Molecular Ecology (1992) 1, 145–154

Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis

L. GARNERY, J.-M. CORNUET and M. SOLIGNAC*
Laboratoire de Neurobiologie Comparée des Invertébrés, INRA-CNRS URA 1190, La Guyonnerie, 91440 Bures-sur-Yvette and
**Laboratoire de Biologie et Génétique Evolutives, CNRS, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France*

Abstract

Variability of mitochondrial DNA (mtDNA) of the honey bee *Apis mellifera* L. has been investigated by restriction and sequence analyses on a sample of 68 colonies from ten different subspecies. The 19 mtDNA types detected are clustered in three major phylogenetic lineages. These clades correspond well to three groups of populations with distinct geographical distributions: branch A for African subspecies (*intermissa, monticola, scutellata, andansonii* and *capensis*), branch C for North Mediterranean subspecies (*caucasica, carnica* and *ligustica*) and branch M for the West European populations (*mellifera* subspecies). These results partially confirm previous hypotheses based on morphometrical and allozymic studies, the main difference concerning North African populations, now assigned to branch A instead of branch M. The pattern of spatial structuring suggests the Middle East as the centre of dispersion of the species, in accordance with the geographic areas of the other species of the same genus. Based on a conservative 2% divergence rate per Myr, the separation of the three branches has been dated at about 1 Myr BP.

Keywords: mitochondrial DNA, length polymorphisms, honey bee, genetic diversity, phylogeography

Received 20 March 1992; accepted 20 May 1992

Introduction

Before European settlers brought it to New World continents, the western honey bee *A. mellifera* L., was naturally found in Africa, Near and Middle East and Europe. This wide area of distribution combined with a low migratory ability (except in tropical areas) and small population size led to a high level of genetic differentiation which is reflected in the large number of recognized subspecies (Ruttner 1988). These subspecies have been defined primarily on morphometrical grounds, but they can also differ in their behaviour and ecology.

Morphometry is a powerful technique for discriminating populations (Tomassone & Fresnaye 1971; Cornuet, Fresnaye & Tassencourt 1975; Ruttner, Tassencourt & Louveaux 1978), but is not suitable for inferring phylogenetic relationships between them. Allozyme variation is more appropriate, but is limited by the low level of variability detected among isozymes of honey bee subspecies.

Mitochondrial DNA (mtDNA) has proven to be a valuable tool in phylogeographic studies at the species or subspecies level (Avise *et al.* 1987). The maternal inheritance of mtDNA makes it even more suited to studies on honey bees for two reasons. First, because all the individuals of the colony are the progeny of the queen, their mtDNA is identical; the colony is then the relevant 'individual' in analyses. This still holds when the queen dies or leaves the hive during swarming, because she is replaced by one of her daughters. Secondly, in colonization processes, which are important in species' life histories, the important role is played by swarms, i.e. colonies, and not by males.

MtDNA of *Apis mellifera* is a circular molecule, 16 500–17 000 bp long. The comparison of restriction maps and partial sequence data indicate that, except for some tRNA genes, the genetic map is similar to that of *Drosophila* (Crozier, Crozier & McKinlay 1989; Cornuet & Garnery 1991).

MtDNA variability within and between subspecies has been previously investigated in *A. mellifera* by restriction site analysis (Moritz *et al.* 1986; Smith 1988; Smith & Brown 1988, 1990; Smith *et al.* 1991) and by sequence

Correspondence: J.-M. Cornuet.

analysis (Cornuet, Garnery & Solignac 1991; Itenov & Pedersen 1991; Garnery *et al.* 1991; Koulianos & Crozier 1991). Most of these studies have been focused on the discrimination of various subspecies and subsidiarily on the africanization process of the American populations (Smith, Taylor & Brown 1989; Hall & Muralidharan 1989; Hall & Smith 1991). In this paper, we present additional data (restriction maps of new subspecies, new sequence data) and we examine their contribution to investigation of the evolutionary history of *Apis mellifera*.

Materials and methods

Sampling

Restriction enzyme variability of mtDNA was investigated in a sample of 68 colonies belonging to 10 different *Apis mellifera* subspecies: *mellifera* (France 1–19), *carnica* (Austria 1–6), *caucasica* (Caucasus 1–3), *ligustica* (Italy 1–9), *intermissa* (Morocco 1 and 2, Algeria 1), *capensis* (Cape 1), *iberica* (Spain 1–15), *monticola* (Malawi 1–9), *scutellata* (Malawi 10 and 11) and *adansonii* (Congo 1). Subspecies assignment was postulated according to the geographic origin of colonies.

Partial mtDNA sequences were obtained for seven colonies chosen from the preceeding ones and belonging to seven different subspecies: *caucasica* (Caucasus 2), *adansonii* (Congo 1), *intermissa* (Algeria 2), *monticola* (Malawi 6), *capensis* (Cape 1), *carnica* (Austria 3), *mellifera* (France 18).

One *Apis cerana* colony from Dacca (Bangladesh) has been taken as outgroup for restriction and sequence analyses, because this species is the closest relative to *Apis mellifera* (Garnery *et al.* 1991).

