DNA Synthesis and Assembly Pipeline

Samuel Deutsch[1], Sarah Richardson[1], Angela Tarver[1], Matthew Hamilton[1], David Robinson[1], Sangeeta Nath[1], Matthew Bendall[1], Miranda Harmon-Smith[1,3], Lisa Simirenko[1], Nathan J Hillson[1,2], and Jan-Fang Cheng[1]

1 Department of Energy Joint Genome Institute, Walnut Creek, CA
2 Joint BioEnergy Institute, Emeryville, CA
3 Lawrence Livermore National Laboratory, Livermore, CA

[a]*To whom correspondence may be addressed. E-mail:* sdeutsch@lbl.gov

March 25, 2013

ACKNOWLEDGMENTS:

The work conducted by the US Department of Energy (DOE) Joint Genome Institute is supported by the Office of Science of the DOE under Contract Number DE-AC02-05CH11231. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government, or any agency thereof, or the Regents of the University of California.

DISCLAIMER:

[**LBNL**] This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California

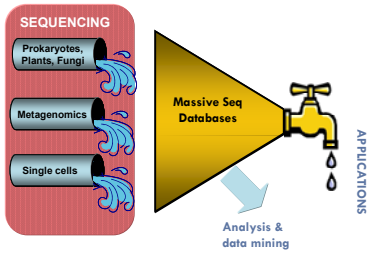
[**LLNL**] This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

DNA Synthesis and Assembly Pipeline

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

Samuel Deutsch, Sarah Richardson, Angela Tarver, Matthew Hamilton, David Robinson, Sangeeta Nath, Matthew Bendall, Miranda Harmon-Smith, Lisa Simirenko, Nathan J Hillson, and Jan-Fang Cheng

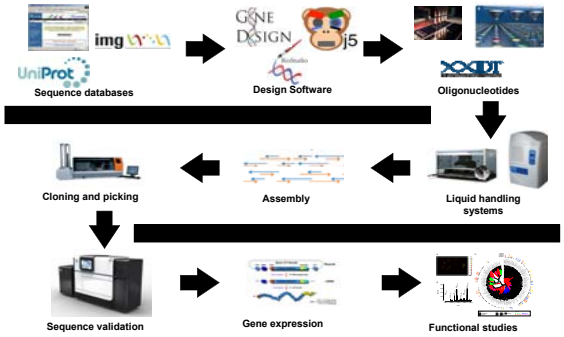
Adding Value to Sequence Information



Sequence information specially from metagenomes and single cell genomes is essentially **digital information** that cannot be easily accessed for biological hypothesis testing. This represents a limiting factor in the development of sequence based applications.

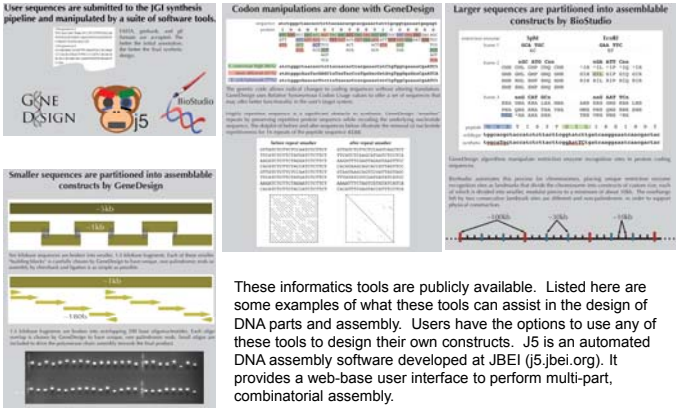
DNA synthesis allows digital sequence information to be 'translated' into biology by generating DNA in a template-independent manner.

DNA Synthesis Workflow at JGI



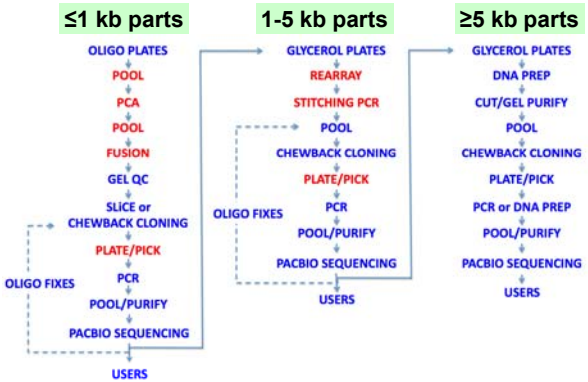
There are many steps involved in the DNA synthesis and assembly. We have implemented a suite of informatics tools to assist the design of oligos, the assembly of larger constructs, and potential strategies for combinatorial libraries. For the lab processes we have utilized robotic workstations, next-generation sequencing, and novel enzymology solutions to significantly improve the scale and reduce the cost.

