Molecular Ecology (2005) **14**, 93–106 doi: 10.1111/j.1365-294X.2004.02399.x

Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe

ANNETTE B. JENSEN,*†‡ KELLIE A. PALMER,*† JACOBUS J. BOOMSMA† and BO V. PEDERSEN*
*Institute of Biology, Department of Evolutionary Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark, †Institute of Biology, Department of Population Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark

Abstract

The natural distribution of honeybee subspecies in Europe has been significantly affected by human activities during the last century. Non-native subspecies of honeybees have been introduced and propagated, so that native black honeybee (*Apis mellifera mellifera*) populations lost their identity by gene-flow or went extinct. After previous studies investigated the remaining gene-pools of native honeybees in France and Spain, we here assess the genetic composition of eight northwest European populations of the black honeybee, using both mitochondrial (restriction fragment length polymorphisms of the intergenic transfer RNAleu-COII region) and nuclear (11 microsatellite loci) markers. Both data sets show that *A. m. mellifera* populations still exist in Norway, Sweden, Denmark, England, Scotland and Ireland, but that they are threatened by gene flow from commercial honeybees. Both Bayesian admixture analysis of the microsatellite data and *Dra*I-RFLP (restriction fragment length polymorphism) analysis of the intergenic region indicated that gene-flow had hardly occurred in some populations, whereas almost 10% introgression was observed in other populations. The most introgressed population was found on the Danish Island of Læsø, which is the last remaining native Danish population of *A. m. mellifera* and the only one of the eight investigated populations that is protected by law. We discuss how individual admixture analysis can be used to monitor the restoration of honeybee populations that suffer from unwanted hybridization with non-native subspecies.

Keywords: Bayesian admixture analysis, conservation, hybridization, microsatellite, population genetics

Received 30 June 2004; revision received 03 September 2004; accepted 06 October 2004

Introduction

Molecular marker techniques have become a standard tool to address questions of phylogeography, population structuring, hybridization, and introgression (Avise 1994, 2000). In particular for conservation purposes, it is necessary to understand the extent to which local populations are genetically isolated (Franklin & Frankham 1998). However, when rare populations are threatened by hybridization with closely related taxa, effective isolation measures are a requirement for successful conservation. Such hybridization threats because of human-mediated introductions are a rapidly increasing problem for the conservation of wild populations and of original races of domesticated animals and plants (Allendorf *et al*. 2001; Rhymer & Simberloff 1996).

Domesticated honeybees have been kept for ages and have followed human migrations. The colonization by Europeans of the New World in the 16th and 17th century was the beginning of the present worldwide distribution of honeybee subspecies (Cornuet 1986) and the genetic mixing of these subspecies started shortly thereafter. In contrast to other domesticated animals, mating is very difficult to control in honeybees, so that gene flow between honeybee subspecies is common (Franck *et al*. 1998; Garnery *et al*. 1998a,b; Sheppard *et al*. 1991a,b) and introgression can

Correspondence: Annette Bruun Jensen, ‡Present address: Department of Ecology, Zoological Section, The Royal Veterinary and Agricultural University, Thorvaldsensvej 40, DK-1871 Frederiksberg C. Denmark. Tel.: + 45 35 28 26 66; Fax: + 45 35 28 26 60; E-mail: abj@kvl.dk

proceed very fast, as the recent spread of Africanized honeybees through the New World shows (Clarke *et al*. 2002; Pinto *et al*. 2004). Where Africanized bees are a clear case of unwanted traits sweeping through populations of domestic honeybees, the present study addresses the opposite: Southern European races that have replaced native honeybees in northwest Europe because of their supposedly superior performance in honey production.

Africa, Europe and western Asia were once the natural distribution area of the western honeybee, *Apis mellifera* L. (Ruttner 1988), which existed in several locally adapted races and subspecies. Based on morphology and ecological traits, *A. mellifera* L. can be grouped into four evolutionary lineages (Ruttner 1988) which have recently been validated (Garnery *et al*. 1992; Estoup *et al*. 1995; Franck *et al*. 2000b): an African lineage (A), a Middle East lineage (O) and two European lineages (C) and (M). The current natural distribution of the European honeybees originated during the last glaciation, when honeybees retreated to the Iberian Peninsula (M-lineage) and the Balkan peninsula (C-lineage) (Ruttner 1988; Hewitt 1996). About 10 000 years ago, the honeybees recolonized Europe with the M-lineage (including *Apis mellifera iberica* and *Apis mellifera mellifera*) becoming established in North and West Europe and the C-lineages (including *Apis mellifera ligustica*, *Apis mellifera carnica*, *Apis mellifera cecropia* and others) in Central Europe. Geographical barriers, such as the Alps, maintained the isolation of these lineages leading to the different subspecies as we know them today.

Since the beginning of the 20th century, commercial bee breeding has been dominated by introduced 'superior' honeybees, especially *A. m. ligustica* from Italy and *A. m. carnica* from former Yugoslavia, into northwest Europe. As a consequence of the direct replacement and the gene flow between native and commercially honeybee populations over longer distances (Peer 1957; Jensen *et al*. 2004), native honeybees are considered to be extinct in many parts of Europe. In Germany, for example, massive introductions have led to the almost complete replacement of *A. m. mellifera* by *A. m. carnica* (Kauhausen-Keller & Keller 1994; Maul & Hähnle 1994). In the Scandinavian countries and on the British Isles, most professional and hobby beekeepers today keep *A. m. ligustica*, *A. m. carnica* or synthetic strains such as the buckfast bee. The natural range of *A. m. mellifera* is thus likely to have become significantly reduced in recent years.

Microsatellites and RFLP (restriction fragment length polymorphism) have previously been used to analyse population structuring and hybridization of honeybees (Clark *et al*. 2001; Clark *et al*. 2002; Franck *et al*. 2000a). These earlier studies of *A. m. mellifera* have been focusing mainly on French and Iberian populations. The French *A. m. mellifera* populations hybridized with the C lineage, *A. m. ligustica* near the Italian border and with *A. m. carnica* near the German border (Garnery *et al*. 1998a,b), whereas a north–south cline of introgression with the African (A) lineage has been documented on the Iberian peninsula. (Garnery *et al*. 1995; Franck *et al*. 1998). The objective of the present study was to analyse the genetic composition of *A. m. mellifera* populations in the northwestern regions of Europe with both nuclear and mitochondrial markers to: (1) complete the genetic inventory of native *A. m. mellifera* in western Europe, (2) evaluate the genetic variability and differentiation between the extant *A.m. mellifera* populations, and (3) elucidate the degree of hybridization between and introgression within *A. m. mellifera* by imported anthropogenic honeybee subspecies.

Materials and methods

Sampling and DNA extraction

We obtained samples of *Apis mellifera mellifera* (preserved in 96% ethanol) via contacts with beekeeper associations in Scotland, England, Ireland, Sweden, Norway and Denmark and a single population of *Apis mellifera ligustica* from Denmark was included as a reference population (Fig. 1). Sample sizes varied from 30 to 52 colonies per population,

Fig. 1 The approximate natural distribution of *Apis mellifera* evolutionary lineages in Europe and the sites of the eight northwest European populations of *Apis mellifera mellifera* (Colonsay, Whitby, Sheffield, East Midtlands, Ireland, Flekkefjord, Hammerdal, Læsø), and the single *Apis mellifera ligustica* population (Jutland) that were sampled for the present study.

but only a single bee per colony was used for both nuclear microsatellite analysis and mitochondrial haplotype analysis. Total DNA was extracted from thorax muscles using a 5% Chelex solution (Walsh *et al*. 1991).

mtDNA analysis

The mtDNA region including the $tRNA_{leu}$ gene, the COI-COII intergenic region and 5′-end, and the COII subunit gene were amplified with the primers E2 and H2 (Garnery *et al*. 1991) using the conditions described by Garnery *et al*. (1993) with minor modifications. The length of PCR (polymerase chain reaction) products was determined on 1.5% agarose gels. Subsequently, the PCR products were subjected to restriction with *Dra*I enzyme and the lengths of the digested PCR fragments were determined on a 5% Nusieve agarose gel. Individual bees, each representing a separate colony, were classified according to their mtDNA haplotype. The *Dra*I-test can easily distinguish between the major branches of *A. mellifera* and has previously been used to investigate the genetic variability of many honeybee populations (Garnery *et al*. 1993). However, within *Dra*I haplotypes there may be small deletions, insertions or single point mutations in the nonrestriction sites, which can only be detected by direct sequencing. In the present study we sequenced representatives of each major haplotype to confirm their identity or uniqueness in comparison to previously published haplotypes (Franck *et al*. 1998).

