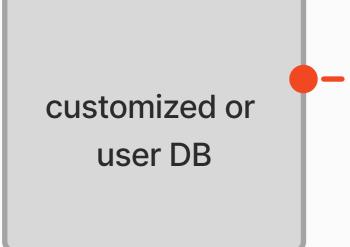
- Virus Identification (Genomad, VS2, CT3)
- pre-processing analysis output
- Motif discovery output
- Recombination analysis output

fasta files



Phylogenetic Analysis

Alignment

- input: fasta
- + MAFFT

Trimming

- input: aligned files
- + trimAl

Phylogenetics

- input: aligned fasta
- + IQ-TREE2

```
output/.

├── alignment.log  # Log file

├── metadata.csv  # Metadata with protein_id, protein_description, family,

├── my_sequences_aligned_sequences.fasta  # Sequence alignment FASTA file

├── my_sequences_aligned_trimmed_sequences.fasta  # Trimmed sequence alignment FASTA file

└── my_sequences_merged.fasta  # Merged FASTA file (all sequences used in alignment and
```

```
output/tree/.
my_sequences_aligned_trimmed_sequences_sanitized_name_table.tsv
    # Table with modified sequence names (spaces replaced by " " to avoid issues in downstream analyse
— my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta
    # Sequences used to build the phylogenetic tree
                                            # Module log file
├── build tree.log
my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.bionj
    # BIONJ format tree (an improved version of the NJ algorithm)
— my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.ckp.gz
    # IQ-tree checkpoint file
 — my sequences aligned trimmed sequences sanitized sequences.fasta.contree
    # Consensus tree with branch supports (branch lengths optimized on the original alignment)
my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.log
    # Log file for the entire tree-building run
my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.mldist
    # Pairwise distance matrix (for downstream analyses)
my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.model.gz
    # IQ-TREE model checkpoint file
— my sequences aligned trimmed sequences sanitized sequences.fasta.splits.nex
    # Support values (percentages) for splits, computed from bootstrap trees
my_sequences_aligned_trimmed_sequences_sanitized_sequences.fasta.iqtree
    # Main IQ-tree report file (self-readable; see computational results)
—— my sequences aligned trimmed sequences sanitized sequences.fasta.treefile
    # ML tree in NEWICK format (compatible with viewers like FigTree or iTOL)
```