Apidologie 38 (2007) 141–155

DOI: 10.1051/apido:2007007

Original article

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 Europe*

Irati MIGUEL[a], Mikel IRIONDO[a], Lionel GARNERY[b], Walter S. SHEPPARD[c], Andone ESTONBA[a]

[a] Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and Technology, University of the Basque Country, PO Box 644, 48080 Bilbao, Spain
[b] Laboratoire Population, Génétique et Évolution, CNRS, avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France
[c] Department of Entomology, Washington State University, Pullman, WA 99164-6382, USA

Received 30 December 2005 – Revised 28 July 2006 – Accepted 2 August 2006

Abstract – We present a population genetic study focused on the two subspecies of the M evolutionary lineage, *A. m. mellifera and A. m. iberiensis*. Nuclear and mtDNA variation was analysed in 27 bee populations from the Iberian Peninsula, France and Belgium. Microsatellite data provides compelling evidence of a barrier to neutral gene flow at the Pyrenees. In addition, they suggest isolation by distance between populations of the M lineage. Mitochondrial data support the hypothesis that the Iberian Peninsula served as glacial refugia for the honeybees of western Europe. They show two paths of post-glacial re-colonization in the extremes of the Pyrenees and suggest that the western path was more significant in the post-glacial re-colonization process. Thus, we report here on three main factors for *mellifera* and *iberiensis* subspecies differentiation: the Pyrenean barrier, isolation by distance and the post-glacial re-colonization process.

honeybees / subspecies / mtDNA / microsatellite / population genetics

1. INTRODUCTION

In the honey bee, *Apis mellifera* L., over two dozen morphological subspecies have been described through a range that includes Africa, Europe and western Asia (Ruttner, 1988; Sheppard et al., 1999; Sheppard and Meixner, 2003). Based on morphological affinities and paleoclimatic considerations, these subspecies have been grouped into five branches presumably reflecting evolutionary lineages: A (African), C ("*carnica*" – subspecies east and south of the Alps including those along the northern Mediterranean), M ("*mellifera*" – subspecies of western Europe), O (Oriental – subspecies from the eastern end of the range of the species) and Y (Yemenitica from Ethiopia; Ruttner et al., 1978; Ruttner, 1988; Franck et al., 2000a, 2001). The M lineage, on which this study is based, includes two subspecies: *A. m. mellifera*, extending from France through northern Europe eastward to the Ural mountains and *A. m. iberiensis* (Engel, 1999) (formerly *A. m. iberica*), a subspecies located in the Iberian Peninsula. These morphological subspecies have also been the subject of a number of DNA studies (Garnery et al., 1992; Estoup et al., 1995; Arias and Sheppard, 1996; Franck et al., 2000a).

Some questions remain unresolved regarding M branch subspecies. First, the Pyrenees

Corresponding author: I. Miguel, ggbmimai@lg.ehu.es
* Manuscript editor: Stefan Fuchs

have been considered to be a major obstacle to gene flow for the honey bees of western Europe and, therefore, the geographic boundary between *A. m. mellifera* and *A. m. iberiensis*. However, the barrier effect of the Pyrenees to honey bee gene flow has not been demonstrated. Moreover, multiple works show different genoclines from the south of the Iberian Peninsula to the north of Europe (Cornuet, 1982; Smith and Glenn, 1995; Garnery et al., 1995, 1998a; Franck et al., 1998). In addition, based on microsatellite data, Garnery et al. (1998b) illustrate an uncertain taxonomic attribution for some northern Iberian populations. Among them, the honeybee population from San Sebastian (at the Iberian western end of the Pyrenees) is not clearly classified within *A. m. iberiensis* but appears to be intermediate between Iberian and French populations.

Second, it is assumed that the Iberian Peninsula was a refugia for the western European honey bee during the last ice age (Ruttner, 1952, 1988), and recent works suggest later migration to the north in the warm post-glacial period (Garnery et al., 1998a,b; Franck et al., 1998, 2000b). Nevertheless, this post-glacial re-colonization process and the genetic legacy of the ice ages are not well known.

To address these questions, this study focuses on the genetic variability of honey bee populations in western Europe, emphasizing the contact region between *A. m. mellifera* and *A. m. iberiensis*. In all, 1398 colonies were analysed for variation in 10 microsatellite loci and the COI-COII intergenic region of mtDNA. These colonies were distributed among 27 populations in the Iberian Peninsula, France and Belgium.

2. MATERIALS AND METHODS

2.1. Sampling

A total of 1014 colonies (one bee per hive) from 19 populations of *A. mellifera* were sampled (Fig. 1). Samples were collected in ethanol from points in Portugal (Evora, EVO; Lisbon, LIS), Spain (Sevilla, SEV; Cáceres, CAC; Granada, GRA; Águilas, AGU; Alicante, ALI; Valencia, VAL; Tortosa, TOR; Badalona, BAD; Gerona, GER; Segovia, SEG; Zaragoza, ZAR; Bizkaia, BIZ; Araba, ARA; Oñati, OÑA; Goizueta, GOI; and Erronkari, ERR), and France (Foix, FOI). In addition, previously published data by Garnery et al. (1998b) were included in the analyses: the genotypes for 7 microsatellite loci in 384 samples from eight populations of France (Baiona, BAI; Sabres, SAB; Saintes, SAI; Ouessant, OUE; Valenciennes, VAN; Annecy, ANN; and Montfavet, MON) and Belgium (Chimay, CHI). To complete the microsatellites panel from 7 to 10, Ap55, Ap81 and Ap66 microsatellite loci were genotyped in these 384 samples.

2.2. DNA Extraction

DNA was extracted from the homogenised head and thorax of each bee using either a phenol-chloroform extraction method, a chelex-based protocol (Garnery et al., 1993) or the Perfect gDNA Blood Mini Kit (Eppendorf). In this last method, a series of modifications to the manufacturer's protocol were made to adjust the extraction to the type of material used: the head and thorax was homogenised in 200 μL of H_2O, the amount of proteinase K was doubled and incubation times were adjusted.

2.3. Microsatellite loci

We present here a new combination of primers that allows the analysis of 10 microsatellites loci based on two PCR multiplexes. In the multiplex named *Apis*1: B124, A88, A28, A24 (Estoup et al., 1995) and Ap66 (Perrier et al., 2003) microsatellites were analysed simultaneously. In multiplex *Apis*2, simultaneous analysis of A113, A7 (Estoup et al., 1995), Ap43 (Garnery et al., 1998b), Ap55 and Ap81 (Perrier et al., 2003) took place. PCR amplifications were carried out in 10 μL, with 1.5 units of AmpliTaq Gold®, 50 ng of DNA template, 0.3 mM of dNTP Mix of GeneAmp® and 10 μg of BSA. $MgCl_2$ concentration was adjusted to 1.5 mM for multiplex *Apis*1 and 1.2 mM for multiplex *Apis*2. The annealing temperature was 54 °C for both multiplex reactions. AmpliTaq Gold was activated for 7 minutes at 95 °C and one hour of elongation at 72 °C was added at the end of the PCR program. The primer concentrations used were 0.2 pmol for A28 and A24, 0.4 pmol for B124, A88, A113, A7, Ap43 and Ap55, 0.5 pmol for Ap81 and 0.6 pmol for Ap66. One each of the primers of A88, Ap66,

