Original article

Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA

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 – Variability of mitochondrial deoxyribonucleic acid (mtDNA) has been studied in 973 colonies from 23 populations of the west European honey bees (lineage M) using restriction profiles of a polymerase chain reaction (PCR) amplified DNA fragment of the COI-COII intergenic region. Although populations are almost always introgressed by two other mtDNA lineages (A and C), results confirmed that the original haplotypes in western Europe are those of mtDNA lineage M. Iberian populations (*Apis mellifera iberica*) are characterised by a extended cline between haplotypes A and M, the former being almost fixed in south Spain and Portugal, and the latter almost pure in northeastern populations. This introgression is most likely attributable to humans and is probably ancient. French populations (*A. m. mellifera*) exhibit various levels of introgression by the C mtDNA lineage. Introgression is rather low in regions with a dominance of amateur bee-keeping while it reaches very high values in regions where professional bee-keepers regularly import foreign queens (mainly *A. m. ligustica* and *A. m. caucasica*). When discarding introgressed haplotypes, French populations group in two clusters, one for the northeastern part of France, and the other one for all other populations, including Swedish and northeastern Spanish populations. © Inra/Elsevier, Paris

honey bee / mitochondrial DNA / PCR / introgression

* Correspondence and reprints

Résumé – Diversité génétique de l'abeille ouest européenne (*Apis mellifera mellifera* et *A. m. iberica*). I. ADN mitochondrial. La variabilité de l'ADN mitochondrial a été étudiée chez l'abeille ouest européenne (lignée évolutive M) à partir de 973 colonies représentant 23 populations, grâce à l'analyse des profils de restriction des produits de PCR de la région intergénique COI-COII. Cette étude confirme que l'haplotype M est caractéristique de cette lignée, même si certaines populations sont lourdement introgressées par des haplotypes des deux autres lignées évolutives (A et C). Les populations de la péninsule ibérique (*Apis mellifera iberica*) sont caractérisées par un cline de fréquence des haplotypes A et M, le premier étant virtuellement fixé en Espagne du sud et au Portugal, le second à peu près pur dans le nord de la péninsule. Cette introgression est attribuée à des activités humaines passées. Les populations françaises (*A. m. mellifera*) montrent divers degrés d'introgression par des haplotypes C. Cette introgression reste très modérée dans les régions d'apiculture traditionnelle mais atteint de très forts pourcentages dans les zones où les apiculteurs professionnels importent massivement et régulièrement des reines étrangères appartenant principalement aux races *A. m. ligustica* et *A. m. caucasica.* Un arbre phylogénétique des populations, fondé sur les seuls haplotypes M, fait apparaître un groupe correspondant aux populations du nord de la France et un autre regroupant les autres populations, englobant celles de Suède et du nord-est de l'Espagne. © Inra/Elsevier, Paris

abeille / ADN mitochondrial / PCR / introgression

1. INTRODUCTION

When describing the genetic variability of the honey bee, an unusually large number of hierarchical levels can be examined: subfamilies, colonies, apiaries, populations, ecotypes, subspecies and evolutionary branches (Estoup et al., 1995a). Central to this hierarchy is the population level which can be defined as a set of colonies whose fertile members (queens and drones) mate in the same drone congregation area(s). At the local geographic scale, the genetic structure is determined mainly by biological factors, such as social life, polyandry, hive density and aggregation, and at a larger scale by geographical and historical factors.

In their original distribution area (Europe, Africa and Near and Middle East), honey bee populations are relatively homogeneous over wide territories corresponding to the subspecies or the geographical race. The honey bee species has thus been split into 24 races according to morphometric criteria. These 24 subspecies have been grouped into larger sets corresponding to evolutionary branches. Three such branches have been distinguished: an African (A) branch including all African subspecies, a north Mediterranean (C) branch in the central and east Mediterranean and in central Europe, and a west European (M) branch, from Spain to Sweden and Poland. These three branches were first identified through morphometry (Ruttner et al., 1978; Ruttner, 1988) and were approximately confirmed through deoxyribonucleic acid (DNA) studies (Cornuet and Garnery, 1991; Smith, 1991a and b; Garnery et al., 1992; Estoup et al., 1995; Arias and Sheppard, 1996).

The French indigenous subspecies, *Apis mellifera mellifera*, belongs to branch M. It extends widely beyond the country limits since it was naturally found everywhere in Europe north of the Alpine Arc (Ruttner, 1988). In the Iberian Peninsula, a related subspecies, *A. m. iberica* (Goetze, 1964), has been recognised although it has long been considered as a local form of

A. m. mellifera. The only other subspecies that is naturally present on a French border is the Italian *A. m. ligustica*, a member of branch C, with which *A. m. mellifera* hybridises over the Alps and beyond (Badino et al., 1983; Sheppard and Berlocher, 1985). However, another subspecies, *A. m. carnica* (also belonging to branch C) has been massively imported in Germany and has almost replaced the former *A. m. mellifera* populations (Kauhausen-Keller and Keller, 1994; Maul and Hähnle, 1994). The consequence is that there is now a possibility of hybridisation of *A. m. mellifera* with *A. m. carnica* in French areas close to the German border.

