Title

US DOE Joint Genome Institute Microbial Sequencing: Genomes to Life Projects

Authors

Bruce, David
Brettin, Tom
Chain, Patrick
et al.

Publication Date

2006-02-01

LBNL-60336 Abs.

US DOE Joint Genome Institute Microbial Sequencing: Genomes to Life Projects

David Bruce[1]*, Tom Brettin[1], Patrick Chain[3], Cliff Han[1], Loren Hauser[5], Nikos Kyrpides[1], Miriam Land[5], Alla Lapidus[2], Frank Larimer[5], Jeremy Schmutz[4], Paul Gilna[1], Eddy Rubin[2], Paul Richardson[2]

[1]JGI-Los Alamos National Laboratory, [2]JGI-Production Genomics Facility and Lawrence Berkeley Laboratory, [3]JGI-Lawrence Livermore National Laboratory, [4]JGI-Stanford Human Genome Center, and [5]JGI-Oak Ridge National Laboratory

The US DOE Joint Genome Institute (JGI) sequences microbial and metagenomic projects through three main programs: DOE Microbial Genome Program (MGP), JGI Community Sequencing Program (CSP) and DOE Genomes to Life Program (GTL). The principle goal of the MGP is to fund sequencing projects related to DOE interests, the principle goal of the CSP is to fund sequencing projects from a broad range of disciplines that may not be covered in the MGP, and the principle goal of the GTL sequencing projects is to fund sequencing projects in direct support of the GTL program. The JGI is responsible for sequencing, assembling, annotating microbial genomes, and publishing sequence and annotation in GenBank and the DOE JGI Integrated Microbial Genomics web based system. The JGI has sequenced nearly 250 microbes and metagenomic samples to draft quality and completely finished over 70 microbes. Most microbial projects are targeted for finishing. The overall capacity is now approximately 100-125 microbial projects per year through draft sequencing and finishing. Virtually all microbial projects are sequenced by the whole genome shotgun method. To being the sequencing process, the Library group randomly shears the purified DNA under different conditions and selects for three size populations. Fragments are end repaired and selected for inserts in the range of 3kb, 8kb, and 40kb. These are cloned into different vector systems and checked for quality by PCR or sequencing. The libraries are sequenced by the Production group to approximately 8.5X coverage. The resulting reads are trimmed for vector sequences and assembled. The assembly is quality checked, automatically annotated by the Annotation group, and released to the collaborating PI as the initial Quality Draft assembly. For finishing, the draft assembly is assigned to a Finishing group. The Finishing group closes all sequence gaps, resolves all repeat discrepancies, and improves all low quality regions. The final assembly is then passed to the Quality Assurance group to assess the integrity and overall quality of the genome sequence. The finished sequence then receives a final annotation and this package is used

as the basis for analysis and publication in GenBank and the DOE JGI Integrated Microbial Genomics web based system. The JGI is made up of affiliates from a number of national laboratories including Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Los Alamos National Laboratory, Oak Ridge National Laboratory, and the Stanford Human Genome Center.

This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231 and Los Alamos National Laboratory under contract No. W-7405-ENG-36.