matrix of five loci, 84 taxa and 3997 nucleotides (22% of missing data, See Tables S1 and S2). All sequence editing and analysis were performed in Geneious R8. The dataset was partitioned following a scheme separating individual genes (28S rRNA and CO1) and introns and exons from the protein coding genes (EF-1alpha, LW-Rhodopsin and Wingless) totalizing 12 partitions. This scheme was considered in all phylogenetic analyses and divergence time estimation, except when indicated.

Maximum Likelihood (ML) analysis was performed in the software RAxML (Stamatakis, 2006) using the graphical interface RaxML GUI (Silvestro and Michalak, 2012) with simultaneous bootstrap support analysis with 1000 replicates. Bayesian tree searches were implemented in MrBayes v. 3.2 (Ronquist et al., 2012). The Markov chain Monte Carlo (MCMC) was run for 50 million generations, with trees sampled every 1000 generations. Default MrBayes priors were used, except for the implementation of a mixed evolutionary model of nucleotide substitution. Chain convergence was assessed in Tracer 1.7 (Rambaut et al., 2014) and 25% of trees obtained prior to convergence were discarded as burn-in. A 50% majority rule consensus tree was constructed with the remaining trees. The resulting trees were visualized and edited in FigTree v. 1.3 (Rambaut, 2009).

2.3. Fossil calibrations and molecular dating

For dating analysis, we relied on the same matrix submitted to phylogenetic analysis composed by 84 taxa of Tapinotaspidini plus other Apinae species as outgroups. Relative branching times were estimated under uncorrelated, lognormal, relaxed-clock model and a Yule tree prior in BEAST 2.4.4 (Bouckaert et al., 2014). BEAST analyses were run on the CIPRES server (Miller et al., 2010), for 100 million generations, sampled every 10000th generations. Analysis performance and effective sample size (ESS) values were evaluated using Tracer 1.7 (Rambaut et al., 2014). After a 25% burn-in, maximum clade credibility and 95% confidence interval were calculated in TreeAnnotator (BEAST package). Final chronogram was visualized and edited in FigTree 1.3 (Rambaut, 2009).

In preliminary analyses, BEAST could not recover clades that we found in our phylogenetic analysis in RaxML and MrBayes, also repeatedly recovered as well-supported in previous studies (Bossert et al.,

2019; Cardinal et al., 2010; Martins et al., 2014), thus these were constrained as monophyletic in the present divergence time estimation in BEAST: Ceratinini + Allodapini, Xylocopine + Eucerine lines, Apine + Xylocopine + Eucerine lines. We also applied an age constraint at the root of the tree based on the age previously estimated for this node (101.26; 95% HPD: 98.81–111.03 ; (Martins et al., 2014), using a normal prior distribution (Mean 105; Stdev: 4; HPD: 95% 97–112). As there is no fossil attributed to the Tapinotaspidini and their relatives in the eucerine line, we chose reliable fossils from other lines in Apinae: one *Apis* from the Miocene, *Boreallodape* from the Eocene and *Paleohabropoda* from the Paleocene. The fossil *Apis lithohermaea* from the middle Miocene of Chojabaru Formation, Japan, is the first fossil assigned to the crown-group of *Apis*, in the *dorsata* group, by its enlarged body size, unique to this group and is probably from the Langhian age (13.8–16 Mya) (Engel, 2006). We placed this fossil at the node uniting *A. dorsata* and *A. cerana* applying a lognormal prior (offset = 15; M = 1.26; S = 0.8). The monotypic fossil tribe Boreallodapini from the Baltic amber, comprises three species of *Boreallodape*, and considered closely related to the extant Ceratinini and Allodapini (Engel, 2001). The three tribes share at least three synapomorphies and phylogenetic analysis including Boreallodapini fossils indicated the sister relationships between Boreallodapini and Allodapini (Engel 2001). For this reason we placed this fossil at the crown of Ceratinini + Allodapini applying a lognormal prior (offset = 44; M = 1.0, S = 1.1). *Paleohabropoda oudardi* Michez & Rasmont from the Palaeocene of Menat (Puy-de-Dôme, France), one of the oldest bee fossils clearly belongs to the tribe Anthophorini as demonstrated by several morphological characters (Michez et al., 2009). This fossil has been used to calibrate the node uniting all Anthophorini (Cardinal et al., 2010; Cardinal and Danforth, 2013) by its possible association to the extant *Habropoda*. But since the relationship between *Habropoda* and *Paleohabropoda* is given by a relatively weak character (wing shape), it is uncertain if it really belongs to the stem of this tribe, thus we applied a normal distribution prior (Mean = 60; Sigma = 5). All calibration points and constraints are depicted in Figure S3.

2.4. Ancestral range estimation

For estimation of ancestral range probabilities we used a multi-model approach in BioGeoBEARS package in R (Matzke, 2013a; R Core Team, 2016), using the three available models: DEC (Ree and Smith, 2008), the maximum likelihood interpretation of BayArea (Landis et al., 2013), namely BAYAREALIKE, and the maximum likelihood implementation of DIVA (Ronquist, 1997), namely DIVALIKE. The founder speciation event parameter *j* was also added to all analyses, creating the models DEC + *j*, DIVALIKE + *j*, BAYAREALIKE + *j*, adding the possibility of a new cladogenesis event, where an individual 'jumps' to an area completely outside the ancestral range, founding a new genetically isolated lineage (Matzke, 2014; Matzke, 2013b). As input we used a modified version of the maximum clade credibility tree produced in TreeAnnotator in which we excluded most outgroup taxa and kept only Tapinotaspidini plus Emphorini representatives (i.e. the closest outgroup) using the function *drop.tip* in the R package ape (Paradis and Schliep, 2018). We assessed the overall fitness of the models conducting likelihood ratio tests and evaluating AICc scores.

We specified nine biogeographical regions based on the neotropical regionalization system proposed by Morrone (2014), modified to encompass the geographical distribution of Tapinotaspidini bees: (A) Atlantic Forest, (B) Chaco, (C) Cerrado, (D) Caatinga, (E) Pampa, (F) Amazon, (G) Central America, (H) Andean Forests and (I) Monte/Chile. For ingroup species only areas of occurrence of terminal taxa were included. We were not able to include specimens occurring in the Caribbean (one species of *Xanthopedia*), thus this region is not represented in our analysis. For the outgroup Emphorini we chose a groundplan approach in which all terminal taxa were assigned to the same area codification, including all open vegetation areas (Chaco, Cerrado,

Caatinga, Pampa), since current evidence indicates that the few members found in humid forested environments are the result of recent biogeographical events.

