Genetic diversity of the west European honey bee (Apis mellifera mellifera and A. m. iberia). II.Microsatellite loci

L. Garnery, Pierre Franck, E. Baudry, D. Vautrin, Jean-Marie Cornuet, M. Solignac

Original article

Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). II. Microsatellite loci

Lionel Garnery[a], Pierre Franck[b], Emmanuelle Baudry[a],

Dominique Vautrin[a], Jean-Marie Cornuet[b], Michel Solignac[a*]

[a] Laboratoire populations, génétique et évolution, Centre national
de la recherche scientifique, 91198 Gif-sur-Yvette cedex, France
[b] Laboratoire de modélisation et de biologie évolutive, Institut national
de la recherche agronomique, 488, rue de la Croix Lavit,
34090 Montpellier cedex, France

Abstract – The genetic variability and differentiation of west European honey bee populations (*Apis mellifera mellifera* and *A. m. iberica*) have been investigated using 11 microsatellite loci. These two subspecies are characterised by a lower genetic variability than most other studied subspecies and several tests are indicative of a recent increase of the population size. Moreover, the genetic profiles are rather homogeneous from southern Spain to Scandinavia. French populations are more or less introgressed (a few percent up to 57 %) by genes from the north Mediterranean lineage which provides most of the imported queens. The inferred percentage of introgressed nuclear genes is generally well correlated with the proportion of alien mitochondrial deoxyribonucleic acid (mtDNA) haplotypes detected in the same populations. The level of introgression is the main source of genetic distances among populations. When introgressed genes are disregarded, however, populations cluster in two groups which correspond to both subspecies (*iberica* and *mellifera*), giving full support to the taxonomy of this lineage. © Inra/Elsevier, Paris

honey bee / microsatellites / population genetics / introgression / conservation

Résumé – Diversité génétique de l'abeille ouest européenne (*Apis mellifera mellifera* et *A. m. iberica*). II. Locus microsatellites. La variabilité génétique et la différenciation entre populations a été étudiée pour 11 locus microsatellites dans 15 populations de l'abeille ouest européenne. Les deux sous-espèces qui constituent ce rameau (*Apis mellifera mellifera* et *A. m. iberica*) ont une variabilité génétique

---

* Correspondence and reprints

plus réduite que la plupart des autres populations étudiées avec les mêmes outils moléculaires et plusieurs tests statistiques sont révélateurs d'une réduction récente de leur effectif. De plus, les profils génétiques des populations sont très comparables, du sud de l'Espagne à la Scandinavie, Les populations françaises sont plus ou moins introgressées (de quelques pourcentages jusqu'à 57 %) par des gènes de la lignée nord-méditerranéenne, les races *ligustica* et *carnica* étant les principales sources de reines importées. Le pourcentage de gènes nucléaires provenant de ces races est généralement équivalent à celui des haplotypes détectés dans les mêmes populations. Les degrés différents d'introgression sont la principale source de la distance génétique observée entre les populations ouest européennes ; cependant, lorsque ces gènes sont retirés, les populations se regroupent en deux ensembles qui correspondent aux deux sous-espèces (*iberica* et *mellifera*). Ce résultat apporte donc un soutien génétique à la taxinomie de cette lignée. © Inra/Elsevier, Paris

abeille / microsatellites / génétique des populations / introgression / conservation

1. INTRODUCTION

The use of microsatellites for population genetics studies is expanding exponentially. While these markers are very useful for the study of polymorphism in a variety of species, there are some organisms, namely Hymenoptera and particularly their social taxa, for which they do not mark simply a renewal but rather the emergence of a formal and population genetics. Indeed, when enzyme polymorphism emerged 30 years ago and was applied to innumerable species, the genetic variability detected in Hymenoptera remained desperately low (Sylvester, 1976; Cornuet, 1979; Nunamaker, 1980; Badino et al., 1983; Sheppard and Berlocher, 1984). Significant advances were only achieved with the introduction of DNA technologies, mitochondrial deoxyribonucleic acid (mtDNA) first and then nuclear markers, namely random amplified polymorphic DNA (RAPD), anonymous sequences and microsatellites.

Recently, RAPD markers have been used to map the honey bee genome (Hunt and Page, 1995) but, in spite of their variability and a relatively favourable situation in the species (where haploid drones can be used to baffle the drawbacks of dominant transmission at these loci), they remain of little interest for population genetics analysis where diploid workers are the main source of DNA. Similarly, a hypervariable probe has been developed for population analysis, but a single locus may not be sufficient to characterise genetic diversity of populations (Hall, 1990; McMichael and Hall, 1996). Until now, microsatellites have been mainly used in *Apis mellifera* in the fields of molecular evolution (Estoup et al., 1993, 1995b), theoretical models of mutations (Estoup et al., 1995a; Cornuet and Luikart, 1997) and reproductive behaviour and sociobiology (Estoup et al., 1994).

With the noticeable exception of mitochondrial DNA (see Part I, accompanying article), the variability of the honey bee has been developed for a long time at the phenetic level, mainly through morphometry (Ruttner et al., 1978; Ruttner, 1988). These data are of excellent quality and have been of great value in guiding subsequent molecular analyses. Morphometry has revealed that the 24 recognised subspecies in the Old World (the original geographic area of *Apis mellifera* before its dispersion around the world by humans) can be grouped in three evolutionary lineages: M for the west European honey bees, A for the

African continent and C for the north Mediterranean (see Part I, accompanying article for more details). Roughly, the same three lineages have been found with mitochondrial DNA (Smith, 1991; Garnery et al., 1992; Arias and Sheppard, 1996) and a study with seven microsatellite loci confirmed their profound genetic differentiation for nuclear genes (Estoup et al., 1995a). Recently, the study has been extended to some Iberian populations to investigate the evolutionary origin of the M lineage (Franck et al., 1998). The present work is a contribution to a more comprehensive knowledge of the genetic variability analysed in 15 populations of the two west European subspecies *Apis mellifera mellifera* and *A. m. iberica* using 11 microsatellite loci.

These populations are characterised by a low genetic variability, a low level of genetic differentiation and a variable level of introgression, mainly by alleles from the C lineage. Three populations from localities where conservatories of the local 'black' honey bees (*A. m. mellifera*) are established or to be established are included in this study.

2. MATERIALS AND METHODS

2.1. Sampling

A total of 571 honey bee workers from 17 populations were investigated (see Part I, *figure 1* of the accompanying article). Most of them have been sampled in the geographic area of the subspecies *Apis mellifera mellifera* (eight from France: Sabres, Landes; Saintes, Charentes; Angers, Maine-et-Loire; Ouessant, Finistère; Avignon, Vaucluse; Annecy, Savoie; Fleckenstein, Bas-Rhin; Valenciennes, Nord; one from Belgium: Chimay; one from Sweden: Umeå), and of the subspecies *A. m. iberica* (two from Spain: Toledo/Segovia, Castilla; Sevilla, Andalucia; one from Portugal: Porto); two Pyrenean populations from the putative contact zone (Bayonne, Atlantic Pyrenees and San Sebastian, Basque Country) are unassigned to one or the other subspecies. The populations of Al Hoceima (Morocco, *A. m. major*) and Chalkidiki (Greece, *A. m. macedonica*), belonging to the A and C lineages respectively, have been used as references. All these populations have been previously characterised for mtDNA (see Part I, accompanying article). For each population, a single worker per colony was sampled for a total of 21 to 50 individuals per population in a radius of 10 km.

2.2. Molecular analyses

DNA extraction was performed according to Kocher et al. (1989), with slight modification as described by Garnery et al. (1993) (see Part I, accompanying article for a more complete description). Radioactive polymerase chain reaction (PCR) amplifications of microsatellite loci were carried out in 10 μL as previously described in Estoup et al. (1995a) except that 0.15 μCi of α^{33}P-dATP was used as radioactive source. Eleven pairs of primers were used to amplify the loci. Seven of them are those already described by Estoup et al. (1995a): *A43*, *B124*, *A88*, *A113*, *A28*, *A24* and *A7* and another one by Franck et al. (unpublished report): *A8*. The sequences of the new ones (with indication of the optimal annealing temperature and concentration in $MgCl_2$) are given in *table I*.

Table I. Polymerase chain reaction primers and conditions for loci *Ap33*, *Ap36* and *Ap43*.

Locus	Primers	Annealing temperature (°C)	MgCl$_2$ (mM)
Ap33	5'-TTTCTTTTTGTGGACAGCG-3'		
	5'-AAATATGGCGAAACGTGTG-3'	54	1.2
Ap36	5'-CTACGCGCTTACAGGGCA-3'		
	5'-GCCGAAATTCAACGCTCA-3'	56	1.2
Ap43	5'-GGCGTGCACAGCTTATTCC-3'		
	5'-CGAAGGTGGTTTCAGGCC-3'	58	1.2

2.3. Statistical analyses

Unbiased estimates of gene diversity (heterozygosity) of microsatellite loci (*Hd*) were calculated according to Nei (1978). The number of alleles or allelic diversity of a sample depends on the sample size, which varies among populations and loci. In order to make valid comparisons, allelic diversity has to be adjusted to a common sample size according to the following formula (El Mousadik and Petit, 1996):

$$Da(m) = \sum_{i=1}^{k}[1 - C_{n-n_i}^{m}/C_n^{m}] \tag{1}$$

where $Da(m)$ is the allelic diversity for a sample size m, k the total number of alleles, n the observed sample size ($n > m$) and n_i the number of alleles i, ($1 \leq i \leq k$).

