

U.S. Department of Energy
Office of Science
Office of Biological and Environmental Research

ACTION: Recommendations for Sequencing Targets in Support of the Bioenergy Mission of the Office of Biological and Environmental Research (OBER)

SUMMARY: This **Notice** seeks the input and nominations of interested parties for candidate microbes, microbial consortia, and 100Mb-or-less-sized organisms for draft genomic sequencing in support of Office of Biological and Environmental Research (BER) Bioenergy Program element within its Genomics:GTL Program. **Nominated candidates must be relevant to the BER mission to determine genomic sequences of microorganisms involved in energy production (particularly conversion of lignocellulosic material to ethanol, or hydrogen, or other biofuels).** Other Genomics:GTL missions, e.g., environmental processes including waste remediation, and carbon management either in the oceans or terrestrially **will not** be considered under this Notice. **Relevance to the Bioenergy mission will be an explicit criterion for review.** This announcement is not an offer of direct financial support for research on these organisms. Those nominations selected will result in the DNA sequence of the prioritized organisms being determined at a draft level (typically ~8X coverage) at the Production Genomics Facility (PGF) at the DOE's Joint Genome Institute (DOE-JGI), (<http://www.jgi.doe.gov>). It is the intention that most or all of the selected single organisms will have their genome sequences finished. This announcement is designed to assist BER in determining and prioritizing a list of microbes, microbial consortia, or modest-genome sized (not more than 100Mb) organisms (including eukaryotes) that address the BER mission of identifying and sequencing genomes from bioenergy relevant microbes. Following merit review, and a determination of satisfactory programmatic relevance, draft sequencing will be carried out at the PGF. **Henceforth, this announcement will be incorporated within the DOE-JGI's Community Sequencing Program (<http://www.jgi.doe.gov/CSP/>), i.e., this announcement will not be repeated in future years.**

DATES: To assure consideration, nominations in response to this notice should be **received by 4:30 p.m. (E.D.T.), August 10, 2006**, to be accepted for merit review. It is anticipated that review will be completed early in the fall of 2006 with draft sequencing at the DOE-PGF to commence early in 2007, conditional upon the provision of high quality DNA.

ADDRESSES: Nominations responding to this notice should be sent to Dr. Daniel W. Drell, Office of Biological and Environmental Research, SC-23.1, Office of Science, U.S. Department of Energy, 1000 Independence Ave., SW, Washington, DC 20585-1290; email is acceptable and encouraged for submitting nominations using the following address: bioenergy.microbes@science.doe.gov .

FOR FURTHER INFORMATION CONTACT: Dr. Daniel W. Drell, SC-23.1, Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy, 1000 Independence Ave. SW, Washington, DC 20585-1290, phone: (301) 903-4742, email: daniel.drell@science.doe.gov . The full text of this notice is available via the Internet using the following Web site address:
http://www.sc.doe.gov/ober/bioenergy_microbes_call.pdf

SUPPLEMENTARY INFORMATION: The DOE Office of Biological and Environmental Research supports fundamental research for a variety of missions (http://www.science.doe.gov/ober/ober_top.html) among them the Genomics:GTL Program, the Climate Change Research Program, the Terrestrial Carbon Processes Program, the Environmental Remediation Sciences Program (ERSP), and the Program for Ecosystem Science. An additional program of relevance is the Biofuels element of the Biomass Program (http://www.eere.energy.gov/RE/bio_fuels.html) of the Office of Energy Efficiency and Renewable Energy. The purpose of this request for nominations is to support one key BER mission by generating and providing initial analyses of microbial DNA sequence information to further the understanding and application of microbiology relating to energy production, with emphases on biological degradation of biomass (e.g., lignocellulose) and resulting generation of ethanol, hydrogen, or other biofuels. Microbial sequence information may come from individual microbes or microbial communities or other single-cell organisms with individual genomes not more than 100Mb in size.

For more than ten years, sequencing of a variety of microorganisms that live in diverse environments has provided a considerable information base for scientific research related not only to DOE missions, but also to other federal agency missions and U.S. industry. (<http://www.jgi.doe.gov>; <http://www.microbialgenome.org>; <http://www.genomesonline.org>, <http://DOEGenomestolive.org>) Nonetheless, most of our current knowledge of microbiology still is derived from individual species that either cause disease or grow easily and readily as monocultures under laboratory conditions and are thus easy to study. The preponderance of species in the environment remains largely unknown to science. Many are thought to grow as part of interdependent consortia in which one species supplies a nutrient necessary for the growth of another. Little is known of the organization, membership, or functioning of these consortia, especially those involved in bioenergy generating processes of BER interest.

For this notice, groups of microbes that may comprise consortia or include species refractory to laboratory culture that play important roles in the generation of energy compounds (e.g., ethanol, hydrogen or other biofuels) may be proposed. The candidate(s) being proposed must mediate or catalyze metabolic events that contribute to these processes.

