

Tools and Data for Metabolic Modeling in KBase

José P. Faria*² (jplfaria@anl.gov), **Christopher Henry**², Janaka Edirisinghe², **Adam P. Arkin**¹, **Bob Cottingham**³, Nomi Harris¹ and the KBase Team at the following institutions

¹Lawrence Berkeley National Laboratory, Berkeley, CA; ²Argonne National Laboratory, Argonne, IL; ³Oak Ridge National Laboratory, Oak Ridge, TN; ⁴Brookhaven National Laboratory, Upton, NY; ⁵ Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

<http://kbase.us>

Project Goals: The DOE Systems Biology Knowledgebase (KBase) is a free, open-source software and data platform that enables researchers to collaboratively generate, test, compare, and share hypotheses about biological functions; analyze their own data along with public and collaborator data; and combine experimental evidence and conclusions to model plant and microbial physiology and community dynamics. KBase's analytical capabilities currently include (meta)genome assembly, annotation, comparative genomics, transcriptomics, and metabolic modeling. Its web-based user interface supports building, sharing, and publishing reproducible, annotated analysis workflows with integrated data. Additionally, KBase has a software development kit that enables the community to add functionality to the system.

KBase offers “one-stop shopping” for a growing range of integrated analysis tools and datasets that enable assembly, annotation, construction, and comparison of genome-based metabolic models. One of the most powerful and popular workflows in KBase starts with genomic sequence reads, goes through pipelines for assembly and annotation, and culminates in metabolic models that can be used to evaluate an organism's metabolic capabilities by simulating growth under different conditions. KBase's metabolic modeling toolkit supports the reconstruction, prediction, and design of metabolic networks in microbes and plants (and their communities). These tools can help advance efforts to optimize microbial production of a certain biofuel, find the minimal media conditions under which that fuel is generated, or predict species interactions within a soil or industrial microbiome.

In KBase, genome-scale metabolic models are primarily reconstructed from functional annotations produced by the KBase annotation apps (see figure). When a genome is functionally annotated, its metabolic genes are mapped onto biochemical reactions. This information is integrated with data about reaction stoichiometry, subcellular localization, biomass composition, estimation of thermodynamic feasibility (directionality of reactions) to produce a detailed stoichiometric model of metabolism.

KBase users have applied the system to address a range of scientific problems. Some examples of published studies that used KBase's metabolic modeling toolkit include finding patterns in the presence of core metabolic pathways across the microbial phylogenetic tree; identifying inconsistent growth conditions in *Klebsiella* KPPR1; and several studies that analyzed the interactions within communities of microbes and plants. The Narratives that researchers have

chosen to share publicly (see <http://kbase.us/narrative-library>) can be viewed, copied, and re-run, with varying parameters or new datasets.



Figure 1. KBase’s metabolic modeling tools (green) as well as some other analysis tools that interoperate with them.

Resources for learning more about KBase's metabolic modeling suite include:

- Home page about modeling in KBase: <http://kbase.us/metabolic-modeling-in-kbase/>
- Narrative tutorial: “Microbial Metabolic Model Reconstruction and Analysis” (<https://narrative.kbase.us/narrative/ws.18302.obj.61>). You can copy the Narrative and rerun the steps yourself.
- Video tutorial: visit the KBase YouTube channel (https://www.youtube.com/channel/UCTX9Pn_WCCCPoER9-3AGL_A)
- Modeling FAQ: <http://kbase.us/metabolic-modeling-faq/>
- Narrative Library: <http://kbase.us/narrative-library>

KBase is funded by the Genomic Science program within 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.