Nuclear and mitochondrial patterns of introgression into native dark bees (*Apis mellifera mellifera*) in Poland

Andrzej Oleksa*1, Igor Chybicki1, Adam Tofilski2 and Jarosław Burczyk1

1Department of Genetics, Institute of Experimental Biology, University of Kazimierz Wielki, Chodkiewicza 30, 85-064 Bydgoszcz, Poland.

2Department of Pomology and Apiiculture, Agricultural University, 29 Listopada 54, 31-425 Kraków, Poland.

Received 30 December 2009, accepted subject to revision 10 November 2010, accepted for publication 21 March 2011.

*Corresponding author: Email: olek@ukw.edu.pl

Summary

The genetic diversity of the north and western European subspecies of honey bee, *Apis mellifera mellifera* (the "dark bee") is severely endangered due to hybridization with introduced bees of evolutionary branch C. Genetic variability of native honey bees in the north-eastern part of Poland, including a special isolated breeding zone in the Augustów Forest, has been investigated using mitochondrial DNA and nuclear microsatellites. These involve analysis for alien haplotypes of the tRNAleu-COII region and presence of diagnostic alien alleles respectively, in conjunction with a Bayesian model based approach. We found that approximately 10 to 30% of the nuclear gene pool and 3 to 50% of mitochondria in the studied populations were derived from non-native bees. Our data revealed the presence of hybrids in populations formerly considered to be the most pure populations of dark bees in Poland. We suggest that the Bayesian analysis of admixture based on nuclear microsatellites provides a reliable tool for measuring introgression in dark bees, which should be routinely used for evaluation during conservation programmes.

Introduction

The original distribution range of the honey bee (*Apis mellifera*) included Africa, Europe (except the northern part) and the Near East. Within this large area the species was differentiated into several evolutionary branches. Up to 29 subspecies have been distinguished on the basis of morphological traits (Ruttner et al., 1978; Ruttner, 1988; Sheppard and Meixner, 2003) and these were originally grouped into three main branches: African (Branch A); north-Mediterranean (Branch C); European (Branch M). Additionally, the lines Middle-East (O) and Yemen (Y) were identified (Franck et al., 2000) and some clades were rearranged within existing branches (Garnery et al., 1992; McMichael and Hall 1996;
These distributions were generally confirmed on the basis of allozyme variation (Badino et al., 1982; Sheppard and Huettel, 1988; Bouga et al., 2005; Ivanova et al., 2007), sequences of mtDNA (Garnery et al., 1992, 1993, 1998a; De la Rúa et al., 2004; Harizanis et al., 2006; Kandemir et al., 2006; Haddad et al., 2009), unique sequences of nuclear DNA restriction fragment-length polymorphisms (RFLPs) (McMichael and Hall, 1996), microsatellites (Estoup et al., 1995; Frank et al., 1998; Garnery et al., 1998b) and single nucleotide polymorphisms (Whitfield et al., 2006).

Despite the wide distribution of the honey bee and the large areas occupied initially by evolutionarily discrete units recognized within the species, the genetic diversity of local varieties of this ecologically and economically very important species is currently severely endangered due to introduction of alien genotypes by man (De la Rúa et al., 2009). In Europe there were originally three evolutionary branches: Branch M (subspecies A. m. mellifera, the “Dark European honey bee”) living in the British Isles (Carreck, 2008) and western and northern Europe from the Atlantic coast to the Ural Mountains and Branch C (A. m. ligustica, A. m. carnica and A. m. macedonica) existing south of the Alps and the Carpathian mountain ranges (Ruttnner, 1988; Sušnik et al., 2004; Dall’Olio et al., 2007; Muñoz et al., 2009). The third evolutionary branch, A (African) is represented in the Mediterranean region as a result of secondary contact between European and African subspecies of bees (Cánovas et al., 2008). The problem of extinction of the west and north European subspecies A. m. mellifera due to hybridization has been noted in several European countries, resulting in the establishment of organizations aiming to protect this subspecies, such as the Societas Internationalis pro Conservatione Apis Melliferae Melliferae (SICAMM). In Germany, where the introduction of new lines was carried out especially efficiently, the autochthon form of honey bee no longer exists, and all individuals are currently hybrids of subspecies: A. m. mellifera, A. m. carnica, A. m. ligustica and A. m. caucasica (Kauhausen-Keller and Keller, 1994). In Poland such admixture was somewhat delayed, compared to Western Europe (Gromisz, 1997). The area of Poland was originally inhabited mainly by A. m. mellifera, of evolutionary branch M. Only in the southern part of the country bees of evolutionary branch C, A. m. carnica occur (Gromisz, 1967). There was, however, a strong diversification of local races, differentiated not only by morphological traits, but also by various behavioural and breeding characters important for present day breeders. These include, honey gathering efficiency and resistance to negative environmental impacts (Prabucki and Chuda–Mickiewicz, 1996; Wilde et al., 2002a; b; 2003). In recent decades (especially since the 1960s) bees from other subspecies were introduced for bee breeding, mainly from south-eastern Europe. These were predominately Carniolan bees, A. m. carnica and their import resulted in severe disruption of the species’ gene pool. Genetic variation of A. mellifera in Poland has been only poorly investigated, and mainly for morphological and economic traits. Investigations of 389 colonies in the 1950s indicated that, only 35% of them displayed the morphological criteria of A. m. mellifera (Gromisz, 1997). Recently, Meixner et al. (2007) examined a large sample of bees from the eastern part of Poland using morphometrical analyses and concluded that bees from the northern part of this area could be unambiguously classified as A. m. mellifera. The proportion of uncertain allocations increased towards the south, where some samples were classified as hybrids between A. m. mellifera and subspecies of the C branch. Until now, however, no efforts have been made to quantify the introgression of genes from the C branch into the gene pool of honey bees in Central and Eastern Europe. Gromisz (1972; 1981; 1997) suggested that the only source of native genetic variants in Poland is specific breeding lines, selected for morphological traits connected with the A. m. mellifera phenotype. Four lines of such bees are included in the conservation programme of native dark bees in Poland. Two, Augustowska (in the Augustów Forest near the Polish-Lithuanian border) and Kampinoska (in the Kampinos National Park), have been preserved in their original habitats since the 1970s, whilst two others, Północna and Asta, were improved without loss of the characteristics of their native ancestors.