Microsatellite analysis

Eleven microsatellite markers (A7, A8, A24, A28, A43, A88, A113, Ap36, Ap43, B124, A79), previous developed for *A. mellifera* and *Bombus terrestris* (for primer sequences see Estoup *et al*. 1994, 1995; Franck *et al*. 1998) were amplified. The PCRs were performed in 20 μL volumes containing 50–200 nM of each primer, 100 μM of each dNTP, 1.2–1.5 mM $MgCl_2$, 1 × buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 0.5 Unit *Taq* polymerase (Applied Biosystems) and 2 μL DNA extract. The PCR conditions consisted of an initial denaturation for 3 min at 94 °C, followed by 30 cycles with denaturation for 30 s at 94 °C, annealing for 30 s at 55 °C–60 °C, extension for 30 s at 72 °C and a final extension for 30 min at 72 °C. The PCR products were visualized and sized on an Applied Biosystem 377 DNA sequencer using the internal size-standard (ROX 500 Genesize) and subsequently analysed with GENESCAN v 3.1 and GENOTYPER v 2.5.

Analysis of mtDNA data

Unbiased estimates and standard deviations of gene diversity (D) of mtDNA were calculated according to Nei & Tajima (1981). Population differentiation based on haplotype frequencies was analysed with Fisher Exact tests and F-statistics. Bonferroni corrected P-values were used when multiple comparison were made (Rice 1989). The haplotype variation was categorized in AMOVA tests (Analysis of Molecular Variance). Programs included in the software package ARLEQUIN v 2.000 (Schneider *et al*. 2000) were used for these calculations. Genetic distances between populations based on the haplotype frequencies were calculated using the Cavalli Sforza & Edwards's chord distances (1967) as implemented in the software package PHYLIP v 3.6 (Felsenstein 1993). The obtained distances were used to construct a Neighbour-joining tree. We did these analyses both on our own data set and on a larger data set including all published *A. m. mellifera* haplotype data and additional data from *A. m. ligustica* and *Apis mellifera iberica* (Franck *et al*. 1998, 2000a; Garnery *et al*. 1993, 1998a), to obtain a more complete overview of the genetic structure of *A. m. mellifera* throughout Europe. To exclude any possible effects of introgressed genes from other honeybee races, we also analysed the M-haplotype data in isolation. Finally, we analysed the cyto-nuclear disequilibrium for each locus-haplotype combination with a Fisher exact test after generating probabilities with a Markov chain analysis (Asmussen *et al*. 1987) in GENEPOP v 3.3c (Raymond & Rousset 1995).

Analysis of microsatellite data

Unbiased estimates and standard deviations of gene diversity (H_E) were calculated according to Nei (1973). Exact test for Hardy–Weinberg equilibrium, genotypic linkages equilibrium and population differentiation (genic and genotypic) were performed with GENEPOP. Bonferroni corrections of the P-values were applied when multiple comparison were made (Rice 1989). The microsatellite variation was categorized in AMOVA tests and estimations of population subdivision (F_{ST}) were performed. Programs included in the software package ARLEQUIN were used for all calculations.

A neighbour-joining tree of individuals was constructed based on the proportion of shared alleles and on genetic distances D_s (Nei 1972) between populations based on the allele frequencies at each locus. Bootstrap values were computed based on 2000 replications. The procedures used (GENDIST, NEIGHBOUR and SEQBOOT) were all from the PHYLIP software package.

The genetic structure of the populations was further examined with two Bayesian clustering methods: BAPS v 2.0 (Corander *et al*. 2003) and STRUCTURE v 2.0 (Pritchard *et al*. 2000). BAPS estimates hidden population substructure by clustering populations into panmitic groups based on expected Hardy–Weinberg equilibrium and linkage equilibrium between loci within each of the observed populations. The method relies on geographical sampling information being available and assumes that no substructuring exists in addition to the sampling locations. The

results are based on simulations of 50 000 burn-in steps and 100 000 MCMC (Markov Chain Monte Carlo algorithm) iterations. We have run five such iterations to check for consistency. STRUCTURE clusters individuals into *K* panmictic groups by minimizing deviations from Hardy–Weinberg and gametic phase equilibrium without any prior information on the population-origin of the individuals included in the sample. To do this, STRUCTURE requires assumptions on the ancestry of the populations (ancestry models) and on the association of the allele frequencies between the populations (allele frequency models). Three models of ancestry without a priori information are available: The no admixture model, which assumes that each individual originates from one of the *K* populations, the admixture model, which assumes that each individual (*i*) has inherited some fraction of its genome from ancestors in all *K* populations, and the linkages model, which can be used if some of the loci are weakly linked. Because the bee populations are not very discrete, we used the admixture model, also because this model should be superior in hybridization studies. In addition to these models, there are two basic allele frequency models: The independent model, which assumes that allele frequencies differ at random across populations and the correlated model, which assumes that allele frequencies in the populations are viscous and can be parameterized in terms of F_{ST}. Because of the supposed recent common ancestry of all *A. m. mellifera* populations and their human mediated migration we have used the correlated allele frequency model, which is also the superior model for detecting structure among closely related populations. STRUCTURE provides probabilistic estimations of admixture coefficients (*Q*), i.e. the proportion of an individual's genotypes originating from a given subpopulation (*K*). The results are based on simulations of 80 000 burn-in steps and 1 000 000 MCMC (Markov Chain Monte Carlo algorithm) iterations. We used five iterations per run in order to check for consistency and chose the one with the highest estimated probability of the data under each model assuming between one and 12 subpopulations (*K*). The level of hybridization, i.e. the proportional introgression of *A. m. ligustica* nuclear alleles into each of the *A. m. mellifera* populations, was investigated by using the *A. m. ligustica* population from Northern Jutland (Denmark) as a standard (assuming $K = 2$). The two individuals that fell into the *A. m. mellifera* cluster in the neighbour-joining tree based on the proportion of shared alleles were excluded. Posterior distributions of individual admixture coefficients and their 90% probability intervals were estimated for each combination of the reference *A. m. ligustica* population with the eight *A. m. mellifera* populations. The pairwise analyses were also performed using the nonadmixed ancestry model and the independent allele frequency models in order to explore introgression under more strict conditions.

Individual assignments test based on the distance method, DA of Nei *et al*. (1983) were performed using the program GENECLASS v 2 (Piry *et al*. 2004). An individual was excluded from a given candidate population if its probability of belonging to a particular population was lower than 5%. To avoid possible bias as a result of 'self assignment' the 'leave one out' procedure was followed, which excludes the tested individual when calculating the allele frequency distribution of their own population. Potential problems caused by nonsampling of microsatellite alleles (when a population has a zero estimated frequency for a given allele, all individuals having that allele will be assigned as having a zero likelihood of originating from that population), were taken into account by assuming an allele frequency of 0.01 in the event of an observed zero frequency. The likelihood of each individual's genotype being found in a population was determined as described by Cornuet *et al*. (1999) and the calculation of probabilities was based on their re-sampling algorithm with 100 000 simulations.

