Mitochondrial and nuclear diversity of colonies of varying origins: contrasting patterns inferred from the intergenic tRNA<sub>leu</sub>-cox2 region and immune SNPs

Dora Henriques<sup>a</sup>, Ana R. Lopes<sup>a</sup>, Nor Chejanovsky<sup>b</sup>, Anne Dalmon<sup>c</sup>, Mariano Higes<sup>d</sup>, Clara Jabal-Uríel<sup>d</sup>, Yves Le Conte<sup>c</sup>, Maritza Reyes-Carreño<sup>e</sup>, Victoria Soroker<sup>b</sup>, Raquel Martín-Hernández<sup>e</sup> and M. Alice Pinto<sup>a</sup>

<sup>a</sup>Instituto Politécnico de Bragança, CIMO, Centro de Investigação de Montanha, Bragança, Portugal; <sup>b</sup>Agricultural Research Organization -Volcani Institute, Rishon LeZion, Israel; <sup>c</sup>Unité Abeilles et Environnement, INRAE, Avignon, France; <sup>d</sup>Laboratorio de Patología Apícola, Centro de Investigación Apícola y Agroambiental (CIAPA), Consejería de Agricultura de la Junta de Comunidades de Castilla-La Mancha, IRIAF, Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal, Marchamalo, Spain; <sup>e</sup>Instituto de Recursos Humanos para la Ciencia y la Tecnología (INCRECYT-FEDER), Fundación Parque Científico y Tecnológico de Castilla—La Mancha, Albacete, Spain

**ABSTRACT**

In this study, we gathered sequence data from the tRNA<sub>leu</sub>-cox2 intergenic mitochondrial (mtDNA) region concurrently with single nucleotide polymorphism (SNP) data from 91 loci of nuclear DNA (ncDNA). The data was obtained from 156 colonies sampled in six apiaries from four countries. The full dataset was analysed and discussed for genetic patterns with a focus on cytonuclear diversity and admixture levels.

The honey bee, *Apis mellifera* L., natural variation is represented by over 24 subspecies belonging to four lineages: African (A), Middle eastern (O), Western European (M), and Eastern European (C) (Ruttner, 1988). The long-term sustainability of *A. mellifera* is largely dependent on locally adapted populations. However, during the last decades, globalization stimulated the introduction of commercial breeds in many regions of the *A. mellifera* native range which have hybridized with the local populations (De la Rúa et al., 2009).

In this work, the cytonuclear background of honey bee colonies sampled in four countries was assessed using the tRNA<sub>leu</sub>-cox2 intergenic mitochondrial region and 91 ncDNA immune SNPs. A total of 156 samples were collected from single apiaries in Portugal, Spain, France, and Israel (Figure 1; Supplementary Table S1). In France, the samples were from Avignon (N = 24) and the island of Ouessant (N = 12). In Portugal, they were from Bragança (N = 23) and the island of São Miguel, Azores (N = 23). Spain and Israel were represented by samples collected from an apiary in Marchamalo (N = 34) and in Beit Dagan (N = 40), respectively. Each sample was composed of a pool of 10 workers collected from a single colony. Total DNA was extracted from pools of 20 forelegs using Ron’s Tissue DNA Mini Kit (Bioron).

The tRNA<sub>leu</sub>-cox2 region was PCR-amplified with the primers E2 and H2 using the reaction conditions of Garnery et al. (1993). PCR products were Sanger-sequenced by STABVIDA (Portugal). The alignment of the sequences was performed with MEGA 6.06 (Tamura et al., 2013). Sequences of two samples were discarded from the mtDNA analysis because they failed quality standards.

Twenty haplotypes (27 variants) were identified in the 154 samples, of which 11 (13 variants) belonged to lineage A, seven (10 variants) to lineage M, and two (four variants) to lineage C (Supplementary Table S1). One haplotype (M81; accession number MZ767556), as well as eight variants, are reported for the first time (Supplementary Table S2). This novel variation was detected in Portugal and Spain, further supporting the view of Iberia as a hotspot of
Most (87%) colonies sampled in Portugal belonged to lineage A, with haplotypes A1 (26%) and A11b (13%) predominating in the mainland and A14 being the single haplotype detected in the apiary of São Miguel (Figure 1; Supplementary Table S1). In contrast, in Spain, 82% of the colonies belonged to lineage M of which 53% carried M4, the most common haplotype in northeastern Iberia (Chávez-Galarza et al., 2017). These results are consistent with earlier maternal surveys in Iberia, reporting a cline formed by haplotypes of M-lineage in the northeastern part of Spain and A-lineage in Portugal and southwestern part of Spain and virtually no C-lineage introgression in the native Apis mellifera iberiensis (Chávez-Galarza et al., 2017; Garnery et al., 1998; Miguel et al., 2007). In São Miguel, despite previously reported high levels of C-lineage introgression, a single A14 haplotype (the novel A14c variant) was detected in the sampled colonies and this is the most common and widespread African haplotype in the Azores (Ferreira et al., 2020).

