

GENOME-WIDE SCANS DETECTED SIGNATURES OF SELECTION IN GENES RELATED WITH VISION, XENOBIOTIC METABOLISM, AND IMMUNITY IN THE IBERIAN HONEY BEE GENOME



Chávez-Galarza J¹, Henriques D¹, Johnston JS², Rufino J¹, Pinto MA¹

¹Mountain Research Centre (CIMO), Polytechnic Institute of Bragança, Campus de Sta. Apolónia, Apartado 1172, 5301-855 Bragança, Portugal.

²Dept. of Entomology, Texas A&M University, College Station, Texas 77843-2475, USA.

Correspondence: Maria Alice Pinto, CIMO, Instituto Politécnico de Bragança, Portugal. E-mail: apinto@ipb.pt

Introduction

The Iberian Peninsula is considered as one of the most important Pleistocene glacial refugia in the European continent. Several phylogeographical studies of flora and fauna describe the Iberian Peninsula as a cradle of genetic differentiation and species repository, explained by its diverse habitats with varied microclimates, which have influenced the demographic processes and local adaptation of many species, and Iberian honey bee is no exception. Currently, population genomics has made possible the study of local adaptation by using methodologies designed to dissect neutral from adaptive variation. Hence, genome-wide scan approaches to detect signatures of selection have become popular tools to study local adaptation at the molecular level and single nucleotide polymorphisms (SNPs) are seemingly the most appropriate marker for that endeavor. The objective of this work is to identify genomic regions under selection, using SNPs, and to determine the putative function of genes located in these regions.

Analysis of Selection

Candidate loci under selection were identified using four F_{st} -based outlier approaches: LOSITAN (Antão *et al.* 2008), ARLEQUIN (Excoffier *et al.* 2009), BAYESFST (Beaumont and Balding 2004), and BAYESCAN (Foll and Gaggiotti 2008). In addition to F_{st} -based outlier approaches, univariate logistic regression analysis was performed by the matSAM Program (Joost *et al.* 2007) to determine the degree of association between the frequencies of each allele and the values of the environmental variables.