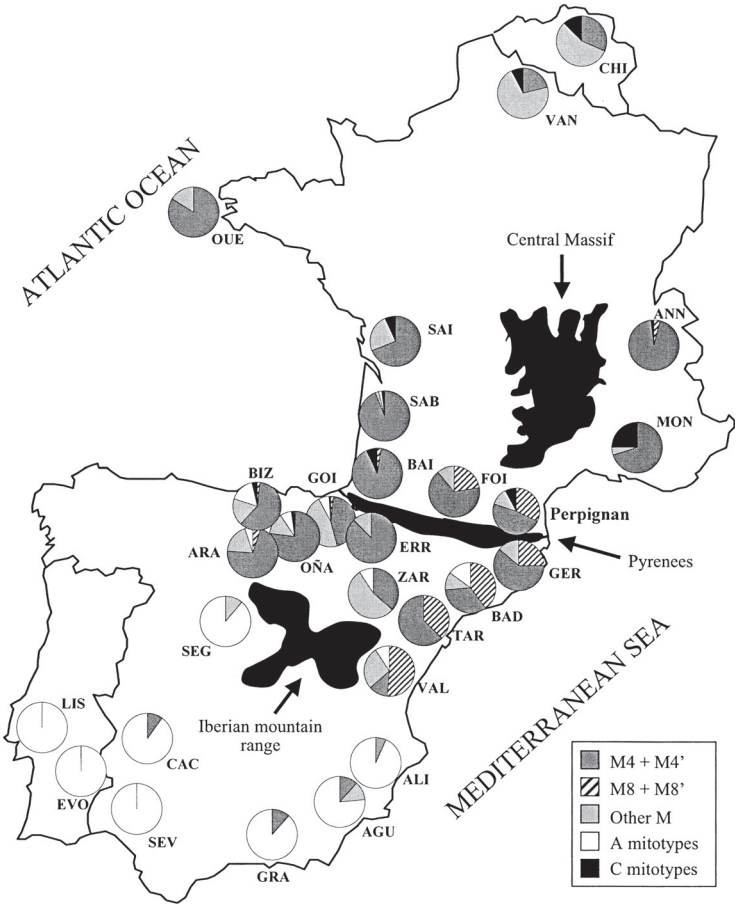

Figure 1. Geographical distribution of mtDNA haplotypes in the populations studied. Data from Belgium (CHI) and France (BAI, SAB, SAI, OUE, VAN, ANN, and MON) except FOI were obtained from Garnery et al. (1998). mtDNA data from Perpignan have been included, but this population has not been used in microsatelitte analysis because no data are available.

B124, A113, Ap55 and Ap81 was marked with 6-FAM (blue), A7 and A24 with HEX (green) and A28 and Ap43 with NED (yellow). DNA fragments were identified using an ABI PRISM 3100 Genetic Analyser and the GeneScan® Analysis 3.7.1 software and were labeled using the Genotyper® 3.7 software. Allele nomenclature was standardized by using reference samples from Estoup et al. (1995), Franck et al. (1998) and Garnery et al. (1998b).

2.4. Mitochondrial DNA

The COI-COII intergenic region was amplified through the use of PCR (Garnery et al., 1993). Fragment length was determined by electrophoresis in 1.4% agarose gel. The *Dra*I restriction pattern was identified by 10% PAGE stained with ethidium bromide. The sequences of new mtDNA haplotypes (mitotypes) were determined by sequencing (dRhodamine Terminator Sequencing Reaction®, Applied Biosystems) COI-COII amplification products previously purified with the purification kit QIAquick (QIAGEN) on an automatic sequencer (ABI PRISM 310 Genetic Analyser). The sequences were edited and aligned by BioEdit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Alignment was achieved using the algorithm ClustalW Multiple Alignment (Thompson et al., 1994).

2.5. Statistical analysis

Heterozygosity and standard error for haploid genomes was calculated according to Nei and Tajima (1981) and following Nei (1978) for diploid polymorphisms. D_A genetic distance matrix for microsatellites (Nei et al., 1983), neighbor-joining tree and bootstrap procedures (1000 iterations) were performed using Populations 1.2.24 software (Langella, 2002). Analyses of molecular variance among and within groups (AMOVA; Excoffier et al., 1992) were carried out using Arlequin ver. 1.1 software (Schneider et al., 1997). Using all the nodes of the neighbor-joining tree two groups of populations were established, each representing one side of the node. In addition, multiple pairs of new groups were established by interchanging populations surrounding these nodes between these groups. AMOVA analysis was performed repetitively for every pair of groups. The most remarkable results were represented using a Gabriel graph. Since AMOVA analysis may be influenced by geographic distance, first-order correlations were used to explore whether geographic and subspecies distance matrices were useful predictors of the genetic distance matrix. As geographic and subspecies distance matrices may also be correlated, partial correlations were calculated (Smouse et al., 1986). In these cases, SPSS 9.0 software was used to obtain residual matrices. Both full and partial matrix comparisons by Mantel's test were performed with the NTSYS program (Rohlf, 1988) and significance was obtained after 10 000 iterations (Smouse et al., 1986). Spatial autocorrelation analysis using Moran's I coefficient in 10 distance classes (Sokal and Oden, 1978) and correlogram construction were carried out with SAAP software (Wartenberg, 1989). The upper bound limits (in km) of the distance classes were 179, 306, 418, 508, 575, 672, 754, 911, 1068 and 1740. The COCOPAN method for subspecies (contiguity-constrained permutational ANOVA; Legendre et al. 1990) and Gabriel graphs were performed using R Package software (Legendre and Vaudor, 1991). When necessary, Bonferroni's correction was applied (Weir, 1996).

3. RESULTS

3.1. Mitochondrial DNA

Within the Iberian Peninsula, *Dra*I RFLP analysis of the COI-COII intergenic region revealed 34 different mtDNA haplotypes (mitotypes): 24 of lineage M, nine of lineage A and one of lineage C (Tab. I). Two mitotypes, M36 and M37, are reported here for the first time. Their frequencies varied between 1% and 2% in populations of the western Pyrenees, GOI and ONA. Sequence analysis showed a single nucleotide substitution for each new mitotype when compared to mitotype M7 sequence. That is, T→A transversion in nucleotide site 862 for M36; and A→T in nucleotide site 701 for M37. Nucleotide enumerations were based on the M7 sequence of the cloned fragment of the honeybee mtDNA of Cornuet et al. (1991). These changes create new digestion patterns compared to M7. Mitotype M36 contains a new *Dra*I digestion, which converts the 131 bp M7 fragment into two M36 fragments (94 and 37 bp). Mitotype M37 shows a single base displacement of the digestion site at one Q region, yielding two new M37 fragments of 96 and 64 bp instead of 95 and 65 bp M7 fragments (Fig. 2).

Based on the geographic distribution of the different mtDNA haplotypes identified on the Iberian Peninsula (Fig. 1), three main features can be highlighted. First, A mitotypes appear in every Iberian locality sampled except TOR, GER, and ERR. Second, the proportion of M mitotypes is less than 10% in populations south of the Iberian Mountain Range, but higher than 80% north of it. Third, in contrast to the high values of mtDNA introgression of the C branch found by Garnery et al. (1998a) in French and Belgian populations, Iberian Peninsula populations show null or very low values (Tab. I). C mitotypes were detected in only three populations close to France, BIZ (3.6%), GOI (1.3%) and ONA (2.2%).