Although bee-keepers have long been expert in managing hives to get the best productions out of them, they generally do not control the entire biological cycle of honey bees, especially the mating of queens and drones. Hence, the current genetic biodiversity in the west part of the Old World, the original distribution area of the species, is still structured in a way that expresses its evolutionary heritage, i.e. it is determined by past demography, adaptation to local conditions, duration of isolation and natural migrations. However, one has to take into account several technical improvements more or less recently introduced in hive management and which may have interfered with the natural evolution of populations. First, queen breeding, when operated on a large scale, artificially reduces the effective size of populations, and can result in a loss of genetic variability. Second, importation of foreign queens can modify the genetic pool of local bees through hybridisation. This 'genetic pollution' is mainly due to professional bee-keepers who import foreign subspecies for their own qualities and/or in the hope of producing superior hybrids with the local subspecies (Fresnaye et al., 1974; Cornuet and Fresnaye, 1979). In France, *A. m. ligustica* and *A. m. caucasica* are the most imported subspecies, mainly because their hybrids with *A. m. mellifera* have demonstrated superior qualities for honey yield (Fresnaye et al., 1974; Fresnaye and Lavie, 1976). Other synthetic strains, such as the English 'buckfast' (Adam Br, 1966) or the American 'midnite' and 'starline' (Witherell, 1976) have also been imported, but to a lesser extent. A third factor is the practice of moving hives several times through the year to increase and diversify the honey production. According to when and where the mating season occurs, this practice may have effects similar to those of importing queens. In addition to change in the genetic pool of populations, this can artificially increase the genetic diversity in two ways: by introducing new alleles and by increasing the effective population size.

Morphometry has long been the only way to describe the genetic diversity of the honey bee. It has also been a very powerful way, since it has not only enabled the distinction of subspecies and branches, but also of ecotypes in the sense of Louveaux (1969), i.e. sets of populations from the same geographic area and characterised by a specific biological cycle of colonies. Allozymes, which have been very effective in many other species, have been of little help in honey bees. Only recently developed molecular markers can help validate and extend the knowledge of the biodiversity of the species. These molecular markers include mitochondrial DNA (mtDNA) and microsatellites. Our long-term objective is to use these markers for a general survey of all subspecies in a manner similar to that achieved with morphometry (Ruttner et al., 1978).

MtDNA has already shown its power to reveal the large-scale phylogeography of *Apis mellifera*. Through restriction of the entire molecule and sequencing

of small fragments, this marker has confirmed the existence of three main evolutionary branches, which coincide rather well with morphometry conclusions (Smith, 1991a, b; Garnery et al., 1992). The development of a simple polymerase chain reaction (PCR)-based test (Garnery et al., 1993) extends the interest of mtDNA to large surveys of colonies and hence to the analysis of genetic structure on a much finer scale. This article presents the results of the application of this test to 23 west European populations of honey bees and the accompanying paper gives the results obtained on the same populations through the analysis of microsatellite loci.

2. MATERIALS AND METHODS

2.1. Sampling

A total of 1 067 individual restriction profiles of mtDNA were analysed for honey bees collected in 25 populations (*table I* and *figure 1*). Half of the results bear on new populations and sampling has been enlarged for some of the populations already analysed. Most belong to the subspecies *Apis mellifera mellifera* and *A. m. iberica* of the west European lineage (M), with the population of Al Hoceima (Morocco, *A. m. major*) and Chalkidiki (Greece, *A. m. macedonica*) being used for comparison.

Table I. Number of specified haplotypes (see *figure 2*) in each studied population of honey bees.

		Haplotypes A												Haplotypes C	
		1	2	3	4	8	9	10	11	12	16	20	21	1	2
Umeå	41													3	
Chimay	48													6	
Valenciennes	76													5	
Vallouise	31						2							1	
Saintes	29													2	
Brest	49													1	
Ouessant	50														
Angers	49													42	1
Sabres	50						1							1	
Bayonne	41													3	
Perpignan	25													2	
Fleckenstein	30													21	
Annecy	50													1	
Versailles	51													23	4
Orsay	51													23	6
Montfavet	40													10	
Avignon	29													4	
Huesca	37						1								
San Sebastian	45													2	
Segovia	26		6	2		3	2								
Toledo	16	1	7			1	1		1						
Sevilla	59	2	39	2	1	1	2		1	3					
Porto	50	1	2	2		1	2		11	1	27	1	1	1	
Chalkidiki	30													30	
Al Hoceima	54	2				18	38	6							

For each population, a single worker has been sampled per colony, two or three colonies per apiary and as many apiaries as necessary, in a radius of 10 km to obtain, when possible, at least 30 individuals.

2.2. DNA extraction

Each bee was preserved separately in 1.5 mL of absolute ethanol. Before DNA extraction, bees were rinsed in 2 mL of insect Ringer (128 mM NaCl, 1.5 mM CaCl$_2$, 5 mM KCl, pH 7.4) and vacuum dried overnight. DNA extraction was performed according to Kocher et al. (1989) modified as in Garnery et al. (1993). Briefly, DNA was extracted by grinding the head and thorax in 0.1 M Tris HCl (pH 8.0), 10 mM EDTA, 100 mM NaCl, 0.1 % SDS, 50 mM DTT and 0.25 mg/mL proteinase K. Extracts were incubated for 2 h at 50 °C, then centrifuged for 5 min at 5 000 g. Deproteination was performed with one phenol followed by one phenol-chloroform extraction, and DNA was ethanol precipitated overnight at −20 °C, pelleted (30 min at 18 000 g), rinsed twice in 70 % ethanol, vacuum dried, resuspended in 200 µL of sterile water and diluted one tenth for use.

