

Text with Link to files to be used in tutorial

The tutorial contains 5 unknown samples. They are assembled genomes in FASTA format. You can start analysing them directly without conducting *de novo* assembly.

Hint:

- 5 unknown samples
- 3 of them are outbreak strains (clonal strains)

Your tasks:

- Identity what they are (species)
- Identify subtype; ST-type, resistance genes, plasmid replicons and pMLST
- Construct SNP phylogenetic to see how closely related they are (You do not have to upload data to CSI Phylogeny. The phylogeny results are provided at the end of the page)
- Make the final decision which are the three outbreak strains

You can use an excel template to summarise your results (link to the excel, <https://www.coursera.org/learn/wgs-bacteria/resources/oF3Z9>).


Here is the link to all the 5 unknown samples; <https://www.coursera.org/learn/wgs-bacteria/resources/xOSBf>

Please unzip the sequence files prior uploading to the web tools.

**** CSI Phylogeny is overused ****

Please find results from CSI Phylogeny (phylogenetic tree based on SNPs, a table of SNP difference and concatenated SNP in FASTA) from the following zip file;

 **CSI_phylogeny_output**
ZIP File

[Download file](#) 

List of programs for phylogenetic tree visualisation;

FigTree (stand-alone tool)

<http://tree.bio.ed.ac.uk/software/figtree/>

iTol (web-based tool)

<https://itol.embl.de>