

Analysing nanopore sequencing data

Alp Aydin QIB





Science Health
Food Innovation

MinIT and EPI2ME demo



Examples of alternatives

Check quality/metadata:

- NanoComp
- MinionQC

Classification:

- Centrifuge (what EPI2ME uses)
- Kraken

Mappers/aligners:

- Minimap2 (what EPI2ME uses)
- BWA
- Bowtie2

Example of another integrated; tool:

- NanoOK





Docs

Windows

Azure

Visual Studio Office Microsoft 365 .NET More

Search

Docs / Install on Windows 10

Bookmark

Feedback

Edit

Share

Theme

Sign in

Filter by title

[About the Windows Subsystem for Linux](#)[Install on Windows 10](#)[Install on Windows Server](#)[Initialize Distro](#)[User accounts](#)[Distro management](#)[File System Permissions](#)[Windows Subsystem for Linux for Enterprise](#)[Interoperability](#)[Build a Custom Distro](#)[Frequently Asked Questions](#)

> WSL 2

Windows Subsystem for Linux Installation Guide for Windows 10

07/23/2018 • 2 minutes to read • +14

Install the Windows Subsystem for Linux

Before installing any Linux distros for WSL, you must ensure that the "Windows Subsystem for Linux" optional feature is enabled:

1. Open PowerShell as Administrator and run:

PowerShell

Copy

```
Enable-WindowsOptionalFeature -Online -FeatureName Microsoft-Windows-Subsystem-Linux
```

2. Restart your computer when prompted.

Is this page helpful?

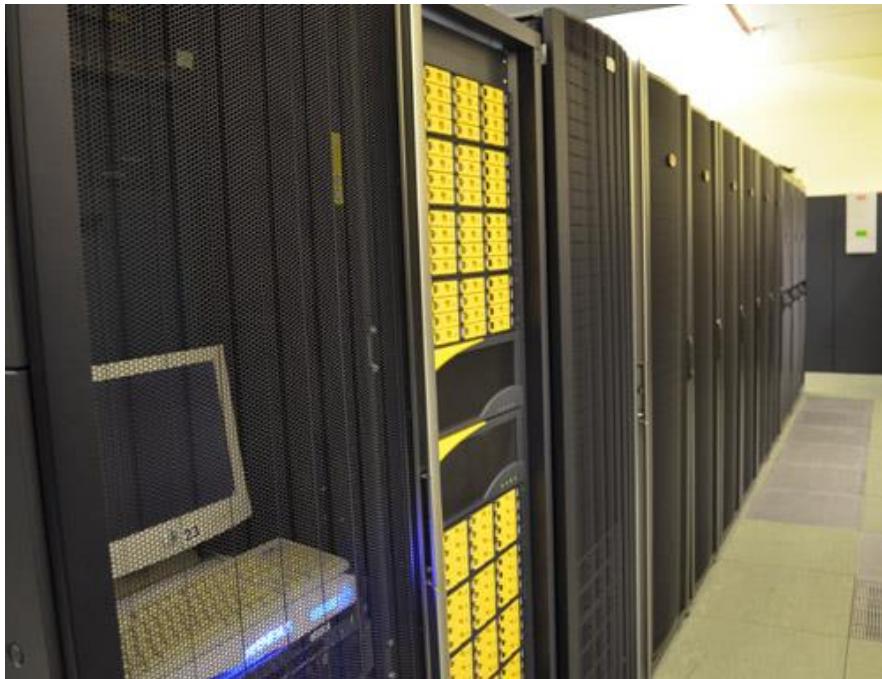
Yes No

In this article

[Install the Windows Subsystem for Linux](#)

[Install your Linux Distribution of Choice](#)[Complete initialization of your distro](#)[Troubleshooting:](#)

CLIMB – Cloud Infrastructure for Microbial Bioinformatics



MRC | Cloud Infrastructure
for Microbial
Bioinformatics

Community

Register for a CLIMB group account



Downstream analysis: Assemblies

Why assemble?