Table I. *Continued.*

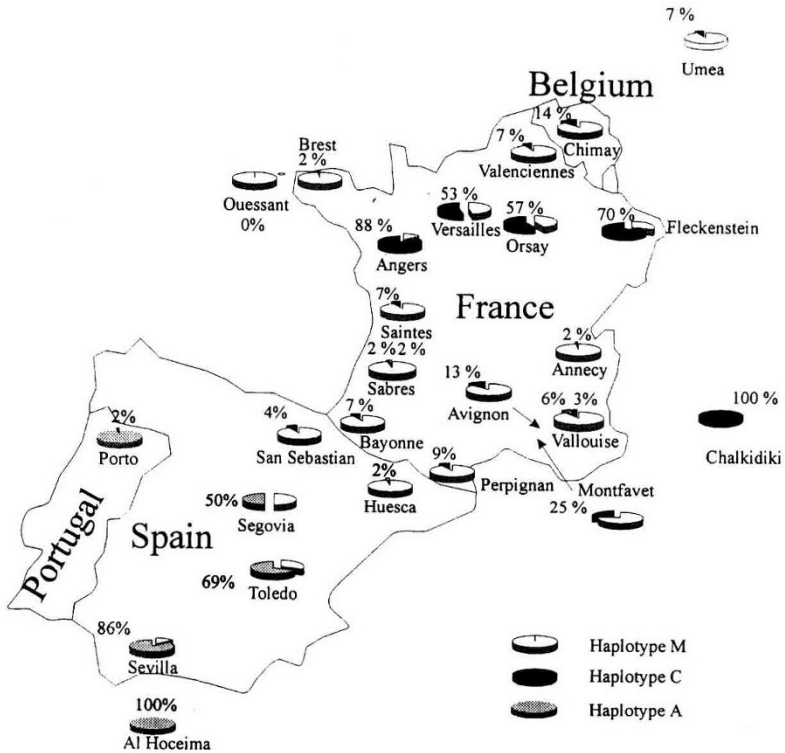| | Haplotypes M | | | | | | | | | | | | | | | | | | | | |
	1	3	4	4'	4''	5	6	7	7'	8	8'	9	10	11	12	13	14	15	16	17	17'
Umeå			20	11		7															
Chimay	1		12	3																25	1
Valenciennes	1		13	4	1		1				1			1						37	12
Vallouise			28																		
Saintes			20				3						1	1	1	1					
Brest	2		36	2			7												1		
Ouessant			42				8														
Angers			5	1																	
Sabres			39	8			1														
Bayonne			27	10						1											
Perpignan			12				1			7	1						1	1			
Fleckenstein			3																	5	1
Annecy			44	3						2											
Versailles			22				2														
Orsay			15	4			2									1					
Montfavet			26	2			2														
Avignon			18	3			4														
Huesca			22	4				3		6	1										
San Sebastian			21	13				5		3									1		
Segovia			5	1				3	2	1	1										
Toledo			2					2	1												
Sevilla		1	5					1	1												
Porto																					
Chalkidiki																					
Al Hoceima																					

Figure 1. Sampling sites of 25 populations of *Apis mellifera*. Pie chart of the frequencies of lineage M, C and A haplotypes (see text). The geographic locations of Umeå (Sweden), Chalkidiki (Greece) and Al Hoceima (Morocco) are out of the range of the map.

2.3. DNA amplification and restriction

Several primers have been designed to amplify the COI–COII intergenic region. We have used the pair E_2 (5′–GGCAGAATAAGTGCATTG–3′) and H_2 (5′–CAATATCATTGATGACC–3′). The 25 μL PCR reaction mixture contained 2.5 μL of *Taq* $10 \times$ buffer, 1.5 mM $MgCl_2$, 25 pM of primers E_2 and H_2, 25 nM of each DNTP, 0.6 unit of *Taq* polymerase and 1 μL of DNA extract. Reactions were submitted to 30 cycles of 30 s at 92 °C, 90 s at 48 °C and 2 min at 63 °C. The size of amplified DNA was determined with 5 μL of the PCR product load on a 1 % agarose gel. *Dra* I digestion was performed in 25 μL (remaining 20 μL of PCR solution and 5 μL of sterile water) at 37 °C overnight with five units of enzyme (Boehringer, Mannheim) without added buffer. Restriction fragments were separated in 5 % and 10 % polyacrylamide gels and ultraviolet visualised after soaking in ethidium bromide.

2.4. Statistical analyses

Haplotype diversities and their standard errors were calculated according to Nei and Tajima (1981).

Phylogenies derived from mtDNA data are generally based on nucleotide divergence between haplotypes. This is not relevant for the present data because the number of sites is far too small to estimate accurately the nucleotide distances, and length variations are not suitable for that purpose. Consequently, we have used the modification of the shared allele distance (Das) for haploid genomes (distance DA of Garnery et al., 1995). The distance between two haplotypes only takes two values, zero if the two haplotypes are identical and one if they are different, whatever the nature of the difference (different branches, number of Q repeats, *Dra* I sites). The distance between two populations is taken as the average of the distances between haplotypes taken pairwise, one in one population, the other in the second one. Population trees are reconstructed according to the neighbour-joining algorithm of Saitou and Nei (1987). Robustness of nodes is evaluated by a bootstrap procedure in which individuals (colonies) within populations are resampled.

