Genome Puzzle Master (GPM): an integrated pipeline for building and editing pseudomolecules from fragmented sequences

2016
Posted by : webmaster
Posted on : 2016/6/19 6:10:00

Jianwei Zhang1,2,*, Dave Kudrna2, Ting Mu1, Weiming Li1, Dario Copetti2,3, Yeisoo Yu2,†, Jose Luis Goicoechea2, Yang Lei1 and Rod A. Wing2,3,*

Bioinformatics (2016)
doi: 10.1093/bioinformatics/btw370
First published online: June 17, 2016

Abstract

Motivation: Next generation sequencing technologies have revolutionized our ability to rapidly and affordably generate vast quantities of sequence data. Once generated, raw sequences are assembled into contigs or scaffolds. However, these assemblies are mostly fragmented and inaccurate at the whole genome scale, largely due to the inability to integrate additional informative datasets (e.g. physical, optical and genetic maps). To address this problem, we developed a semi-automated software tool—Genome Puzzle Master (GPM)—that enables the integration of additional genomic signposts to edit and build ‘new-gen-assemblies’ that result in high-quality ‘annotation-ready’ pseudomolecules.

Results: With GPM, loaded datasets can be connected to each other via their logical relationships which accomplishes tasks to ‘group,’ ‘merge,’ ‘order and orient’ sequences in a draft assembly. Manual editing can also be performed with a user-friendly graphical interface. Final pseudomolecules reflect a user’s total data package and are available for long-term project management. GPM is a web-based pipeline and an important part of a Laboratory Information Management System (LIMS) which can be easily deployed on local servers for any genome research laboratory.

Availability and Implementation: The GPM (with LIMS) package is available at https://github.com/Jianwei-Zhang/LIMS

Contacts: jzhang@mail.hzau.edu.cn or rwing@mail.arizona.edu

http://bioinformatics.oxfordjournals.org/ ... To8iVJz20F5De&keytype=ref