In partnership with

RNSBB 2024 Spring-Workshop

Bilbao, 13 March 2024

Proceedings

Edited by Melanie Parejo and Cecilia Costa

Re-examining DNA microsatellites with Next Generation Sequencing

Mahsa Vafadar (1) Jamal Momeni (2) Karina Kalmayer Jaquet (2) Cecilia Costa (3) Per Kryger (4)
Author Affiliation
1) University of Teheran; Iran; https://orcid.org/0009-0003-1685-5847 2) Eurofins Genomics A/S Denmark 3) CREA Agricoltura ed Ambiente; Italy; https://orcid.org/0000-0001-9985-2729 4) Aarhus University; Denmark; https://orcid.org/0000-0001-5306-3230

Abstract
Molecular biology has enhanced our understanding of honey bee biology. However; despite the shift to genome-wide association studies; progress in marker-assisted selection only is emerging slowly. Studying genetic diversity with diallelic SNP loci is challenging; requiring the use of multiple closely linked SNPs and complex bioinformatics methods. In this study; our objective is to compare next-generation sequencing (NGS) of pooled DNA from over 90 individual bees with the results derived from individual bees' microsatellite loci scored using PCR and electrophoresis. Two populations were chosen; *A. m. mellifera* bees from Læsø and *A. m. ligustica* bees from Italy; since these two subspecies are known to have rather distinct evolutionary history. With NGS we explored variations introduced by single nucleotide variants and interrupted repeats in the sequences. These factors amend the observable genetic diversity beyond what is identified through PCR and electrophoresis. The results obtained give confidence in our datasets of yesteryear. We have uncovered some additional single nucleotide variation that could not be discerned simply by the length variation of PCR products. For example; analyzing the data from the A008 microsatellite; it is evident that genetic variation between *A. m. mellifera* and *A. m. ligustica* are substantial. At the locus A008 locus (Solignac et al. 2003) in the NCBI-Genbank a sequence of 181 bp appears. In our A. m ligustica sequence data we find that allele 177 bp is the most common allele; whereas in *A. m. mellifera*; it is a rarity. The PCR electrophoresis data find allele 179 bp in *A. m. ligustica* with the highest frequency; whereas in *A. m. mellifera;* it is absent with PCR; but occurs rarely in the pooled sequence. The comparison of PCR and sequencing results confirms that a length of 165 bp is the shortest and predominant allele in *A. m mellifera.* Additionally; within the sequencing results of both subspecies; for allele 165; appears a single nucleotide variant (c/t) that were undetectable through electrophoresis. The outcomes of the A113 locus for sequencing and PCR methods; the two subspecies appear similar. The allele with the highest frequency; aligning with a length of 220 bp; was determined through pool sequencing for both *A. m. mellifera and A. m. ligustic*a. The same allele is the most common in the electrophoresis results. The A113 locus was already used by Estoup et al. (1995) as an example of size homoplasy; the occurrence of identical allele length; in spite of sequence differences; due to the imperfect nature of the repeat elements. Our observations confirm the existence of these variant types. Relaunching DNA microsatellites for genotyping through NGS seems feasible. Solignac et al. 2003 https://doi.org/10.1046/j.1471-8286.2003.00436.x Estoup et al. 1995 https://doi.org/10.1093/oxfordjournals.molbev.a040282