3. RESULTS

3.1. Distribution of mtDNA haplotypes

The intergenic region analysed in this work is located between the COI-tRNA^{Leu} genes and the COII gene. Its structure has been described in great detail by Cornuet et al. (1991). The first unit of the COI-COII intergenic region, absent in some haplotypes, is the P sequence, about 50 bp long, which probably originated through an ancient duplication of the tRNA^{Leu} gene. This unit is followed by one to four Q sequences, each about 200 bp long, themselves composed of three subregions derived from the duplication of the block which encompasses the 3′ end of the COI gene, the tRNA^{Leu} and the P sequence. Both P and Q sequences are very A + T-rich. In haplotypes where the Q sequence is duplicated, repeats are organised in direct tandem and all copies are virtually identical. The shortest haplotypes found in the C lineage lack the P sequence and possess a single Q sequence.

The variability of this region results from the superimposition of length variation (presence/absence of the P sequence, number of reiterated Q sequences, possible small deletions in both) and nucleotide substitutions. Structure and part of the nucleotide changes are accessible through PCR amplification of the entire region and analysis of the restriction profiles obtained with the enzyme *Dra* I (TTTAAA). The restriction map was reconstructed with the help of various published sequences (Crozier et al., 1989; Cornuet et al., 1991) in order to locate new and lost *Dra* I sites (Garnery et al., 1992).

This procedure allows identification of the evolutionary branch to which the haplotype belongs (i.e. A, M or C branch). The geographical distribution of haplotypes and their relative frequencies in the sampled populations are given in *figure 1*. In Western Europe, the vast majority of haplotypes belong to the M lineage but a high proportion of A haplotypes are noted in southern Spain

and Portugal and all French populations, but Ouessant displays C haplotypes in varying amounts.

A noticeable length and site variability can also be detected within each lineage with this procedure, except in the C branch, which is almost invariant. *Figure 2* provides the physical map of all haplotypes encountered in this work. The number of individuals and the frequency of each haplotype for every studied population are given in *table I* and *figure 3*, respectively.
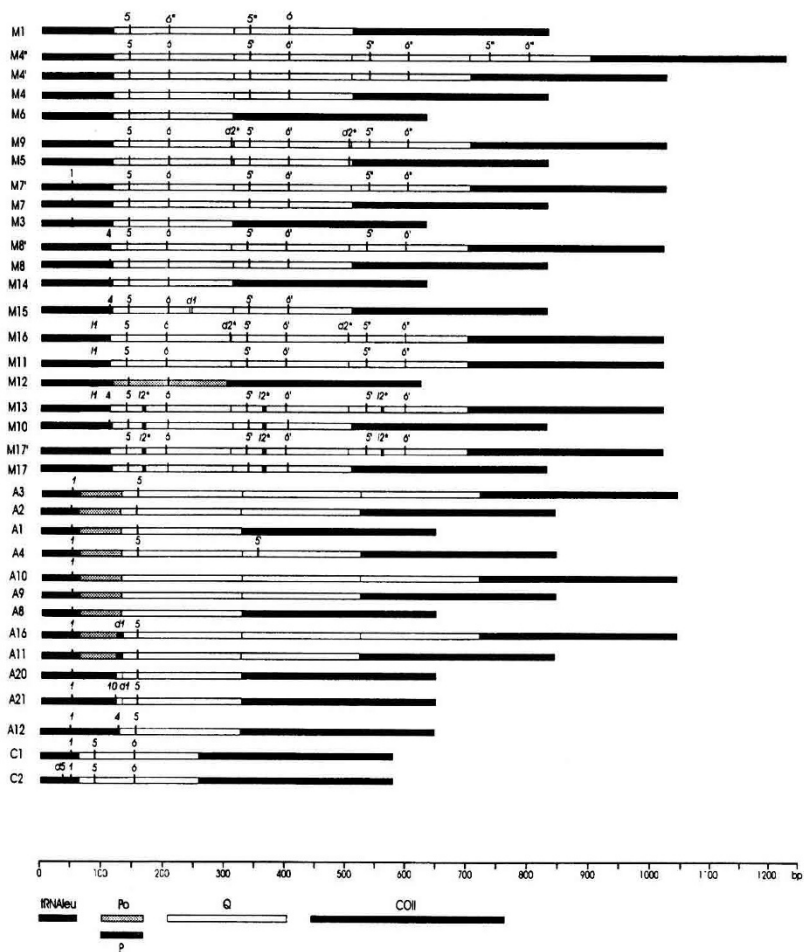

