







# Simulated exercise on CRE outbreak – Klebsiella

# EURGen-RefLabCap Virtual multidisciplinary training workshop 02 May 2023 Jette S. Kjeldgaard & Faisal Khan (jetk@food.dtu.dk – fakh@food.dtu.dk)



Title

2





# Simulated exercises - background

- Series of multidisciplinary training workshops
  - Sept/Oct 2022 introduction to SNP analysis and cgMLST for cluster analysis (WS1)
  - May 2023 Simulated exercise on outbreak analysis (Klebsiella pneumoniae; WS1)
  - Sept 2023- Simulated exercise on outbreak analysis (WS2)
  - 2024: Simulated exercise on outbreak analysis (WS2)

WS1: CCRE/ *E. coli* and *Klebsiella* spp. WS2: CPO/ *Pseudomonas aeruginosa* and *Acinetobacter baumannii* 



# Purpose of the workshop

- To build capacity to work with outbreak investigations
  - background information about bacterial subtyping and cluster analysis
  - suggestions for online available analytical tools to get started on bacterial comparison and outbreak detection
    - In continuation of last virtual training workshop Sept/Oct 2022
- To work with larger amounts of sequencing data and metadata and analyse outputs from SNP analyses
- See more examples of typing and characterisation options by online tools and try working with online tools by yourself





# Learning objectives

- You will be able to perform a cluster analysis of bacteria to look into possible relatedness in a dataset
- You will learn about other relevant tools for sequence analysis and work with a real outbreak data set to characterise and analyse
- You will apply the results from the cluster analysis and the additional analyses to elucidate a possible hospital outbreak of CCRE



# Agenda

#### Exercise schedule:

- Day 1: Tuesday 2 May 13:00~ 14:30
- Introduction and purpose of exercise
- Introduction to typing methods; short review of species ID, AMR, MLST and phylogeny tools
- *Klebsiella* hyper-virulence identification (K-PAM)
- Information about exercise setup, data, tasks and question surveys
- Time for questions

#### • Day 2: Friday 12 May 13:00-15:00

- Presentation of exercise results (outbreak investigation)
- Use of different tools and examples of additional analyses
- Introduction to bioinformatics tool *Kleborate*
- Questions/comments from participants

Your input is very welcome!!

DTU



# **Workshop overview**







# **Resources – where to find what?**

- Slack
  - Slides and video recordings from previous multidisciplinary workshop
    - MLST/cgMLST, SNP calling tools (CSIPhylogeny and MinTyper)
  - Links to data injects
  - Slides and link to video recording from this workshop
- Sciencedata
  - Data injects:
    - Sequence data (fastq and fasta) for each inject
    - Description and additional information about the data injects
- Email
  - Links to data injects



# Tasks - analyses

Analyses

- Characterisation:
  - Species identification
  - Resistance mechanisms (NDM pos/neg)
  - Hypervirulence
- Subtyping:
  - MLST

- Suggestions for tools
- KmerFinder and/or rMLST
- ResFinder/CARD/AmrFinderPlus
- K-PAM
- MLST at CGE (or Pasteur)

• Cluster analysis

CSIPhylogeny & MinTyper

These analyses will be performed on a limited number of isolates Overall cluster analysis will be available as link and newick + SNP matrix files

9





#### # 10-workshops-and-courses ~ Wednesday, April 26th ~ 00 🙌 🤤 🗐 🎝 Д Jette Sejer Kjeldgaard 9:50 AM Dear all, here you can find the links to the recorded presentations from the previous workshop on outbreak scenarios (Sept/Oct 2022), also with presentations on SNP calling tools; CSIPhylogeny and Mintyper, and with some background on typing tools, incl. MLST and cgMLST. CCRE outbreak detection - part 2 Multidisciplinary workshop on CCRE outbreak detection - part 1 Panopto EURGen-RefLabCap\_1st\_virtual\_multidisciplinary\_training\_resolv Here you can also find the slides from the previous course in Sept/Oct 2022 part2 (14 kB) -4 files ▼ Microsoft Teams EURGen\_WS\_1\_Intro\_and CSIPhyl... EURGen\_WS\_1\_MinION\_typing\_H... PDF PDF EURGen\_WS\_1\_SNPvscgMLST\_Ha... EURGen\_WS1\_Exercise\_part\_2\_Je... × PDF PDF EURGen-RefLabCap: 1. virtual multidisciplinary training to resolve a... 2022-10-10 11:00 UTC Jette Sejer Kjeldgaard