- Reference for a new species
- Compare to reference for variation
- Many bioinformatic tools use it
- Study structure of genomes (arrangements, repeats etc.)



Reference alignment vs. De novo assembly

Use existing data

A)



'New'

B)

De Novo Assembly

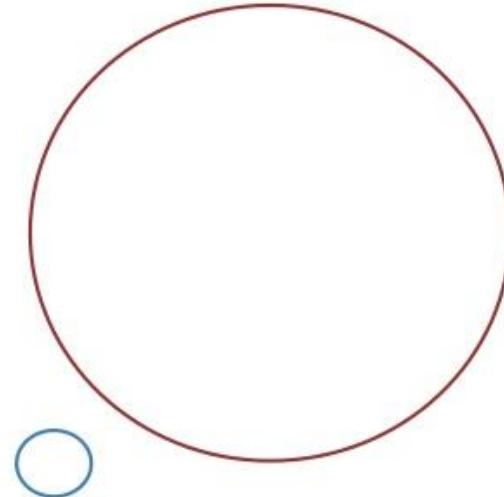
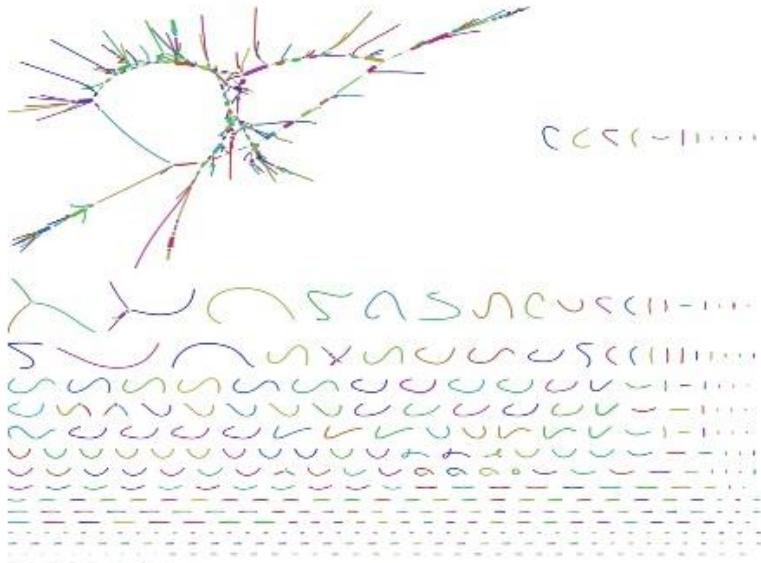
Reads



Contig -1
Contig -2
Contig -3
Contig -4
Contig -5
Contig -6



Long read advantage



Repeat regions

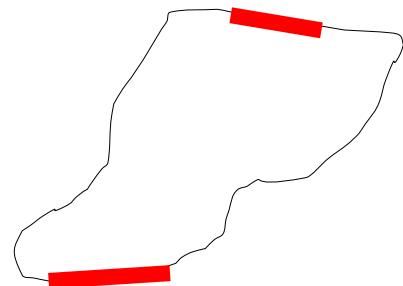
Short reads - where is the overlap?:

GCACGATTGTC **AATT AATT AATT**
AATT AATT AATT GCAGTACGT

GCACGATTGTC **AATT AATT AATT**
AATT AATT AATT GCAGTACGT

Long read:

GCACGATTGTC **AATT AATT AATT AATT AATT AATT AATT** GCAGTACGT



How would you assemble?

