# Consensus-Based Detection of Biosynthetic Gene Clusters with Application to RiPPs from Antarctic Bacteria

Vítková Helena1, Jakubíčková Markéta1, Bezdíček Králová Stanislava2

1 Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic

2 Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Brno, Czech Republic

Bacterial secondary metabolites represent a rich source of bioactive compounds, yet much of their biosynthetic potential remains unexplored. One particularly promising group are ribosomally synthesized and post-translationally modified peptides (RiPPs), which form a structurally diverse and pharmaceutically promising class of natural products. Therefore, our study focuses on uncovering novel RiPP biosynthetic gene clusters (BGCs) from extremophilic bacteria inhabiting Antarctic environments, which offer access to unique and largely uncharacterized microbial diversity. A new computational pipeline is being developed to enable robust and comprehensive RiPP detection. The proposed pipeline integrates several specialized tools for detecting RiPPs specifically and BCGs in general, including DeepRiPP, RRE-Finder, decRiPPter, antiSMASH, and DeepBGC. Selected tools employ various algorithmic approaches, primarily utilizing machine learning methods such as bidirectional long short-term memory recurrent neural networks, support vector machines, or hidden Markov models, resulting in varying outputs. These outputs are then combined through a consensus-based approach to generate unified predictions, improving the accuracy and reliability of RiPP identification. A key part of our workflow is the use of long-read sequencing to enable high-quality reconstruction of microbial genomes directly from complex metagenomic samples. This strategy allows us to recover biosynthetic pathways from previously inaccessible and uncultivated microbial lineages, thus broadening the scope for natural product discovery. Identified candidate clusters will undergo comparative genomics and functional annotation to prioritize targets for downstream characterization and synthetic biology applications. Our approach aims to expand the known chemical space of RiPPs and provide a robust resource for future biotechnological exploitation.

Acknowledgement: This work was supported by GA25-17343S.