The exact test for Hardy-Weinberg equilibrium, genotypic linkage disequilibrium and genic and genotypic structure were computed using the GENEPOP package version 1.2 (Raymond and Rousset, 1995). F-statistics were estimated according to Weir and Cockerham (1984).

Phylogenetic relationships between populations were established using the neighbour-joining (N-J) algorithm (Saitou and Nei, 1987) and the chord distance of Cavalli-Sforza and Edwards (1967), which is among the best genetic distances for recovering the correct tree topology according to Takezaki and Nei (1996). The distances of Goldstein et al. (1995) and Shriver et al. (1993), based on allele size, were also used for comparison. Bootstrap values were computed over 2 000 replications (Hedges, 1992) resampling loci or individuals. Trees of individuals were constructed with the Das (shared allele) distance (Chakraborty and Jin, 1993) and the N-J algorithm according to Bowcock et al. (1994) and Estoup et al. (1995a).

To approximate nuclear introgression in the populations, we estimated the frequency of alien alleles in *mellifera* populations for nine loci (all studied loci except *B124* and *Ap33*) for which one or several alleles seem to be diagnostic between the M and C lineages. The proportion of introgressed alleles (*IR*) was calculated by locus as follows:

$$IR = \sum_{i \in D} p_i / \sum_{i \in D} q_i \tag{2}$$

where p_i and q_i are the allelic frequencies in the different M populations and in the Chalkidiki population, respectively, and D is the set of diagnostic alleles at the locus. We tested the homogeneity of the introgression of diagnostic alleles per locus and per population using Fisher's exact test obtained through a Markov process according to Raymond and Rousset (1995).

3. RESULTS

3.1. Genetic diversity

Allelic frequencies for all locus × sample combinations are detailed in *table II* which also gives observed (Hp) and calculated (Hd) heterozygosities. The number of alleles per locus, the allelic diversities Da and the effective allele diversities $Da(m)$ are summarised in *table III* considering either the raw data or the data corrected for introgressed alleles. Compared to Chalkidiki and Al Hoceima, the *mellifera* and *iberica* populations show lower variability.

Among 187 (17 × 11) locus × sample combinations, 11 significant departures from the Hardy-Weinberg equilibrium were detected when nine are expected by chance alone at the 5 % level. Two significant tests were obtained for each of the following four populations: Fleckenstein, Angers, Avignon and Umeå.

Within populations, 70 pairs of loci present a significant linkage disequilibrium over 825 (55 × 15) tests. Two or three linkage disequilibria are expected to be significant by chance alone among the 55 tests performed between loci taken pairwise within each population. This value was exceeded, sometimes considerably, in four populations: Avignon (6), Bayonne (8), Sabres (10) and Angers (19). Ten cases of cytonuclear disequilibrium were detected over the 181 tests performed. The multiple probability test by pair of loci across all populations (Fisher's method) was significant for 12 nuclear pairs and one cytonuclear pair $(B124)$ over, respectively, 55 and 11 tests performed (populations, global, GENEPOP; Raymond and Rousset, 1995).

3.2. Diagnostic alleles

Most of the French populations of *mellifera* are slightly to heavily introgressed by C mitochondrial haplotypes (see Part I, accompanying article). We looked for a concomitant introgression of nuclear genes. For nine loci, it appears that some alleles that exist at intermediate or high frequency in the Chalkidiki population, the reference population of the C lineage, are also present with a variable frequency in *A. m. mellifera* populations. This is, for example, the case for the large alleles (158 to 166) at locus *A8* or for allele 141 at locus *A43* (*table II*). These alleles are generally located at one extreme of the size distribution of the M lineage. Moreover, they have a very high frequency in *A. m. mellifera* populations which also exhibit a high frequency of C mtDNA haplotypes, namely Angers and Fleckenstein (Garnery et al., accompanying paper) and a low frequency in the other *mellifera* populations which have a low frequency of C haplotypes, whereas they are absent or nearly so in *A. m. iberica*, which are not introgressed by C haplotypes. It would be interesting to confirm

Table II. Allele frequency at 11 loci in 17 honey bee populations.

Locus B124 Allele (pb)	Castilla (n=42)	Sevilla (n=56)	SanSebas. (n=50)	Porto (n=56)	Sabres (n=100)	Bayonne (n=84)	Saintes (n=58)	Angers (n=98)	Ouessant (n=98)	Montfavet (n=60)	Annecy (n=100)	Fleckenstein (n=60)	Valenciennes (n=58)	Chimay (n=60)	Umeå (n=44)	Chalkidiki (n=60)	Al Hoceima (n=56)
212					0.010												
214					0.010												
216			0.040		0.020											0.100	
218				0.018		0.012	0.121	0.020			0.040	0.017	0.034	0.017	0.045	0.133	0.071
220	0.048		0.160		0.080	0.179	0.034	0.184			0.040	0.367	0.017			0.083	
222	0.381	0.357	0.320	0.286	0.090	0.095	0.155	0.092	0.082	0.433	0.240	0.167	0.241	0.333	0.045	0.233	0.232
224	0.167	0.143	0.160	0.250	0.290	0.143	0.086	0.214	0.204	0.100	0.120	0.100	0.086	0.067	0.114	0.067	0.125
226	0.143	0.036	0.060	0.232	0.110	0.167	0.034	0.122		0.067	0.180	0.033	0.155	0.117	0.159	0.167	0.036
228	0.048	0.161	0.060	0.089	0.070	0.083	0.034	0.010	0.041	0.033	0.080	0.067	0.017	0.067	0.023	0.050	0.071
230	0.071	0.089	0.060	0.036	0.100	0.060	0.121	0.071	0.194	0.050	0.070	0.083	0.086	0.100		0.017	0.107
232	0.095	0.089	0.100	0.018	0.040	0.095	0.052	0.061	0.071	0.133	0.040	0.067	0.069	0.017	0.136	0.033	0.018
234	0.024	0.071	0.020	0.018	0.080	0.036	0.034	0.102	0.286	0.100	0.060	0.017	0.086	0.050	0.205	0.017	0.054
236	0.024	0.018		0.018	0.030	0.071	0.069	0.071		0.033	0.070	0.033	0.069	0.050	0.182	0.017	0.036
238		0.036		0.018	0.020	0.036	0.190	0.020	0.122	0.017	0.020	0.033	0.069	0.050	0.068	0.067	0.018
240					0.030	0.012	0.052	0.010		0.033	0.030		0.017	0.050		0.017	0.089
242			0.020		0.020	0.012					0.010	0.017	0.034	0.017			0.054
244							0.017	0.010	0.010				0.017	0.017			0.018
246																	0.018
250																	0.018
Hp	0.810	0.821	0.800	0.929	0.900	0.833	0.931	0.837	0.755	0.800	0.820	0.733	0.828	0.800	0.864	0.967	0.929
Hd	0.806	0.817	0.840	0.805	0.872	0.894	0.904	0.880	0.819	0.777	0.878	0.821	0.893	0.857	0.878	0.884	0.905

Locus A8 Allele (pb)	Castilla (n=42)	Sevilla (n=56)	SanSebas. (n=50)	Porto (n=56)	Sabres (n=100)	Bayonne (n=84)	Saintes (n=58)	Angers (n=98)	Ouessant (n=100)	Montfavet (n=60)	Annecy (n=100)	Fleckenstein (n=60)	Valenciennes (n=58)	Chimay (n=60)	Umeå (n=44)	Chalkidiki (n=60)	Al Hoceima (n=56)
116																0.017	
148	0.929	0.893	0.920	0.946	0.910	0.929	0.862	0.776	0.960	0.917	0.790	0.267	0.793	0.783	0.727		0.089
150				0.018													0.018
152			0.020														
154												0.083					0.214
155												0.033	0.017				0.446
156	0.071	0.107	0.040	0.036	0.070	0.036	0.017	0.061	0.020	0.017	0.110	0.300	0.138	0.133	0.227	0.100	0.143
157												0.050					
158*						0.024	0.034	0.061		0.017	0.050	0.117				0.350	0.089
159			0.020									0.017			0.023		
160*					0.020	0.012	0.034	0.071		0.017	0.010	0.100	0.017	0.050		0.317	
162*							0.034	0.020	0.020	0.017	0.020		0.017	0.033		0.117	
164*							0.017	0.010		0.017	0.020	0.017	0.017		0.023	0.083	
166*												0.017				0.017	
Hp	0.143	0.071	0.160	0.107	0.140	0.119	0.276	0.388	0.080	0.167	0.320	0.667	0.414	0.367	0.364	0.733	0.750
Hd	0.136	0.195	0.154	0.105	0.168	0.137	0.257	0.389	0.078	0.161	0.364	0.818	0.357	0.371	0.428	0.759	0.731

Table II. Continued.

