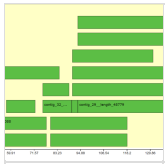


Mar 29, 2021

Protocol for assembling SARS-Cov-2 runs with s-aligner

Juanjo Bermúdez[1]

[1]Contignant Technologies SL

1 Works for me dx.doi.org/10.17504/protocols.io.bsbvnan6

Coronavirus Method Development Community Contignant

Juanjo Bermúdez

SUBMIT TO PLOS ONE

ABSTRACT

This is a protocol for using Contignant s-aligner to de-novo assemble SARS-Cov-2 genomes. S-aligner outperforms common open-source software for de novo-assembly of viruses by 110% increased performance.

DOI

dx.doi.org/10.17504/protocols.io.bsbvnan6

PROTOCOL CITATION

Juanjo Bermúdez 2021. Protocol for assembling SARS-Cov-2 runs with s-aligner. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bsbvnan6

CREATED

Feb 11, 2021

LAST MODIFIED

Mar 29, 2021

PROTOCOL INTEGER ID

47189

PARENT PROTOCOLS

In steps of

untitled protocol

Indexing and fast-testing

1

Index the reads in the run. You can use as input a FASTA file (preferred) or a FASTQ. The file can also be compressed with gzip, having a .gz extension. The script will uncompress the file and pass it to FASTA format if it's a FASTQ file.

./index.sh your_run_id /mnt/c/your_run.fastq.gz

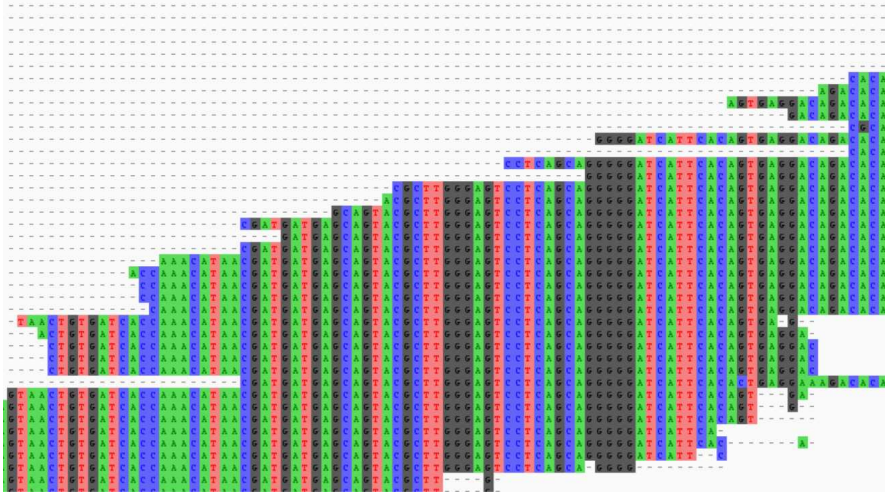
2

Assemble

./assemble.sh your-run-id -output alignments 500 > results/your-run-id-500.fa

Assemble up to processing 500 reads

results/your-*run*-id-500.fa should have the largest contig bigger than 1.200bp and at least one of the files containing aligned reads to form a contig should look like that:



Step 2 includes a Step case.

<1.200bp

>1200bp

step case

<1.200bp

The largest contig in your-*run*-id-500.fa is shorter than 1200bp

3

Map your reads to a reference genome to see what's going wrong.

./map.sh your-run-id sequences/your-run-id/sra_data.part-71.fa reference.fa 4000 > results/map-your-run-id.fa

Step 3 includes a Step case.

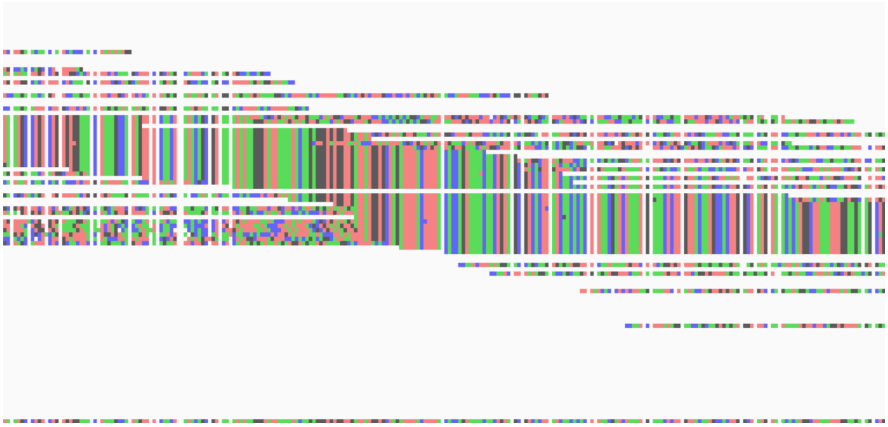
Data not trimmed

Citation: Juanjo BermÃÂºdez (03/29/2021). Protocol for assembling SARS-Cov-2 runs with s-aligner. https://dx.doi.org/10.17504/protocols.io.bsbvnan6

Too many chimeras
No overlapping
Runs are short
Contamination

step case

Data not trimmed



If the data is not trimmed and has adaptors at the extremes of the runs it will look like the image above.

4 Trim your data and start again.

go to step #1

Citation: Juanjo BermÃÂºdez (03/29/2021). Protocol for assembling SARS-Cov-2 runs with s-aligner. https://dx.doi.org/10.17504/protocols.io.bsbvnan6