

Using Data Mining and High-Throughput Cultivation Methods to Better Understand Bacterial:Fungal Interactions

Geoffrey L. House^{1*} (ghouse@lanl.gov), Armand E. K. Dichosa¹, Debora F. Rodrigues³, Hang N. Nguyen³, Saskia Bindschedler², Jean F. Challacombe¹, Jamey D. Young⁴, Pilar Junier², and **Patrick S. G. Chain¹**

¹Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico; ²Institute of Biology, University of Neuchâtel, Neuchâtel Switzerland; ³Civil and Environmental Engineering, University of Houston, Houston, Texas; ⁴School of Engineering Vanderbilt University, Nashville, Tennessee

Interactions between bacteria and fungi are important determinants of ecosystem function, yet little is known about these interactions or how they operate. This is a critical knowledge gap as these interactions are important in addressing multiple DOE priorities including developing renewable energy sources, understanding the possible effects of Earth system change, and understanding how these interactions may help overcome energy and environmental challenges. Here we outline a range of research questions that we are beginning to address through a new SFA in order to better understand the diversity and function of these bacterial:fungal interactions. Using bioinformatics-based data mining of existing fungal genome sequencing data coupled with single cell isolation, microfluidics, and cultivation techniques, we are beginning to understand the diversity of bacteria that form associations with fungi, and how these associations affect both fungal and bacterial phenotypes.

There is an increasing amount of fungal genome and metagenome sequencing information available in public repositories such as JGI's MycoCosm and NCBI's sequence read archive (SRA), driven in part by the 1000 Fungal Genomes project. We have started to mine this genome sequencing data, beginning with raw sequence data obtained from fungal isolates. Our goal is to screen these sequences for any that appear to have bacterial origins. We propose to screen both assemblies and raw data since reads and/or contigs with bacterial signatures may have been discarded from the original fungal genome assembly. By using this approach across the phylogenetically broad sampling represented in the 1000 Fungal Genomes project, we will determine: 1) the range of bacterial diversity found across sequenced fungi, and 2) how this diversity is organized, (e.g. whether specific bacterial clades are general associates with many fungi or whether they associate specifically with particular fungal groups). After analyzing the genomic data of fungal isolates, we will take a similar approach to analyze metagenome samples from soil where both the fungi and the bacteria present in the sample are unknown. Because this latter step involves more complex samples and larger amounts of data, we will leverage an existing funded Exascale computing partnership between LANL, JGI, and LBNL. A critical capability for this work is the ability to better determine whether identified bacterial sequences represent bacterial contamination or true associations between bacteria and fungi. To make this determination as confidently as possible, we will use a combination of bioinformatics and experimental approaches. We will identify bacterial candidates for these associations by looking for strong phylogenetic associations between specific bacterial and fungal groups, as well as identifying any bacterial taxa that are recurrently encountered during the genome sequencing of

multiple fungi. We will then use targeted field soil collections and single cell cultivation techniques to better determine whether these identified candidate bacteria do form associations with fungi or whether they may have been contaminants in the genome sequencing.

In order to better characterize the growth responses of both partners in bacterial:fungal interactions, we will utilize a novel method, high-throughput screening of cell-to-cell interactions (Hi-SCI). Hi-SCI integrates the co-cultivation of bacteria and fungi within individual gel microdroplets (GMDs) with the efficient screening of the millions of different bacterial:fungal interactions represented within the GMDs. The GMDs allow either the random or the targeted capture of bacterial and fungal partners, and millions of GMDs can be grown simultaneously, each within its own cultivation chamber on a chip. The GMDs can then be rapidly screened and sorted using flow cytometry based on morphological or physiological signals, and then used for downstream genomics analysis. We have begun capturing fungal spores in GMDs to measure their speed of germination and other physical characteristics. We will then begin co-capturing fungal spores together with various known bacterial associates of fungi to determine how they affect fungal growth both alone and in different combinations.

The information gained about the diversity of bacteria associating with fungi and the information about the growth responses of fungi and bacteria when grown in combination within the GMDs can be tightly integrated. Better understanding the bacterial diversity associated with fungi can help guide the targeted co-capture of specific bacteria with specific fungi to access their growth responses. Conversely, better understanding of morphological and physiological interactions using GMD cultivation can guide molecular screens for specific functions or possible interactions between bacterial and fungal genomes.

We will use the results of the data mining approach to help inform the next experimental phases of this SFA aimed at better understanding bacterial:fungal interactions and their function in a range of growth conditions, and to interrogate gene, protein, and metabolite interactions. Together, these measurements and approaches will help give a better understanding of the complex interactions between bacteria and fungi in the soil, and how these interactions may exert important effects on plant growth and other important ecosystem services.

This SFA is supported under the Computational Biosciences Program of the Office of Biological and Environmental Research in the DOE Office of Science.

LA-UR-18-20257