2.5. Ancestral state reconstruction

To assess shifts in floral host plant and habitat across the evolutionary history of Tapinotaspidini bees, we performed analyses of ancestral state reconstruction. The analyses were performed in a maximum likelihood implementation using the function *ace* in the R package ape (Paradis and Schliep, 2018). Based on the oil foraging preferences of our sampled species, we coded them in three categories: (0) Malpighiaceae (for species that present a strong preference for Malpighiaceae); (1) Malpighiaceae plus other oil producing plant families (for species that show no preference for a particular oil producing plant family) and (2) only other oil producing plant families (for species that do not visit Malpighiaceae but instead visit one or more other oil host families). Information about which oil host plant is visited by each species was retrieved from the literature (Aguiar 2012, Carneiro et al. 2015, Cocucci 1991, Cocucci and Vogel 2000, Martins and Alves-dos-Santos, 2013, Renner and Schaefer, 2010, Vogel 1974) and from personal observations in the field (AJC Aguiar and GAR Melo, pers. obs.) and are presented in Table S7. For analyses of habitat, we based the categories on the same area matrix used in the BioGeoBEARS analysis (see above). In this case, however, we merged areas A = Atlantic Forest, F = Amazon, G = Central America and H = Andean forests into a single "forested" category and areas B = Chaco, C = Cerrado, D = Caatinga, E = Pampa and I = Monte/Chile into a single "open" category, creating a matrix of two trait states: (1) "open vegetation"; and (2) "forested vegetation".

2.6. Trait dependent diversification analyses

We explored the heterogeneity in diversification rates through time and across the evolution of habitat (open or forested vegetation) and oil producing host plants, using the same character matrices generated for the ancestral states reconstruction analyses (see above). We tested if shifts into new habitats significantly changed diversification rates in Tapinotaspidini bees, using Binary State Speciation and Extinction models (BiSSE, Fitzjohn, 2012). BiSSE models simultaneously analyze rates of character state change along with speciation and extinction in each state. Likewise, to explore if shifts in oil producing host plant families affected diversification rates in Tapinotaspidini bees we relied on Multi-State Speciation and Extinction models (MuSSE, Fitzjohn, 2012), a similar framework as BiSSE but that allows analyses of multistate characters. Both analyses were performed using functions of the R package diversitree (Fitzjohn, 2012). In BiSSE, a sampling fraction of 0.45 (for species in open vegetation) and 0.19 (for species in forested vegetation) was incorporated. In MuSSE, a sampling fraction of 0.27 (for species that visit exclusively Malpighiaceae), 0.53 (for species that visit Malpighiaceae and other plant families) and 0.37 (for species that visit only other plant families) was used. These sampling fractions were incorporated to smooth the effect of sampling bias across different character states in both analyses. To calculate sampling fractions we use total valid species per lineage.

In applications of both BiSSE and MuSSE models, we first tested the fit of allowing different constraints on the parameters (i.e. allowing variation in both *lambda* and *mu* or having them both fixed). The best AIC score for both BiSSE and MuSSE was a model where extinction, speciation and transition rates are all equal (i.e. lambda0 ~ lambda1 and mu0 ~ mu1 for the BiSSE model and lambda1 ~ lambda2 ~ lambda3 and mu1 ~ mu2 ~ mu3 for the MuSSE model; with all transition probabilities in the q matrix fixed) (see Supplementary Tables S5 and S6 for AIC comparisons). Nevertheless, we used both models in their "full" versions (i.e. allowing parameters *lambda* and *mu* to vary between trait states) to set priors for each

parameter in the Bayesian BiSSE and MuSSE analyses. MCMC were run for parameter estimations in both models (10.000 steps for BiSSE, 20.000 steps for MuSSE) and ESS > 200 were confirmed in both analyses. When no significant changes in speciation or extinction associated to traits states were recovered, we did not include tests of diversification associated to hidden states (i.e. HiSSE) or character independent diversification ("CID") advocated by Beaulieu and O'Meara (2016). These are mainly indicated to avoid type I errors when a trait state significantly change diversification in a group.

3. Results

3.1. Phylogenetic relationships and divergence time estimates

Our alignment resulted in a concatenated matrix of 3997 characters for the five amplified regions (28S, CO1, EF, Opsin, Wg), available at Dryad. Fig. 1 and Supplementary Figures S1 and S2 show respectively the chronogram and the phylogenetic trees. We demonstrated the monophyly of the eucerine line (Bootstrap support (BS) = 99; Posterior Probability (PP) = 1) with the following internal relationships: (*Ancyloscelis* + Exomalopsini) ((Ancylaini + Eucerini) (Emphorini + Tapinotaspidini)). This lineage originated in the Cretaceous, *circa* 76 Ma (million years ago) (95% Highest posterior density (HPD): 65–86) (Fig. 1). Maximum likelihood and Bayesian inference approaches resulted both in a highly supported sister relationship between Tapinotaspidini and Emphorini (BS = 87, PP = 97), originating in the beginning of the Cenozoic era, around 62 Ma (HPD = 53–72).

We recovered Tapinotaspidini as monophyletic in all analyses, although with moderate support (BS = 72, PP = 0.91). The group was estimated to have originated in the late Paleocene, around 57 Ma (HPD = 47–66). Our analysis strongly supports all genera as monophyletic, except *Chalepogenus, Tropidopedia* and *Paratetrapedia*. *Chalepogenus* comprises two different clades in our phylogeny: the first, *Chalepogenus* sensu stricto, contains the type species *Chalepogenus muelleri* and most species occurring in the Chaco and Monte regions, as well as in Chile; the second contains four species related to *Lanthanomelissa* and occurs in the Pampa vegetation. *Tropidopedia* is not monophyletic due to the position of *T. nigrita*, sister to the clade containing the remaining species of *Tropidopedia* and other genera. The position of *Paratetrapedia rutilipes* as sister to *Lophopedia* renders the speciose genus *Paratetrapedia* polyphyletic.

Monoeca appears as sister to all extant Tapinotaspidini, followed by *Caenonomada* and *Tapinotaspis* + *Tapinotaspoides* consecutively. The remaining genera group in a large clade, where *Chalepogenus* s.s is sister to a group formed by *Trigonopedia* and the *Arhysoceble* + *Lanthanomelissa* lineage (including the remaining *Chalepogenus*) plus the *Paratetrapedia* lineage (sensu Michener and Moure, 1957). The latter assemblage, composed by *Urbanapis, Xanthopedia, Paratetrapedia, Nasutopedia, Tropidopedia* and *Lophopedia*, is strongly supported as monophyletic in all analyses (BS = 100, PP = 1). This lineage arose in the Oligocene at 30 Ma (HPD = 24–35). Crown node ages for all genera are younger than 20 Ma and most are younger than 10 Ma, i.e. all genera are estimated to have a Miocene origin.