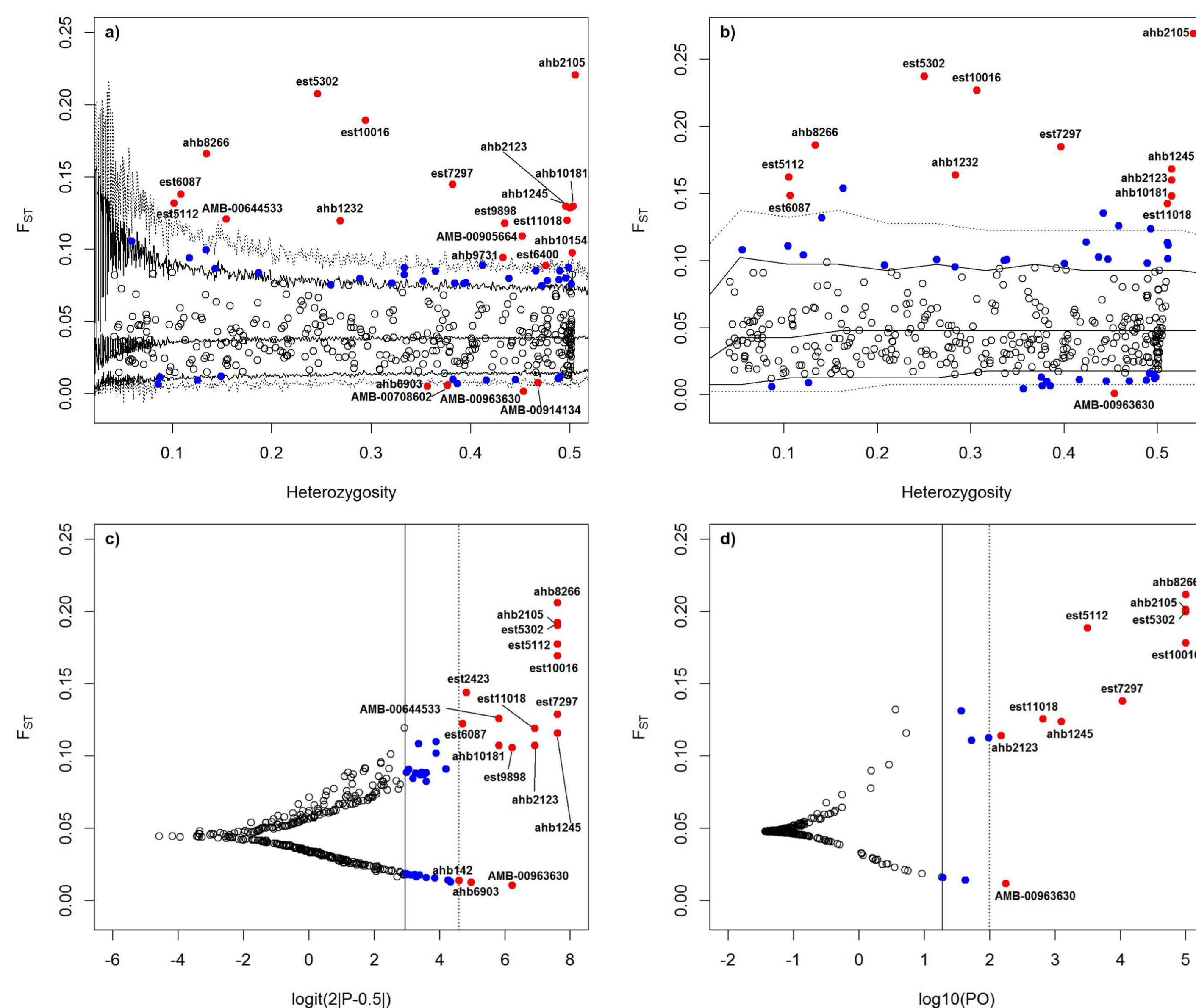


Fig. 2(a) LOSITAN, (b) ARLEQUIN, (c) BAYESFST, and (d) BAYESCAN plots showing loci under selection. Solid and dashed lines represent 95% and 99% confidence intervals, respectively. The middle line in (a) and (b) depicts the median value. Outlier directional (upper) and balancing (lower) SNP loci with P -value < 0.005 and posterior probability > 0.99 and are labeled in each plot

References

- Antão T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: A workbench to detect molecular adaptation based on F_{st} -outlier method. *BMC Bioinformatics* 2008, 9: 323.
- Beaumont MA, Balding DJ. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology* 2004, 13: 969-980.
- Excoffier L, Hofer T, Foll M. Detecting loci under selection in a hierarchically structured population. *Heredity* 2009, 103: 285-298.
- Foll M, Gaggiotti OE. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 2008, 180: 977-993.
- Joost S, Bonin A, Bruford MW *et al.* A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology* 2007, 16: 3955-3969.

ACKNOWLEDGMENTS

We owe special thanks to numerous beekeepers from Spain and Portugal for their valuable help in obtaining samples. DNA extractions and SNP genotyping was performed by Colette Abbey, who was deeply committed to this project. Special thanks to Dr. Clare Gill for providing the facilities for SNPs genotyping and for assistance with analysis. Julio Chávez-Galarza is supported by Fundação para a Ciência e Tecnologia through the scholarship SFRH/BD/68682/2010. This research was funded by Fundação para a Ciência e Tecnologia and COMPETE/QREN/EU through the project PTDC/BIA-BEC/099640/2008.

Sampling

A total of 711 Iberian honey bee individuals (each representing a single colony) was collected in 2010 across three North-South transects in the Iberian Peninsula (Fig. 1).