Restriction data

Mitochondria were purified from newly emerged workers by differential centrifugation on sucrose gradient (Solignac & Monnerot 1986) and lysed in SSC (150-mmol/l NaCl, 15-mmol/l sodium citrate, pH7) containing 1% SDS. Proteins were removed by two phenol-chloroform extractions and nucleic acids were precipitated overnight in sodium acetate/ethanol at −20°C. Aliquots of mtDNA were digested with 16 restriction endonucleases (*Acc*I, *Ava*I, *Ava*II, *Bcl*I, *Bgl*II, *Cfo*I, *Cla*I, *Eco*RI, *Eco*RV, *Hin*cII, *Hin*dIII, *Nde*I, *Pst*I, *Pvu*II, *Spe*I, *Xba*I) using the conditions recommended by the supplier (Boehringer Mannheim). The digested fragments were end-labelled with α^{32}P deoxynucleotides, separated on 0.8% agarose, 3.6 and 5% polyacrylamide gels and visualized by autoradiography. Restriction maps of mtDNA have been independently established for four colonies (Italy 1, France 1, Spain 1, Caucasus 1), using the double digestion method. MtDNA of the other colonies have been mapped by comparing their digestion profiles with those of the mapped mtDNA types. Every additional restriction site was mapped by analysis of appropriate double digestions.

Sequence data

Total DNA was extracted according to Kocher *et al.* (1989) with slight modifications (Garnery *et al.* 1991). The DNA templates were produced by the polymerase chain reaction (PCR). The conditions for PCR were as described previously (Garnery *et al.* 1991). The sequenced region encompasses the 3′ end of the CO-I gene, the $tRNA^{leu}$ gene, an intergenic region of variable length and the 5′ end of the CO-II gene. Single-stranded DNA was obtained through strand specific digestion by lambda exonuclease (Higuchi & Ochman 1989) and sequenced using the dideoxy chain-termination method (Sanger, Nicklen & Coulsen 1977) with the Pharmacia sequencing kit. Five different primers were used:

E1: 5′-GGAGTAAATCTAACTTTC-3′;
E2: 5′-GGCAGAATAAGTGCATTG-3′;
E3: 5′-ATACCACGACGTTATTCAGA-3′;
H2: 5′-CAATATCATTGATGACC-3′;
H3: 5′-ATATGAATCATGTGG-3′.

The positions of the 3′ end of the above primers on a previously published sequence (Cornuet *et al.* 1991, Fig. 2) are 258, 576, 308, 1562 and 1266, for E1, E2, E3, H2 and H3 respectively.

Phylogenetic analyses

To infer phylogenetic relationships between the mtDNA types, we applied a maximum parsimony method and a distance method. In both cases, trees were rooted by taking the *Apis cerana* colony as an outgroup. Maximum parsimony trees were obtained with PAUP software (Swofford, 1989) using the heuristic approach for both restriction and sequence data and considering characters as unordered. We used the Neighbor–Joining algorithm (Saitou & Nei 1987) for the distance method. The distances were calculated from sequence data according to Kimura (1980) and those from the restriction maps were expressed as the mean number of nucleotide substitutions per site (Nei & Li 1979). Sampling variances for these parameters have been estimated with the jack-knife procedure (Efron 1979; Nei & Miller 1990). All the computations (distance, diversity and Neighbor–Joining tree) for restriction site data have been performed with the RESTSITE v1.1 software (Miller 1991).

Results

Restriction site variability

In order to facilitate comparisons with formerly published maps (Smith & Brown 1988, 1990), mtDNA was linearized on the same site, *PvuII* 1 (Fig. 1). Of the 45 sites mapped with the 16 restriction enzymes assayed, 23 were found in all samples (constant sites). The remaining 22 variable sites (Fig. 1) distinguished 19 mtDNA types, noted I to XIX (Table 1).

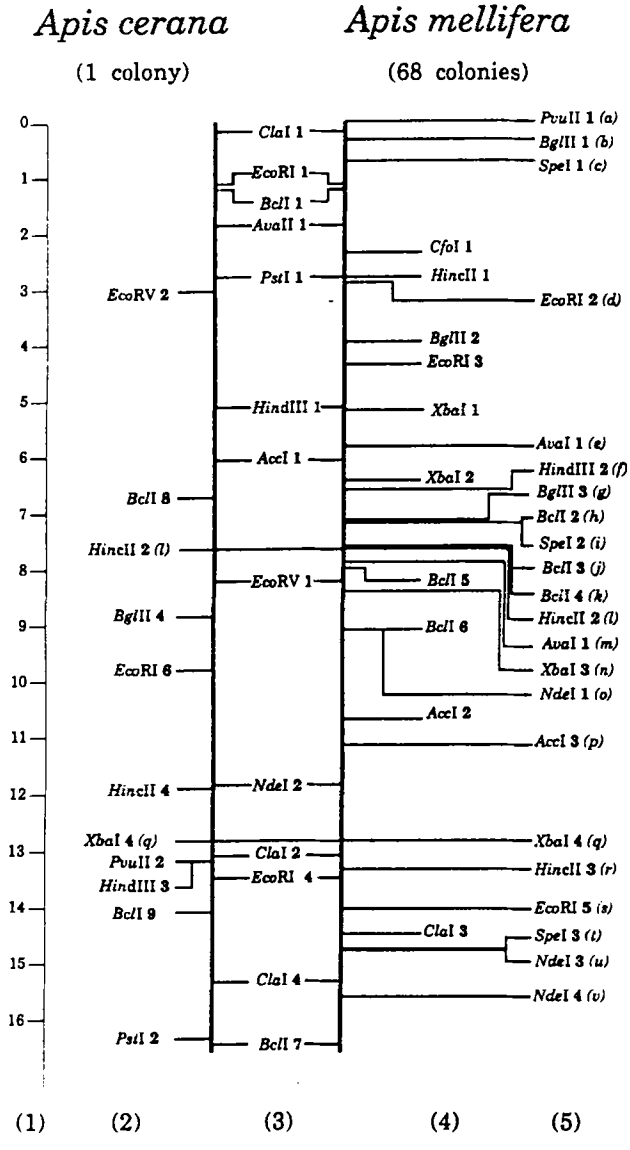

Fig. 1 Restriction map of *Apis mellifera* and *Apis cerena* mtDNAs. From left to right: (1) a scale in kbp, (2) restriction sites found in the *Apis cerana* colony and not found in all *Apis mellifera* colonies, (3) restriction sites found in all colonies of both species, (4) restriction sites constant in *Apis mellifera* and not found in *Apis cerana* and (5) restriction sites variable in *Apis mellifera*. Two sites, *Hin*cII2 (*l*) and *Xba*I4 (*q*), have been found in *Apis cerana* and in some colonies of *Apis mellifera*. Restriction sites annotated with an italic letter between parentheses are polymorphic in *Apis mellifera* (see Table 1).