In the present study, we examine the genetic variability of the Augustowska and Północna (Northern) breeding lines that exist in the north-eastern part of Poland. Both lines were established from stock obtained within the region, believed to be one of the few areas of occurrence of native dark bees in Poland. Because of massive importation of bees from other parts of Europe and the Caucasus, protected populations are subjected to possible introgression of genes from bees of Branch C. In order to examine the level of admixture with non-native genes we used nuclear and mitochondrial DNA markers (PCR-RFLP analysis of mtDNA and nuclear microsatellites) since the two types of markers could show discordant patterns of differentiation in honey bees (Franck et al., 2001; Kraus et al., 2007). Our results were intended to serve as a measure of the effectiveness of conservation in a lowland area where isolation of breeding populations is achieved only by distance.

**Materials and methods**

**Sampling and DNA extraction**

Sampling of material was designed to comprise three zones differing in conservation regime: 1. the core area of protection in the very heart of the Augustów Primeval Forest; 2. the isolation zone with respect to the core area, of approximately 10 km diameter, where breeding of bees other than A. m. mellifera is forbidden by law; 3. other samples of bees in the vicinity, from the areas where importing of exotic material is allowed; 4. additionally, a single population of A. m. carnioca (10 colonies from the stock kept at Research Institute of Pomology and Floriculture in Pulawy) was included as a
reference population of C Branch (Table 1). In 2006, approximately 25 worker bees were collected from each of 83 bee colonies from 12 apiaries (localities) in north-eastern Poland (Fig. 1). Bees were collected directly from hive frames, killed by immersion in 90% ethanol and stored in a freezer at –20°C before DNA extraction from the metathoracic leg. In some apiaries (samples of Północna line), we analyzed material sampled for evaluation of phenotypic traits of tongues and wings, carried out as a part of the breeding programme. In this material, DNA had to be extracted from a smaller quantity of tissue, because only the studied bee parts (wings and tongues) were taken and preserved in 90% ethanol for future examination. Control reciprocal analyses showed, however, that the source of tissue for DNA extraction has no effect on genotyping results. DNA from all samples was extracted with a standard Chelex 100 protocol (Walsh et al., 1991). Extracted DNA was stored at –20°C until used for amplifications.

**Mitochondrial DNA analysis**

The main aim of the study was to assess changes in the gene pool of native Polish bees, representing Branch M, due to introduction of individuals from Branch C. Considering this introduction was presumably the result of importation of queens, the analysis of maternally inherited mitochondria should provide relevant information. Using PCR, the mtDNA region including the tRNALeu gene, the COI-COII intergenic region and the 5’ end of the COII gene was amplified with primers E2 (5’-GGCAGAATAAGTGCATTG-3’) and H2 (5’-CAATATCATTGATGACC-3’) according to the well established protocol described by Cornuet et al. (1991). This mtDNA region is especially appropriate for distinguishing between C and M branches because its variability results mainly from a varying number of sequences called P and Q (presence or absence of the P sequence, number of repeated Q sequences, possible small deletions; Garnery et al., 1998b). Within the C branch, the corresponding fragment size (ca. 570 bp) is relatively small due to absence of a P sequence and presence of only a single Q sequence. In general, bees of the M branch have a larger amplicon size, resulting from the combined presence of a P sequence and presence of only a single Q sequence. In the most common European haplotype, M4, there is one P sequence and two Q, so the total size is 825 bp.

