Novel quantitative proteomic methods to discover and localize protein complexes

Uma K Aryal1, 2, Yi Xiong3, Eileen Mallery1, Mark C. Hall2, Jun Xie4, Daisuke Kihara3,5, Peter N Baker6, Daniel B. Szymanski1,5

Departments of Agronomy 1, Biochemistry2, Computer Science3, Statistics4, Biology5, Cyber Center5, Purdue University, West Lafayette, IN 47907

Summary

Characterization of plant protein complexes provides useful insights about biological processes. This information is important for an efficient translation of basic knowledge into improved crop traits. This project aims to predict and validate composition of 100 protein complexes in Arabidopsis thaliana leaf cytosol, and their conservation in Glycine max. Here, we outline details of our approaches to characterize protein complexes and present some of our preliminary results. Protein complexes isolated under non-denaturing conditions are separated by multiple chromatography techniques, and identified by liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. Complexes are predicted based on cluster analysis, bioinformatics and existing “omics” data sets. We expect to demonstrate the feasibility of protein correlation profiling to solve protein complex localization and composition.

Methods

Computational strategies to predict protein complex composition

Golden standard data sets

Prediction methods

Benchmarked

Genome-based – Gene neighborhood 
- Phylogenetic profile

Sequence-based – Machine learning

Novel-based – Sequencing and template

Preliminary Results

Fig 7. Comparison of proteome between in solution (global) vs. gel-based fractionation.

Fig 8. Western blot and MS1 feature quantification using Progenesis LC-MS software. (A) Western blot of SEC fractions probed with an anti-PHCP antibody. (B) Topographical view of an aligned feature for PHCP peptide 4207 using Progenesis LC-MS software. Red polygon shows feature space for peptide 4207. Multiple isotopes of the peptide are detected. (C) Peptide quantification of PHCP2 peptide (score >15) across all SEC fractions. (D) Peptide quantification for 3 PHCP2 peptides (score >25). (E) Histogram of coefficient of variation (CV) of triplicate run of SEC fraction 6.

Fig 9. SEC fraction volume vs. the ratio of the apparent to the actual molecular weight (Rapp). The SEC fractions were analysed using LC-MS.

Fig 10. Relative protein abundance (A) and cross correlation analysis (B) of the proteins using spectral count data.

Preliminary Outcomes

- To develop an orthogonal set of chromatographic methods to reproducibly separate endogenous protein complexes.
- To create robust, sensitive, and quantitative mass spectrometry methods for label-free protein abundance profiling.
- To use protein profile data and clustering analysis to predict protein complex composition in Arabidopsis.
- To create new computational methods that can evaluate and ultimately generate predictions of protein complex composition.
- To test for conserved protein complexes in the cytosol of Glycine max

Milestones


References