In France, all colonies from the Avignon apiary were of C-lineage ancestry (91% C2 and 9% C1), contrasting with the apiary of Ouessant where all colonies carried M-lineage mitochondria (58% M17e and 42% M4). These results are consistent with reports of high C-lineage introgression into the native Apis mellifera mellifera in many regions of mainland France and with Ouessant acting as a refuge for this subspecies (Miguel et al., 2007; Pinto et al., 2014; Requier et al., 2019). As for Avignon, in the apiary of Beit Dagan, Israel, all colonies were of C-lineage ancestry, with haplotype C2 also predominating (95%) over C1 (5%). While Apis mellifera syriaca (lineage O) is the native subspecies of Israel (Ruttner, 1988), our data is consistent with the claim that it has largely been replaced by commercial stocks of C-lineage ancestry (Soroker et al., 2018). However, according to anecdotal reports, the C-lineage A. m. ligustica has been favoured by Israeli beekeepers (Soroker et al., 2018) and the finding of mostly C2 in Beit Dagan suggests that most colonies in this apiary likely derived from Apis mellifera carnica or Apis mellifera caucasia imports, similar to Avignon. This assertion stems from the fact that while C2 is the predominant haplotype in the distributional ranges of those two subspecies, A. m. ligustica is characterized by harbouring haplotype C1 (Franck et al., 2000; Kandemir et al., 2006; Muñoz et al., 2009; Özdil et al., 2009; Pinto et al., 2014). We caution, however, against any over-interpretation of the data presented herein as they were collected from single apiaries and might therefore suffer from biases linked to management practices such as colony splitting or queen rearing.

The 156 pooled samples were genotyped in the MassARRAY® MALDI-TOF platform using a recently developed assay containing 91 immune SNPs (Henriques et al., 2021). The membership proportions (Q-values) were inferred from these SNPs with ADMIXTURE (Alexander et al., 2009) using the same settings and reference populations of Henriques et al. (2021). GENEALEX 6.50 (Peakall & Smouse, 2006) was used to estimate the mean number of
haplotypes, the effective number of haplotypes, and unbiased diversity whereas allelic richness was calculated with ADZE v1 (Szpiech et al., 2008).

In the ADMIXTURE analysis, the most likely number of clusters was K = 3 (cross-validation test), coinciding with lineages A, M and C detected with mtDNA in the four countries. At K = 4, it is possible to distinguish not only lineages A and C (the purple and orange cluster, respectively), but also the two M-lineages subspecies A. m. mellifera (green cluster) and A. m. iberiensis (blue cluster) represented in Figure 1 (see Supplementary Table S1 and Figure S5 in Henriques et al., 2021, for details).

The membership proportions inferred from the SNP data in ADMIXTURE are concordant with the mtDNA data in what concerns C-lineage introgression. Colonies from Portugal, Spain, and Ouessant were virtually free of C-lineage cytonuclear introgression. In contrast, colonies from Avignon and Beit Dagan, which only harboured C-lineage haplotypes, exhibited membership proportions into the C-lineage cluster >32% (Figure 1, Supplementary Table S1). However, these colonies also displayed higher levels of genetic diversity for ncDNA but not for mtDNA, which contrasts with the other mainland colonies from Iberia showing the opposite trend (Table 1).

The finding of elevated nuclear diversity in admixed populations relative to parental populations is consistent with predictions from population genetics theory and has been confirmed by empirical data in honey bee studies (Harpur et al., 2012). Of note is that this finding is also a by-product of the biased nature of the SNP assay employed here (Henriques et al., 2021), which was tailored to include immune-related SNPs segregating between the two most divergent lineages: M and C. This means that both alleles will co-occur in admixed populations whereas only one alternative allele is expected to be found in parental populations.