3.2. Ancestral range estimation and diversification associated to shifts into forested areas

The model DIVALIKE + J yielded the highest likelihood and best AICc scores for ancestral range estimation for the Tapinotaspidini phylogeny (see Table S4 for statistics). The most recent common ancestor (MRCA) of Tapinotaspidini had the Cerrado as its most probable primeval habitat range, followed by other open vegetation habitats such as Chaco and Pampa. All the earliest nodes of Tapinotaspidini, including the stem nodes of *Monoeca, Caenonomada, Chalepogenus* s.s., *Lanthanomelissa* + *Arhysoceble*, had either Cerrado, Chaco or Pampa in their ancestral range distribution.
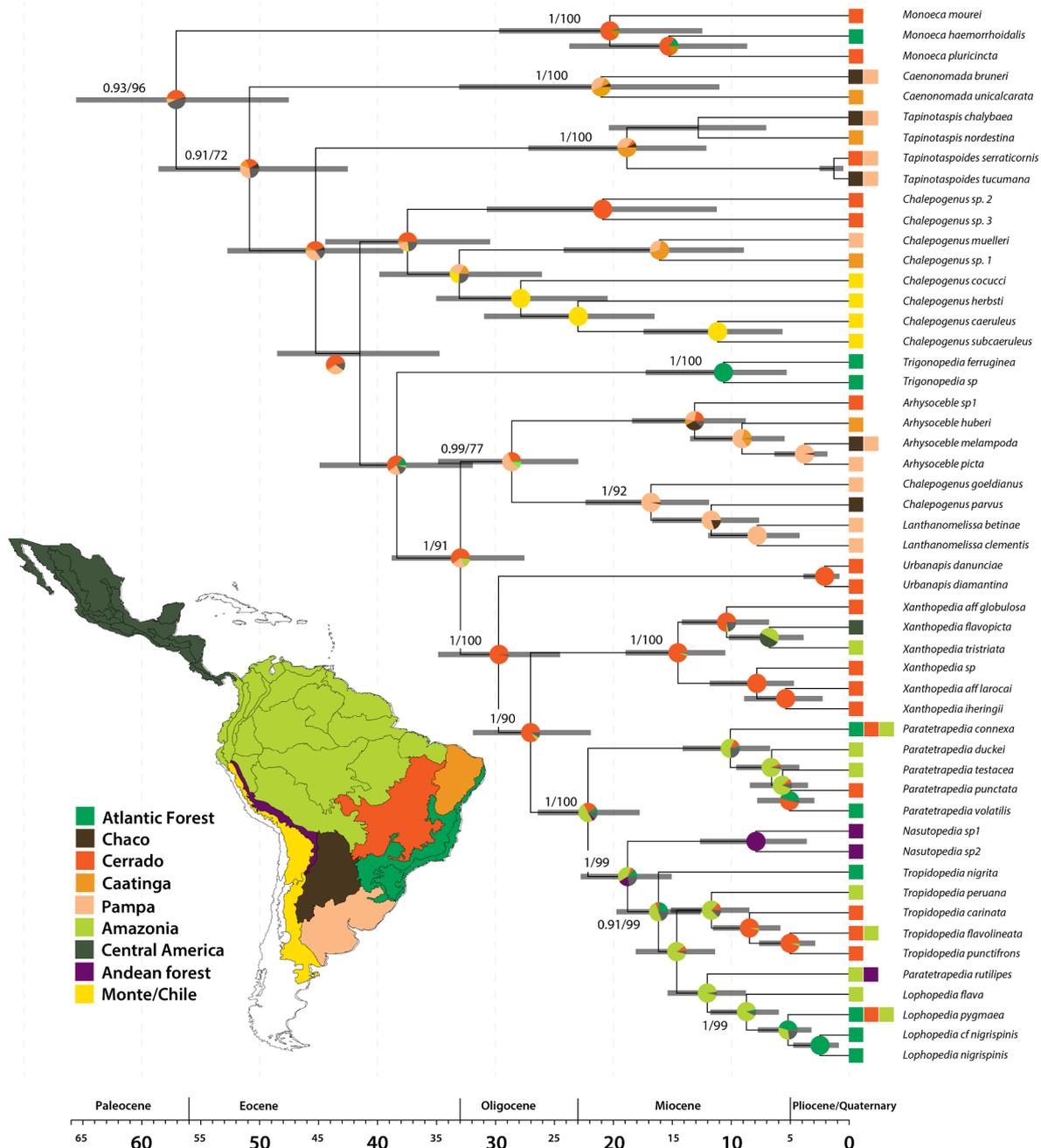
Fig. 1. Time-calibrated tree and ancestral range estimation. The maximum clade credibility tree resulted from divergence time estimation in BEAST obtained with the matrix of 84 taxa and 3997 nucleotides for Tapinotaspidini bees and outgroups of the subfamily Apinae. Bars represent the 95% Highest Posterior density interval. Pie charts on nodes represent ancestral range estimation under the highest likelihood model DIVALIKE + J analyzed in BioGeoBEARS package. Biogeographic areas represent the geographical distribution range of living species: (A) Atlantic Forest, (B) Chaco, (C) Cerrado (D) Caatinga (E) Pampa, (F) Amazon, (G) Central America, (H) Andean Forests and (I) Monte.

Ancestral state estimation indicated open vegetation areas including savannas as the most probable vegetation type occupied by Tapinotaspidini bees. Even though a strong preference for open vegetation was recovered in the early diversification of the Tapinotaspidini, at least four shifts to forested areas are inferred if Amazon and Brazilian Atlantic forest are lumped together in a single forested category. Shifts to forested areas first occurred in the MRCA of *Paratetrapedia* and relatives, which are estimated to have originated in the Amazon (crown age about 30 Ma) and in the MRCA of *Trigonopedia* whose origin is estimated to be the Brazilian Atlantic forest (crown age about 10 Ma). These multiple shifts appear to increase diversification rates in lineages that have forests as their habitat (Fig. 2A). The BiSSE model was shown

to fit best than the HiSSE and CID ones (AICcs: BiSSE: 448.8433; CID: 463.1021; BiSSE: 482.3197), and it shows that extinction rates are similar for lineages in both types of vegetation, but lineages in forested areas have slightly higher speciation rates associated to this trait state. This agrees with our original expectation that higher speciation would have occurred in forested groups, which include *Paratetrapedia* the most speciose genus in the tribe, with 33 valid species.