Length polymorphisms

Length polymorphisms have been detected in at least two regions of the mtDNA genome. These polymorphisms are apparent in the restriction profiles of the enzymes *BclI* (fragment *BclI* 5–*BclI* 6) and *XbaI* (fragment *XbaI* 1–*XbaI* 2).

Four different sizes (1150, 1200, 1400 and 1600 bp) have been found for the fragment *BclI* 5–*BclI* 6. This variability is explained by the assemblage, in various combinations, of two short sequences, a sequence P/P_0 (P_0 and P are two variants differing by a central 15 bp insertion/deletion, see Fig. 2) and a sequence Q (194–196 bp long), located between the $tRNA^{leu}$ and the CO-II genes (Cornuet *et al.* 1991). The shortest *BclI* fragment is deprived of the sequence P and possesses a single sequence Q; it is characteristic of colonies of the subspecies *ligustica, carnica* and *caucasica*. In the *mellifera* colonies, three fragment lengths, (1200, 1400 and 1600 bp) were observed and they correspond to the intergenic structures PQ, PQQ and PQQQ respectively. In African colonies (*intermissa, monticola, capensis, scutellata* and *adansonii*), the same length variability has been observed and the intergenic structure is P_0Q, P_0QQ or P_0QQQ (Table 1).

The second region variable in length (fragment *XbaI*1–*XbaI*2) overlaps the A+T-rich region (Cornuet & Garnery 1991). Two sizes of fragments have been found with a 100-bp difference. The largest size is characteristic of *mellifera* (Table 1).

Sequence of the COI–COII region

Nucleotide sequences of mtDNA from seven colonies of *A. mellifera* and one colony of *A. cerana* are given in Fig. 2, in addition to two already published sequences, one (France 15) from Cornuet *et al.* (1991) and the other (*ligustica*) from Crozier *et al.* (1989). These sequences correspond to the 5′ end of the CO-I gene (236 bp), the $tRNA^{leu}$ gene (71 bp), one sequence Q (197 bp) and the 3′ end of the CO-II gene (239 bp), i.e. a total of 743 bp. In colonies which have two or three Q repetitions (France 15 and 18, Congo 1, Malawi 6 and Cape 1), the sequences of the different repeats were found identical. Within the species *A. mellifera* (i.e. excluding the *A. cerana* sequence), 26 sites are variable, 20 of which are phylogenetically informative. Every pair of colonies exhibits differences, even those of the same subspecies such as France 15 and France 18 which were not distinguishable by their restriction maps (Table 1).

Table 1 Site and length polymorphisms in honey-bee mtDNA. Presence of restriction site is noted by 1 and the absence by 0. Explanation for the length polymorphism of the *BclI*5–*BclI*6 fragment is given in the text

Type	Samples	Restriction sites numbered as in figure 1	Length polymorphisms		Lineage
		a b c d e f g h i j k l m n o p q r s t u v	*XbaI* 1 - *XbaI* 2	*BclI* 5 - *BclI* 6	
I	France15-18 *(mellifera)*	1 1 1 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 1 1 1 0	+100 bp	$P(Q)_2/P(Q)_3$	M
II	France1-14, Spain1-8 *(mellifera, iberica)*	1 1 1 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 1 0 1 0	+100 bp	$PQ/P(Q)_2/P(Q)_3$	M
III	France19 *(mellifera)*	0 1 1 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 1 1 1 0	+100 bp	$P(Q)_2$	M
IV	Spain9 *(iberica)*	1 1 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 1 0 1 0	+100 bp	$P(Q)_2$	M
V	Morocco1-2, Algeria1, Congo1, Malawi11, Spain10-15 *(intermissa, adansonii, scutellata, iberica)*	1 0 1 0 1 1 0 1 0 1 0 0 1 0 1 1 1 0 0 0 0 0		$P_0Q/P_0(Q)_2$	A
VI	Cape1 *(capensis)*	1 0 1 0 1 1 0 1 0 0 1 0 1 0 1 1 1 0 0 0 0 0		$P_0(Q)_3$	A
VII	Malawi10 *(scutellata)*	1 0 1 0 1 1 0 1 0 1 0 0 1 C 0 1 1 0 0 0 0 0		P_0Q	A
VIII	Malawi7-9 *(monticola)*	1 0 1 0 1 1 0 1 0 1 0 0 1 0 1 1 0 1 0 0 0 0		$P_0Q/P_0(Q)_2$	A
IX	Malawi5-6 *(monticola)*	1 0 1 0 1 1 0 1 0 1 0 0 1 0 1 1 0 0 0 0 0 0		$P_0(Q)_2$	A
X	Malawi1-4 *(monticola)*	1 0 1 0 1 1 0 1 1 1 0 0 1 0 1 1 1 1 0 0 0 0		$P_0(Q)_2/P_0(Q)_3$	A
XI	Caucasus1 *(caucasica)*	1 1 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 1 0		Q	C
XII	Caucasus2-3, Austria5-6 *(caucasica, carnica)*	1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 1 1 1 0 0 0 1 0		Q	C
XIII	Austria1 *(carnica)*	1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 1 1 0 0 0 0 1 0		Q	C
XIV	Austria2 *(carnica)*	1 1 0 1 1 0 0 1 1 1 0 0 1 1 1 1 1 1 0 0 0 1 1		Q	C
XV	Austria4 *(carnica)*	1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 1 1 1 0 0 0 1 1		Q	C
XVI	Italy1-5, Austria3 *(ligustica, carnica)*	1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 1 1 0 0 0 0 1 1		Q	C
XVII	Italy6 *(ligustica)*	1 1 1 1 0 1 0 1 1 1 0 0 0 1 1 1 0 0 0 0 1 1		Q	C
XVIII	Italy7-8 *(ligustica)*	1 1 0 1 0 1 0 1 1 1 0 0 0 1 1 1 0 0 0 0 1 1		Q	C
XIX	Italy9 *(ligustica)*	1 1 0 0 1 1 0 1 1 1 0 0 1 1 1 1 0 0 0 0 1 1		Q	C
	Apis cerana	0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0		+89 bp	

