Simulated Outbreak Exercise

1. Introduction to the exercise

This exercise is based on a fictional outbreak scenario, described below, that requires your team to work together to perform cluster analysis and additional analyses to solve a potential hospital outbreak of carbapenem- and/or colistin resistant *Enterobacteriaceae* (CCRE).

2. Exercise Scenario

A recent rise in cases of carbapenemase producing *E. coli* in several regional hospitals indicates one or more ongoing outbreaks. Patients include both domestic and travel-related cases, and a batch of samples has already been sequenced using Illumina. From these sequences, subtyping by MLST was performed and a selection (12 *E. coli* isolates) of the most predominant MLST (ST410) isolates has been transported to your laboratory for further analysis. Your laboratory has just finalized setting up Oxford Nanopore Technology (ONT) MinION sequencing, and you wish to use this occasion to work with both types of sequencing data.

The currently available metadata for the samples can be seen in Table 1 below:

Sample name	Species	Date	Region of isolation	Travel	MLST	Carba genotype (PCR)
Eco001	E. coli	2015	Copenhagen	Pakistan	ST410	OXA-48-like
Eco002	E. coli	2015	Copenhagen	Thailand	ST410	OXA-48-like
Eco003	E. coli	2015	Jutland - M	India	ST410	NDM
Eco004	E. coli	2015	Copenhagen	Lebanon	ST410	OXA-48-like
Eco005	E. coli	2016	Zealand	No	ST410	NDM, OXA-48-like
Eco006	E. coli	2016	Zealand	No	ST410	NDM, OXA-48-like
Eco007	E. coli	2017	Copenhagen	Pakistan	ST410	OXA-48-like
Eco008	E. coli	2018	Jutland - N	Thailand	ST410	NDM
Eco009	E. coli	2018	Zealand	No	ST410	NDM, OXA-48-like
Eco010	E. coli	2018	Zealand	No	ST410	NDM, OXA-48-like
Eco011	E. coli	2018	Zealand	No	ST410	NDM
Eco012	E. coli	2018	Zealand	No	?	OXA-48-like

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

3. Exercises

a. MLST

The MLST sequence type (ST) is missing for Eco012.

Use MLST 2.0 to find out the ST for Eco012.

b. Resfinder

The carba genotype has already been described using PCR.

Use Resfinder to find out the carba genotype using WGS-data for all 12 E. coli isolates. You can work with the Illumina fasta files.

What other antimicrobial resistance genes are there?

There is an excel sheet where you can insert your results.

The results have also been uploaded to an excel file that is available on the ScienceData site.

c. MobileElementFinder

Now the carba genotype has been described with PCR and WGS data.

Use MobileElementFinder to find out if any of the carbapenemase genes are on plasmids. You can work with the Illumina fasta files.

There is an excel sheet on the ScienceData sheet where you can insert your results.

The results have also been uploaded to an excel file that is available on the ScienceData site.

d. SNP analysis

Below there are a series of analyses to perform SNP analysis on the 12 E. coli isolates with both the Illumina and ONT data.

For each analysis, the tool, reference, pruning and data you should use is stated. The result of the analysis is also provided as a link.

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

Analysis 1

Tool: CSI Phylogeny

Reference: KmerFinder reference

Prune: 10

Data: Illumina draft genomes (all 12 isolates)

Analysis 2

Tool: CSI Phylogeny

Reference: KmerFinder reference

Prune: 100

Data: Illumina draft genomes (all 12 isolates)

Analysis 3

Tool: CSI Phylogeny

Reference: Best Reference

Prune: 100

Data: Illumina draft genomes (all 12 isolates)

Analysis 4

Tool: CSI Phylogeny

Reference: KmerFinder reference

Prune: 100

Data: Illumina raw data (all 12 isolates)

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

Analysis 5

Tool: CSI Phylogeny

Reference: KmerFinder reference

Prune: 100

Data: Illumina raw data (9 closest related isolates only)

Analysis 6

Tool: CSI Phylogeny

Reference: Best reference

Prune: 100

Data: Illumina raw data (9 closest related isolates only)

Analysis 7

Tool: CSI Phylogeny

Reference: Best reference

Prune: 100

Data: ONT assembly data (9 closest related isolates only)

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

The following exercises use MinTyper:

Analysis 8

Tool: MinTyper

Reference: Best reference

Prune: 100

Data: Illumina raw data (All 12 isolates)

Analysis 9

Tool: MinTyper

Reference: Best reference

Prune: 100

Data: ONT raw data (All 12 isolates)

Analysis 10

Tool: MinTyper

Reference: Best reference

Prune: 100

Server run time (approximately): 20-30 minutes

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

Analysis 11

Tool: MinTyper

Reference: Best reference Prune: 100

Data: Illumina and ONT raw data (All 2 x 12 isolates)

4. Questions

Do you find any clusters between the 12 E. coli isolates in the dataset?

What is the number of SNPs within the cluster(s)?

Is there any difference if you used genomes or raw data in CSIphylogeny?

Did exclusion of less related isolates change the number of SNPs?

Did the number of SNPs change when using the optimal reference compared to the kmer-reference?

Was it possible to combine Illumina data and ONT MinION data in one analysis?

Which OXA- and NDM- genes were found in the bacteria?

Did clustering strains contain the same type(s) of carbapenemase genes?

Did any of the carbapenemase genes seem to be on a plasmid? If yes, are there other resistance genes on the same plasmid?

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

Based on the sum of all your analysis, which isolates do you think belongs to the same hospital outbreak?

Is there any of the molecular data (SNPs, Resistance genes, Plasmids ect) which is not supporting what would be expected from a clonal outbreak? What can be the explanation(s)?

5. Links to CGE tools

KmerFinder
https://cge.food.dtu.dk/services/KmerFinder/

This tool is for species identification. You can upload a single fasta file or one/two fastq files.

MLST
https://cge.food.dtu.dk/services/MLST/

This tool is for performing MLST typing on bacteria. You can upload a single fasta file or one/two fastq files.

ResFinder
http://genepi.food.dtu.dk/resfinder

This tool identifies AMR genes and point mutations. You can upload a single fasta file or one/two fastq files.

MINTyper
https://cge.food.dtu.dk/services/MINTyper/

This tool is for SNP calling on ONT data. You can upload a reference file as a fasta file (optional), and upload at least two closely related fastq files as your input.

CSI Phylogeny
https://cge.food.dtu.dk/services/CSIPhylogeny/

This tool is for SNP calling on Illumina data. You can upload a reference file as a fasta file, and upload at least two closely related fasta or fastq files as your input.

cgMLSTFinder
https://cge.food.dtu.dk/services/cgMLSTFinder/

Technical University of Denmark
National Food Institute
joam@food.dtu.dk

This tool performs Core genome Multi Locus Sequence Typing (cgMLST) from a set of reads. You can upload a single fasta file or one/two fastq files.

MobileElementFinder
https://cge.food.dtu.dk/services/MobileElementFinder/

This tool identifies mobile genetic elements and their relation to antimicrobial resistance genes and virulence factors. You can upload a single fasta file as input.

Technical University of Denmark
National Food Institute
joam@food.dtu.dk