Figure 2. Physical map of the COI-COII region of M, A and C haplotypes deduced from *Dra* I restriction profiles and reference sequences. Short deletions are denoted by d and short insertions by i. The exact place of some *Dra* I sites (noted*) in repeated sequences Q is ambiguous. Each variant profile is identified by a letter (corresponding to the branch) and a number. Haplotypes which apparently exhibit the same restriction profile but in fact possess different numbers of Q repeat units (and hence possible multiple bands for serial homologous regions) are distinguished by a ′ and a ″ after the number (M4, M4′, M4″ for example).
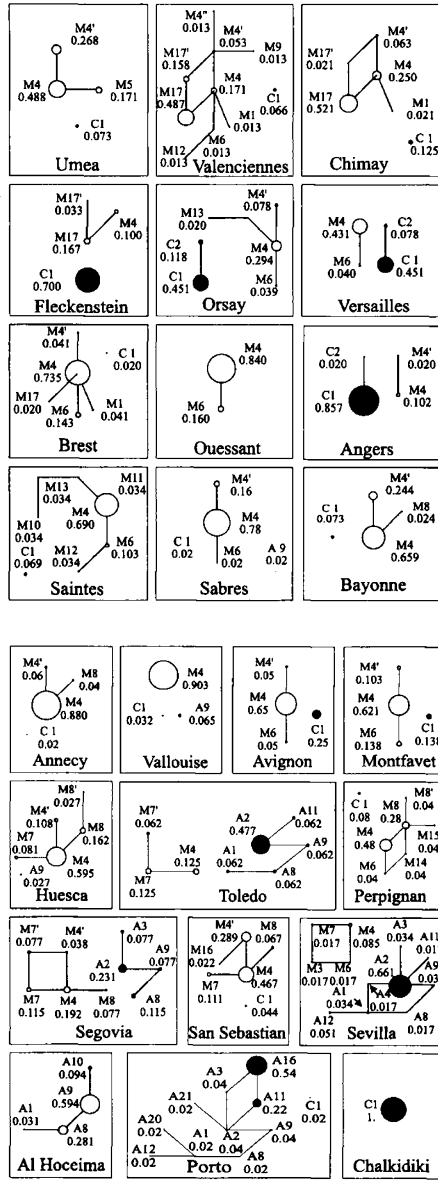
Figure 3. Distribution of haplotypes in the populations of *Apis mellifera*. Frequencies of haplotypes are proportional to the circle radius. Haplotypes A, M and C from the same population are represented in separate networks. Links between two adjacent nodes correspond to a single difference in the restriction maps; vertical links denote a change in the number of repeats of the Q sequence (the upper one being the longer) and horizontal/oblique links a *Dra* I site change or a short insertion/deletion. White, black and grey for lineages M, C and A respectively.

Table II. Mitochondrial DNA haplotype diversity (D) and standard errors (SD) in the population of honey bees of the west European lineage. For Spanish populations, D and SD have been calculated separately for the A and M haplotypes and the total.

Gene diversity	Total		M only		A only	
	D	SD	D	SD	D	SD
Umeå	0.66	0.049	0.62	0.049	–	–
Chimay	0.66	0.054	0.57	0.063	–	–
Valenciennes	0.71	0.044	0.67	0.044	–	–
Vallouise	0.19	0.090	0.00	0.000	0.00	0.000
Saintes	0.52	0.108	0.45	0.114	–	–
Brest	0.45	0.082	0.42	0.082	–	–
Ouessant	0.27	0.071	0.27	0.071	–	–
Angers	0.26	0.078	0.33	0.215	–	–
Sabres	0.37	0.078	0.32	0.075	–	–
Bayonne	0.51	0.072	0.44	0.071	–	–
Perpignan	0.71	0.073	0.66	0.079	–	–
Fleckenstein	0.49	0.097	0.64	0.126	–	–
Annecy	0.22	0.077	0.19	0.073	–	–
Versailles	0.62	0.036	0.16	0.093	–	–
Orsay	0.70	0.043	0.52	0.111	–	–
Avignon	0.52	0.072	0.25	0.099	–	–
Montfavet	0.59	0.091	0.46	0.105	–	–
Huesca	0.62	0.079	0.60	0.081	–	–
San Sebastian	0.70	0.049	0.67	0.050	–	–
Segovia	0.90	0.033	0.81	0.077	0.74	0.091
Toledo	0.81	0.093	0.80	0.164	0.62	0.164
Sevilla	0.56	0.076	0.64	0.184	0.41	0.087
Porto	0.67	0.063	–	–	0.65	0.064
Al Hoceima	0.57	0.049	–	–	0.57	0.049
Chalkidki	0.00	0.000	–	–	–	–

Total haplotype diversities have been calculated for every population, and, for populations containing an admixture of A and M haplotypes, separately for haplotypes of each lineage (*table II*). It should be noted that a higher variability is observed in the Iberian populations for total diversity but also for that of M and A haplotypes considered separately. For French populations, the values are rather erratic. Note that the lowest variability ($H = 0.27$) is observed in Ouessant where the number of colonies is artificially limited to 50.

Within each population the different haplotypes can be linked in a network where, irrespective of the nature of the difference (i.e. length or site variation), two adjacent types are assumed to differ by a single mutational step. However, it would be meaningless to plot all the haplotypes on the same diagrams because previous analyses – restriction maps of the total molecule (Smith, 1991a and b; Garnery et al., 1992) as well as partial sequences of various genes (Cornuet and Garnery, 1991; Arias and Sheppard, 1996; Koulianos and Crozier 1996) – have shown that A, M and C mtDNAs are highly divergent. Their occurrence in the same population is necessarily the consequence of a secondary admixture. Consequently, they have been considered separately in the networks given in *figure 3*.

The M4 haplotype has generally an internal and often a central position in the M networks of *figure 3*. In addition, it is present in all populations carrying an M haplotype and its frequency is always among the highest except in populations from Valenciennes and Chimay. This haplotype could be ancestral for this lineage and all other haplotypes would derive from M4, with usually one or two mutational steps, except in the Saintes population. The spectrum of other haplotypes varies among populations, some being private, and a few being circumscribed to a limited region such as M17 and M17′ in eastern and northern populations from France or M7 and M7′ in Spain.

