

DTU

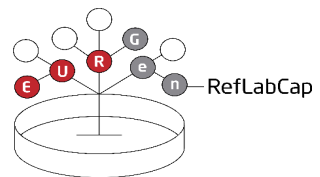


Simulated exercise on multispecies CPE outbreak

2nd Meeting – Conclusion and Discussion

EURGen-RefLabCap
Virtual multidisciplinary training workshop
22 November 2024

Carmen Espinosa-Gongora, Jette S. Kjeldgaard & Faisal Khan, Belen Martin
(cesg@food.dtu.dk jetk@food.dtu.dk fakh@food.dtu.dk)



Agenda for today

- **Scenario Background**
 - Overview of CPE study in UK
 - Scenario
 - Injects
- **Plasmid analysis using Pling**
- **Presentation of exercise results (outbreak investigation)**
 - Survey 1
 - Survey 2
- **Questions/comments from participants**

Simulated exercises - Scenario

- Scenario was based on a retroactive study of CPE in Cambridge University Hospital, UK
 - CPE collected between 2014 to 2020 (n=165)
 - Mostly nosocomial infections
 - Possible horizontal transfer of carbapenemases between isolates
- Sequence data is available at
 - <https://www.ebi.ac.uk/ena/browser/view/PRJEB30134>

MICROBIAL GENOMICS

RESEARCH ARTICLE

Roberts et al., *Microbial Genomics* 2023;9:001048
DOI 10.1099/mgen.0.001048



Long-read sequencing reveals genomic diversity and associated plasmid movement of carbapenemase-producing bacteria in a UK hospital over 6 years

Leah W. Roberts^{1,2}, David A. Enoch³, Fahad Khokhar⁴, Grace A. Blackwell⁵, Hayley Wilson², Ben Warne^{2,4}, Theodore Gouliouris^{2,3,6}, Zamin Iqbal^{1,*} and M. Estée Török^{2,3,6,*}

Abstract

Healthcare-associated infections (HCAs) affect the most vulnerable people in society and are increasingly difficult to treat in the face of mounting antimicrobial resistance (AMR). Routine surveillance represents an effective way of understanding the circulation and burden of bacterial resistance and transmission in hospital settings. Here, we used whole-genome sequencing (WGS) to retrospectively analyse carbapenemase-producing Gram-negative bacteria from a single hospital in the UK over 6 years (n=165). We found that the vast majority of isolates were either hospital-onset (HAI) or HCAI. Most carbapenemase-producing organisms were carriage isolates, with 71% isolated from screening (rectal) swabs. Using WGS, we identified 15 species, the most common being *Escherichia coli* and *Klebsiella pneumoniae*. Only one significant clonal outbreak occurred during the study period and involved a sequence type (ST)78 *K. pneumoniae* carrying *bla*_{NDM-1} on an IncFIB/IncHI1B plasmid. Contextualization with public data revealed little evidence of this ST outside of the study hospital, warranting ongoing surveillance. Carbapenemase genes were found on plasmids in 86% of isolates, the most common types being *bla*_{NDM}- and *bla*_{OXA}-type alleles. Using long-read sequencing, we determined that approximately 30% of isolates with carbapenemase genes on plasmids had acquired them via horizontal transmission. Overall, a national framework to collate more contextual genomic data, particularly for plasmids and resistant bacteria in the community, is needed to better understand how carbapenemase genes are transmitted in the UK.

<https://doi.org/10.1099/mgen.0.001048>

Scenario

- Hospital X has faced an increasing incidence of carbapenemase-producing Enterobacterales (CPE) and wants to study CPE infections from past two years
- **Objectives**
 - Characterize 26 CPE from Hospital X
 - Identify genetic mechanisms and find prevalent clones
 - Investigate the role of plasmids and horizontal gene transfer in the spread of carbapenemases in hospital

Inject 1

- WGS data of 26 CPE isolates was provided
 - *E. coli* (n=9), *K. pneumoniae* (n=15), *K. quasipneumoniae* (n=1), *Enterobacter hormaechei* (n=1)

Objective

- Characterise the CPE isolates and complete the missing information
 - Species identification
 - MLST
 - Identification of carbapenemase gene
- Identify prevalent clones using cluster analysis and metadata

Inject 2

- Plasmid typing data and sequences of reconstructed plasmids from MOB-suite were provided.

Objective

- Investigate potential transmission within and between CPE species
 - Formulate hypothesis based on plasmid typing data from MOB-typer and MGE Finder
 - Prove or disprove hypothesis based on the analysis of plasmid sequencing data
 - Pling
 - Other methods

Plasmid Genomes

Rapid rearrangements and structural changes:

- Gain/loss of genes, rearrangements
 - Recombination mediated by repeat elements (within a cell)
 - Cointegrate to form a new combined plasmid
-
- The analysis/tool should take evolutionary changes into account

Pling: Tool for plasmid phylogenies

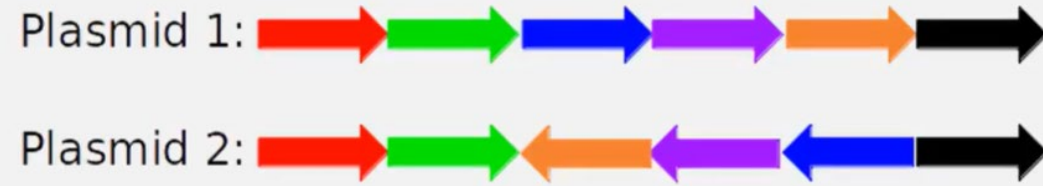
- Uses the Double Cut and Join (DCJ) Indel model, which calculates the minimum **number** of DCJs (insertions, and deletions) to transform one plasmid into another.
- It combines:
 - 1. Containment distance (CD)**, representing the level of shared DNA sequence
 - If CD is below a threshold (default 0.5, or **0.3** for transmission analysis), it continues to calculate the DCJ-Indel distance
 - If CD is above the threshold, the strain is no further included in the network analysis
 - 2. The DCJ-Indel distance** represents how many structural changes differentiate one plasmid from another