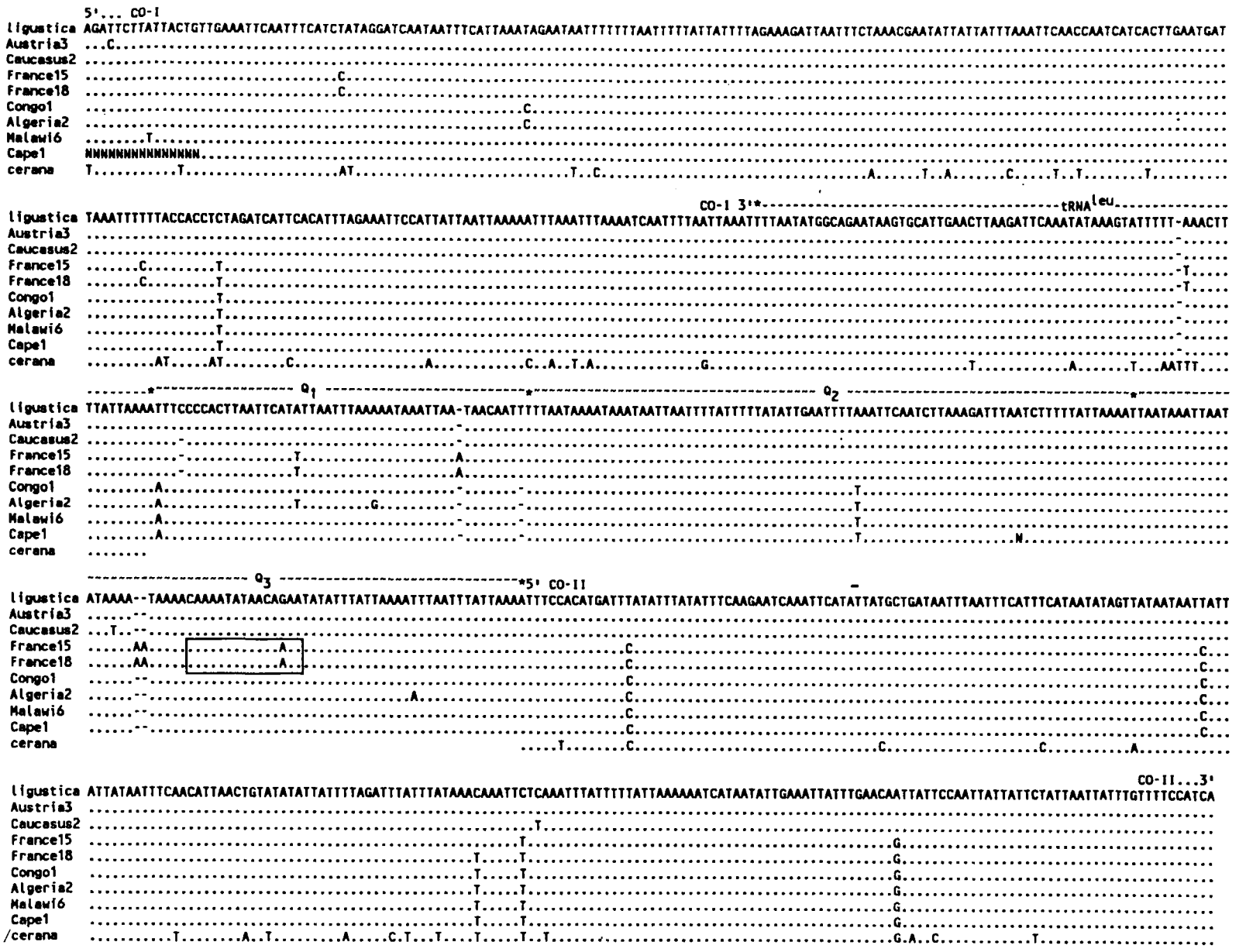

Fig. 2 Partial nucleotide sequence of mtDNA from nine *Apis mellifera* colonies and one *Apis cerana* colony in the COI-COII region of mtDNA, encompassing the 3'end of the COI gene, the $tRNA^{leu}$ gene, one copy of sequence Q and the 5'end of the COII gene. The first sequence (*ligustica*), from literature (Crozier *et al.* 1989), is taken as the reference to which other sequences are compared: a dot indicates an identical nucleotide, a dash corresponds to a deleted site and 'N' indicates an undetermined nucleotide. *Apis cerana* mtDNA contains an intergenic sequence of 89 bp, already published (Cornuet *et al.* 1991), which is homologous to both portions Q_2 and Q_3 of sequence Q. The sequence P_0, found in all African samples and located between the $tRNA^{leu}$ gene and Q_1, is exactly identical to the sequence of Q_3 of the same samples. The sequence P, found in French samples and located at the same position, is identical to the sequence Q_3 found in the same genome, apart from a 15-bp deletion identified by a box.

