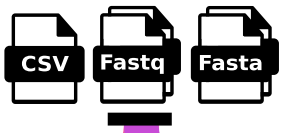
PathogenSurveillance

nf-core



Input data:
Metadata
Raw reads
Optional references

Select and download references

Fast initial ID
bbmap sendsketch

Search for assemblies
Entrez esearch

Download assemblies
datasets download

Create signatures of assemblies
sourmash sketch

Select references for variant calling

Filter reads by kmer abundance
khmer trim_low_abund

Make signatures of user references
sourmash sketch

Make signatures of reads
sourmash sketch

Compare signatures
sourmash compare

Assign references
custom script

Genome assembly and annotation

Trim adapters
fastp

Assemble genomes
spades

Filter assemblies
custom script

Annotate genomes
bakta

Variant calling

Align reads
bwa mem

Identify duplicates
picard MarkDuplicates

Call variants
graphtyper genotype

Combines VCF files
graphtyper vcf_concatenate

Filter variants
vcflib vcffilter

Infer phylogeny
iqtree

Read2Tree phylogeny (Eukaryotes)

Extract BUSCO genes
busco

Make Read2Tree reference database
custom script

Infer phylogeny
read2tree

Core genome phylogeny (Prokaryotes)

Identify shared orthologs
pirate

Extract core genes
custom script

Make gene multiple sequence alignments
mafft

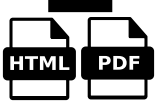
Infer phylogeny
iqtree

Make read quality report
multiqc

Create reports
custom script

Assess read quality
fastqc

Assess assembly quality
quast



- Find and download references
- Select reference for variant calling
- Genome assembly and annotation
- Variant calling
- Read2Tree phylogeny
- Core genome phylogeny
- Quality control and reporting

nextflow