

output_directory	
aligned_reads	
<reference_id>--<sample_id>.bam	: Unfiltered read alignments to references in BAM format
<reference_id>--<sample_id>.formatted.bam	: Quality filtered read alignments to references in BAM format
<reference_id>--<sample_id>.formatted.bam.csi	: Index for the filtered BAM files produced by samtools index
<reference_id>--<sample_id>.formatted.MarkDuplicates.metrics.txt	: Output from picard MarkDuplicates containing summary data tables
annotations	
<method>	
<sample_id>.embl	: Genome annotations and sequences in (multi) EMBL format
<sample_id>.faa	: CDS/ORF amino acid sequences in FASTA format
<sample_id>.ffn	: Feature nucleotide sequences in FASTA format
<sample_id>.fna	: Replicon/contig DNA sequences in FASTA format
<sample_id>.gbff	: Annotations of genomes in (multi) GenBank format
<sample_id>.gff3	: Annotations of genomes in GFF3 format
<sample_id>.hypotheticals.faa	: Hypothetical protein CDS amino acid sequences in FASTA format
<sample_id>.hypotheticals.tsv	: Information on hypothetical protein CDSs in TSV format
<sample_id>.tsv	: Information for protein CDSs in TSV format
<sample_id>.txt	: Genome annotation summary statistics in plain text format
assemblies	
flye	
<sample_id>.assembly_graph.gfa.gz	: Genome assembly graphs in gzipped GFA format
<sample_id>.assembly_graph.gv.gz	: Genome assembly in gzipped Graphviz format
<sample_id>.assembly_info.txt	: Statistics on contigs produced by Flye in TSV format
<sample_id>.assembly.fasta.gz	: Genome assemblies in gzipped FASTA format
<sample_id>.flye.log	: Log file produced by Flye in plain text format
<sample_id>.params.json	: Parameters used to run Flye in JSON format
spades	
<sample_id>.assembly.gfa.gz	: Genome assembly graphs in gzipped GFA format
<sample_id>.scaffolds.fa.gz	: Genome assemblies in gzipped FASTA format
<sample_id>.spades.log	: Log file produced by Spades in plain text format
<sample_id>.warnings.log	: Warnings produced by Spades in plain text format
busco	
<sample_id>--<database>-busco	: BUSCO output directory
<sample_id>--<database>-busco.batch_summary.txt	: BUSCO analysis summary statistics in TSV format
<sample_id>--<database>-busco.log	: Log file produced by BUSCO in plain text format
downloads	
annotations	
<reference_id>.gff.gz	: Genome annotations downloaded by the pipeline in gzipped GFF format
assemblies	
<reference_id>.fasta.gz	: Genome assemblies downloaded by the pipeline in gzipped FASTA format
reads	
<reference_id>_<orientation>.fastq.gz	: Reads downloaded by the pipeline in gzipped FASTQ format
fastp	
<sample_id>.fastp.html	: FASTP report with read quality and adapter trimming statistics in HTML format
<sample_id>.fastp.json	: FASTP report with read quality and adapter trimming statistic in JSON format
<sample_id>.fastp.log	: Log file produced by FASTP
metadata	
reference_metadata.tsv	: Reference metadata in TSV format
sample_metadata.tsv	: Sample metadata in TSV format
pipeline_info	
execution_report_<date>.html	: Statistics like time and ram usage for pipeline execution as a HTML report
execution_trace_<date>.txt	: Statistics like time and ram usage for pipeline execution in TSV format
messages.tsv	: Errors, warnings, and messages produced by the pipeline in TSV format
params_<date>.json	: Parameters used when running the pipeline in JSON format
pathogensurveillance_run_info.yml	: Pipeline run metadata in YML format
pipeline_dag_<date>.html	: Flowchart of the pathogensurveillance pipeline automatically generated by Nextflow
version_info.yml	: Versions of software used in the pipeline in YML format
pirate	
<report_id>_results	: Pirate output directory with gene orthology information
pocp	
<report_id>_pocp.tsv	: Percentage of conserved proteins in TSV format
quality_control	
fastqc	
<reads_id>_fastqc.html	: FastQC report on read quality in HTML format
<reads_id>_fastqc.zip	: Zipped directory of FastQC output
multiqc	
<report_id>_multiqc	: Directory of MultiQC output with intermediate files used to make the MultiQC report
multiqc_report.html	: MultiQC report combining the summary quality control output in HTML format
nanoplot	
<sample_id>_nanoplot-report.html	: Nanoplot report on read quality in HTML format
<sample_id>_NanoStats.txt	: Nanoplot report on read quality in plain text format
quast	
<sample_id>.tsv	: Quast summary statistics on genome assembly quality in headerless TSV format
<sample_id>_report.html	: Quast report with data on genome assembly quality in HTML format
reference_data	
considered	
<taxon_id>.tsv	: Metadata of available assemblies on NCBI in TSV format
downloaded	
<sample_id>.tsv	: Metadata of references selected for download from NCBI in TSV format
selected	
<report_id>_busco_references.tsv	: IDs of references selected for multigene phylogenies in TSV format
<report_id>_core_references.tsv	: IDs of references selected for multigene phylogenies in TSV format
<report_id>_mapping_references.tsv	: IDs of references selected for multigene phylogenies in TSV format
reports	
<report_id>_pathsurveil_report.html	: Primary report for pathogensurveillance in HTML format
sendsketch	
<sample_id>.txt	: Table returned by BMap sendsketch with initial identifications in TSV format
shared_genes	
<method>	
<gene_id>.aligned.fas	: Aligned sequences for each gene shared among a group of organisms in FASTA format
sketch_comparisons	
ani_matrices	
<report_id>_comp.csv	: Estimated ANI similarity matrix made by sourmash compare in CSV format
sketches	
<id>.sig	: FracMinHash signature of the given sequence made by sourmash sketch in JSON
trees	
busco	
<report_id>--<cluster_id>.treedata	: Trees inferred from multigene phylogeny in Newick format
core	
<report_id>--<cluster_id>.treedata	: Trees inferred from multigene phylogeny in Newick format
snp	
<report_id>--<reference_id>.treedata	: Trees inferred from variant analysis in Newick format
variants	
<report_id>--<reference_id>.fasta	: Filtered variants in FASTA format
<report_id>--<reference_id>.vcf.gz	: Unfiltered variants in VCF format
<report_id>--<reference_id>.vcffilter.vcf.gz	: Filtered variants in VCF format