Phylogenetic relationships

Phylogenetic trees of the mtDNAs (Fig. 3) have been deduced from restriction (A and B) and sequence (C and D) data, using either a distance method (A and C) or a parsimony method (B and D). The topology of the four trees support the existence of three major mtDNA lineages, namely:

- a west Mediterranean lineage including all mtDNA types from French populations (henceforth abbreviated M, for *mellifera*);
- an African lineage including all mtDNA types from colonies of African origin (and consequently named the lineage A);
- a north Mediterranean–Caucasian lineage corresponding to colonies from Caucasus, Austria and Italy (lineage C).

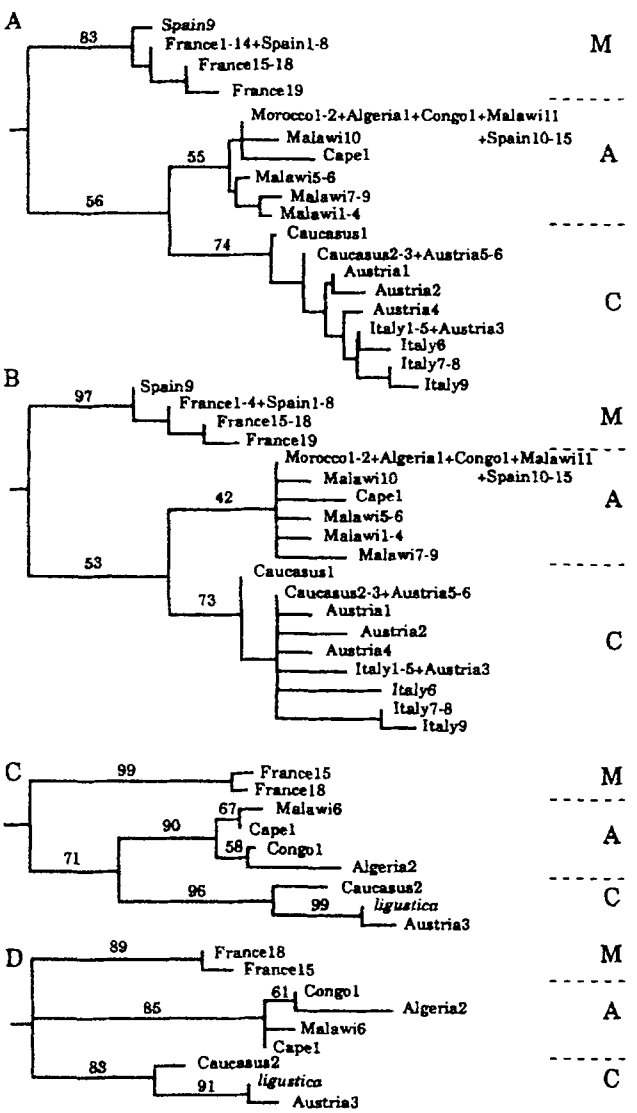

Fig. 3 Phylogenetic trees of *Apis mellifera* mtDNA based on restriction (trees A and B) and sequence (trees C and D) data. Trees have been obtained either through the Neighbor-Joining method (trees A and C) or the Parsimony method (trees B and D) taking *Apis cerana* as outgroup. Bootstrap scores (percentages over 200 replicates for trees A and B or 1000 replicates for trees C and D) are indicated for the main branches. In trees C and D, *ligustica* refers to the mtDNA sequence of Crozier *et al.* (1989).

Polytomies observed in trees B and D result from equally parsimonious trees produced by the PAUP software. In tree B, this concerns the internal topology of lineages A and C. In tree D, the relationship between the three major lineages is not resolved at the level of the deepest nodes.

MtDNA variability in A. mellifera

Results from restriction data are based on a large number of colonies, but the sequence data concern a larger sample of nucleotides. The average number of nucleotide substitutions per site inferred from restriction maps varies from zero to as much as 4.13% (between France 19 and Italy 1). For nucleotide sequences, Kimura's distances vary by up to 2.2%. Within the species *A. mellifera*, the mean (± SD) divergence between individual genomes taken pairwise (i.e. the nucleotide diversity) is 1.9 ± 0.5% for restriction data and 1.3 ± 0.7% for sequence data. Within-lineage nucleotide diversities based on restriction data are equal to 0.25%, 0.30% and 0.07% for lineages A, C and M, respectively.

Table 2 provides raw and net average nucleotide divergences between the three major lineages within *A. mellifera* (the net divergence being equal to the raw divergence minus the average of the intralineage nucleotide diversities). The divergence between *A. mellifera* and *A. cerana* is also indicated.

