CSP FY2025 Call (status: CLOSED)

The JGI's Community Science Program (CSP) is now accepting **Letters of Intent** for large-scale genomic science projects that fall broadly within the DOE mission. Of particular interest are proposals that address the following areas of emphasis and exploit the diversity of JGI capabilities:

I. Genes to Function

Today's ability to generate sequence data far outpaces the rate at which gene functions can be validated. Gene annotation tools are becoming more sophisticated but still rely on experimental characterization for comparative analyses. Proposals building a broader understanding of gene function in plants and microorganisms, especially genes of unknown function conserved across diverse organisms, are encouraged to improve gene and genome annotations for energy and environment applications. Characterizations of metabolic and functional capabilities of uncultivated microorganisms of high phylogenetic diversity are of particular interest. Projects may include:

Annotation of gene function using a combination of advanced computational analyses, functional genomics, DNA synthesis, and/or metabolomics.
Generation of genome-wide CRISPR gRNA libraries that enable gene function characterization.
Improvements to predictively design biosystems that combine large-scale sequencing, metabolomics, data integration, and DNA synthesis.
Characterization of genes of unknown function conserved across groups of organisms, such as fungi,
Characterization of secondary metabolite biosynthetic gene clusters, their products and the role of these products in environmental processes, spanning (meta)genome mining, computational predictive analyses, elucidation of regulatory networks, cluster expression, metabolomics and functional studies.
Characterization of metabolites, including novel environmental metabolites,for linkage to ecological processes and genomic analyses. For example, exometabolite impacting soil carbon cycling and studies of gene expression, enzymes, mutants, etc. to provide direct biochemical evidence of gene function.

Function-driven microbial single-cell genomics and metagenomics, e.g. sequencing of stable isotope-labeled DNA or selectively sorted cells or populations to assign functional roles to populations within communities. Population-focused studies that evaluate gene function and diversity within given species.

II. Plant and Algal Functional Genomics

The JGI has produced several "flagship plant genomes" for plants of interest including *Brachypodium*, *Chlamydomonas*, *Physcomitrella*, *Miscanthus*, *Populus*, *Setaria*, *Glycine*, *Sorghum*, and *Panicum*. These genera are of special interest as potential biofuel feedstocks or as comparators to provide insight into feedstock evolution and phenotype. Projects that directly relate to these genomes are encouraged. For all plant proposals, priority will be given to multi-organism proposals that seek to 1) compare among plants and/or analyze plant-microbiome interactions, and/or 2) connect sequence to function in key arenas including carbon sequestration, mycorrhizal associations, secondary metabolism, or molecular foundations of the ecology of natural systems. For algal genomics, diversity-focused projects are encouraged. Projects of interest may fall into one of the following four categories:

1. *High quality de novo genomes & pan-genomes* – We invite proposals for whole genome sequencing of species using our advanced pipelines. The scientific focus can include comparative or evolutionary genomics studies with the JGI flagship species or clusters of comparative plant species that would inform fundamental plant biology and lead to new insights into plant gene function. Examples of projects might include multi-genotype, high-quality pan-genomes for comparison to diversity resources, abiotic stress-tolerant species, exudation, secondary metabolite producers, or a concentration of genomes in an under-sampled area of the plant phylogeny with compelling biology that is relevant for DOE's mission. For proposals that include non-flagship species, the relevance and application to DOE plant science should be clearly demonstrated.
2. *Large-scale germplasm resequencing* – We invite germplasm resequencing projects aimed at 1) understanding natural population variation of relevant genera/species, 2) creating a foundation for large scale GWAS projects for gene discovery, 3) sequencing mutant collections to create JGI-supported community resources for functional genomics, or 4) developing views of the pan- and core genomes to determine a complete picture of gene content within a relevant genus/species.
3. *Comparative transcriptomes and functional assays* – Proposals are encouraged that expand experimental conditions for JGI-sequenced plant genomes with an emphasis on plant abiotic stress responses, useful biological properties (e.g. bioaccumulation, biosynthesis of useful or unusual chemicals, etc.) and extend these functional studies beyond straightforward transcriptomics (e.g.,

metabolomics). Proposals that use phylogenetics (e.g. evolution of biosynthetic pathways), temporal dynamics, and multiple assays to create data that can be used to find associations and create networks to identify candidate genes, pathways, and regulatory sequences are encouraged(see core capabilities below for available assays).