1. Filter (e.g. NanoFilt)
2. Trim (e.g. Porechop)
3. Assemble (e.g. Canu)
4. Polish (e.g. Nanopolish)
5. Visualise (e.g. Bandage)

Hybrid assembly: Unicycler

Canu command: `canu -p -d genomeSize=? -nanopore-raw`



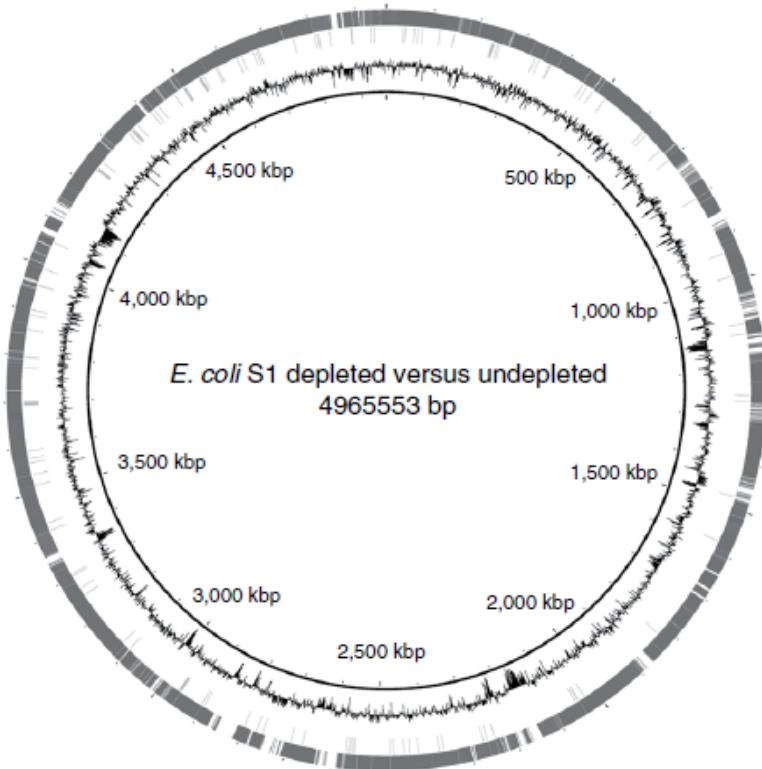


Science Health
Food Innovation

Assembly demo



Example assembly with Canu



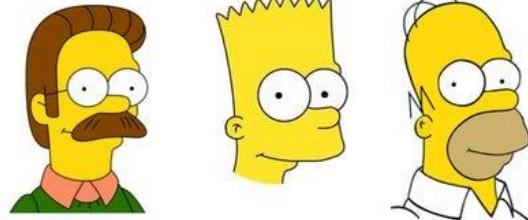
Depleted:
33.5x coverage after 2 hours

Undepleted:
0.2x coverage after 2 hours



Typing

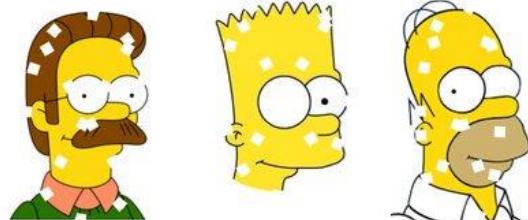