Containment distance (CD)

- Represents the level of **shared sequence**.
- It measures the number of matches between the sequence of two plasmids, and the specified length of the matches, considering the total plasmid length

Double Cut and Join (DCJ) model

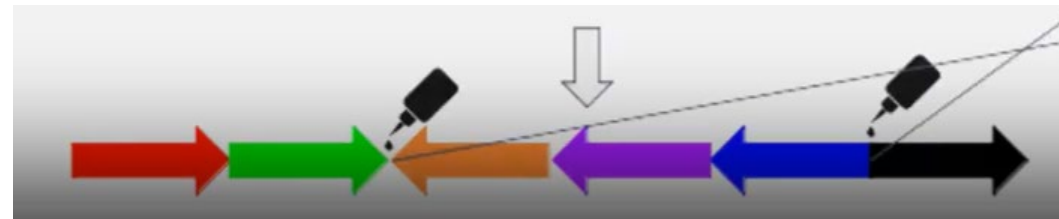
- A coarse (zoomed out) sequence analysis
- Genomes represented as sequence of blocks (genes or aligned blocks)



2 plasmids differed by one inversion event
DCJ distance = 1



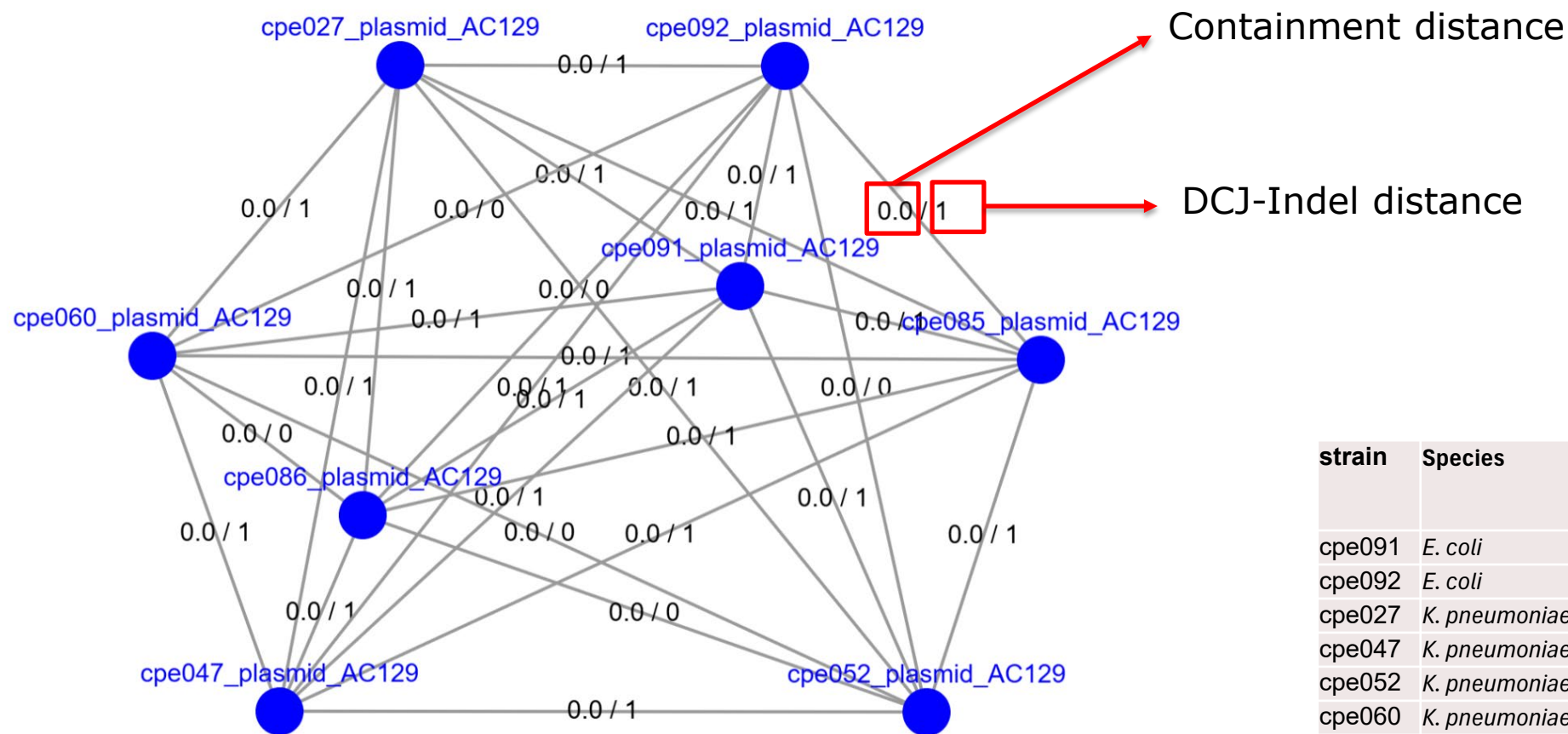
Step 1: Double cut



Step 2: Join