4. *Algal genomics* – Algae are important primary producers with tremendous diversity, long evolutionary history, and exceptional potential for DOE science and applications. Significant and rapid advances in the fundamental knowledge of algal biology, the entire biomass-to-bioenergy supply chain, and algal cultivation strategies are dependent on genetic, biochemical and phenotypic information which is currently lacking. Proposals are encouraged that will expand genomic knowledge across algal diversity, that will build fundamental knowledge of algal metabolism and physiology, and which will provide insights into algal associations with other microbes and viruses as they relate to production of biofuels and bioproduct research.

III. Inter-organismal interactions

A key focus for the JGI is understanding the mutualistic, competitive, or antagonistic interactions among bacteria, archaea, fungi, algae, protozoa, plants, and viruses, as these in turn impact global biogeochemical cycles of carbon and other nutrients, as well as plant health. Proposals aimed at characterizing secondary metabolite biosynthetic pathways in plants and/or associated microbes and exometabolites mediating ecological interactions are specifically encouraged, as are hypothesis-driven projects deciphering functional and phylogenetic changes of natural or synthetic communities upon manipulation of the host and/or host environment. We encourage applicants to consider using the DOE Systems Biology Knowledgebase (KBase) to model interactions. Projects that could address this focus include:

Study of bipartite or multi-partite interactions involving plants, fungi, algae, bacteria, archaea, protists and/or viruses in DOE mission-relevant systems
Multi-omics-enabled interrogations of symbioses (non-animal associated) from terrestrial, freshwater, estuarine, and intertidal environments, and including symbiosis model systems.
Investigation of the genomic basis of microbial mutualism and microbe-microbe interactions in stable model communities, e.g. enrichment cultures or synthetic communities.
Investigations using EcoFAB devices (https://eco-fab.org/) supplied by the JGI to conduct experiments to uncover the mechanisms underlying the interactions between plants and their root microbiomes.
Chemical ecology of metabolites: Integration of large-scale genome sequencing, genome mining and predictive analyses, with functional activation and exometabolomic analysis including of secondary metabolites to infer function in

inter-organismal interactions, including mechanisms of resource competition and cross-feeding.
Multi-omics-based and/or single-cell enabled investigation of coupled virus-microbe and/or virus-virus dynamics, activity, and/or evolution.
Genome-wide CRISPR gRNA libraries combined with specific genetic tools (e.g., CRAGE) that enable functional characterization of genes involved in plant root colonization and plant growth promotion.
Exploration of viral molecular mechanisms for host cell take-over and metabolic reprogramming using advanced computational analyses, functional genomics, DNA synthesis, and/or metabolomics.
Metabolomic analysis of plant exudates, soil organics, environmental metabolite dynamics, and microbial use and production of exometabolites.

IV. Understanding natural communities important for carbon storage, nutrient cycling, and climate change

Natural communities, made up of interacting species of plants, fungi, algae, bacteria, archaea, protists and viruses, play a critical role in carbon cycling, including cabon sequestration or production of carbon dioxide and methane. Understanding these communities could bring insight into approaches for measuring and mitigating the climate impact of human activities such as farming (including enteric fermentation in ruminants), and forest management. While progress has been made in increasing understanding of the individual impacts of specific species, a holistic and systematic understanding of the function of interacting hosts (i.e. plants, fungi) and microbes in important environments such as forests, agricultural lands, grasslands, soils, terrestrial/aquatic interfaces, and rumen is lacking. Similarly, our understanding of how exometabolites mediate ecological and environmental processes such as carbon sequestration is limited to relatively few well-characterized molecules. Proposals are encouraged to develop integrative genomic resources (e.g., combining metagenomics, metatranscriptomics and/or metabolomics, as well as isolate and/or single cell genome sequencing) focused on these communities to enable continued investigation and increase understanding of how hosts, microbes, and metabolites influence the carbon cycle.