Of the nine A mitotypes detected in the Iberian Peninsula, mitotype A2 predominated in populations south of the Iberian Mountain Range. Mitotype A2 is common on some Mediterranean islands (Garnery et al., 1993; Franck et al., 2000b; De la Rúa et al., 2001), while it is very rare in African populations (Garnery et al., 1998a). Similarly, mitotype A16 occurred frequently in southwest Iberia, yet has not been found in African populations. In addition, typically African mitotypes were mainly observed in the southern part of the

Table I. Number of specified mitotypes in each studied population of honey bees.

	Mitotypes M																										Mitotypes A									C
POP	1	3	4	4'	4''	5	6	7	7'	8	8'	9	10	11	11'	12	13	17	17'	19	20'	26	27	27'	29	32	36	37	1	2	3	4	15	8	9	13	16	C1
LIS																													3		4		2	1			19	
EVO																													9	12	3			5	11	2	3	
SEV																													9	5		1		5				
CAC			1																											1	2			4			2	
GRA			3																										3	11	2	7						
AGU			1	1				1										1												3		10						
ALI			3																										4	36	2							
SEG								1																						3				4			1	
VAL			4					9		17																				3								
TOR			13	9						11	2																											
BAD			16	2		2		2		19	2											2							1	7								
GER			20	2		3				8	1											1			1												
ZAR		1	4					2	3																					1								
ARA			67	16	4		1	11	1	5	1			1		3					1									2				3	1			
BIZ			21	12	1			2			1		1	2	1	1				1				1					1	5	2			1				2
ONA			114	25	1			14	1		1															3	6			3				5			5	4
GOI			99	7			5	38	7	2	2					6				47			2			3	3			7	8				1			3
ERR			29					2												2																		
BAI			27	10						1																												3
FOI			12							4												2																
SAB			39	8			1																												1			1
SAI			20				3						1	1		1	1																					2
MON			26	2			2																															10
ANN			44	3						2																												1
OUE			42				8																															
VAN	1		12	4	1		1					1				1		37	12																			6
CHI	1		12	3														25	1																			6

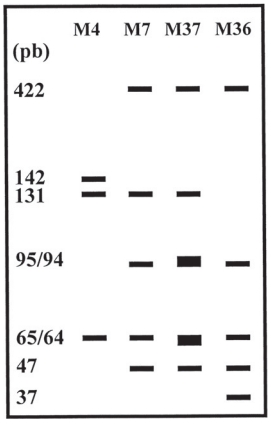

Figure 2. Banding pattern of the new mtDNA COI-COII haplotypes M36 and M37, compared with M7 and M4 mitotypes.

Iberian Peninsula, such as, A1, A4, A8 and A9 (Franck et al., 2001). Finally, mitotype A15, previously reported exclusively from the Canary Islands (De la Rúa et al., 1998), was also found in a Portuguese population (LIS, 7%).

M4 was the most common M mitotype in Spanish populations, as previously reported for France and Belgium. However, some remarkable features were found in the geographic distribution of the other 23 distinct M mitotypes detected in western Europe. First, 14 mitotypes detected in northern Iberia were absent in French and Belgian populations (M3, M5, M7, M7', M8', M11', M19, M20', M27, M27', M29, M32, M36, M37). However, only four mitotypes were exclusive to France and Belgium (M1, M9, M13 and M17'). Second, six common M mitotypes (M4", M10, M6, M11, M11' and M12) previously described by Garnery et al. (1998a) in French and Belgian populations were also present in populations of the Basque Country (western Pyrenees), but were not detected in the rest of the Iberian Peninsula. In the same direction, M7' was present only in the western and central part of the Pyrenees, where M7 occurred in high frequency. By contrast, the eastern side of the Pyrenees (Mediterranean) showed a very low incidence of mitotype M7, and mitotype M8 occurred at a high frequency, ranging from 20% to 52%, yet it was rare in the western Pyrenees.

Table II. Mitochondrial DNA haplotype diversity (H) and nuclear heterozygosity and standard deviations (SD).

	$H_{MITOCHONDRIAL}$			$H_{NUCLEAR}$		
POP	N	H	SD	2N	H_e	SD
LIS	29	0.586	0.099	60	0.513	0.218
EVO	45	0.825	0.026	90	0.475	0.275
SEV	20	0.705	0.061	42	0.508	0.295
CAC	10	0.822	0.097	34	0.491	0.283
GRA	26	0.745	0.057	50	0.442	0.279
AGU	17	0.647	0.119	34	0.476	0.306
ALI	45	0.354	0.087	92	0.516	0.267
SEG	9	0.751	0.112	16	0.466	0.285
VAL	33	0.657	0.061	68	0.480	0.245
TOR	35	0.715	0.032	74	0.458	0.266
BAD	53	0.770	0.036	98	0.443	0.249
GER	36	0.647	0.073	86	0.468	0.265
ZAR	11	0.818	0.082	62	0.460	0.288
ARA	117	0.650	0.002	142	0.453	0.273
BIZ	55	0.805	0.042	110	0.486	0.297
ONA	182	0.583	0.040	204	0.459	0.264
GOI	240	0.764	0.020	204	0.494	0.276
ERR	33	0.227	0.093	66	0.469	0.273
BAI	41	0.513	0.072	84	0.357	0.300
FOI	18	0.523	0.112	42	0.444	0.293
SAB	50	0.372	0.078	100	0.401	0.224
SAI	29	0.522	0.108	58	0.428	0.294
MON	40	0.523	0.072	60	0.454	0.263
ANN	50	0.224	0.077	100	0.443	0.310
OUE	50	0.274	0.071	100	0.279	0.275
VAN	76	0.712	0.044	58	0.394	0.293
CHI	48	0.659	0.054	60	0.38	0.344

The mtDNA diversity levels found in Iberian populations were greater than those found in France or Belgium (Tab. II). The only exception is ERR (H = 0.23), possibly due to the effect of a small population size (33 colonies).

3.2. Microsatellites

Nuclear heterozygosity values based on microsatellites data are shown in Table II. In

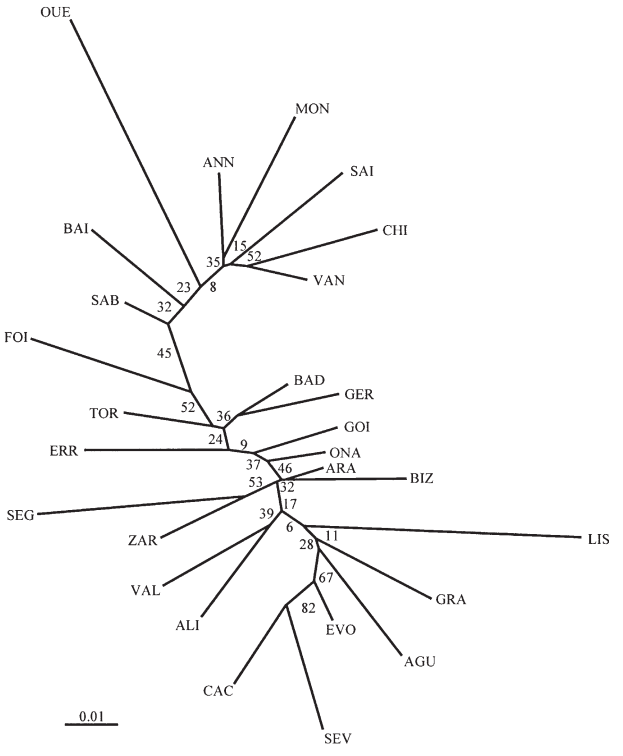

Figure 3. Neighbor-joining tree of the 27 Iberian, French and Belgian populations based on DA genetic distance matrix (10 microsatellites).

contrast to mitochondrial diversity data, the nuclear markers showed similar heterozygosity values in *A. m. mellifera* and *A. m. iberiensis* subspecies.