We amplified the studied mtDNA region from three bees representative of each colony in a PCR reaction of total volume 15 μl (7.5 ml of Qiagen PCR Master Mix, BSA, primers E2 and H2, and deionised water to the total volume) and ran 5 μl of the product on 1% agarose gel to estimate the total size of the amplified fragment. The remaining part of the product was digested with the restriction enzyme Dral, and the resultant fragments were separated on 2% agarose gels. Banding patterns were photographed with CCD camera under UV light and analyzed using a computerized gel documentation system (Quantity One ver. 4.6.5, Bio Rad; USA).

Most haplotypes of the M branch could be easily interpreted through Dral analysis. Nevertheless, the length difference between

<table>
<thead>
<tr>
<th>Id</th>
<th>Locality</th>
<th>Geographic coordinates</th>
<th>Number of sampled colonies</th>
<th>Number of sampled individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Płaska</td>
<td>23°19’14”E, 53°53’54”N</td>
<td>30</td>
<td>630</td>
</tr>
<tr>
<td>2</td>
<td>Danowskie</td>
<td>23°6’11”E, 53°5’729”N</td>
<td>4</td>
<td>123</td>
</tr>
<tr>
<td>3</td>
<td>Jasionowo</td>
<td>23°18’28”E, 53°45’48”N</td>
<td>3</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>Muly</td>
<td>23°20’14”E, 53°54’3”N</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>Rudawka</td>
<td>23°31’3”E, 53°51’48”N</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>Sucha Rzeczka 1</td>
<td>23°11’13”E, 53°53’25”N</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>Sucha Rzeczka 2</td>
<td>23°12’17”E, 53°52’56”N</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td><strong>Σ 54</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Barciány</td>
<td>21°21’10”E, 54°13’6”N</td>
<td>7</td>
<td>210</td>
</tr>
<tr>
<td>9</td>
<td>Dolivy</td>
<td>22°21’16”E, 54°3’2”N</td>
<td>6</td>
<td>177</td>
</tr>
<tr>
<td>10</td>
<td>Kobiela</td>
<td>20°44’29”E, 54°5’46”N</td>
<td>5</td>
<td>150</td>
</tr>
<tr>
<td>11</td>
<td>Olecko</td>
<td>22°30’16”E, 54°2’13”N</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>12</td>
<td>Skandawa</td>
<td>21°15’8”E, 54°16’20”N</td>
<td>10</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td><strong>Σ 29</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Pulawy</td>
<td>21°58’03”E, 51°24’31”N</td>
<td>10</td>
<td>200</td>
</tr>
</tbody>
</table>

**Table 1.** Geographic locations and sample sizes of the studied dark bees (Augustowska and Północna lines) and Carniolan bees.
haplotypes C1 and C2 results only from the presence / absence of a cytosine at position 3428 (Franck et al., 2000; Sušnik et al., 2004; Özdil et al., 2009). Thus, the resolution need is beyond that of a standard agarose gel. To ensure that the interpretation of PCR-RFLP patterns is correct, we sequenced mtDNA region of five individuals and compared it with published C haplotypes (Franck et al., 2000; Sušnik et al., 2004; Muñoz et al., 2009). The DNA template for sequencing was prepared directly from PCR-products and purified by ultra-centrifugation on silica filter columns with the Clean-up kit (A&A Biotechnology; Poland) following the manufacturer’s instructions. The tRNA^{tRNA-COII} region was then sequenced with the primer E2 using the Fluorescent BigDye 3.1 Terminator Sequencing kit (Applied Biosystems) according to the manufacturer’s instructions. Unincorporated terminators were removed by ultra-centrifugation on silica filter columns with the Ex Terminator Kit (A&A Biotechnology; Poland). Products of sequencing PCR were re-suspended in 10 ml of formamide and then run on an ABI3130xl genetic analyser (Applied Biosystems) and processed with the Sequencing Analysis 3.7 program. Sequences were manually checked for size calling and aligned with published sequences (Franck et al., 2000; Sušnik et al., 2004; Özdil et al., 2009) with MEGA ver. 4 (Tamura et al., 2007).

Microsatellite analysis

In the study we used nine nuclear microsatellite loci: A7, A24, A28, A88, A113, Ap43, Ap55, Ap66, and Ap81 (Solignac et al., 2003), amplified in two multiplex reactions (Strange et al., 2008). Forward primers for these loci were 5’ labelled with fluorescent dyes. PCR was performed in a total volume of 10 µl containing 20 ng of extracted DNA, 1x Qiagen PCR Buffer, 0.2 mM dNTP mixture, 100–400 nM of each primer, 0.5 mg/ml bovine serum albumin and 0.25 units of Taq polymerase (Qiagen). The separation of fragments was carried out on automated sequencer ABI PRISM 3130xl (Applied Biosystems) using the internal size standard (LIZ 600, Applied Biosystems). Resulting electropherograms were scored using GeneScan ver. 3.7 and Genotyper ver. 3.7 software (Applied Biosystems). In order to make our results comparable with the previous results of Garnery et al. (1998b) we recalculated the estimated fragment sizes, assuming that the most common allele in other dark bee populations (usually frequency > 0.5) was also the most common allele in the studied bees of NE Poland, and calculating the sizes of other alleles in relation to that most common allele. We confirmed such an approach by sequencing of alleles (see above) and comparing sequences with previously published microsatellite sequences (for Accession Numbers, see Solignac et al., 2007).

Genetic structure parameters were calculated using 93 queen genotypes, inferred from worker genotypes, based on a maximum likelihood procedure available in MLTR software ver. 3.2 (Ritland, 2002). Because queen genotypes were identified based on the assessment of bees of a given colony there is a small probability of not detecting a queen allele because of sampling effect. Even when sampling as few as 10 bees per colony, however, the probability of inferring a queen’s actual heterozygous genotype as homozygote, is small and equals e = 0.5^{(10-1)} = 0.00195. This means that genotyping 123 colonies based on nine loci, assuming an average heterozygosity not larger then 0.64, it is expected that on average only 1.38 heterozygous genotypes will be wrongly identified as homozygotes.Allele frequencies, observed and expected heterozygosity, and polymorphism information content (PIC) values were computed using GeneAlEx (Peakall and Smouse, 2006). Hardy-Weinberg ratios were computed separately for the four studied “populations” or zones: Augustowska from the core area of protection, Augustowska from the isolation zone, Pólnocna and carnica. Deviations from Hardy-Weinberg equilibrium and population differentiation were assessed using exact tests implemented in GenePop ver. 4.0 (Rousset, 2008). Genetic structuring (F_{ST}) according to the method of Weir and Cockerham (1984) was computed with FSTAT ver. 2.9.3 (Goudet, 2001). The analysis of molecular variance (AMOVA) was performed using the program Arlequin ver. 2.0 (Excoffier et al., 1995). Finally, we inferred relations between studied populations with cluster analysis. Allelic frequencies were used to compute the matrix of genetic distances (chord distance of Cavalli-Sforza and Edwards, 1967) and a neighbour-joining tree was inferred with PHYLIP package (Felsenstein, 2009) and visualized in MEGA (Tamura et al., 2007).

Numbers of sampled colonies differed among the localities. To deal with the problem that large samples are expected to have more alleles than small samples, we employed rarefaction, a statistical technique allowing for comparison of samples of different size (Hurlbert, 1971). Estimation of allelic richness was performed with HP-Rare ver. 1.1 (Kalinowski, 2005).

Estimation of degree of admixture

For the mitochondrial genome, introgression at a population level could be simply measured as the proportion of colonies with haplotype C of the tRNA{tRNA-COII} region. In contrast to haplotypes M and A, haplotype C possesses no P sequence and only one Q sequence. There is no evidence that such a completely different haplotype originally existed within M branch, thus its presence within M branch should be interpreted exclusively as a result of introgression (Franck et al., 1998).

Hybridization between native and introduced honey bees was also examined in the nuclear genome (microsatellite loci) by the Bayesian statistical method implemented in the STRUCTURE ver. 3.2.1 (originally described in Pritchard et al., 2000). The program implements a model-based clustering method for inferring population structure using genotype data. The model assumes that there are K populations, each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are
probabilistically assigned to one of the populations, or jointly to two or more populations if their genotypes indicate they are admixed. This procedure accounts for the presence of Hardy–Weinberg and linkage disequilibrium by introducing population structure, and attempts to find population groupings that (as far as possible) are not in disequilibrium. The Markov Chain Monte Carlo method can allow the posterior probability distribution to be computed for estimated parameters. In order to estimate the proportional introgression of nuclear alleles from the C branch, we performed analysis on a dataset composed of genotypic data on studied populations and a reference population of A. m. carnica. STRUCTURE allows for different assumptions on the ancestry of the population (i.e., ancestral populations could be admixed or not admixed) and the association of allele frequencies (correlated or not correlated between populations). We used the admixture model, which assumes that each individual (i) has inherited some fraction of its genome from ancestors in all K populations, and the correlated allele frequency model. Such assumptions seem to be the most reliable in the case of the honey bee, but were also proved to be superior for detecting structure between closely related populations (Falush et al., 2003). In order to compare introgression estimates obtained under different assumptions, the analyses were also performed under a non-admixture model with the assumption of independent allele frequency. Five runs for each K-value, ranging from 1-10, were used to find the most likely uppermost level of structure, with an ad hoc statistic of delta K (Evanno et al., 2005). As a result the number of populations that has contributed to the gene pool of the studied population was estimated as K = 2. We thus report results only for K = 2. A burn-in of 50,000 iterations, followed by an MCMC (Markov Chain Monte Carlo algorithm) of 1,000,000 iterations was applied. In the analysis only queen genotypes, inferred from workers genotypes, as described above, were included. Mean individual admixture proportions, qi, and their 90% credible limits were estimated for each individual. Bees with qi not significantly different from 0 and significantly smaller than 0.5 were classified as A. m. mellifera; bees with qi not significantly different from 1 and significantly greater than 0.5 as A. m. carnica; all other intermediate cases were treated as hybrids. We used only one sample of bees from C branch as a reference for the source population of introgression. It was, however, possible that different bees representing C branch were imported and usefulness of such a single reference population could be problematic. Therefore, we verified our results from Bayesian clustering with an independent approach based on the frequency of microsatellite alleles diagnostic for the C evolutionary branch according to the formula \(IR = \frac{\Delta P}{\Delta q}\), where \(\Delta P\) is the frequency of \(\Delta q\) diagnostic allele in the studied population and \(q\) is the corresponding frequency in C-branch population (Garnery et al., 1998b). For computation, we used a set of diagnostic alleles designated by Garnery et al. (1998b) and A. m. macedonica (Chalkidiki, Greece) as a reference population (allelic frequencies from Garnery et al., 1998b). Calculations were carried out separately for the set of diagnostic alleles at six loci (A113, A24, A28, A7, A88 and Ap43) for which diagnostic alleles were indicated. \(IR\) was computed based on allelic frequencies for each bee colony separately, then these values were averaged to attain a corresponding value for each locality and breeding line.

**Results**

**Mitochondrial DNA analysis**

Three mtDNA haplotypes were revealed by DraI restriction digestion of the tRNAleu-COII intergenic region (Fig. 1). In the carnica sample, only haplotype C2 was recorded. The most common haplotype M4 occurred in 64.6% of the bee colonies of NE Poland. It was followed by C2 (32.9%) and M4’ (2.5%). More detailed analysis of C haplotypes sequences showed that two types of C2 haplotype (C2c and C2d) were present in the studied populations. Sequencing of a larger sample would however be necessary to estimate their frequencies. There were significant differences in geographic distributions of tRNAleu-COII haplotypes between the two breeding lines (Fig. 1). In general, the number of colonies with haplotypes originating from the M branch was higher in the Augustów primeval forest. As much as 76.5% of colonies from the Augustówka breeding line represented M haplotypes, while in the Pónocna breeding line, native and alien haplotypes were observed at equal frequency.

The pairwise multilocus \(F_{ST}\) estimates based on mitochondrial haplotype frequencies showed a different pattern from those based on microsatellites (Table 3). Mitochondrial DNA analysis revealed significant differentiation of the Augustówka line of the core protection area from all other bees of NE Poland (\(F_{ST} = 0.3\)), including bees of the same breeding line from the isolation zone of the protected area. Bees in the isolation zone were not differentiated in terms of their mtDNA from those of the Pónocna line (\(F_{ST}\) estimate was not significantly different from 0). All the A. m. mellifera bees from NE Poland were highly differentiated from A. m. carnica bees. These results are well supported by neighbour-joining cluster analysis based on genetic distances (Fig. 2). When the same analysis was performed with microsatellite data, bees from the two parts of the Augustów Forest clustered together very closely. The analysis based on haplotype frequencies showed that bees from the isolation zone of the core area of protection fall closer to the Pónocna population due to a similar degree of introgression of C haplotypes. Results from AMOVA showed that more than 40% of the total haplotype differentiation was attributable to differences within
Introgression in *A. m. mellifera* in Poland

When the *carnica* sample was excluded from the AMOVA, 81% of the total variance was due to haplotype differentiation within the population; and no differentiation was observed between the Augustowska and Północna breeding lines.

**Nuclear microsatellite DNA analysis**

All nine microsatellite loci were polymorphic in the material studied. In the sample of queen genotypes, there were 73 alleles at nine loci, where 42 alleles could be described as rare, with a frequency of 5%.

---

**Fig. 1.** Location of studied bee populations in NE Poland: Pn – Północna breeding line; Au – Augustów breeding line (lower panel); and frequencies of COI-COII haplotypes (upper panel). Radius of circles is proportional to the number of sampled colonies.

**Fig. 2.** Neighbour-joining trees based on chord distances (Cavalli-Sforza and Edwards, 1967): a. based on mtDNA haplotypes; b. based on nuclear microsatellites. Each tree is rooted in a midpoint of the longest branch.
<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>Augustowska, core area of protection (n=60)</th>
<th>Augustowska, isolation zone (n=48)</th>
<th>Pólnocna (n=58)</th>
<th>A. m. carnica (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A7</td>
<td>118*</td>
<td>0.013</td>
<td>0.022</td>
<td>0.022</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>120*</td>
<td>0.017</td>
<td>0.021</td>
<td>0.022</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>122*</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>124</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>132</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>134</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A88</td>
<td>143</td>
<td>0.067</td>
<td>0.091</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>146</td>
<td>0.700</td>
<td>0.750</td>
<td>0.759</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>148</td>
<td>0.150</td>
<td>0.114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>152*</td>
<td>0.017</td>
<td>0.034</td>
<td>0.050</td>
<td></td>
</tr>
<tr>
<td></td>
<td>154*</td>
<td>0.050</td>
<td>0.045</td>
<td>0.103</td>
<td>0.800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>156</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A24</td>
<td>98</td>
<td>0.667</td>
<td>0.705</td>
<td>0.500</td>
</tr>
<tr>
<td></td>
<td>102</td>
<td>0.091</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>106</td>
<td>0.222</td>
<td>0.136</td>
<td>0.241</td>
<td>0.500</td>
</tr>
<tr>
<td></td>
<td>108*</td>
<td>0.111</td>
<td>0.068</td>
<td>0.259</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS</td>
<td>0.542</td>
<td>0.495</td>
<td>0.588</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ho</td>
<td>0.607</td>
<td>0.583</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIC</td>
<td>0.504</td>
<td>0.470</td>
<td>0.560</td>
</tr>
<tr>
<td></td>
<td>A28</td>
<td>130</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>132</td>
<td>0.733</td>
<td>0.771</td>
<td>0.603</td>
<td>0.200</td>
</tr>
<tr>
<td></td>
<td>134</td>
<td>0.050</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>136</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>138*</td>
<td>0.250</td>
<td>0.229</td>
<td>0.397</td>
<td>0.700</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HS</td>
<td>0.406</td>
<td>0.361</td>
<td>0.487</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ho</td>
<td>0.467</td>
<td>0.458</td>
<td>0.448</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIC</td>
<td>0.332</td>
<td>0.291</td>
<td>0.364</td>
</tr>
<tr>
<td></td>
<td>A7</td>
<td>107</td>
<td>0.167</td>
<td>0.063</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>109</td>
<td>0.017</td>
<td>0.083</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>0.043</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>111</td>
<td>0.600</td>
<td>0.646</td>
<td>0.717</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>112</td>
<td>0.017</td>
<td>0.021</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>113</td>
<td>0.042</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>114</td>
<td>0.150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>116*</td>
<td>0.017</td>
<td>0.104</td>
<td>0.065</td>
<td>0.300</td>
</tr>
</tbody>
</table>
Allele frequencies for all loci and studied populations are shown in Table 2, which also provides the observed (Ho) and expected (He) heterozygosities. Two breeding lines of A. m. mellifera and A. m. carnica bees had nearly the same allelic richness when we accounted for different sample sizes by sampling 20 genes with rarefaction (average number of alleles per locus ranged from 3.81 to 4.15, the differences were not significant).

The dark bee (A. m. mellifera) population showed lower heterozygosity than the Carniolan. Expected heterozygosity averaged over loci ranged from 0.49 in the Augustowska line, to 0.64 in carnica bees, with the intermediate value of 0.55 in the Pólnocna line. Observed heterozygosity showed similar patterns: the smallest value in the Augustowska line, at 0.57, intermediate in the Pólnocna line, at 0.63, and the highest in carnica bees, at 0.68. No departures from Hardy-Weinberg equilibrium were significant (with \( p < 0.05 \)) in the dark bee populations, but two departures (in loci A113 and Ap43) were observed in the reference carnica population.

Genetic differentiation estimates based on microsatellite frequencies were low between the studied localities in NE Poland and more prominent when dark bees were compared with reference population of carnica. In general, Fisher exact tests for population differentiation (detailed results not shown) and pairwise multilocus FST values (Table 3) produced concordant patterns of significances.

Differentiation of bees from the core protection area of the Augustów line and its isolation zone was not significantly different from 0 when measured with the aid of FST and barely significant when examined with the Fisher’s exact test (\( p = 0.047 \)). All other pairwise estimates of differentiation were significant when assessed with both methods, with FST values 0.03 (comparison of Augustowska and Pólnocna lines).

or less (\( p < 0.05 \)). Allele frequencies for all loci and studied populations are shown in Table 2, which also provides the observed (Ho) and expected (He) heterozygosities.

Two breeding lines of A. m. mellifera and A. m. carnica bees had nearly the same allelic richness when we accounted for different sample sizes by sampling 20 genes with rarefaction (average number of alleles per locus ranged from 3.81 to 4.15, the differences were not significant).

The dark bee (A. m. mellifera) population showed lower heterozygosity than the Carniolan. Expected heterozygosity averaged over loci ranged from 0.49 in the Augustowska line, to 0.64 in carnica bees, with the intermediate value of 0.55 in the Pólnocna line. Observed heterozygosity showed similar patterns: the smallest value in the Augustowska line, at 0.57, intermediate in the Pólnocna line, at 0.63, and the highest in carnica bees, at 0.68. No departures from Hardy-Weinberg equilibrium were significant (with \( p < 0.05 \)) in the dark bee populations, but two departures (in loci A113 and Ap43) were observed in the reference carnica population.
and values approximately one order of magnitude larger (approximately 0.3) for comparisons of A. m. mellifera and carnica bees.

AMOVA based on microsatellite data (Table 4) revealed a high differentiation between bees from NE Poland and the carnica sample (24% of the total variance), although most of the genetic variation was observed within populations. When carnica bees were excluded from the analysis an even greater proportion of the total variation (>90%) was found within populations, while differentiation between the two studied breeding lines of dark bee (Augustowska and Północna) was not significant.

Estimation of introgression

The frequency of colonies with mtDNA haplotypes originating from the C evolutionary branch was differentiated between the two breeding lines of dark bee (for detailed results, see Fig. 1 and the previous section). Alien haplotypes of the C evolutionary branch were encountered even in the core area of the closed breeding region in the Augustów primeval forest in one out of 30 colonies. In one, westernmost apiary of the Północna breeding line, C haplotypes were found exclusively. All other apiaries of both lines showed variable proportions of C haplotypes, ranging from 1/4 to 2/3. Overall the proportion of the colonies with C haplotypes outside the most strictly protected area of Augustowska bees was 45% in the isolation zone of the closed breeding region, and 50% among the Północna bees.

A Bayesian method of clustering performed with STRUCTURE showed evidence of introgression in all studied populations of dark bees from NE Poland. The analysis, however, revealed differences between the level of introgression in the Augustowska and Północna breeding lines. When the admixture model with correlated allele frequency was assumed, the average proportion of introgressed microsatellite alleles amounted to 11% in the core protection area of the Augustowska line, 8% in its isolation zone, and 32% in the Północna bees. The majority of queens from the Augustowska line could be assigned as A. m. mellifera when criterion qi < 0.2 was applied (see materials and methods). Bees of the Północna line have a wider range of membership coefficients. The distribution of individual admixture coefficients is shown in Fig. 3. Different assumptions on the ancestry of populations and correlation in allele frequencies yielded similar estimates (results not shown), but regardless of the model assumed, certain queens were consistently found to have highly admixed ancestry.

When introgression was measured with respect to the C diagnostic alleles at microsatellite loci related to their frequency in C
Estimates of $q$ and IR were highly correlated ($r = 0.700$, $p < 0.001$), respectively, but the difference was not significant by U-test.

Among the Augustowska bees, colonies from the core area of the closed breeding region (locality Płock) had slightly higher value of IR than bees from the isolation zone (mean values 14.7% and 22.5%, respectively; $p < 0.05$).

This is the first research based on molecular markers, to confirm the presence of a viable population of *A. m. mellifera* in Poland, although there have been numerous morphometric studies (Gromisz, 1967; 1972; 1997; Meixner et al., 2007) which concluded that bees from northeastern part of the country could be classified as representatives of the native subspecies. Presence of the M haplotype in Poland is mentioned only in one abstract (Pedersen, 2002) and two publications (Garnery et al., 1998 a; Jensen, 2005 a) as a reference to unpublished data. The Polish population of *A. m. mellifera* may be an important complement to the few other surviving populations of this subspecies in France, Belgium (Estoup et al., 1995; Franck et al., 1998;
much smaller (ca. 10 km), thus being within the flight range of drones, given the flat, lowland topography. Nuclear DNA markers could have remained in the studied population at a high proportion because daughters were kept in breeding. On the other hand, nuclear genes, which are bi-parentally inherited, could have been efficiently diluted through mating with local males (in the honey bee, queens mate with a nearly panmictic sample of males; Oldroyd et al., 1995; Baudry et al., 1998). This is because beekeepers can control the genetic identity of queens, but cannot control the matings these queens experience. In previous studies of protected European honey bee populations (Jensen et al., 2005a; b; Soland-Reckeweg et al., 2009) they were isolated mainly by topography (mountain valleys) or water (islands). Genetic study of the efficiency of spatial isolation in mountain valleys showed that even 15 km between drone and queen colonies is not enough to eliminate gene flow (Jensen et al., 2005 b). The radius of the closed region of breeding in the Augustów Forest is about 11%. A similar pattern was reported by Garnery et al. (1998 b) and it could be a result of ancestral importation of queens, before the protection of native bees had started. In such a case, alien mtDNA could have remained in the studied population at a high proportion because daughters were kept in breeding. On the other hand, nuclear genes, which are bi-parentally inherited, could have been efficiently diluted through mating with local males (in the honey bee, queens mate with a nearly panmictic sample of males; Oldroyd et al., 1995; Baudry et al., 1998). We can therefore expect these to yield reliable estimates of introgression.