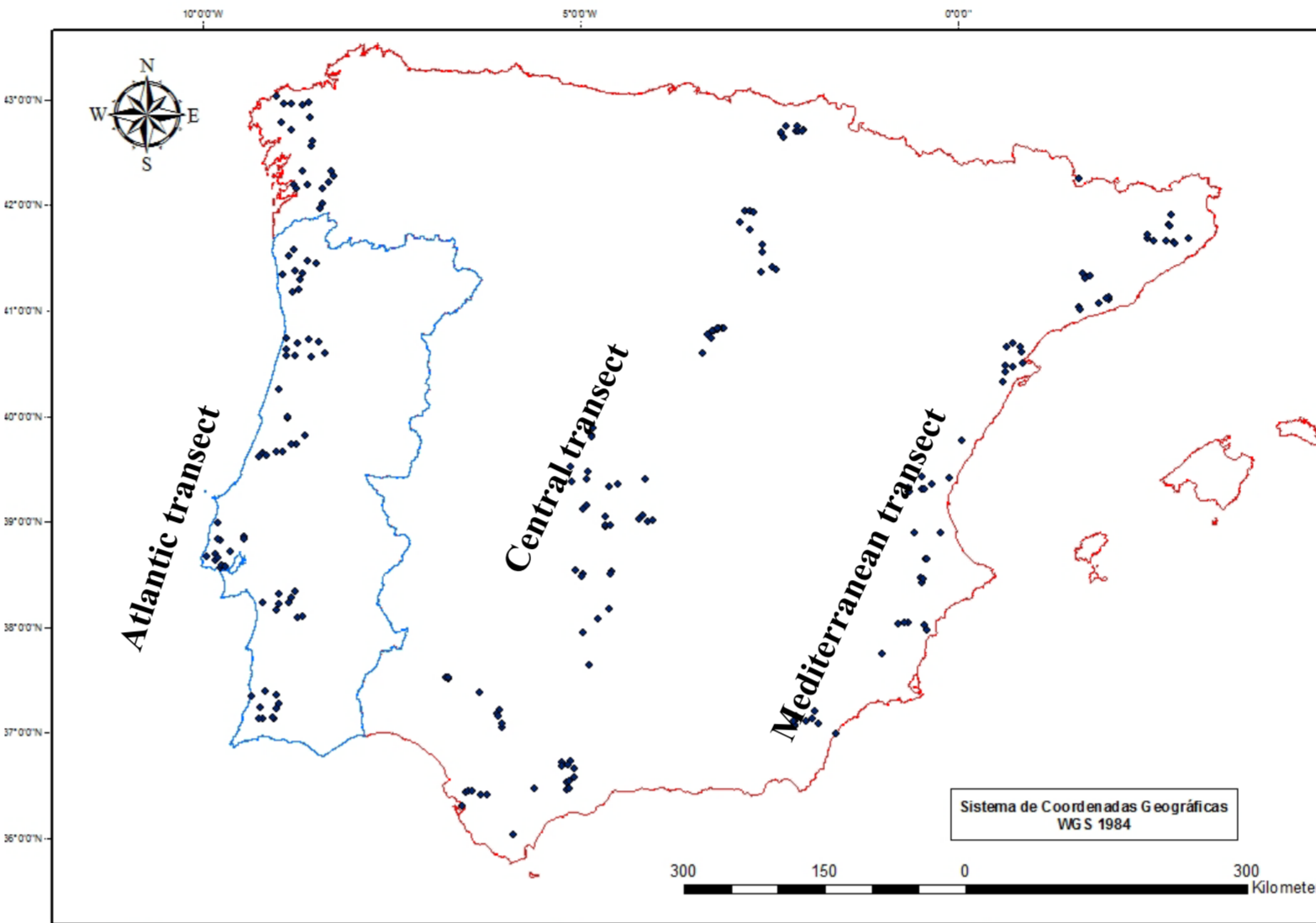


Fig. 1 Sampling transects. Each dot represents an apiary. Three colonies were sampled in each apiary.

Environmental Data

Environmental data were obtained for each apiary. This dataset consisted of Altitude (WorldClim database), Precipitation, Minimum temperature, Mean temperature, Maximum temperature and Cloud cover (Climatic Research Unit), Relative humidity and Insolation (OPENEI), and Land cover (European Environment Agency). All climatic data were extracted yearly, seasonally and monthly.

Genotyping

Individuals were genotyped for a panel of 1536 SNP's with Illumina BeadStation 500G using a custom Oligo Pool Assay. Individuals were scored using Illumina's Genome Studio software. The final number of loci was 383, after removing monomorphic loci (cutoff 2%) and failed genotype calls.

Results

Table 1 Genomic information for the SNP loci related to xenobiotic detoxification, vision, immunity obtained from NCBI, BEEBASE, and FLYBASE.

SNP code	Linkage group	Gene product	Selection	matSAM
Detoxification				
ahb1245	1	Gst-mic2 - microsomal glutathione S-transferase 2	Directional	Long, Prec, Tmin, Ins
est5302	6	UDP-glucosyltransferase (UDP-glucosyltransferase 35b, Ugt35b) [§]	Directional	Lat, Prec, Tmean, Tmax, Cld
est10016	13	Cytochrome P450 - CYP6AS7 (Cyp6a14) [§]	Directional	Lat, Prec, Tmean, Cld, Ins
Vision				
ahb8266	6	Teneurin 3 - like isoform 1 (Tenascin major, Ten-m) [§]	Directional	Lat, Prec, Tmean, Tmax, Cld, Ins
est7297	10	15-hydroxyprostaglandin dehydrogenase [NAD ⁺]-like (Photoreceptor dehydrogenase, Pdh) [§]	Directional	Long, Lat, Prec, Ins
est2423	2	Retinol dehydrogenase 11-like	Directional	Ins
ahb142	1	Sema 1 - Semaphorin 1A (Sema-1a) [§]	Balancing	
ahb4188	14	Blop - blue-sensitive opsin (Rhodopsin, Rh5) [§]	Balancing	
Immunity				
est11018	15	NimC2 - nimrod C2 (nimrod C2, nimC2) [§]	Directional	
ahb6903	4	Dscam - Down syndrome cell adhesion molecule (Down syndrome cell adhesion molecule, Dscam) [§]	Balancing	

[§]Names and/or symbols within parentheses correspond to orthologous genes of *Drosophila melanogaster* as in FLYBASE Longitude(Long), Latitude (Lat), precipitation (Prec), Minimum temperature (Tmin), Mean temperature (Tmean), Maximum temperature (Tmax), Cloud cover (Cld) and Insolation (Ins).

- Over 57 loci under selection were identified by LOSITAN, 49 ARLEQUIN, 41 BAYESFST and 17 BAYESCAN (Fig 2). Analysis with MatSAM identified 33 loci associated to environmental variables (Table 1).
- Over 10 outlier SNPs presented the strong signal of selection (ahb1245, est5302, est5112, ahb8266, ahb2123, est7297, ahb2105, est10016, est11018, and AMB-00963630).
- Precipitation was the environmental variable that presented more associations with outlier SNP.
- Outlier SNPs were located in the 16 linkage groups from honey bee genome marking genes with a diverse array of putative functions including signaling, structural, metabolism, regulation, transport, and immunity.
- Three biological processes (xenobiotic detoxification, vision and immunity) were represented by genes exhibiting a signature of selection (Table 1).



FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR