

Maternal structure of Iberian honey bees inferred from whole mitochondrial genomes



Dora Henriques^{1,2}, Julio Chávez-Galarza^{1,2}, John C. Patton³, José Rufino¹, M. Alice Pinto¹

¹Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Campus de Santa Apolónia, Apartado 1172, Bragança 3201-855, Portugal.

²Centro de Biologia Molecular e Ambiental (CBMA), Universidade do Minho, Campus de Gualtar, Braga 4710-057, Portugal.

³Dept. of Entomology, Texas A&M University, College Station, Texas 77843-2475, USA

Introduction

- The maternally inherited mitochondrial DNA has been the marker of choice for assessing Iberian honey bee variation, particularly the intergenic tRNA^{leu}-cox2 region. The data generated by massive sampling of this region confirmed early findings of coexistence of African (A) and western European (M) lineages, forming a southwestern–northeastern cline, and revealed unparalleled levels of haplotype diversity and complexity. Accordingly, it has been suggested that Iberia served as a glacial refuge, and as a place of secondary contact between European and African lineages.
- While we have learned a great deal with this region, due to its high levels of variation and repetitive structure, there are evolutionary questions that can only be properly addressed using other mitochondrial regions. In this study, we used NGS technology to sequence the mitogenomes of 92 individuals and analyzed the data using two phylogenetic methods.

Samples

- A total of 92 mitogenomes were sequenced, as follows: 70 *A. m. iberiensis*, 7 *A. m. intermissa*, 8 *A. m. mellifera*, 3 *A. m. carnica*, 4 *A. m. ligustica*. The *A. m. iberiensis* samples were collected along 3 transects: Atlantic (AT), Central (CT) and Mediterranean (MT) (Figure 1).