Despite the limited geographical sampling undertaken in each country, this study further corroborates the (i) genetic integrity of A. m. mellifera in Ouessant and A. m. iberiensis, (ii) high levels of mitochondrial diversity in Iberia, supported by the detection of a novel haplotype and many variants, and (iii) presence of C-lineage introgression in the native range of M-lineage (mainland France) and O-lineage (Israel) subspecies. Moreover, this study detected contrasting levels of diversity in the cytonuclear compartments. The higher nuclear diversity in the colonies from mainland France and Israel can be explained by admixture. While admixture may lead to increased diversity (Harpur et al., 2012), as measured by the classical parameters (Table 1), it may also lead to disruption of gene complexes that have been fine-tuned by selection acting on local variation across micro-evolutionary time scale (De la Rúa et al., 2013). This type of diversity may be important for the populations to face the challenges of a rapidly changing world.

Acknowledgments

We are deeply indebted to Frank Aguiar, Luís Silva, Edgardo Melo, João Martins, João Melo, Manuel Moura, Manuel Viveiros, and Ricardo Sousa from "Direção Regional da Agricultura e Desenvolvimento Rural dos Açores" (Portugal), and to Laura Garreau, Laurent Maugis, Pascale Sauvage and Jacques Kermagoret, from “Association Conservatoire de l’Abeille Noire Bretonne” (France), for sampling the apiaries in São Miguel, Santa Maria, and Ouessant islands. Genotyping was outsourced to the Epigenetics and Genotyping laboratory, Central Unit for Research in Medicine (UCIM), University of Valencia, Spain. Data analyses were performed using computational resources at the Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança.

Disclosure statement

No potential competing interest was reported by the authors.

Funding

This research was funded through the projects BEEHAPPY (POCI-01-0145-FEDER-029871, FCT and COMPETE/QREN/EU) and BEEHEAL. BEEHEAL was funded by the ARIMNet2 2016 Call by the following agencies: INIA (Spain), MOARD (Israel), ANR (France) and FCT (Portugal). ARIMNet2 (ERA-NET) received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 618127. Ana Rita Lopes is supported by a PhD scholarship (SFRH/BD/143627/2019) from the Foundation for Science and Technology (FCT), Portugal. FCT provided financial support by national funds (FCT/MCTES) to CIMO (UIDB/00690/2020).

ORCID

Dora Henriques http://orcid.org/0000-0001-7530-682X
Ana R. Lopes http://orcid.org/0000-0002-1888-2262
Nor Chejanovsky http://orcid.org/0000-0002-6455-6521
Anne Dalmon http://orcid.org/0000-0001-5044-4724
Mariano Higes http://orcid.org/0000-0001-7338-5851

Table 1. Diversity estimates for mtDNA (left column) and SNPs (right column) for each sampled apiary from Portugal (PT), Spain (SP), France (FR) and Israel (IL). Na - mean number of alleles; Ne - number of effective alleles; uh - unbiased diversity; Ar - allelic richness.

| Location       | Na   | Ne   | uh   | Ar   |
|----------------|------|------|------|------|
| São Miguel (PT)| 3.00 | 1.66 | 1.00 | 1.25 |
| Bragança (PT)  | 12.00| 7.90 | 2.22 | 0.89 |
| Marchamalo (SP)| 9.00 | 3.73 | 1.23 | 0.74 |
| Ouessant (FR)  | 2.00 | 1.17 | 1.17 | 0.51 |
| Avignon (FR)   | 4.00 | 1.80 | 1.30 | 0.64 |
| Beit Dagan (IL)| 2.00 | 1.60 | 1.60 | 0.10 |
References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Chávez-Galarza, J., Garnery, L., Henríques, D., Neves, C. J., Loucif-Ayd, W., Jonhston, J. S., & Pinto, M. A. (2017). Mitochondrial DNA variation of *Apis mellifera iberiensis*: Further insights from a large-scale study using sequence data of the tRNA leu-cox2 intergenic region. *Apidologie*, 48(4), 533–544. https://doi.org/10.1007/s13592-017-0498-2

De la Rúa, P., Jaffé, R., Dall’Olio, R., Muñoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honeybees. *Apidologie*, 40(3), 263–284. https://doi.org/10.1051/apido/2009027

De la Rúa, P., Jaffé, R., Muñoz, I., Serrano, J., Moritz, R. F., & Kraus, F. B. (2013). *Conserving genetic diversity in the honeybee: Comments on Harpur et al.* (2012). Wiley Online Library.

Ferreira, H., Henriquez, D., Neves, C. J., Machado, C. A., Azevedo, J. C., Francoy, T. M., & Pinto, M. A. (2020). Historical and contemporaneous human-mediated processes left a strong genetic signature on honey bee populations from the Macaronesian archipelago of the Azores. *Apidologie*, 51(3), 316–328. https://doi.org/10.1007/s13592-019-00720-w

Franck, P., Garnery, L., Celebrano, G., Solignac, M., & Cornuet, J. M. (2000). Hybrid origins of honeybees from Italy (*Apis mellifera ligustica*) and Sicily (*A. m. sicula*). *Molecular Ecology*, 9(7), 907–921. https://doi.org/10.1046/j.1365-294x.2000.00945.x

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.-M., & Solignac, M. (1998). Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*) 1. *Genetics Selection Evolution*, 30(Suppl 1), S31–S47. https://doi.org/10.1186/1297-9686-30-51-S31

Garnery, L., Solignac, M., Celebrano, G., & Cornuet, J.-M. (1993). A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera L. Experientia*, 49(11), 1016–1021. https://doi.org/10.1007/BF02125651

Harpur, B. A., Minaei, S., Kent, C. F., & Zayed, A. (2012). Management increases genetic diversity of honey bees via admixture. *Molecular Ecology*, 21(18), 4414–4421. https://doi.org/10.1111/j.1365-294x.2012.05614.x

Henriques, D., Lopes, A. R., Chejanovsky, N., Dalmon, A., Higes, M., Jabal-Uriel, C., Le Conte, Y., Reyes-Carreño, M., Soroker, V., Martín-Hernández, R., & Pinto, M. A. (2021). A SNP assay for assessing diversity in immune genes in the honey bee (*Apis mellifera L.*). *Scientific Reports*, 11(1), 15317. https://doi.org/10.1038/s41598-021-94833-x

Kandemir, I., Kence, M., Sheppard, W. S., & Kence, A. (2006). Mitochondrial DNA variation in honey bee (*Apis mellifera L.*) populations from Turkey. *Journal of Apicultural Research*, 45(1), 33–38. https://doi.org/10.1080/00218839.2006.11101310

Miguel, I., Iriondo, M., Garnery, L., Sheppard, W. S., & Estomba, A. (2007). Gene flow within the M evolutionary lineage of *Apis mellifera*: Role of the Pyrenees, isolation by distance and post-glacial re-colonization routes in the western. *Apidologie*, 38(2), 141–155. https://doi.org/10.1051/apido:2007007

Muñoz, I., Dall’Olio, R., Lodesani, M., & De la Rúa, P. (2009). Population genetic structure of coastal Croatian honeybees (*Apis mellifera carnica*). *Apidologie*, 40(6), 617–626. https://doi.org/10.1051/apido/2009041

Özdil, F., Yildiz, M. A., & Hall, H. G. (2009). Molecular characterization of Turkish honey bee populations (*Apis mellifera*) inferred from mitochondrial DNA RFLP and sequence results. *Apidologie*, 40(5), 570–576. https://doi.org/10.1051/apido/2009032

Peacock, R., & Smouse, P. E. (2006). *GENALEX 6*: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pinto, M. A., Henríques, D., Chávez-Galarza, J., Kyger, P., Garnery, L., van der Zee, R., Dahlé, B., Soland-Reckeweg, G., de la Rúa, P., Dall’ Olio, R., Carreck, N. L., & Johnston, J. S. (2014). Genetic integrity of the Dark European honey bee (*Apis mellifera*) from protected populations: A genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research*, 53(2), 269–278. https://doi.org/10.3896/IBRA.1.53.2.08

Pinto, M. A., Henríques, D., Neto, M., Guedes, H., Muñoz, I., Azevedo, J. C., & De la Rúa, P. (2013). Maternal diversity patterns of Ibero-Atlantic populations reveal further complexity of Iberian honeybees. *Apidologie*, 44(4), 430–439. https://doi.org/10.1051/apido/2013592-013-0192-y

Requier, F., Garnery, L., Kohl, P. L., Njou, H. K., Pirk, C. W., Crewe, R. M., & Steffan-Dewenter, I. (2019). The conservation of native honey bees is crucial. *Trends in Ecology & Evolution*, 34(9), 789–798. https://doi.org/10.1016/j.tree.2019.04.008

Ruttner, F. (1988). *Biogeography and taxonomy of honeybees. Springer*, Germany.

Soroker, V., Yossi, S., & Chejanovsky, N. (2018). Apiculture in Israel. In Maria Bouga (Ed.), *Asian beekeeping in the 21st century* (pp. 95–109). Springer.

Szpiech, Z. A., Jakobsson, M., & Rosenberg, N. A. (2008). ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics (Oxford, England)*, 24(21), 2498–2504. https://doi.org/10.1093/bioinformatics/btn478

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. https://doi.org/10.1093/molbev/ms3197