

Title: A Bioinformatic Pipeline to Identify and Classify Potential Microbial Signatures from Fungal Sequencing Data Integrated with a Searchable Database of Described Interactions

Authors: Michal Babinski^{1*} (mbabinski@lanl.gov), Aaron J. Robinson^{1*} (arobin@lanl.gov), Karen W. Davenport,¹ and **Patrick S. Chain¹**

Institutions: ¹Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico

Website URL: <https://www.lanl.gov/science-innovation/science-programs/office-of-science-programs/biological-environmental-research/sfa-bacteria-fungal.php>

Project Goals: The vast taxonomic diversity and the complexity of interactions within the soil microbiome present many challenges. Many of the interactions between soil-dwelling bacteria and fungi are not yet well-understood, and a more comprehensive understanding of these relationships and their response to environmental pressures would lead to substantial agricultural, environmental, and energy-focused advancements. These potential developments align with the foci of the DOE, and would influence multiple scientific disciplines. The aim of this Science Focus Area (SFA) is to better understand the diverse and abundant interactions within the soil rhizosphere, specifically between fungi and bacteria, and decipher the mechanisms behind their communication.

Abstract Text: Bacteria and fungi are dominant microorganisms in soil and terrestrial systems and are responsible for an array of ecological functions. Recent research by our group demonstrated that interactions between bacteria and fungi are more widespread and diverse than previously anticipated¹. Bacterial associations with fungi are largely discovered through sequencing-based approaches, either amplification of 16S sequences directly from fungal isolates or the capture of bacterial sequences during sequencing of fungal genomes. Currently, there exist no standardized bioinformatic pipelines or tools to identify or classify potential BFI from diverse fungal sequencing datasets. In this presentation we will outline our efforts to develop a bioinformatic pipeline to detect and analyze potential bacterial associations from genomic and transcriptomic fungal sequencing datasets. Additionally, we will describe our efforts to expand the capability of this pipeline to identify other organisms and components which comprise the microbiomes of fungal isolates, such as viruses, archaea and even other fungi. After taxonomic classification, binned sequences will also undergo assembly and annotation through sub-pipelines that are optimized for each individual kingdom. The outputs of this pipeline include concise summaries of the diversity of potential associations detected in the datasets, relevant statistics and metrics of detection and access to all sequences assembled or annotated during the pipeline. While this pipeline can be used to examine data from a single or small number of fungal sequencing projects, we have utilized it to screen all publicly available fungal genomic and transcriptomic datasets available through the NCBI SRA database. This resulted in optimization of the pipeline to allow for examination of large datasets in a reasonable time and manner. Furthermore, we have also developed the first comprehensive database of previously

described or observed BFI, which will be integrated with the pipeline to allow users to compare their results to previous knowledge and allow for the curated addition of newly discovered associations. Both the pipeline and database were designed in a manner that allows for smooth integration to platforms frequently used for microbial and microbiome research such as KBase and the NMDC. We believe this pipeline will provide a valuable resource not only for researchers in the BFI field, but for the larger fields of microbiology and microbial ecology through the provided means to assess potential impacts of microbial interactions in future or previously obtained datasets.

References/Publications

1. Robinson, A.J., House, G.L., Morales, D.P., Kelliher, J.M., Gallegos-Graves, L.V., LeBrun, E.S., Davenport, K.W., Palmieri, F., Lohberger, A., Bregnard, D. and Estoppey, A., 2021. Widespread bacterial diversity within the bacteriome of fungi. *Communications biology*, 4(1), pp.1-13.

Funding Statement: This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Division, under award number LANLF59T.