

Design and -omics exploration of synthetic communities

Alexis L. Marsh^{1*} (almarsh@iastate.edu), Jia Wang,² Dale A. Pelletier,² Myra B. Cohen,¹ Paul E. Abraham,² David J. Weston,² Christopher S. Henry,³ Janaka N. Edirisinghe,³ Paramvir S. Dehal,⁴ Robert Cottingham,² Mitchel J. Doktycz,² and **Priya Ranjan**²

¹Iowa State University, Ames, IA; ²Oak Ridge National Laboratory, Oak Ridge, TN; ³Argonne National Laboratory, Lemont, IL; and ⁴Lawrence Berkeley National Laboratory, Berkeley, CA

<https://pmiweb.ornl.gov/> ; <https://kbase.us>

Project Goals: Synthetic communities (SynComs) offer a key experimental platform for examining how environmental changes affect the structure of microbiome and host physiology and productivity. As part of this effort, we propose adding new functionality and workflows to KBase, in support of the Plant-Microbe Interfaces SFA at Oak Ridge National Laboratory. With these new tools in KBase, microbes can be identified with beneficial properties that have the ability to coexist as communities. We propose new KBase Apps to simplify various steps of microbial community design for SynCom experiments. Additionally, we will add functionality to help validate SynCom experiments.

Synthetic communities (SynComs) can be used to obtain a deeper understanding of specific mechanisms that drive community assembly, how multiple biotic interactions take place, and the expression of functional properties that occur when microbes, plants, and the environment interact simultaneously in time and space. Furthermore, the capability to obtain correlated data on transcript, protein, and metabolite abundance from a designed synthetic community experiment facilitates linking functional processes to molecular level information and helps improve the quality of annotation and predictive models. For Plant-Microbe Interfaces Science Focus Area (PMI SFA) experiments, SynComs are crucial for taking advantage of the extensive set of plant and microbial resources assembled in order to further investigate and understand in more precise detail the productive relationship between a plant and its microbiome.

As experimental data and design complexity grow, computational modules and applications are required to support rational design of SynComs based on desired functional and phenotypic characteristics and the ability of members of the community to survive together. A more comprehensive interpretation of the resulting experimental -omics data is needed to accelerate the design, build, test, and learn cycles associated with SynCom experiments. KBase, a DOE Genomic Sciences Program funded, public, and freely accessible software and data science platform, is ideal for developing such applications as it already provides a wide range and increasing number of relevant tools for functional annotation, metabolic modeling, auxotrophy prediction, substrate utilization and production of by products, taxonomy information, and prediction of microbial traits (*e.g.*, nitrogen fixation) that are important considerations in the rational design of SynComs and its members.

We propose to add *Populus*-associated microbiome and isolate genome data generated as part of PMI SFA, and Apps to KBase that leverage existing KBase capabilities for rational SynCom design. The -omics data resulting from the designed experiments will be integrated back into KBase and stored as abundance matrices. The four Apps will: **1) Summarize functional**

potential of genome set, 2) Generate synthetic communities functionally similar to a metagenome, 3) Design synthetic microbial communities with desired characteristics, and 4) Compare groups of genomes from SynCom experiments. The resulting workflow will advance a more productive SynCom research effort by accelerating the design, build, test, and learn cycle in the PMI SFA, and serve the growing community that leverages this powerful research platform.

Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKPA39.