Metabolic and Community Modeling of Phototroph-Heterotroph Community Interactions

William Cannon[1]* (william.cannon@pnl.gov), Gretta Serres[2], Sebastian Jaramillo[1], and Don Bryant [3,4]

[1] Computational Biology and Bioinformatics Group, Pacific Northwest National Laboratory, Richland, WA 99352; [2] Marine Biological Laboratory, Woods Hole, MA 02543; [3] Dept. of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802; [4] Dept. of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717.

Project Goals:

Understanding interactions between microbes in the environment is a daunting task. Because of the complexity of such interactions, it is necessary to use computational models to track the data and proposed interactions, and to develop intuition, hypotheses and predictions. With this in mind, the goals of the modeling component of the PNNL FSFA are two-fold. First, in the short-term we will evaluate the existing metabolic model of *Synechococcus* sp. PCC 7002 by incorporating metabolomic and proteomic data. We will initially utilize the model to evaluate whether the newly discovered complete but alternative cyanobacterial TCA cycle maximizes the utilization of 2-oxoglutarate for production of biosynthetic precursors for growth, and to evaluate the growth and production of extracellular metabolites. Second, in the long term, we will more generally identify the key metabolic interactions between the autotrophic and heterotrophic populations of Yellowstone hot spring mats and consolidate the quantitative data describing these interactions within a context of extended phototroph–heterotroph model, including metabolite predictions.

Abstract

Building useful models requires several tasks. First, initial models are constructed based on genome sequences. These initial models only tell us about metabolic potential, rather than actual metabolism because genome annotation is not complete and regulatory information is missing. Enzymes and transporters that carry out reactions of interest may not be expressed under all conditions, predicted metabolites may not actually be produced, and more frequently unpredicted metabolites are observed. Thus a second critical aspect to building useful models is to have data analysis tools that analyze the data thoroughly. Multiple data types – proteomics, metabolomics and transcriptomics to name a few – may then be interpreted with respect to the model. Based on the data, the model is adjusted. This is the process frequently referred to as data integration. Finally, the model is curated to the point at which simulations can be carried out, the output of which can be directly compared to experimental data.

To enable process of model building, the PNNL SFA has developed 11 pathway genome databases of phototrophic and heterotrophic species being studied on this project. These databases are linked to external pathway databases of the same

organisms for comparative studies. Using this platform, experimental data can be interpreted and used to refine the model.

Using this approach, we evaluated the protein complement of *Synechococcus* spp. 7002 grown under conditions of phototrophic growth under atmospheric CO_2 concentrations. Not surprisingly, pathways related to photosynthesis and carbon fixation were identified most frequently. However, initial analysis implied that a specific ICT family CO_2 transporter, SYNPCC7002_A0690, was not expressed, implying that either the annotation was incorrect or an alternative transporter is used. Since transporters set the boundary conditions that determine the energetic feasibility of metabolic processes, such adjustments to the metabolic model are critical. However, a new, highly accurate computational method that we applied found that the ICT transporter SYNPCC7002_A0690 was not only expressed but appears to be one of the most highly expressed proteins in the cell [1].

Pathway databases also serve as sources of hypotheses. Using pathway databases, we have begun to investigate pathways that have been inferred to be incomplete by computational methods. Investigation into the incomplete TCA cycle of cyanobacteria has demonstrated that, in fact, the cycle is complete and uses an alternative enzyme for the conversion of 2-oxoglutarate to succinate [2]. A principle difference between the standard pathway and the newly discovered alternative pathway is that the formation of succinate is no longer coupled to the highly favorable reaction involving coenzyme-A. Due to the removal of this highly favorable step, we hypothesize that the alternative pathway leads to an increased steady-state level of 2-oxoglutarate. This could be significant in that 2-oxoglutarate is the precusor used for assimilation of nitrogen, and the *Synechococcus* species in the Yellowstone mat system have been observed to dramatically increase nitrogen fixation during a brief 2-hour period in the morning. This burst of activity may serve as a nitrogen source for other members of the microbial community, and yet it is not clear how such a dramatic increase in nitrogen fixation can occur in such a short time period. Both laboratory and simulation based tests of these hypotheses are being conducted.

The simulation model starts with the framework of the previously determined genome-scale and flux-based model of *Synechococcus* sp. PCC 7002. We will model the TCA cycle in kinetic detail by incorporating aspects of the kinetic model by Singh and Ghosh. This will allow us to predict 2-oxoglutarate levels for both the standard and alternative TCA cycles under multiple conditions.

Funding: This research is supported by the Genomic Science Program (GSP), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Foundational Scientific Focus Area.

[1] W. R. Cannon, M. M. Rawlins, D. J. Baxter, S. J. Callister, M. S. Lipton, and D. A. Bryant, "Large Improvements in MS/MS-Based Peptide Identification Rates using a Hybrid Analysis," *J Proteome Res,* Mar 30 2011.

[2] S. Zhang and D. A. Bryant, "The Tricarboxylic Acid Cycle in Cyanobacteria," *Science,* vol. 334, pp. 1551-1553,2011. http://www.sciencemag.org/content/334/6062/1551.abstract