A system-level view on the function of natural eukaryotic biomes through taxonomically resolved metabolic pathway profiling

JAYSON GUTIERREZ[1], PASCAL HABLÜTZEL[1]

[1] *Flanders Marine Institute (VLIZ), InnovOcean Site, Oostende, Belgium*

High-throughput sequencing of environmental samples has dramatically improved our understanding of the molecular activities of complex microbial communities. For instance, by enabling taxonomic profiling and differential gene expression analysis, microbiome studies have revealed intriguing associations between community structure and ecosystem functions. In this work, we conducted a systems-level interrogation of environmental samples, which can effectively augment the insights obtained through traditional gene-centric analysis. To achieve this, we utilized the popular HUMAnN pipeline, which has proven effective at delineating taxon-specific metabolic pathways actively contributing to the functioning of a microbiome. However, the effectiveness of sequence data analysis to characterize the functioning of microbial ecosystems at the systems level has thus far been limited by the quality and scope of reference sequence databases, which are optimized for widely studied prokaryotic microbiomes. To overcome this limitation, we leveraged publicly available genome/gene sequences for a wide array of (mostly eukaryotic) planktonic organisms to build a customized protein sequence database. To test the efficacy of our database customization, we reanalyzed previously published metatranscriptome datasets derived from different marine environments. We found that database customization can substantially improve our ability to assess core metabolic processes across taxonomically diverse eukaryotic microbiomes, which have remained largely uncharacterized at the systems level. By further expanding on the taxonomic and functional complexity of our database with newly released genomes and gene catalogs, we aim to improve our ability to map the molecular traits that drive changes in the composition and functioning of marine planktonic networks through space and time.