Bacteria, archaea, fungi, algae, protozoa, and viruses are also directly involved in nutrient and other biogeochemical cycles that impact long-term climate processes. While a nascent understanding of nutrient and biogeochemical cycling in marine environments exists, understanding of the involvement of plants and microbes in these complex cycles and processes in terrestrial, freshwater, and coastal environments (here defined by the area within 185 km of land) has lagged behind. Proposals are encouraged that will provide insight into plant and microbial interactions and activities controlling cycles of nitrogen, phosphorus, sulfur, iron, and other micronutrients from a broad range of terrestrial, freshwater, and coastal environments (including terrestrial-aquatic interfaces such as peat bogs, wetlands, marshes, and hyporheic zones) and the abiotic and biotic controls on the

dynamics of exometabolites. In addition, developing multi-omics datasets to enable modeling of regulatory and metabolic processing of these elements in model microbes and microbial systems is encouraged. Similarly, advancing our knowledge of fermentation and methanogenesis in the rumen will be imperative for any efforts to mitigate climate change in the next 30 years.

Time-series genomic and metabolomic studies of microbial communities have provided unprecedented insights into microbial community structure and ecosystem function, and stability. Such data offer a foundation for the modeling of community behaviors under DOE-relevant scenarios, such as climate change. Groups such as the Long Term Ecological Research Network (LTER), which is focused on studying ecological processes over significant temporal and spatial scales at twenty-eight LTER sites covering diverse ecosystems, provide unique access to temporal sampling. To enable time-resolved microbial community genomic, transcriptomic, and metabolomic analysis, proposals leveraging temporal samples from LTER, National Ecological Observatory Network (NEON), or other such sites and resources, are specifically encouraged.

V. Biofuels, biomaterials, and bioproducts

Replacement of petroleum with biological feedstocks for the production of biofuels and bioproducts is critical for energy security and climate stability. Proposals are encouraged that are aimed at characterizing biological processes (including those novel pathways generated by synthetic biology approaches) that are relevant to biofuel, biomaterial, and bioproduct generation, and connecting these processes to omics-based analyses for DOE-relevant plants, microbes, viruses and microbial communities. Relevant biological processes include biosynthesis and deconstruction of plant biomass, especially lignocellulose, and production of metabolites that are precursors of biofuels, biomaterials, and/or non-pharmaceutical bioproducts. Specific topics include the discovery and characterization of enzymes and metabolic pathways for polymer breakdown and/or conversion to novel products, the microbially mediated construction and deconstruction of plastics, secure synthetic biology, genome-enabled material synthesis, and investigations into organisms and/or biological products involved in plant-microbial interactions that impact biofuel and bioproduct feedstock productivity.

VI. Viruses, mobile elements, and new lineages of life

The JGI has recently identified two research areas, Virus EcoGenomics and Applications (“VEGA”) and New Lineages of Life (“NeLLi”), as critical to lead to a more comprehensive understanding of microbiome processes and enable innovative research in the areas of bioenergy and biotechnology. Here, “VEGA” area encompasses all viruses of microbes including bacteriophages, archaeal viruses, and eukaryotic viruses, i.e. viruses of protists, algae, and fungi, as well as other similar mobile genetic elements such as plasmids. “NeLLi” area covers new and uncultivated microbial taxa across bacteria, archaea, and microeukaryotes, including novel phyla and other underexplored lineages. Proposals leading to a more systematic and contextualized understanding of the diversity, distribution,

function and dynamics of these elements and microorganisms are encouraged. Projects may include:

- Large-scale targeted ‘omics exploration of specific elements/organisms within an ecosystem, e.g. metagenomics or single cell genomics targeting viral, plasmid, or ultra-small bacteria.
- Discovery, exploration, and functional characterization (*in silico* and/or *in vitro*) of novel genes encoded on VEGA and/or NeLLi elements/organisms, e.g. genes that can be linked to lifestyle and/or habitat transitions.
- Development of new approaches to recover high-quality genomes for uncultivated lineages within the VEGA and/or NeLLi area, especially from understudied ecosystems and/or organisms and elements underrepresented in current databases.
- Omics-enabled characterization of novel isolates or co-cultures, that expand existing culture collections with representatives of so-far uncultivated lineages.

Project Structure

CSP projects are expected to generate publicly available data that will answer important questions relevant to the target organism or environment, as well as provide a substrate for broader use by the DOE research community. CSP projects have historically provided a means for user communities to assemble and interact in collaborative ways. Proposals are encouraged that involve some or all of the following features: 1) a scale and complexity that exceeds the capacity of a single lab, 2) engaging a large group of collaborators, 3) requiring JGI capabilities that reach beyond genome sequencing, 4) generating data of high value to the scientific community, and 5) plans to analyze and distribute data and results through KBase and the National Microbiome Data Collaborative (NMDC).

All proposals may request up to 10 Tbp of sequence data. For multi-PI projects generating data of broad utility to the scientific community, requests of up to 20 Tbp will be considered. Larger Tbp totals (up to 50 Tbp) will be considered for plant Illumina resequencing and large-scale metagenome projects only, but such proposals will be evaluated separately with the anticipation that only 1 or 2 would be approved. Requests for Pacific Biosciences long-read sequencing are capped at 1Tbp and 50 samples (up to 200 samples for bacterial/archaeal isolate genomes), but larger requests will be considered for multi-PI projects of high value to the scientific community.

Proposals may request DNA synthesis up to 500 kbp of synthesized DNA, and up to 1,500 kbp of synthesized DNA for consortium proposals with co-PIs from at least 3 different institutions.

Requests for metabolomics are capped at 200 samples for polar metabolite analysis and 500 samples for nonpolar metabolite analysis (e.g., lipids and secondary metabolites). Larger requests may be considered for a very limited number of multi-PI projects of high

value to the scientific community, at JGI's discretion and contingent to successful scientific and technical review.

Requests for DAP-seq should include a minimum of 92 transcription factors, and ideally consist of an even multiple of 92 transcription factors up to a maximum of 500kb total. This is a 96-well plate based protocol where 4 wells in each plate are reserved for negative controls.

For EcoFAB experiments, up to 50 EcoFAB can be requested.

The JGI provides extensive data analysis pipelines. Applicants should present a plan for all data analysis that may be required beyond these standard pipelines. Users are encouraged to consider and describe in their proposal how KBase may be used to advance their analyses and data management/sharing plans. KBase integrates a variety of community data and analysis tools into an easy-to-use platform that leverages scalable computing infrastructure to perform sophisticated systems biology analyses, while also creating reproducible analysis workflows for sharing and publication. Additionally, applicants interested in collaborating with the National Microbiome Data Collaborative (NMDC) to develop project plans aligned with making their data findable, accessible, interoperable, and reusable (FAIR) should indicate so in their proposal.

JGI Capabilities

All proposals should justify why JGI capabilities are critical to success. The JGI employs an evolving suite of sequencing platforms, currently consisting of short read Illumina (NovaSeq X) as well as single molecule long-read Pacific Biosciences technology (Revio). The capabilities available for this call are listed below. While individual proposals may draw from one or more of these capabilities as needed to fulfill project goals, within the overall cap, the final scope is ultimately at the discretion of the DOE JGI. Successful projects frequently utilize a combination of capabilities listed below; more details can be found here.

