LBNL-61436 Abs

Pyrosequencing Strategies for cDNA Libraries.

Feng Chen, Dean Ng, Edward Kirton, Jamie Jett, Erika Lindquist, and Paul M. Richardson

The US DOE Joint Genome Institute (JGI) is a high-throughput genomics facility involved in sequencing a variety of organisms. A major effort at JGI is the sequencing of genomes and microbial community samples of relevance to the DOE missions of carbon sequestration, bioremediation and energy production. cDNA/EST sequencing is an integral part of genomic sequencing because it provides crucial information for gene models and genome annotation. 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.

Several strategies have been designed and carried out at JGI to use the 454 platform for cDNA/EST sequencing. cDNA libraries constructed by conventional methods were subjected to direct 454 sequencing. In addition, special primers and adaptors were also designed for library construction so the directional sequencing feature of the 454 technology can be used to sequence a particular end of the cDNA/EST fragments. Adaptor sequences used by 454 library construction can be incorporated into polyT primer, cap primer and/or random primer for cDNA/EST library construction. The 454 sequencing platform can deliver 200 to 400 thousand cDNA/EST reads from a single run and does not require cloning step, potentially improving the coverage obtained through traditional Sanger sequencing. The large numbers of short reads generated by the 454 platform can be aligned to genome assemblies to extend and confirm gene models. Results from different strategies of library construction combined with 454 sequencing will be presented. The coverage of the library and the novelty rate are compared with traditional Sanger sequencing. The possible assembly problems caused by short reads with slightly higher error rate from 454 will also be addressed.

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. DE-AC52-06NA25396.