16S mtDNA variation in *Apis mellifera* detected by PCR-RFLP*

Thaís COLLETA, Maria Cristina ARIASb, Marco Antonio DEL LAMAA

a Departamento de Genética e Evolução, Universidade Federal de São Carlos, CP 676 CEP 13.565-905, São Carlos, SP, Brazil
b Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, CEP 05508.900, São Paulo, Brazil

Received 13 December 2005 – Accepted 6 June 2006

Abstract – Phylogeographic and morphometric evidence can be used to cluster *Apis mellifera* subspecies into evolutionary lineages or branches. Mitochondrial DNA sequence and restriction site analyses have shown similar clustering of subspecies groups. Thus, mtDNA variation can be used to infer honey bee evolutionary relationships. In this paper, we describe three 16S mtDNA PCR-RFLP patterns, each one completely associated with a previously determined A, M, or C *Dra* I restriction pattern of the COI-COII region. These results indicate that the COI-COII and the 16S genes have had a very closely linked evolutionary history. Although distinct patterns were obtained with *Eco* RI, *Alu* I, *Hinc* II and *Taq* I, the best differentiation among the three patterns was observed with *Dra* I and *Vsp* I enzymes. Nucleotide sequence analysis of the 16S gene fragment displayed 10 sites of base substitution (1.35%) among the three patterns and two insertions in the *A. m. scutellata* pattern.

*Apis mellifera* / 16S patterns / mitochondrial DNA / PCR-RFLP

1. INTRODUCTION

Phylogeographic and morphometric evidence supports the clustering of *Apis mellifera* L. subspecies into four evolutionary lineages or branches: M (west European), C (east European), O (Near and Middle Eastern) and A (African) (Ruttner et al., 1978; Ruttner, 1988). A similar cluster of subspecies groups can be derived from mitochondrial DNA sequence or restriction site analyses, with some exceptions. More recently, mitochondrial lineages “Y” from Ethiopia and “O” from Egypt have been reported (Franck et al., 2000b, 2001).

Honeybees of the A, M and C mitochondrial classification includes the primary subspecies that were introduced worldwide. Today, hybrid descendents of these lineages can be found in many areas of the world, owing to widespread transport and introduction of *A. mellifera* populations by beekeepers.

In the last twenty-five years the genetic composition of introduced honeybee populations in Argentina (Sheppard et al., 1991), Tasmania, Australia (Oldroyd et al., 1995), Balearic and Canary Islands (de la Rua et al., 2001a, b), Yucatan, Mexico (Clarke et al., 2001), Brazil and Uruguay (Lobo et al., 1989; Del Lama et al., 1990; Diniz et al., 2003), Peru (Quezada-Euán et al., 2003) and Chile (Del Lama et al., 2004) has been analyzed.

The maternal origin of honeybee colonies has been identified by mitochondrial markers. Cytochrome b (CytB) gene amplification followed by *Bgl* II restriction allows assignment of tested samples to European and Middle Eastern (M and C+O) or African (A) evolutionary lineages (Crozier et al., 1991). Further analysis using the the *Hinc* II site for COI
Table I. Source of the Apis mellifera samples analyzed and association with Dra I (COI-COII) haplotypes (according to Garnery et al., 1993).

<table>
<thead>
<tr>
<th>Source</th>
<th>COI-COII</th>
<th>16S</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M3 M4 M6 M7 C1 C2 A1 A4 A26 A28 A29 A30</td>
<td>M C A</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>15 45</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Chile</td>
<td>2 4</td>
<td>17</td>
<td>19 36</td>
</tr>
<tr>
<td>Colombia</td>
<td>23 29 25 3 2</td>
<td>23 59 82</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>1 2 1</td>
<td>3 1</td>
<td>4</td>
</tr>
<tr>
<td>USA</td>
<td>12 1</td>
<td>13 13</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>5 10 2 26 1</td>
<td>17 26 1 44</td>
<td></td>
</tr>
<tr>
<td>Uruguay</td>
<td>1 4 4 33 1 5</td>
<td>1 4 43 48</td>
<td></td>
</tr>
<tr>
<td>Venezuela</td>
<td>2 2 27 41</td>
<td>4 68 72</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8 14 1 15 80 9 75 145 4 1 2 5</td>
<td>38 89 232 359</td>
<td></td>
</tr>
</tbody>
</table>

COI-COII data from 1 Collet, 2004; 2 Ferreira, 2002; 3 Prada, 2004; 4 Souza, 2002; 5 unpublished data.

discriminates the M lineage from the C+O lineage in the non-African group (Hall and Smith, 1991).

Currently, digestion of the mitochondrial COI-COII intergenic region with Dra I (Garnery et al., 1993) provides the most resolved characterization of subspecies assignment to mitochondrial lineage (A, M, C, Y or O) (Moritz et al., 1994; Garnery et al., 1995; Franck et al., 1998, 2000a,b, 2001). The COI-COII region is normally composed of two distinct nucleotide sequences, named P and Q. According to Garnery et al. (1993), each evolutionary branch includes several variants of the P sequence combined with a different copy number of the Q sequence, resulting in a length polymorphism of this mtDNA region. Additional polymorphism can be resolved following digestion of the length variants with Dra I enzyme.

Analysis of populations via molecular markers (e.g., mtDNA sequence or RFLP variation) is a particularly useful method for elucidating historical relationships underlying population differentiation (Avise, 2000). Moreover, the mitochondrial genome yields a great amount of data that can be used to deduce evolutionary relationships (Gray et al., 1999). Therefore, haplotype networks generated by data taken from polymorphic loci could provide a valuable resource for examining pathways and mechanisms of mtDNA genome evolution. This study presents such data by examining a polymorphism found at the 16S locus. This analysis was conducted on samples that had their maternal origin previously identified as A, M, or C, based on COI-COII patterns. Three 16S patterns were observed, each one completely associated with the A, M, or C DraI restriction patterns of the COI-COII region, showing that COI-COII and 16S genes have a very closely linked evolutionary history.

2. MATERIALS AND METHODS

2.1. Samples and DNA extraction

Africanized populations from Brazil (n = 39), Uruguay (n = 6), Colombia (n = 24) and Venezuela (n = 5) were analyzed. Previously, Africanized populations had been shown to be primarily composed of African mtDNA (A) and of African A. m. scutellata (A) and European, A. m. mellifera (M) and A. m. ligustica (C) nuclear genes (Lobo et al., 1989; Del Lama et al., 1990; Hall and Smith, 1991; Diniz et al., 2003). Three hundred and fifty-nine colonies, whose maternal origin had previously been identified according to the Dra I pattern for the COI-COII region (Ferreira, 2002; Souza, 2002; Collet, 2004; Prada, 2004), were analyzed. European COI-COII patterns were obtained in samples from Chile (free from Africanized swarms), Pullman (USA), Italy (Milano), and the Galician region (Spain) (see Tab. I).

Worker honeybees were stored at ~20 °C until analysis. Total DNA was extracted by
the phenol-chloroform method (Sheppard and McPheron, 1991) from the thorax of one worker from each colony.

2.2. PCR amplification and endonuclease digestion

Restriction analyses of the CytB (Bgl II), COI (Hinc II) and COI-COII (Dra I) mitochondrial DNA regions were performed, to provide a comparative molecular characterization of the maternal origin of the bees sampled. This previous characterization was carried out as a mean to assign these bees to the African lineage (absence of Bgl II site in the CytB region), or to European subspecies (only A. m. mellifera has the Hinc II site in the COI region) or to characterize our samples according to their COI-COII (Dra I) patterns.

The primers utilized for CytB, COI and 16S were described by Hall and Smith (1991) and the amplification and restriction of the intergenic COI-COII region followed Garnery et al. (1992). The PCR was performed in a total volume of 25 µL, containing 10× reaction buffer, 250 µM of each dNTP, 2.5 mM MgCl₂, 1 µM of each primer, 1 µL of DNA, 1 U of Taq Polymerase (Promega) and 16 µL of sterile water. PCR amplification for CytB, COI-COII and 16S was performed as follows: 30 cycles of 94 °C for 30 s, 54 °C for 15 s and 62 °C for 2 min. Reactions for COI were submitted to an initial denaturation step of 3 min at 94 °C followed by 3 cycles of 94 °C for 1 min, 50 °C for 2 min and 72 °C for 3 min; then 35 cycles of 94 °C for 1 min, 2 min at 50 °C and 1.5 min at 72 °C, and a final extension step of 5 min at 72 °C. After amplification, 2 µL of PCR products were electrophoresed in 8% polyacrylamide gels and 2 µL were digested with the following restriction enzymes: Eco RI, Alu I, Hinc II, Taq I, Vsp I and Dra I for the 16S fragment; Bgl II for CytB, Hinc II for COI and Dra I for COI-COII. The restriction reactions were kept at 37 °C for 4 h (except for Taq I which requires 60 °C). Restriction fragments were separated on 10% polyacrylamide gels.

A 16S double digestion was performed with Eco RI and Vsp I restriction enzymes to produce a more visible differentiation among the patterns.

