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

Joshua N. Adkins*, Kristin E. Burnum-Johnson, Carrie D. Nicora, Qibin Zhang, Stephen J. Callister, Scott R. Kronewitter, Aaron T. Wright, Ronald J. Moore, Eric Huang, Yi Qu, Weijun Qian, Young-Mo Kim, Erin S. Baker, Therese R. W. Clauss, Karl K. Weitz, Daniel J. Orton, Sam Payne, Matthew E. Monroe, Anil K. Shukla, Rui Zhao, Roslyn Brown, **Gordon A. Anderson**, **Thomas O. Metz, Mary S. Lipton,** and **Richard D. Smith** (Pacific Northwest National Laboratory)

Pacific Northwest NATIONAL LABORATORY

Project Goals

Achieving a predictive systems level understanding of plants, microbes and microbial communities requires an integrative experimental approach to ultimately enable solutions to energy, environment, and climate challenges. We are applying advanced mass spectrometry (MS)-based capabilities for comprehensive molecular characterizations (proteomics – including post-translational modifications – metabolomics, lipidomics, and glycomics) of biological systems. Evaluation of a common solvent-based extraction protocol indicates that it can be used to isolate proteins, metabolites, and lipids from the same sample for subsequent omics analyses. Additional advancements include targeted proteomics methods (activity-based protein profiling and selected reaction monitoring) and elucidation of protein proteoforms through integrated top-down, bottom-up, and post-translational modifications measurements. This pan-omics approach allows new biological insights by providing a holistic view of complex phenotypic relationships between environmentally important microorganisms and higher organisms, as well as metabolic activities within microbial communities.

An integrated pan-omics strategy

Sample extraction and processing

Extraction procedures for

- Communities (soils, microbial mats, fungal gardens, water column, etc.)
- Plants (root hairs, woody tissue)

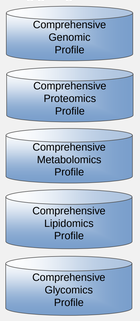
Extracted components

- Proteins (sub-cellular fractions, activity-based enrichments, post-translational modifications and intact protein analysis)
- Metabolites
- Lipids
- Glycans

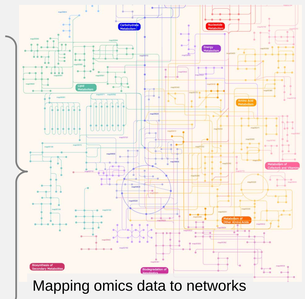
Quantitative approaches

- Integrated peak intensities, labeling strategies, and spectrum counts

Data extraction, processing, and aggregation



Integrative interpretation and model population



Mapping omics data to networks

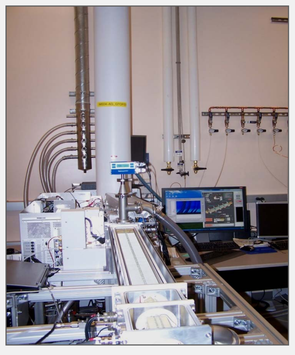
- Metabolic
- Signaling
- Regulatory

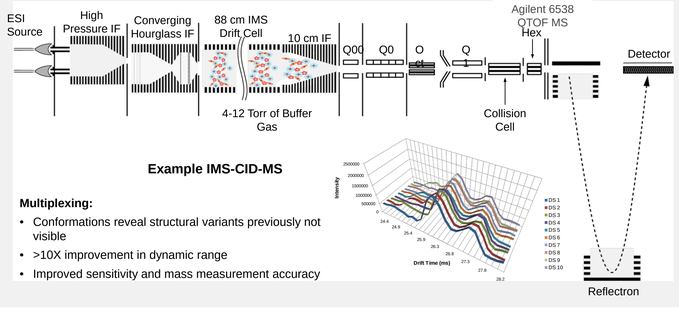
Technology developments to enable current and future pan-omics measurements

Higher throughput and more sensitive measurements from smaller samples

- Improved electrospray-mass spectrometry interfaces
- Instrument modifications to enable better accumulation of ions
- Improved instrument data acquisition speeds

Informatics approaches that improve quantification and omics coverage (e.g. at a given FDR)





