Transforming our understanding of chloroplast-associated genes through comprehensive characterization of protein localizations and protein-protein interactions

Lianyong Wang, Princeton University
Bibin Paulose, Michigan State University
Danny J. Schnell, Michigan State University
Martin C. Jonikas, Princeton University

Our project aims to generate a map of protein localizations and protein-protein interactions for 4,262 genes associated with the chloroplast, the energy-producing organelle which is a hallmark of plants. We consider these genes a high priority set because of the organelle's central roles in photosynthesis, metabolism and intracellular signaling, all of which are targets of ongoing biofuels crop engineering efforts. Furthermore, chloroplast-associated genes are particularly underrepresented in existing systems-level datasets because most high-throughput studies to date were performed in model systems that lack chloroplasts. As demonstrated in yeast, protein localization and protein-protein interaction data transform our understanding of the genes under study by immediately generating specific hypotheses about the mechanism of action of their protein products.

This project consists of two synergistic elements, one in the unicellular model green alga *Chlamydomonas reinhardtii* and the other in the dedicated biofuels oilseed crop *Camelina sativa*. Objectives 1 and 2 are to generate a searchable online resource of protein localizations and protein-protein interactions for nearly all chloroplast-associated proteins. We will achieve these objectives by leveraging high-throughput protein tagging, microscopy and affinity purification-mass spectrometry in Chlamydomonas. Objective 3 is to illustrate the value of this resource to biofuel crops by validating high-priority localizations and protein-protein interactions in Camelina and by building on the newly generated knowledge to advance our understanding of protein interaction networks that impact yield and stress resistance.

The project is based on extensive preliminary data from the PI and Co-I demonstrating feasibility, quality and value of the work. Significant progress has already been made towards tagging and localizing many of the target proteins. The team has a strong track record of developing large-scale resources for the community and advancing our basic understanding of chloroplast biology.

This study will transform our understanding of photosynthetic eukaryotes and will open a broad range of engineering opportunities. The localization data has the potential to reveal novel classes of protein localization, providing insight into the sub-organellar structure of chloroplasts. The localization and protein-protein interaction data will provide key information on the functions of thousands of uncharacterized proteins, many of which have no recognizable protein motifs. The data will improve annotations for thousands of other genes. The project will also have a long-term impact as the scientific community utilizes the resource of strains, constructs and data.