Designing a computational system to predict protein-protein interactions in *Arabidopsis Thaliana*

Kamaldeep Singh,[1] Yanjun Qi,[2] and Judith Klein-Seetharaman[3]

[1]Kamaldeep Singh, School of Engineering and Computer Science, University of the Pacific, Stockton, CA 95211

[2]Yanjun, Qi, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

[3]Mentor: Dr. Judith Klein-Seetharaman, Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260

Proteins are unique amongst organic compounds in supporting every reaction occurring in biological systems. It is important to identify the structures and understand the dynamics and functions of proteins. Further, it becomes essential to understand the mechanism by which a set of proteins communicate and collaborate toward a common cellular function. There are several genomic and proteomic databases that maintain valuable information for model organisms. These databases can be used to ease and optimize the task of predicting protein-protein interactions. Moreover, various computational methods are utilized to identify interactions. Recent studies conducted by Yanjun Qi, Ziv Bar-Josepth, and Judith Klein-Seetharaman indicate that the Random Forest (RF) classifier performed the best in terms of accurately predicting interactions in yeast. The results from previous studies and online-databases such as The Arabidopsis Information Resource (TAIR) and Biomolecular Interaction Network Database (BIND) will be employed to construct a proteome of the model organism *Arabidopsis Thaliana*.