Jan 15, 2016

MetaVelvet : An extension of Velvet assembler to de novo metagenome assembly from short sequence reads

DOI
https://dx.doi.org/10.17504/protocols.io.d7q9mv

Afiahayati[1], Sato K[1], Namiki T[1], Hachiya T[1], Tanaka H[1], Sakakibara Y.[1]

[1]Hurwitz Lab

VERVE Net | Hurwitz Lab

Bonnie L Hurwitz
University of Arizona

DOI: https://dx.doi.org/10.17504/protocols.io.d7q9mv

Collection Citation: Afiahayati , Sato K, Namiki T, Hachiya T, Tanaka H, Sakakibara Y. 2016. MetaVelvet : An extension of Velvet assembler to de novo metagenome assembly from short sequence reads. **protocols.io** https://dx.doi.org/10.17504/protocols.io.d7q9mv

Protocol status: Working

Created: November 17, 2015

Last Modified: January 19, 2018

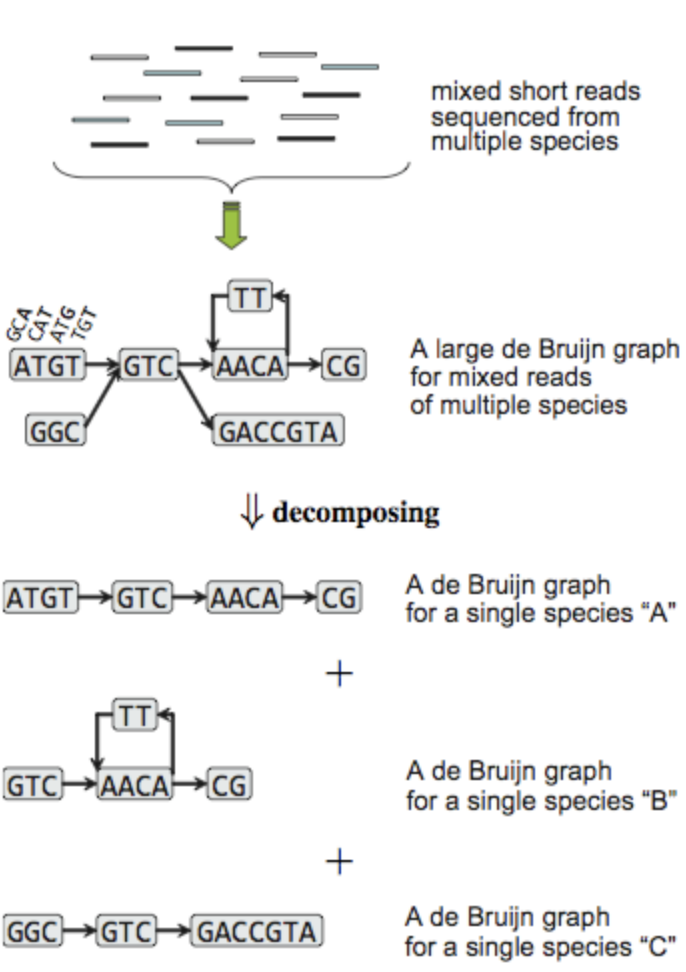
Collection Integer ID: 2000

Keywords: de novo metagenome assembly from short sequence, de novo metagenome assembly, metagenomic, species genome, metagenome assembly, assembly of multiple genome, mixed sequence reads of multiple species, genome assembler, multiple genome, single genome, metagenome, genome, multiple species, mixed sequence read, microbial community, different species, de novo, conserved sequence, chimera contig, bruijn graph, mixed short read

Abstract

Motivation: An important step of "metagenomics" analysis is the assembly of multiple genomes from mixed sequence reads of multiple species in a microbial community. Most conventional pipelines employ a single-genome assembler with carefully optimized parameters and post-process the resulting scaffolds to correct assembly errors. Limitations of the use of a single-genome assembler for *de novo* metagenome assembly are that highly conserved sequences shared between different species often causes chimera contigs, and sequences of highly abundant species are likely mis-identified as repeats in a single genome.

Methods: We modified and extended a single-genome and de Bruijn-graph based assembler, **Velvet**, for *de novo* metagenome assembly. Our fundamental ideas are first decomposing de Bruijn graph constructed from mixed short reads into individual sub-graphs and second building scaffolds based on every decomposed de Bruijn sub-graph as isolate species genome.



Files

SEARCH

Protocol

NAME

Installation and Getting Started

VERSION 1

CREATED BY

Bonnie L Hurwitz
University of Arizona

OPEN →

Protocol

NAME

Manual setting of k-mer coverage peaks

VERSION 1

CREATED BY

Bonnie L Hurwitz
University of Arizona

OPEN →

Protocol

NAME

Use Bambus2 scaffolding module

VERSION 1

CREATED BY

Bonnie L Hurwitz
University of Arizona

OPEN →

Protocol

NAME

How to use paired-end information for graph decomposition

VERSION 1

CREATED BY

Bonnie L Hurwitz
University of Arizona

OPEN →