In an attempt to establish the geographic boundary between *A. m. mellifera* and *A. m. iberiensis* subspecies, several statistical tests were applied to assess the genetic relationships and hypothetical groupings within western European honeybee populations. A neighbor-joining tree was constructed based on D_A genetic distance matrix (Fig. 3). The tree does not show clear groups, but the locations of the populations in the tree mirror their north-south geographic distribution, suggesting that isolation by distance is an evolutionary mechanism of genetic differentiation among honeybee populations in western Europe. This analysis does not provide evidence of any significant barrier to gene flow between *A. m. mellifera* and *A. m. iberiensis* subspecies. Subsequently, we performed a hierarchical analysis of molecular variance (AMOVA) using each of the nodes from the neighbor-joining tree as a hypothetical line of separation between *A. m. mellifera* and *A. m. iberiensis*. For each node, variance between groups (F_{CT}) and within groups (F_{SC}) was estimated (Tab. III). The aim of this repetitive analysis was to establish two population groups, which would maximise the variance among groups (F_{CT}) and minimise within-group variance (F_{SC}). The results of this analysis are presented on a Gabriel graph (Fig. 4): each thick line on the graph represents the geographical line that separates the two groups of populations tested via AMOVA, and the figure shown alongside is the F_{CT}/F_{ST} ratio between those groups. Results for all the pairs of groups obtained using all the nodes of the neighbor-joining tree are depicted in Figure 4a. From a geographic point of view, separation lines mainly show east-west orientation and, subsequently, comparisons are made between groups of populations

Table III. Results of the hierarchical analyses of molecular variance (AMOVA) using each of the nodes from the neighbor-joining tree as a hypothetical line of separation between *A. m. mellifera* and *A. m. iberiensis*. Variance among groups (F_{CT}) and within groups (F_{SC}). The node that separates western Europe populations exactly at the Pyrenees showes the maximum variance between groups (F_{CT}) and minimum within-group variance (F_{SC}) and the optimal hierarchical classification. LIS (1), EVO (2), SEV (3), CAC (4), GRA (5), AGU (6), ALI (7), SEG (8), VAL (9), TOR (10), BAD (11), GER (12), ZAR (13), ARA (14), BIZ (15), ONA (16), GOI (17), ERR (18), BAI (19), FOI (20), SAB (21), SAI (22), MON (23), ANN (24), OUE (25), VAN (26), CHI (27).

FSC	FST	FCT	FSC(%)	FCT(%)	group1	group2
0.034	0.054	0.021	63.1%	38.2%	26–27	1–25
0.033	0.052	0.019	64.5%	36.7%	22, 26–27	1–21, 23–25
0.035	0.047	0.013	74.0%	27.0%	23–24	1–22, 25–27
0.031	0.051	0.021	60.1%	41.1%	22–24, 26–27	1–21, 25
0.027	0.054	0.028	50.2%	51.2%	22–27	1–21
0.025	0.054	0.030	46.5%	54.9%	19, 22–27	1–18, 20–21
0.024	0.053	0.030	45.0%	56.3%	19, 21–27	1–18, 20
0.023	0.053	0.030	43.7%	57.6%	19–27	1–18
0.024	0.050	0.026	48.6%	52.6%	10, 19–27	1–9, 11–18
0.036	0.036	0.000	101.2%	–1.3%	11–12	1–10, 13–27
0.027	0.045	0.019	58.5%	42.7%	10–12, 19–27	1–9, 13–18
0.027	0.044	0.018	61.4%	39.7%	10–12, 18–27	1–9, 13–17
0.027	0.045	0.019	59.0%	42.1%	10–12, 17–27	1–9, 13–16
0.027	0.045	0.019	60.2%	41.0%	10–12, 16–27	1–9, 13–15
0.036	0.035	–0.001	104.1%	–4.3%	8, 13	1–7, 9–12, 14–27
0.037	0.035	–0.001	103.4%	–3.6%	14–15	1–13, 14–27
0.029	0.048	0.020	60.8%	40.4%	8, 10–27	1–7, 9
0.035	0.044	0.010	78.9%	21.8%	7, 9	1–6, 8, 10–27
0.031	0.050	0.019	62.4%	38.8%	7–27	1–6
0.034	0.046	0.013	72.1%	28.8%	1, 7–27	2–6
0.034	0.048	0.014	71.0%	30.0%	1, 5, 7–27	2–4, 6
0.034	0.047	0.013	72.6%	28.4%	1, 5–27	2–4
0.035	0.047	0.012	75.9%	25.0%	1–2, 5–27	3–4

located north and south of the lines. The pair of groups at Figures 4b, c are based on the previous groups in Figure 4a, but modified by interchanging populations surrounding the nodes. Only the most representative results are presented. In Figure 4d, unlike the previous figures and for comparison purposes, populations are subdivided into two groups based on a north-south barrier. The low F_{CT}/F_{ST} values observed in this last figure (Fig. 4d) in comparison to the other three are not worth highlighting. Based on the neighbor-joining tree they have clearly east-west separation lines.

Of the more than one hundred groups tested, the geographical line that separated western European populations at the Pyrenees showed the maximum F_{CT} value with respect to F_{ST} (F_{ST} = 0.05253; F_{CT} = 0.03026, –57.6%–; F_{SC} = 0.02296, –43.7%–).