Results

Population structure based on mitochondrial data

The *Dra*I RFLP analysis of the COI-COII intergenic region produced seven different mitochondrial haplotypes in the 379 colonies examined (Table 1). Six of the seven haplotypes were known from previous studies (Garnery *et al*. 1993, 1998a; Franck *et al*. 1998) but the seventh was new. The new haplotype now called M28 is similar to M4 but has a single bp deletion in the tRNAleu and a six bp deletion in the P element. Six of the haplotypes, including the new one, are known to belong to the west European M lineage (Garnery *et al*. 1993, 1998a; Franck *et al*. 1998). These haplotypes are characterized by the presence of the P element

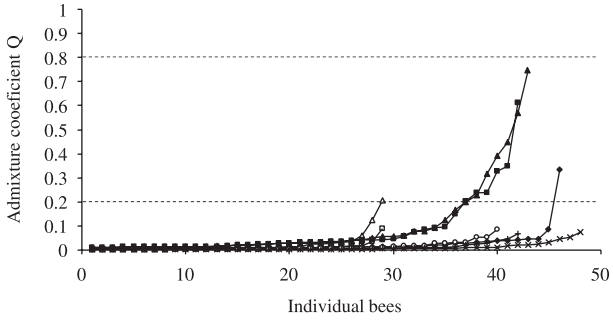

Fig. 2 Posterior distribution of individual admixture coefficients (Q) based on microsatellite markers in pairwise comparisons of the eight *Apis mellifera mellifera* populations with the single *Apis mellifera ligustica* population. Q-values of all individuals in a sample have been ranked from left to right. Colonsay (×), Whitby (◆), Sheffield (■), East Midtlands (△), Ireland (○), Flekkefjord (+), Hammerdal (□), Læsø (▲). Individuals inbetween the two dotted lines ($0.2 < Q > 0.8$) are highly admixed with *A. m. ligustica*.

Table 1 Population-wide haplotype frequencies of mtDNA according to the *Dra*I RFLP test of the intergenic region of COI-COII (Garnery *et al*. 1993). The composition of the six *Apis mellifera mellifera* haplotypes (M) and the single *Apis mellifera ligustica* haplotype (C) based on the presence and number of P and Q elements (Cornuet & Garnery 1991) is given below each recovered haplotype. *N* = the number of individuals analysed; D = the unbiased estimate of haplotype diversity ± SD (Nei 1987). DK = Denmark, IE = Ireland, NO = Norway, SE = Sweden, UK = United Kingdom

Population	Subspecies	*N*	M4 P_0QQ	M4′ P_0QQQ	M4″ P_0QQQQ	M7 P_0QQ	M28 P_0QQ	M3 P_0Q	C1 Q	D
Colonsay (UK)	*A. m. mellifera*	45	0.386	0.614						0.485 ± 0.037
Whitby (UK)	*A. m. mellifera*	46	0.717	0.174	0.043				0.065	0.459 ± 0.079
Sheffield (UK)	*A. m. mellifera*	48	0.375	0.563	0.021				0.042	0.552 ± 0.043
East Midlands (UK)	*A. m. mellifera*	28	0.464	0.107	0.393	0.036				0.640 ± 0.053
Ireland (IE)	*A. m. mellifera*	47	0.766	0.170			0.064			0.389 ± 0.078
Flekkefjord (NO)	*A. m. mellifera*	50	0.880	0.100	0.020					0.220 ± 0.073
Hammerdal (SE)	*A. m. mellifera*	30	1.000							0.000 ± 0.000
Læsø (DK)	*A. m. mellifera*	42	0.881					0.024	0.095	0.220 ± 0.080
Jutland (DK)	*A. m. ligustica*	43				0.023			0.977	0.047 ± 0.044

Table 2 Pairwise multilocus F_{ST} estimates of eight *Apis ellifera mellifera* populations and a single *Apis ellifera ligustica* population based on haplotype frequencies above the diagonal and on microsatellite frequencies below the diagonal. After Bonferroni corrections (Rice 1989) all F_{ST} values were significantly different from zero, except for the bold-face printed ones

		A. m. mellifera								*A. m. ligustica*
		Colonsay	Whitby	Sheffield	East Midlands	Ireland	Flekke-fjord	Hammer-dal	Læsø	Jutland
Colonsay	*A. m. mellifera*		0.234	**0.017**	0.262	0.272	0.418	0.560	0.463	0.732
Whitby	*A. m. mellifera*	0.055		0.196	**0.136**	**0.007**	**0.033**	**0.153**	**0.061**	0.724
Sheffield	*A. m. mellifera*	0.050	0.013		0.217	0.237	0.373	0.503	0.413	0.678
East Midlands	*A. m. mellifera*	0.050	**0.004**	0.020		0.191	0.284	0.412	0.297	0.703
Ireland	*A. m. mellifera*	0.065	0.008	0.026	**0.011**		**0.016**	**0.135**	**0.064**	0.776
Flekkefjord	*A. m. mellifera*	0.115	0.027	0.040	0.028	0.049		**0.060**	**0.023**	0.861
Hammerdal	*A. m. mellifera*	0.170	0.084	0.078	0.081	0.111	0.039		**0.060**	0.973
Læsø	*A. m. mellifera*	0.197	0.113	0.083	0.137	0.148	0.080	0.057		0.854
Jutland	*A. m. ligustica*	0.253	0.242	0.187	0.256	0.275	0.276	0.240	0.201	

with a deletion of 13 bp positioned 20 bp upstream. After the P element, one (M3), two (M4, M7, M28), three (M4′) or four (M4″) copies of the Q sequence follow. The C1 haplotype is known to belong to the Mediterranean C lineage and has no P element and only a single Q sequence. No haplotypes from the African (A) or the Middle East (O) lineages were found. M4 was the most common haplotype, present in all *Apis mellifera mellifera* populations and the most frequent haplotype in six of the eight populations. The exceptions, Colonsay and Sheffield, had M4′ as the most common haplotype, which was the second most common haplotype overall. Haplotype M4″ occurred in high frequency in the East Midland population and in low frequencies in three other populations. The M3, M7 and new haplotype were only found in few populations and with low frequencies. M7 is characteristic for the West European M lineages but has been found in most Italian *Apis mellifera ligustica* populations, where it is supposed to be a sign of an old hybridization event (Franck *et al*. 2000a). The occurrence of M7 in the *A. m. ligustica* population from northern Jutland, Denmark is therefore unlikely to be the result of recent introgression. The Mediterranean haplotype C1 was detected with a high frequency in the Danish *A. m. ligustica* population and in low frequencies in three *A. m. mellifera* populations. The unbiased estimates of haplotype diversity (*D*) varied between populations from zero to 0.64 (Table 1). Haplotype diversity in the British and Irish *A. m. mellifera* populations was higher than in the Scandinavian populations and similar to the haplotype diversity of French and Belgian populations, but not as diverse as Iberian populations (Franck *et al*. 1998; Garnery *et al*. 1995).

Fisher exact tests for population differentiation and pairwise multilocus F_{ST} values produced the same pattern of significances, with only in a single population pair being not concordant (Whitby and East Midlands). However, the respective *P*-values for this comparison were just above and just below the significance level and thus implied no real mismatch. We therefore present only the F_{ST} results (Table 2).

