A molecular phylogeny of the stingless bee genus Melipona (Hymenoptera: Apidae)

Santiago R. Ramírez¹*, James C. Niehb, Tiago B. Quentala, David W. Roubikc,d, Vera L. Imperatriz-Fonsecad, Naomi E. Piercea

¹ Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA
²Division of Biological Sciences, Section of Ecology, Behavior, and Evolution, University of California, San Diego, 0116, La Jolla, CA 92093, USA
³Smithsonian Tropical Research Institute, MRC 0580-12, Unit 9100 Box 0948, DPO AA 34002-9998, USA
⁴Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 14040-901 São Paulo, Brazil

A R T I C L E   I N F O

Article history:
Received 30 March 2009
Revised 2 April 2010
Accepted 28 April 2010
Available online xxxx

Keywords:
Relaxed molecular clock
Penalized likelihood
Eusociality
Stingless bees
Honey bees
Melipona
Apis
Referential communication

A B S T R A C T

Stingless bees (Meliponini) constitute a diverse group of highly eusocial insects that occur throughout tropical regions around the world. The meliponine genus Melipona is restricted to the New World tropics and has over 50 described species. Melipona, like Apis, possesses the remarkable ability to use representational communication to indicate the location of foraging patches. Although Melipona has been the subject of numerous behavioral, ecological, and genetic studies, the evolutionary history of this genus remains largely unexplored. Here, we implement a multigene phylogenetic approach based on nuclear, mitochondrial, and ribosomal loci, coupled with molecular clock methods, to elucidate the phylogenetic relationships and antiquity of subgenera and species of Melipona. Our phylogenetic analysis resolves the relationship among subgenera and tends to agree with morphology-based classification hypotheses. Our molecular clock analysis indicates that the genus Melipona shared a most recent common ancestor at least ~14–17 million years (My) ago. These results provide the groundwork for future comparative analyses aimed at understanding the evolution of complex communication mechanisms in eusocial Apidae.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The stingless bee genus Melipona contains at least 50 species of medium-sized (8–15 mm), robust, and often hairy bees inhabiting forests of tropical America, from Mexico to Argentina (Schwarz, 1932; Michener, 2007). Most species of Melipona inhabit lowland wet forests, with the greatest species diversity concentrated in the Amazon Basin (Moure and Kerr, 1950). These bees are highly eusocial, which means they exhibit reproductive division of labor, cooperative brood care, and overlap of generations (Wilson, 1971).

Similar to honey bees (Apis), Melipona are remarkable for insects, in their ability to recruit nest mates to specific foraging sites (von Frisch, 1967; Michener, 1974; Dyer, 2002; Nieh, 2004). All Apis use a form of referential communication known as the waggle dance, whereby returning foragers inform colony members about newly discovered resource sites (von Frisch, 1967; Seeley, 1995; Dyer, 2002). The waggle dance communicates distance and direction (von Frisch, 1967; Gould, 1976; Michelsen et al., 1992; Esch et al., 2001; Dyer, 2002; Sherman and Visscher, 2002). The communication mechanisms of Melipona are less studied, but experimental evidence indicates functional referential communication in some species (Esch, 1967; Aguilar and Briceno, 2002; Nieh, 2004), but not in others (Hrncir et al., 2006). Upon returning to the nest, successful M. panamica and M. seminigra foragers may perform short piloting flights outside of the nest in the direction of the resource (Nieh, 1998; Nieh and Roubik, 1998), while inside the nest, they produce sound pulses while distributing food samples to potential recruits (Esch, 1967; Nieh, 2004). The average duration of sound pulses correlates with, and thus potentially encodes, distance to food sources relative to the location of the nest (Esch, 1967; Nieh and Roubik, 1998). Additionally, there are differences in the ability to communicate different spatial dimensions among species of Melipona, which correlate well with spatial distribution of floral resources in their current environment (Nieh et al., 2003). Whether and how this information is actually utilized by nest mates is still a subject of intense investigation, as was the case for decades in Apis.