Since the AMOVA analysis may be influenced by geographic distance, so that the observed classification could be artificial and caused by the arbitrary creation of boundaries, we performed a correlation analysis and obtained statistical significance through Mantel's test. We compared the D_A genetic distance matrix (GEN) with the geographical distance matrix between pairs of populations measured as a straight-line distance between populations (GEO) and a third matrix of distances

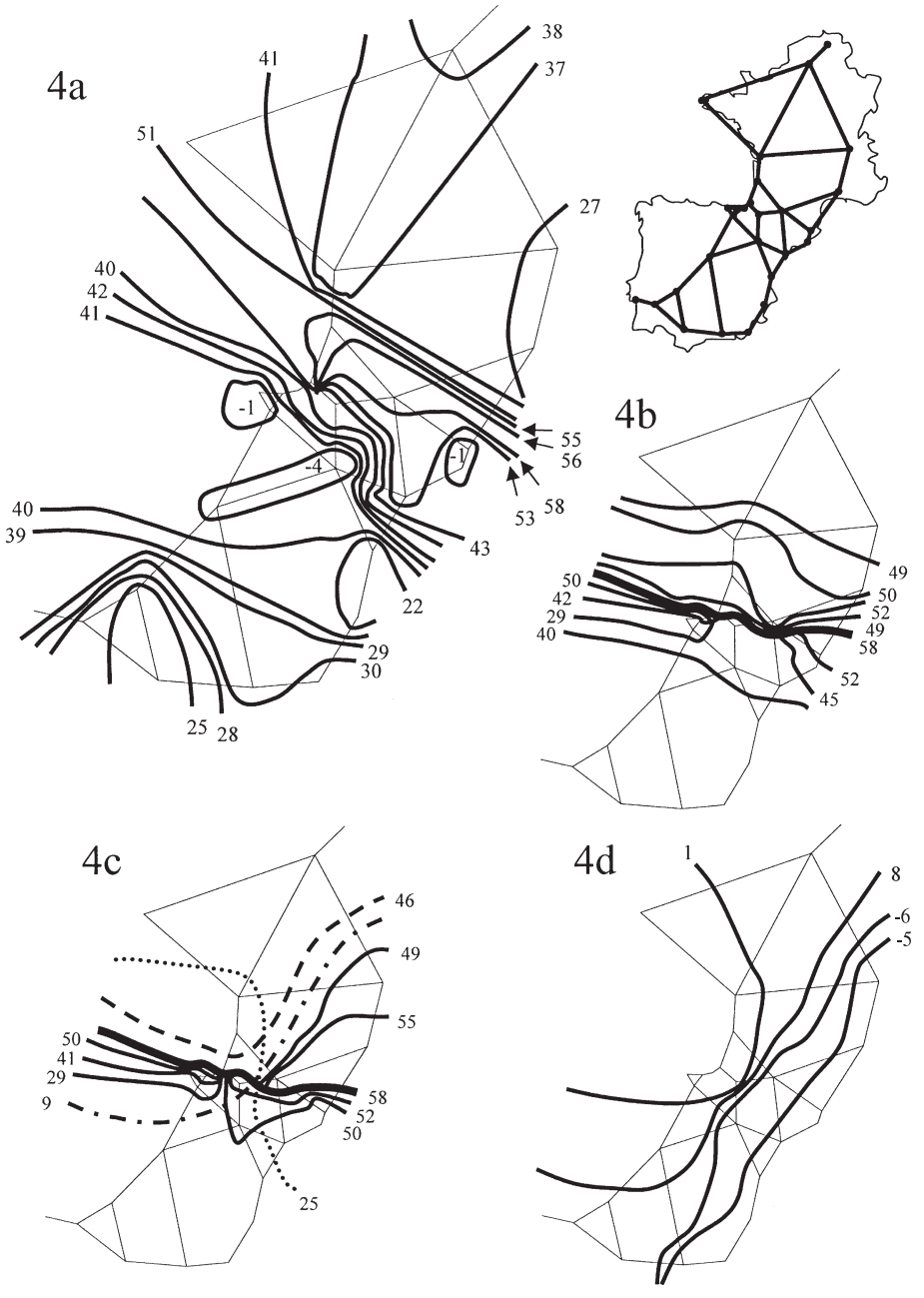

Figure 4. Representation of F_{CT}/F_{ST} values of the groups of populations based on the nodes of the neighbor-joining tree (a), and some of other groupings tested (b–d).

representing the grouping of western honeybee populations in two subspecies separated at the Pyrenees (SUBSP). To create this last matrix we allocated a distance of 0 to pairs of populations from the same subspecies and 1 to pairs from different subspecies. Simple matrix correlations proved highly significant for genetic and geographical distance matrices (GEN,GEO r = 0.678; $P < 0.0001$) and for genetic and subspecies matrices (GEN,SUBSP r = 0.634; $P < 0.0001$). Similarly, SUBSP and GEO distances were found to be closely correlated (r = 0.418; $P < 0.0001$), so we performed a partial correlation analysis to determine the level of correlation of SUBSP or GEO matrices with the genetic distance matrix when one of them remains constant. The geographic partial correlation value was $r_{GEN,(GEO.SUBSP)} = 0.588$; $P < 0.0001$, and subspecies partial correlation $r_{GEN,(SUBSP.GEO)} = 0.525$; $P < 0.0001$. Oden and Sokal (1992) recommend the use of a limit value of $P = 0.001$ in tests of this type. Thus, our results are statistically significant for the correlation between residuals in the two cases. So both geographical and subspecies distances can be considered to be useful predictors of the genetic distance matrix.

To test for the possibility that spatial autocorrelation might influence this analysis, we established correlograms for the 44 alleles with allele frequencies greater than 2% by dividing the geographic area under study into 10 distance classes. When Bonferroni's correction is applied 14 out of 44 correlograms show a significant probability ($P < 0.0012$). 4 of the 14 alleles with significant correlograms show patterns like those described by Sokal et al. (1989) as *clinal* correlograms, with Moran's I values monotonically decreasing from positive to negative autocorrelation (Fig. 5a). The same clinal structure is observed in mtDNA, since two of the four mitotypes with significant correlograms show a similar pattern. These would all seem to form part of the correlation observed previously between genetics and geography. Figure 5b includes the correlograms for a further seven alleles which have patterns with Moran's I values which are positive and significant in distance classes 2−5, and highly negative at long distances. This result is consistent with the existence of two areas with similar allele frequencies which are distinct from each other. Figure 5c shows a type of correlogram that is different from those mentioned above, and includes the remaining nuclear alleles plus the other two mitotypes. In most of them there are also negative values of Moran's I at long distances, though a few, such as mitotype M7, display a structure similar to the *regional patches* described by Sokal et al. (1989). The spatial autocorrelation observed may bias the earlier AMOVA analysis, so it becomes necessary to use a method that can test genetic structuring of honeybee populations in western Europe, eliminating possible autocorrelation effects. We therefore applied the COCOPAN method (contiguity-constrained permutational ANOVA). The individual probabilities of the 44 alleles studied were corrected using Bonferroni's procedure. Four alleles (Ap43*135, Ap43*137, Ap81*128 and B124*222) yielded significant probability levels when tested for the existence of a boundary at the Pyrenees between the two subspecies. In earlier spatial autocorrelation analyses, three of them showed significant correlograms. This suggests the existence of two areas with similar allele frequencies which are different from each other (Fig. 5b). The detailed analysis of their frequency distribution marks a clear step at the Pyrenees.

