Maternal structure of Iberian honey bees inferred from whole

INSTITUTO POLITÉCNICO DE BRAGANÇA

mitochondrial genomes

TEXAS A&M



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).

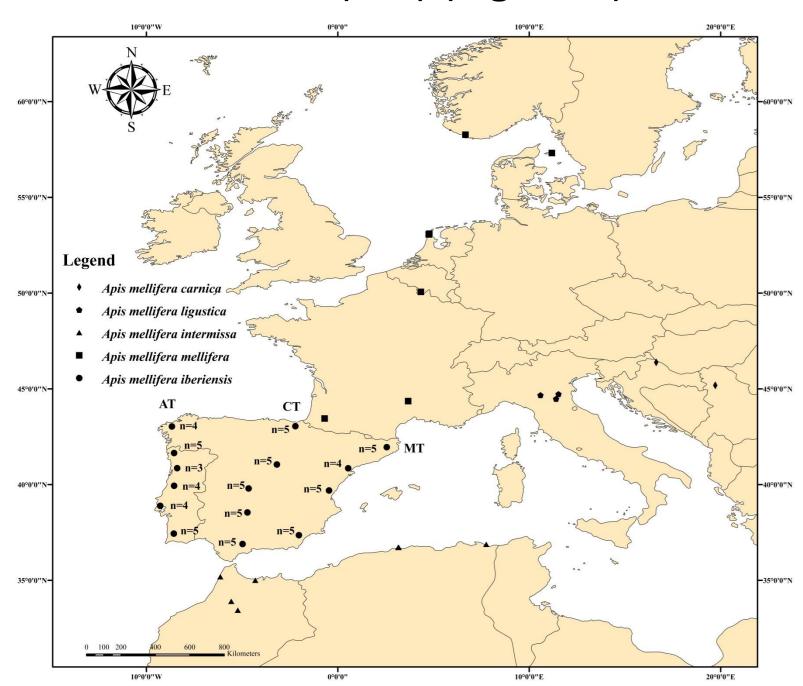


Figure 1 - Sampling locations.

Sequencing

- ➤ Libraries were generated using Illumina's TruSeqTM 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.

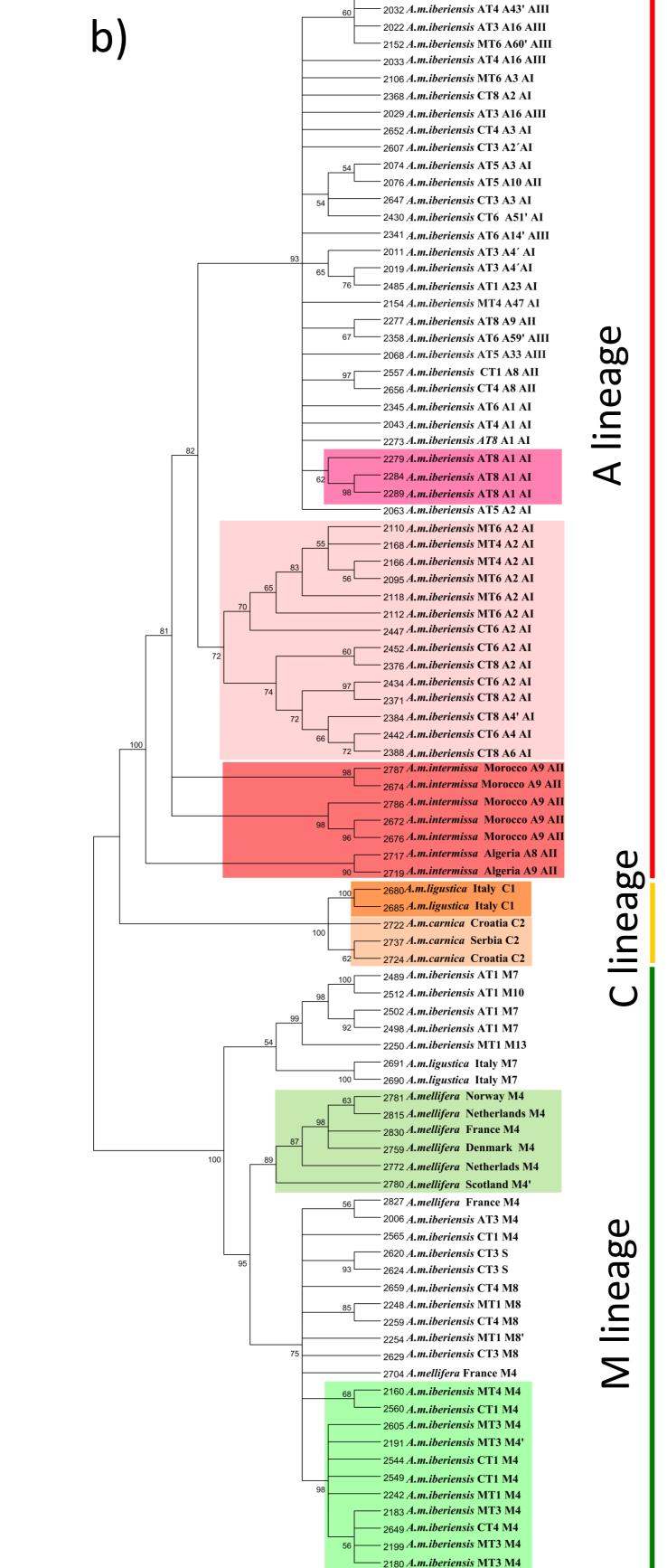
Results

- 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, AII) 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 *DraI* test (Garnery et al. 1992). AI, AII, and AIII refer to the African sublineages (*sensu* Franck et al. 2001).

