MADis: Genomic Analysis Tool for the Revelation of Multiple Alleles Within a Single Gene

Kaňovská I1, Biová J1, Bilyeu K2, Škrabišová M1

*maria.skrabisova@upol.cz*

1 Department of Biochemistry, Faculty of Science, Palacký University, Olomouc, Czech Republic

2 United States Department of Agriculture-Agricultural Research Service, Plant Genetics Research Unit, Columbia, MO, USA

Genome-to-phenome research in agriculture aims to improve and accelerate crop breeding. Genome-wide association study (GWAS) has been widely used to identify genomic loci underlying important traits. Numerous post-GWAS analyses were developed to narrow down the associated genomic regions. However, in many cases, they are still unable to identify candidate genes or causative mutations (CMs). Natural and artificial selection alter allele frequencies in genes controlling phenotypes. This raises the likelihood that numerous alleles with independent CMs would be present in a single gene, which poses an issue for GWAS analyses.

As of yet, there has been no association method addressing the issue of multiple alleles. Therefore, we developed a tool that computes a score for a combination of variant positions in a single candidate gene and based on the highest score identifies the best number and combination of CMs. The tool is available as a Python package on GitHub and as a web-based Multiple Alleles discovery (MADis) tool specifically designed for soybeans [1].

We demonstrated the utilization of MADis on an example of a recently cloned gene *L1* [2]. We successfully identified a candidate gene for the soybean pod color *L2* locus and predicted the existence of multiple alleles that potentially cause loss of pod pigmentation in soybeans. In this work, we demonstrated the use of genomic analysis in the exploration of the natural and artificial selection of multiple alleles. The MADis tool can be applied to other species to aid in the discovery of genes under selection for improved breeding.

***References:***

[1] Biová et al., 2024. Front. Genet. 14:1320652.

[2] Lyu et al., 2023. Mol. Plant 16(7), 1178-1191.