2.3. DNA cloning and sequencing

One sample of each 16S pattern identified was sequenced. PCR fragments of the three 16S restriction patterns were cloned, using the T-Easy cloning kit (Promega), and used to transform competent E. coli DH-5α cells. Positive clones were selected. The recombinant vectors were recovered and sequenced, following the protocols suggested by Applied Biosystems (http://www.appliedbiosystems.com). An automated sequencer ABI-3100 (Applied Biosystems) was used to sequence the samples. Two clones of each pattern were sequenced from both directions.

3. RESULTS

After 16S amplification and restriction analysis, three patterns were observed. They were named A, M and C because each of them showed complete association with one of the COI-COII Dra I patterns (A, M and C) (Tab. I).

The sequences obtained for the 16S gene fragment showed a total size of 740 bp in the C and M patterns, and 742 bp in the A pattern. Pairwise and multialignment sequence comparisons of the DNA sequences revealed nucleotide substitutions at 10 nucleotide sites (1.35%), eight transitions (1 A→G and 7 C→T) and two transversions (1 A→T and

<table>
<thead>
<tr>
<th>13770</th>
<th>13811</th>
<th>13961</th>
<th>13990</th>
<th>14018</th>
<th>14025</th>
<th>14097</th>
<th>14098</th>
<th>14182</th>
<th>14190</th>
<th>14352</th>
<th>14385</th>
<th>14405</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. m. ligustica*</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>-</td>
<td>-</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>A. m. ligustica</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A. m. mellifera</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>A. m. scutellata</td>
<td>C</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Figure 1. Variable sites of the mtDNA 16S region. The numbers correspond to the nucleotide position of A. m. ligustica (*) sequence published by Crozier and Crozier (1993). Position 14025 represents the difference between ours and Crozier and Crozier’s (1993) sequence. The dots indicate nucleotide identity and (-) gaps. The A pattern from Africanized samples showed two insertions at the 14097 and 14098 positions (respectively 390 and 391 in our 742 bp fragment).
of the M, C and A patterns. No variation was observed within each pattern in these samples.

Some base substitutions or insertions occurred at restriction sites. The Vsp I site at position 391 was lost in the A pattern, due to two insertions. We also detected the loss of a Dra I site at position 477 in the C pattern, due to a base-pair substitution. Double digestion with Eco RI and Vsp I restriction enzymes allowed a more noticeable differentiation among the three patterns (Fig. 3).

The Vsp I, Dra I and Eco RI restriction patterns for the 16S region were compared with those obtained by using Bgl II (CytB) and Hinc II (COI) patterns. For most of the samples, the two molecular approaches agreed well. However, some differences were observed with the Hinc II site in the COI gene. For example, two colonies (corresponding to 5.6%) from Chile were identified as M according to Hinc II (COI), but showed the C pattern by Vsp I, Dra I and Eco RI (16S). Distinct results were also obtained in 13% (six colonies) of all the Italian samples. Furthermore, a colony from Italy (Seregno 5) was typed as the African pattern according to Vsp I, Dra I and Eco RI (16S) and Bgl II (CytB), but for Hinc II (COI) the pattern was M (see Tab. II).

Figure 2. 16S amplification and restriction patterns detected in honeybees after 10% polyacrylamide gel electrophoresis and silver staining. The enzymes utilized and their respective patterns (C, M and A) are indicated above. A: Lanes 10 and 11 correspond to 50 bp and 25 bp size standards, respectively. Lanes 18, 19 and 20 represent the amplified fragments. B: Dra I restriction patterns from 16S region.

Figure 3. C, M and A 16S patterns resulting from Eco RI and Vsp I double digestion.
Table II. Comparative results observed with the mtDNA markers utilized for maternal origin identification. Cyt B (Bgl II) differentiates African (A) from European (E) samples (Crozier et al., 1991) while COI (Hinc II) differentiates melli fera (M) from the east European bees (C) (Hall and Smith, 1991).

<table>
<thead>
<tr>
<th>Source</th>
<th>Colony</th>
<th>Cyt B COI</th>
<th>16S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>Seregno 5</td>
<td>A M A A</td>
<td>Bgl II Hinc II Vsp I Dra I Eco RI</td>
</tr>
<tr>
<td>Italy</td>
<td>Meda 3</td>
<td>E C M M</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Meda A</td>
<td>E M C C</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Meda 2</td>
<td>E C M M</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Meda 7</td>
<td>E M C C</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Meda 10</td>
<td>E M C C</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>Coyhaique</td>
<td>M C</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>Olmoé</td>
<td>M C</td>
<td></td>
</tr>
</tbody>
</table>

4. DISCUSSION

As each one of the three 16S patterns showed a complete correspondence with an M, C or a Dra I pattern of the COI-COII region, we named the 16S patterns according to their associated COI-COII patterns. The substitution rate we found for the 16S fragment was slightly lower than the 1.67% rate reported from a region encompassing the tRNA ILE and part of the ND2 gene of the mtDNA of the same subspecies (Arias and Sheppard, 1996).