3.3. Changes in evolutionary dynamics associated to host plant families

The results from the analyses of ancestral character reconstruction yielded Malpighiaceae as the most probable oil producing family
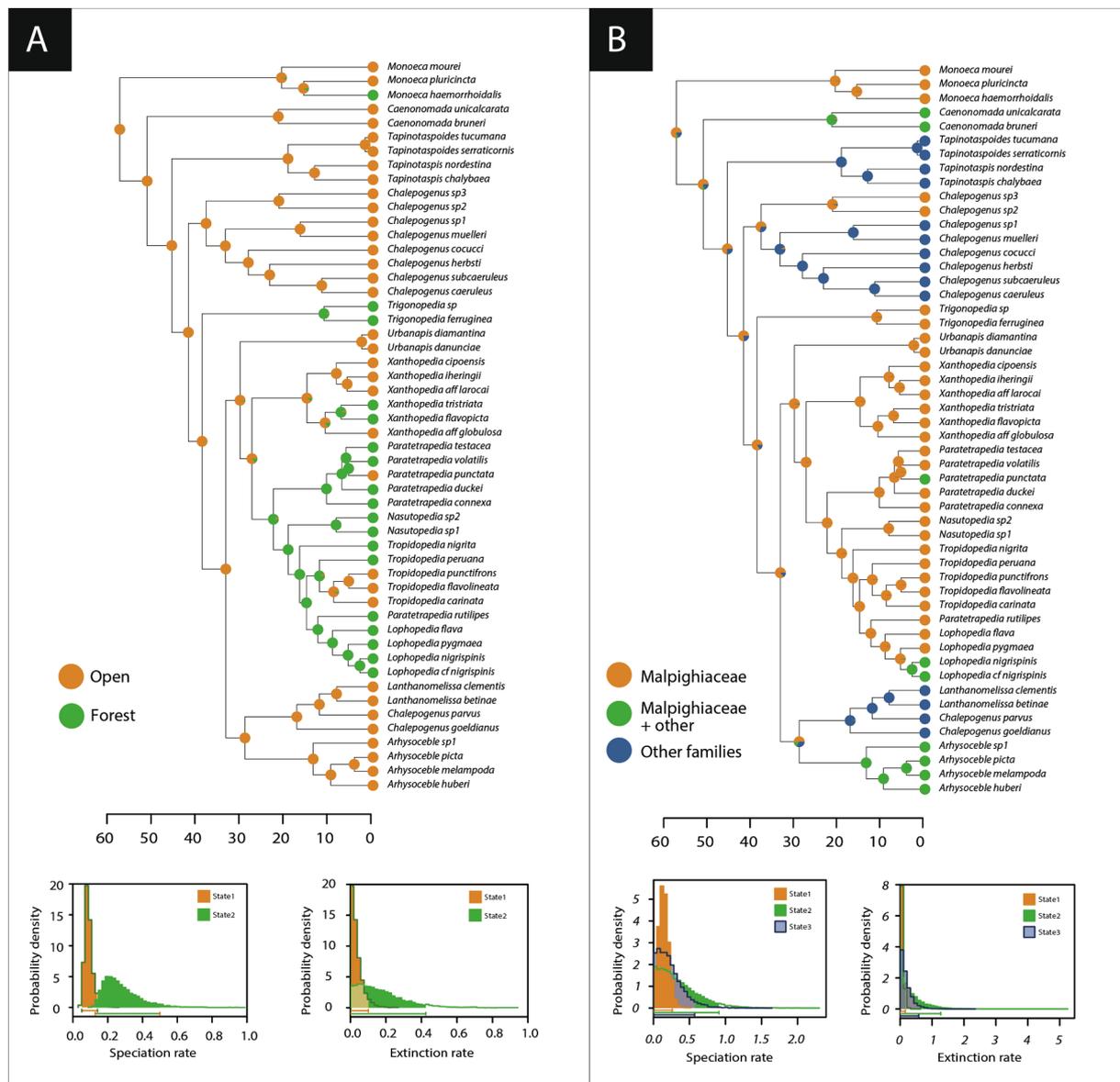
Fig. 2. Analyses of ancestral state reconstruction and state dependent speciation and extinction for habitat and floral host plant across time. X-axis are in scale of millions of years before present. A. ACE and BiSSE results regarding habitat occupied: open and forested, showing speciation and extinction rates associated to each character state. B. ACE and MuSSE results regarding floral host categories: Mapighiaceae, Malpighiaceae plus other families and only other families.

ancestor for Tapinotaspidini (Fig. 2b). Shifts from Malpighiaceae to other plant families occurred in *Chalepogenus*, *Tapinotaspoides* and *Lanthanomelissa*, which today harvest for oil in Iridaceae, Solanaceae, Calceolariaceae, and Plantaginaceae. Host broadening, i.e. inclusion of other oil host families occurred in *Caenonomada*, *Arhysoceble*, *Paratetrapedia* and *Lophopedia*, which today collects oil on Malpighiaceae, Plantaginaceae, Krameriaceae, Iridaceae, Solanaceae. Some lineages, for example *Monoeca*, *Xanthopedia* and *Tropidopedia*, continued to forage exclusively on malpigh species flowers for oil. MuSSE analysis showed that speciation and extinction rates are similar for all host categories (Fig. 2B).

4. Discussion

4.1. The first molecular phylogeny of tapinotaspidine bees confirms the monophyly of the tribe

We reconstructed for the first time a broadly sampled molecular phylogeny of Tapinotaspidini bees and used biogeographic and

macroevolutionary analyses to build a scenario of evolution in time and space for this Neotropical group. We confirmed the monophyly of the tribe, corroborating previous hypotheses based on morphology (Roig-Alsina and Michener, 1993), and its sister relationship to Emphorini, both nested in the eucerine line (Bossert et al., 2019; Cardinal et al., 2010; Cardinal and Danforth, 2013). Our results also corroborate the relationships among the tribes that compose the eucerine line – Ancylaini, Eucerini, Emphorini, Exomalopsini and Tapinotaspidini (Bossert et al., 2019), including the isolated position of the Neotropical group *Ancyloscelis* (usually placed in Emphorini) within the eucerine line. The tribe name Ancyloscelidini has been already proposed by Roig-Alsina and Michener (1993), and indeed deserves to be recognized again, in order to remove the paraphyletic status of Emphorini under its traditional composition.

The position of *Monoeca* and *Caenonomada* as consecutive sister lineages to all remaining Tapinotaspidini supports the morphological evidence indicating them as the early branching genera of the tribe (Roig-Alsina 1997). Our results, however, were mostly not congruent with the topologies obtained by Roig-Alsina (1997) in his

morphological phylogenetic study of the tribe. Most of the genera appear as monophyletic, except for *Chalepogenus*, *Tropidopedia* and *Paratetrapedia*. The position of *Tropidopedia nigrita* as sister to *Tropidopedia*, *Lophopedia* and *P. rutilipes* renders *Tropidopedia* paraphyletic (see also Aguiar and Melo, 2007). Similarly, the position of *Paratetrapedia rutilipes* sister to *Lophopedia* renders the highly diverse *Paratetrapedia* paraphyletic (Aguiar and Melo, 2011). And *Chalepogenus* is composed by at least two lineages, one related to *Arhysoceble* + *Lanthanomelissa* and another related to *Trigonopedia* plus the more derived lineages of *Tapinotaspidini*. The paraphyletic nature of *Chalepogenus* indicate this as the most important target of taxonomic revision among the tapinotaspidine bees.

4.2. Biogeography of Tapinotaspidini indicates the presence of ancient savannas in South America

The ancestral range estimation indicated Cerrado-like savannas as the most probable ancestral range of Tapinotaspidini bees in the Paleocene of South America. These bees continued to diversify in open vegetation environments through the Eocene and Oligocene. The association of the ancestor of Tapinotaspidini with savannas is reinforced by the present-day distribution of their sister group, the tribe Emphorini, and the whole eucerine line. Most of the ancient lineages in the tribes composing the eucerine line are distributed in open vegetation biomes, with the groups associated with humid habitats representing more recent clades, as detected here for the Tapinotaspidini. Their ground nesting behavior is probably directly linked to this habitat preference, allied to floral resource utilization. The present study brings new evidence postulating presence of savanna-like biomes in South America tracing back to Paleocene.