Example IMS-CID-MS

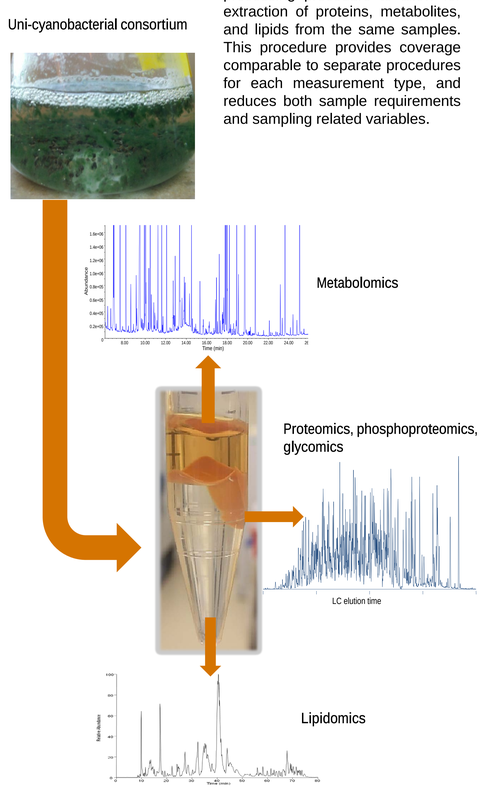
Multiplexing:

- Conformations reveal structural variants previously not visible
- >10X improvement in dynamic range
- Improved sensitivity and mass measurement accuracy

Beyond Genome Potential - Applications of Pan-Omics Approaches for Microbial Communities

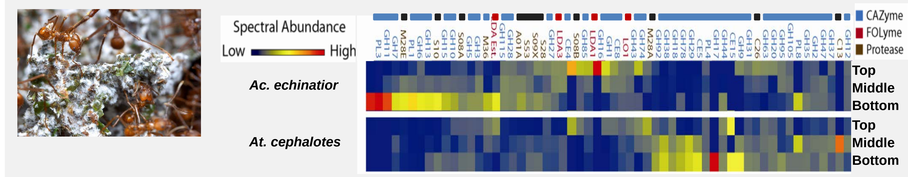
Pan-omic Sample Preparation to Increase Biological Insights and Reduce Variability

Illustration of pan-omic sample processing procedure to allow for extraction of proteins, metabolites, and lipids from the same samples. This procedure provides coverage comparable to separate procedures for each measurement type, and reduces both sample requirements and sampling related variables.

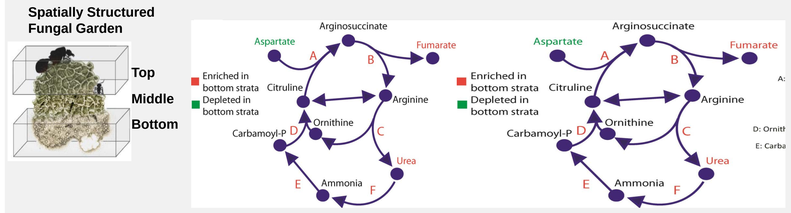


Spatially structured cellulose degradation in symbiotic microbial-fungal gardens*

Collaborators: Frank Aylward & Cameron Currie, University of Wisconsin-Madison



Spatial classification of the *L. gongylophorus* proteome in two different fungal gardens (*Ac. echinatior* and *At. Cephalotes).* The heatmap represents the relative number of mass spectra matching to CAZyme (blue), FOLyme (magenta), and Proteases (brown) protein families in the top (T), middle (M), and bottom (B) strata .

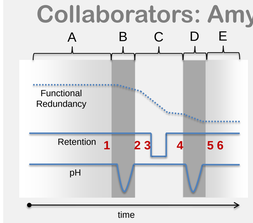


Meta-proteomics and meta-metabolomics results suggests *that L. gongylophorus* transitions from using urea in the top stratum to producing it in the bottom

*Research *in press* at *Environmental Microbiology and Environmental Microbiology Reports*

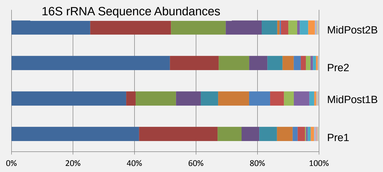
Functional resilience of a rumen community during repeated pH perturbations

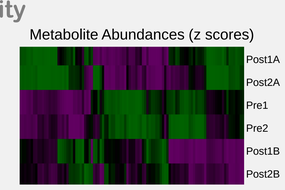
Collaborators: Amy Boaro & Birgitte Ahring, Washington State University



Sampled Time Points

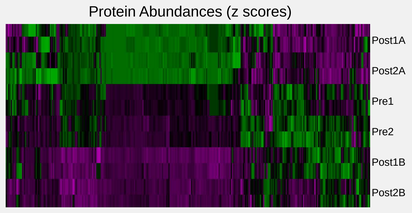
1. Pre-Perturbation 1 (Pre1)
2. Post-Perturbation 1A (Post1A)
3. Post-Perturbation 1B (Post1B)
4. Pre-Perturbation 2 (Pre2)
5. Post-Perturbation 2A (Post2A)
6. Post-Perturbation 2B (Post2B)





Five stages of the perturbation experiment

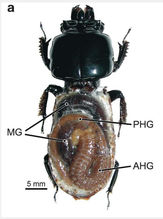
1. Establishment of steady state
2. Pulse with H_3PO_4 to ΔpH 6.8 to 5.5, with NaOH to pH 6.8.
3. Decrease hydraulic retention time from 4 days to 2 days to add second stress to result in a hypothesized decrease in functional redundancy. Restore hydraulic retention time.
4. Repeat pulse perturbation.
5. Re-establishment of steady state operation.