The usefulness of molecular genetic markers in discrimination of purebreds from hybrids should be interpreted with caution. The precision of estimates of introgression from unlinked loci depends on their number and distribution throughout the genome. The nine microsatellite loci studied could be considered as physically unlinked. Loci Ap55, A24, Ap81, A28 and A113 are on different chromosomes (1, 7, 9, 14 and 16, respectively), while four others are on the same chromosomes (Ap43 and Ap66 on 3, and A7 and A88 on 8), but at relatively large distance from one another (87.6 and 65.4 cH, respectively, Solignac et al., 2007). We can therefore expect these to yield reliable estimates of introgression.

Data from DNA markers, both mtDNA and microsatellites, provide information that can be actively used as a tool in conservation management of preserved breeding populations that have experienced hybridization with non-native bees (Jensen et al., 2005). Currently, bees from the protected populations of Poland and most other populations in Europe are evaluated by morphometric analyses. Our results show that morphometry may be insufficiently precise to allow for correct identification of hybrid individuals. Genetic markers used in this study have the potential to increase the genetic purity of introgressed A. m. mellifera populations because colonies with high levels of introgression can be readily identified and eliminated from breeding stock.

The potential pitfall of marker assisted selection is that the relation between markers and the genes responsible for phenotypic traits is unknown. We must keep in mind that the final goal of conservation is the preservation and restoration of indigenous populations in their original habitats. This means breeding individuals with both characteristic phenotypes and the capability of transmitting them to their offspring. Selection based only on neutral markers may result in loss of genes important for shaping phenotype, if diagnostic alleles for the M lineage and alleles responsible for characteristic dark bee phenotype are not in fact linked. A combination of phenotypic and marker assisted selection is therefore necessary.

Acknowledgements

We thank Mrs Wioletta Naruszewicz and Mrs Łucja Skonieczna (breeding apiaries of Animals Inseminating Center in Bydgoszcz) for kindly providing samples of bees from NE Poland, and Dr Małgorzata Bieńkowska (Research Institute of Pomology and Floriculture in Pulawy) for providing a sample of Carniolan bees. Katarzyna Kowalkowska, Ewa Szuciak-Piątek and Artur Dziulak helped us with laboratory analyses of bees. We further like to thank Dorian J Pritchard for helpful discussions and the lingual revision. We also thank the anonymous reviewers for their helpful comments and suggestions on the earlier version of the manuscript. The study was financially supported by a grant from the Polish Ministry of Sciences, No N311 029 32/2255.

References


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.


JENSEN, A B; PALMER, K A; BOOMSMA, J J; PEDERSEN, B V (2005a) Varying degrees of Apis mellifera ligustica introgression in protected populations of the black honey bee, Apis mellifera mellifera, in northwest Europe. Molecular Ecology 14: 93-106. DOI: 10.1111/j.1365-294X.2004.02399.x

JENSEN, A B; PALMER, K A; CHALINE, N; RAINÉ, N E; TOFILOSKI, A; MARTIN, S J; PEDERSEN, B V; BOOMSMA, J J; RATNIJEK, F L W (2005b) Quantifying honey bee mating range and isolation in semi-isolated valleys by DNA microsatellite paternity analysis. Conservation Genetics 6: 527–537. DOI: 10.1007/s10592-005-9007-7


MEIXNER, MD; WOBOK, M; WILDE, J; FUCHS, S; KOENIGER, N (2007) Apis mellifera mellifera in eastern Europe - morphometric variation and determination of its range limits. Apidologie 38: 191-197. DOI: 10.1051/apido:2006068

MORITZ, R F A; HÄRTEL, S; NEUMANN, P (2005) Global invasions of the western honey bee (Apis mellifera) and the consequences for biodiversity. Ecocene 12: 289-301. DOI: 10.2980/11195-6860-12-3-289.1


ÖZDİL, F; YILDIZ, M A; HALL, H G (2009) Molecular characterization of Turkish honey bee populations (Apis mellifera) inferred from mitochondrial DNA RFLP and sequence results. Apidologie 40: 570–576. DOI: 10.1051/apido/2009032


WILDE, J; BRATKOWSKI, J; SIUDA, M (2002a). Defensive behaviour of three breeds of *Apis mellifera* L. In 5th International Conference on the Black Bee *Apis mellifera* mellifera. 2-6 September 2002 Wierza, Poland. Reports and summaries. pp 85-90.


WILDE, J; SIUDA, M; RYKOWSKI, D; BRATKOWSKI, J (2002b) Flight activity of three subspecies of honey bee *Apis mellifera* depending on time of day and air temperature. In 5th International Conference on the Black Bee *Apis mellifera* mellifera. 2-6 September 2002 Wierza, Poland. Reports and summaries. pp 91-97.