For this notice, candidate organisms (either individual organisms, consortia of organisms, or eukaryotes with small genomes) comprised of archaea, bacteria, fungi, algae, protists, and other eukaryotes with genome sizes not greater than 100 Mbp can be proposed for draft sequencing. **For all candidates proposed for genomic sequencing, the amount of sequencing and the degree of sequence coverage being requested must be explicitly**

stated. For a current list of microbes that have been and are being sequenced by BER, see <http://www.ornl.gov/microbialgenomes/organisms.html> and http://genome.jgi-psf.org/mic_curl.html. For a more comprehensive list of genomes (including microbes) being sequenced in the public sector, see <http://www.genomesonline.org>. **Each nomination should begin with a statement of the case for the relevance of the nominated organism(s) for bioenergy production.** This should encompass the nature and forms of the energy produced biologically by the organism(s) and how knowledge of the sequence would enable or accelerate advances in understandings of biological energy generation. The nomination should then proceed to address the criteria listed below, either for single organisms or consortia:

Aims: This request for nominations of candidate sequencing targets has two broad foci:

(1) **Single organisms.** These may be bacteria, archaea, fungi, microalgae or multicellular organisms with genomes not larger than 100Mb. The criteria that will be used to evaluate proposed candidates for draft sequencing will include:

- (a) The candidate has significant relevance to the BER mission noted above;
- (b) To assess suitability for whole genome shotgun sequencing, preliminary data on genome size, repeat content, genome structure, GC content, polymorphism, and other characteristics are provided, especially for larger genomes;
- (c) The source of genomic DNA (i.e., strain or isolate, and researcher) is identified, and a clonal sample (or samples with low and characterized polymorphism) is available;
- (d) A brief description of post sequencing follow-up work (e.g., a data use plan and how will data be annotated to enable rapid and open use) is included, although funding will not be provided for these activities;
- (e) The availability of a DNA/gene transfer system supporting genetic analyses is known;
- (f) Biological novelty or uniqueness (i.e., unusual genetically determined characteristics pertinent to BER missions) is described;
- (g) Place in the currently understood, 16s RNA based, “Tree of Life” is identified, e.g., is the proposed candidate in a sparsely populated or more heavily populated section of the tree?
- (h) A brief description of the scientific user community for the sequence information, and for the organism in general, is given;
- (i) The potential impact on science is large;
- (j) Explicit commitment to a data-release schedule, consistent with the guidelines given below is provided.
- (k) Explicit commitment to the deposition of a culture of the proposed microbe to a publicly accessible repository is made.

(2) **Environmental consortia.** The review criteria that will be used to evaluate proposed candidates for draft sequencing will include most of the criteria listed above for single organisms (with less emphasis on genome size/structure, presence/absence of a genetic system, or position in the “Tree of Life” since it is recognized that few data on these attributes will be available), but in addition, the following considerations will be included:

(a) Some measure of the “complexity” of the target consortium being proposed, e.g., approximate number of species, size(s) of genomes, and proportions of different members (it is understood that in most cases, only estimates of these parameters may be available) is discussed. When the environmental consortia are complex, approaches should be described to normalize the DNA libraries in order to reduce the amount of sequencing required and assure adequate sampling of the complexity of the consortia. Additionally, the proposer(s) should be prepared to work together with JGI scientists to optimize the yield from the sequencing effort required;

(b) Past attempts to cultivate consortium members are described, e.g., have any members of this consortium been successfully cultured;

(c) Some spatial/temporal/hydrochemical/geochemical or other characterization of the environment is given, e.g., the physicochemical parameters of the site from which the selected community is derived; a description of the site contaminants; the accessibility of the site for future sampling; the adequacy of site documentation;

(d) If proposed, technical approaches and technology development specific for defining and isolating members of a given consortium are described;

(e) Some indication of the biological function of the relationships, within consortium members where available, along with a discussion of the scientific and programmatic importance of understanding these relationships is given;

(f) Information where available is given about the phylogenetic position(s) of the members of the consortium and what is known about closely related organisms.

(g) Available informatics tools and annotation plan (e.g., for annotating genes from a consortium analysis or grouping identified genes into a putative “consortium phenotype” within the chosen environment) are described;

(h) Explicit commitment to a data-release schedule, consistent with the guidelines given below is provided.

With respect to access to sequencing data, the usual and customary practice for the DOE-JGI is to put all sequencing data up on its web site (<http://www.jgi.doe.gov/>) at frequent and periodic intervals with deposition in a publicly available database within a maximum of one year. However, for the purposes of this notice, BER’s policy is that after **no more than 3 months** from the completion of a “first assembly” of the sequence for a single-genome microbial project, the data will be released on the DOE-JGI web site or to a publicly accessible database with no use restrictions (see <http://www.jgi.doe.gov/sequencing/collaborators/datarelease.html>.) For microbial community projects, the DOE-JGI will conduct normal QA/QC assessments on the sequence output (at approximately 2 x coverage), then review with the proposer(s) and with BER staff the amount of requested sequencing to achieve a satisfactory genomic “view” of the selected microbial community. From the time of initiation of this discussion, not more than 3 months will be permitted to elapse before unconditional release of these data. Proposers should clearly understand that the priority in the sequencing queue that a selected project is given may be linked to the willingness of the proposer(s) to shorten this “embargo” period. BER stresses its general commitment to the rapid release, without any use restrictions, of sequence data into publicly accessible databases.

