LBNL-59003 Abs

Comparison of 454 Sequencing Platform with Traditional Sanger Sequencing: a Case Study with de novo Sequencing of Prochlorococcus Marinus NATL2A Genome

Feng Chen, Joseph Alessi, Edward Kirton, Vasanth Singan, and Paul Richardson

Feng Chen, Ph. D.
fchen@lbl.gov
(925) 296-5733
Genomic Technologies
DOE Joint Genome Institute
2800 Mitchell Dr. B400
Walnut Creek, CA 94598

The US DOE Joint Genome Institute (JGI) is a high-throughput sequencing center involved in a myriad of sequencing projects. A major effort at JGI is the sequencing of microbial genomes of relevance to the DOE missions of carbon sequestration, bioremediation and energy production. The JGI Microbial Program is responsible for the generation of over 200 microbial genomes and we are interested in utilizing new technologies to increase capacity. The 454 sequencing platform is an integrated system of emulsion-based PCR amplification of hundreds of thousands of DNA fragments linked to high throughput parallel pyrosequencing in picoliter-sized wells. The 454 sequencing platform can deliver over 20 million base pairs (mbp) from a single run, however, the quality of resulting assembly has not been well characterized. We sequenced *Prochlorococcus marinus* NATL2A (DNA kindly provided by Penny Chisolm and Claudia Steglich, Department of Biology, MIT) with one run on 454 platform, and generated over 36 mbp of data. The 332,387 reads of average length 109 bp were assembled with 454's Newbler assembly tool which generated contigs of consensus sequence. We compared the assembly results from the 454 platform with sequence previously finished at JGI using traditional Sanger sequencing technology. The methods were compared for error rates, misassemblies, coverage and depth. We will also discuss the feasibility of using the 454 platform for de novo whole genome shotgun sequencing and the optimal ratio of data from 454 and Sanger sequencing to achieve high quality finished genome sequences in a time and cost effective manner.

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.