

A comprehensive investigation of metagenome assembly by linked-read sequencing

Lu Zhang
Xiaodong Fang
Herui Liao
Zhenmiao Zhang
Xin Zhou
Lijuan Han
Yang Chen
Qinwei Qiu
Shuai Cheng Li

Video Byte

Keywords: metagenome assembly, linked-reads, short-reads, PacBio CCS long-reads, parameter space, next-generation sequencing, NGS, Microbiome

Posted Date: February 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-275909/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Linked-read sequencing is helping microbiome researchers remove some of the ambiguities produced by short-read metagenomic sequencing, but how the parameters of linked-read sequencing affect metagenome assembly quality has remained unclear. Using simulated data, a mock community and real metagenomic data, a new study recommends best practices for linked-read sequencing of microbial genomes. Results showed that assemblies of medium-abundance microbes could be substantially improved by choosing a deep read depth. Indeed, selecting a read depth between 120X and 140X is recommended to optimize assembly quality. A tradeoff was observed between read depth per fragment and DNA fragment physical depth, with a deep read depth per fragment generating more high-quality bins. and DNA fragment physical depth controlling total assembly length. Furthermore, while a large average fragment length allows for DNA fragments spanning distant contigs, it is not necessary to produce extremely long fragments for microbial assembly genomes. Finally, a small number of fragments per partition can reduce off-target reads and boost the efficiency of local assembly. Overall, the best approach for using linked reads for metagenome assembly could be to merge linked reads from multiple libraries. where each has a sufficient read depth per fragment but a smaller amount of input DNA.