High-Throughput Determination of a Subcellular Metabolic Network Map of Plants

Kevin Radja[1], Suryatapa Ghosh Jha[1], Tallyta Silva[2], Charles Hawkins[1], Angela Xu[1], Bo Xue[1], Christine Aquino[1], Edward J. Wolfrum[3], Markita P. Landry[4], David W. Ehrhardt[1], Jenny C. Mortimer[2], **Seung Y. Rhee[1]*** (srhee@carnegiescience.edu)

[1]Carnegie Institution for Science, Stanford, CA; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; [3]National Renewable Energy Laboratory, Golden, CO; [4]University of California, Berkeley, CA

Project Goals: The goal of this project is to build an integrated pipeline to characterize metabolic interactions and pathways at a cellular level, using a combination of computational prediction, metabolic network modeling, and high-throughput experimental testing. This pipeline will be divided into three stages in order to develop a high-resolution subcellular map of small molecule metabolism in Sorghum and Brachypodium: a) generating localization predictions using bioinformatic algorithms, b) testing those predictions using nanotechnology mediated transformation of fluorescently tagged target proteins and high-sensitivity confocal imaging, and c) using the experimental data to generate new compartmentalized metabolic network models as well as refining existing pathway models. This project will initiate the creation of a repository for subcellular locations of metabolic enzymes, yielding important insight into the structure and function of metabolic networks in model systems as well as economically important crop species.

Advances in our understanding of plant metabolism have underpinned many traits that contribute towards improving plant productivity. To identify (by predictive modeling and experimentation) and engineer desirable metabolic traits, such as maximizing biomass production under suboptimal conditions or reallocation of biomass from carbohydrates to lipids, we must decode the complex metabolic networks. Subcellular compartmentation of metabolic reactions through the locations of enzymes is critical to understanding, modeling, and engineering plant metabolism. Yet, the localization of the majority of the predicted enzymes are not yet known. The paucity of experimentally validated information in most plants, especially in the DOE flagship bioenergy plants, severely limits scientists and engineers to assess the performance and translatability of computational tools and resources.

In this new project, a trans-disciplinary team with expertise in plant cell biology, genomics, metabolic modeling, algorithm development, synthetic biology, geochemistry, nanotechnology, and analytical chemistry will develop an integrated pipeline that combines computational prediction, metabolic network modeling, and high-throughput experimental testing using state of the art technologies in live confocal imaging, nanomaterial-mediated plant transformation with target metabolic enzymes, and metabolic network modeling. Using the pipeline, the team will create a high-quality subcellular map of small molecule metabolism as well as accurately compartmentalized metabolic network models in Sorghum and Brachypodium. The models will be experimentally validated by measuring a series of outputs in response to environmental challenges, and by knocking out gene expression in somatic tissue using CRISPR technology. Here, we will outline our approach to this project, as well as some initial progress on transient sorghum transformation and data analysis.

The pipeline will rely on the existing metabolic pathway database SorghumBicolorCyc. This database was created by using the E2P2 software to predict enzymatic function of the proteins in the sorghum genome

sequence, then using the Pathologic and SAVI software to call the presence in sorghum of metabolic pathways from the Metacyc database. The pipeline implements a novel network-based classifier to infer compartmentalization of the Sorghum metabolic pathway database. The network-based classifier utilizes an existing classifier which predicts protein subcellular localization based on sequence information as training on the partially compartmentalized network.

Based on the predictions, target metabolic genes will be chosen to validate the localization in cells. To achieve this, Gateway and In-Fusion molecular cloning technology will be used to fuse these candidates with known and effective fluorescent proteins, and subsequently expressed in the plants using high-throughput nanomaterial-based plant transformation techniques. Utilizing high resolution confocal live imaging (EMCCD spinning disk and Leica HyD point scanning), the locations of these enzymes will be validated *in planta.* The dataset collected from these validations will be used for developing the network maps and refine current models.

Overall, this project aims to holistically decipher the complexity of plant metabolic networks in order to engineer pathways to tackle the impending problems of changing climate, food security and availability of sustainable energy sources.

This research is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program grant no. DE-SC0020366