Locus B24 Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 84)	Saintes (n = 58)	Angers (n = 98)	Ouessant (n = 100)	Montfavet (n = 60)	Annecy (n = 100)	Fleckenstein (n = 56)	Valenciennes (n = 58)	Chimay (n = 60)	Umeå (n = 44)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
97																	0.036
98	0.905	0.929	0.960	0.982	0.910	0.833	0.862	0.653	0.920	0.917	0.780	0.393	0.810	0.850	0.864		0.679
101																	0.071
102																	0.018
104																0.017	0.054
106	0.095	0.071	0.040	0.018	0.090	0.143	0.138	0.194	0.080	0.050	0.140	0.304	0.138	0.117	0.091	0.467	0.125
108*						0.024		0.153		0.033	0.080	0.304	0.052	0.033	0.045	0.517	0.018
Hp	0.191	0.143	0.080	0.036	0.140	0.214	0.276	0.571	0.160	0.100	0.420	0.643	0.276	0.233	0.182	0.567	0.571
Hd	0.176	0.135	0.078	0.036	0.165	0.288	0.242	0.518	0.149	0.159	0.369	0.673	0.327	0.267	0.250	0.524	0.523

Locus A43 Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 84)	Saintes (n = 58)	Angers (n = 98)	Ouessant (n = 100)	Montfavet (n = 60)	Annecy (n = 100)	Fleckenstein (n = 60)	Valenciennes (n = 56)	Chimay (n = 60)	Umeå (n = 44)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
126												0.100					0.018
127*																	
128	0.929	0.893	0.880	0.964	0.800	0.810	0.810	0.602	0.790	0.750	0.740	0.367	0.804	0.900	0.682	0.083	0.286
130										0.017							
132																	0.018
133				0.018						0.017							
134	0.024	0.036															0.054
135			0.020														
136																	
138																	0.071
139																0.033	0.018
140	0.048		0.080	0.018	0.190	0.143	0.155	0.173	0.210	0.150	0.180	0.150	0.143	0.083	0.295		0.054
141*		0.071	0.020		0.010	0.048	0.034	0.184		0.050	0.060	0.350	0.054	0.017	0.023	0.850	
142								0.020				0.017					
143								0.020			0.010	0.017					
144																0.017	
145											0.010						0.143
146										0.017						0.017	
148																	0.268
150																	0.071
Hp	0.143	0.214	0.240	0.071	0.320	0.333	0.379	0.490	0.340	0.400	0.380	0.700	0.321	0.200	0.455	0.267	0.821
Hd	0.138	0.200	0.223	0.071	0.327	0.326	0.324	0.579	0.335	0.419	0.420	0.722	0.337	0.186	0.458	0.273	0.824

Table II. Continued.

Locus Ap33 / Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 84)	Saintes (n = 58)	Angers (n = 98)	Ouessant (n = 100)	Montfavet (n = 60)	Annecy (n = 100)	Fleckenstein (n = 60)	Valenciennes (n = 58)	Chimay (n = 60)	Umeå (n = 42)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
205	0.024							0.010									0.054
223				0.018										0.017			
225			0.020	0.071							0.020			0.150			
227	0.238	0.125	0.120	0.036	0.220	0.440	0.241	0.133	0.020	0.250	0.130	0.117	0.138	0.017	0.190		
229	0.119	0.071	0.140	0.018	0.020	0.036		0.031		0.017	0.030	0.017					0.036
231	0.024	0.036		0.089		0.012				0.017	0.010	0.050		0.033		0.017	0.018
233	0.024	0.036	0.020							0.017	0.010	0.017					0.054
235		0.018	0.020		0.010	0.012	0.017	0.031		0.100		0.083	0.017		0.024	0.033	0.018
237	0.024	0.018	0.040	0.071	0.040	0.024	0.017	0.010	0.120	0.050	0.010	0.233	0.034	0.017		0.133	0.018
239	0.095	0.161	0.040	0.071	0.070	0.083	0.052	0.061	0.030	0.050	0.030	0.117	0.017	0.017	0.071	0.133	0.143
241	0.167	0.143	0.060	0.018	0.100	0.095	0.069	0.133	0.330	0.067	0.120	0.100	0.034	0.050	0.024	0.217	0.161
243	0.048	0.036	0.040	0.089	0.150	0.095	0.069	0.051	0.260	0.083	0.160		0.172	0.167	0.167	0.167	0.161
245	0.048	0.125	0.220	0.339	0.090	0.036	0.155	0.184	0.210	0.167	0.090		0.172	0.150	0.024	0.117	0.125
247	0.167	0.107	0.220	0.125	0.140	0.060	0.155	0.184	0.010	0.133	0.100	0.067	0.138	0.200	0.143	0.017	0.071
249		0.107	0.040	0.054	0.130	0.083	0.086	0.031		0.017	0.160	0.200	0.121	0.100	0.143	0.033	0.036
251	0.024	0.018	0.020		0.030	0.024	0.052	0.071		0.017	0.090		0.086	0.067	0.048	0.117	0.054
253							0.017	0.031	0.020		0.030		0.052		0.071		0.036
255							0.069	0.010			0.010				0.071		0.018
257													0.017	0.017	0.024	0.017	
259								0.020									
263								0.010									
Hp	0.878	0.906	0.875	0.848	0.875	0.776	0.881	0.890	0.771	0.879	0.897	0.868	0.890	0.883	0.897	0.874	0.910
Hd	0.857	0.964	0.880	0.714	0.820	0.810	0.931	0.959	0.800	0.867	0.940	0.900	0.828	0.867	0.905	0.900	0.929

Table II. Continued.

Locus A7 Allele (pb)	Castilla (n=42)	Sevilla (n=56)	SanSebas. (n=50)	Porto (n=56)	Sabres (n=100)	Bayonne (n=84)	Saintes (n=58)	Angers (n=98)	Ouessant (n=100)	Montfavet (n=60)	Annecy (n=100)	Fleckenstein (n=60)	Valenciennes (n=54)	Chimay (n=60)	Umeå (n=44)	Chalkidiki (n=60)	Al Hoceima (n=56)
98																0.017	0.018
103												0.017					0.018
104		0.018								0.017							
105								0.020				0.017				0.033	0.018
106	0.048			0.018		0.048							0.019	0.050		0.017	0.109
107	0.167	0.143	0.140	0.089	0.070			0.031		0.033	0.040	0.017					0.055
108						0.012											0.236
109	0.143	0.071	0.100	0.161		0.048					0.010	0.017			0.045		0.073
110						0.024											
111	0.595	0.732	0.720	0.732	0.810	0.810	0.828	0.653	1.000	0.783	0.880		0.907	0.917	0.864		0.309
113	0.048	0.036			0.090	0.048	0.034	0.102		0.117	0.050		0.037		0.068		0.055
114																0.017	
115														0.033			
116*							0.017	0.061				0.233				0.517	0.036
118*					0.020		0.034	0.082		0.017		0.150	0.019			0.133	
120*												0.033	0.019		0.023	0.017	
122*						0.012						0.050				0.033	
123*							0.017	0.020								0.017	
126*																0.050	
128*											0.010					0.017	
130*							0.017			0.017						0.050	
132*			0.020				0.017									0.017	
134								0.010								0.017	
135*							0.034										
136																0.017	
137*					0.010												
140								0.010			0.010						
142*								0.010								0.017	
148										0.017							
150																	
156*								0.010								0.017	
158																	0.036
160																	0.036
Hp	0.524	0.500	0.440	0.536	0.320	0.262	0.276	0.449	0.000	0.400	0.200	0.800	0.111	0.167	0.136	0.633	0.815
Hd	0.607	0.445	0.460	0.438	0.334	0.341	0.316	0.556	0.000	0.377	0.223	0.768	0.178	0.159	0.253	0.717	0.841

Table II. Continued.