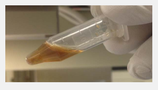
pH perturbation had a negative impact on *Fibrobacter* and a positive impact on *Prevotella* based on changes in protein abundances. We speculate that based on the pH resistance and nature of cellulose/cellobiose degradation that *Prevotella* has a short-term advantage during the pulse. The balance quickly recovers following the return to 'stabile' conditions. Utilizing metabolomics and 16s rRNA give complementary insights, providing a basis for hypothesis generation and testing.

Preliminary evaluation of the host and community pan-ome in the gut of lignocellulose degrading Passilid beetles

Collaborators: Eoin Brodie[1], Javier Ceja-Navarro[1], & Jennifer Pett-Ridge[2], [1]LBNL and [2] LLNL

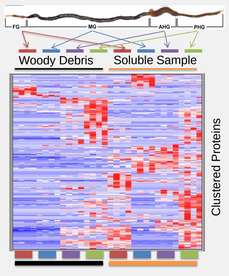


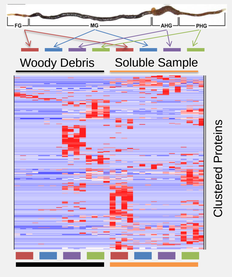
Ceja-Navarro et al, ISME J, 2014



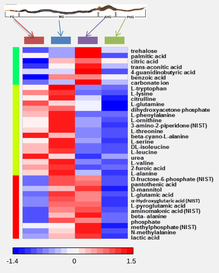
Passalid Beetles – Preliminary Analysis

1. >900 species world-wide
2. 4.5X it weight in wood is processed/day
3. 4 gut regions being sequenced to identify community members
4. Studied soluble extracted proteins and proteins from residual woody debris
5. Initial proteomics analysis utilized a distantly related genome for the beetle and an extensive microorganism database





Generally, good recovery of both beetle and microbial community proteins was observed. Future efforts will require some fractionation and database refinement. Beetle proteins were more likely to be observed across conditions, whereas community proteins were more gut region restricted.



Given previous experience with cellulose degrading systems, there are fewer cellulose related metabolites present. This may be related to the rapid consumption and digestion of wood and the colony environment.

Presentations of GSP Projects Utilizing Pan-omics Capabilities

PI - Title

Richardson, Ruth - Development of Quantitative Protein Biomarker Assays for Enzymes Involved in Bacterial Iron and Uranium Reduction
Richardson, Ruth - Identification of Iron Reductases Using Top-down Proteomics and Heterologous Expression]
Richardson, Ruth - Investigations into the Proteome of Gram-positive Metal-reducing Bacterium Desulfotomaculum reducens MI-1
Qian, Wei-Jun - Quantitative Site-Specific Redox Proteomics on Protein Thiols and Broad Light/Dark Modulation of Thiol Oxidation in Cyanobacteria
Payne, Samuel - Algorithms Enabling Metaproteomics
Worden, Alexandra - Examining the post-transcriptional program governing the metabolic proteome of Micromonas pusilla
Worden, Alexandra - Phytochromes in widespread photosynthetic algae reveal origins of plant signaling proteins]
Worden, Alexandra - Responses of the Widespread Green Alga Micromonas to Elevated CO2 at Different Levels of Phosphorus Availability
Beliaev, Alexander - Characterization of Redox Signaling Pathways in Cyanobacteria during Nutrient Limitation
Firestone, Mary - Mapping soil carbon from cradle to grave #1: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C
Firestone, Mary - Mapping soil carbon from cradle to grave #2: Multi-level omics analyses for parameterization of trait-based models of rhizosphere microbial community function
Pett-Ridge, Jennifer - A Systems Biology Characterization of the Biotechnological Potential Stored in the Wood-Feeding Beetle Odontotaenius disjunctus
Fredrickson, James - FSFA Poster #2: Metabolic Interactions Structuring Microbial Communities
Baker, Scott - Identifying biological pathway targets for lipid production in Yarrowia lipolytica
Lindemann, Stephen - Interrogating the interspecies interactions that drive microbial community assembly: from phylotype to function

U.S. DEPARTMENT OF ENERGY

Acknowledgements

This research was funded by the U. S. Department of Energy (DOE) Genomic Science Program. Work was performed in the Environmental Molecular Sciences Laboratory (EMSL), a DOE/BER national scientific user facility at Pacific Northwest National Laboratory.

Contact Information
Richard D. Smith, Ph.D.
Dick.Smith@pnnl.gov

omics.pnl.gov