Quantifying Metabolic Enzyme Variation in Non-Model Yeasts

Arjuna Subramanian[1,3*] (arjunas@princeton.edu), Meera Gupta[2,3], W. Lance Martin[3], and **Martin Wühr**[1,2,3]

[1]Department of Chemistry, [2]Department of Chemical and Biological Engineering, [3]Department of Molecular Biology & the Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ

https://www.igb.illinois.edu/DOEcenter

Project Goals: Pinpointing the role of variable enzyme expression across different yeast species is an essential aspect of understanding and exploiting yeast metabolism. To this end, we are developing high-throughput proteomics methods for deep quantification of metabolic enzyme abundances in three non-model yeast species (*Issatchenkia orientalis*, *Rhodosporidium toruloides*, and *Yarrowia lipolytica*) of interest for bioproducts synthesis, using *Saccharomyces cerevisiae* as a reference comparison. Analysis of all four yeasts will include experimental measurements of both absolute and relative protein abundances, suitable for integration with metabolomics and transcriptomics data towards complete genome-scale metabolic models for each organism.

The non-model yeasts *Issatchenkia orientalis*, *Rhodosporidium toruloides*, and *Yarrowia lipolytica* have shown substantial promise as microbial factories for synthesis of alcohols, organic acids, fatty acids, and other desirable bioproducts. The expression levels of the enzymes that catalyze reactions relevant to producing such small molecules are essential inputs for understanding, predicting, and engineering metabolic behavior in these organisms. We have adapted proteomics protocols for sample preparation, data collection,[1,2] and data analysis[3,4] to be compatible with the physiology and protein characteristics of these non-standard yeasts. Of particular importance was the development of robust and reproducible lysis methods for yeast cell walls, which are notoriously difficult to disrupt completely. Applying these tools, we confidently quantify over 4700 proteins in *S. cerevisiae*, 4800 proteins in *I. orientalis*, and 5600 proteins in *R. toruloides*, while mapping ~3000 orthologs between *S. cerevisiae* and each non-model species via reciprocal similarity. Using spike-in protein standards and precursor-based MS1 quantification we estimate that intracellular concentrations of individual detectable proteins vary over the ~10 pM-10 μM range, roughly seven orders of magnitude, in all three yeast species under study. Expression levels of orthologous proteins are correlated between each non-model species and *S. cerevisiae* with R^2-values between 0.487 and 0.587; ortholog correlation data is further resolvable at the level of individual metabolic pathways, including glycolysis, the citric acid cycle, and fatty acid metabolism. Additionally, we report shallow tandem mass tag (TMT)-based relative protein quantification of *I. orientalis*, *R. toruloides*, and *S. cerevisiae* in unconstrained, carbon-limited, nitrogen-limited, and phosphorus-limited growth conditions, reaching ~1700 proteins per species, ideal for integration with metabolomics measurements in order to gain a refined understanding of metabolic trends and tradeoffs in these non-standard organisms. The absolute and relative quantification data collected will be used to accelerate

collaborative genome-wide integrative ‘omics metabolic modeling of *I. orientalis*, *R. toruloides*, and *Y. lipolytica*; furthermore, the rigorous proteomics pipeline established may be generalized to more yeast species of interest as they arise.

References

1. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal Chem* **2014**, *86*, 7150-7158. DOI: 10.1021/ac502040v

2. Sonnett, M.; Yeung, E.; Wühr, M. Accurate, sensitive, and precise multiplexed proteomics using the complement reporter ion cluster. *Anal Chem* **2018**, *90* (8), 5032-5039. DOI: 10.1021/acs.analchem.7b04713

3. Wühr, M.; Freeman, R. M.; Presler, M.; Horb, M. E.; Peshkin, L.; Gygi, S. P.; Kirschner, M. W. Deep proteomics of the Xenopus laevis egg using an mRNA-derived reference database. *Curr. Biol.* **2014**, *24*, 1467-1475. DOI: 10.1016/j.cub.2014.05.044

4. Peshkin, L.; Gupta, M.; Ryazanova, L.; Wühr, M. Bayesian confidence intervals for multiplexed proteomics integrate ionstatistics with peptide quantification concordance. *bioRxiV* **2019**. DOI: 10.1101/210476

This work was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-SC0018260.