Upon selection of a nominated microbial sequencing target, BER expects that Principal Investigators will collaborate with the DOE-JGI by providing high quality, high molecular weight genomic DNA for library construction as well as assisting in annotating the draft sequence data until a sufficiently complete annotation is achieved, understanding that this will be sensitive to hypothetical gene predictions and the usual uncertainties of annotation. (See Information for Collaborators, <http://www.jgi.doe.gov/sequencing/collaborators/index.html>.) Notification of successful proposals will take place sometime in October 2006 at which time the successful proposer is expected to work out a detailed and mutually satisfactory User Agreement with the DOE-JGI (see: <http://www.jgi.doe.gov/sequencing/collaborators/>, click on “sample_agreement.doc”). This User Agreement will specify the expectations of the proposer and the commitment of DOE-JGI resources to the project, including the amount of sequencing capacity or megabases to be sequenced. Sequencing will NOT begin prior to the finalization of this User Agreement. Thus proposers are strongly encouraged to begin preparation of DNA BEFORE notification of successful proposals because from the date of the conclusion of the User Agreement, the proposer is expected to furnish the DOE-JGI with usable, high quality DNA within 6 months or risk losing her/his place in the sequencing. A separate communication with the detailed requirements for DNA will be sent to proposers whose nominations are accepted for sequencing. Following data acquisition and annotation, DOE expects that Principal Investigators whose nominations have been sequenced will make good faith efforts to publish the results of their subsequent work in the open scientific literature, including both the genome sequences of the organisms sequenced under this notice as well as the annotation. BER also expects the Principal Investigator of a selected effort to either deposit a culture of the microbe or consortium into a publicly accessible collection or repository, or make it available directly so others can have access. These parties are encouraged to create process- and cost-effective partnerships that will maximize data production and analysis, data dissemination, and progress towards understanding basic biological mechanisms that can further the aims of this effort. Additionally, it must be explicitly understood that DOE-JGI will provide an assembled and computationally annotated draft (roughly 8X; carried out in a paired-end sequencing approach) sequence of the microbe(s) selected, but that research using that sequence data should be funded from separate sources and/or separate solicitations. Finally, there is no commitment to finish a given drafted sequence, although it is the intention of the DOE-JGI to carry each single-genome project through to completion.

Submission Information: Interested parties should submit a brief white paper to either of the two categories (single microbes or environmental consortia) described above, consisting of not more than 5 pages of narrative exclusive of attachments (which should be kept to a minimum) responding to each of the specific criteria set forth. Electronic submission to (bioenergy.microbes@science.doe.gov and daniel.drell@science.doe.gov) is strongly encouraged. It is expected that the Principal Investigator will serve as the main point of contact for additional information on the nominated microbe. Nominations must contain a very short abstract or project summary and a cover page with the name of the applicant, mailing address, phone, fax, and email. The nomination should include 2-page curriculum vitae of the key investigators; letters of intent (or e-mails) from collaborators

(suggesting the size of the interested community) are permitted. Finally, the proposer must indicate the biosafety level of the source culture (microbial or environmental) and whether the source culture has international, federal or state shipping or possession restrictions regardless of absence of restrictions on the purified nucleic acids.

Nominations will be reviewed relative to the scope and research needs of the BER programs cited above. A brief response to each nomination will be provided electronically following merit and programmatic reviews.

Other useful web sites include:

DOE-JGI Microbial Sequencing Priorities for FY2005:
<http://microbialgenome.org/organisms.shtml>

DOE-JGI Microbial Web Page— <http://genome.jgi-psf.org/microbial/index.html>

GenBank Home Page— <http://www.ncbi.nlm.nih.gov>

DOE Genomics:GTL— <http://DOEGenomestolife.org>

DOE EERE Biomass Program—Sugar Platform:
http://www1.eere.energy.gov/biomass/sugar_platform.html

DOE EERE Biomass Program: Ethanol --
<http://eereweb.ee.doe.gov/biomass/ethanol.html>

DOE EERE Biomass Program: Biofuels: http://www.eere.energy.gov/RE/bio_fuels.html

DOE National Renewable Energy Laboratory Biomass Program --
<http://www.nrel.gov/biomass/>

BER Genomics:GTL Roadmap Appendix A (Energy Security) --
http://doegenomestolife.org/roadmap/pdf/GTL05_05Energy.pdf

Making Bioethanol Cost Competitive Brochure (PDF) --
http://doegenomestolife.org/pubs/Bioethanol10_27_05_8.5.pdf