# **Brush-up of suggested tools** (previously demonstrated)

Antimicrobial	Class	WGS-predicted phenotype	Genetic background
vancomycin	glycopeptide	No resistance	
mupirocin	pseudomonic acid	No resistance	
tobramycin	aminoglycoside	Resistant	armA (armA_AY220558), aac(6')-lb-cr (aac(6')-lb-cr_DQ303918)
hygromycin	aminoglycoside	No resistance	
isepamicin	aminoglycoside	Resistant	armA (armA_AY220558)
virginiamycin s	streptogramin b	Resistant	msr(E) (msr(E)_FR751518)







# **Species identification (December 2022 course)**

Species identification/confirmation by WGS

- KmerFinder
  - Full genome match to database

#### **KmerFinder**

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

- Shows the most identical match to database
- genome wide plus additional matches
- Can show contamination
  - No threshold for 'acceptable' level of other species
- Can be used to find similar bacteria as reference for SNP analysis (Accessible by accession # in match)

#### **Ribosomal MLST – rMLST**

(https://pubmlst.org/species-id)

• Typing based on the 53 genes encoding the bacterial ribosome protein subunits

## - rMLST

- Typing based on ribosomal genes
  - (*rps* genes)



# **Resistance mechanisms**

- ResFinder, CARD, AMRFinder+, etc
- Confirmation of NDM-gene and variant
- https://cge.food.dtu.dk/services/ResFinder/

Antimicrobial	Class	WGS-predicted phenotype	Genetic background
vancomycin	glycopeptide	No resistance	
mupirocin	pseudomonic acid	No resistance	
tobramycin	aminoglycoside	Resistant	armA (armA_AY220558), aac(6')-lb-cr (aac(6')-lb-cr_DQ303918)
hygromycin	aminoglycoside	No resistance	
isepamicin	aminoglycoside	Resistant	armA (armA_AY220558)
virginiamycin s	streptogramin b	Resistant	msr(E) (msr(E)_FR751518)





# **Subtyping - MLST**

- Use CGE tool, PubMLST or Pasteur website
  - <u>https://cge.food.dtu.dk/services/MLST/</u>

## **MLST 2.0**

Service

Instructions Output

Article abstract Citations

Software version: 2.0.9 (2022-05-11)

Database version: (2023-05-01)

MLST allele sequence and profile data is obtained from PubMLST.org.

Momentanously, the species Lactococcus Lactis is unavailable.

#### Select MLST configuration

Klebsiella pneumoniae 🛛 🛩

Locus	Identity	Coverage	Alignment Length	Allele Length	Gaps	Allele
gapA	100	100	450	450	0	gapA_3
infB	100	100	318	318	0	infB_4
mdh	100	100	477	477	0	mdh_6
pgi	100	100	432	432	0	pgi_1
phoE	100	100	420	420	0	phoE_7
rpoB	100	100	501	501	0	rpoB_4
tonB	100	100	414	414	0	tonB_38





# SNP analysis- CSIPhylogeny & MinTyper

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

## CSI Phylogeny 1.4 (Call SNPs & Infer Phylogeny)

Illumina reads and assemblies

CSI Phylogeny calls SNPs, filters the SNPs, does site validation and infers a phylogeny based on the concatenated alignment of the high quality\* SNPs.

Output for overall analyses on fastq reads will be provided! PLEASE don't re-run these You can run subsets of samples – preferably **fasta** format!

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

ONT (and Illumina) reads and assemblies

## MINTyper 1.0

SNP distance matrice and phylogenetic tree with long and short raw sequencing reads.





## CSI Phylogeny 1.4 (Call SNPs & Infer Phylogeny)

CSI Phylogeny calls SNPs, filters the SNPs, does site validation and infers a phylogeny based on the concatenated alignment of the high quality\* SNPs.

#### Upload reference in fasta – e.g. first patient (kp1)

Change pruning to 100 (optional)

Select additional files (preferably fasta to avoid overload of server) Upload all

#### Input data Comment (to yourself) This comment will appear unaltered on your output page. It has no effect on the analysis. Upload reference genome (fasta format) Note: Reference genome must not be compressed. Vælg fil Der er ikke valgt nogen fil Use altered FastTree (more accurate) Include reference in final phylogeny. Note: Read more here Select min. depth at SNP positions Upload read files and/or assembled genomes (fasta or fastg format) 10x V Note: Read files must be compressed with gzip (compressed files often ends with .gz). If you get an "Access forbidden. Error 403": Make sure the start of the web adress is https and not just http. Fix it by clicking here. Select min. relative depth at SNP positions 10 % ~ Select minimum distance between SNPs (prune) Isolate File 10 bp V Select min. SNP quality Name Size Progress 30 V Select min. read mapping quality 25 ~ Select min, Z-score 1.96 V O Upload Remove ----



#### MINTyper 1.0



SNP distance matrice and phylogenetic tree with long and short raw sequencing reads.

