

1. What kind of sequences does the Bacterial Analysis Pipeline accept ? (please chose one or more correct options)

- ☒ FASTQ - single end
- ☒ FASTQ - paired end
- ☒ Assembled genome in FASTA format
- ☐ Genbank

2. What are correct about the Bacterial Analysis Pipeline ? (please chose one or more correct options)

- ☒ MLST, ResFinder and PlasmidFinder are included in the Bacterial Analysis Pipeline
- ☒ It requires users to submit excel metadata
- ☐ It requires users to specify species in the excel metadata
- ☒ The Bacterial Analysis Pipeline allows user to rerun their samples without re-uploading.

3. What is assumed when creating SNP trees with CSI Phylogeny?

- ☐ Only assemblies have been uploaded
- ☐ Only raw data have been uploaded
- ☒ The SNPs are random and independent
- ☐ The entire genome of all analyzed isolates has been sequenced

4. CSI Phylogeny filters SNPs using these criteria:

- ☒ Coverage, quality, and z-score
- ☐ Coverage, quality, and a p-value
- ☐ Coverage, quality, and location of SNP on the read
- ☐ Coverage, quality, and the type of SNP

5. What type of data can be used as input to CSI Phylogeny (Please chose all options that apply)?

- ☐ VCF files
- ☒ A mix of FASTA files and FASTQ files
- ☒ FASTQ files
- ☒ FASTA files

6. What are the output from CSI Phylogeny (Please chose all options that apply)?

- ☒ SNPs in VCF format
- ☒ A matrix or table shows SNP differences
- ☐ Genbank File
- ☒ Phylogenetic tree based on SNPs

7. From the output page, a pseudo-alignment in FASTA format can be downloaded. What can this file be used for?

- ☒ As input to other tree building algorithms
- ☐ To see codon changes
- ☐ To get an overview of the called SNPs
- ☐ To get number of synonymous SNPs

8. What is the database for MyDBFinder ?

- ☐ A single FASTA file containing one DNA sequence
- ☐ A single FASTA file containing multiple DNA sequences
- ☐ A single FASTA file containing multiple genes
- ☒ All above

9. *Vibrio cholera* O1 contains *rfvB* gene but *Vibrio cholera* O139 contains *wbfz* gene. Using 'Vibrio_typing_database' and 'Vibrio_unknown_sample' as inputs in MyDBFinder, which is the serogroup of the *Vibrio* unknown sample ?

Vibro_typing_database.fasta.zip

Vibrio_unknown_sample.fasta.txt.zip

Please note that the sequence files need to be **unzipped** prior uploading to the web tool.

- ☒ O1
- ☐ O139
- ☐ Non-typeable
- ☐ The result is not clear to identify serogroup

10. What are the input data (unknown input) for MyDBFinder ? (please chose one or more correct options)

☐ Amino acid sequences in FASTA format

☒ FASTQ - single end

☒ FASTQ - paired end

☒ Assembled genomes in FASTA format