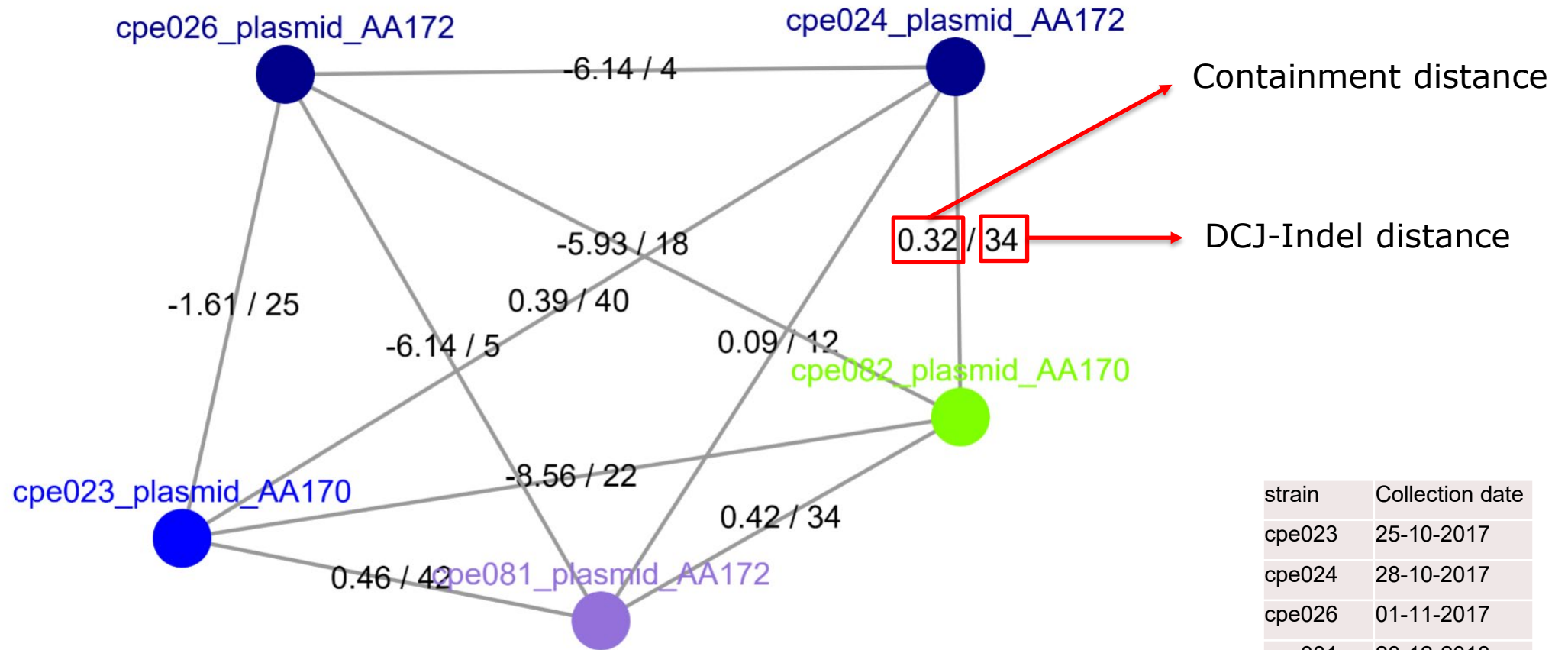
1 double cut join operation =
DCJ 1

HT1: three clones (one *E. coli* and two Kpn) OXA-232, CoIKP3



strain	Species	Collection date
cpe091	<i>E. coli</i>	10-04-2019
cpe092	<i>E. coli</i>	11-04-2019
cpe027	<i>K. pneumoniae</i>	01-11-2017
cpe047	<i>K. pneumoniae</i>	04-05-2018
cpe052	<i>K. pneumoniae</i>	21-07-2018
cpe060	<i>K. pneumoniae</i>	11-05-2017
cpe085	<i>K. pneumoniae</i>	14-02-2019
cpe086	<i>K. pneumoniae</i>	19-02-2019

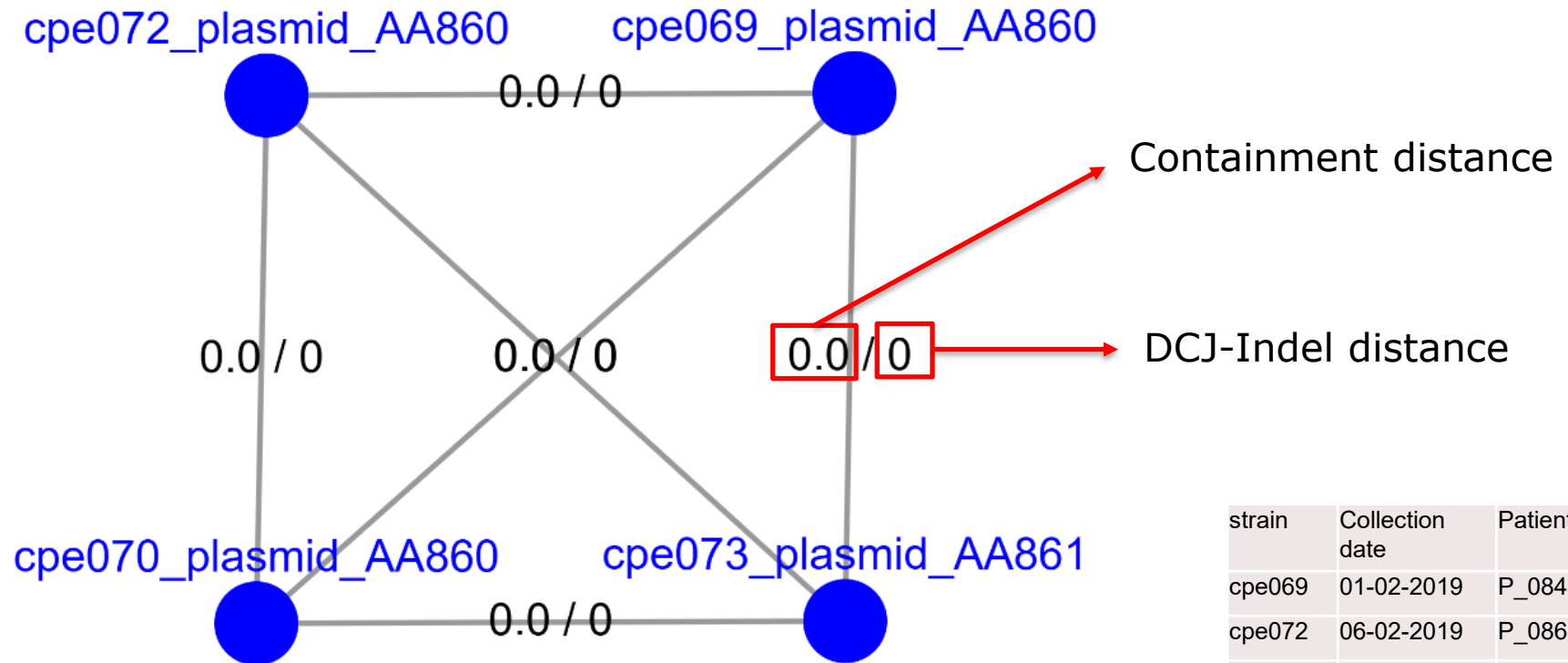
HT2: different *E. coli* backgrounds NDM-5, IncFIA, IncFII