4. DISCUSSION

In this study, the analysis of 10 microsatellite markers in 27 local honey bee populations in the Iberian Peninsula, France and Belgium highlights two factors of genetic isolation in western Europe: geographic distance and the geographical barrier of the Pyrenees. Analysis of simple and partial correlations indicates that geographic distance is an isolation factor that explains a large amount of the nuclear variation of honey bee populations from western Europe. In fact, the shorter the geographic distance between populations, the greater their genetic similarity is. In this sense, the location of populations on the neighbor-joining tree coincides largely with their geographic distribution as a north-south continuum, with no clear

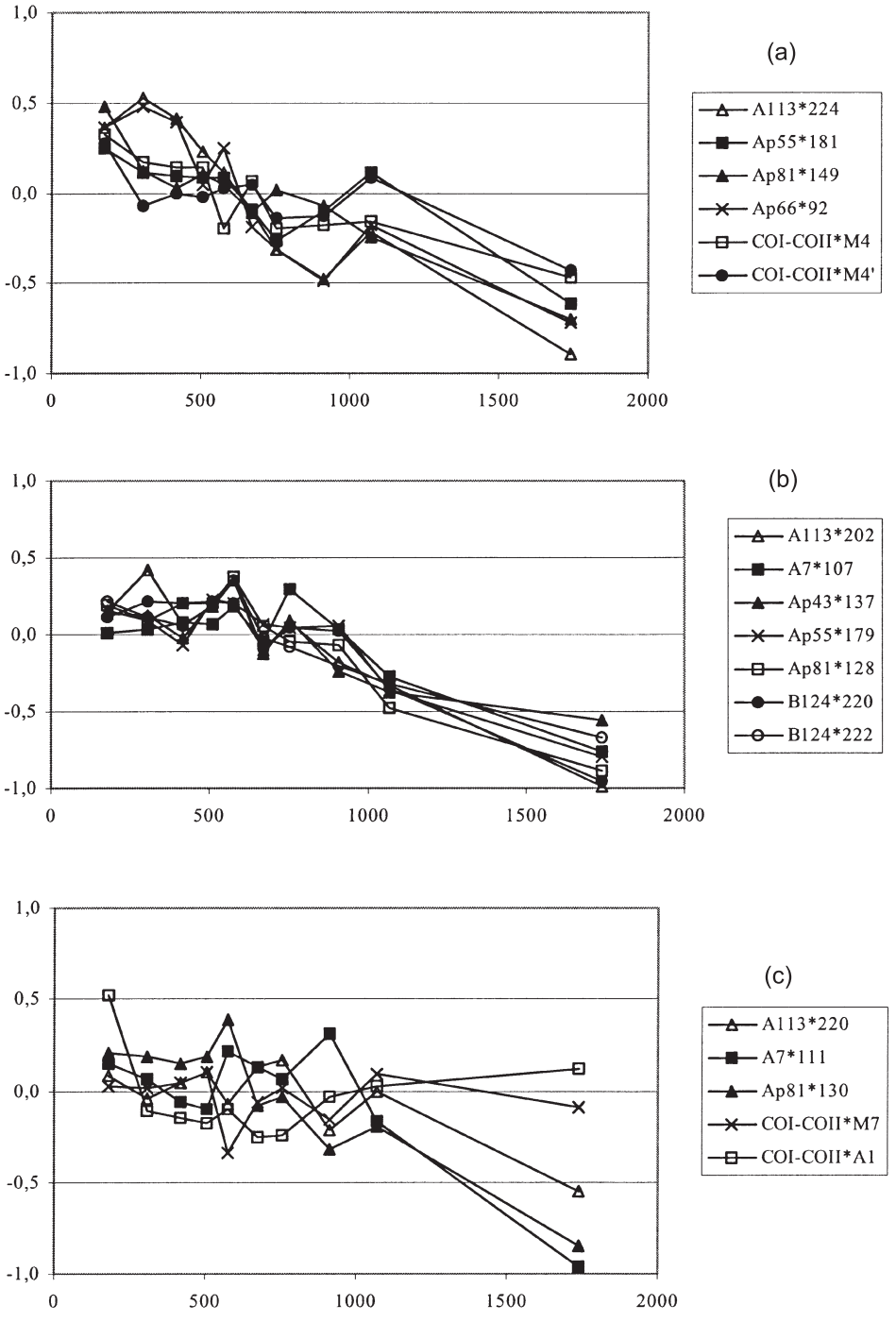

Figure 5. Spatial correlograms for the 14 microsatellite alleles and 4 mitotypes which yield significant patterns.

break interrupting this tendency (Fig. 3). Nevertheless, AMOVA and COCOPAN methods applied to microsatellite data, provide compelling evidence of a barrier to neutral gene flow at the Pyrenees. This confirms that the Pyrenees delimit the geographic boundary between *A. m. mellifera* and *A. m. iberiensis* subspecies, as suggested by Ruttner (1988).

Likewise, mtDNA variation shows that A mitotypes appear in all the Iberian localities sampled except for TOR, GER, and ERR, while north of the Pyrenees A mitotypes are very rare. Moreover, with regard to mtDNA haplotype distribution south of the Pyrenees, several authors (Garnery et al., 1995, 1998a; Franck et al., 1998) describe a clinal transition from mitotype A to M from the south-western to north-eastern regions of the Iberian Peninsula. Our study analyses 18 Iberian populations rather than the 6 of previous studies, and our results indicate that despite the fact that mitotypes show a clinal tendency, a marked step appears near the Iberian Mountain Range, where A mitotype frequencies vary from less than 10% north to more than 80% south. This observation shows at least two classes within the *A. m. iberiensis* subspecies: northeastern and southwestern groups. A step-like demarcation consistent with our findings was also reported from a fine-scale evaluation of morphological, mtDNA and allozyme variation in honey bee colonies from southern France, Spain, Portugal and northern Morocco (Arias et al., 2006).

On the other hand, the distribution of M lineage mtDNA haplotypes in western Europe is consistent with the hypothesis that the Iberian Peninsula served as a refugia for the western European bee during glacial periods and that consecutive founder effects occurred during the process of re-colonization of northern Europe. Thus, M lineage mitotypes are clearly predominant around the Iberian Pyrenees, France and Belgium (Tab. I). In all, M4 is the most common M mitotype. However, 14 out of the 24 different mitotypes M were detected exclusively in the Iberian Peninsula. Moreover, overall M lineage mitochondrial diversity is significantly higher in Iberian than French or Belgian populations. The reduction of variability due to genetic drift is a characteristic of rapid re-colonization processes (Hewitt, 1996). However, this reduction of variability is not found in microsatellite data; nuclear genetic diversity is similar between Iberian, French and Belgian populations. This may be explained by higher mutation rates of the microsatellites, which would increase the initial diversity and would mask the effects of the re-colonization process in a relatively short time. So the Iberian Peninsula, especially the area between the Iberian Mountain Range and the Pyrenees, could be considered a genetic reservoir for M maternal lineages in western Europe and a valuable resource for conservation purposes.

The role of the Iberian Peninsula as an ice-age refugia has been described for multiple taxa, including mammals, amphibians, arthropods and plants (Hewitt, 1996, 1999, 2000; Lunt et al., 1998, Taberlet et al., 1998). Moreover, the ability of different animal species to expand from an Iberian refuge across the Pyrenees is quite a common feature (Lunt et al., 1998; Taberlet et al., 1998) and the most probable origins of re-colonization of northern Europe are the western and eastern ends of these mountains. In that sense, our results show that south and north of the two ends of the Pyrenees, bee populations share M mitotypes. This reflects gene flow at both ends of the Pyrenees and confirms these areas as paths for bee populations. However, the two ends of the Pyrenees exhibit clear differences from each other based on mitotype distribution. This enables us to discern the influence of each Pyrenean path on the post-glacial re-colonization of northern Europe. That is, in the north- and south-eastern Pyrenees, the Mediterranean area, M8 is found in high frequencies, but it is rare at the western end and in the rest of Europe (Fig. 1). Conversely, the western extreme of the Pyrenees shows not only the typical French M4 mitotype, but multiple mitotypes also detected among French and Belgian populations, but absent from Mediterranean populations. Overall, the mitotype distribution of the French populations is predominantly a subset of that found in the western Pyrenees. Interestingly, Atlantic vegetation and climate are found from the Iberian western Pyrenees to northern Europe, whereas the eastern Pyrenees

and southeastern France have Mediterranean vegetation and climate (Fig. 1). It is likely that these climatological differences could explain, at least in part, the M mitotype variation pattern in western Europe, as bee populations in the western Pyrenees may have been more competitive than their eastern Pyrenees counterparts during the re-colonization process. In addition, the French Central Massif, located north of the Pyrenees (Fig. 1) could act as a barrier to the eastern migratory path.