3.2. Genetic relationships among populations

A first tree has been calculated, using the neighbour-joining algorithm and the Das matrix (*figure 4a*). Populations from southern Spain and Portugal are close to the Moroccan one (pure A branch) because they share some A haplotypes. The Greek population of Chalkidiki (pure C branch) is clustered with Angers and then Fleckenstein, Orsay and Versailles which, in this order, show a decreasing frequency of the C haplotype. The Valenciennes and Chimay populations, which are geographically close, carry a particular haplotype (M17) at a high frequency and are also clustered together. Clustering of the other populations is far less clear and the very low bootstrap values dispense with considering their topology.

We reanalysed the branching pattern of the populations after exclusion of C haplotypes and separation of M and A for the Iberian Peninsula (*figure 4b*). Although some samples are now rather small, a very different picture emerges from the tree. This tree is supposed to reflect relationships between populations prior to their introgression by alien haplotypes. The population of Fleckenstein, previously branched with the Greek sample (Chalkidiki), is now related to its nearest (Chimay and Valenciennes). These three eastern populations are isolated in a separate cluster, a reflection of their high frequency for the M17 haplotype. Similarly, Iberian populations, which were separated into two groups (one possessing A haplotypes, the other deprived of this haplotype or possessing it at a very low frequency), are clustered together and with populations from southern France. French populations strongly introgressed by the C haplotypes now return to a more homogeneous group that is not resolved.

4. DISCUSSION

The study of mitochondrial DNA is of a particular interest in honey bees since it is the ideal marker of the colony – all individuals in the colony, queen, workers and drones, sharing the same haplotype (excluding mutations). Therefore, by studying 1 067 individuals, we obtain information on 1 067 colonies. This property, combined with the haploidy of the genome and the high variability of some of its regions, confers a high power of resolution. Thus, it enables precise detection of foreign haplotypes in populations.

Our results concern a large part of the territory of the honey bees of Western Europe (*Apis mellifera mellifera* and *A. m. iberica*) belonging to the M branch. Two types of populations can be distinguished (*figure 1*): 1) populations from Spain and Portugal which show a decreasing ratio-cline of African haplotypes
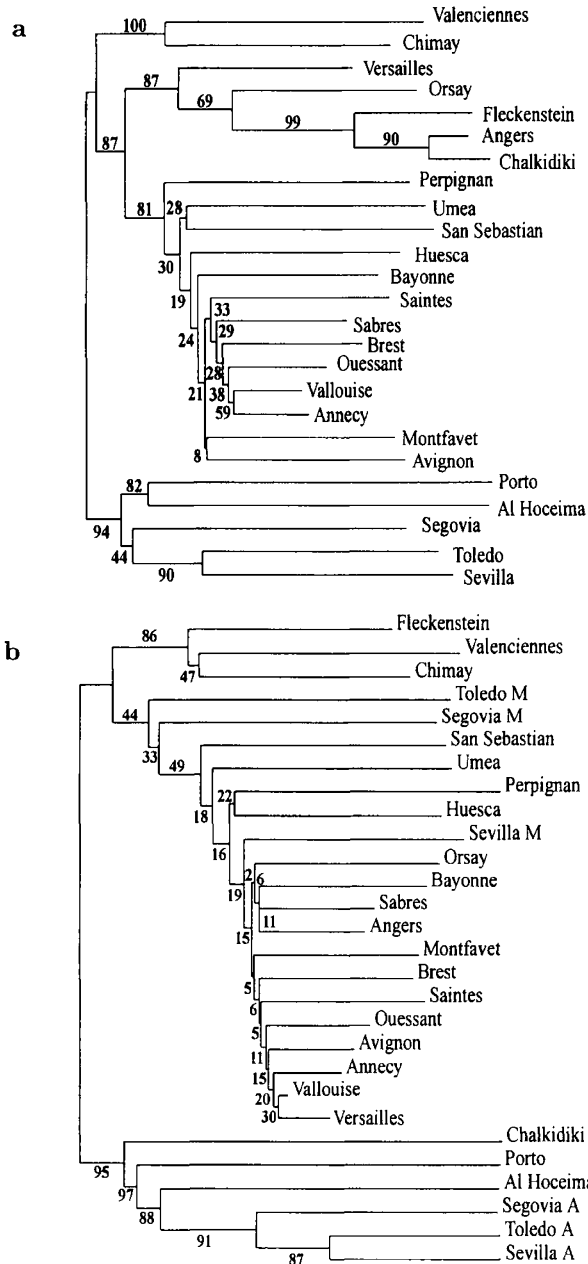
Figure 4. Neighbour-joining of the mitochondrial DNA haplotypes for the 25 populations studied using the Das distance (see text). **a**, Rough data: the frequencies of haplotypes are those observed in the samples; **b**, interpretative data: A and M haplotypes have been separated in Spanish populations and C haplotypes (considered as introgressing) have been removed in the other populations of the branch M.

progressively replaced by M haplotypes when going from southwest to northeast of the peninsula; 2) populations from San Sebastian to Belgium through France and probably extending as far as Sweden, where M haplotypes coexist with the C mitochondrial genome in all populations but Ouessant.

