Introduction

Human activities have been shaping the distribution of honey bee subspecies in Europe. In fact, during the last decades there has been an extensive introduction of the beekeepers’ favorite eastern European (lineage C) subspecies *A. m. ligustica* (Italian) and *A. m. carnica* (carniolan) into western Europe. Because of these introductions, there has been gene flow, and even replacement, of the native western European subspecies *A. m. mellifera* (black honey bee), which belongs to lineage M. Assessing levels of introgression is an important activity in breeding programs, especially when conservation of native subspecies is a major concern. Previous surveys of the *A. m. mellifera* populations estimated the introgression of lineage C into lineage M honey bees by using mtDNA and microsatellite markers (STR). Others markers, such as SNPs, have some advantages as they provide a genome wide coverage, higher quality data, and at the same time they are suitable for automatic and standardization in high throughput technologies. Previous studies indicate that the discriminatory power of SNPs to detect population structure is lower than microsatellites; about 100 SNPs are needed to provide the same power of 10-20 microsatellites. In this study we will compare introgression levels between microsatellites and SNPs in a black honey bee collection originating from several countries across western Europe.

Samples

A total of 77 *A. m. mellifera* individuals from France (18), Denmark (10), the Netherlands (15), Switzerland (6), Scotland (10), Norway (10) and England (8) were collected. Samples of *A. m. carnica* (12) from Croatia and Serbia, and *A. m. ligustica* (17), from Italy, were included as reference populations of C-lineage.

Genotyping

The 113 individuals were genotyped using SNPs and microsatellite. The 1536 SNP loci were scored using Illumina’s BeadArray Technology and the Illumina microsatellite. The 113 individuals were genotyped using SNPs and microsatellite. The positions was ascertained using the Map Viewer tool available in NCBI. Genomic position was ascertained using the Map Viewer tool available in NCBI.

The introgression levels were first estimated using both full datasets. Then, to have similar discriminatory power between the 12 microsatellite and SNPs, we used between 60 and 120 SNP loci from the initial 1183 SNP dataset by selecting SNPs located nearby the microsatellite loci.

Data sets

Of the 1536 a total of 1183 SNPs were available for analysis after removing monomorphic loci (cutoff 2%) and non-calls.

To obtain genomic position for both kind of markers, sequences were mapped to the honey bee Assembly 4.5 using BLAST in NCBI. Genomic position was ascertained using the Map Viewer tool available in NCBI.

The introgression levels were first estimated using both full datasets. Then, to have similar discriminatory power between the 12 microsatellite and SNPs, we used between 60 and 120 SNP loci from the initial 1183 SNP dataset by selecting SNPs located nearby the microsatellite loci.

Mean deviation from the Q value for the 7 *A. m. mellifera* populations

When analyzing the mean Q for the 7 *A. m. mellifera* populations for different datasets, it is possible observe that the highest differences are between the 12 STR and 1183 SNPs datasets. The populations that show higher differences are France, Denmark and Switzerland. While the 12 STR dataset has higher Q values for France, the 1183 SNPs dataset has higher values for Denmark and Switzerland populations (Fig. 4).

Results and Discussion

STRUCTURE analysis performed using different dataset, show that introgression level are variable across the 77 individuals sampled in the black honey bee race (Fig. 1).

The populations that show individuals with more differentiated membership proportion between datasets are France and England (Fig 1; Table 2).

The outside membership proportion (Q) in the yellow cluster is very variable dataset to dataset. While in the 12 STR dataset the membership proportion vary between 0.009 and 0.951 in the 1183 SNP dataset vary between 0.001 and 0.6902 (Table 1).

Mean deviation from the Q value for the 7 *A. m. mellifera* populations

Taking the 12STR dataset as reference, deviation between the datasets was calculate and the dataset that had the highest deviation was the 1183 SNPs (0.08), while the 60 SNPs dataset display the lowest value (0.076) (Fig. 2).

While over than 42 individuals for each dataset have deviation values lower than 0.05, there are some individuals that have deviation values of higher deviation than 0.20 (Table 2). The maximum deviation was found for an individual from England that have a deviation over 0.57  for all dataset (Table 2).

Conclusion

Results obtained by microsatellites are different from the results obtained by SNPs.

While there are more than 50% of individuals that have a deviation lower than 0.05, there are over 10 individuals that have values of deviation higher than 0.20, with an individual from England with a deviation higher than 0.5.

In spite of the differences found between all datasets, the 12STR are more similar with 120 SNP and 60 SNPs than with 1183 SNP dataset.