Table 3 Quantitative estimates and partitioning of haplotype and microsatellite variation in different AMOVA designs. The different groups were designed as follows: Two groups only distinguishing subspecies (*Apis mellifera mellifera* vs. *Apis mellifera ligustica*), two groups within *A. m. mellifera* characterizing the main geographical split between Britain and Scandinavia, and four groups within *A. m. mellifera* representing the genetic resolution obtained in the 'Structure' analysis: Denmark (Læsø), Norway + Sweden, England + Ireland, Scotland

	Among groups		Populations within groups		Within populations	
AMOVA design	Total variance	% of total	Total variance	% of total	Total variance	% of total
Mitochondria						
Two subspecies: *mellifera*, *ligustica*	0.637	58.05	0.123	11.24	0.337	30.71
Two groups within *A. m. mellifera*: British, Scandinavia	0.089	16.57	0.075	13.92	0.374	69.51
Four groups within *A. m. mellifera*: Læsø, England and Ireland, Scotland, Norway and Sweden	0.055	10.83	0.808	15.83	0.374	73.34
Microsatellites						
Two subspecies: *mellifera*, *ligustica*	0.725	22.76	0.168	5.26	2.294	71.98
Two groups within *A. m. mellifera*: British, Scandinavia	0.138	5.81	0.955	4.00	2.151	90.19
Four groups within *A. m. mellifera*: Læsø, England and Ireland, Scotland, Norway and Sweden	0.169	7.14	0.041	1.74	2.151	91.12

Twelve of the 28 possible pairwise F_{ST} comparisons of the *A. m. mellifera* population pairs were not significantly different, but the *A. m. ligustica* population was significantly differentiated from all *A. m. mellifera* populations and displayed significantly higher genetic distance in the F_{ST} comparisons (Table 2). More than half of the total haplotype variation observed (58%) was the result of differences between *A. m. mellifera* and *A. m. ligustica*. (Table 3). However when *A. m. ligustica* was excluded from the AMOVA, 69%–73% of the total variance was because of differentiation within populations. In the neighbour-joining tree of the examined populations *A. m. ligustica* is separated by a very long branch from a cluster composed of all the *A. m. mellifera* populations. Within the *A. m. mellifera* cluster, the Læsø population was closest to *A. m. ligustica* followed by the two other Scandinavian populations, whereas the British and Irish populations branched out in the 'canopy' of the tree (Data not shown).

Population structure based on nuclear data (microsatellite markers)

The number of alleles at the 11 loci varied across populations, from one (locus A28 in East Midlands) to 16 (locus B124 in Withby)(Table 4). The average number of alleles in the *A. m. mellifera* populations varied between 3.9 (Hammerdal) and 6.6 (Whitby), but the reference population of *A. m. ligusitca* had a higher level of polymorphism with 6.9 alleles per locus on average (Table 5). The average gene diversity measured as expected heterozygosity ranged from 0.391 (Ireland) to 0.525 (Sheffield) for the *A. m. mellifera* populations, and was again higher in the *A. m. ligusitca* reference population (0.693). F_{IS} estimates varied from slightly positive to slightly negative but none of them reached statistical significance.

The English samples came from three localities in Yorkshire and Derbyshire. Fisher's exact tests and F_{ST} analysis (Table 2) showed significant population differentiation for genotype and allele frequencies ($P < 0.05$), so that the three British populations were treated as separate populations in the further analyses. The Irish samples came from six localities (with a radius of approximate 15 km) covering eight different counties in the two provinces of Munster and Leister. For these sites, however, Fisher's exact tests for genotypic and genic differentiation showed no significant differentiation ($P > 0.05$ in all comparisons) and therefore the Irish samples were treated as one population in the further analyses.

Seven of the 99 locus-population combinations showed significant departures from Hardy–Weinberg expectation involving four different loci and six populations. However after applying the Bonferroni corrections only one locus-population combinations remained significant and across all loci none of the populations showed significant departures from Hardy–Weinberg equilibrium, indicating that the populations were panmictic overall. Seven of the 55 locus-pairs across all populations showed significant linkage disequilibrium, but only one locus pair (A43 and A79) remained significantly linked after Bonferroni correction.

Table 4 Genetic variation at microsatellite loci in eight *Apis mellifera mellifera* populations and a single *Apis mellifera ligustica* population from northwest Europe. N = sample size, n = the number of alleles detected, H_O and H_E are the observed and expected heterozygozity, respectively. The latter is also known as gene diversity. DK = Denmark, IE = Ireland, NO = Norway, SE = Sweden, UK = United Kingdom

	A. m. mellifera								*A. m. ligustica*
Locus	Colonsay (UK)	Whitby (UK)	Sheffield (UK)	East Midlands (UK)	Ireland (IE)	Flekkefjord (NO)	Hammerdal (SE)	Læsø (DK)	Jutland (DK)
A7									
N	50	48	48	30	49	54	30	45	44
n	5	6	8	6	5	5	3	7	9
H_O	0.62	0.417	0.688	0.267	0.469	0.352	0.067	0.578	0.841
H_E	0.686	0.425	0.625	0.249	0.475	0.313	0.066	0.531	0.855
A8									
N	47	49	48	29	46	52	29	42	43
n	3	5	6	4	4	3	2	5	8
H_O	0.489	0.347	0.521	0.31	0.152	0.423	0.276	0.429	0.698
H_E	0.449	0.32	0.501	0.304	0.164	0.347	0.242	0.423	0.793
A24									
N	50	49	51	29	49	52	30	48	44
n	2	3	3	3	3	3	3	3	3
H_O	0.3	0.265	0.451	0.31	0.265	0.308	0.533	0.604	0.545
H_E	0.258	0.24	0.391	0.275	0.241	0.41	0.581	0.516	0.574
A28									
N	50	49	50	29	50	52	30	48	44
n	3	3	3	1	4	2	2	3	2
H_O	0.18	0.163	0.18	0	0.1	0.019	0.233	0.188	0.432
H_E	0.203	0.155	0.201	0	0.098	0.019	0.21	0.207	0.411
A43									
N	49	46	50	30	46	45	29	46	44
n	2	2	2	2	2	3	2	4	3
H_O	0.122	0.478	0.48	0.4	0.239	0.467	0.517	0.283	0.568
H_E	0.116	0.39	0.424	0.325	0.276	0.517	0.487	0.349	0.558
A88									
N	48	46	51	29	48	52	30	48	43
n	2	4	4	2	2	2	2	3	7
H_O	0.042	0.217	0.137	0.207	0.146	0.019	0.033	0.188	0.581
H_E	0.041	0.238	0.166	0.189	0.137	0.019	0.033	0.173	0.605
A113									
N	50	46	44	27	47	54	30	43	44
n	5	8	9	5	3	7	3	5	7
H_O	0.56	0.543	0.568	0.37	0.362	0.444	0.233	0.488	0.432
H_E	0.531	0.532	0.612	0.378	0.429	0.413	0.264	0.565	0.525
Ap36									
N	49	46	49	30	42	42	28	45	37
n	9	12	9	7	10	9	6	11	13
H_O	0.776	0.717	0.816	0.767	0.833	0.857	0.857	0.733	0.892
H_E	0.77	0.846	0.847	0.842	0.815	0.824	0.794	0.764	0.903
Ap43									
N	50	49	43	30	46	54	30	44	40
n	2	6	7	4	5	4	4	6	6
H_O	0.52	0.49	0.535	0.7	0.413	0.37	0.267	0.432	0.775
H_E	0.502	0.534	0.571	0.562	0.476	0.401	0.389	0.454	0.817
B124									
N	45	47	47	21	43	47	29	45	38
n	9	16	12	10	14	10	10	11	9
H_O	0.844	0.851	0.915	0.952	0.814	0.809	0.931	0.822	0.737
H_E	0.853	0.897	0.901	0.869	0.789	0.801	0.842	0.842	0.789
A79									
N	50	50	49	28	49	54	28	48	44
n	7	8	7	6	7	7	6	4	10
H_O	0.46	0.28	0.592	0.393	0.347	0.148	0.357	0.729	0.818
H_E	0.52	0.307	0.537	0.4	0.4	0.286	0.594	0.652	0.817

Table 5 Multilocus microsatellite variability in the eight *A. m. mellifera* populations and the single *A. m. ligustica* population analysed. Values are averages ± SD. N is the mean sample size across all loci, n is the mean number of alleles per locus, H_O and H_E are the observed and expected average heterozygozities, F_{IS} is the inbreeding coefficient. DK = Denmark, IE = Ireland, NO = Norway, SE = Sweden, UK = United Kingdom