Locus A88

Allele (pb)	Castilla (n=42)	Sevilla (n=56)	SanSebas. (n=50)	Porto (n=56)	Sabres (n=100)	Bayonne (n=84)	Saintes (n=58)	Angers (n=98)	Ouessant (n=100)	Montfavet (n=60)	Annecy (n=100)	Fleckenstein (n=60)	Valenciennes (n=58)	Chimay (n=60)	Umeå (n=44)	Chalkidiki (n=60)	Al Hoceima (n=56)
138																0.017	
143	0.024				0.010		0.017	0.163		0.033	0.040					0.217	0.089
146	0.952	0.982	0.980	0.982	0.960	1.000	0.948	0.806	0.970	0.950	0.880	0.400	0.983	0.967	1.000	0.033	0.339
148				0.018								0.017					
150												0.050					0.304
152*								0.010			0.010	0.500		0.033		0.133	0.036
153																	0.161
154*	0.024		0.020		0.030		0.034	0.020	0.030	0.017	0.070		0.017			0.600	0.071
155		0.018										0.033					
156																	
Hp	0.095	0.036	0.040	0.036	0.080	0.000	0.104	0.204	0.060	0.100	0.240	0.533	0.034	0.067	0.000	0.567	0.786
Hd	0.094	0.036	0.040	0.036	0.078	0.000	0.101	0.326	0.059	0.098	0.221	0.596	0.034	0.066	0.000	0.584	0.766

Locus A113

Allele (pb)	Castilla (n=42)	Sevilla (n=56)	SanSebas. (n=50)	Porto (n=56)	Sabres (n=100)	Bayonne (n=84)	Saintes (n=58)	Angers (n=98)	Ouessant (n=100)	Montfavet (n=60)	Annecy (n=100)	Fleckenstein (n=60)	Valenciennes (n=58)	Chimay (n=60)	Umeå (n=44)	Chalkidiki (n=60)	Al Hoceima (n=56)
202	0.024	0.018	0.100		0.100	0.119	0.103	0.092	0.120	0.267	0.150	0.017	0.172	0.083	0.024		0.196
204										0.017							0.018
208						0.012	0.017										0.071
210																	0.089
212					0.030	0.012											
214*					0.010												0.446
216		0.018					0.017	0.020	0.020	0.033	0.050	0.500	0.034	0.033		0.800	
218	0.048	0.018	0.020	0.036	0.020		0.017	0.010	0.010			0.017					
220	0.667	0.571	0.820	0.607	0.680	0.714	0.621	0.643	0.810	0.517	0.720	0.433	0.724	0.833	0.857	0.017	0.054
222	0.095	0.107	0.020	0.107	0.040	0.071	0.017	0.041		0.033	0.030		0.017		0.024		0.054
224	0.071	0.232	0.020	0.143	0.060		0.034	0.010		0.033	0.010			0.017	0.048		
226	0.071	0.036	0.020	0.071			0.034	0.031	0.040	0.017					0.048		
228					0.030	0.024	0.069	0.071		0.033	0.020	0.017	0.034	0.017		0.050	0.054
230					0.020	0.036	0.052	0.051			0.020	0.017	0.017			0.017	0.018
232						0.012	0.017									0.067	
234	0.024			0.018	0.010			0.020		0.017				0.017		0.033	
236										0.017							
238										0.017						0.017	
240								0.010									
Hp	0.619	0.714	0.320	0.571	0.580	0.405	0.586	0.531	0.340	0.733	0.400	0.467	0.483	0.300	0.190	0.367	0.750
Hd	0.546	0.617	0.322	0.603	0.525	0.474	0.603	0.573	0.331	0.672	0.459	0.571	0.451	0.302	0.266	0.357	0.753

Table II. Continued.

Locus Ap36 Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 76)	Saintes (n = 58)	Angers (n = 95)	Ouessant (n = 100)	Montfavet (n = 58)	Annecy (n = 100)	Fleckenstein (n = 60)	Valenciennes (n = 56)	Chimay (n = 60)	Umeå (n = 44)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
125*			0.020		0.030	0.026	0.034	0.095						0.050	0.091	0.817	0.089
127	0.024										0.050	0.500	0.018	0.017		0.017	0.143
129																	0.071
131																	0.036
133			0.020					0.042									0.089
137																	
139																	
141					0.010												0.089
143	0.310	0.411	0.020	0.482	0.080	0.013	0.052	0.053	0.050		0.040	0.050	0.089	0.033			0.054
145	0.310	0.143	0.120	0.304	0.010	0.039		0.021	0.010			0.017		0.017			0.357
147	0.095	0.054	0.040					0.042		0.103		0.017	0.018	0.017			0.018
149							0.017	0.032			0.010				0.023		
151				0.018							0.010					0.050	
153	0.024		0.020	0.018		0.013	0.034	0.021		0.017						0.050	0.018
155		0.018	0.020	0.018	0.030		0.017									0.033	
157														0.017		0.033	
159	0.048		0.060	0.036	0.100	0.079	0.034	0.021	0.060	0.069	0.040	0.017	0.125	0.017	0.045		
161		0.071	0.080	0.018	0.100	0.132	0.138	0.105	0.060	0.069	0.100	0.017	0.161	0.100	0.182		
163	0.119	0.071	0.160	0.018	0.230	0.276	0.276	0.074	0.260	0.103	0.200	0.117	0.107	0.117	0.364		0.036
165	0.048	0.125	0.180	0.054	0.180	0.276	0.276	0.137	0.140	0.276	0.260	0.067	0.179	0.117	0.205		
167	0.024	0.036	0.140	0.018	0.110	0.053	0.086	0.189	0.180	0.103	0.170	0.083	0.179	0.183			
169			0.100	0.018	0.060	0.066	0.034	0.137	0.030	0.034	0.090	0.050	0.089	0.150			
171		0.036	0.020		0.010	0.026		0.032	0.210	0.103	0.020	0.050	0.018	0.083			
173		0.036			0.020						0.010			0.017	0.023		
175												0.017		0.017	0.068		
177					0.010					0.017							
179													0.018	0.050			
181					0.010												
195																	
Hp	0.714	0.821	0.880	0.643	0.840	0.684	0.759	0.837	0.860	0.793	0.800	0.700	0.893	0.933	0.727	0.333	0.679
Hd	0.798	0.792	0.902	0.681	0.879	0.824	0.828	0.847	0.834	0.874	0.848	0.728	0.882	0.907	0.795	0.331	0.832

Table II. Continued.

Locus Ap43 Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 84)	Saintes (n = 58)	Angers (n = 98)	Ouessant (n = 100)	Montfavet (n = 58)	Annecy (n = 100)	Fleckenstein (n = 60)	Valenciennes (n = 58)	Chimay (n = 60)	Umeå (n = 42)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
101																	
113																	
125								0.010									
127		0.018										0.017					0.036
129																	0.054
131		0.018								0.017							0.143
133										0.017							0.107
134	0.024																
135	0.286	0.339	0.260	0.518	0.200	0.190	0.241	0.286	0.020	0.207	0.190	0.183	0.328	0.400	0.190	0.133	0.179
136					0.020	0.012											
137	0.405	0.429	0.600	0.268	0.750	0.690	0.672	0.541	0.980	0.672	0.680	0.333	0.569	0.483	0.690	0.017	0.179
139	0.071	0.018	0.020		0.010	0.048	0.017	0.051		0.017	0.030	0.017	0.052	0.033		0.033	0.071
141								0.051				0.117					0.036
143*							0.017			0.017	0.020		0.017	0.033		0.267	
144																0.017	
145*							0.034				0.040		0.034	0.050	0.119	0.433	
147*					0.020	0.012										0.067	
149				0.036								0.283					
151	0.024			0.018								0.050					0.018
153	0.024		0.020														0.018
155											0.010						
157	0.071	0.054	0.060	0.054													
159	0.024	0.018		0.018													0.018
161	0.024			0.036													0.018
163	0.048	0.018	0.040	0.018		0.012	0.017	0.010									0.018
165		0.018		0.018				0.010									0.018
167		0.018				0.012		0.010									0.018
169								0.010		0.017						0.017	0.036
171																	0.018
173				0.018												0.017	
177											0.010						
181		0.036						0.010			0.010						0.018
183																	
185		0.018			0.012	0.012				0.017	0.010						
187																	
195																	
Hp	0.757	0.707	0.578	0.664	0.401	0.490	0.496	0.626	0.040	0.511	0.503	0.771	0.575	0.612	0.484	0.729	0.906
Hd	0.714	0.750	0.600	0.750	0.500	0.548	0.586	0.531	0.040	0.586	0.500	0.800	0.724	0.533	0.428	0.600	0.857

Table II. Continued.

Locus A28	Allele (pb)	Castilla (n = 42)	Sevilla (n = 56)	SanSebas. (n = 50)	Porto (n = 56)	Sabres (n = 100)	Bayonne (n = 84)	Saintes (n = 58)	Angers (n = 98)	Ouessant (n = 100)	Montfavet (n = 60)	Annecy (n = 100)	Fleckenstein (n = 56)	Valenciennes (n = 58)	Chimay (n = 60)	Umeå (n = 44)	Chalkidiki (n = 60)	Al Hoceima (n = 56)
	126	0.024																0.018
	127	0.024																0.018
	128																	0.107
	129																	0.054
	130									0.050								0.054
	131																	0.196
	132	0.905	0.982	0.900	1.000	0.880	0.810	0.879	0.735	0.840	0.783	0.870	0.357	0.966	0.900	0.977	0.100	0.232
	133			0.060		0.010												0.125
	134	0.048	0.018	0.020		0.030	0.024	0.052	0.163		0.067	0.090	0.643	0.034	0.100	0.023	0.883	0.161
	135						0.060										0.017	0.036
	138*			0.020		0.080	0.107	0.069	0.102	0.110	0.150	0.040						
	140																	
	144																	
	H_p	0.143	0.036	0.200	0.000	0.240	0.357	0.241	0.347	0.320	0.433	0.180	0.571	0.069	0.200	0.045	0.200	0.893
	H_d	0.182	0.036	0.189	0.000	0.220	0.333	0.224	0.428	0.283	0.366	0.236	0.468	0.068	0.183	0.045	0.213	0.862

The number of sampled chromosomes is given in parentheses. H_d: heterozygosity observed for each locus; H_p: heterozygosity assuming the Hardy-Weinberg equilibrium (gene diversity). Asterisks indicate putative diagnostic alleles.

