WROCŁAW DOCTORAL SCHOOL
OF INSTITUTES OF POLISH ACADEMY OF SCIENCES

Allelic Diversity and Functional Heterozygosity of the Honey Bee (*Apis mellifera*) Complementary Sex Determiner Gene

Robert Mroczek

Supervisor:
dr hab. Małgorzata Cebrat

This doctoral dissertation is based on experimental work performed in the Laboratory of Molecular and Cellular Immunology, Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences

Wrocław 2025

Abstract

In haplodiploid species like honey bees, sex is determined genetically, with males developing from unfertilized haploid eggs and females from fertilized diploid eggs. A key genetic mechanism in honey bees is Single Locus Complementary Sex Determination (sl-CSD), where sex is dictated by heterozygosity at the complementary sex determiner (*csd*) gene locus. Heterozygous individuals develop as females, while hemizygous into males. Homozygous diploids develop into sterile diploid males, which introduces a genetic load, especially in inbred or small populations, which makes a genetic diversity one of key features for population survival.

The study introduces a novel genotyping method for analysing *csd* gene diversity in *Apis mellifera* colonies by targeting worker bees rather than drones. The worker-based method captures greater allelic diversity but requires large sample sizes to reduce non-sampling error and ensure rare patrilines are represented. To increase throughput, method based on Terminal Restriction Fragment Length Polymorphism (T-RFLP) was developed.

Established genotyping method allowed for analysis of 2668 worker bee larvae revealing 137 unique *csd* alleles in samples from 19 honey bee colonies across three distinct locations: HIRSZ apiary in Wrocław city, ASP apiary in the city center of Wrocław, and PISZ apiary in the rural region of Piszkawa. Functional analysis of *csd* alleles pairings was aimed at detections of genotypes that are underrepresented, potentially due to lethality or partial non-functionality in determining femaleness. 118 unique allele pairings were identified, most of which were equally represented. However, two genotypes with minimal sequence divergence (1 and 4 amino acids) were experimentally confirmed to be functional based on femalespecific *dsx* transcript expression. This discovery provides new insights into criteria of functional heterozygosity between *csd* alleles.

To explore *csd* function at the protein level, AlphaFold2 structure predictions showed that the N-terminal and arginin-serin Rich (RS) domains consistently form α-helices, while the Hyper Variable Region (HVR) and Proline rich Region (PR) remain largely unstructured in monomers. Homotrimer formation predictions revealed significant structural rearrangements, including β-strand formation and increased hydrogen bonding, particularly in the HVR. In contrast, heterotrimers generally had fewer hydrogen bonds in this region. Interestingly, two exceptional alleles formed heterotrimers with higher bonding than their homotrimers, indicating potential allele-specific mechanisms of functional heterozygosity.

In parallel, *csd* HVR mutability using deep sequencing of drone sperm DNA was investigated. Illumina sequencing revealed high rates of PCR-induced artifacts, limiting the detection of low-frequency variants. In contrast, Oxford Nanopore sequencing with minimal PCR amplification provided clearer profiles, with most reads aligning closely to reference alleles. Using DADA2 algorithm a few rare Amplicon Sequence Variants (ASVs) that could arise during replication error were identified.