Population	Subspecies	*N*	*n*	H_O	H_E	F_{IS}
Colonsay; Scotland (UK)	*A. m. mellifera*	48.91 ± 1.64	4.5 ± 2.8	0.447 ± 0.260	0.448 ± 0.267	0.003
Whitby; North Yorkshire (UK)	*A. m. mellifera*	47.73 ± 1.56	6.6 ± 4.2	0.433 ± 0.213	0.444 ± 0.242	0.024
Sheffield; Yorkshire (UK)	*A. m. mellifera*	48.18 ± 2.64	6.4 ± 3.1	0.535 ± 0.234	0.525 ± 0.230	−0.018
East Midlands; Derbyshire (UK)	*A. m. mellifera*	28.36 ± 2.62	4.5 ± 2.6	0.425 ± 0.275	0.399 ± 0.265	−0.066
Ireland; Leinster and Munster (IE)	*A. m. mellifera*	46.82 ± 2.56	5.4 ± 3.7	0.376 ± 0.249	0.391 ± 0.243	0.037
Flekkefjord (NO)	*A. m. mellifera*	50.73 ± 4.15	5.0 ± 2.8	0.383 ± 0.273	0.395 ± 0.258	0.031
Hammerdal (SE)	*A. m. mellifera*	29.36 ± 0.81	3.9 ± 2.5	0.391 ± 0.293	0.409 ± 0.275	0.044
Læsø (DK)	*A. m. mellifera*	45.64 ± 2.16	5.6 ± 2.9	0.498 ± 0.219	0.498 ± 0.209	0.007
Jutland (DK)	*A. m. ligustica*	42.27 ± 2.65	6.9 ± 3.4	0.663 ± 0.164	0.693 ± 0.165	0.043

All populations were differentiated from each other ($P < 0.05$) in the Fisher's exact comparisons. This agreed rather well with the significance tests of the pairwise multilocus F_{ST} values (Table 2), where all except for three populations pairs (Whitby, East Midland and Ireland) were significantly different. The *A. m. ligustica* population was, as in the mitochondrial haplotype analysis, significantly differentiated from all the *A. m. mellifera* populations and had significantly higher genetic distance measured by the F_{ST} values (Table 2). The greatest amount of the total microsatellite variation (> 70%) was found within the populations. When *A. m. ligustica* was excluded from the AMOVA this fraction increased to more than 90% (Table 3).

The individual bees did fall into two main clusters based on the proportion of shared alleles representing the subspecies *A. m. mellifera* and *A. m. ligustica*, respectively. A few *A. m. mellifera* from the Læsø and Sheffield populations did fall into the *A. m. ligustica* cluster and two *A. m. ligustica* did fall into the *A. m. mellifera* cluster. No geographical pattern was observed within the *A. m. mellifera* individuals (Supplementary material, Figure S1). The overall topology of the microsatellite neighbour-joining tree based on population allele frequencies was very similar to the mitochondrial tree. The branching order in the crown of the trees varied little between the two marker systems (data not shown).

The Bayesian clustering methods allowed us to investigate the genetic architecture of the entire assembly of populations. In the BAPS (Corander *et al*. 2003) analyses, which used the geographical information given by the sampling design, the highest posterior probability ($P = 0.99999$) for population structure was found for a partitioning into five clusters, a cluster with the three English populations plus the Irish population, a cluster with the Norwegian and Swedish populations, and three clusters with the Colonsay, the Læsø, and the Jutland population, respectively. In STRUCTURE (Pritchard *et al*. 2000) the posterior probability of the data was highest when using a model that grouped the data into six populations. In the $K = 2$ model, the populations clustered in two distinct groups, the combined *A. m. mellifera* populations and the single *A. m. ligustica* population. With values of $K > 3$ the populations became arranged in four to six clusters of variable distinctness. The population from the Island of Colonsay (Scotland, UK), and the *A. m. ligustica* population consistently fell into two separate groups. The three English populations uniformly grouped together, whereas the relationship between the Scandinavian and the Irish populations differed between the analyses. However, the Swedish and Norwegian populations usually fell in between the Danish and the Irish populations. The estimated *K*-memberships were fractioned for most of the individuals because of shared alleles between the different populations, especially within the *A. m. mellifera* populations, which may indicate a fairly recent common ancestry or a high degree of admixture between the extant populations.

Introgression and hybridization

The M and C haplotypes are highly divergent (Arias & Sheppard 1996; Cornuet & Garnery 1991; Garnery *et al*. 1992; Smith 1991; Smith *et al*. 1991) so that their occurrence in the same population must be a consequence of admixture, as a result of intentional or accidental introduction of queens from the Mediterranean (C) lineage. Three of the *A. m. mellifera* populations, Sheffield (UK), Whitby (UK) and Læsø (DK) displayed mitochondrial introgression varying from 4%–10% (Table 1).

The proportion of introgressed *A. m. ligustica* alleles at microsatellite loci (defined as the mean estimated proportion of admixture (Q) per population) varied among the *A. m. mellifera* populations and exact estimates also differed between the different models (Table 6). Under the admixture model all *A. m. mellifera* populations showed evidence of *A. m. ligustica* nuclear introgression, varying

Table 6 Average percentage of introgressed *Apis mellifera ligustica* microsatellite alleles per individual in each of the eight *Apis mellifera mellifera* populations. Pairwise analyses were performed in 'STRUCTURE' (Prichard *et al*. 2000) exploring both admixed and nonadmixed ancestry models and correlated and independent allele frequency models

Ancestry model	Frequency model	Colonsay	Whitby	Sheffield	East Midlands	Ireland	Flekkefjord	Hammerdal	Læsø
Nonadmixture	Independent	—	0.5%	2.7%	—	—	—	—	4.6%
Nonadmixture	Correlated	—	0.6%	3.1%	—	—	—	—	4.8%
Admixture	Independent	0.8%	2.5%	7.8%	1.7%	1.1%	0.9%	0.9%	9.3%
Admixture	Correlated	1.0%	2.5%	7.7%	2.4%	1.3%	1.1%	1.0%	9.9%

Table 7 The fraction of individuals from each population that were assigned to their own or a different populations in a GENECLASS assignments test (Piry *et al*. 2004)

Assigned to Individuals from		*A. m. mellifera*								*A. m. ligustica*
		Colonsay	Whitby	Sheffield	East Midlands	Ireland	Flekkefjord	Hammerdal	Læsø	Jutland
Colonsay	*A. m. mellifera*	**0.74**	0.06		0.03	0.04		0.03		
Whitby	*A. m. mellifera*	0.06	**0.29**	0.06	0.03	0.16	0.02	0.03	0.02	
Sheffield	*A. m. mellifera*	0.20	0.65	**0.88**	0.70	0.71	0.67	0.30	0.09	
East Midlands	*A. m. mellifera*			0.02	**0.20**					
Ireland	*A. m. mellifera*				0.03	**0.10**	0.02			
Flekkefjord	*A. m. mellifera*						**0.13**	0.03		
Hammerdal	*A. m. mellifera*			0.02			0.09	**0.23**		
Læsø	*A. m. mellifera*						0.07	0.37	**0.83**	0.02
Jutland	*A. m. ligustica*			0.02					0.06	**0.98**

from 0.8%–9.9%, whereas the three populations Sheffield (UK), Whitby (UK) and Læsø (DK) even showed evidence of nuclear introgression (0.5%–4.8%) when using the non-admixture models.

The distribution of the individual admixture proportions (using the admixture and correlated allele frequency model) are shown in Fig. 2. Intermediary individuals ($0.2 < Q < 0.8$) were observed in four population (Læsø, Sheffield, Whithby and East Midtlands), the same populations that also had the highest proportion of introgressed alleles per population.

Results of the GENECLASS assignment tests are given in Table 7. The *A. m. mellifera* individuals not correctly assigned to the population of sampling were all assigned to another *A. m. mellifera* population except for one individual in the Sheffield population and three individuals in the Læsø population that were assigned to the *A. m. ligustica* population. For the single *A. m. ligustica* population, all except a single individual were correctly assigned.

Association of nuclear and mitochondrial markers

Cyto nuclear linkage disequilibria were detected in six instances out of 99 combinations involving three different populations and four different loci, but none remained significant after Bonferroni correction. Across all populations, the 11 cyto nuclear combinations were all in linkage equilibrium. However, at the higher evolutionary level, not distinguishing between the different haplotypes within the two clades, cyto nuclear linkage disequilibria were observed in two and three instances in the Læsø and Sheffield population, respectively, indicating recent introductions.