Table III. Allelic diversity at 11 microsatellite loci in 17 honey bee populations with rough data and their average (*Da*).

Locus	Cast n = 42	Sevi n = 56	San S n = 50	Port n = 56	Sabr n = 100	Bayo n = 84	Sain n = 58	Ange n = 98	Oues n = 100	Avig n = 60	Anne n = 100	Flec n = 60	Vale n = 58	Chim n = 60	Umeo n = 44	Chal n = 60	Al Ho n = 56
B124	9	9	10	11	15	13	13	14	7	10	13	12	11	14	10	13	15
A88	3	2	2	2	3	1	3	4	2	3	4	5	2	2	1	5	6
A8	2	2	4	3	3	4	6	6	3	6	6	10	6	4	4	7	6
B24	2	2	2	2	2	3	2	3	2	3	3	3	3	3	3	3	7
A43	3	3	4	3	3	3	3	5	2	6	5	6	3	3	3	5	10
Ap33	12	13	13	12	11	12	12	16	8	14	15	10	12	13	12	11	15
A7	5	5	5	4	5	7	8	10	1	7	6	11	5	3	4	17	12
A113	7	7	6	7	10	8	11	11	5	11	7	6	6	6	5	7	9
Ap36	9	10	14	11	16	11	11	14	9	11	12	12	11	16	8	6	11
Ap43	10	12	6	10	5	9	6	11	2	9	9	7	5	5	3	9	18
A28	4	2	4	1	4	4	3	3	3	3	3	2	2	2	2	3	10
Da	6.00	6.09	6.36	6.00	7.00	6.82	7.09	8.82	4.00	7.55	7.55	7.63	6.00	6.45	5.00	7.82	10.82
Da Ef	6.00	5.59	5.95	5.30	5.62	6.65	6.48	6.83	4.38	6.18	6.12	7.22	6.04	5.58	4.96	7.67	10.68
Da Cor	5.91	6.09	5.82	6.00	6.18	6.091	5.727	7.273	3.727	6.273	6.273	5.818	5.000	5.545	4.273	7.82	10.82
Da CEf	5.91	5.59	5.59	5.31	5.19	5.15	5.68	5.72	3.38	5.69	5.05	ND	4.73	5.01	4.26	6.82	10.02

Da Cor: average after correction for introgressed alleles; *Da* Ef and *Da* CEf: two preceding values corrected for the sample size.

this interpretation by more powerful analyses (e.g. sequences of flanking regions of microsatellites). In addition, there is also a good correlation ($r = 0.74$, $P < 0.01$) between the frequency of C haplotypes and C nuclear genes. Several other features, which emerge when these alleles are removed (see later, linkage disequilibrium and phylogeny of populations), suggest that these alleles were secondarily introduced into *mellifera* populations.

Consequently, as long as the population of Chalkidiki is representative of the donor populations, these alleles are indicative of introgression by nuclear genomes from lineage C. They provide a conservative value of the frequency of nuclear introgression, because other alien alleles may have remained undetected.

Considering 29 diagnostic alleles between the M and C lineages (noted with an asterisk in *table II*), we estimated the proportion of introgressed alleles (*IR*) in the *iberica* and *mellifera* populations by locus and by population (*table IV*). The mean *IR* values by population are generally congruent with the proportions of mitochondrial haplotypes C, even if the latter are usually larger. The introgression rates (*IR*) were significantly different among populations and loci (Fisher's exact test, $P < 10^{-3}$ for both tests), mainly because the Fleckenstein population is more introgressed than the other populations and because of loci *A88* and *A113* (Fisher's exact test, $P = 0.075$ without Fleckenstein and $P = 0.176$ without *A88* and *A113*, respectively).

To test the influence of introgressed alleles on the populations, the 29 diagnostic alleles were disregarded in the matrix of individual multilocus genotypes. A re-analysis of linkage disequilibrium for the *mellifera* and *iberica* populations showed a general decrease of linkage disequilibrium: 29, i.e. 3.9 %, were significant instead of 70, i.e. 8.5 % (the number of possible tests was reduced from 825 to 740), and the global test by pair of loci across all populations was significant for only one pair of loci instead of 12 pairs. The allelic diversity was also re-calculated, ignoring diagnostic alleles (*Da* Cor, *table III*). In most cases, resulting genetic profiles in *mellifera* populations were closer to those observed in *iberica* samples which are supposed to be free of C genes.

3.3. Genetic distances and population relationships

The N-J tree of populations obtained with raw data (not corrected for introgression) shows several features (*figure 1*). Only two groups are supported by high bootstrap values. First, the four Spanish and Portuguese populations (*A. m. iberica*) form a cluster in which the population from San Sebastian branches basically. Second, Valenciennes and Chimay populations branch together, providing another example of the close relationship between small genetic distance and geographic proximity. The nodes of the other French populations are generally not well resolved but it is striking to observe that the more introgressed they are, the deeper they branch in the tree. Fleckenstein, the most heavily introgressed population, is located out of the *mellifera–iberica* set and clusters with Chalkidiki.

Disregarding diagnostic alleles in *mellifera* populations (excluding Fleckenstein) results in a new N-J tree in which the preceding two robust clusters (Iberian populations and Chimay/Valenciennes) are conserved and the other

Table IV. Frequency ($\times 1\,000$) of introgressed alleles in 13 populations of the M lineage for the nine loci that exhibit diagnostic alleles.

Population	Locus									Mean	C mtDNA
	A8	A24	A43	A7	A88	AA13	Ap36	Ap43	A28		
Castilla	0	0	0	0	33	0	0	0	0	0	0
Sevilla	0	0	0	0	0	0	0	0	0	0	0
San Sebastian	23	0	21	21	27	0	24	0	23	15	40
Porto	0	0	0	0	0	0	0	0	0	0	20
Sabres	23	0	11	32	41	13	37	26	34	24	20
Bayonne	41	46	51	13	0	0	32	15	68	30	70
Saintes	134	0	36	145	46	21	42	65	59	57	70
Angers	183	296	197	185	41	25	116	130	185	151	880
Ouessant	23	0	0	0	41	25	0	0	0	10	0
Avignon	69	64	54	36	23	41	0	43	76	45	140
Annecy	113	155	64	21	109	63	61	77	102	85	20
Fleckenstein	284	588	482	498	705	625	612	574	728	566	700
Valenciennes	58	101	58	41	23	43	22	43	39	48	70
Chimay	94	64	18	0	47	41	61	106	113	60	140
Umeå	52	87	25	25	0	0	111	152	26	53	70
Total	1 097	1 401	1 018	1 017	1 136	897	1 118	1 231	1 453	1 148	

The average of C microsatellites is highly correlated with the frequency of C haplotypes (last two columns).
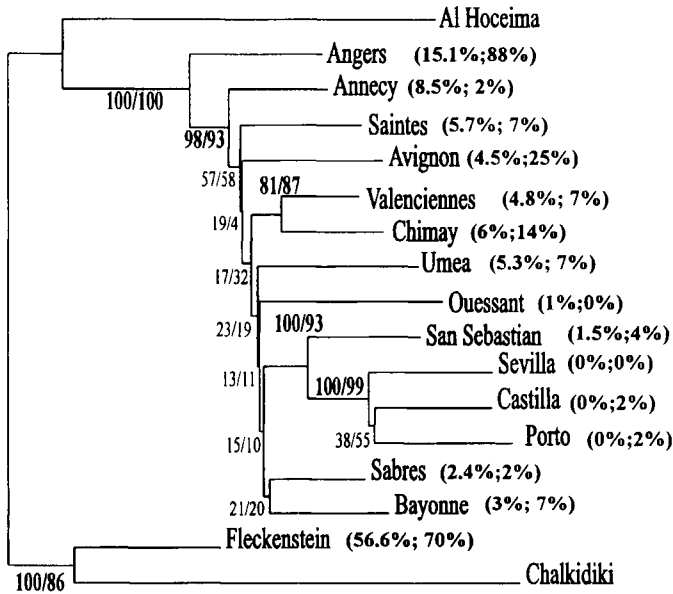
Figure 1. Phylogenetic tree of the populations. The tree has been constructed using observed allelic frequencies (rough data), the chord distance (Cavalli-Sforza and Edwards, 1967) and the neighbour-joining algorithm (Saitou and Nei, 1987). Values on the nodes indicate the bootstrap scores (2 000 iterations) resampling loci/individuals. The frequency of microsatellite C alleles and C haplotypes is given for M populations.
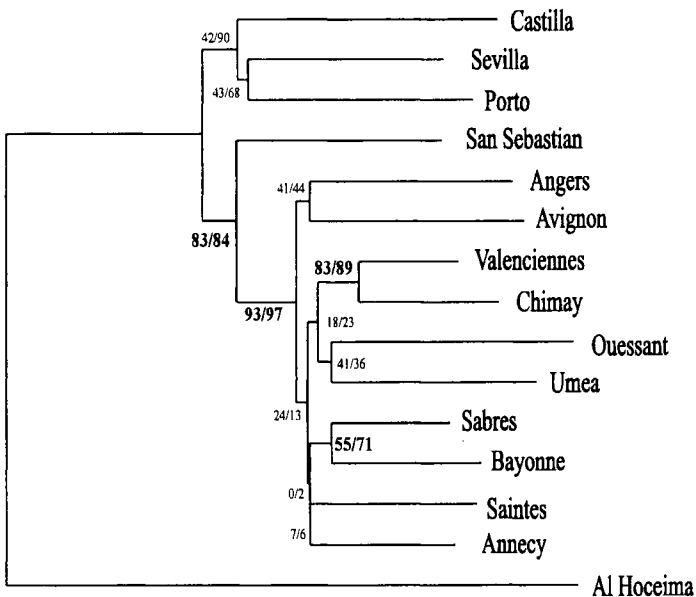