Discussion

The three major mitochondrial lineages in A. mellifera

Restriction and sequence data obtained in this work on *A. mellifera* mtDNA show the existence of three distinct clades. The monophyly of lineages C and M is supported by their high bootstrap scores, whatever the data and the method of phylogeny reconstruction are, whereas the monophyly of the lineage A appears less robust with this test. An additional character supporting this tripartite distribution of the colonies is the pattern of length polymorphisms. The length of the *BclI5–BclI6* fragment indicates three groups of colonies: the intergenic sequences Q, $P_0(Q)_n$ and $P(Q)_n$ are strictly associated with the lineages C, A, and M, respectively. Lineage M is also characterized by a larger *XbaI1–XbaI2* fragment. In conclusion, the present data strongly sustain the existence of three mtDNA

Table 2 Nucleotide distances in percent between the three major mtDNA lineages (A, C and M) of *Apis mellifera* and between western and eastern honey bees, *Apis mellifera* and *Apis cerana*. The number of sequences per lineage does not allow a meaningful estimate of intra-lineage diversity

Nucleotide distance	*Restriction data*		Raw sequence data
	raw	net	
Lineages A–M	2.8 ± 0.8%	2.6 ± 0.8%	1.9 ± 0.1%
Lineages A–C	2.1 ± 0.9%	1.8 ± 0.8%	1.5 ± 0.2%
Lineages C–M	3.3 ± 0.9%	3.1 ± 0.8%	1.9 ± 0.2%
A. mellifera–A. cerana	13.6 ± 3.2%	11.7 ± 3.2%*	9.5 ± 0.4%

*The diversity of *A. cerana*, which has not been estimated, has been taken to be equal to that of *A. mellifera*.

lineages in *A. mellifera*, as did preliminary results of Smith (1991).

The nucleotide divergence between lineages A and C is lower than between any of them and the lineage M (Table 2). This is reflected in the closer relatedness of lineages A and C in three out of four phylogenetic trees (the fourth one being unsolved for this question). It is then probable that the lineage M diverged first. However, a definite proof is lacking and we will hereafter adopt the conservative hypothesis that the three lineages diverged simultaneously.

Phylogeography of A. mellifera *inferred from mtDNA variability*

The high level of nucleotide divergence of mtDNA within *A. mellifera*, combined with the strong geographic structuring of its variability, clearly indicates that this species belongs to the phylogeographic category I defined by Avise *et al.* (1987). The first explanation, given by these authors, for such a pattern is a long-term geographic isolation of populations, resulting in independent evolution. In such a case, populations from different geographical origins occupy recognizable branches on an intraspecific evolutionary tree. An alternative explanation is a foundation from a highly polymorphic population followed by a loss of different variants in the various regions. The current level of intrapopulational or intraracial variability is low and does not favour the existence of a highly polymorphic ancestral population. In addition, within each geographical region, the major clades are subdivided into related subtypes and at least one of them (lineage C) shows a clear substructure, the most derived types being located in the western part of the geographic distribution (Fig. 3A). The long-term isolation of populations appears then as the most plausible hypothesis.

Additional information about the time at which the population groups characterized by the three major mtDNA types separated may be tentatively drawn from mtDNA phylogenies. Attention must be paid to the fact that gene trees and population trees are not always exactly superimposed (Pamilo & Nei, 1988). However, the separation of populations must come after the initial divergence of the genes, so that their current divergence gives an upper limit of the time estimate for the cleavage of the ancestral population. Since, as assumed above, the three lineages diverged almost simultaneously, we will use the average of the three interlineage distances (raw estimate in Table 2) which is equal to 2.7% for restriction data.

On the other hand, the existence of related subtypes within each of the three major mtDNA types is a strong indication that the corresponding population groups were geographically separated prior to the intralineage diversification. Therefore, the lower limit for the geographic separation of three major population groups is given by the beginning of the diversification of the mitochondrial genome within each of them. This last value can be estimated in each lineage from the average divergence between pairs of mtDNA types that are connected through the deepest node of the lineage. The average for the three lineages is equal to 0.66%.

Between these two limits, the most probable dating is given by that corresponding to the net distance between lineages (Wilson *et al.* 1985; Nei 1987; Ball, Neigel & Avise 1991) which amounts to 2.5% on average.

In the absence of any calibration of the mtDNA evolutionary rate in this taxon, the absolute dating of these events is difficult to estimate with sufficient accuracy. The application of the mammal (Brown, George & Wilson 1979) and *Drosophila* (DeSalle *et al.* 1987; Monnerot, Solignac & Wolstenholme 1990) rate of divergence of mtDNA (2% per Myr) to the three above values in the honey bee would give 1.35 Myr and 0.33 Myr as upper and lower limits and 1.25 Myr as the more plausible estimate. These values are only indicative: they could be easily half or twice as high.

Evolution of honey bee populations

To date, the interpretations of the evolution of the *A. mellifera* species have been largely based on morphometrical analysis and palaeogeography. Ruttner,

Tassencourt & Louveaux (1978) have shown that the various subspecies are distributed according to a Y-shape in the first plane of a discriminant analysis. The three branches of the Y have been interpreted as three evolutionary lines leading respectively to African subspecies (branch A), west Mediterranean subspecies (branch M) and north Mediterranean subspecies (branch C). Although morphometric characters are not suited to infer phylogenetic links between taxa (the genetic determinism is complex and homoplasy is frequent), this hypothesis, in agreement with palaeogeography, received support from the geographic distribution of allelic frequencies at the MDH locus (Badino, Celebrano & Manino 1982; Cornuet 1982, 1983) and from a tree based on more allozyme loci (Sheppard & Huettel 1988).

MtDNA analysis also supports most of Ruttner's conclusions: it allows us to distinguish, as morphometry did, three major lineages. However, although we have adopted Ruttner's abbreviations, some differences must be pointed out. The most important discrepancy is the set of subspecies included in the branch M. According to Ruttner *et al.* (1978), the branch M is composed of the 'chain of races' ('Rassenkreise') *intermissa–iberica–mellifera*. Our results indicate that the populations of north Africa (*intermissa*) are more closely related to the other African subspecies than to *mellifera*. The Iberian population presents two very distinct mtDNA types (see Table 1 and Fig. 3), which strongly evokes a recent contact between *mellifera* and *intermissa* populations with secondary intergradation. Furthermore, the *mellifera*-like type predominates in the north of Spain whereas the *intermissa*-like type predominates in the south (Smith *et al.* 1991).