Although the genus has been the focus of behavioral, genetic, ecological, and pollination studies (Roubik, 2006), only partial phylogenetic analyses have been carried out to date (Rego, 1990; Costa et al., 2003; Fernandes-Salomão et al., 2005; Rasmussen and Cameron, 2010). The stingless bee genus Melipona is clustered within
the Neotropical Meliponini (Rasmussen and Cameron, 2010), and its monophyly is well supported (Rego, 1990; Costa et al., 2003; Fernandes-Salomão et al., 2005). A recent global phylogenetic analysis of the entire tribe Meliponini supported a Miocene (~24 My) origin for *Melipona*, but only 20 of the 50 described species were sampled and the internal relationships were not well resolved (Rasmussen and Cameron, 2010). Here, we present the first comprehensive species-level phylogenetic analysis of *Melipona* coupled with a molecular clock analysis.

2. Materials and methods

2.1. DNA sequencing and taxonomic sampling

We sequenced ~4.5 kb of DNA from five different fragments including mitochondrial *COI* (~1.2 kb), ribosomal 16S (~0.6 kb), nuclear *EF1-α* (~1.2 kb), ArgK (~0.7 kb), and *Pol-II* (0.8 kb). DNA was extracted from individual bee specimens from either leg or thoracic muscle tissue using Qiagen DNA Extraction Kits (Qiagen Inc., Valencia, California). Polymerase Chain Reactions (PCRs) were carried on a Bio-Rad DNA Engine Dyad™ Peltier thermal cycler (Bio-Rad Laboratories Inc., Hercules, California) in 25 μL reactions with 2.5 mmol/L MgCl₂, 2.5 mmol/L PCR buffer, and Taq polymerase (Qiagen Inc., Valencia, California) using various primer pairs (Danforth et al., 2004; Supplementary Table 1). We purified PCR products by incubating samples at 37 °C for 35 min using *Escherichia coli* Exonuclease I enzyme (New England Biolabs, Hanover, Maryland) and subsequently raising the temperature to 80 °C for 20 min. Purified products were cycle-sequenced using BigDye™ Terminator v3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, California). Samples were directly sequenced on an Applied Biosystems Inc., 3100 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA). Both forward and reverse strands were sequenced for each of the five markers; complementary strands were assembled using the software Sequencher™ v4.2 (Gene Codes Corp., Ann Arbor, MI).

All major lineages within *Melipona* and *Apis* were sampled for this study, including four subgenera, 35 species, and 51 individuals of approximately 50 described species of *Melipona* representing all main species groups, and three subgenera, six species, and 10 individuals of *Apis*. Additionally, we sampled 30 taxa within the corbiculate bees, including bumble bees, stingless bees, and orchid bees, and two outgroups (*Epicharis* and *Centris*). We include a total of 88 terminals. GenBank accession numbers are provided in Supplementary Table 3.

2.2. Phylogenetic analyses

A single DNA matrix containing five loci was assembled using MacClade v4.06 (Maddison and Maddison, 2003). Parsimony analyses were implemented in the software package Paup+ v4.0b (Swofford, 2003) with all characters weighted equally and transitions assumed unordered. We performed 100 random addition sequences using the TBR algorithm, and estimated node support via non-parametric bootstrapping (100 replicates). A Maximum Likelihood (ML) phylogenetic analysis was performed in the software package GARLI (Zwickl, 2006) with model parameters estimated over the specified number of runs. Bootstrap support values were estimated in GARLI with 100 heuristic tree searches using the same parameters as those implemented during tree searches. Additionally, Bayesian analyses were implemented in the software package MrBayes v3.1.1 (Ronquist and Huelsenbeck, 2003). Bayesian tree searches were made assuming both single (GTR+F+I) and multiple models of sequence evolution for each locus (see Supplementary Table 2). In addition, we ran a tree search where models of sequence evolution were partitioned by codon positions, with parameters estimated separately for first, second, and third codon positions of nuclear coding genes. Markov chain Monte Carlo (MCMC) searches were run for 10,000,000 generations, sampling every 1000 generations for a total of 10,000 trees; model parameters were estimated during the run. Three parallel runs were carried, and for each run one unheated and three incrementally heated chains were used. We checked for convergence within tree searches by plotting tree likelihood values against the number of generations, and among searches by comparing resulting topologies. Bayesian posterior probabilities were estimated as the proportion of trees containing each node over the trees sampled during runs. The trees corresponding to the first 1000 generations were discarded (“burn-in”).