These observations suggest that the M haplotype, found in most populations, is characteristic of the west European evolutionary branch of the honey bee species. Previous studies of the mitochondrial genome (Smith 1991a and b; Garnery et al., 1992) have already shown that, on a large geographic scale, the M haplotypes are the most abundant in this branch whereas only A haplotypes have been found in Africa (reference population: Al Hoceima; *figure 1*) and C haplotypes in the central and eastern part of the Mediterranean Sea and in Central Europe (reference population: Chalkidiki; *figure 1*). Although the M haplotypes are the most abundant in west European populations, it is important to consider whether the presence of other haplotypes is the persistence of an ancient, and hence natural, polymorphism or if it is the result of some kind of management by bee-keepers. The two types of populations distinguished here (M + A and M + C) have to be discussed separately.

As long as the mitochondrial genome is considered alone, populations of the Iberian Peninsula could evoke a wide and natural cline resulting from the contact between African and European blocks (Smith et al., 1991a and b). In fact, this cline involving two very divergent genomes (about 2.5 % of nucleotide differences; Garnery et al., 1992) cannot be the persistence over time of an ancient polymorphism (such a divergence would imply a coalescence one million years ago or so; Garnery et al., 1992) for populations with effective sizes estimated around 1 000 (the coalescence time which is equal to $2N_f$ for mtDNA would correspond to a few thousand years). Whereas the Iberian A haplotypes are very close to African ones, they are not strictly identical. In addition, the haplotype diversity is higher in Spanish and Portuguese populations than in Africa. The network of the population of Al Hoceima (*figure 3*) shows only four haplotypes, two adjacent ones distinguished by a single difference (either one repetition of the Q unit or a *Dra* I site). It is representative of all the networks observed for seven areas in Morocco, whatever the race (*major, intermissa* or *sahariensis*). In north Morocco (race *major*), the only part of the country which can be considered as a plausible natural donor, haplotypes *A8* and *A9* are largely predominant (15/19) whereas they are very rare in the Iberian Peninsula. Similarly, the most abundant haplotypes in Spain and Portugal (*A2* and *A16*), are either very rare and observed only in the West Atlas (*A2*) or totally absent from all Morocco populations (*A16*) (Garnery et al., 1995). This strongly suggests that the Iberian A haplotypes do not originate from natural migrations through the Strait of Gibraltar but rather from multiple introductions of colonies bearing A haplotypes by humans, possibly from different geographical origins. The other conclusions that can be drawn for the mitochondrial structure of these populations need to be discussed in parallel with nuclear gene polymorphism (see Part II, accompanying article).

For the second set of populations, we can distinguish the Fleckenstein population from the others. In this locality the frequency of the C haplotype is very high (70 %). Ten years ago, after the decline of populations due to invasion by the mite *Varroa jacobsoni*, the bee-keepers massively imported bees of the *carnica* subspecies (belonging to the C branch) to replace local colonies. In

fact, they did as the German bee-keepers did 50 years ago for other reasons (replacement of *mellifera* by *carnica*), and they entirely modified the genetic profile of their populations: for instance at Oberursel, near Frankfurt, all the 96 individuals analysed belonged to the C lineage (unpublished results). The high density of the C haplotype of *carnica* at Fleckenheim may be the result of the superimposition of these recent massive importations and of genetic flow across the German border with other *carnica* which have themselves been imported.

In the other populations, which are far from any natural contact with other lineages, the presence of the C haplotype varies greatly in frequency in France (note the almost total absence of the C haplotype in Spain), namely from 0 % for Ouessant to 88 % for Angers. Here again, this polymorphism is not in agreement with the persistence of an ancient polymorphism (for the same reasons that in Spain the M and C haplotypes are also highly divergent) and its level is closely related to the rate of queen importation, mainly the subspecies *caucasica* and *ligustica*, both bearing the C haplotype. Two situations have to be distinguished: C haplotype frequency under or over 10 %. The first situation corresponds to populations of extensive bee-keeping by amateurs who principally use local honey bees. The second situation corresponds to regions (Angers, Versailles, Orsay, Avignon, Montfavet) where professional bee-keepers or experimental apiaries regularly import foreign queens. The preferential localisation of highly introgressed populations on the same latitude (*figure 1*) is in this respect only incidental. Some other commercial bees (buckfast, midnite, starline, etc.) are imported but their genetic profiles are not precisely defined. For the present, the precise geographical origin of the contaminant populations cannot be accurately given because the variability of the COI-COII region in this branch is virtually null. We are currently developing a promising test on the A + T-rich region (control region) to obtain more information.