Figure 2. Phylogenetic tree of the populations with data corrected for introgression with the same methods as in *figure 1*.

populations remain rather unresolved (*figure 2*). However, the new feature is the basal divergence between the two racial sets of populations (*iberica* and *mellifera*), with San Sebastian in the middle. Within the *mellifera* cluster, no particular structure emerges, apart from the already cited Chimay/Valenciennes association. The same results, distinction between *iberica* and *mellifera* populations and absence of within-subspecies structure, were obtained with other genetic distances, including those based on allele size (i.e. Shriver et al., 1993; Goldstein et al., 1995).

Multilocus *Fst* calculated with raw allelic frequencies (Weir and Cockerham, 1984) are generally significant at the 5 % level (only four of 105 pairs are not significant, namely Valenciennes/Chimay, Annecy/Valenciennes, Castilla/Sevilla, Saintes/Valenciennes), but their values are rather low (*table V*). Two populations are particularly differentiated: Fleckenstein ($0.13 < Fst < 0.31$) and Ouessant ($0.04 < Fst < 0.18$, excluding Fleckenstein), the former because it is highly introgressed and the latter because it is an artificial, selected, tiny and isolated population. The A and C outgroup populations are very distinct from the M lineage. All loci contribute to the genetic differentiation, the lower contribution being that of loci *A8*, *A88*, *A24* and *A28*. With corrected data, changes in Fst are small for the least introgressed populations, but the decrease is important for Fleckenstein. The pair Sabres/Annecy and the four loci cited earlier are not significantly differentiated.

Another way to analyse genetic structure is to consider the relationships of individuals on a tree (*figure 3*), which has been constructed only with a subset of populations to facilitate legibility. If individuals of the A (outgroup) and C lineage form two compact clusters, those of the M lineage are intermingled in their branches, whatever their race (*mellifera* or *iberica*), even if there are some groups of individuals from the same population. It can also be seen that terminal branches which lead to individuals and internal ones which connect them are very short: this reflects the low genetic variability in these populations and the small genetic differentiation among them.

4. DISCUSSION

This discussion bears mainly on the genetic variability observed for microsatellite loci but it also considers them in relation to results obtained with mitochondrial DNA described in Part I, the accompanying article.

As aforementioned, a significant but unequal part of the variability within populations and of the differentiation among populations is the direct consequence of their levels of genetic introgression. Consequently, an attempt to estimate the genetic profiles of populations prior to the introduction of alien genes has to be considered before discussing other points.

4.1. Genetic introgression

The estimation of introgression in *mellifera* populations is based on the study of the nine loci (over 11) possessing putative diagnostic alleles. How these alleles were identified has been given in the results section. Several features suggest that this estimation of introgression is roughly correct: 1) there

Table V. Multilocus F_{st} for populations taken pairwise.

	Castilla	Sevilla	San Sebas.	Porto	Sabres	Bayonne	Saintes	Angers	Ouessant	Avignon	Annecy	Fleckenstein	Valenciennes	Chimay	Umeå	Chalcidi.	Alhocei.
Castilla		0.003 §	0.020**	0.021*	0.021*	0.051**	0.051**	0.051**	0.143**	0.064**	0.058**	0.221**	0.054*	0.056**	0.073**	0.463**	0.203**
Sevilla	0.003 §		0.032**	0.018*	0.048**	0.064**	0.051**	0.061**	0.133**	0.063**	0.058**	0.238**	0.047**	0.056**	0.072**	0.479**	0.216**
San Seb	0.020**	0.032**		0.060**	0.017*	0.026**	0.020*	0.036*	0.077**	0.030*	0.021*	0.231**	0.017*	0.017*	0.039**	0.484**	0.233**
Porto	0.021*	0.018*	0.060**		0.088**	0.109**	0.094**	0.088**	0.182**	0.105**	0.100**	0.271**	0.080**	0.081**	0.128**	0.506**	0.242**
Sabres	0.021*	0.048**	0.017*	0.088**		0.007*	0.007**	0.030*	0.041**	0.024**	0.007*	0.236**	0.012*	0.025**	0.017*	0.485**	0.246**
Bayonne	0.051**	0.064**	0.026**	0.109**	0.007*		0.008*	0.021**	0.073**	0.019*	0.013**	0.230**	0.021**	0.028*	0.026*	0.478**	0.267**
Saintes	0.051**	0.051**	0.020*	0.094**	0.007**	0.006*		0.022**	0.046**	0.013**	0.006*	0.202**	0.006 §	0.016*	0.019*	0.455**	0.236**
Angers	0.051**	0.061**	0.036*	0.088**	0.030*	0.021**	0.022**		0.068**	0.027**	0.016**	0.130**	0.022**	0.031*	0.038**	0.350**	0.163**
Ouessant	0.143**	0.133**	0.077**	0.182**	0.041**	0.074**	0.062**	0.084**		0.070**	0.002 §	0.308**	0.063**	0.084**	0.059**	0.551**	0.330**
Avignon	0.064**	0.063**	0.030*	0.105**	0.024**	0.022**	0.013**	0.027**	0.026**		0.017*	0.209**	0.018*	0.029**	0.051**	0.451**	0.206**
Annecy	0.058**	0.058**	0.021*	0.100**	0.007*	0.013**	0.006*	0.016**	0.067**	0.017*		0.192**	-0.001 §	0.009**	0.015*	0.438**	0.217**
Fleckenst	0.221**	0.238**	0.231**	0.271**	0.236**	0.088**	0.202**	0.130**	0.056**	0.209**	0.067**		0.208**	0.213**	0.211**	0.461**	0.172**
Valencien	0.054*	0.047**	0.017*	0.080**	0.012*	0.025**	0.006 §	0.022**	0.096**	0.018*	0.002 §	0.056**		-0.006 §	0.035*	0.467**	0.244**
Chimay	0.056**	0.056**	0.017*	0.081**	0.025**	0.034**	0.016*	0.031*	0.068**	0.029**	0.009**	0.074**	-0.006 §		0.019*	0.471**	0.263**
Umeå	0.073**	0.072**	0.039**	0.128**	0.017*	0.033**	0.019*	0.038**	0.035**	0.051**	0.015*	0.066**	0.030*	0.057**		0.502**	0.256**
Chalcidi	0.463**	0.479**	0.484**	0.506**	0.485**	0.501**	0.476**	0.350**	0.559**	0.451**	0.438**	0.430**	0.490**	0.505**	0.502**		0.245**
Al Hocei	0.203**	0.216**	0.233**	0.242**	0.246**	0.267**	0.236**	0.163**	0.330**	0.206**	0.217**	0.172**	0.244**	0.263**	0.256**	0.245**	

The values above the diagonal are calculated without correction for introgressed alleles and values below the diagonal after correction. *, **, §: level of significance $P < 0.05$, $P < 10^{-5}$ and P not significant, respectively.