The three populations (Whitby, Sheffield and Læsø) that revealed *A. m. ligustica* derived mitochondria in the sampled material were the same three populations that displayed the highest proportion of introgressed *A. m. ligustica* nuclear alleles (Fig. 3). The Whitby population did, however, not differ much from the other *A. m. mellifera* populations for the nuclear markers, whereas the Sheffield and the Læsø populations had significantly higher proportions of both *A. m. ligustica* mitochondrial and nuclear genes. In these two populations individuals harbouring *A. m. ligustica* mitochondia had significantly higher proportions of *A. m. ligustica* nuclear alleles as well (Fig. 4).

Discussion

Generally, the nuclear and mitochondrial marker systems used in this paper yielded concordant results. They both showed that relatively pure *Apis mellifera mellifera* populations still exist in northwest Europe, but also that introgression of *Apis mellifera ligustica* alleles has occurred in varying degrees, in particular and in the single Danish population

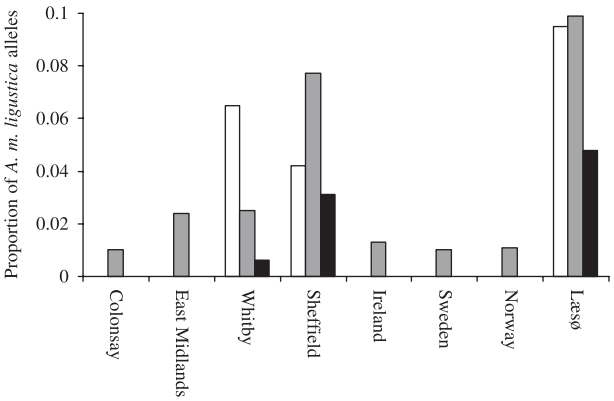

Fig. 3 The proportion of *Apis mellifera mellifera* individuals having *Apis mellifera ligustica* mitochondrial haplotypes (white bars) and introgressed *A. m. ligustica* alleles at microsatellite loci, using the admixture model (grey bars) and the nonadmixture model (black bars).

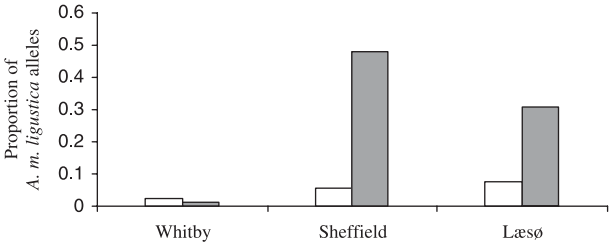

Fig. 4 Association of nuclear and mitochondrial markers. The mean proportion of *Apis mellifera ligustica* nuclear alleles in individuals having *Apis mellifera mellifera* (white bars) and *A. m. ligustica* (grey bars) mitochondrial haplotypes, respectively.

and in one of the English populations. The nuclear microsatellite markers were more variable than the mitochondrial marker and were therefore superior in detecting population differentiation and population structuring.

The existence of native honeybees in northwest Europe and their degree of differentiation

The original *A. m. mellifera* distribution once covered all of Europe north, west and east of the Alps and even the Ukraine (Ruttner 1988). The reductions that gave rise to the present scattered distribution of remnant populations has mostly taken place in just a single century by importation of and replacement by queens of other *Apis* subspecies. Currently, *A. m. mellifera* is known to have maintained populations in the British Isles and Scandinavia (this paper), and in Spain, France, and Belgium (Estoup *et al*. 1995; Franck *et al*. 1998; Garnery *et al*. 1995, 1998a,b). Recently sequence analysis has also confirmed the presence of *A. m. mellifera* populations in Switzerland and Poland (B.V. Pedersen, unpublished data).

In a formal analysis including all published haplotype data and the present results, it appeared that the Spanish populations of *A. m. mellifera* (which are mixed with the hybrid-subspecies *Apis mellifera iberica*) grouped together and so did the Italian and the Danish *A. m. ligustica* populations (Fig. 5). All the French, Belgian, British and Nordic *A. m. mellifera* populations clustered in the middle. The populations that had *c*. 50% or more C-haplotypes were all of French origin and clustered near the *A. m. ligustica* clade. However, in a similar analysis that only used the M-haplotypes these same populations clustered in between all the other *A. m. mellifera* populations (Fig. 5b). The analysis failed to detect a clear phylogeographical pattern of the Scandinavian, British, Belgian and French populations of *A. m. mellifera*, irrespective of whether introgression of A and C-haplotypes was taken into account (Fig. 5a) or not (Fig. 5b). This is because the extant *A. m. mellifera* populations mostly have the same spectrum of haplotypes, consistent with a fairly recent common ancestry (Garnery *et al*. 1998a).

The microsatellite markers were more effective in detecting population differentiation and population structuring than the mitochondrial markers. The populations from the British Isles were generally more closely related to each other than to the Scandinavian population. Interestingly, the beekeepers that provided bees from these populations, except for the Scottish one, are all members of Bee Improvers and Bee Breeders' Association, BIBBA, so that the most obvious explanation is that queens had been shared between members. The Scottish population on the island of Colonsay was more isolated and differs from the other British populations. Beekeepers have described bees from this region as being bigger, browner and having specific behavioural characters, which confirm their distinctness and suggest local adaptation (Cooper 1986; Ruttner *et al*. 1989).

Hybridization and introgession

Introgression of *A. m. ligustica* haplotypes and nuclear alleles was detected in varying degrees. In some populations, hardly any introgression was detected, whereas mitochondrial and nuclear introgression of almost 10% was recorded in other populations. The *A. m. mellifera* populations included in our study never reached the same high level of introgression of *A. m. ligustica* haplotypes and nuclear alleles as was detected in some of the French populations (Franck *et al*. 1998; Garnery *et al*. 1998a,b). However, only populations assumed to be *A. m. mellifera* were included in our analysis and most of these are involved in conservation programs. The population of *A. m. mellifera* on Læsø, which had the highest level of *A. m. ligustica* introgression, was only a subsample of the bees present on the Island. The beekeepers that keep *A. m. ligustica* or hybrids illegally, presumably more than half of the number of *A. m. mellifera* colonies present, refused to cooperate with samples so that bees from their colonies could not be included in this

