

Genome survey

Hansheng Zhao

Abstract

The understanding of genome characteristics for a given new species, i.e., genome size and heterozygosity, facilitates customizing a specific sequencing and assembling strategy. Thus, the genome size was estimated using four independence methods, i.e., a script of KmerSpectrumPlot.pl in ALLPATHS-LG (version r52488), GCE (Genome Characteristics Estimation, released 20150107, <ftp://ftp.genomics.org.cn/pub/gce>), JELLYFISH (version 2.0) and flow cytometry.

Citation: Hansheng Zhao Genome survey. **protocols.io**

[dx.doi.org/10.17504/protocols.io.rf3d3qn](https://doi.org/10.17504/protocols.io.rf3d3qn)

Published: 03 Jul 2018

Protocol

Data preprocessing

Step 1.

In the data preprocessing, low-quality reads (the proportion of the base of Q<13 more than 40% in a given reads) were filtered using NGS QC Toolkit (version 2.3.3) with default parameters.

 [SOFTWARE PACKAGE \(LINUX - \)](#)

NGS QC Toolkit, 2.3.3 [↗](#)

Genome survey

Step 2.

The understanding of genome characteristics for a given new species, i.e., genome size and heterozygosity, facilitates customizing a specific sequencing and assembling strategy. Thus, the genome size was estimated using four independence methods, i.e., a script of KmerSpectrumPlot.pl in ALLPATHS-LG (version r52488), GCE (Genome Characteristics Estimation, released 20150107, <ftp://ftp.genomics.org.cn/pub/gce>), JELLYFISH (version 2.0) and flow cytometry.

 [SOFTWARE PACKAGE \(LINUX - \)](#)

Jellyfish, 2.0 [↗](#)