Later, in a re-examination of his morphometrical data, Ruttner (1988) introduced a fourth branch (O) composed of the subspecies *meda–anatoliaca–caucasica*. This is not supported by mtDNA data which place *caucasica* samples at the base of branch C in the phylogenetic trees. In addition, when two Iranian samples belonging to the *meda* subspecies (G. Celebrano, personal communication) are added to our sample, they branch at the very base of the C lineage, and they do not form any particular branch with our *caucasica* samples.

It has been proposed by Ruttner *et al.* (1978) that the centre of dispersion of the three branches was located in north-eastern Africa. Their M branch, which includes all north-western African subspecies (*intermissa, sahariensis* and *major*), colonized western Europe through the Straight of Gibraltar. If this was true, a closer phylogenetic relationship should exist between our M lineage and *intermissa* mtDNA samples, a statement which is clearly incompatible with our molecular tree.

Considering that *Apis mellifera* originated in Asia (Deodikar, Thakar & Tonapi 1958) where all other *Apis* species are distributed, and that the Alpine arc, the Black Sea, the Caucasus and the Caspian Sea have been natural obstacles to colonization by bees, the simplest hypothesis is that the branch M results from an initial westward progression north to these obstacles whilst the other two branches progressed more southward (along the coast of the Persian Gulf). This leads to the location of the initial centre of dispersion of the proto-*mellifera* species in the Middle East (Fig. 4). Another separation should have occurred when the species came to the Mediterranean Sea, northern populations skirting it by the north (branch C) and southern populations invading Africa. Once separated by distance and natural obstacles, populations diverged progressively.

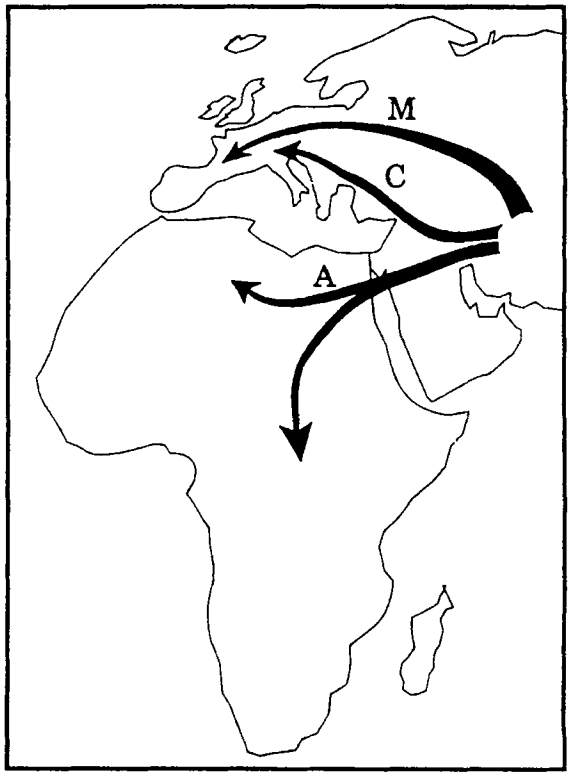

Fig. 4 Hypothesis for the origin and evolution of current *Apis mellifera* populations suggested by the relationship between the geographical origin of mtDNA types and their position in phylogenetic trees.

The westward progression of branch C can still be traced in the phylogenetic trees (Fig. 3). A comparable situation has not been found in the branch M, even when adjoining complementary mtDNA data from other parts of its current distribution area (Smith & Brown 1990). However, this can be explained by the glaciations during which populations have either become extinct or drawn to south-western refuges (Spain, southern France).

Acknowledgements

We wish to thank Dr J. David, Dr D. Lachaise, Dr R.F.A. Moritz , R. Aïteur and P. Carré for help in sample collection, and Dr M. Monnerot for critically reading the manuscript.

References

Avise JC, Arnold J, Ball RM (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Badino G, Celebrano G, Manino A (1982) Genetic variability of *Apis mellifera ligustica* Spin. in a marginal area of its geographical distribution. *Experientia*, **38**, 540–541.

Ball RM, Neigel JE, Avise JC (1991) Gene genealogies within the organismal pedigrees of random-mating populations. *Evolution*, **44**, 360–370.

Brown WM, George MJr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **76**, 1967–1971.

Cornuet JM (1982) The MDH polymorphism in some west-Mediterranean honeybee populations. *Proceeding of the International Union for the Study of Social Insects Congress*. Boulder, Colo.

Cornuet JM (1983) Reproduction, génétique et sélection de l'abeille. *Bulletin Technique Apicole*, **10**, 13–36.

Cornuet JM, Fresnaye J, Tassencourt L (1975) Discrimination et classification de populations d'abeilles à partir de caractères biométriques. *Apidologie*, **6**, 145–187.

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

Cornuet JM, Garnery L (1991) Phylogenetic relationship in the genus *Apis* inferred from mitochondrial DNA sequence data. *Apidologie*, **22**, 627–642.

Crozier RH, Crozier YC, Mackinlay AG (1989) The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates. *Molecular Biology and Evolution*, **6**, 399–411.

Deodikar G, Thakar CV, Tonapi KV (1958) Evolution in the genus *Apis* and its bearing on breeding better strains on Indian bees. *Proceeding of the International Beekeeping Congress*, **17**, 245–250.

