

LBNL-59859 Abs

ASSEMBLER COMPLEMENTATION TOOL

Vasanth Singan, Stephan Trong, Eugene Goltsman, Alla Lapidus
Joint Genome Institute – Production Genomics Facility,
2800 Mitchell Dr, Walnut Creek, CA-94598

A plethora of Genome Assemblers exist, that use various techniques to generate the consensus. In most cases, no one specific assembler can give the ideal assembly. With the recent study on Assembler Comparison conducted at the JGI / PGF, we have showed that assemblers that perform well for certain genomes not necessarily perform in the same manner with all the projects. In order to facilitate the finishers with choosing the optimal assembly, we proposed an initial draft Assembly with more than one assembler so that the finishers can choose the best assembly to start with.

The Assembler Complementation tool is a graphical visualization program that will not only serve to decide the best assembly but also help to collectively use the best consensus for comparative finishing. This tool will aid in identifying misassemblies, gaps, and rearrangements in assemblies relative to one another. Gaps can be visually identified by comparing a fragmented assembly against a better assembled one and conversely, misassemblies can be identified by comparing a conserved assembly against a less stringent one.

The contigs of the two assemblies to be compared are blasted against one another and based on the reference genome selected, for each of the contig in the reference genome, hit from contigs of the query genome are shown in a tiling view format with a different color for each contig. Information regarding the region of misassemblies, hits, etc can also be obtained interactively.

This work was performed under the auspices of the US DOE of Science, Biological and Environmental Research Program, and by the University of California, LLNL under Contract No. W-7405-Eng-48, LBNL under Contract No. DE-AC02-05CH11231 and LANL under Contract No. W-7405-ENG-36.