Core Capabilities include:

De novo sequencing of fungal, algal, bacterial, archaeal, viral (including giant viruses), and plant genomes.
Resequencing for variation detection.
Metagenomes and metatranscriptomes, i.e. microbial and/or viral community shotgun DNA/RNA sequencing (no amplicon sequencing of 16S rRNA or other genes).
Prokaryotic whole genome DNA methylation analysis via PacBio sequencing.
Transcriptome analysis including coding transcript annotation and expression profiling.
In vitro transcription factor binding site mapping by DNA affinity purification sequencing (DAP-seq). DNA/gene synthesis should also be requested for construction of affinity-tagged transcription factor clones used in the assay.

Stable Isotope Probing (SIP) metagenomic studies, including density centrifugation and fraction collection carried out by the JGI.
Fluorescence activated cell sorting for targeted metagenomics, single-cell genomics (bacteria and archaea only) and microbial aggregates.
Flow cytometric sorting and genomic analysis of metabolically active microbes labeled via Bio-Orthogonal Non-Canonical Amino acid Tagging (BONCAT).
DNA/gene synthesis linked to sequence data generation, including codon optimization, refactoring, and assembly of biosynthetic pathways into appropriate vector systems for expression in native or heterologous hosts. This also includes synthesis of combinatorial pathway libraries. More details
Whole-genome or partial-genome CRISPR-based gRNA library construction and QC.
Mass spectrometry-based metabolomics and exometabolomic analysis of primary and secondary metabolites from plants, microorganisms, and environmental samples (e.g. soil metabolomics, etc) using targeted and untargeted approaches.
Analysis pipelines for the datasets above.

The JGI also has limited capacity for the following developing or resource-intensive capabilities. *Since these are emerging technologies, proposers should contact JGI staff for technical details when developing proposals.*

Custom analysis of JGI datasets.
EcoFAB pilot projects. The JGI can provide a limited number of EcoFAB devices (https://eco-fab.org/) to study plant-microbiome interactions. These devices allow for non-destructive root imaging and sampling of the growth media while maintaining a sterile environment. In addition, the JGI can provide a standardized defined microbial community that colonizes plant roots and *Brachypodium* germplasm, if desired. Users would conduct experiments using these resources and return samples to the JGI for analysis by existing JGI capabilities e.g., metabolomics, transcriptomics and metatranscriptomics.
Long-read (PacBio) metagenomes.

Mechanism of Review

Letters of intent will only be accepted electronically and should be submitted at https://proposals.jgi.doe.gov/. The CSP Call is open to anyone with the understanding that CSP data are made publicly available one year after completion, without exception. Applicants will be advised approximately six weeks prior to the proposal submission deadline whether to prepare a full proposal. Guidance for submitting full proposals will be included in the email notification to invited applicants. A full schedule is below.

Proposals will be independently peer-reviewed and ranked following given review criteria. Final decisions will be made by JGI senior management with final approval given by DOE

program management. All projects will begin as soon as User Agreements are finalized, targeted for October 2024.

Contacts

For general questions, please contact Christa Pennacchio, Project Management Office. For questions about the appropriateness of projects, program specifics, or experimental design, please contact Tanja Woyke, Deputy for User Programs.

Proposal Schedule

To respond to the annual CSP call, a Letter of Intent (LOI) is required before submitting a proposal. LOIs will only be accepted electronically and should be submitted at https://proposals.jgi.doe.gov/. Guidance for submitting full proposals will be included in the email notification to invited applicants.

The full FY25 schedule is below:

Begin accepting Letters of Intent	February 19, 2024
Letters of intent received	April 18, 2024
Invitation of proposals	May 15, 2024
Proposals received	July 2, 2024
Technical and scientific review	week of August 12, 2024 (exact dates TBD)
Approval and rejection notices sent	by September 13, 2024
Projects start	October 1, 2024, or as soon as the User Agreement is finalized