Genomes



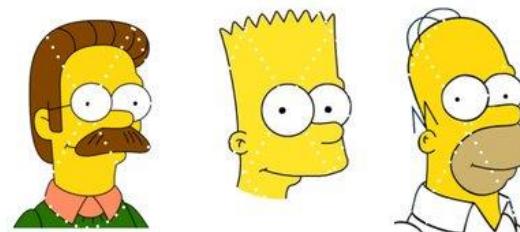
Multi Locus Sequence Typing



Core/Whole Genome MLST



Whole genome SNPs



e.g. Krocus -
MLST from
sequencing
data



Variant calling

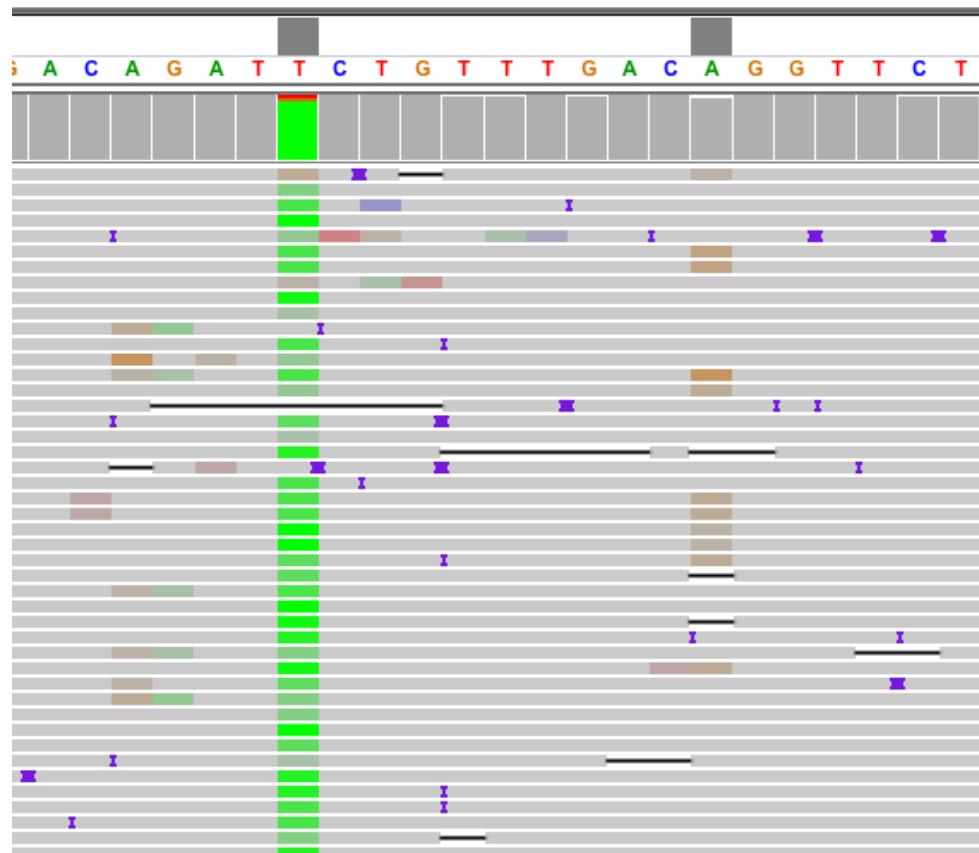


```
minimap2 -ax map-ont ref.fa ont.fq.gz > aln.sam
```

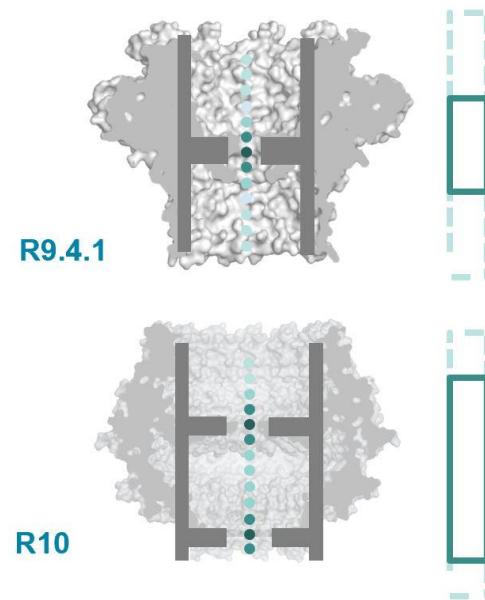
```
nanopolish variants --reads reads.fa --bam alignments.bam --genome genome.fa
```

Variant calling – Nanopore data

Nanopolish
(Variant)



