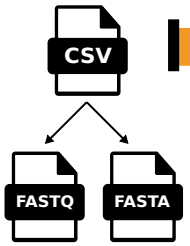
nf-core/pathogensurveillance

Input data:
Metadata
Sequence data
References (optional)



- Input preparation
- Assembly and sketching
- Core genome phylogeny
- Variant calling
- BUSCO phylogeny
- Report creation

Standardize and download input
custom R script, entrez eutils

Fast initial ID
bbmap sendsketch

Select and download references
NCBI datasets

Trim adapters
fastp

Assemble genomes
spades, flye

Sketch assemblies and references
sourmash sketch

Compare all sketches
sourmash compare

Pick contextual references
Custom R script

Annotate genomes
bakta

Identify shared orthologs
pirate

Align shared orthologs
mafft

Infer phylogeny
iqtree2

Pick mapping reference
Custom R script

Align reads
bwa mem

Call variants
graphtyper, picard

Filter variants
vcflib

Infer phylogeny
iqtree2

Pick contextual references
Custom R script

Extract BUSCO genes
busco

Align BUSCO genes
mafft

Infer phylogeny
iqtree2

Make quality control reports
multiqc, fastqc, quast

Make main report
Custom Quarto script, psminer

Output data:
Interactive reports
Intermediate files

USDA
Oregon State University
nextflow