In addition to previous results, this study adds valuable information about the pattern of expansion of honey bees in Europe during the warm post-glacial period. Altogether, the data support a unique place of origin of northern European bees: the Iberian Peninsula. Based on morphological evidence, Ruttner (1988) included the north European bee, *A. m. mellifera*, in the M evolutionary lineage together with *A. m. iberiensis,* while the subspecies of Italy (*A. m. ligustica*) and the Balkans (*A. m. carnica*) were classified in the C evolutionary lineage. Genetic data support this classification. Even though a recent study detected some C lineage mitotypes in northern honeybee populations (Norway, Sweden, Denmark, England, Scotland and Ireland), their presence is consistent with commercial importation by humans (Jensen et al., 2005). The absence of Balkan subspecies from northern Europe could be explained by an earlier invasion by Iberian honeybees into northern Europe that blocked the expansion of the Balkan honeybee. The same reason could explain the lack of expansion of the Italian bee, but in this case the barrier effect of the Alps (Hewitt, 1999) is a more likely reason. So all northern European bees included in the M evolutionary lineage seem to have originated from the Iberian refugia. Thus the "honeybee" re-colonization pattern does not seem to fit any of the three main post-glacial expansion patterns described by Hewitt (1999, 2000) for other taxa. Moreover, we cannot exclude the input of bees from other cryptic refugia elsewhere in Europe as suggested by Stewart and Lister (2001) for several species. The presence of mitotypes M1, M9, M13 and M17' in France and Belgium, but not in Iberia, would be consistent with alternative small refugia.

In summary, our data suggest that the re-colonization of northern Europe was carried out by bees that left the Iberian Peninsula from both the western and eastern ends of the Pyrenees, although the western Pyrenean populations were more influential. The post-glacial re-colonization process, isolation by distance between populations and the barrier constituted by the Pyrenees are three factors that were involved in subsequent differentiation of the subspecies *A.m. mellifera* and *A. m. iberiensis*.

ACKNOWLEDGEMENTS

This study was funded by the Basque Government Departments of Agriculture and Fisheries, and Education, Universities and Research and by the INIA, in the "Conservation of the Genetic Resources of Food-processing Interest" framework of the Spanish national R&D&i plan. (Project No. RZ02-009-C2-1). We also thank the numerous beekeepers and colleagues that provided invaluable assistance with the collection of honey bees for this research. We are grateful to Gipuzkoako Erlezainen Elkartea and to Gipuzkoako Foru Aldundia.

Flux génique au sein de la lignée évolutive M d'*Apis mellifera* : rôle des Pyrénées, isolation par la distance et routes de recolonisation post-glaciaires en Europe occidentale.

***Apis mellifera* / génétique des populations / microsatellite / sous-espèce / différenciation / AD-Nmt / Europe**

Zusammenfassung – Genfluss in der evolutionären Linie M von *Apis mellifera*: Rolle der Pyrenäen, Isolation durch Entfernung und Wege nacheiszeitlicher Wiederbesiedlung im westlichen Europa. Die Unterarten der Honigbiene (*Apis mellifera)* werden in die 5 evolutionäre Linien A (African), C (northern Mediterranean), M (western Europe), O (Oriental) and Y (Yemenitica) gruppiert. Die in dieser Studie untersuchte evolutionäre Linie M enthält zwei Unterarten: *A. m. mellifera* wird von Frankreich bis zu den Bergen des Ural gefunden und *A. m. iberiensis* ist auf der iberischen Halbinsel verbreitet. Obwohl allgemein angenommen wird, dass die Pyrenäen ein bedeutendes Hindernis für den Genfluss zwischen *A. m. mellifera* und *A. m. iberiensis* darstellt, wurde ein solcher Barriereeffekt bislang nicht nachgewiesen. Die Existenz genetischer Gefälle vom Süden der Iberischen Halbinsel bis zum nördlichen Europa und

eine ungewisse taxonomische Zuordnung einiger Populationen in den Pyrenäen tragen zu der Unsicherheit bezüglich der Rolle der Pyrenäen als genetische Barriere bei. Andererseits wird seit längerem angenommen, dass die Iberische Halbinsel während der Eiszeit als Refugium für die westliche Honigbiene diente (Ruttner, 1952, 1988) und eine nacheiszeitliche Wiederbesiedlung von Nordeuropa wird von mehreren Autoren unterstützt (Garnery et al., 1998a,b; Franck et al., 1998, 2000b). Gegenstand dieser Untersuchung war es, einen potentiell isolierenden Effekt der Pyrenäen nachzuweisen, neue Daten zu dem Differenzierungsprozess der zwei Unterarten beizutragen und den Ablauf der Wiederbesiedlung zu untersuchen. Wir untersuchten 1398 Völker aus 27 Populationen der Iberischen Halbinsel sowie aus Frankreich und Belgien auf Variation an 10 Mikrosatellitenloci und der COI-COII intergenischen Region der mtDNA. Wir verwendeten verschiedene Arten statistischer Analysen wie die D_A genetische Distanzmatrix, neighbor-joining trees, Korrelationen erster Ordnung und partielle Korrelationen, AMOVA, Analyse räumlicher Autokorrelationen und COCOPAN für Mikrosatellitendaten. Die Ergebnisse zeigten eine Isolation zwischen den verschiedenen Populationen der M Linie durch die Entfernung auf und lieferten sehr deutliche Hinweise auf eine Barriere für den neutralen Genfluß bei den Pyrenäen. Die Verteilung der mtDNA Haplotypen bestätigte das Iberische Refugium der Westeuropäischen Honigbiene in der Eiszeit. Wir konnten zwei verschiedene Wege der nacheiszeitlichen Wiederbesiedlung von der Iberischen Halbinsel aus an den beiden Enden der Pyrenäen ableiten. Es gab deutliche Unterschiede in der Verteilung der Mitotypen zwischen den westlichen und östlichen Enden der Pyrenäen, diese legten nahe, dass der westliche Weg für den nacheiszeitlichen Widerbesiedlungsprozess wichtiger war. Nach der in dieser Untersuchung beobachteten hohen Variabilität der M Mitotypen südlich der Pyrenäen könnten diese eine nützliche genetische Ressource für die Konservation der Westeuropäischen Honigbienen darstellen. Der nacheiszeitliche Wiederbesiedlungsverlauf, die Isolation durch die Entfernung und die von den Pyrenäen gebildete Verbreitungsbarriere sind Einflüsse, die zu der Ausbildung der Unterarten *A. m. mellifera* und *A. m. iberiensis* beigetragen haben.

