ALLPATHS-LG Genome Assembly and Hi-C mapping strategy

Hansheng Zhao

Abstract

Citation: Hansheng Zhao ALLPATHS-LG Genome Assembly and Hi-C mapping strategy. **protocols.io** dx.doi.org/10.17504/protocols.io.phmdj46
Published: 16 Apr 2018

Protocol

Quality control of WGS data using NGS QC Toolkit (version 2.3.3) and FastUniq (version 1.1)
Step 1.
SOFTWARE PACKAGE (Linux -)
FastUniq, 1.1
WGS Assembly: 1. Calculating insert size using Bowtie 2 (version 2.2.9) and Picard (2.7)
Step 2.
SOFTWARE PACKAGE (Linux -)
Picard, 2.7
WGS Assembly: 2. Constructing contigs and scaffolds using ALLPATHS-LG (version r52488)
Step 3.
SOFTWARE PACKAGE (Linux -)
ALLPATHS-LG, r52488
WGS Assembly: 3. Linking the scaffolds with SSPACE (version 1)
Step 4.
SOFTWARE PACKAGE (Linux -)
SSPACE, 1
WGS Assembly: 4. Filling the gaps with GapFiller (version 1.11) and GapCloser (version 1.12)
Step 5.
SOFTWARE PACKAGE (Linux -)
GapCloser, 1.12
Hi-C mapping: 1. Evaluating and qualifying Hi-C data with HiC-Pro (version 2.8.0_devel)
Step 6.
SOFTWARE PACKAGE (Linux -)
HiC-Pro, 2.8.0_devel
Hi-C mapping: 2. Hi-C assembling with 3D-DNA pipeline (version 170123)
Step 7.
SOFTWARE PACKAGE (Linux -)
3D-DNA pipeline, 170123
Hi-C mapping: 3. The contact map was visualized by Juicerbox (version 1.5.2)
Step 8.
SOFTWARE PACKAGE (Linux -)
Juicerbox, 1.5.2

Hi-C mapping: 4. To evaluate the Hi-C assembly, the chromosome sequences of moso bamboo were aligned with rice (Oryza sativa) genome using Lastz (version 1.02.00) [13] with the parameters: T=2 C=2 H=2000 Y=3400 L=6000 K=2200. Syntenic blocks longer than 2kb were plotted using Circos (version 0.69)

Step 9.

SOFTWARE PACKAGE (Linux -)

Circos, 0.69