Agile BioFoundry CRADA Project – Agilent

Title: Development and implementation of high-throughput proteomic and metabolomics assays by using advanced chromatographic and mass spectrometric systems

Project Partners: Pacific Northwest National Laboratory, Lawrence Berkeley National Laboratory, Sandia National Laboratories

Relevant ABF Capabilities:

- Test: Multi-omics
- Learn: Machine Learning
- Learn: Deep Learning

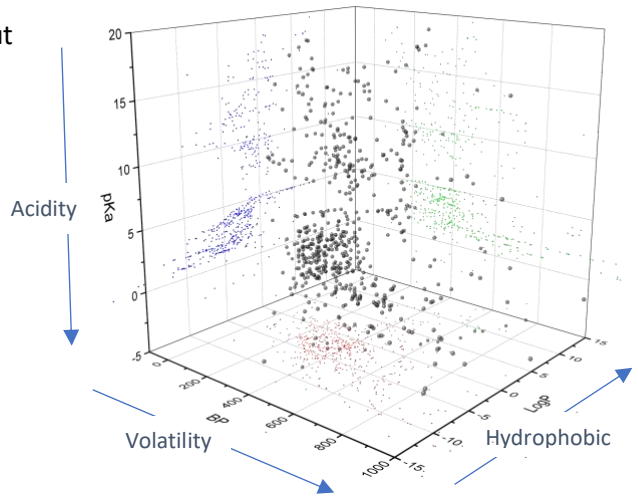

Figure 1. The mass spectrometry data generated in this work will be used to generate a carefully selected subset of possible high throughput methods that can span the chemical space (acidity, volatility, hydrophobicity) of metabolites.

Description:

Elucidating multiple types of information about engineering biosynthetic pathways and the host biological processes is necessary to generate groundbreaking insights critical to driving bio-based chemical optimization processes. Significant advances in synthetic biology, genome editing, and DNA synthesis capabilities have propelled the ability to design and construct novel strains for biomanufacturing research. Generating thousands of unique strains for a given target molecule covering multiple pathways, tuned protein expression, and targeted genome modification is now routine. Yet, analytical tools to measure metabolites and proteins do not match this throughput, resulting in significant bottlenecks in strain testing efforts. There is great need for high-throughput analytical workflows that reduce time and resource needs to enable synthetic biology research at the Agile BioFoundry (ABF). Agilent Technologies' developments in high-throughput Liquid Chromatography-Mass Spectrometry (LC-MS/MS) methods, coupled with advanced software to predict which methods will be successful, offer valuable components for these types of workflows. Interfacing these technologies with ABF software (i.e., Experimental Data Depot (EDD)) and workflows for strain design and construction would have a significant impact on foundry operations, strain development processes, and mathematical modeling efforts.

Agilent Technologies' prototype system, which can process hundreds of samples per day, was developed using ultra-high performance liquid chromatography-quadrupole time-of-flight (UHPLC-QTOF) mass spectrometry-based targeted metabolomic methods. In this CRADA, we will extend this technology to targeted triple quadrupole and drift tube ion mobility spectrometry (DTIMS) mass spectrometry

analyses. Both mass spectrometry platforms will incorporate targeted metabolomic methods with chromatographic methods developed by Agilent Technologies. Implementation on these mass spectrometry systems offers selectivity, increased sensitivity, compound structural information, and direct integration with ABF software systems such as the EDD. By developing triple quadrupole and DTIMS methods, additional information will be available to inform analytical method selection, allowing a much larger metabolomic space to be analyzed (Fig. 1). The resulting data will support further development of methods like Random Forests, Artificial Neural Networks and Gradient Boosted Decision.

Trees will be combined with Locally Interpretable Machine-Agnostic Explanations to generate high accuracy, human interpretable decision paths for matching metabolites to analytical methods, thereby advancing Agilent techniques' software prediction.

ABF Contacts: Alex Apffel (Agilent), Anya Tsalenko (Agilent), Kristin Burnum-Johnson (PNNL), Jon Magnuson (PNNL), Chris Petzold (LBNL), Corey Hudson (SNL)

Performance period: 06/01/19-06/01/21

Resulting publication(s)/patent(s): None to date