2.3. Divergence time estimation

Divergence times were estimated using a fully resolved topology obtained by applying a 50% Majority-Rule (MR) consensus to all the trees obtained from a Bayesian phylogenetic search; the remaining polytomies (six) were resolved randomly using the R software package APE v2.3. Using a Likelihood Ratio Test (LRT) we estimated this tree had a significantly lower score value (~0.00) when a molecular clock was enforced than when the assumption was relaxed (~0.00). We calculated branch lengths on the 50% MR consensus tree via maximum likelihood in the software package Paup+, optimized under the model of sequence evolution GTR+F+I (molecular clock not enforced). Node divergence times were estimated with Penalized Likelihood (PL) using the Truncated-Newton algorithm in the software package r8s v1.71 (Sanderson, 1997). Mean ages ± SD were calculated using non-parametric bootstrapping.

We used two sets of calibration ages, corresponding to the youngest and oldest estimates of the ages of the fossils used as node age constraints. A total of five different ages were used to calibrate our molecular clock trees (indicated by letters in Fig. 1): A, maximum root age (80–100 My, based on oldest stem bee fossil (Poinar and Danforth, 2006) and molecular clock analysis done by Hines (2008)); B, *Cretotrigona prisca* (65–70 My, Michener and Grimaldi, 1988; Engel, 2000) used as a minimum age calibration; C, *Euglossa moronei* (15–20 My, Engel, 1999b) as a minimum age calibration; D *Apis lithohermaea* (14–16 My, Engel, 2006) used as a minimum age calibration; and *E. propyleia dominicana* (15–20 My, Wille and Chandler, 1964; Camargo et al., 2000) as a minimum age calibration. Although the age of *C. prisca* has been the subject of controversy (Michener and Grimaldi, 1988; Engel, 2000), this fossil exhibits synapomorphic characters that unambiguously place it within crown Meliponini. Thus, we used its age as a minimum age calibration for all Meliponini. The placement of *E. moronei* within extant (crown) *Euglossa* is justified by the presence of elongated mouthparts, labrum shape, and pubescence (Engel, 1999b). The phylogenetic position of *A. lithohermaea* within extant *Apis* is justified by the enlarged body size, elongated metabasitarsus, wing venation, and infuscated wing membrane (Engel, 2006). The placement of *P. dominicana* within extant Neotropical Meliponini is justified by the short trapezoidal clypeus, triangular shape of forewing medial cell, and shape of titiae and basitarsi (Camargo et al., 2000). The concordance among calibration points was assessed with the cross-validation method (Near et al., 2005; Supplementary Fig. 3). Since our phylogenetic sampling included divergent extant lineages within *Apis* (Raffiudin and Crozier, 2007), *Euglossa* (Ramírez et al., in press) and Neotropical Meliponini (Rasmussen and Cameron, 2010) we used fossil ages as minimum age constraints, even though in some cases lineage sampling was incomplete (e.g. *Euglossa*).
3. Results and discussion

3.1. Phylogenetic relationships

Our maximum parsimony, maximum likelihood, and Bayesian phylogenetic analyses, based on five loci, resolved relationships within and between Melipona, Apis, and related clades of corbiculate bees. We obtained well-resolved and supported phylogenetic trees (Figs. 1 and 2) that are congruent with each other under different optimization schemes and model parameters (Figs. 1, 2 and Supplementary Figs. 1, 2). Parsimony analyses yielded 32 shortest trees (TL = 6344), with a strict consensus almost identical to a max-