Community = $CD < 0.5$ (or 0.3) + $DCJ\text{-}Indel < 4$

strain	Collection date
cpe023	25-10-2017
cpe024	28-10-2017
cpe026	01-11-2017
cpe081	28-12-2018
cpe082	30-12-2018

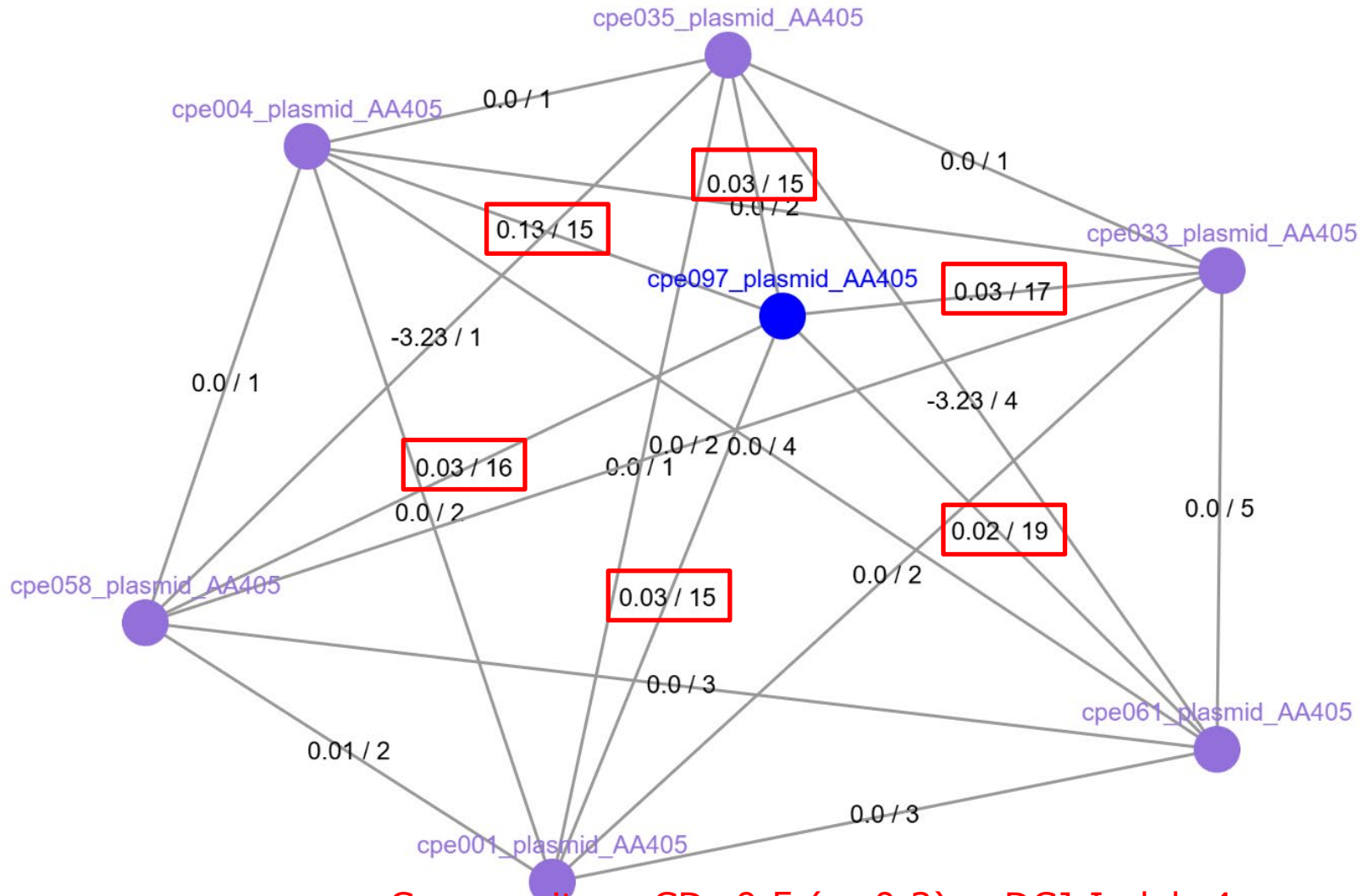
HT3: IP1 (*E. coli* and *K. pneumoniae*) and IP2 (*E. coli* and *K. pneumoniae*) NDM-1, IncA/C2



strain	Collection date	Patient
cpe069	01-02-2019	P_084
cpe072	06-02-2019	P_086
cpe070	01-02-2019	P_084
cpe073	06-02-2019	P_086

Community = $CD < 0.5$ (or 0.3) + $DCJ\text{-}Indel < 4$

HT4: All *K. pneumoniae* + *K. quasipneumoniae* NDM-1, IncHI1B/IncFIB



strain	Collection date
cpe001	06-07-2016
cpe004	09-12-2016
cpe033	24-12-2017
cpe035	27-12-2017
cpe058	20-02-2017
cpe061	13-10-2018
cpe097	01-10-2019

Community = $CD < 0.5$ (or 0.3) + $DCJ\text{-}Indel < 4$

Results task 1 + 2 (Survey 1)

- Thank you for submitting your answers!
- 21 responses
 - 15 countries
 - 2 anonymous
 - Overall very nice concordance 😊