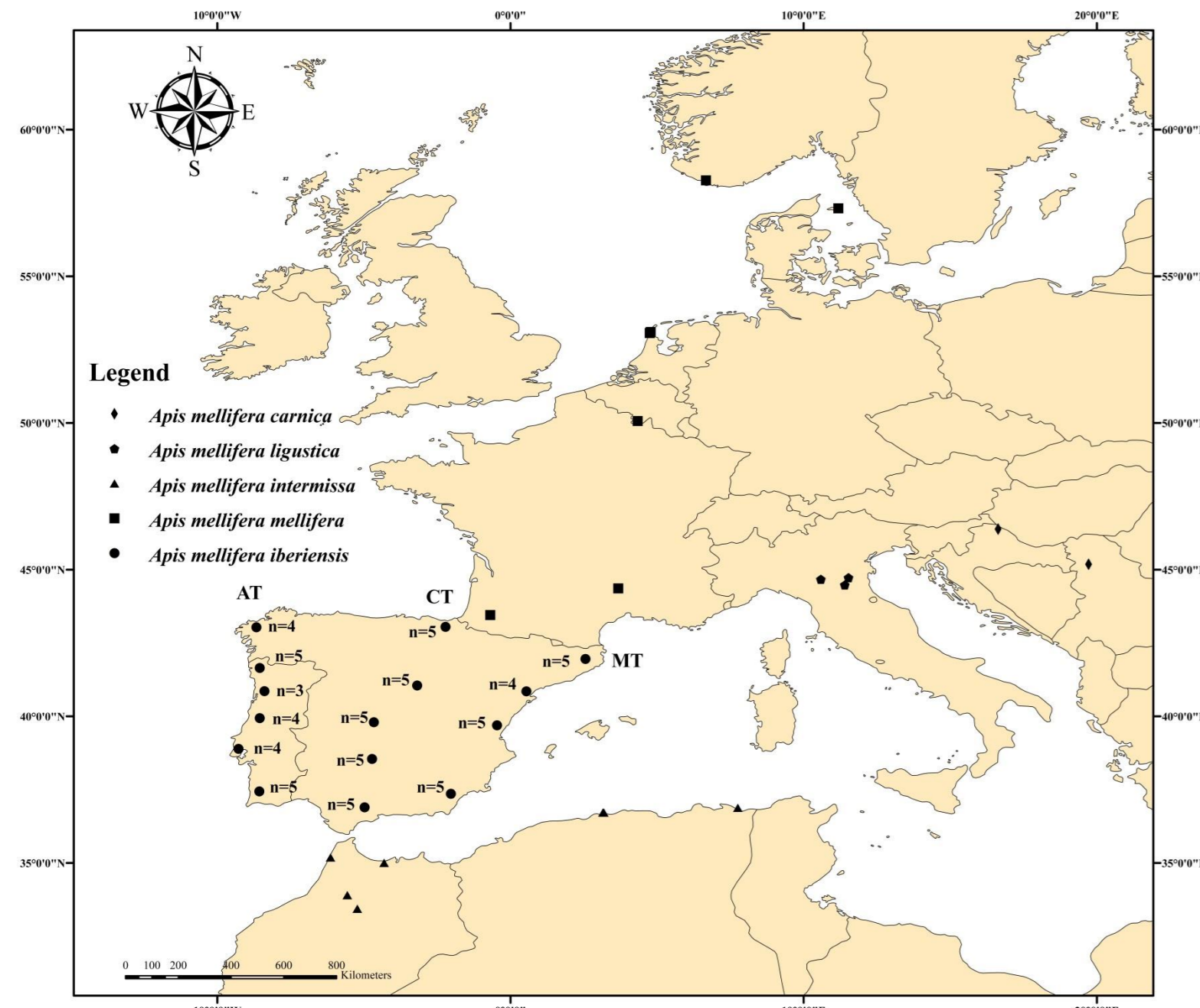


Figure 1 - Sampling locations.

Sequencing

- Libraries were generated using Illumina's TruSeq™ Sample Preparation kit.
- Sequencing was performed in Illumina HiSeq 2500 platform.

Quality filtering

- Adapters were removed using Trimmomatic and/or fastx_clipper.
- Poor quality bases (less than Phred-20) were removed from both the 5' and 3' ends of the unaligned reads.
- Reads below a length of 30 bases were discarded.

Mapping

The reads were aligned using Bowtie2 against the reference genome *A. m. ligustica*.

