161.   Analysis of Microbes

Paramvir Dehal*[1](pdehal@lbl.gov), Chris Henry[2], Jason Baumohl[1], Aaron Best[1], Jared Bischof[2], Ben Bowen[1], Tom Brettin[2], Tom Brown[2], Shane Canon[1], Stephen Chan[1], John- Marc Chandonia[1], Dylan Chivian[1], Ric Colasanti[2], Neal Conrad[2], Brian Davison[3], Matt DeJongh[6], Paramvir Dehal[1], Narayan Desai[2], Scott Devoid[2], Terry Disz[2], Meghan Drake[3], Janaka Edirisinghe[2], Gang Fang[7], José Pedro Lopes Faria[2], Mark Gerstein[7], Elizabeth M. Glass[2], Annette Greiner[1], Dan Gunter[1], James Gurtowski[5], Nomi Harris[1], Travis Harrison[2], Fei He[4], Matt Henderson[1], Chris Henry[2], Adina Howe[2], Marcin Joachimiak[1], Kevin Keegan[2], Keith Keller[1], Guruprasad Kora[3], Sunita Kumari[5], Miriam Land[3], Folker Meyer[2], Steve Moulton[3], Pavel Novichkov[1], Taeyun Oh[8], Gary Olsen[9], Bob Olson[2], Dan Olson[2], Ross Overbeek[2], Tobias Paczian[2], Bruce Parrello[2], Shiran Pasternak[5], Sarah Poon[1], Gavin Price[1], Srivdya Ramakrishnan[5], Priya Ranjan[3], Bill Riehl[1], Pamela Ronald[8], Michael Schatz[5], Lynn Schriml[10], Sam Seaver[2], Michael W. Sneddon[1] , Roman Sutormin[1], Mustafa Syed[3], James Thomason[5], Nathan Tintle[6], Will Trimble[2], Daifeng Wang[7], Doreen Ware[5], David Weston[3], Andreas Wilke[2], Fangfang Xia[2], Shinjae Yoo[4], Dantong Yu[4], **Robert Cottingham**[3], **Sergei Maslov**[4], **Rick Stevens**[2], **Adam P. Arkin**[1]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA, [2]Argonne National Laboratory, Argonne, IL, [3]Oak Ridge National Laboratory, Oak Ridge, TN, [4]Brookhaven National Laboratory, Upton, NY, [5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY [6]Hope College, Holland, MI, [7]Yale University, New Haven, CT, [8]University of California, Davis, CA, [9]University of Illinois at Champaign-Urbana, Champaign, IL, [10]University of Maryland, College Park, MD

http://kbase.us

Project Goals: The KBase project aims to provide the capabilities needed to address the grand challenge of systems biology: to predict and ultimately design biological function. KBase enables users to collaboratively integrate the array of heterogeneous datasets, analysis tools and workflows needed to achieve a predictive understanding of biological systems. It incorporates functional genomic and metagenomic data for thousands of organisms, and diverse tools for (meta)genomic assembly, annotation, network inference and modeling, allowing researchers to combine diverse lines of evidence to create increasingly accurate models of the physiology and community dynamics of microbes and plants. KBase will soon allow models to be compared to observations and dynamically revised. A new prototype Narrative interface lets users create a reproducible record of the data, computational steps and thought process leading from hypothesis to result in the form of interactive publications.

The Microbial Sciences component of the KBase project has three overall goals: 1) to enable the generation of predictive models for metabolism and gene regulation to facilitate the manipulation of microbial function; 2) to vastly increase the capability of the scientific community to communicate and utilize existing data; and 3) to enable the planning of effective experiments and to maximize our understanding of microbial system functions. To achieve these goals we have focused on unifying existing 'omics datasets and modeling toolsets within a single integrated framework that will enable users to move seamlessly from the genome assembly and annotation process through to a reconciled metabolic and regulatory model that is linked to all existing experimental data for a particular organism. The results are hypotheses for such things as gene-function matching and the use of comparative functional genomics to perform higher quality evidence-based annotations. KBase offers tools for reconciling the models

against experimental growth phenotype data, and using them to predict phenotypes in novel environments or under genetic perturbations.

To prioritize the development of the microbial science area and enable new science, we will focus on accomplishing prototype science workflows rather than general tasks. To date we have developed KBase capabilities and demonstrations workflows for: (1) genome annotation and metabolic reconstruction, (2) regulon reconstruction, (3) metabolic and regulatory model reconstruction, and (4) reconciliation with experimental phenotype and expression data.