Legend:
- Castilla
- Sabres
- Saintes
- Valenciennes
- Angers
- Fleckenstein
- Chalkidiki
- Al Hoceima
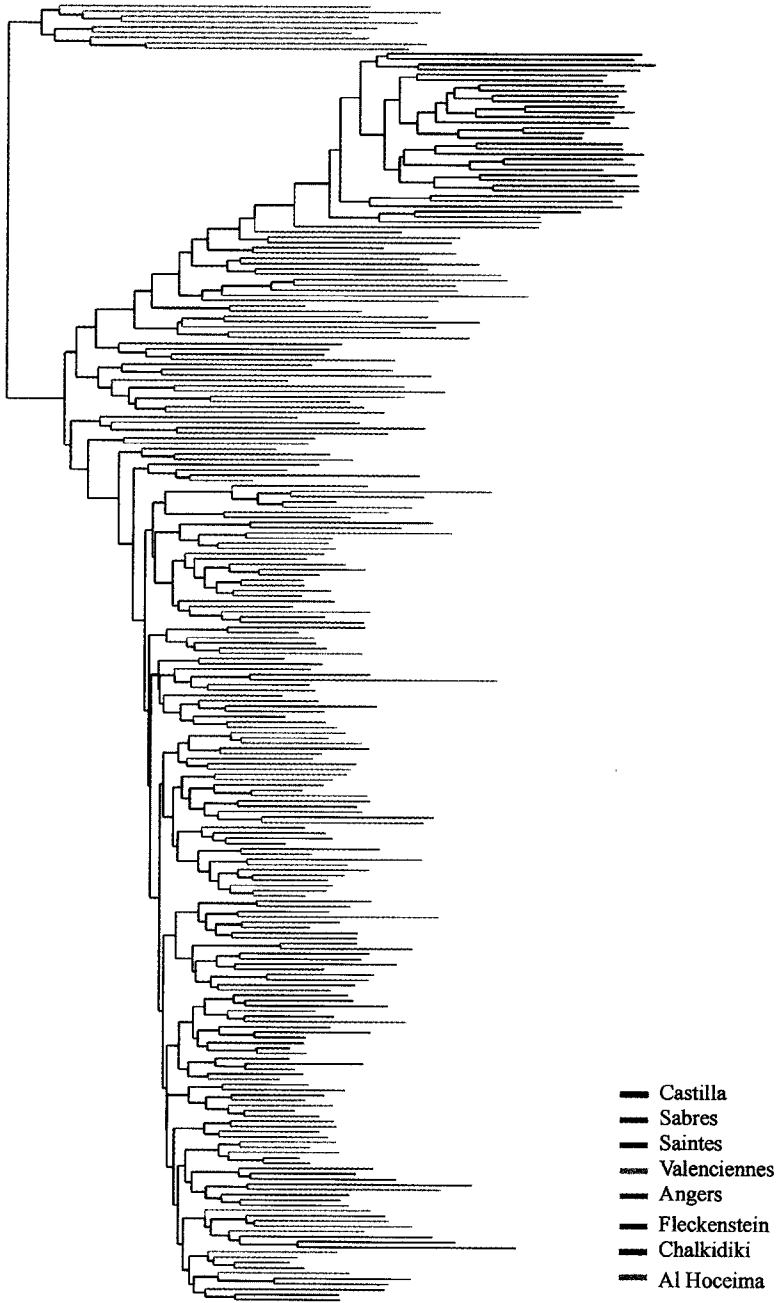
Figure 3. Branching pattern of individuals. The tree is based on the *Das* distance (Chakraborty and Jin, 1993) and the neighbour-joining algorithm (Saitou and Nei, 1987) for a subset of populations.

is a good correlation between the frequency of C haplotypes and C nuclear genes and both are also congruent with what is known about bee-keeping in these populations; 2) most of the linkage disequilibria disappear after the removal of the diagnostic alleles; 3) the genetic parameters (allelic diversity and heterozygosity) of *mellifera* and *iberica* populations (the latter being free of C haplotypes and hence supposedly free of C nuclear alleles) are more similar after correction and their genetic distance is lower.

Introgression was calculated with the allelic frequencies observed in the Chalkidiki population (Greece, lineage C); they are not necessarily identical in the other C populations from which foreign bees have been imported (mainly *ligustica* from Italy and *carnica* from Central Europe). A more accurate correction would have implied the use of allelic frequencies of the real donor population(s) for each recipient population, but they are generally unknown except that they belong to the C lineage. In addition, the genetic profiles for C populations are poorly known, only three samples being available (Estoup et al., 1995a) for six of the nine loci used. The average frequency of diagnostic alleles for these loci is 0.795 in the Chalkidiki population (*macedonica*) and only 0.616 for Berlin (*carnica*) and 0.688 for Forli (*ligustica*). A lower frequency in the last two populations is possibly due to not considering some of their alleles as diagnostic with lineage M, because they are absent in the Chalkidiki sample. This would result in a 10 to 20 % underestimation of nuclear introgression for some populations.

French, Belgian and Swedish populations analysed in this work show various degrees of introgression, mainly by alleles from the C lineage. The percentage of introgressed genes is roughly correlated with that of C mtDNA haplotypes previously established (see Part I, accompanying article and *table IV*) but is generally slightly lower. One possible reason is that some form of selection acts against the introgression of nuclear genes. Another reason could be a differential diffusion attributable to the role of the two sexes in their transmission. Still another is more likely: nuclear introgression could have been slightly underestimated as indicated earlier. There are, however, two exceptions. In the Annecy population, the relative excess of C nuclear genes observed (compared to C haplotype frequency) could be the result of a long-term equilibrium in a natural zone of contact between lineages M and C. For Angers, the percentage of C mtDNA haplotypes outnumbers that of the C microsatellite alleles whereas no such excess of mitochondrial genes is observed in the population of Fleckenstein, also heavily introgressed for mtDNA (see later for possible explanations).

Spanish and Portuguese populations seem to be almost deprived of nuclear alien genes, their genetic profiles being very similar to the least introgressed French populations. The absence of C nuclear genes is in agreement with the absence of C mtDNA but that of A nuclear genes is more surprising because the most southern populations carry a variety of A haplotypes at high frequency. This has been interpreted as the consequence of reiterated importation from various origins of queens belonging to the A lineage, and a differential introgression of cytoplasmic and nuclear genes attributable to the different swarming behaviour of African and European honey bees and a possible selection in favour of A haplotypes (Franck et al., 1998).

4.2. Genetic variability within populations

There is no general standard for the variability of microsatellite loci, but the comparison of the 15 populations of lineage M to populations of the C and A lineages (Greece and Morocco, respectively) shows that heterozygosity and the allele number are higher or far higher in non-M populations. A single population has been used as reference for the C and A lineages but, as far as we know, they are representative of the variability in these lineages (Estoup et al., 1995a; Franck et al., 1998). The lower genetic variability of M populations has been interpreted as a consequence of lower effective population sizes due to ecological and ethological differences (Estoup et al., 1995a).

It has also been shown (Estoup et al., 1995a) that the populations of the west European honey bees are not, whatever the mutation model (IAM or SMM), at mutation/drift equilibrium but, for almost all loci, given the number of alleles, they exhibit a deficit of estimated genic diversity. This is indicative of a recent increase in the population size (Franck et al., 1998). Moreover, as the genetic differentiation between populations, although often significant, is generally low, it can be inferred that the bottleneck was common for all populations (or two groups of populations, see section 4.4), probably during the last ice age, its effects still being perceptible in the current populations. It is difficult to invoke an alternative explanation, homogenisation by migration over more than 5 000 km, the west European honey bees being considered as poor migrants (compared to African populations). An important consequence is that the low level of variability is not the consequence of recent domestication of honey bees by humans but rather that of the natural history of populations.

4.3. Population differentiation and relationships

Whatever the means used to infer genetic similarity among populations of the west European honey bees (phylogeny of populations, *Fst*, tree of individuals) using microsatellite data, it appears that they are very similar to and very different from populations of the other two lineages A and C (see also Estoup et al., 1995a). In fact, the level of introgression is responsible for most of the differentiation observed (*figures 2* and *3*). The effect of alien alleles appears clearly in the comparison of population trees where the topology is largely modified between corrected and uncorrected data. It also appears in the topology of the tree constructed with individual genotypes: alien genes increase the length of the branches (populations or individuals) when their number is limited, then they displace the branch at the root of their own aboriginal lineage when this number increases, and finally transfer the branch at the root of the population from which introgressing genes come when it is high. The consequences of introgression are important to discuss but, whatever the purpose (natural history of the species or conservation of its populations), it is also interesting to consider the relationships among populations on the basis of their genetic structure prior to importation.

Fleckenstein and Angers are characterised by a particularly high level of introgression. In Fleckenstein (directly branched on the C lineage), the population belonged originally to the M lineage, but massive importation of

carnica, following reduction in the population size due to the parasite mite *Varroa* (P.P. Merck, personal communication), has completely modified their genetic profile, which is currently characterised by a majority of *carnica* genes (mitochondrial and nuclear). However, population genetics parameters (few departures from the Hardy-Weinberg equilibrium and few linkage disequilibria) indicate that local and imported honey bees currently constitute a single panmictic population. The population from Angers has been sampled in an area of traditional apiculture (i.e. using mainly local honey bees) but in the vicinity of a professional bee-keeper who has continuously imported *ligustica* queens. The introgression of mitochondrial genes is massive (88 %) but that of nuclear genes remains comparatively low (15 %). In addition, linkage disequilibrium is widespread (19 significant tests of 55). The discrepancy between these two populations could have several origins: 1) introgressing genes have not the same racial origin (*ligustica* for Angers and *carnica* for Fleckenstein) and could have been better detected in one population than in the other; 2) selection could act differentially and be strong against alien nuclear genes in Angers; 3) in the latter population a strong mitochondrial gene flow could be ensured by swarms escaped from the professional bee yards, but most queens were inseminated by local males in drone congregation areas. To a lesser degree, the same can be observed in the Avignon population (where *ligustica* have been imported for many years). In conclusion, it seems that when importations are reiterated every year, the mtDNA genes show a propensity to introgress more than nuclear genes, whereas introgression is equivalent for the two compartments when they have been massive over a short period of time.