The presence of savannas in the Paleocene of South America contrasts with the warm climate and wide distribution of tropical forests at that time (Hoorn et al., 2010; Zachos et al., 2001). Fossil evidence corroborates the dominance of tropical forests in South America at least until the end of the Miocene (Hoorn et al., 2010; Hughes et al., 2013; Ortiz-Jaureguizar and Cladera, 2006; Romero, 1993). Accompanying the global climatic cooling that begins by the end of the Mid-Miocene Climatic Optimum open vegetation biomes expanded in many parts of the world including South America, a process that was facilitated by the expansions of C4 grasses until the dominance of open vegetation environments in Pliocene (Iglesias et al., 2011; Linder et al., 2018; McElwain, 2018; Simon et al., 2009; Pennington and Hughes, 2014).

Paleopalinological evidence shows that the first plant elements characteristic of savannas or sclerophyllous forests (such as Melastomataceae (*Heterocolpites palaeocenica*) and Poaceae pollen) were already present in the Paleocene of South America, although they became common only in the Eocene (Burnham and Johnson, 2014; Romero, 1993). Around the Eocene climatic optimum (52 Mya) the warm climate favored the tropical elements to reach their southernmost distribution (Wilf et al., 2003). In late Eocene floristic elements typical of Chaco (i.e. *Schinopsis*, Anacardiaceae) become conspicuous components in southern South America beginning a trend of increasing presence of elements associated with savanna and Chaco accompanying climate cooling (Ortiz-Jaureguizar and Cladera, 2006). In the fossil record from the Oligocene many floristic elements that dominate the present dry diagonal of open vegetations were already present, *e.g.* Anacardiaceae, Leguminosae, Malvaceae, Malpighiaceae (Romero, 1993). The mammal fossil fauna corroborates the floristic evidence that suggests early scattered elements of open vegetations in southernmost parts of South America in the Paleocene, indicated by the appearance of the first prothypsodont mammals, whose morphology adapted to feed on grasses will make them dominant only in the Oligocene (Ortiz-Jaureguizar and Cladera, 2006).

In other words, even though the ancestral range estimation shows a likely early diversification in the savannas for the Tapinotaspidini bees, due to the climatic conditions at that time, their actual ancestral range

may have been in areas further south. It is important to highlight that those environments were not necessarily homologous to the biomes that today dominate the open diagonal in South America, but some floristic elements may have been similar.

4.3. Transitions to forested areas out of open vegetation in the Eocene

Thirty million years after the estimated origin of the ancestor of the tapinotaspidine bees, the first incursion into tropical rainforests took place, represented here by the ancestral range estimate reconstructed for the *Paratetrapedia* lineage (i.e. *Paratetrapedia* and closely related genera), whose members are mostly associated with the Amazon forests. Brazilian Atlantic rainforest was further colonized, at least five times independently in the genera *Monoeca*, *Trigonopedia*, *Xanthopedia*, *Paratetrapedia*, and *Lophopedia*. Reversals to open vegetation biomes, specially to Cerrado, occurred within the genera *Tropidopedia* and *Paratetrapedia* only very recently (ca. 5 ma.).

As storing perishable food for larvae (mix of pollen and nectar) in ground nests of tropical forests is challenging, groups that succeed in these humid environments present adaptations to cope with these unfavorable conditions. In the case of Tapinotaspidini, the most diverse forest-inhabiting clade, the *Paratetrapedia* lineage, differs from the extant members of the group in their nesting biology: they use cavities in wood as substrate in opposition to the ancestral ground nesting behavior (Aguiar et al., 2004). In other oil-collecting groups, like *Centris* and *Epicharis*, the use of floral oils of Malpighiaceae allowed them to efficiently waterproof their nest cells, thus making them successful inhabitants of humid forests, where they probably originated (Martins and Melo, 2016). While there is little or no evidence of the use of floral oils in nest cell lining in ground nesting tapinotaspinine bees (Rozen et al. 2006), females of at least some genera of the *Paratetrapedia* lineage, such as *Paratetrapedia* and *Lophopedia*, line the cell walls with a varnish-like substance that seems to be hardened floral oils (GAR Melo and AJC Aguiar, unpubl. data). The habit of nidification in wood was also accompanied by a key morphological trait, a highly modified pygidial plate. Females in the *Paratetrapedia* lineage exhibit a pygidial plate composed of a broad basal portion followed by a narrow, spatulate apex (see drawings in Silveira et al. 2002 and microphotographs in Aguiar and Melo 2007, 2011). These peculiar plates might have evolved as an adaptation to nesting in wood and use of oil in cell lining. Given the general preference for open areas by most groups of bees, the possibility of colonizing forested vegetation may have led to increased speciation rates in some clades of the Tapinotaspidini, as these were able to explore a possibly empty niche.

Biotic exchanges occurred vastly among South American biomes, predominantly from forested/wet regions to open/dry habitats (Antonelli et al., 2018b). Apparently Cerrado savanna has been porous to the recruitment of species from adjacent biomes, from humid forests to subtropical grasslands and seasonally dry tropical forests, contributing to the striking species richness observed today in this region (Simon et al. 2009). The interchange between forested regions through adjacent open vegetation biomes is facilitated by the presence of riverine forests (Oliveira-Filho and Ratter, 1995). Many animal and plant groups present a disjunct pattern between forested areas in South America caused by the in-between presence of open areas (Costa, 2003; Fiaschi and Pirani, 2009). At the same time, several other organisms are distributed mainly in open vegetation biomes avoiding the forested humid environments (Plana et al., 2004; Werneck et al., 2012). Nevertheless a few groups, in addition of the tapinotaspidine bees, offer us the opportunity to understand the colonization of both basic kinds of vegetation at different times and involving adaptations to do so, like the plant genera *Myrcia* (Vasconcelos et al., 2019) and *Galeandra* (Martins et al., 2018). In the case of *Galeandra*, shift from forested to open vegetation areas was accompanied by both habit change (epiphytic to terrestrial) and pollination mode (from long tongued euglossini bees to other bee groups or even autopollination) (Martins et al. 2018). For

both plant groups, forest adapted species use riverine forests (also called gallery forests) as corridors to interchange among forested biomes (Vasconcelos et al. 2019, Martins et al. 2018).