2354 A.m.iberiensis AT6 A16 AIII 2032 *A.m.iberiensis* AT4 A43' AIII 2022 A.m.iberiensis AT3 A16 AIII **a**) 2152 A.m.iberiensis MT6 A60' AIII 2033 A.m.iberiensis AT4 A16 AIII 2029 A.m.iberiensis AT3 A16 AIII 2652 A.m.iberiensis CT4 A3 AI 2341 A.m.iberiensis AT6 A14'AIII 2011 A.m.iberiensis AT3 A4' AI - 2019 A.m.iberiensis AT3 A4'AI 2485 A.m.iberiensis AT1 A23 AI 2277 A.m.iberiensis AT8 A9 AII 2154 A.m.iberiensis MT4 A47 AI 2106 A.m.iberiensis MT6 A3 AI 2368 Alm.iberiensis CT8 A2 AI 2607 A.m.iberiensis CT3 A2 AI - 2074 A.m.iberiensis AT5 A3 AI 2076 A.m.iberiensis AT5 A10 AII 2647 A.m.iberiensis CT3 A3 AI 2430 A.m.iberiensis CT6 A51' AI 2068 A.m.iberiensis AT5 A33 AIII 2557 A.m.iberiensis CT1 A8 AII 2656 A.m.iberiensis CT4 A8 AII 2345 A.m.iberiensis AT6 A1 AI 2043 A.m.iberiensis AT4 A1 AI ₂₂₇₃ A.m.iberiensis AT8 A1 AI 2284 A.m.iberiensis AT8 A1 AI 2289 A.m.iberiensis AT8 A1 AI 2063 A.m.iberiensis AT5 A2 AI - 2110 A.m.iberiensis MT6 A2 AI 2168 A.m.iberiensis MT4 A2 AI 66 A.m.iberiensis MT4 A2 AI 2005 A.m.iberiensis MT6 A2 AI 2118 A.m.iberiensis MT6 A2 AI 2112 A.m.iberiensis MT6 A2 AI ₀₄₄₇ A.m.iberiensis CT6 A2 AI 2452 A.m.iberiensis CT6 A2 AI 2376 A.m.iberiensis CT8 A2 AI 2434 A.m.iberiensis CT6 A2 AI 2371 A.m.iberiensis CT8 A2 AI 2384 A.m.iberiensis CT8 A4'AI 2442 A.m.iberiensis CT6 A4 AI 2388 A.m.iberiensis CT8 A6 AI – ₂₆₇₄ A.m.intermissa Morocco A9 AII ₂₇₈₆ A.m.intermissa Morocco A9 AII 2672 A.m.intermissa Morocco A9 AII - 2676 A.m.intermissa Morocco A9 AII 2717 A.m.intermissa Algeria A8 AII - 2719 A.m.intermissa Algeria A9 AII 2680 A.m.ligustica Italy C1 2685 A.m.ligustica Italy C1 2722 A.m.carnica Croatia C2 - 2737 A.m.carnica Serbia C2 2724 A.m.carnica Croatia C2 2489 A.m.iberiensis AT1 M7 2512 A.m.iberiensis AT1 M10 2502 A.m.iberiensis AT1 M7 2498 A.m.iberiensis AT1 M7 2691 A.m.ligustica Italy M7 2690 A.m.ligustica Italy M7 2781 A.mellifera Norway M4 2830 A.mellifera France M4 2759 A.mellifera Denmark M4 2772 A.mellifera Netherlands M4 2827 A.mellifera France M4 2006 A.m.iberiensis AT3 M4 2565 A.m.iberiensis CT1 M4 - 2620 A.m.iberiensis CT3 S 2624 A.m.iberiensis CT3 S 2248 A.m.iberiensis MT1 M8 - 2259 A.m.iberiensis CT4 M8 M line 2254 A.m.iberiensis MT1 M8' 2629 A.m.iberiensis CT3 M8 2659 A.m.iberiensis CT4 M4 2704 A.mellifera France M4 2160 A.m.iberiensis MT4 M4 2560 A.m.iberiensis CT1 M4 2605 A.m.iberiensis CT3 M4 2191 A.m.iberiensis MT3 M4' 2544 A.m.iberiensis CT1 M4 2549 A.m.iberiensis CT1 M4 2242 A.m.iberiensis MT1 M4 2183 A.m.iberiensis MT3 M4 2649 A.m.iberiensis CT4 M4 2199 A.m.iberiensis MT3 M4

2180 A.m.iberiensis MT3 M4



– 2354 A.m.iberiensis AT6 A16 AIII

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.