\* For large datasets (>50 isolates), consider running the analysis locally, as uploading large quantities of data to the webserver may be troublesome. For a local installation of MINTyper, please see MINTyper on github.

View the version history of this server.

Single reference of your choosing Note: If you would like to choose a Vælg fil Der er ikke valgt nogen fil

Select the host database Bacteria organisms (KmerFinder DB)

Motif masking Volume Vo

Prune significance	
Significant calls only	~

#### Pruning length:

The pruning length should be non-negative - the default is 10

Cluster length:

Maximum SNP distance to determine if two isolates belongs to the same cluster.

10

Input files: fastq files (gzipped also suppoerted). Fasta file assemblies will be ignored and excluded from analysis. Note: 2 or more samples are required as input!

#### Please do not upload more than 50 isolates.





# DTU

# **Visualisation of newick files**

- Several tools available
  - FigTree (http://tree.bio.ed.ac.uk/software/figtree/)
  - iTOL (https://itol.embl.de/)
  - Microreact online tool
    - <u>https://microreact.org/upload</u>





# **Survey questions**

- Identifier: Name of contact/Country/Institution
- Species ID
- AMR genes (NDM)
- MLST's
- Hypervirulence
- Part of a cluster?
  - How many SNPs

DTU





# Scenario

Location: Hospital X, EU

"Recently, there has been a rise in infections caused by K. pneumoniae in the intensive care unit (ICU) of hospital X. Initial phenotypic and PCR testing has shown the presence of NDM-positive isolates from patients in the ICU of hospital X and patient-to-patient spread is suspected. This has led to whole genome sequencing (WGS) of three isolates from three unrelated patients (by 'time, place and person' analysis). Cluster analysis and SNP- analysis of the WGS-data has revealed high clonal relatedness between the three isolates".







# **Scenario-Roles**

## **Outbreak Management Team (OMT)**

- Inter-department communication
- Patient health records
- Environmental assessment (e.g. identification of contaminated food or food handling equipment, infection control breaches, cleaning, environmental sampling)
- epidemiological data (e.g. movements and contacts of cases)
- Laboratory data (e.g. whole genome sequencing)

## Your Role

Your (the exercise participants) role is to support OMT in the analyses of epidemiology and laboratory data (including WGS data) to generate a hypothesis of the most likely exposure that has caused the outbreak and develop effective prevention and control measures.





# Injects 1.1 to 1.3







# Inject 2



Further K. pneumoniae identified at ICU, Hospital X

Actions: MLST, cluster analysis

Inject 2





# Injects 3.1 to 3.3







# **Injects 4.1 to 4.2**







May 10: Submit answers for survey 2 ©

# Inject 5



In hospital X, K. pneumoniae isolates from the archive were also sequenced

Actions: cluster analysis

Inject 5

## 



# Hypervirulent clones of Klebsiella pneumoniae

- Hypervirulent K. pneumoniae (hvKp) is more virulent than the classical K. pneumoniae (cKp)
  - rmpA & rmpA2 are responsible for hypermucoidity
  - iucA plays role in the synthesis of aerobactin (siderophores)
  - iroB plays role in the synthesis of salmochelin (siderophore)
  - peg-344 (a metabolic transporter of unknown function)
- K-PAM (<u>https://www.iith.ac.in/K-PAM/predhv.php</u>)
  - Tool to predict *Klebsiella* serotype and hypervirulence using WGS data (Tutorial in Inject 1.1, Annex 2)





# ScienceData

DTU

• Link to the workshop data:

https://sciencedata.dk/shared/dtu.dk/3392574dc38c8b5bbb629e024eeda0a0

Science Data	Science Data		
Multidisciplinary Workshop May 2023	Multidisciplinary Workshop May 2023 Inject 1.2		
Inject 1.1 (Particpant instructions) Inject 1.2	□ Name ▲		
Inject 1.3 (Survey 1)	CSIPhylogeny		
Inject 3.1	🗆 📔 FASTA files		
Inject 3.2 Inject 3.3	🗆 🔚 FASTQ files		
🗆 💼 Inject 4.1	D Inject 1.2.pdf		
Inject 4.2 (Survey 2) Inject 5 Inject 5	3 folders and 1 file		