154. Integrated Pan-omics Measurements for Systems Level Characterization of Biological Systems

Presenter: Joshua N. Adkins*, Kristin E. Burnum, Carrie D. Nicora, Qibin Zhang, Stephen J. Callister, Scott R. Kronewitter, Aaron T. Wright, Ronald J. Moore, Eric Huang, Yi Qu, Zeping Hu, Xiaoli Gao, Weijun Qian, Erin S. Baker, Therese R. W. Clauss, Karl K. Weitz, Daniel J. Orton, Matthew E. Monroe, Anil K. Shukla, Rui Zhao, Roslyn Brown, **Gordon A. Anderson**, **Thomas O. Metz, Mary S. Lipton,** and **Richard D. Smith (PI;** rds@pnnl.gov**)**

Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352

http://www.pnl.gov/

Project Goals: Achieving a predictive systems level understanding of plants, microbes and microbial communities requires the integration of developments to enable solutions to energy, environment, and climate challenges. We are applying advanced mass spectrometry (MS)-based capabilities for comprehensive molecular characterization (proteomics including post-translational modifications, metabolomics, lipidomics, and glycomics) of biological systems. Automating and speeding key steps in pan-omics sample processing, high resolution separations combined with high mass accuracy MS measurements affords large gains in measurement quality and throughput. Additional advancements include targeted proteomics methods (activity-based protein profiling and multiple reaction monitoring) and elucidation of protein proteoforms through integrated top-down and bottom-up and post-translational modifications measurements. This pan-omics approach provides new insights by elucidating complex phenotypic relationships between environmentally important microorganisms and higher organisms, as well as metabolic activities within microbial communities.

Central to developing a complete understanding of biological systems is the paradigm of pan- omics: the ability to comprehensively characterize the range of biomolecules from individual samples. To that end, we are developing MS-based approaches and technologies, and applying their capabilities in the context of GSP collaborations. One benefit of the pan-omic technology and approaches that we are developing is the ability to analyze and obtain data on a range biomolecular classes from the same sample, eliminating key sources of biological variation while greatly reducing the amount of sample required. In one aspect of these efforts, we have developed and initially evaluated an integrated biomolecule extraction method that uses a chloroform/methanol extraction to isolate proteins, as well as polar and non-polar metabolites, from the same sample. For both a model bacterium and a unicyanobacterial microbial consortium, we have also shown nearly identical results (e.g. reproducibility and proteome coverage) using these analyses compared with a standard proteomics sample workflow.

Important challenges for pan-omics approaches involve both the biological complexity of the systems being studied, which include the microbial communities of environmentally crucial ecosystems. The study of such microbial communities can provide an understanding of the manner in which microbes affect and are affected by their environment.

Examples of resent pan-omics applications and a few of the resulting insights include:

Fungus *L. gongylophorus* dominates during cellulose degradation while symbiotic bacteria play supporting roles. Studies of fungus-growing ant–microbe symbiosis are paradigmatic of organic

complexity generated through symbiotic association. We have demonstrated in-depth profiling of the fungal garden complete with bacteria (including fungus alone, isolated bacteria, and the actual fungal garden) to understand the relationship between the fungus and the bacterial protectors. Proteomics and metabolomic studies have revealed that the fungus *L. gongylophorus* plays a dominant role in breaking down cellulose and other plant polymers, while the bacteria turn the partially digested sugars into a variety of nutrients that support the fungal and ant growth.

Gluconeogensis dominates in phototrophic mats in the early morning. Phototrophic mats rely on photosynthetic organisms for carbon capture, storage and nutrient cycling to nourish heterotrophic community members. As such, these photosynthetic organisms are dependent on storing high energy nutrients during the day that can be used after the sun goes down. A pan- omics based study of phototrophic microbial mats from Yellowstone National Park revealed that during early morning, *Synechococcus* sp are engaged in gluconeogenesis rather than glycolysis. We observe pools of glucose-6-phosphate and the presence of bifunctional fructose 1,6- bisphosphatase which catalyzes the reversed reaction of PFK and phosphoenolpyruvate synthase, which catalyzes the reversed reaction of the other major control point of glycolysis.

Pan-omics reveals community elasticity comes at a price of functional redundancy. For many microbial communities presently of interest, the metagenome is unavailable; presenting a significant challenge since proteomics is often interpreted using genome sequence(s). However, the pan-omics approaches we are developing can still provide insights into the ecology of such microbial communities. In a study of a community isolated from cow rumen grown in an engineered bioreactor in the absence of metagenome data, our pan-omic approach (including proteomic, transcriptomics, and metabolomics studies) revealed redundant patterns and correlations between groups of proteins and metabolites with the community structure. Additionally, we found that after perturbation of the community, the pan-ome for the community returned to a new steady state, but the functional redundancy was decreased.

Revealing the 'true proteome' of periplasmic proteins. An important aspect of our approach is the ability to provide extensive information on the actual proteoforms present, and where use of conventional bottom-up proteomics approaches are generally ineffective. The periplasm of Gram-negative bacteria is a dynamic and physiologically important subcellular compartment where the constant exposure to potential environmental insults amplifies the need to protect function, and thus the proteoforms present are functionally important. Hence, the evaluation of the periplasmic fraction for *Novosphingobium aromaticivorans* revealed a large array of proteoforms for 55 proteins in the periplasm. The proteoforms found included post translational modifications due to signal peptide removal, N-terminal methionine excision, acetylation, glutathionylation, pyroglutamate, and disulfide bond formation.

Acknowledgements: This research is supported by the Office of Biological and Environmental Research (OBER) of the U.S. Department of Energy (DOE). Portions of this research were performed in the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the DOE's OBER. Pacific Northwest National Laboratory is operated for the DOE by Battelle Memorial Institute through Contract No. DE-AC05-76RLO 1830.