Although the distinction among the cytoplasmic lineages is particularly easy, it would be hasty to deduce the impact of importation directly from the frequency of foreign haplotypes. First, drift is not likely to be responsible for these differences because this queen traffic is a very recent one in France (it could be older in Spain and Portugal). Second, the mitochondrial genome is probably subject to a direct selection (Ballard and Kreitman, 1994; Rand et al., 1994). For the honey bee, we have some evidence that, in spite of massive importation of *mellifera* during the French colonisation in Morocco, the M haplotypes have left no trace in current populations (Garnery et al., 1995; Franck et al., unpublished report). This does not necessarily mean that the M haplotype itself was counterselected but that maternal lineages derived from these imported queens were unsuccessful. Third, the Italian subspecies *ligustica*, which represents the main source of importation in various regions, is, perhaps, at least in the southeast of the Alpine Arc, itself introgressed by the M haplotype (the rate of importation will thus be underestimated, except if Italian populations have a diagnostic M haplotype). Lastly, mitochondrial introgression may be symmetrical or asymmetrical (in one taxon but not in the other) and parallel or differential (depending on the genes considered, nuclear and mitochondrial) (see Aubert and Solignac, 1990, for more details). These differences reflect the different roles played by the two sexes in the transmission of the mtDNA molecule, the behaviour of drones and queens and a possible ecological disadvantage of hybrids of different generations. Consequently, the

results obtained with mtDNA have to be compared to those of nuclear markers (see Part II, accompanying article).

To summarise, the west European lineage of honey bees, before its management by man, probably bore a pure M haplotype, except on its margins in the regions of contact with other lineages. Its variability within populations shows a close relationship between the different haplotypes which differ only by a few mutational steps, with rare variants linked to a frequent haplotype (mainly M4) by one or two steps. This suggests that they form for each population, a single evolutionary group. In addition, most populations from southern Spain to Sweden show roughly the same panel of haplotypes. This is the consequence of their common origin from the same population rather than strong homogenising migrations. They were propagated, without major changes, probably from a Spanish refuge, by colonisation at the end of the ice age. Since that time, genetic drift (natural or consecutive to apiculture) was probably moderate because populations are presently very homogeneous. However, the Hennig comb structure, joined to a decrease in the branch length from Spain to France in the tree for M haplotypes (*figure 4b*) which parallels a decrease in haplotype diversity (*table II*), suggests that a part of variability has been lost during the process of colonisation. Finally, the existence of particular haplotypes (M17 and M17′) in the east of France (and in Poland and northern Italy; work in progress) is perhaps indicative of a second and more oriental refuge during the last ice age.

REFERENCES

Adam O.B.E. Br, In Search of the Best Strains of Bees, Northern Bee Books, Walman Verlag, Zell-Weierbach, 1966.

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

Aubert J., Solignac M., Experimental evidence for mitochondrial DNA introgression between *Drosophila* species, Evolution 44 (1990) 1272–1282.

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

Ballard J.W.O., Kreitman M., Unraveling selection in mitochondrial genome of Drosophila, Genetics 138 (1994) 757–772.

Cornuet J.-M., Fresnaye J., Production de miel chez les hybrides interaciaux d'abeilles (*Apis mellifera* L.) lors de générations successives de rétrocroisement sur la race locale, Apidologie 10 (1979) 3–15.

Cornuet J.-M., Garnery L., Mitochondrial DNA variability in honeybees and its phylogeographic implications, Apidologie 22 (1991) 627–642.

Cornuet J.-M., Garnery L., Solignac M., Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA, Genetics 128 (1991) 393–403.

Crozier R.H., Crozier Y.C., Mackinlay A.G., The CO-I and CO-II region of the honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates, Mol. Biol. Evol. 6 (1989) 399–411.

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 (1995) 679–695.

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

Fresnaye J., Lavie P., Sélection et hybridation de l'abeille en France. Symp. Génétique, sélection et reproduction de l'abeille, Bull. Tech. Apic. OPIDA 3 (1976) 15–20.

Fresnaye J., Lavie P., Boesinger E., La variabilité de la production de miel chez l'abeille de race noire (*Apis mellifera* L.) et chez quelques hybrides interaciaux, Apidologie 5 (1974) 1–20.

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.

Goetze G., Die Honigbiene in natürlicher und künstlicher Zuchtauslese, Parey, Hamburg, 1964.

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.

Kauhausen-Keller D., Keller R., Morphometrical control of pure race breeding in the honeybee (*Apis mellifera* L.), Apidologie 25 (1994) 133–143.

Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Pääbo S., 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.

Koulianos S., Crozier R.H., Mitochondrial DNA sequence data provides further evidence that the honeybees of Kangaroo Island Australia are of hybrid origin, Apidologie 27 (1996) 165–174.

Louveaux J., L'importance de la notion d'«écotype» chez l'abeille, in : 22e Congr. Int. Apic., Munich, 1969, pp. 499–501.

Maul V., Hähnle A., Morphometric studies with pure bred stock of *Apis mellifera carnica* Pollmann from Hessen, Apidologie 25 (1994) 119–132.

Nei M., Tajima F., DNA polymorphism detectable by restriction endonucleases, Genetics 97 (1981) 145–163

Orantes-Bermejo F.J., Garcia-Fernadez P., Morphological variability of *Apis mellifera iberica* in different apiaries of southern Spain, J. Apicult. Res. 34 (1995) 23–30.

Rand D.M., Dorfsman M., Kann L.M., Neutral and non-neutral evolution of Drosophila mitochondrial DNA, Genetics 138 (1994) 741–756.

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., New allozyme variability in Italian honey bees, J. Hered. 76 (1985) 45–48.

Smith D.R., Mitochondrial DNA and honey bee biogeography, in: Smith D.R. (Ed.), Diversity of the Genus Apis, Westview Press, Oxford and IBH Publishing Co., 1991a, pp. 131–176.

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

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

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.

Witherell P.C., A story of success: the Starline and Midnite hybrid bee breeding programme, Am. Bee J. 116 (1976) 63-64, 82.