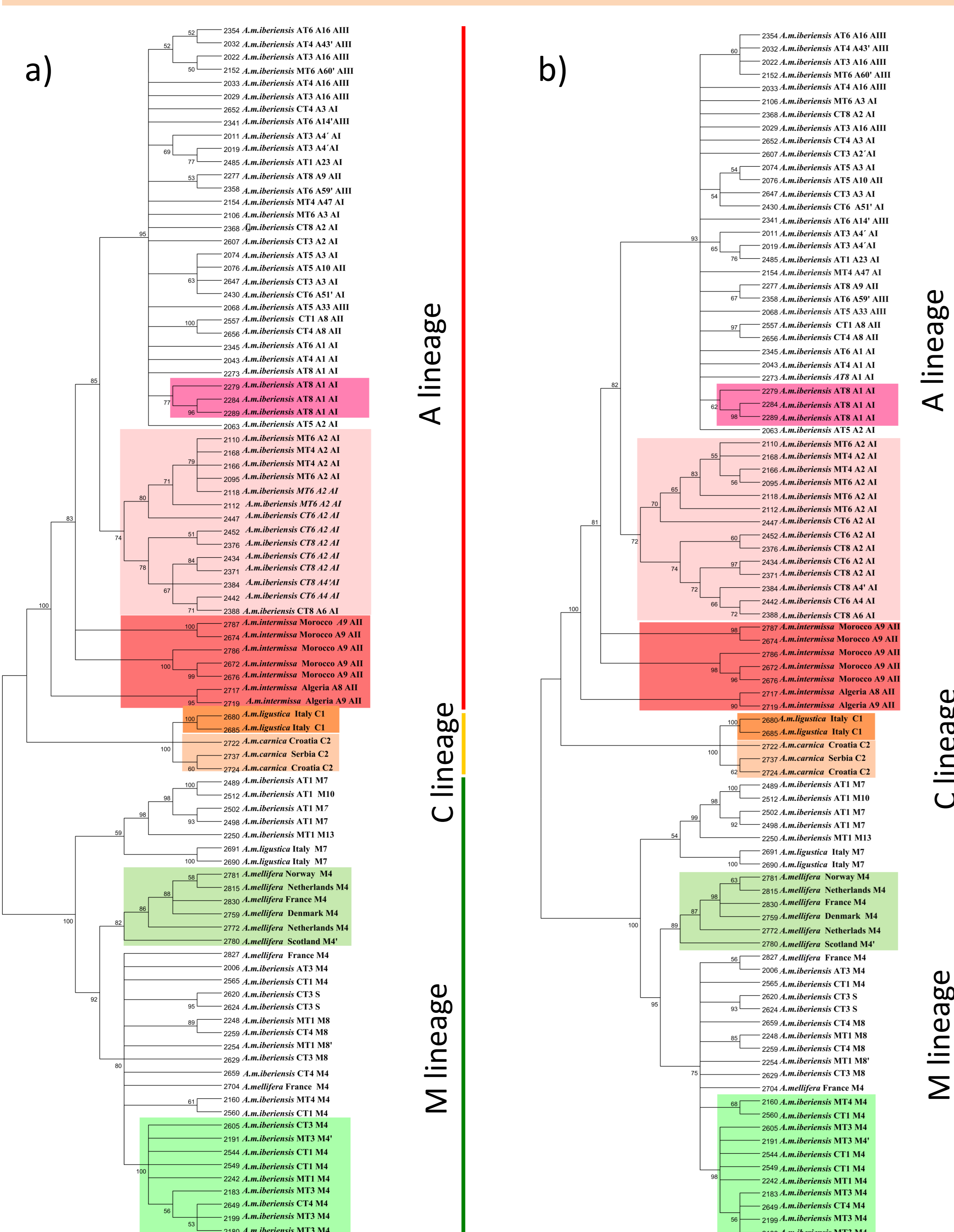
Variant calling

- The variant calling was done using Freebase. Samtools and VCFtools softwares were used to convert formats.

Phylogenetic analysis

- The mitogenomes were aligned using MAFFT.
- The GTR+G+I substitution model was used, as indicated by the jModelTest software.
- The phylogenetic trees were inferred using Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods implemented in MEGA software. The bootstrap consensus trees were inferred from 100 replicates. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. All positions containing gaps and missing data were eliminated. There were a total of 16055 positions in the final dataset.

Results



- The mitogenome data (Figure 2 a,b) clearly separate the 3 lineages (A, M, and C) identified by others using the tRNA^{leu}-cox2 intergenic region.
- There is no support for the three African sublineages (AI, AII, AIII) proposed by Franck et al. (2001).
- As expected, both A-derived and M-derived mitogenomes were found in Iberia, with the former clustering with the northern African *A. m. intermissa* and the latter with the western European *A. m. mellifera*.
- A great amount of variation was observed within haplotypes named according to the *Dral* test. For example, while most A2 samples form a coherent cluster, there are also samples dispersed in the tree.
- Iberian samples harboring haplotypes A8 and A9 do not cluster with *A. m. intermissa* samples of A8 and A9 ancestry.
- Iberian samples harboring M4 haplotype do not cluster with most *A. m. mellifera* samples of M4 ancestry.

Figure 2 - Unrooted phylogenetic tree inferred by NJ (a) and ML (b) methods. Bootstrap values are shown on the branches. AT, CT and MT refer to Iberian sampling transects (Figure 1). The 29 haplotype names (A1, A4', M4, M7, C1, C2 etc.) are based on the *Dral* test (Garnery et al. 1992). AI, AII, and AIII refer to the African sublineages (*sensu* Franck et al. 2001).

References

Franck P, Garnery L, Loiseau A, Oldroyd B, Hepburn H, Solignac M, Cornuet J. M. (2001). Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data, *Heredity* 52, 1119-1134.

Garnery L, Solignac M., Celebrano G., Cornuet J. M. (1992). A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. *Experientia* 49, 1016-1021.

Acknowledgments

We are deeply grateful to numerous people that collaborated in this study. Beekeepers from Spain and Portugal helped obtaining Iberian samples. Antonio Pajuelo provided the contacts of Spanish beekeepers. Margarida Neto, Andreia Brandão and Irene Muñoz collaborated in the sampling. Pilar de la Rúa, Wahida Loucif, Per Kryger, Bjorn Dahle, Lionel Garnery, Raffaele Dall' Olio, and Romée van der Zee provided the reference samples. Phillip San Miguel and Paul Parker sequenced the whole genomes and Rick Westerman performed the mapping (Purdue University). Dora Henriques and Julio Chávez-Galarza are supported by Fundação para a Ciência e Tecnologia through the scholarships SFRH/BD/84195/2012 and SFRH/BD/68682/2010, respectively. This research was funded by Fundação para a Ciência e Tecnologia and COMPETE/QREN/EU through the project PTDC/BIA-BEC/099640/2008.