DeSalle R, Freedman T, Prager EM, Wilson AC (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian Drosophila. *Journal of Molecular Evolution*, **26**, 157–164.

Efron B (1979) Boostrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1–26.

Garnery L, Vautrin D, Cornuet JM, Solignac M (1991) Phylogenetic relationships in the genus *Apis* inferred from mitochondrial DNA sequence data. *Apidologie*, **22**, 87–92.

Hall HG, Muralidharan K (1989) Evidence from mitochondrial DNA that African honey bees spread as continuous maternal lineages. *Nature*, **339**, 211–213.

Hall HG, Smith DR (1991) Distinguishing African and European honeybee matrilines using amplified mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **88**, 4548–4552.

Higuchi RG, Ochman H (1989) Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucleic Acids Research*, **17**, 5965.

Itenov K, Pedersen BV (1991) Nucleotide sequence differences in the mithochondrial CO-I gene in the honey-bee subspecies *Apis mellifera mellifera* and *Apis m. ligustica*. *Tidsskr. Planteavl*, **95**, 101–103.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions. *Journal of Molecular Evolution*, **16**, 11–120.

Kocher TD, Thomas WK, Meyer A *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA*, **86**, 6196–6200.

Koulianos S, Crozier RH (1991) Two ancient mitochondrial alleles in Australian honeybees. *Apidologie*, **22**, 621–626.

Miller, JC (1991) Restsite: a phylogenetic program that sorts raw restriction data. *Journal of Heredity*, **82**, 262–263.

Monnerot M, Solignac M, Wolstenholme DR (1991) Discrepancy in divergence of the mitochondrial and nuclear genomes of *Drosophila teissieri* and *D. yakuba*. *Journal of Molecular Evolution*, **30**, 500–508.

Moritz RFA, Hawkins CF, Crozier RH, Mackinlay AG (1986) A mitochondrial DNA polymorphism in honeybees (*Apis mellifera* L.). *Experientia*, **42**, 322–324.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Miller JC (1990) A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics*, **125**, 873–879.

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the USA*, **76**, 5269–5273.

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–583.

Ruttner F (1988) *Biogeography and Taxonomy of Honeybees*. Springer-Verlag, Berlin.

Ruttner F, Tassencourt L, Louveaux J (1978) Biometrical-statistical analysis of the geographic variability of *Apis mellifera* L. 1 Material and methods. *Apidologie*, **9**, 363–381

Saitou N, Nei M (1987) The neighbor–joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminator inhibitors. *Proceedings of the National Academy of Sciences of the USA*, **74**, 5463–5467.

Sheppard WS, Huettel MD (1988) Biochemical genetic markers, intraspecific variation, and population genetics of the honey bee, *Apis mellifera*. In: Needham GR, Page RE Jr, Delfinado-Baker M, Bowman C (eds) *Proceedings of the Conference on Africanized Honey Bees and Bee Mites*, Columbus, Ohio, New York. March 30–April 1, 1987, Ellis Horwood, New York. pp. 281–286.

Smith DR (1988) Mitochondrial DNA polymorphisms in five Old World subspecies of honey bees and in New World hybrids. In: Needham GR, Page RE Jr, Delfinado-Baker M, Bowman (eds) *Proceeding of the Conference on Africanized Honey Bees and Bee Mites*, Columbus, Ohio. March 30–April 1, 1987, Ellis Horwood, New York. pp. 303–312.

Smith DR (1991) African bees in the Americas: Insights from biogeography and genetics. *Trends in Ecology and Evolution*, **6**, 17–21.

Smith DR, Brown WM (1988) Polymorphisms in mitochondrial DNA of European and Africanized honeybees (*Apis mellifera*). *Experientia*, **44**, 257–260.

Smith DR, Brown WM (1990) Restriction endonuclease cleavage site and length polymorphisms in mitochondrial DNA of *Apis mellifera mellifera and A. m. carnica* (Hymenoptera: Apidae). *Annals of the Entomological Society of America*, **83**, 81–88.

Smith DR, Palopoli MF, Taylor BR, *et al.* (1991) Geographical overlap of two mitochondrial genomes in Spanish honeybees (*Apis mellifera iberica*). *Journal of Heredity*, **82**, 96–100.

Smith DR, Taylor OR, Brown WM (1989) Neotropical Africanized honeybees have African mitochondrial DNA. *Nature*, **339**, 213–215.

Solignac M, Monnerot M (1986) Race formation, speciation, and introgression within *Drosophila simulans*, *D. mauritiana*, and *D. sechellia* inferred from mitochondrial DNA analysis. *Evolution*, **40**, 531–539.

Swofford DL (1989) *Phylogenetic analysis using parsimony (PAUP)*, Release 3. Illinois Natural History Survey, Champaign, Ill.

Tomassone R, Fresnaye, J (1971) Etude d'une méthode biométrique et statistique permettant la discrimination et la classification de populations d'abeilles (*Apis mellifica* L.). *Apidologie*, **2**, 49–65.

Wilson AC, Cann R, Carr SM *et al.* (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, **26**, 375–400.

This paper presents a large part of the work performed by L. Garnery to obtain his Doctor in Sciences grade (PhD), under the joint scientific authority of J.-M. Cornuet and M. Solignac. J.-M. Cornuet, director of research at the National Institute of Agronomical Research (INRA) has specialized in the quantitative and molecular genetics of honey bees. M. Solignac, professor at Paris XI University, is involved in the study of evolutionary genetics by mean of molecular technics focused around mitochondrial DNA and ribosomal RNA.