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

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

[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; and [4]Howard Hughes Medical Institute

Project Goals: Develop an *en masse* yeast two-hybrid screening system using next generation sequencing to rapidily identify all protein-protein interactions in biofuels feedstocks, and generate comprehensive interactome maps that represent the findings.

In order to keep up with global energy demands, it is imperitive we acquire more knowledge of biofuel feedstocks for improving their cultivation and energy yield. Knowledge of protein-protein interaction (PPI) networks that promote robust plant growth or that are perturbed by pathogens causing disease could progress strategies for improving cultivation. However, current technologies available for obtaining PPI data are insufficient and unrealistic for non-model organisms because of time, cost, and sensitivity constraints. Even the largest high quality PPI map for the model plant *Arabidopsis thaliana* (Arabidopsis Interactome 1 or AI-1), that we generated, contains only 2% of all potential interactions, and took upwards of 5 years and $8 million to finish[1]. To address this problem, we are developing a next-generation sequencing integrated yeast two-hybrid (Y2H) system that will greatly improve the rate at which PPI data can be obtained and will be applicable to virtually any cell from which RNA can be extracted.

Our system enables *en masse* pooling and massively paralleled sequencing for the identification of interacting proteins by exploiting Cre-lox recombination. Screening of Y2H plasmids containing mutant loxP sequences in a yeast strain expressing the reporter gene for Cre Recombinase has shown only interacting proteins can induce Cre-mediated recombination of plasmids. The newly formed double mutant loxP makes an irreversible linkage of each protein's corresponding coding sequence, and has allowed us to identify protein interactions using Illumina paired-end sequencing. Preliminary testing of a positive and random reference set (PRS/RRS) consisting of ~300 open reading frames (ORFs) in a 1 ORF vs 1 ORF screen has shown a sensitivity and reproducibility similar to our previous Y2H screen used to make AI-1. Our next generation sequencing and analysis pipeline has shown a 98% overlap with interactions detected by Sanger sequencing. *En masse* Y2H screening has been attempted on the PRS/RRS and also on a subset of ~4,500 Arabidopsis ORFs that we thoroughly assayed when creating AI-1. Since all expected interactions have not yet been identified because the inherently low percentage of informative DNA (< 0.003%), we are working to improve assay sensitivity by enriching for double-mutant-loxP-sequence-containing DNA fragments with an inverse PCR strategy and/or an in-solution RNA probe capture approach. To more stringently select for diploids containing only hybrid plasmid, we added an antibiotic resistance marker to the plasmids that becomes functional only when plasmids are Cre-recombined. Once we are confident that our assay is detecting all interacting pairs that occurred, we will move forward in testing ORF libraries from feedstocks by shotgun cloning into our Y2H plasmids and carrying out our *en masse* screening pipeline. The ability to rapidly construct large PPI networks will yield deeper insight into a variety of molecular processes and pathways that will potentially allow improvement of feedstock productivity and sustainability.

References

1. Arabidopsis Interactome Mapping Consortium, Science 333, 6042 (2011).

We gratefully acknowledge the U.S. DOE Office of Biological and Environmental Research in the U.S. DOE Office of Science for funding this project; DOE-DE-SC0007078.