---

**Fig. 1.** Strict consensus cladogram of 32 equally short maximum parsimony trees showing both maximum likelihood and maximum parsimony parametric bootstrap values; asterisks denote unavailable values. Five different ages were used to calibrate a molecular clock (indicated by letters): A, maximum root age (80–100 My, estimated based on Hines (2008) and Poinar and Danforth (2006)); B, minimum age constraint, *Cretotrigona prisca* (65–70 My, Michener and Grimaldi, 1988; Engel, 2000); C, minimum age constraint, *Euglossa moronei* (15–20 My, Engel, 1999b); D, minimum age constraint, *Apis lithohermaea* (14–16 My, Engel, 2006); and E, minimum age constraint, *Proplebeia dominicana* (15–20 My, Wille and Chandler, 1964; Camargo et al., 2000).
imum likelihood analysis (Fig. 1). Bayesian phylogenetic analyses also yielded well-supported trees that varied little whether using single (GTR+Γ+I), or partitioned (locus-specific and codon-specific) models of sequence evolution (Supplementary Figs. 1 and 2). All phylogenetic methods returned (i) *Melipona* as sister to the other Neotropical meliponine taxa included in our study (14 genera); (ii) *Melipona* + Neotropical Meliponini as sister to the African stingless bee genus *Meliponula*; and (iii) all stingless bees (Meliponini) + bumble bees (Bombini) as a monophyletic clade. These results concord with previous molecular studies (Cameron and

Fig. 2. Relaxed-clock chronogram of *Melipona* and related groups; the tree corresponds to the maximum a posteriori topology resulting from a Bayesian tree search. Divergence times were calculated (in millions of years, My) using the software package r8s v1.71. Node bars represent the 95% confidence intervals of the age of each node. A single model of sequence evolution was fitted to all loci (GTR+Γ+I) to estimate branch lengths. Statistics for numbered nodes are indicated in Table 1.
Mardulyn, 2001; Rasmussen and Cameron, 2007, 2010; Kawakita et al., 2008). In a recent study by Rasmussen and Cameron (2010), *Melipona* was recovered as sister to most Neotropical genera, except *Celetrigona*, *Dolichotrigna*, *Trigoniscia*, and *Leurotrigna*. Although our study did not include these taxa, our results are congruent with their hypothesis about the placement of *Melipona*. The placement of both honey bees (*Apis*) and orchid bees (*Euglossini*) were inconsistent in our study, depending on methodology (Supplementary Figs. 1 and 2). The parsimony and codon-partitioned Bayesian analyses supported the topology of (*Euglossini* (*Apini*, (*Bombini, *Meliponini*))), whereas the Likelihood, Bayesian (optimized with both single and gene-partitioned models) supported the topology of (((*Euglossini*, *Apini*), (*Bombini, *Meliponini*))). Although our analyses recovered the clade *Apini* + *Euglossini* as monophyletic only in some analyses, we obtained strong support for the clade *Bombini* + *Meliponini*, a group that has been controversial (Kawakita et al., 2008). Bombini is primarily boreal and temperate in distribution and Meliponini is restricted to tropical latitudes. Overall, our results differ from earlier morphology-based hypothesis about the relationships of corbiculate bees (Michener, 1944; Engel, 2001; Schultz et al., 2001) and, perhaps not surprisingly, support hypotheses based on molecular data (Cameron and Mardulyn, 2001; Kawakita et al., 2008).

Our phylogenetic analyses indicate that three of the four subgenera recognized by morphology within *Melipona* (*Melikerria*, *Melipona* s. str., *Michmelia*) are monophyletic, but one (*Eomesa*) is polyphyletic (Fig. 2). All *Melipona* species currently unassigned to a specific subgenus (*Incertae sedis*)—except *M. fuliginosa* (Camargo and Pedro, 2007, 2008)—form a monophyletic clade, sister to the subgenus *Michmelia* (Fig. 2). In our MP analysis, *M. amazonica* was sister to the rest of species in the genus, but this placement was supported by a low bootstrap value (67). On the other hand, our Bayesian analysis supported the placement of *M. amazonica* as sister to *M. marginata* and *M. bicolor*, which was supported by a high Bayesian posterior probability (98). *M. amazonica* constitutes a problematic taxon and additional genetic fragments may be required to resolve its placement. Additionally, some of the internal branches within *Melipona* were not resolved or were not well supported. We note that multiple branches within *Melipona* are relatively short, particularly in the subgenus *Michmelia* and thus a rapid lineage diversification may explain the observed low support values. The results from our phylogenetic analysis may guide future taxonomic studies, particularly on the delineation of subgenera and assignment of unplaced species.

