Plant-Microbe Interfaces: Dynamic data analytics through integrated knowledgebase and LIMS

Guruprasad Kora[2*] (koragh@ornl.gov), Susan Holladay[1], Tatiana Karpinets[1], Mustafa Syed[1], Christopher Schadt[1], **Ed**ward **Uberbacher**[1], and **Mitchel J. Doktycz**[1]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2]Computer Science and Mathematics, Oak Ridge National Laboratory, Oak Ridge, TN

http://PMI.ornl.gov

Project Goal: Understand the genome-dependent molecular and cellular events involved in establishing and maintaining beneficial interactions between plants and microbes. *Populus* and its associated microbial community serves as an initial experimental system for understanding how these molecular events manifest themselves within the spatially, structurally, and temporally complex scales of natural systems. To achieve this goal, we will focus on 1) characterizing the natural variation in *Populus* microbial communities within complex environments, 2) elucidating *Populus*-microbial interactions at the molecular level and dissecting the signals and pathways responsible for initiating and maintaining microbial relationships, and 3) performing metabolic and genomic modeling of these interactions to aid in interpreting the molecular mechanisms shaping the *Populus*-microbial interface.

Data management, efficient tracking and effective visualization are essential aspects of any field-study related biological process. Each of these aspects pose numerous challenges as size and complexity of the data and field-study campaigns grow. Establishing a disciplined data management infrastructure at all levels of field-study campaigns enables effective life-cycle (data collection to analysis) management of any biological processes.

The Plant-Microbe Interfaces (PMI) project has developed an easy to use, web-based data management and tracking portal as part of their Knowledgebase for PMI field-studies that enables users to track and manage the entire life-cycle of the collected data. PMI LIMS and Knowledgebases are coupled together to form an integrated data analytics platform to perform statistical as well as data mining operations dynamically. The Knowledgebase augments data analytics infrastructure with scientific visualization techniques (visual data analysis). It enables users to use graphical representation of data as a means of gaining understanding and insight into the data. It helps researchers to comprehend spatial and temporal relationships between collected data. It provides efficient interactive techniques for researchers to focuses on exploratory, comparative analytics and visualization.

The PMI Knowledgebase has an established data acquisition workflow that enabled field-studies carried out in 2009-2011 to be effectively managed. Currently, the system supports various types of data from raw soil readings to tree specification to post-processed data like 454 sequences, isolates etc. The system supports various datatypes from image files, ascii-based text files, binary blobs etc. The workflow interface enables users to access the underlying data management layer with an easy to use and intuitive web interface. The interface seamlessly connects to laboratory-wide LIMS environment and makes day-to-day tasks like raw-data fetching and data summarization extremely efficient and easy. Furthermore, the portal provides a dynamic data analytics environment that facilitates users to perform standard statistical analysis on the LIMS stored data.

The Plant Microbe Interfaces Scientific Focus Area is sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research