New pore

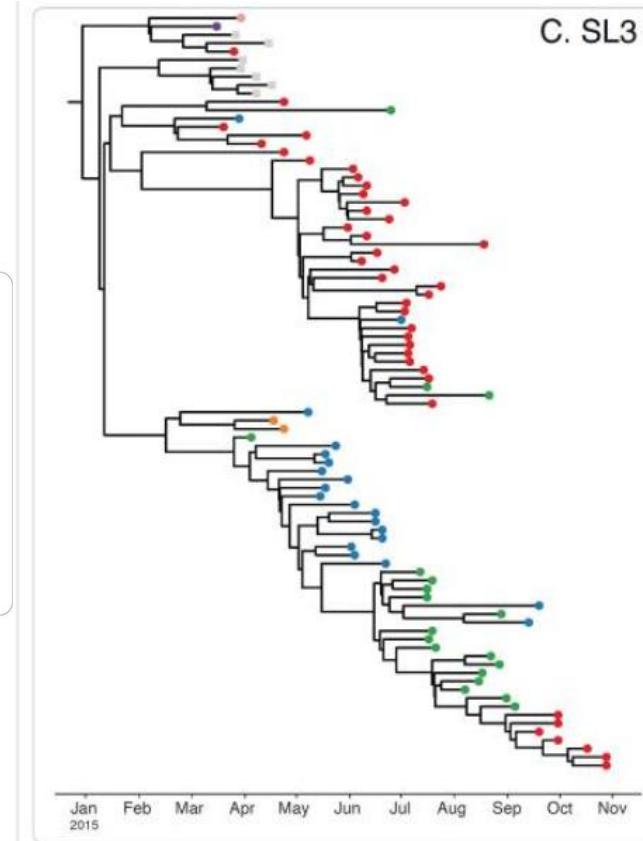


ATCGGAAAAAAATCACGCCACGTCCAAA



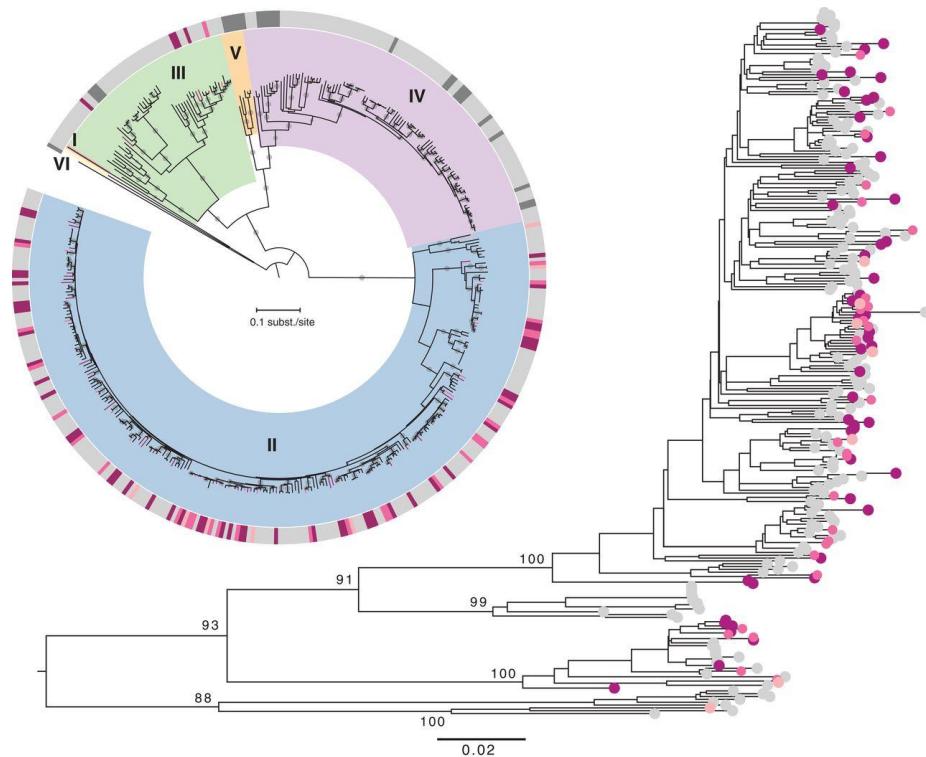
Phylogenetic reconstructions

2014-2016 West African Ebola outbreak



2018 Nigerian Lassa fever outbreak

RAXML - trees
BEAST - timescale





Quadram
Institute

Science • Health •
Food • Innovation