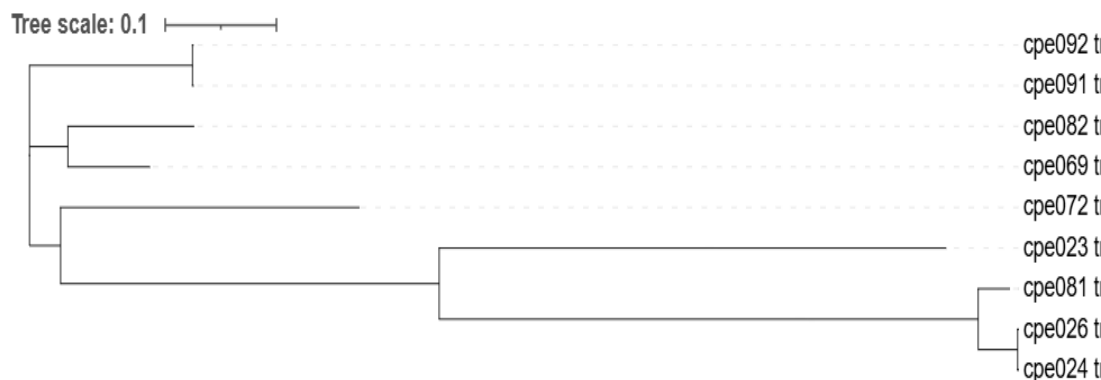
Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Look at the CSI output files (CSI link and newick and snp matrix files provided) and interpret the overall cluster analysis for *E. coli* strains.

Q1A For *E. coli*: Do you identify any clusters? (Mark the best matching option)

Answers	#
Yes, two pair-wise matches and sporadic cases	19
Yes, two clusters and some sporadic cases	1
Yes, several small clusters and pairwise matches	1

Cluster?
Consider meta data -
could be same patient

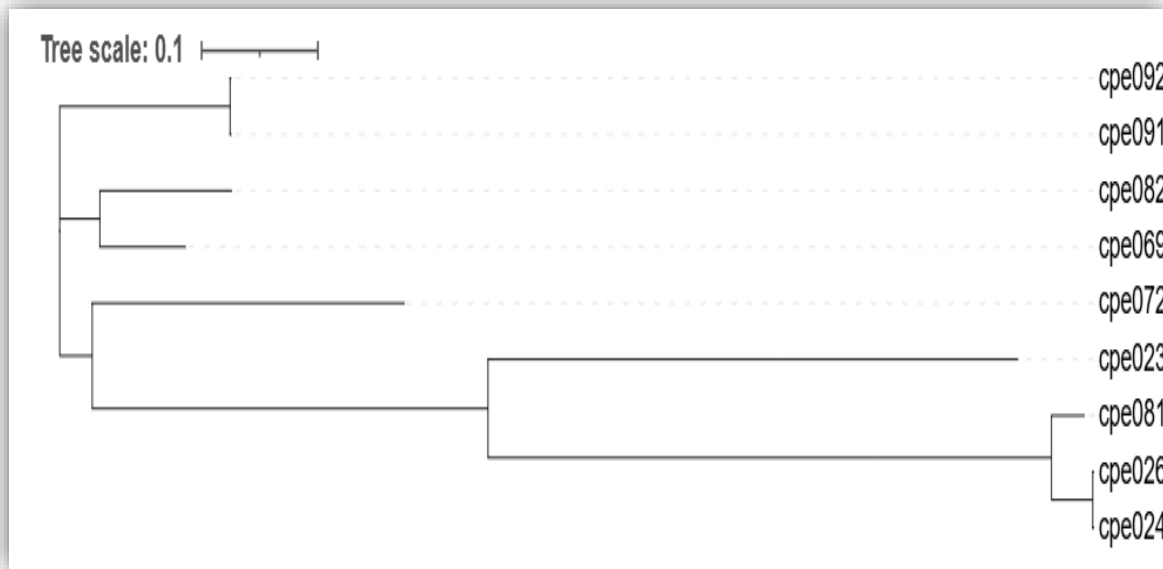


E. coli	cpe023_tri	cpe024_tri	cpe026_tri	cpe069_tri	cpe072_tri	cpe081_tri	cpe082_tri	cpe091_tri	cpe092_tri
cpe023	0	33357	33352	32901	35977	33563	34938	35440	35436
cpe024	33357	0	7	34240	38101	3934	36396	36877	36873
cpe026	33352	7	0	34233	38094	3929	36389	36870	36866
cpe069	32901	34240	34233	0	19799	34053	10520	13198	13194
cpe072	35977	38101	38094	19799	0	37145	19493	20604	20600
cpe081	33563	3934	3929	34053	37145	0	35802	36673	36669
cpe082	34938	36396	36389	10520	19493	35802	0	15107	15103
cpe091	35440	36877	36870	13198	20604	36673	15107	0	6
cpe092	35436	36873	36866	13194	20600	36669	15103	6	0
min: 6 max: 38101									

Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Q1B Do the E. coli clustering align with MLST data?

A: Yes, clusters and/or pairs each have the same MLST type (16) (most...; 5)



Patient #	MLST	Carba-gene
P 094	Unknown (410)	<i>blaOXA-232</i>
P 093	Unknown (410)	<i>blaOXA-232</i>
P 079	Unknown	<i>blaNDM-5</i>
P 084	Unknown	<i>blaNDM-1</i>
P 086	13857?	<i>blaNDM-1</i>
P 037	405	<i>blaNDM-5</i>
P 078	648	<i>blaNDM-5</i>
P 040	648	<i>blaNDM-5</i>
P 038	648	<i>blaNDM-5</i>