Honigbienen / Unterarten / mtDNA / Mikrosatelliten / Populationsgenetik

REFERENCES

Arias M.C., Sheppard W.S. (1996) Molecular phylogenetics of honey bee subespecies (*Apis mellifera*), Mol. Phylogenet. Evol. 5, 557–566.

Arias M.C., Rinderer T.E., Sheppard W.S. (2006) Further characterization of honey bees from the Iberian Peninsula by allozyme, morphometric and mtDNA haplotype analyses, J. Apic. Res. 45, 188–196.

Cornuet J.M. (1982) The MDH polymorphism in some west-mediterranean honeybee populations. International Union for the Study of Social Insects Congress, Boulder.

Cornuet J.M., Garnery L., Solignac M. (1991) Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. Mitocondrial DNA, Genetics 1128, 393–403.

De la Rúa, P., Serrano J., Galian J. (1998) Mitochondrial DNA variability in the Canary Islands honeybees (*Apis mellifera* L.), Mol. Ecol. 7, 1543–1547.

De la Rúa P., Galián J., Serrano J., Moritz R.F.A. (2001) Molecular characterization and population structure of the honeybees from the Balearic Islands, Apidologie 32, 417–427.

Engel M.S. (1999) The taxonomy of recent and fosil honey bees (Hymenoptera: Apidae; Apis), J. Hymenoptera Res. 8, 165–196.

Estoup A., Garnery L., Solignac M., Cornuet J.M. (1995) Microsatellite variation in hone bee (*Apis mellifera*) populations:hierarchical genetic structure and test of infinite models, Genetics 140, 679–695.

Excoffier L., Smouse P.E., Quattro J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data, Genetics 131, 479–491.

Franck P., Garnery L., Solignac M., Cornuet J.M. (1998) The origin of west European subspecies of honeybees (*Apis mellifera*): new insights from microsatellite and mitochondrial data, Evolution 52, 1119–1134.

Franck P., Garnery L., Solignac M., Cornuet J.M. (2000a) Molecular confirmation of a fourth lineage in honeybees from Near East, Apidologie 31, 167–180.

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

Franck P., Garnery L., Loiseau A., Oldfroyd B.P., Hepburn H.R., Solignac M., Cornuet J.M. (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data, Heredity 86, 420–430.

Garnery L., Cornuet J.M., Solignac M. (1992) Evolutionary history of the honey bee (*Apis mellifera* L.) inferred from mitochondrial DNA analysis, Mol. Ecol. 1, 145–154.

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, 1016–1021.

Garnery L., Mosshine E.H., Oldroyd B.P., Cornuet J.M. (1995) Mitochondrial DNA variation in Moroccan and Spanish honey bee populations, Mol. Ecol. 4, 465–471.

Garnery L., Franck P., Baudry E., Vautrin D., Cornuet J.M., Solignac M. (1998a) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA, Genet. Sel. Evol. 30 (Suppl. 1), S31–S47.

Garnery L., Franck P., Baudry E., Vautrin D., Cornuet J.M., Solignac M. (1998b) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). II. Microsatellite loci, Genet. Sel. Evol. 30 (Suppl. 1), S49–S74.

Hewitt G. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation, Biol. J. Linnean Soc. 58, 247–276.

Hewitt G. (1999) Post-glacial re-colonization of european biota, Biol. J. Linnean Soc. 68, 87–112.

Hewitt G. (2000) The genetic legacy of the Quaternary ice ages, Nature 405, 907–913.

Jensen A.B., Palmer K.A., Boomsma J.J., Pedersen B.V. (2005) Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe, Mol. Ecol. 14, 93–106.

Langella O. (2002) Populations, 1.2.28, Copyright (C) 1999, CNRS UPR9034.

Legendre P., Oden N.L., Sokal R.R., Vaudor A., Kim J. (1990) Approximate analysis of variance of spatially autocorrelated regional data, J. Classif. 7, 53–75.

Legendre P., Vaudor A. (1991) The R package: multidimensional analysis, spatial analysis. Département de Sciences Biologiques, Université de Montréal, Canada.

Lunt D.H., Ibrahim K.M., Hewitt G.M. (1998) mtDNA phylogeography and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*, Heredity 80, 633–641.

Nei M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals, Genetics 89, 583–590.

Nei M., Tajima F. (1981) DNA polymorphism detectable by restriction endonucleases, Genetics 97, 583–590.

Nei M., Tajima F., Tateno Y. (1983) Accuracy of estimated phylogenetic trees from molecular data, J. Mol. Evol. 19, 153–170.

Oden N.L., Sokal R.R. (1992) An investigation of 3-matrix permutation tests, J. Classif. 9, 275–290.

Perrier C., Strange J., Langella O., Sheppard S.W., Garnery L. (2003) Diversité génétique, introgressions mitocondriales et nucléaires dans une population d'abeilles des Landes de Gascogne, Actes Bureau Ressources Génétiques 4, 79–100.

Rohlf F.J. (1988) NTSYS-pc: numerical taxonomy and multivariate analysis system for the IBM PC microcomputer (and compatible), version 1.40, Oxford University Press, New York.

Ruttner F. (1952) Alter und Herkunft der Bienenrassen Europas, Österr. Imker (2), 8–10.

Ruttner F. (1988) Biogeography and taxonomy of honeybees, Springer Verlag, Berlin.

Ruttner F., Tassencourt L., Louveaux J. (1978) Biometrical-statistical analysis of the geographic variability of *Apis mellifera* L., Apidologie 9, 363–381.

Schneider S., Kueffer J.M., Roessli D., Excoffier L. (1997) Arlequin ver. 1.1: A software for population genetic data analysis, Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sheppard W.S., Meixner M.D. (2003) *Apis mellifera pomonella*, a new honey bee subspecies from Central Asia, Apidologie 34, 367–375.

Sheppard W.S., Rinderer T.E., Garnery L., Shimanuki H. (1999) Analysis of Africanized honey bee mitochondrial DNA reveals further diversity of origin, Genet. Mol. Biol. 22, 73–75.

Smith D.R., Glenn T.C. (1995) Allozyme polymorphisms in Spanish honeybees (*Apis mellifera iberica*), J. Hered. 86, 12–16.

Smouse P.E., Long J.C., Sokal R.R. (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence, Syst. Zool. 35, 627–632.

Sokal R.R., Oden N.L. (1978) Spatial autocorrelation in biology 1. Methodology, Biol. J. Linnean Soc. 10, 199–228.

Sokal R.R., Harding R.M., Oden N.L. (1989) Spatial patterns of human gene frequencies in Europe, Am. J. Phys. Anthropol. 80, 267–294.

Stewart J.R., Lister A.M. (2001) Cryptic northern refugia and the origins of the modern biota, Trends Ecol. Evol. 16, 608–613.

Taberlet P., Fumagalli L., Wust-Saucy A.G., Cosson J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe, Mol. Ecol. 7, 453–464.

Thompson J.D., Higgins D.G., Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22, 4673–80.

Wartenberg D. (1989) SAAP. A spatial autocorrelation analysis program, Setauket, New York.

Weir B.S. (1996) Genetic data analysis II. Sinauer Associates Inc., Sunderland, MA.