Informatics Tools for Designing Oligos and Assemblies



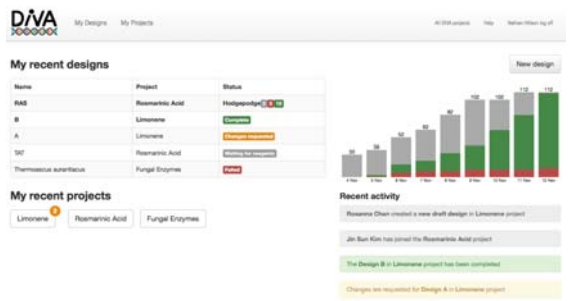
These informatics tools are publicly available. Listed here are some examples of what these tools can assist in the design of DNA parts and assembly. Users have the options to use any of these tools to design their own constructs. J5 is an automated DNA assembly software developed at JBEI (j5.jbei.org). It provides a web-base user interface to perform multi-part, combinatorial assembly.

Laboratory Processes



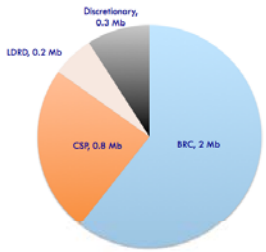
The laboratory processes begin with the construction of basic 1Kb DNA parts. Typically 48-96 basic parts can be synthesized in a week by a single FTE. Several robotic workstations are applied to handle the labor intensive steps shown in red. The basic parts are then assembled to larger size constructs in a similar process.

Synthesis Design and Status Reporting Platform



We plan to integrate the DIVA (Design, Implementation, Validation Automation) platform developed at JBEI with the DNA synthesis pipeline. This will provide a web-based graphical user interface for external users of the JGI's DNA Synthesis Program

Synthesis Allocation FY13

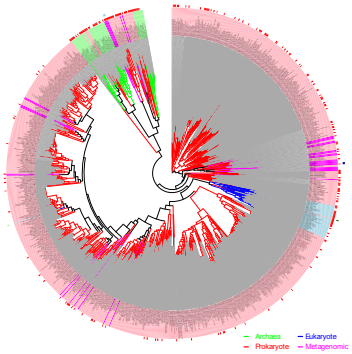


Examples of Proposed Genes

- Glycosyl Hydrolases
- Polyketide Synthases
- Plant Transcription Factors
- Lignin Peroxidases
- Cyanobacterial Carboxysomes
- Mevalonate Pathway
- Nitrogenases

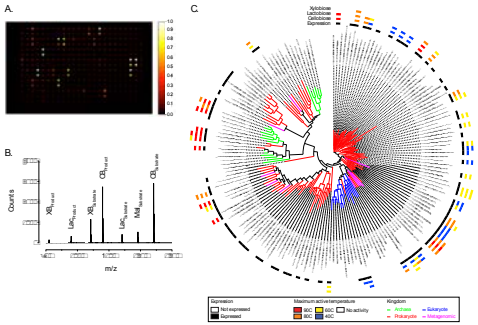
The FY13 synthesis capacity will be used primarily to support Bioenergy Research Centers (BRCs) and our Community Sequencing Programs (CSPs). Additional capacity will be used to develop proof-of-principle projects.

Systematic Survey of GH1 Enzyme Diversity



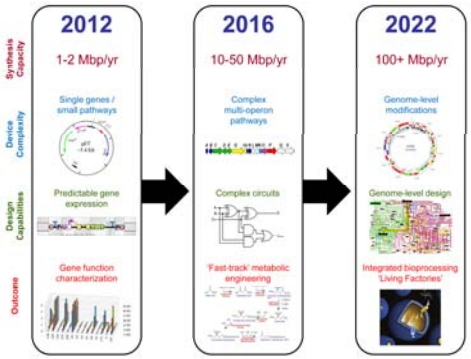
Our initial demonstration project was to functionally characterize the natural diversity of GH1 enzymes. GH1s participate in the last step of biomass breakdown (cellobiose to glucose). We identified GH1 candidate genes from NR, CAZy and JGI metagenomics databases, performed multiple sequence alignments, constructed phylogenetic trees and selected 200 representatives that cover the maximum sequence space (selected GH1 genes are represented by red dots on tree outer circle) .

Synthesis and Biochemical Characterization



We have synthesized and over-expressed our 200 GH1 candidate genes. Initial biochemical characterization at different temperatures and pHs has revealed a wide range of functional properties, validating our phylogenomics approach.

Synthetic Biology Vision at JGI



Synthetic biology approaches combined with large scale genomics information, will result in novel sequence driven applications. This will require substantial output and cost improvements as well as increasingly complex regulation and circuitry tools.

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.