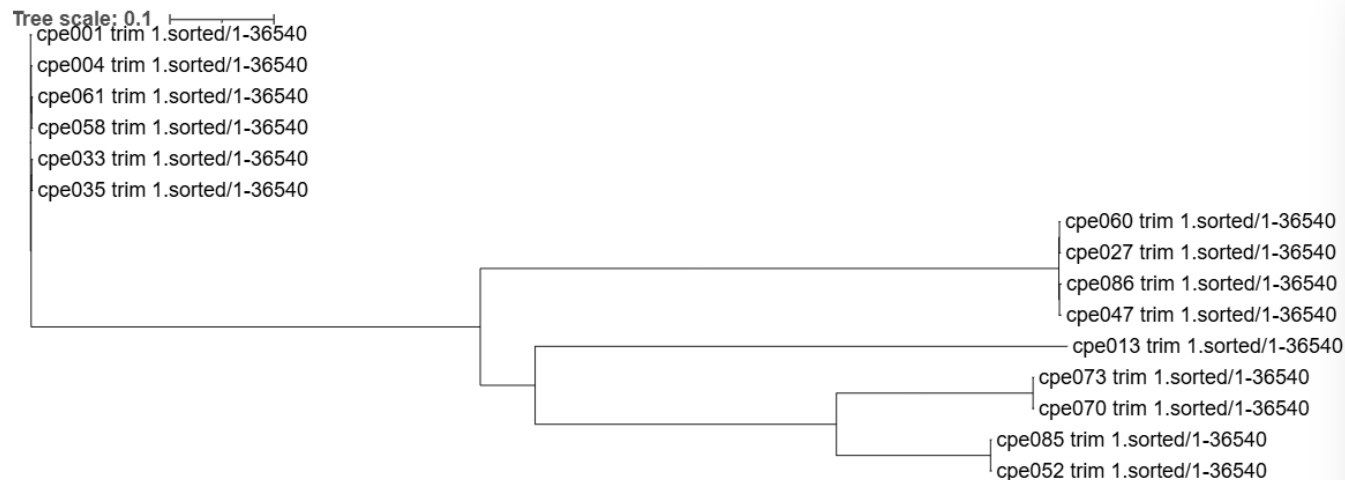
Comments:

- Strain cpe081 has the same ST as other strains but doesn't belong to the same cluster
- We found out that cpe091 and cpe092 have the same MLST 410 on Seqsphere software using the hybrid assembly

Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Look at the CSI output files (CSI link and newick and snp matrix files provided) and interpret the overall cluster analysis for *E. coli* strains.

Q2A For *Klebsiella*: Do you identify any clusters? (Mark the best matching option)

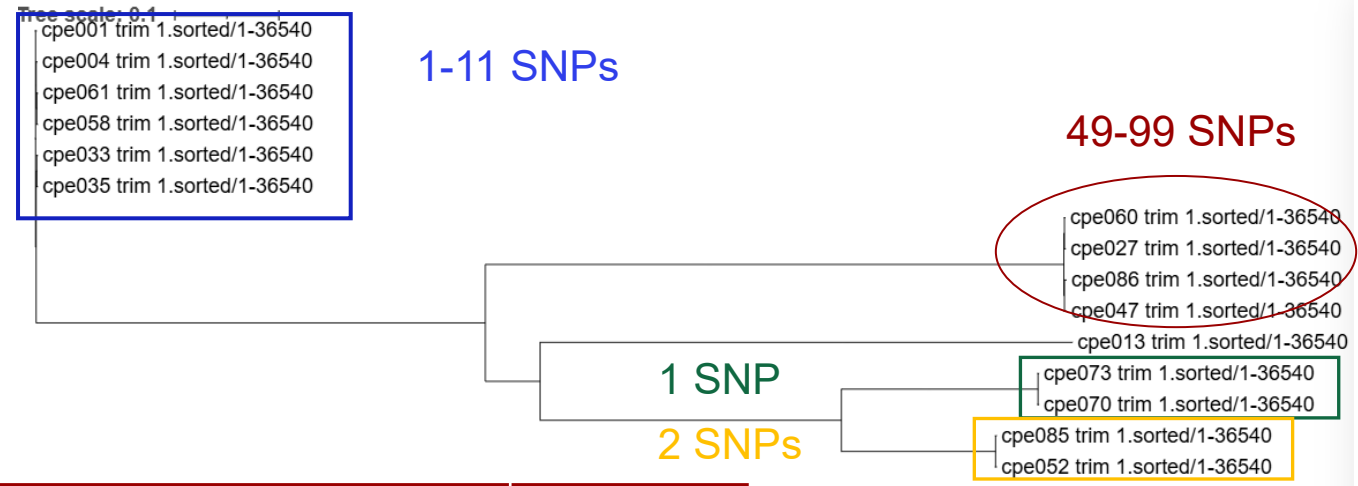


	cpe001/	cpe004/	cpe013/	cpe027/	cpe033/	cpe035/	cpe047/	cpe052/	cpe058/	cpe060/	cpe061/	cpe070/	cpe073/	cpe085/	cpe086/
cpe001/	0	3	18362	17424	5	6	17443	17889	4	17417	7	16888	16887	17889	17450
cpe004/	3	0	18363	17425	6	7	17444	17890	1	17418	4	16889	16888	17890	17451
cpe013/	18362	18363	0	19503	18365	18366	19522	16306	18364	19496	18367	18300	18299	16306	19531
cpe027/	17424	17425	19503	0	17427	17428	83	19176	17426	49	17429	19096	19095	19176	86
cpe033/	5	6	18365	17427	0	3	17446	17892	7	17420	10	16891	16890	17892	17453
cpe035/	6	7	18366	17428	3	0	17447	17893	8	17421	11	16892	16891	17893	17454
cpe047/	17443	17444	19522	83	17446	17447	0	19195	17445	70	17448	19115	19114	19195	99
cpe052/	17889	17890	16306	19176	17892	17893	19195	0	17891	19169	17894	8811	8810	2	19203
cpe058/	4	1	18364	17426	7	8	17445	17891	0	17419	3	16890	16889	17891	17452
cpe060/	17417	17418	19496	49	17420	17421	70	19169	17419	0	17422	19089	19088	19169	85
cpe061/	7	4	18367	17429	10	11	17448	17894	3	17422	0	16893	16892	17894	17455
cpe070/	16888	16889	18300	19096	16891	16892	19115	8811	16890	19089	16893	0	1	8811	19122
cpe073/	16887	16888	18299	19095	16890	16891	19114	8810	16889	19088	16892	1	0	8810	19121
cpe085/	17889	17890	16306	19176	17892	17893	19195	2	17891	19169	17894	8811	8810	0	19203
cpe086/	17450	17451	19531	86	17453	17454	99	19203	17452	85	17455	19122	19121	19203	0
min: 1 max: 19531															

Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Look at the CSI output files (CSI link and newick and snp matrix files provided) and interpret the overall cluster analysis for *E. coli* strains.

