Partnering Mechanisms

Through the Facilities Integrating Collaborations for User Science (FICUS) initiative, JGI and NERSC offer Microbiome Data Science expertise, enabling users to perform metagenomics research and computational genomics, and translating sequence information into biological discovery. Examples include assembly, annotation, and mining isolate genome and metagenome data for genetic features of interest, e.g., metabolic pathways.

bit.ly/JGI-NERSC-FY18

Community Science Program (CSP)

Peer-reviewed selection process for massive-throughput sequencing, DNA synthesis and big data analysis and software development for projects of relevance to sustainable energy production, global element cycling, and biogeochemistry.

bit.ly/CSP-home

Cooperative Research and Development Agreements (CRADAs)

Enable research jointly sponsored by the Berkeley Lab and one or more partners for shared benefit.

Strategic Partnership Projects (SPPs)

Enable research funded by an industry partner to perform a defined scope of work using JGI's unique facilities, equipment, and personnel.

bit.ly/JGI-IEP

Workshops/Meetings

Microbial Genomics & Metagenomics (MGM) Workshops

September 16–20, 2019
Berkeley, CA
Five days of hands-on workshops combining intensive seminars and hands-on tutorials for the IMG/M suite of tools for annotation and comparative analysis of prokaryotic and viral genomes and metagenomes.

mgm.jgi.doe.gov

15th Annual Genomics of Energy & Environment Meeting

March 23–27, 2020
San Francisco, CA
Bioinformatics tutorials and workshops followed by three days of presentations and poster sessions by leading researchers connected to energy and environmental science.

usermeeting.jgi.doe.gov

JGI JOINT GENOME INSTITUTE
A DOE OFFICE OF SCIENCE USER FACILITY

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

NERSC

BioSciences

DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA USA 94598

jgi.doe.gov

18-JG-4963

JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

Engage JGI Data Science To Enable Impactful Bioscience R&D

The U.S. Department of Energy (DOE) Joint Genome Institute (JGI) is the only large scale genomic science user facility dedicated to enabling researchers to efficiently conduct analyses of microbes, microbial communities, plants, fungi, algae and other targets relevant to DOE missions in energy, environmental biology, and global carbon and other nutrient cycling.

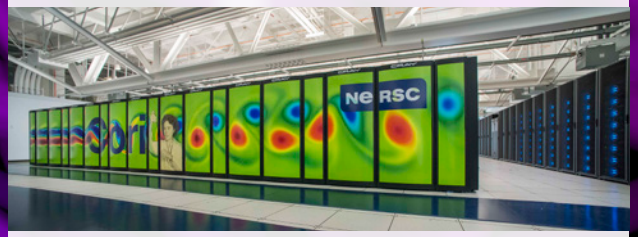
"Let's talk about how the JGI can help advance your research objectives."

– **Kjiersten Fagnan**

JGI Chief Informatics Officer
Data Science and Informatics Leader

kmfagnan@lbl.gov

Bright Minds + Big Iron = Impactful Data Science



The staff of the JGI has well-established expertise in computational biology and bioinformatics — generating, analyzing and managing massive, diverse data sets from the sequencing, synthesis and metabolomics user projects and in-house research initiatives. To optimize algorithms and workflows for the High Performance Computing (HPC) environment, the JGI partners with the National Energy Research Scientific Computing Center (NERSC), another user facility operated by Lawrence Berkeley National Laboratory. The HPC capabilities are provided by a Cray XC40 named Cori, NERSC's newest supercomputer and the 8th most powerful in the world. The JGI generates petabytes of high-quality nucleic acid sequence data and analyses.

This JGI-NERSC partnership combines the power of data science and HPC to enable users to explore the wealth of genomic and metagenomic data generated worldwide and address DOE-relevant questions that require massive-scale data and computing capabilities.

Data Portal Access

The JGI's data are freely available for download and comparative analysis through:

- **JGI Genome Portal**, JGI's central hub for genome data download.

  genome.jgi.doe.gov/portal

- **Integrated Microbial Genomes & Microbiomes** (IMG/M), the largest publicly available integration of assembled metagenomic sequences and isolate genomes.

  img.jgi.doe.gov

- **Phytozome**, plant comparative genomics portal.

  phytozome.jgi.doe.gov

- **MycoCosm**, comparative genomics of fungi.

  bit.ly/MycoCosm-JGI

- **GOLD**, (Genomes OnLine Database) a resource for sequencing projects and associated metadata.

  gold.jgi.doe.gov

KBase
PREDICTIVE BIOLOGY

KBase

The DOE Systems Biology Knowledgebase is the first large-scale bioinformatics platform that enables users to upload their own data, analyze it, build models, share, and publish their workflows and conclusions. Get access to JGI tools and data through the KBase platform. Integrate your own data sets and contribute back to the community. KBase and JGI hold extensive workshops at the JGI Genomics of Energy & Environment Annual Meeting.

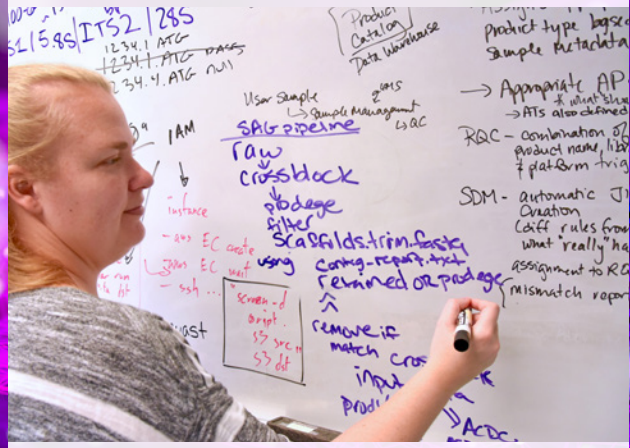
kbase.us

Software Development

At the JGI, over 80 software developers solve data processing and integration challenges across production infrastructure, technology pipelines and scientific disciplines. Recent tools made available include:

- **BBTools**, a suite of fast, multithreaded bioinformatics tools designed for analysis of DNA and RNA sequence data in common sequencing file formats such as fastq, fasta, sam, scarf, fasta+qual, compressed or raw, with autodetection of quality encoding and interleaving.
- **BOOST** (Build Optimization Software Tools for DNA Synthesis), a suite of software tools to streamline the design-build transition in synthetic biology engineering workflows.
- **HipMer**, the first end-to-end *de novo* genome assembler designed for extreme scale analysis via efficient parallelization. The single-genome assembly implementation is a high-performance parallelization and port of the **Meraculous** assembler: jgi.doe.gov/data-and-tools/meraculous.
- **HipMCL**, a high-performance parallel implementation of the Markov clustering algorithm for large-scale networks construction.

bit.ly/JGI-Data-Tools



Explore software licensing opportunities at Berkeley Lab's Intellectual Property Office.

ipo.lbl.gov