4.4. Shifts in host plant did not increase diversity, but were possibly required to change habitat

Ancestral state estimates indicated Malpighiaceae as the most probable ancestral host for floral oil but shifts to other hosts or host broadening occurred after the Eocene. This is in agreement with the antiquity of Malpighiaceae as the first angiosperm family to acquire floral oil glands and the much younger origin of other plant families with a similar syndrome (Martins et al., 2015; Renner and Schaefer, 2010; Vogel, 1974). Malpighiaceae originated in the Paleocene and diversified early in South America (Davis et al., 2002; Xi et al., 2012). *Monoeca*, the sister group of the remaining Tapinotaspidini, is the only legitimate visitor of Malpighiaceae, i.e. they are capable of collecting oil while pollinating the flowers in the same way as the main malpigh pollinators, the Centridini and Epicharitini bees (Sigrist and Sazima, 2004). Other Tapinotaspidini bees are in general not big enough to access the floral oil glands of Malipighiaceae while simultaneously touching fertile parts. On the other hand they are the main pollinators of oil-producing Iridaceae, Solanaceae and possible many orchids (see below).

Present distribution of Malpighiaceae shows a strong preference for forested habitats and savannas, with low presence or complete absence from arid and semiarid regions (Anderson, 1979). Thus, the shifts from open habitats to forested ones during the evolution of tapinotaspidine bees seem to have been accompanied by shifts of most preferred host plant family. Changes from Malpighiaceae to other families are not necessarily advantageous in the sense of promoting diversification but may have been necessary for the Tapinotaspidini bees to explore xeric areas, such as Caatinga and Monte, where this family is rare or absent.

Thus exploitation of other oil sources were important in occupying new habitats, such as the Andean region where *Chalepogenus* exploits Calceolariaceae (Sérsic, 2004), the Pampas where *Lanthanomelissa* visits *Sisyrinchium* (Iridaceae) (Cocucci and Vogel, 2000) and *Tapinotaspis chalybaea* explores *Nierembergia* (Solanaceae) (Cocucci, 1991). Most Tapinotaspidini genera originated in the early Miocene and diversified in the late Miocene. This age estimates are congruent with the origin of distinct oil host plant families like Iridaceae, Plantaginaceae and Krameriaceae, which today are well stablished in the Dry Diagonal of Open vegetation (Martins and Alves-dos-Santos, 2013; Simpson, 1989). The same pattern was detected for *Centris* bees which also broadened their oil plant host spectrum during the Miocene (Martins et al. 2015). In fact, habitat and plant host shifts cannot be totally disentangled since oil producing plant families differ in habitat preferences and habitat shifts in bees are frequently accompanied by oil host shifts.

A first comprehensive phylogenetic hypothesis for this diverse tribe of Neotropical bees was here presented offering a solid ground to further taxonomic revisions and classification. Moreover, we infer the Paleocene origin of these bees possibly associated to open vegetation habits in South America, which, due to warmer climate at that time, were encountered further south than today. The several habitat shifts from open vegetation to forests increased diversification while requiring nesting habitat shifts. Host plant families shifts did not increase diversification on the other hand, but were probably necessary to co-lonization of different habits, since oil producing families are not evenly distributed.

Data availability statement

Genetic sequence data are available in GenBank with the accession numbers included in Table S1 together with the voucher information. Matrices and trees will be available upon request.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106692.

References

Aguiar, A.J.C., 2012. Tapinotaspidini. Catalogue of Bees (Hymenoptera, Apoidea) in the Neotropical Region - online version. URL http://moure.cria.org.br/catalogue?id=28975.

Aguiar, A.J.C., Melo, G.A.R., 2007. Taxonomic revision, phylogenetic analysis, and biogeography of the bee genus *Tropidopedia* (Hymenoptera, Apidae, Tapinotaspidini). Zool. J. Linn. Soc. 151, 511–554.

Aguiar, A.J.C., Melo, G.A.R., 2011. Revision and phylogeny of the bee genus *Paratetrapedia* Moure, with description of a new genus from the Andean Cordillera (Hymenoptera, Apidae, Tapinotaspidini). Zool. J. Linn. Soc. 162, 351–442. https://doi.org/10.1111/j.1096-3642.2010.00678.x.

Aguiar, A.J.C., Melo, G.A.R., Rozen, J.G., Alves-dos-Santos, I., 2004. Synopsis of the nesting biology of Tapinotaspidini bees (Apidae: Apinae). In: in: 8th IBRA International Conference on Tropical Bees and VI Encontro Sobre Abelhas, pp. 80–85.

Almeida, E.A.B., Danforth, B.N., 2009. Phylogeny of colletid bees (Hymenoptera: Colletidae) inferred from four nuclear genes. Mol. Phylogenet. Evol. 50, 290–309. https://doi.org/10.1016/j.ympev.2008.09.028.

Anderson, W.R., 1979. Floral Conservatism in Neotropical Malpighiaceae. Biotropica 11, 219–223.

Antonelli, A., Ariza, M., Albert, J.S., Andermann, T., Azevedo, J., Bacon, C.D., 2018a. Conceptual and empirical advances in Neotropical biodiversity research 1–53. https://doi.org/10.7717/peerj.5644.

Antonelli, A., Sanmartín, I., 2011. Why are there so many plant species in the Neotropics? Taxon 60, 403–414. https://doi.org/10.2307/41317138.

Antonelli, A., Zizka, A., Carvalho, F.A., Scharn, R., Bacon, C.D., Silvestro, D., Condamine, F.L., 2018b. Amazonia is the primary source of Neotropical biodiversity. Proc. Natl. Acad. Sci. USA. https://doi.org/10.1073/pnas.1713819115.

Beaulieu, J.M., O'Meara, B.C., 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. Syst. Biol. 65, 583–601. https://doi.org/10.1093/sysbio/syw022.

Biomatters, 2015. Geneious 8.1.

Bossert, S., Murray, E.A., Almeida, E.A.B., Brady, S.G., Blaimer, B.B., Danforth, B.N., 2019. Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Mol. Phylogenet. Evol. 130, 121–131. https://doi.org/10.1016/J.YMPEV.2018.10.012.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H.C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian

Evolutionary Analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537.

Buchmann, S.L., 1987. The ecology of oil flowers and their bees. Annu. Rev. Ecol. Syst. 18, 343–369.

Burnham, R., Johnson, K.R., 2014. South American paleobotany and the origins of Neotropical rainforests South American palaeobotany and the origins of neotropical rainforests. Philos. Trans. R. Soc. B Biol. Sci. 359, 1595–1610. https://doi.org/10.1098/rstb.2004.1531.

Buschini, M.L.T., 2006. Species diversity and community structure in trap-nesting bees in Southern Brazil. Apidologie 37, 58–66. https://doi.org/10.1051/apido:2005059.

Camargo, J.M.F., Pedro, S.R.M., 2003. Meliponini neotropicais: o gênero Partamona Schwarz, 1939 (Hymenoptera, Apidae, Apinae) - bionomia e biogeografia. Rev. Bras. Entomol. 47, 311–372.