Q2A For *Klebsiella*: Do you identify any clusters? (Mark the best matching option)



Answers	#
Yes, one big cluster and some pair-wise matches	9
Yes, two clusters, pairwise match(es) and sporadic case(s)	11
Yes, several small clusters and pairwise matches	1

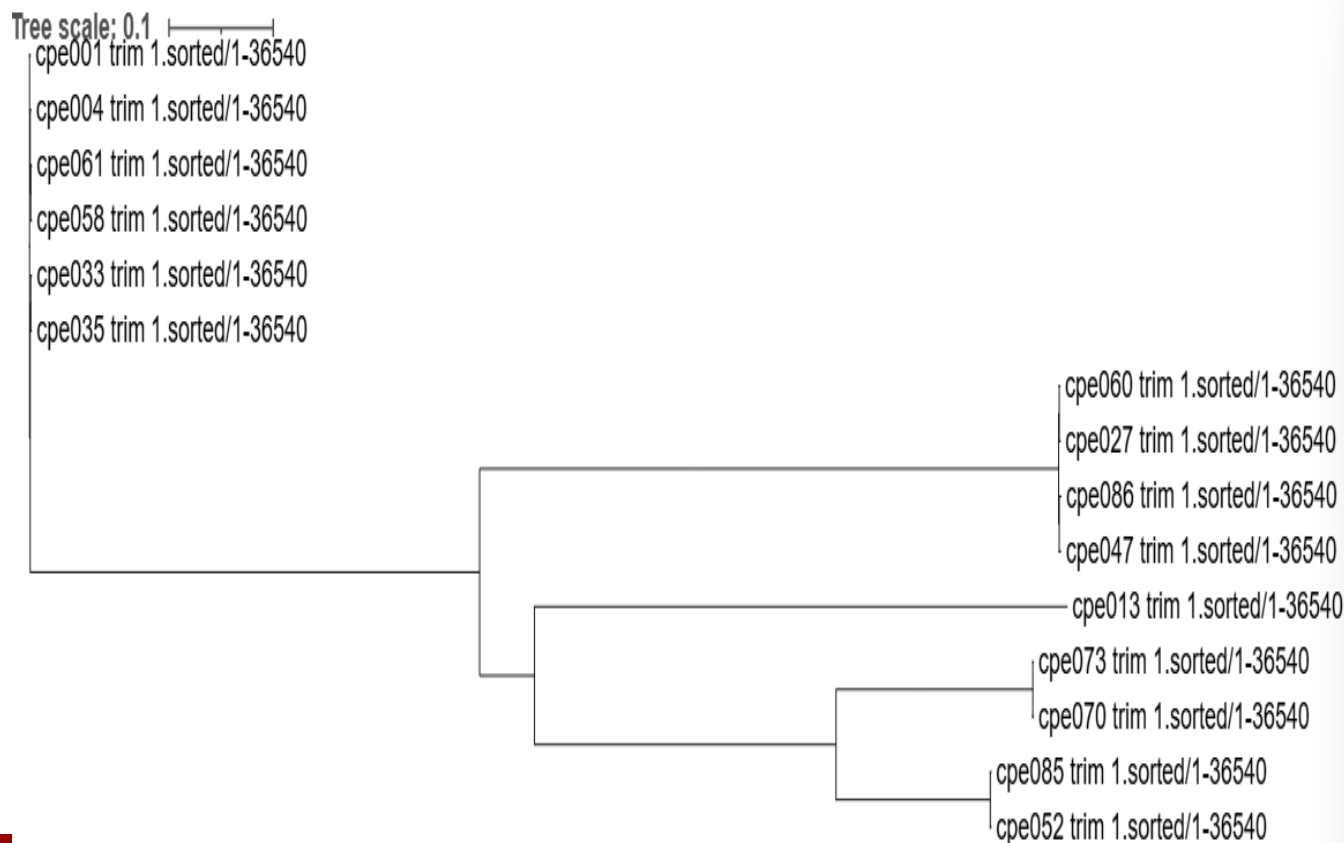
Comments:

- We would actually say we have 3 clusters, 1 big cluster (with 6 isolates) and 2 small ones (with pair wise matches)
- Two major clusters of MLST 78 and 231, two pairwise of MLST 2497 and 11

Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Q2B Do the Klebsiella clustering align with MLST data?

A: Yes, clusters and/or pairs each have the same MLST type (19) (most...; 2)



Patient #	MLST	Carba-gene
P 010	78	<i>bla</i> NDM-1
P 016	78	<i>bla</i> NDM-1
P 018	78	<i>bla</i> NDM-1
P 019	78	<i>bla</i> NDM-1
P 008	78	<i>bla</i> NDM-1
P 008	78	<i>bla</i> NDM-1
P 133	231	<i>bla</i> OXA-232
P 038	231	<i>bla</i> OXA-232
P 088	231	<i>bla</i> OXA-232
P 056	231	<i>bla</i> OXA-232
P 028	392	<i>bla</i> NDM-6
P 086	11	<i>bla</i> NDM-1
P 084	11	<i>bla</i> NDM-1
P 087	2497	<i>bla</i> OXA-232
P 064	2497	<i>bla</i> OXA-232

Task 1: Characterise the carbapenemase-producing Enterobacterales collection

Q3A: Due to poor sequence quality, some MLST results may be inconclusive. Which data did you use and why?

Answers	#
Illumina fasta	13
Illumina fastq	7
ONT assemblies	(2)
Both	2

Q3B: If you tried with several types of data files, which one produced better MLST result?