As expected, the 16S fragment showed a high A + T content. In accordance with Whitfield and Cameron (1998), insects in general and Hymenoptera in particular exhibit a significantly high A+T content, a result also reported for Apis mellifera mtDNA by Crozier and Crozier (1993), Cameron (1993) and Arias and Sheppard (1996). The high transversion to transition ratio we found was similar to that reported for other mtDNA regions in the honey bee (Arias and Sheppard, 1996).

The Galician (Spain) samples presented haplotypes from M and A lineages, a result which corroborates previous studies (Smith et al., 1991; Franck et al., 1998). According to these authors, Spain was a possible secondary contact zone between these two evolutionary branches. Analysis of continental Italian samples showed the presence of both M and C patterns as has been previously reported (Franck et al., 2000a). Their explanation for the widespread distribution of the M pattern within Liguria was based on paleogeographical patterns of refugial isolation during glaciation followed by recolonization. These authors also corroborated the presence of the A pattern in honeybees of Sicily, supporting the conclusion that some Mediterranean island subspecies derived from A lineage ancestors (Arias and Sheppard, 1996; Sheppard et al., 1997).

The Dra I and the Vsp I restriction patterns of the 16S fragment allowed the best distinction and positive identification of the three patterns. This is attributable to the two insertions in the A pattern at positions 390/391 and to the substitution at position 477 of the C pattern, that cause the loss of the Vsp I and Dra I sites, respectively (Fig. 1). Patterns M and C are clearly distinguished with Vsp I digestion due to the difference in mobility of the 129 bp fragment (Fig. 2). In relation to A and M differentiation by Dra I digestion, the insertions in the African mtDNA give rise to a 102 bp fragment which presents a different gel mobility from that of the M 100 bp fragment. Moreover, it is evident that the 116 bp fragments have distinct electrophoretic mobilities (A versus C and M), certainly due to conformational differences arising from nucleotide substitutions. This same explanation can also be applied to the Vsp I 129 bp fragment.

Double digestion with Eco RI and Vsp I resulted in easier differentiation among the three 16S mtDNA patterns. This double digestion is justified by the fact that only the C haplotype has the Eco RI restriction site, facilitating its distinction from the M pattern, whose differentiation by Vsp I is less pronounced (see Fig. 3).

The controls we carried out gave different results between the COI pattern (Hinc II) and those revealed by Dra I (COI-COII), Bgl II (CytB) and Vsp I and Dra I (16S). This indicates that the former marker may lead to an incorrect identification of the maternal origin. Also, the amplification by PCR of a 16S gene region and subsequent digestion with six restriction enzymes showed the characteristic M, C and A mtDNA patterns, previously characterized by Dra I (COI-COII).
Although the variation associated with the 16S locus could be considered an alternative method for the maternal origin of a colony (vis a vis mitochondrial lineage), the methods described here should be more valuable as a means to infer events related to mitochondrial genome evolution. Further, it should be possible to relate these events to *A. mellifera* subspecies and population level differentiation. In this sense, similar studies should be extended to other mitochondrial gene loci and morphologically characterized samples of the different subspecies of *Apis mellifera* to reconstruct this evolutionary history.

ACKNOWLEDGEMENTS

We are grateful to Isabel Cristina de Godoy, Rogério Oliveira Souza, Carlos Prada Quiroga, and Suzy Coelho for their technical assistance. This study was supported by ‘Fundaçao de Amparo à Pesquisa do Estado de São Paulo’ (Fapesp - 03/06342-9) and ‘Conselho Nacional de Desenvolvimento Científico e Tecnológico’ (CNPq).

Variation de l’ADNmt 16S mis en évidence par PCR-RFLP chez *Apis mellifera*.

ADNmt / profil 16S / phylogéographie / évolution du génome / histoire évolutive / PCR-RFLP / *Apis mellifera*


*Apis mellifera* / 16S Muster / mitochondriale DNA / PCR-RFLP

REFERENCES


Franck P., Garnery L., Celebroano G., Solignac M., Cornuet J.M. (2000a) Hybrid origins of honeybees from Italy (Apis mellifera ligustica) and Sicily (A.m. sicula), Mol. Ecol. 9, 907–921.


Souza R.O. (2002) Dinâmica do fluxo gênico em populações de *Apis mellifera* do Chile observada através de marcadores nucleares e mitocondriais, MSc Thesis, Universidade Federal de São Carlos, SP, Brazil, 87 p.