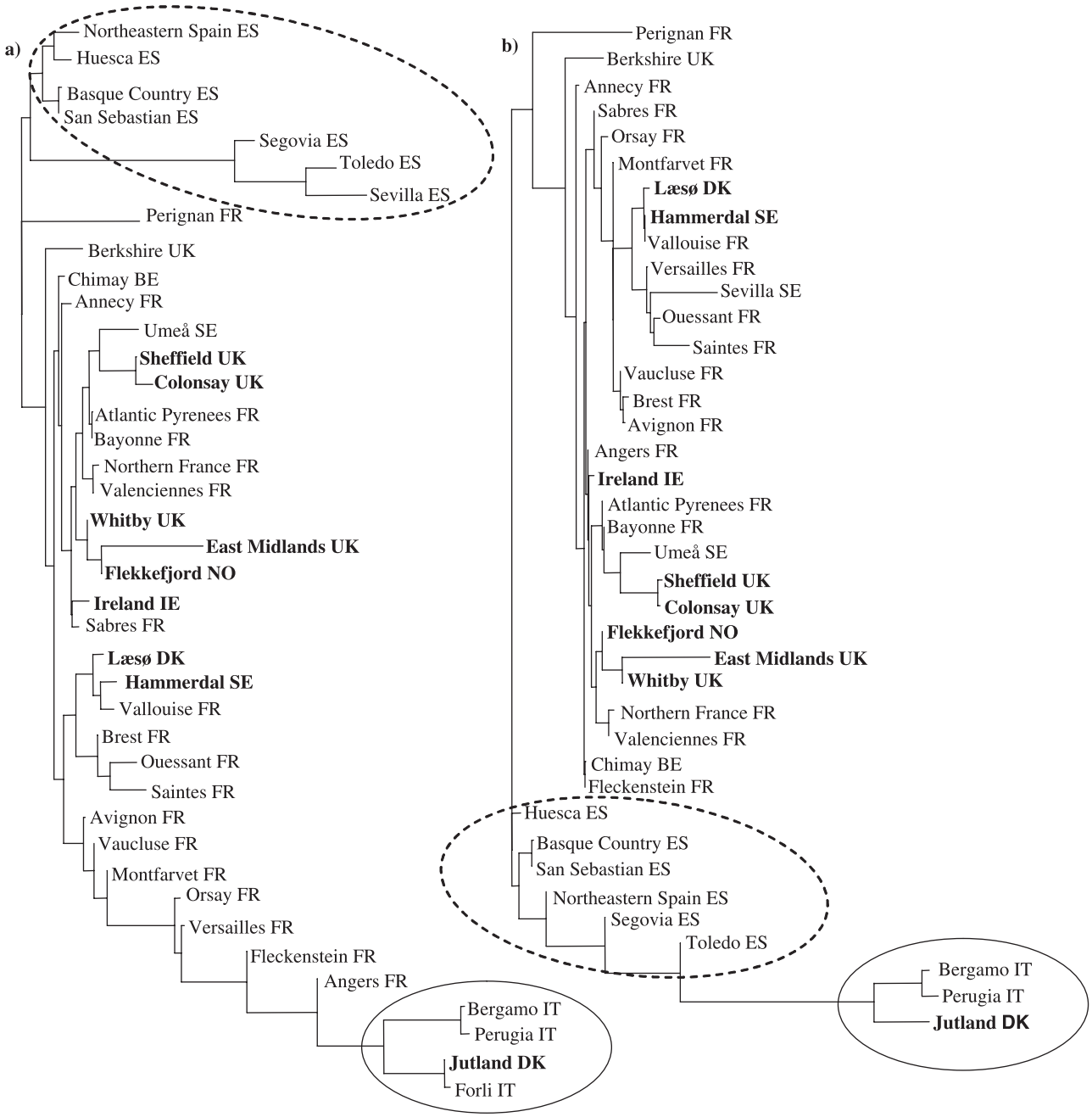

Fig. 5 The overall genetic relationships of western European honey bee populations, mainly *Apis mellifera mellifera*, based on mitochondrial haplotype data. (a) A neighbour-joining tree based on Cavalli Sforza & Edwards's chord distances (1967) estimated from the mitochondrial haplotype frequencies of the three evolutionary lineages present in West Europe (A, M, C; see Fig. 1). (b) A neighbour-joining tree based only on the M-haplotypes in the same populations. The eight *A. m. mellifera* populations (M-type) and the single *Apis mellifera ligustica* populations (C-type) of the present paper are given in bold face print. All other data are from previous studies (Franck *et al*. 1998, 2000a; Garnery *et al*. 1993, 1998a). The dotted circles include *Apis mellifera iberica* composed of a mixture of A- and M-lineages; The solid circles include *A. m. ligustica* from the C-lineage. BE = Belgium, DK = Denmark, ES = Spain, FR = France, IE = Ireland, IT = Italy, NO = Norway, SE = Sweden, UK = United Kingdom.

analysis. A genetic survey of foraging bees across the island in 2003 showed that only 35% of these foragers could be catagorized as *A. m. mellifera*, whereas 55% were hybrids and 10% were *A. m. ligustica* (Jensen & Pedersen 2005). The introgression and hybridization for the entire Læsø population may therefore be substantially higher than the 10% documented here.

In the current paper, only a single population of *A. m. ligustica* was used. This population is a very good reference population for the Læsø population as it is believed to be

the main source of the illegal yellow bees imported into the island and thus of hybridization with and introgression into the native *A. m. mellifera* population. The neighbouring honeybee populations of the other *A. m. mellifera* populations might also include some *A. m. carnica* or synthetic buckfast bees, but most of these belong to the same evolutionary (C) lineage as *A. m. ligustica*. Ideally, neighbouring population should have been sampled for each *A. m. mellifera* population studied but, given the substantial international exchange of commercial honeybee queens we believe that our analysis gives a reasonably accurate picture of the purity of north and west European *A. m. mellifera* populations.

Asymetrical gene flow between the sexes of hybridizing populations will result in cyto nuclear disequilibrium (Arnold 1993) and such disequilibria were indeed observed in the two most hybridized populations on Læsø and around Sheffield (Fig. 4). This nonrandom association of mitochondrial and nuclear genes can probably be explained by active managements of beekeepers involved in the conservation programs, who have relative good control over the genetic identity of queens, but not over the matings that these queens obtain. A recent genetic study of the efficiency of controlled mating in valleys around Sheffield showed that hybridization is still possible even when drone and queen colonies are 15 km apart (Jensen *et al.* 2005).

How to increase the genetic purity of admixed honeybee populations

The results of the present study show that the available genetic markers are so powerful that individual bees with a high degree of introgression can be identified. This information can be actively used as a tool in conservation management of populations that suffer from a high degree of hybridization with non-native bees. In colonies used for drone production only, the queen has to be tested as it is only her genome that is passed on to the haploid drones. A method for genotyping queens from just a tiny fragment of a wing is now available (Châline *et al.* 2004). For the colonies to be used for queen production the situation is more complicated because these queens may mate with drones of the wrong genotype. Assuming that honeybee queens generally mate with 10–20 different drones (Tarpy & Nielsen 2002), the proportion of unwanted father genotypes can be fairly accurately assessed by examining about 50 individual worker progeny. This would allow breeders to exclude queens with too many alien matings before grafting their larvae and rearing new queens, which should have a quite significant effect in just a few generations. However, this procedure would only work effectively when a routine monitoring program is established and most beekeepers in a breeding area (e.g. an island or an isolated valley) collaborate.

Implications for European management and conservation of Apis mellifera mellifera

In the present paper, we have shown that several relatively pure populations of *A. m. mellifera* still exist in northwest Europe. These populations are genetically quite homogenous as most of their genetic variation occurs within populations and not between populations. However, despite this uniformity at neutral genetic markers, *A. m. mellifera* has survived throughout western Europe from Spain to Norway and has developed several different ecotypes (Ruttner *et al.* 1989), indicating that *A. m. mellifera* has a relatively high evolutionary potential for local adaptation. The extant genetic diversity of *A. m. mellifera* throughout Europe therefore implies that this subspecies is a highly valuable gene pool for controlled breeding programs selecting for resistance against honeybee diseases.

It has been suggested that the native British *A. m. mellifera* bees went extinct resulting from the act of the Isle of Wight disease (Brother 1974), but others have argued that remnants of these original populations still exist (Cooper 1986; Ruttner *et al.* 1989). The relative distinctness of the British populations in our comparative study supports the latter argument and would justify and increase support of the British authorities for the conservation of British *A. m. mellifera.* In the Nordic countries, *A. m. mellifera* populations are covered by national conservation programs to a certain extent. In the area around Flekkefjord in southern Norway, bees other than *A. m. mellifera* are banned. This strategy works well in Norway, but has largely failed on the island of Læsø in Denmark, although *A. m. mellifera* is officially protected. On the Island of Læsø, only *A. m. mellifera* bees are allowed and this island would be an ideal conservation site because its distance to mainland Jutland is too huge for drones to reach the island. However, at present, the degree of hybridization with illegal bees of alien origin is increasing (Jensen & Pedersen 2005; Münster-Swendsen 1998, 2000), in spite of strong efforts by *A. m. mellifera* beekeepers on the island to prevent hybrids from breeding through. There is little doubt that effective conservation of remnant populations of black *A. m. mellifera* bees in Europe would require that authorities give them the same status as endangered races of other domesticated animals like cattle or sheep. At present, *A. m. mellifera* is not even included on the FAO's World Watch list for domesticated animals.

