Next-Generation Protein Interactomes for Plant Systems Biology and Biomass Feedstocks Research

Shelly A. Trigg[1,2,3*] (swanamaker@salk.edu), Renee Garza[1], Andrew MacWilliams[1], Joseph Nery[1], Joaquin Reina[1,4], Anna Bartlett[1], Rosa Castanon[1], Adeline Goubil[1], Joseph Feeney[1], Ronan O'Malley[1,2], Mary Galli[1,2] and **Joseph R. Ecker**[1,2,5]

[1]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla CA; [2]Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla CA; [3]Biological Sciences Department, University of California, San Diego; Bioinformatics Program, University of California, San Diego; and [5]Howard Hughes Medical Institute

Project Goals: **Elucidate interactome characteristics giving rise to desirable phenotypic traits in biofuel feedstock crops by rapidly identifying all their protein-protein interactions using the *en masse* next-gen yeast two-hybrid screening system, ProCREate.**

Biofuel crop cultivation is a necessary step in heading towards a sustainable future, making their genomic studies a priority. While technology platforms that currently exist for studying non-model crop species, like switch-grass or sorghum, have yielded large quantities of genomic and expression data, still a large gap exists between molecular mechanism and phenotype. The aspect of molecular activity at the level of protein-protein interactions has recently begun to bridge this gap, providing a more global perspective. Interactome analysis has defined more specific functional roles of proteins based on their interaction partners, neighborhoods, and other network features, making it possible to distinguish unique modules of immune response to different plant pathogens[1]. As we work towards cultivating heartier biofuel crops, interactome data will lead to uncovering crop-specific defense and development networks. However, the collection of protein interaction data has been limited to expensive, time-consuming, hard-to-scale assays that mostly require cloned ORF collections. For these reasons, we have successfully developed a highly scalable, economical, sensitive, and potentially quantitative yeast two-hybrid assay, ProCREate, that can be universally applied to generate proteome-wide primary interactome data. ProCREate enables *en masse* pooling and massively paralleled sequencing for the identification of interacting proteins by exploiting Cre-lox recombination. As a proof of principle, we thoroughly screened a well-validated collection of ~2000 *Arabidopsis* transcription factors in 10 replicate experiments, generating a comprehensive *Arabidopsis thaliana* transcription factor interaction network (AtTFIN1). After observing a high recall rate, a high overlap with literature, and rational novel interactions, we then tested ProCREate's scalability by generating and screening *Arabidopsis* cDNA libraries. Once these recently collected datasets are further validated, ProCREate will be used to screen ORF/cDNA libraries from feedstock plant tissues. The interactome data generated will yield deeper insight into many molecular processes and pathways that can be used to guide improvement of feedstock productivity and sustainability.

References
1. Jiang, Z., Dong, X., and Zhang, Z. *Nature Scientific Reports* **6**, 19149 (2016)