Our phylogenetic analyses agree with the proposed and widely accepted hypothesis of the internal relationships of honey bees: (*Micrapis*, (*Megapis, *Apis* s. str.)) (Arias and Sheppard, 1996, 2005; Oldroyd and Wongsiri, 2006; Raffiudin and Crozier, 2007) and concur with molecular phylogenetic studies of the corbiculate bee tribes based on molecular data (Cameron and Mardulyn, 2001; Thompson and Oldroyd, 2004; Kawakita et al., 2008).

### 3.2. Molecular clock analysis

We performed molecular clock analyses calibrated with five different fossil ages using Penalized Likelihood (PL). Because previous phylogenetic analyses (and our own results) have produced uncertainty in the placement of *Apini* and *Euglossini*, we used two alternative tree topologies that resulted from our Bayesian analyses (Supplementary Fig. 1) that were produced by applying both single and partitioned models of sequence evolution. The main difference between these two alternative topologies was in the relative position of *Apis* and *Euglossini*, where *Apis* was recovered sister to *Euglossini* (single model of sequence evolution), and *Euglossini* was recovered sister to the remaining corbiculate bees (gene and codon partitions). To account for this uncertainty in our subsequent molecular clock analyses, we used these two alternative tree topologies. The results from our PL analysis suggest that whereas extant *Apis* shared a recent common ancestor during the Oligocene, 29 ± 2 to 34 ± 2 My ago, *Melipona* shared a most recent common ancestor during the Miocene, 14 ± 1 to 17 ± 1 My ago, depending on whether we use the oldest or youngest ages of the fossil calibrations (Table 1). These time estimates varied little when using either of the two alternative models of sequence evolution (Table 1) or the two alternative topologies (data not shown).

*Melipona* is one of the two largest (*species* rich) genera in the highly eusocial stingless bees (*Plebeia* is the other genus). *Melipona* is also exclusively Neotropical, and its origin is not known (Rasmussen and Cameron, 2010). The few endemic species on islands in both the Lesser Antilles and Pacific Panama (Camargo and Pedro, 2007) indicate relics of past mainland connections during Miocene times (Roubik and Camargo, unpublished data). Our molecular clock analysis coincides with this scenario.

Our fossil-calibrated molecular clock provides an age estimate for the origin of *Apis*, *Melipona*, and the main clades of corbiculate Apidae. This provides a temporal framework with which to estimate the antiquity of referential communication. Although the genus *Apis* has an extensive fossil record, with the oldest fossil dating to the Oligocene (~25 My ago) (Engel, 1998, 1999a, 2006; Engel et al., 2009), all honey bee fossils known to date (with the exception of *A. liothermae*) are stem relatives of extant *Apis* (Engel, 2006). By implementing molecular clocks, we show that extant *Apis* likely shared a most recent ancestor during the Eocene–Oligocene (~29–33 My ago), whereas *Melipona* appears to have shared a common ancestor more recently, during the Miocene (14–17 My ago). Our age estimates for the most recent common ancestor of *Melipona* differ from those obtained by Rasmussen and Cameron (2010), which suggested that living members of genus shared an ancestor ~25 My ago. In their analysis, Rasmussen and Cameron (2010) specified a maximum age for the root node (*Meliponini*) of 125 My based on the oldest fossil angiosperms (most bees, except roughly 20 meliponine species (*Lestrimelitta, Cleptotrigna*, and the *Trigona hypogea* group) depend on flowering plants for feeding. We used a different age estimate (80–100 My) for the

