How KBase Supports Education

Adam P. Arkin[1], **Robert Cottingham**[3], **Chris Henry**[2], Ellen Dow[1*] (egdow@lbl.gov), Benjamin Allen[3], Jason Baumohl[1], Kathleen Beilsmith[2], David Dakota Blair[4], Jay Bolton[1], Shane Canon[1], Stephen Chan[1], John-Marc Chandonia[1], Dylan Chivian[1], Zachary Crockett[3], Paramvir Dehal[1], Meghan Drake[3], Janaka N. Edirisinghe[2], José P. Faria[2], Jason Fillman[1], Tianhao Gu[2], AJ Ireland[1], Marcin P. Joachimiak[1], Sean Jungbluth[1], Roy Kamimura[1], Keith Keller[1], Vivek Kumar[5], Sunita Kumari[5], Miriam Land[3], Sebastian Le Bras[1], Zhenyuan Lu[5], Filipe Lui[2], Dan Murphy-Olson[2], Erik Pearson[1], Gavin Price[1], Priya Ranjan[3], William Riehl[1], Boris Sadkhin[2], Samuel Seaver[2], Alan Seleman[2], Gwyneth Terry[1], Charles Trenholm[1], Sumin Wang[1], Doreen Ware[5], Pamela Weisenhorn[2], Elisha Wood-Charlson[1], Ziming Yang[4], Shinjae Yoo[4], Qizhi Zhang[2]

[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.

http://kbase.us

The Department of Energy Systems Biology Knowledgebase (KBase) is a knowledge creation and discovery environment designed for both biologists and bioinformaticians. KBase integrates a large variety of data and analysis tools, from DOE and other public services, into an easy-to-use platform that leverages scalable computing infrastructure to perform sophisticated systems biology analyses. KBase is a publicly available and developer extensible platform that enables scientists to analyze their own data within the context of public data and share their findings across the system.

The KBase user interface (UI) enables instructors to work with students to conduct hands-on data science research and analysis without the need for programming skills or computational resources. The KBase team works with instructors and researchers of varying skill and career levels to ensure the transfer of domain knowledge is accompanied by an understanding of bioinformatic tools and techniques. KBase has been supporting education in three main ways: establishing the KBase Educators program, participating in the Environmental Molecular Sciences Laboratory Summer School, and hosting a variety of topic-based workshops and webinars. Each program supports a different cross-section of the BER research community, with the overall goal of improving and expanding the next-generation of data analysis using KBase.

KBase Educators

The KBase Educators program (https://www.kbase.us/kbase-educators/) consists of biological and data science instructors, ranging from High School to Graduate level, that have adapted the KBase platform to their curriculum needs by developing modular, adaptable, and customizable instructional units using KBase Narratives. These instructional modules contain teaching resources, data, analysis tools, and mark-down utility to tailor instructions and learning goals. Each module can be adapted for independent class concepts that involve Genomics, Metagenomics, Phylogenetics, Pangenomics, and/or Metabolic Modeling. The KBase Educators Organization provides access to resources in KBase, and a KBase Users Slack group provides access to a community network of peers, supported by community-driven guidelines,

instructional templates, and KBase staff. Working groups are short-term commitments around topics that the community identifies as important or relevant to their teaching goals. The aim behind community working groups (https://www.kbase.us/educator-working-groups/) is to enable educators using KBase to collaboratively build resources to add to the Educators Org, while dividing the workload and providing accountability to deliver resources. The KBase team helps facilitate working group formation and setting up the initial kick-off meeting to ensure everyone is familiar with the scope and goal, providing support during development and release of materials.

EMSL Summer School

KBase collaborates with and supports instruction for the Environmental Molecular Sciences Laboratory (EMSL) Summer School for early career researchers. National laboratory and academic researchers provide lectures on multidisciplinary research over the course of a week, which are free and open to the public. In addition, post-doctoral researchers and Ph.D. students attended tutorial classes, which provided hands-on instruction and learning. KBase contributes to both lectures and hands-on training on the complex analysis of biological data. During the 2020 (https://www.kbase.us/multiscale-microbial-dynamics-modeling/) and 2021 (https://www.kbase.us/multi-omics-modeling-of-biochemical-pathways/) programs, KBase was used for analysis and sharing data on diverse topics using bacterial and fungal data.

KBase Workshops and Webinars

For special topics and collaborations, KBase hosts workshops that address needs of research groups, educators, and classes. Workshops target a smaller class size to better facilitate hands-on tutorials with attendees and students. For the broader audience of all KBase Users, KBase holds live webinars throughout the year with a live question & answer section, which introduce new features and topics with speakers including community developers and subject matter experts. Webinars are posted on the KBase YouTube channel (https://www.youtube.com/DOEKBase) for anyone to revisit and view after the event and often include public Narratives for users to test out new tools and workflows.

Through each of these approaches, KBase empowers skilled researchers and inspires the next generation of biologists and data scientists by providing a platform that seamlessly enables users to integrate conceptual knowledge with sophisticated systems biology investigative tools.

1. Dow EG, Wood-Charlson EM, Biller SJ, Paustian T, Schirmer A, Sheik CS, Whitham JM, Krebs R, Goller CC, Allen B, Crockett Z and Arkin AP. Bioinformatic Teaching Resources – For Educators, by Educators – Using KBase, a Free, User-Friendly, Open Source Platform. Front. Educ. 2021; 6:711535. doi: 10.3389/feduc.2021.711535

This work is supported as part of the BER Genomic Science Program. The DOE Systems Biology Knowledgebase (KBase) is 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.