KBase Partners with SFA Developers to Increase Functionality of Tools and Analysis

Elisha M. Wood-Charlson,[1 *] (elishaw@lbl.gov), Meghan Drake,[2] Chris Henry,[2] Paramvir Dehal,[1] Dylan Chivian,[1] Robert Cottingham,[3] and **Adam P. Arkin.** [1]

[1]Environmental Genomics and Systems Biology, Lawrence Berkeley National Lab, Berkeley, CA; [2]Argonne National Laboratory, Chicago, IL; [3] Oak Ridge National Lab, Oak Ridge, TN.
*presenting author

http://www.kbase.us

Project Goals: (Limit to 1000 characters)
Several SFAs were selected for training on the KBase Software Development Kit (SDK) to integrate tools and data into the KBase platform. Overall, the collaborative effort was viewed as highly successful. Every SFA project was able to achieve their proposed targets for training and functionality, while also gaining an appreciation for the role KBase plays/could play in their future science plans. Both the SFA and KBase teams felt that continued collaboration would be beneficial, especially if the programs were able to jointly prioritize scope and readiness of future projects. This poster will outline the scope of functionality added by SFA community developers, and highlight points of intersection across these efforts.

Abstract text

In total, 5 SFAs were awarded additional funds to support developer time, with the goal of integrating analysis tools into KBase to support the SFA's science. Across the board, this strategy seemed to be well-received, as several SFAs confirm it can be challenging to justify resources for technology-related goals alongside the often complex and ambitious SFA science goals. Two SFAs (LLNL Biofuels and Soil) reported that they are already actively using the new functionality inside KBase, and Stuart acknowledged that the integration has greatly improved their model development. Other SFAs recognized that this effort focused mainly on the core implementation of longer-term analysis goals (LANL Fungal, LLNL Soil, ORNL Structure). All SFAs recognized the value in these efforts to enable the broader research community to also benefit from the additional functionality in KBase.

Overall, the SFAs reported significant progress towards their proposed goals during the additional funding period. For example, KBase now has the first phage analysis pipeline in the system (LLNL Soil), which was presented as an online webinar to the broader user community in November 2020. KBase has several ways for researchers to improve their fine-scale evaluations of microbial communities by exploring strain-level variation (ENIGMA), improve metabolic models by combining various annotation methods (LLNL Biofuels), and screen small-molecule

ligand binding to protein structures (ORNL Structure). Still coming are additional measures to assess quality of contigs prior to assigning taxonomic and/or functional annotations (LLNL Soil), and a joint SFA effort to provide functional annotation for viral sequences (LLNL Biofuels and Soil). Many of the SFAs reported an added bonus to these efforts - increased knowledge of KBase's resource enhancements and community investment, making them ambassadors for KBase in their own communities (LANL Fungal, LLNL Biofuels and Soil).

This poster will outline the major functional components added by SFA developers. The presenter will be available to answer any questions about possible future collaborations.

This work is supported as part of the Genomic Sciences Program DOE Systems Biology Knowledgebase (KBase) funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-AC02-98CH10886.