Camargo, J.M.F., Vit, P., 2013. Historical Biogeography of the Meliponini (Hymenoptera, Apidae, Apinae) of the Neotropical Region. In: Vit, P., Pedro, S.R.M., Roubik, D.W. (Eds.), Pot-Honey: A Legacy of Stingless Bees. Springer International Publishing, pp. 19–35. https://doi.org/10.1007/978-1-4614-4960-7.

Cappellari, S.C., Melo, G.A.R., Aguiar, A.J.C., Neff, J.L., 2011. Floral oil collection by male Tetrapedia bees (Hymenoptera: Apidae: Tetrapediini). Apidologie 43, 39–50. https://doi.org/10.1007/s13592-011-0072-2.

Cardinal, S., Danforth, B.N., 2013. Bees diversified in the age of eudicots. Proc. R. Soc. B 280, 1–9.

Cardinal, S., Straka, J., Danforth, B.N., 2010. Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. Proc. Natl. Acad. Sci. USA 107, 16207–16211. https://doi.org/10.1073/pnas.1006299107.

Carneiro, L.T., Aguiar, A.J.C., Martins, C.F., Machado, I.C., Alves-dos-Santos, I., 2015. Krameria tomentosa oil flowers and their pollinators: bees specialized on trichome elaiophores exploit its epithelial oil glands. Flora - Morphol. Distrib. Funct. Ecol. Plants 215, 1–8. https://doi.org/10.1016/j.flora.2015.06.002.

Carrijo, T.F., Gonçalves, R.B., Santos, R.G., 2012. Review of bees as guests in termite nests, with a new record of the communal bee, Gaesochira obscura (Smith, 1879) (Hymenoptera, Apidae), in nests of Anoplotermes banksi Emerson, 1925 (Isoptera, Termitidae, Apicotermitinae). Insectes Soc. 59, 141–149. https://doi.org/10.1007/s00040-012-0218-x.

Cocucci, A.A., 1991. Pollination biology of Nierembergia (Solanaceae). Plant Syst. Evol. 174, 17–35.

Cocucci, A.A., Vogel, S., 2000. Oil-producing flowers of Sisyrinchium species (Iridaceae) and their pollinators in southern South America. Flora 196, 26–46.

Condamine, F.L., Rolland, J., Höhna, S., Sperling, F.A.H., Sanmartín, I., 2018. Testing the role of the red queen and court jester as drivers of the macroevolution of apollo butterflies. Syst. Biol. 67, 940–964. https://doi.org/10.1093/sysbio/syy009.

Costa, L.P., 2003. The historical bridge between the Amazon and the Atlantic forest of Brazil a study of molecular phylogeography with small mammals. J. Biogeogr. 30, 71–86. https://doi.org/10.1046/j.1365-2699.2003.00792.x.

Danforth, B.N., Cardinal, S., Praz, C., Almeida, E.A.B., Michez, D., 2013. The impact of molecular data on our understanding of bee phylogeny and evolution. Annu. Rev. Entomol. 58, 57–78. https://doi.org/10.1146/annurev-ento-120811-153633.

Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA 99, 6833–6837. https://doi.org/10.1073/pnas.102175899.

Engel, M.S., 2006. A giant honey bee from the middle Miocene of Japan (Hymenoptera: Apidae). Am. Museum Novit. 3504, 1–12.

Engel, M.S., 2001. A monograph of the Baltic amber bees and evolution of the Apoidea (Hymenoptera). Bull. Am. Museum Nat. Hist. 259, 1–192. https://doi.org/10.1206/0003-0090(2001) 259<0001:AMOTBA>2.0.

Fiaschi, P., Pirani, J.R., 2009. Review of plant biogeographic studies in Brazil. J. Syst. Evol. 47, 477–496. https://doi.org/10.1111/j.1759-6831.2009.00046.x.

Fitzjohn, R.G., 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3, 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x.

Gerling, D., Velthuis, H.H.W., Hefetz, A., 1989. Carpenter Bees of the Genus Xylocopa. Annual Rev. Entomol. 34, 163–190.

Hoorn, C., Wesselingh, F.P.P., ter Steege, H., Bermudez, M.A.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T., Antonelli, A., Sanmartin, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Sarkinen, T., Antonelli, A., 2010. Amazonia Through Time : Andean uplift, climate change, landscape evolution, and biodiversity. Science 330, 927–931. https://doi.org/10.1126/science.1194585.

Hughes, C.E., Pennington, R.T., Antonelli, A., 2013. Neotropical plant evolution: assembling the big picture. Bot. J. Linn. Soc. 171, 1–18. https://doi.org/10.1111/boj.12006.

Iglesias, A., Artabe, A.E., Morel, E.M., 2011. The evolution of Patagonian climate and vegetation from the Mesozoic to the present. Biol. J. Linn. Soc. 103, 409–422. https://doi.org/10.1111/j.1095-8312.2011.01657.x.

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

Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804. https://doi.org/10.1093/sysbio/syt040.

Linder, H.P., Lehmann, C.E.R., Archibald, S.A., Osborne, C.P., Richardson, D.M., 2018. Global grass (Poaceae) success underpinned by traits facilitating colonization, persistence and habitat transformation. Biol. Rev. 93, 1125–1144. https://doi.org/10.1111/brv.12388.

Litman, J.R., Danforth, B.N., Eardley, C.D., Praz, C.J., 2011. Why do leafcutter bees cut leaves? Proc. R. Soc. 278, 3593–3600.

Martins, A.C., Alves-dos_Santos, I., 2013. Floral- oil-producing Plantaginaceae species: geographical distribution, pollinator rewards and interactions with oil-collecting bees. Biota Neotrop. 13, 77–89.

Martins, A.C., Bochorny, T., Pérez-Escobar, O.A., Chomicki, G., Monteiro, S.H.N., Smidt, E.D.C., 2018. From tree tops to the ground: reversals to terrestrial habit in Galeandra orchids (Epidendroideae: Catasetinae). Mol. Phylogenet. Evol. 127, 952–960.

Martins, A.C., Melo, G.A.R., 2016. The New World oil-collecting bees Centris and Epicharis (Hymenoptera, Apidae): molecular phylogeny and biogeographic history. Zool. Scr. 45, 22–33. https://doi.org/10.1111/zsc.12133.

Martins, A.C., Melo, G.A.R., Renner, S.S., 2015. Gain and loss of specialization in two oil-bee lineages, Centris and Epicharis (Apidae). Evol. 69, 1835–1844. https://doi.org/10.1111/evo.12689.

Martins, A.C., Melo, G.A.R., Renner, S.S., 2014. The corbiculate bees arose from New World oil-collecting bees: Implications for the origin of pollen baskets. Mol. Phylogenet. Evol. 80, 88–94. https://doi.org/10.1016/j.ympev.2014.07.003.

Matzke, N.J., 2013a. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in. R Scripts.

Matzke, N.J., 2013b. Probabilistic historical biogeography : new models for founder - event speciation, imperfect detection, and fossils allow improved accuracy and model – testing. Front. Biogeogr. 5, 242–248.

Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63, 951–970. https://doi.org/10.1093/sysbio/syu056.