Most of the other populations appear poorly resolved on the phylogenies but some groups are rather strong. The Iberian populations are always grouped together, San Sebastian emerging first, making this population, in agreement with its geographical location, a plausible intermediate between Iberian and French populations (*iberica* and *mellifera*). Iberian populations are less variable for microsatellite markers than French ones with raw data, but they appear to be the most variable when data from French populations are corrected for introgression and all populations for sample size. This result is expected because they are assumed to live in the refuge of the western honey bees during the ice age.

Another clade emerging from the phylogenies is in good agreement with geographic proximity (Valenciennes/Chimay). This clade was also seen by mtDNA analysis (see Part I, accompanying article).

4.4. Evolutionary origin of west European honey bees

The key geographical position of Iberian populations, their morphological and genetic peculiarities and their possible role in the origin of the M lineage of honey bees, have been discussed in detail elsewhere (Franck et al., 1998) and will only be briefly summarised here. Ruttner (1988) proposed that the west European honey bee *A. m. mellifera* was derived from the North African race *intermissa* through *iberica*, which has a morphology intermediate between these two races for numerous characters. Later, the detection of the superimposition of two very divergent mtDNAs designating a cline in the Iberian Peninsula suggested that this region was not a primary zone of intergradation but a

secondary contact zone between the very divergent haplotypes A and M (Smith et al., 1991; Garnery et al., 1995). Results based on microsatellite loci reject both hypotheses because all the honey bees classified in the M lineage exhibit very similar genetic profiles, with no trace of intergradation between lineages A and M, either primary or secondary, in Spain or Portugal (Franck et al., 1998). The existence of A haplotypes, very diversified and at high frequency in these countries, was reinterpreted as the consequence of multiple importation from different origins by humans and of a possible selective advantage (direct or indirect) of these haplotypes in the Iberian Peninsula.

4.5. Conservation

The genetic variability of west European populations of the honey bee is particularly low, compared to those belonging to clades C and A, both in terms of allelic diversity and heterozygosity. This variability is still lower than actually observed because a variable part, depending on the population, can be attributed to introgression and the standard in the whole lineage is close to that observed in the Iberian Peninsula. In addition, most of the populations are characterised by very similar profiles, the largest differences being observed between two clades matching the taxonomic limits between races. Their low variability is not the consequence of modern apiculture, ecological degradation of the environment or agricultural modifications, but of their recent natural history, probably in relation to paleoclimate variations. In spite of this relative homogeneity, populations have developed two races, several ecotypes and can survive in very different climates (from southern Spain to Sweden), all features which are indicative of high evolutionary potential for local adaptation despite their great similarity for neutral genes.

Quantification of genetic 'pollution' is assuredly a new result and is important to consider. Three factors can contribute to introgression of alien genes: queen import, queen rearing (from imported queens) and hive moving (ordered in decreasing importance). The last factor is difficult to detect because it occurs mainly within subspecies, except when professional bee-keepers move hives headed by imported queens. The importation of queens belonging to races such as *carnica*, *caucasica* and *ligustica* is rare among amateurs and is mainly due to professional bee-keepers who appreciate their quietness and several other supposed or real qualities.

In most localities, introgression remains rather low (below 10 %) and we have found only two heavily introgressed populations (Fleckenstein and Angers) corresponding to identified and recent apicultural managements. The situation in Angers is very asymmetrical for cytoplasmic and nuclear genes but in the south of the Iberian Peninsula it is still more extreme if our interpretation is correct. Consequently, even if the study of mtDNA is easier and more reliable, it is not sufficient in some particular cases to quantify the level of introgression by alien genes.

The two candidates for a conservatory of the local honey bees are located in two different clades of the lineage, one in Chimay/Valenciennes, the other in west Brittany, and both are among the least introgressed and the most divergent populations. Consequently, conservatories would preserve populations that are

only poorly introgressed (this was the first motive of the bee-keepers), as well as a maximum of the genetic diversity and hence of the evolutionary history of the race.

REFERENCES

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

Badino G., Celebrano G., Manino A., Population structure and Mdh-1 locus variation in *Apis mellifera ligustica*, J. Hered. 74 (1983) 443–446.

Bowcock A.M., Ruiz-Linares A., Tomfohrde J., Minch E., Kidd J.R. et al. High resolution of human evolutionary trees with polymorphic microsatellites, Nature 368 (1994) 455–457.

Cavalli-Sforza L., Edwards A.W.F., Phylogenetic analysis: models and estimation procedures, Am. J. Hum. Genet. 19 (1967) 233–257.

Chakraborty R., Jin L., A unified approach to study hypervariable polymorphism: statistical considerations of determining relatedness and population distances, in: Pena S.D.J., Chakraborty R., Eplen J.T., Jeffreys A.J. (Eds.), DNA Fingerprinting: State of the Science, Birkhauser Verlag, Basel, 1993, pp. 153–175.

Cornuet J.-M., The MDH system in honey bees from Guadeloupe, J. Hered. 70 (1979) 223–224.

Cornuet J.-M., Luikart G., Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data, Genetics 144 (1997) 2001–2014.

El Mousadik A., Petit R.J., High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.), Skeels] endemic to Morocco, Theor. Appl. Genet. 92 (1996) 832–839.

Estoup A., Solignac M., Harry M., Cornuet J.-M., Characterization of (GT)n and (CT)n microsatellites in two insect species, *Apis mellifera* and *Bombus terrestris*, Nucleic Acids Res. 21 (1993) 1427–1431.

Estoup A., Solignac M., Cornuet J.-M., Precise assessment of the number of patrilines and of genetic relatedness in honey bee colonies, Proc. Roy. Soc. Lond. [Biol.] 258 (1994) 1–7.

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

Estoup A., Taillez C., Cornuet J.-M., Solignac M., Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae), Mol. Biol. Evol. 12 (1995b) 1074–1084.

Franck P., Garnery L., Solignac M., Cornuet J.-M., The origin of west European subspecies of honey bees (*Apis mellifera*): new insights from microsatellite and mitochondrial data, 1998 Evolution (in press).

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

Garnery L., Solignac M., Celebrano G., Cornuet J.-M., A simple test using PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L., Experientia 49 (1993) 1016–1021.

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

Goldstein D.B., Ruiz Linarez A., Cavalli-Sforza L.L., Feldmann M.W., An evaluation of genetic distances for use with microsatellites, Genetics 139 (1995) 463–471.

Hall H.G., Parental analysis of introgressive hybrids between Africanized and European honey bees using nuclear RFLPs, Genetics 125 (1990) 611–621.

Hedges S.B., The number of replications needed for accurate estimation of the bootstrap P-value in phylogenetic studies, Mol. Biol. Evol. 9 (1992) 366–369.

Hunt G.J., Page R.E., Linkage map of honey bee *Apis mellifera*, based on RAPD markers, Genetics 139 (1995) 1371–1382.

Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Villablanca F.X., Wilson A.C., Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers, Proc. Natl. Acad. Sci. USA 86 (1989) 6196–6200.

McMichael M., Hall H.G., DNA RFLPs at a highly polymorphic locus distinguishing European and African subspecies of the honey bee *Apis mellifera* L. and suggest geographical origins of New World honey bees, Mol. Ecol. 5 (1996) 403–416.

Nei M., Estimation of average heterozygosity and genetic distance from a small number of individuals, Genetics 47 (1978) 253–259.

Nunamaker R.A., Subspecies determination in the honey bee (*Apis mellifera* L.) based on isoelectric focusing of malate dehydrogenase, Ph.D. thesis, University of Wyoming, Laramie, WO, 1980.

Raymond M., Rousset F., GENEPOP (Version 1.2): a population genetics software for exact tests and ecumenicism, J. Hered. 86 (1995) 248–249.

Ruttner F., Biogeography and Taxonomy of Honey Bees, Springer-Verlag, Berlin, 1988.

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

Saitou N., Nei M., The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425.

Sheppard W.S., Berlocher S.H., Enzyme polymorphism in *Apis mellifera* from Norway, J. Apicult. Res. 23 (1984) 64–69.

Shriver M.D., Jin L., Chakraborty R., Boerwinkle E., VNTR allele frequency distribution under the stepwise mutation model, Genetics 134 (1993) 983.

Smith D.R., African bees in the Americas: insights from biogeography and genetics, Trends Ecol. Evol. 6 (1991) 17–21.

Smith D.R., Palopoli M.F., Taylor B.R., Garnery L., Cornuet J.-M., Solignac M., Brown W.M., Geographical overlap of two mitochondrial genomes in Spanish honey bees (*Apis mellifera iberica*), J. Hered. 82 (1991) 96–100.

Sylvester H.A., Allozyme variation in the honey bee (*Apis mellifera*), Ph.D. thesis, University of California, CA, 1976.

Takezaki N., Nei M., Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA, Genetics 144 (1996) 389–399.

Weir B.S., Cockerham C.C., Estimating F-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370.