Acknowledgements

We thank the beekeepers who kindly provided samples, in particular Carl-Johan Junge (Læsø), Willy Svendsen (Jutland), Nils Drivdal (Norway), Ingvar Arvidson (Sweden), Andrew Abrahams (Scotland), Albert Knight, John Dews and Shirley Parkins (Bee Improvement and Bee Breeders Association; BIBBA, UK) and Jacob Kahn (Galtee Bee, BIBBA, Ireland). Pernille Selmer Olsen provided invaluable technical assistance in the laboratory and Rikke de Neergaard and Pilar de

la Rúa contributed valuable discussions. This work was supported by the EU (FW5-ENV) research network 'Beekeeping and *Apis* Biodiversity in Europe' (BABE)(contract EVK2-CT-2000–00068).

Supplementary material

The following material is available from http://blackwellpublishing.com/products/journals/suppmatt/MEC/MEC2399/MEC2399sm.htm

Appendix S1. Genotypes (PCR product sizes) of each individual bee using 11 microsatellite markers from the eight *A. m. mellifera* population examined: Colonsay (C). Whitby (W). Sheffield (S). East Midtlands (B). Ireland (I). Flekkefjord (N). Hammerdal (Sv) and Læsø (L), respectively, and from the *A. m. ligustica* bees from Jutland (JY) are given as a STRUCTURE input file. The data have not been standardized against previous published works. so we do not recommend direct comparison of allele sizes. Samples of DNA will be sent on request for standardizing procedures.

Figure S1. A neighbour-joining tree of individual bees based on the proportion of shared microsatellite allele distances. The solid lines represent *A. m. mellifera* bees from Colonsay (C). Whitby (W). Sheffield (S). East Midtlands (B). Ireland (I). Flekkefjord (N). Hammerdal (Sv) and Læsø (L), respectively. and the dotted lines represent *A. m. ligustica* bees from Jutland (JY). The *A. m. ligustica* individuals are typed in red.

References

Allendorf FW, Leary RF, Spurell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613–622.

Arias MC, Sheppard WS (1996) Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **5**, 557–566.

Arnold J (1993) Cytonuclear disequilibria in hybrid zones. *Annual Review of Ecology and Systematics*, **24**, 521–554.

Asmussen MA, Arnold J, Avise JC (1987) Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. *Genetics*, **115**, 755–768.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Avise JC (2000) *PhyloGeographyraphy: the History and Formation of Species*. Harvard University Press, Cambridge, MA.

Brother A (1974) *Beekeeping at Buckfast Abbey*. Northern Bee Books, Mytholmroyd, Hebden Bridge.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis models and estimation procedures. *Evolution*, **3**, 550–557.

Châline N, Ratnieks FLW, Raine NE, Badcock NS, Burke T (2004) Non-lethal sampling of honey bee, *Apis mellifera*, DNA using wing tips. *Apidologie*, **35**, 311–318.

Clarke KE, Oldroyd BP, Javier J, Quezada-Euan G, Rinderer TE (2001) Origin of honeybees (*Apis mellifera* L.) from the Yucatan Peninsula inferred from mitochondrial DNA analysis. *Molecular Ecology*, **10**, 1347–1355.

Clarke KE, Rinderer TE, Franck P, Quezada-Euan JG, Oldroyd BP (2002) The Africanization of honeybees (*Apis mellifera* L.) of the Yucatan: A study of a massive hybridization event across time. *Evolution*, **56**, 1462–1474.

Cooper BA (1986) *The Honeybees of the British Isles*. British Isles Bee Breeder's Association, Derby, UK.

Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Cornuet J-M (1986) Population genetics. In: *Bee Genetics and Breeding* (ed. Rinderer TE), pp. 235–254. Academic Press Orlando, FL.

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

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

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

Estoup A, Solignac M, Cornuet J (1994) Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies. *Proceedings of the Royal Society of London B*, **258**, 1–7.

Felsenstein J (1993) PHYLIP, *Phylogeny Inference Package, Version 3.6*. Distributed by the Author. Department of Genetics, University of Washington, Seattle.

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

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

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

Franklin IR, Frankham R (1998) How large must populations be to retain evolutionary potential? *Animal Conservation*, **1**, 69–73.

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

Garnery L, Franck P, Baudry E, Vautrin D, Cornuet J-M, Solignac M (1998a) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA. *Genetics Selection Evolution*, **30**, 31–47.

Garnery L, Franck P, Baudry E, Vautrin D, Cornuet J-M, Solignac M (1998b) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). II. Microsatellites. *Genetics Selection Evolution*, **30**, 49–74.

Garnery L, Mosshine EH, Oldroyd BP, Cornuet J-M (1995) Mitochondrial DNA variation in Moroccan and Spanish honey bee populations. *Molecular Ecolology*, **4**, 465–471.

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

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

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biology Journal of the Linnean Society*, **58**, 247–276.

Jensen AB, Palmer KA, Chaline N *et al.* (2005) Quantification of honey bee mating range and isolation in semi-isolated valleys in the Peak District National Park, England, as revealed by paternity analysis using DNA microsatellites. *(Submitted to Conservation Genetics)*.

Jensen AB, Pedersen BV (2004) Honeybee conservation: a case story from Læsø Island, Denmark. In: *Sustainable Bee Breeding* (ed. Lodesani M). Northern Bee Books, Mytholmroyd, Hebden Bridge (in press).

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

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

Münster-Swendsen M (1998) Field captures of honeybees on Læsø in 1998. *Tidsskrift for Biavl*, **12**, 377–379. (In Danish).

Münster-Swendsen M (2000) Bees on Læsø. *Tidsskrift for Biavl*, **6**, 183–186. (In Danish).

Nei M (1972) Genetic distances between populations. *American Naturalist*, **106**, 283–292.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Science of the USA*, **70**, 3321–3323.

Nei M (1987) *Molecular Evolutionary Genetics*. Colombia University Press, New York, NY. USA.

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

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.

Peer DF (1957) Further studies on the mating range of the honey bee. *Canadian Entomologist*, **89**, 108–110.

Pinto AM, Rubink WL, Coulson RN, Patton JC, Johnston SJ (2004) Temporal pattern of Africanization in feral honeybee population from Texas inferred from mitochondrial DNA. *Evolution*, **58**, 1047–1055.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS 2: an assignment software to select or exclude populations as origins of individuals. (submitted to *Molecular Ecology*).

Pritchard JK, Stephens M, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2), population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **7**, 248–249.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Ruttner F (1988) *Biogeography and Taxonomy of Honey Bees*. Springer Verlag, Berlin.

Ruttner F, Milner E, Dews JE (1989) *The Dark European Honey Bee: Apis Mellifera Mellifera Linnaeus 1758*. The British Isles Bee Breeders Association, Derby, UK.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Sheppard WS, Rinderer TE, Mazzoli JA, Stelzer JA, Shimanuki H (1991a) Gene flow between African and Europaen-derived honey bee populations in Argentina. *Nature*, **349**, 782–784.

Sheppard WS, Soares AEE, DeJong D, Shimanuki H (1991b) Hybrid status of honey bee populations near the historic origin of Africanization in Brazil. *Apidologie*, **22**, 643–652.

Smith DR (1991) Mitochondrial DNA and honey bee biogeography. In: *Diversity of the Genus Apis* (eds Smith DR), 131–176.Westview Press, Oxford and IBH Publishing Co.

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

Tarpy DR, Nielsen DI (2002) Sampling error, effective paternity, and estimating the genetic structure of honey bee colonies (Hymenoptera: Apidae). *Annals of the Entomological Society of America*, **95**, 513–528.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, **10**, 506–513.

The research of Annette B. Jensen is focused on population ecology of economical important insects and insect pathogenic microorganisms. Kellie A. Palmer is interested in multiple mating of honeybees. The main research interests of Bo V. Pedersen are phylogeny of European Bumblebees (Bombibi) and DNA variation in mt-genes in various strains of the Black honey bee *Apis mellifera mellifera*. Jacobus J. Boomsma is primarily interested in population biology of conflict and cooperation, i.e. in the evolutionary ecology (a.o. sex allocation and mating systems) and population genetics of insect societies including their social parasites and in the co-evolutionary study of mutualisms.