McElwain, J.C., 2018. Paleobotany and global change: important lessons for species to biomes from vegetation responses to past global change. Annu. Rev. Plant Biol. 69, 761–787. https://doi.org/10.1146/annurev-arplant-042817-040405.

Michener, C.D., 2007. The Bees of the World, second ed. The John Hopkins University Press, Baltimore, Maryland, USA.

Michener, C.D., 1979. Biogeography of the Bees. Ann. Missouri Bot. Gard. 66, 277–347.

Michez, D., De Meulemeester, T., Rasmont, P., Nel, A., Patiny, S., 2009. New fossil evidence of the early diversification of bees: Paleohabropoda oudardi from the French Paleocene (Hymenoptera, Apidae, Anthophorini). Zool. Scr. 38, 171–181. https://doi.org/10.1111/j.1463-6409.2008.00362.x.

Miller, M.A., Pfeiffer, W., Schwartz, T., Schwarz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Gateway Computing Environments. New Orleans, USA, pp. 1–8.

Morrone, J.J., 2014. Biogeographical regionalisation of the Neotropical region. Zootaxa 10.11646/zootaxa.3936.2.3.

Oliveira-Filho, A.T., Ratter, J.A., 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. Edinburgh J. Bot. 52, 141–194. https://doi.org/10.1017/S0960428600000949.

Ortiz-Jaureguizar, E., Cladera, G.A., 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. J. Arid Environ. 66, 498–532. https://doi.org/10.1016/j.jaridenv.2006.01.007.

Paradis, E., Schliep, K., 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Pennington, T.R., Lewis, G., Ratter, J.A., 2006. Neotropical savannas and Dry Forests: Plant Diversity, Biogeography and Conservation. CRC Press Florida.

Pennington, T.R., Hughes, C.E., 2014. The remarkable congruence of New and Old World savanna origins. New Phytol. 204, 4–6. https://doi.org/10.1111/nph.12996.

Plana, V., Gascoigne, A., Forrest, L.L., Harris, D., Pennington, R.T., 2004. Pleistocene and pre-Pleistocene Begonia speciation in Africa. Mol. Phylogenet. Evol. 31, 449–461. https://doi.org/10.1016/j.ympev.2003.08.023.

Core Team, R., 2016. R: a language and environment. for statistical computing.

Rambaut, A., 2009. FigTree: Tree figure drawing tool.

Rambaut, A., Suchard, M.A., Drummond, A.J., 2014. Tracer v1.6, 2003-2013: MCMC trace analysis tool.

Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14. https://doi.org/10.1080/10635150701883881.

Renner, S.S., Schaefer, H., 2010. The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 423–435. https://doi.org/10.1098/rstb.2009.0229.

Rocha-Filho, L.C., Ferreira-Caliman, M.J., Carlos, J., Maria, J., Camargo, F. De, Garófalo, C.A., 2017. Nesting ecology of the oil-collecting bee Centris (Melacentris) conspersa Mocsáry and its potential association with the cleptoparasite Cyphomelissa diabolica Friese (Apidae: Centridini, Ericrocidini). J. Apic. Res. 56, 489–496. https://doi.org/10.1080/00218839.2017.1351906.

Roig-Alsina, A., 1997. A generic study of the bees of the tribe Tapinotaspidini, with notes on the evolution of their oil-collecting structures. Mitt. Munch. Ent. Ges. 87, 3–21.

Roig-Alsina, A., Michener, C.D., 1993. Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). Univ. Kansas Sci. Bull. 55, 123–173.

Romero, E.J., 1993. South American Paleofloras, in: Biological Connections between South America and Africa.

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 195. https://doi.org/10.2307/2413643.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029.

Rozen, J.G., Melo, G.A.R., Aguiar, A.J.C., Alves-dos-Santos, I., 2006. Nesting biologies and immature stages of the Tapinotaspidine bee genera Monoeca and Lanthanomelissa and of their osirine cleptoparasites Protosiris and Parepeolus (Hymenoptera: Apidae:

Apinae). Am. Museum Novit. 3501, 1–60.

Sérsic, A.N., 2004. Pollination biology in the genus *Calceolaria* L. (Calceolariaceae). Stapfia 82, 1–121.

Sigrist, M.R., Sazima, M., 2004. Pollination and reproductive biology of twelve species of neotropical malpighiaceae: stigma morphology and its implications for the breeding system. Ann. Bot. 94, 33–41. https://doi.org/10.1093/aob/mch108.

Silvestro, D., Michalak, I., 2012. raxmlGUI: a graphical front-end for RAxML. Org. Divers. Evol. 12 335–337. https://doi.org/DOI: 10.1007/s13127-011-0056-0.

Simon, M.F., Grether, R., de Queiroz, L.P., Skema, C., Pennington, R.T., Hughes, C.E., 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proc. Natl. Acad. Sci. 106, 20359–20364. https://doi.org/10.1073/pnas.0903410106.

Simpson, B.B., 1989. Krameriaceae, Flora. Neotropica 49, 1–108.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446.

Vasconcelos, T.N.C., Chartier, M., Prenner, G., Martins, A.C., Wingler, A., Lucas, E., 2019. Floral uniformity through evolutionary time in a species-rich tree lineage. New Phytol. 221, 1597–1608. https://doi.org/10.1111/nph.15453.

Vogel, S., 1974. Ölblummen und ölsammelnde Bienen. Trop. und Subtrop. Pflanzenwelt 7, 1–267.

Werneck, F.P., 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. Quat. Sci. Rev. 30, 1630–1648. https://doi.org/10.1016/j.quascirev.2011.03.009.

Werneck, F.P., Gamble, T., Colli, G.R., Rodrigues, M.T., Sites, J.W., 2012. Deep diversification and long-term persistence in the South American "dry diagonal": integrating continent-wide phylogeography and distribution modeling of geckos. Evolution 66, 3014–3034. https://doi.org/10.1111/j.1558-5646.2012.01682.x.

Wilf, P., Cúneo, N.R., Johnson, K.R., Hicks, J.F., Wing, S.L., Obradovich, J.D., 2003. High plant diversity in Eocene South America: evidence from Patagonia. Science 300, 122–125. https://doi.org/10.1126/science.1080475.

Willig, M.R., Kaufman, D.M., Stevens, R.D., 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annu. Rev. Ecol. Evol. Syst. 34, 273–309. https://doi.org/10.1146/annurev.ecolsys.34.012103.144032.

Xi, Z., Ruhfel, B.R., Schaefer, H., Amorim, A.M.A., Sugumaran, M., Wurdack, K.J., Endress, P.K., Matthews, M.L., Stevens, P.F., Mathews, S., Davis, C.C., 2012. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. Proc. Natl. Acad. Sci. USA 109, 17519–17524. https://doi.org/10.1073/pnas.1205818109.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, global rhythms, aberrations in global climate 65Ma to present. Science 292, 686–693. https://doi.org/10.1126/science.1059412.