---

**Table 1**

<table>
<thead>
<tr>
<th>Node</th>
<th>Younger calibrations</th>
<th>Older calibrations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GTR+I+H (μ ± SD)</td>
<td>Locus-specific (μ ± SD)</td>
</tr>
<tr>
<td><em>Apis</em></td>
<td>29.92 ± 1.51</td>
<td>30.57 ± 1.57</td>
</tr>
<tr>
<td><em>Melipona</em></td>
<td>15.43 ± 0.89</td>
<td>14.56 ± 0.82</td>
</tr>
<tr>
<td><em>Euglossini</em></td>
<td>20.25 ± 1.27</td>
<td>20.89 ± 1.34</td>
</tr>
<tr>
<td>Neotropical <em>Meliponini</em></td>
<td>22.04 ± 1.52</td>
<td>21.10 ± 1.71</td>
</tr>
<tr>
<td>1</td>
<td>6.34 ± 0.48</td>
<td>6.32 ± 0.40</td>
</tr>
<tr>
<td>2</td>
<td>8.49 ± 0.65</td>
<td>8.73 ± 0.65</td>
</tr>
<tr>
<td>3</td>
<td>11.66 ± 0.73</td>
<td>12.23 ± 0.78</td>
</tr>
<tr>
<td>4</td>
<td>8.29 ± 0.62</td>
<td>8.22 ± 0.68</td>
</tr>
<tr>
<td>5</td>
<td>NA</td>
<td>10.60 ± 0.83</td>
</tr>
<tr>
<td>6</td>
<td>32.04 ± 2.51</td>
<td>31.69 ± 2.44</td>
</tr>
<tr>
<td>7</td>
<td>17.27 ± 1.42</td>
<td>17.51 ± 1.49</td>
</tr>
<tr>
<td>8</td>
<td>22.32 ± 1.48</td>
<td>22.68 ± 1.54</td>
</tr>
<tr>
<td>9</td>
<td>18.56 ± 1.80</td>
<td>18.89 ± 1.79</td>
</tr>
<tr>
<td>10</td>
<td>77.62 ± 1.14</td>
<td>81.76 ± 1.23</td>
</tr>
<tr>
<td>11</td>
<td>83.87 ± 1.36</td>
<td>81.76 ± 1.23</td>
</tr>
</tbody>
</table>
divergence between corbiculate bees and the outgroup (Centridini). We suspect that the difference between both studies stems from applying different ages to the root node. The discovery of new fossils may shed new light on the time of origin of apid lineages.

Because no morphological characters have been associated with the use of recruitment communication in bees, we cannot infer whether stem fossil Apis, or any other extinct stingless bee lineages (Engel, 2001), exhibited this form of communication behavior. However, since all extant members of Apis use recruitment communication, it is likely that the most recent common ancestor of extant Apis had a form of recruitment communication similar to that exhibited by modern species. If true, this would suggest that recruitment communication in honey bees has been stable since the Eocene–Oligocene. On the other hand, our results suggest that the genus Melipona shared a most recent common ancestor more recently, during the Miocene. Although detailed behavioral observations are available for less than 10 of the 50 Melipona species in our phylogeny (Nieh, 2004), available data suggest that communication abilities in Melipona are more variable (Nieh, 2004). Thus, our study suggests that the traits associated with communication behavior are perhaps younger and more flexible in Melipona than in Apis. This study should guide future comparative analyses of referential communication and of other aspects of life history evolution in Melipona and other eusocial bees.

Acknowledgments

We thank保罗 Nogueira-Neto and Jay Evans for providing crucial samples used in this study, Jennifer Davis and Jessica Girard for assistance in the laboratory, and João M.F. Camargo and Claus Rasmussen for their assistance in the identification of bee specimens. Beth Pringle provided useful comments on the manuscript. This project was supported by grants from ORBS (Opportunities for Research in the Behavioral Sciences Program), the National Science Foundation (NSF-DDIG, DEB 0608409, NSF-IBN 0316697, NSF DEB 0447242, and NSF-IBN 0554586) and grants from the Putnam Expedition Fund to J.C.N., S.R.R., and N.E.P. Samples were collected in Brazil under the Permit No. 057/2003 (IBAMA) and Peru under Permit No. 005141 (AG-IRENA).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.04.026.

References