- Illumina data proved to be better than ONT
- Hybrid & Illumina

Comments:

- We used hybrid assemblies when possible, due to the higher completeness of the genome sequences and the higher accuracy associated with the nucleotide sequences
- We used our own assembly pipeline on the fastq files to be able to compare and see if the results match with the fasta files you delivered, to compare the assembly pipeline
- We used hybrid assemblies on Seqsphere software and we had the MLST result for the unknown ST in the Excel table

Q3C: Rank the following file formats that you expect would generate a phylogenetic analysis of highest quality

Options:

- a. Raw fastq Illumina files
- b. Raw ONT files
- c. Illumina assemblies
- d. ONT assemblies
- e. Hybrid assembly (Illumina and ONT)

1st choice	Last choice
Hybrid assembly (12) (Illumina and ONT)	Raw ONT Files (12)
Raw fastq Illumina (6)	ONT assemblies (7)
Illumina assemblies (2)	Illumina assemblies (1)

Task 2: Investigate resistance determinants

- Q4A: Which species did you determine for strain cpe093?
 - *Enterobacter hormaechei*
- Q4B: Look at the combination of carbapenemase gene and plasmid(s) in strain cpe093, and interpret if this strain is likely a part of a multi-species outbreak
 - **No, the strain does not carry the same gene and plasmid combination as other potential outbreak strains (17)**
 - Yes, the strain carries the same gene and plasmid combination as other potential outbreak strains (**4**)

NDM-1 & IncHI2A/IncHI2

The plasmid differ from all other strains

Task 2: Investigate resistance determinants

- Q4A: Which species did you determine for strain cpe097?
 - *Klebsiella quasipneumoniae* (subsp. Similipneumoniae)
- Q4B: Look at the combination of carbapenemase gene and plasmid(s) in strain cpe093, and interpret if this strain is likely a part of a multi-species outbreak
 - **Yes, the strain carries the same gene and plasmid combination as other potential outbreak strains (20)**
 - No, the strain does not carry the same gene and plasmid combination as other potential outbreak strains (1)

NDM-1 IncHI1B/IncFIB The plasmid and gene combination is identical as with other strains

Analysis suggests possible horizontal transmission of this gene-plasmid combination to other species

Strain	Species (KmerFinder)	Collection date	Isolation source	Patient ID	MLSTFinder	Carbapenemase	Plasmid MGE
cpe001	K. pneumoniae	06-07-2016	BAL	P_010	78	NDM-1	IncHI1B/IncFIB
cpe004	K. pneumoniae	09-12-2016	Screen	P_016	78	NDM-1	IncHI1B/IncFIB
cpe033	K. pneumoniae	24-12-2017	Sputum	P_008	78	NDM-1	IncHI1B/IncFIB
cpe035	K. pneumoniae	27-12-2017	Blood	P_008	78	NDM-1	IncHI1B/IncFIB
cpe058	K. pneumoniae	20-02-2017	Urine	P_019	78	NDM-1	IncHI1B/IncFIB
cpe061	K. pneumoniae	13-10-2018	Screen	P_018	78	NDM-1	IncHI1B/IncFIB
cpe097	K. quasipneumoniae	01-10-2019	Screen	P_103	Unknown	NDM-1	IncHI1B/IncFIB

Task 2: Investigate resistance determinants

Q6A: According to the data received in this inject, is the NDM-1 gene in *E. coli* cpe069 likely located on a plasmid or on the chromosome?

- Plasmid **(20)**
- Analysis suggests possible **horizontal transmission** of this gene-plasmid combination to **other *E. coli* strains**
- Analysis suggests possible **horizontal transmission** of this gene-plasmid combination to **other species**
- Analysis suggests **clonal transmission of an *E. coli* clone** with this gene-plasmid combination

Strain	Species (KmerFinder)	Collection date	Isolation source	Patient ID	MLSTFinder	Carbapenemase	Plasmid MGE	Gene location
cpe069	<i>E. coli</i>	01-02-2019	Screen	P_084	Unknown	NDM-1	IncA/C2	Plasmid
cpe072	<i>E. coli</i>	06-02-2019	Screen	P_086	13857?	NDM-1	IncA/C2	Plasmid
cpe070	<i>K. pneumoniae</i>	01-02-2019	Screen	P_084	11	NDM-1	IncA/C2	Plasmid
cpe073	<i>K. pneumoniae</i>	06-02-2019	Screen	P_086	11	NDM-1	IncA/C2	Plasmid

Task 2: Investigate resistance determinants

- Q7A: For strain cpe082, MGEFinder did not indicate the presence of NDM-5 on a plasmid. How do you interpret this result?
 - **NDM-5 may still be on a plasmid, but MGEFinder could not link it to a plasmid replicon (19)**
 - The result suggest NDM-5 is likely chromosomal in this strain (2)
 - No, this analysis is too unsure
- Q7B: Which type of data files would you prefer use to determine gene locations, and why?
 - Hybrid assembly
 - due to its ability to combine the strengths of different sequencing technologies and produce a more accurate and complete genome assembly.
 - ONT files (raw data or assembly) with their ultra-long reads and low number of contigs. This promotes genomic assembly including complete plasmids

Results Task 3+4 (Survey 2)

- Thank you for submitting your answers!
- More challenging questions and new tool
- 12 responses
 - 10 countries

Task 3 - Formulate hypotheses of transmission pathways

- Here you will plan your analysis to assess the role of clonal transmission and horizontal gene transfer in the spread of resistance within the hospital
 - Perform analysis to assess which strains you hypothesize to be clonally transmitted or may have been involved in horizontal transmission

Strategy for analysis?

Q1A: What type of analysis do you need to confirm the hypotheses you have generated in Task 3?

- PlasmidFinder, MOB-typing
- Plasmid phylogenetic analysis
 - calculate phylogenetic distances between plasmids, **look at patient contact evaluation**
- Plasmid comparison/ Identify and compare the genetic backbone, structural changes and full sequence of plasmids
- cgMLST for clonal transmission.

Indicate at least two methods to obtain the plasmid sequences

- Hybrid assembly
 - MOB-recon tool outputs all plasmid sequences in separate fasta files for every sample
 - MOB-suite/MOB-typer/MOB-recon
 - Pling
 - Bandage
 - plasmidSPAdes
 - PlasmidFinder
-
- Other option is use the Illumina files and perform the combination of BLAST searches and PCR, followed by Sanger sequencing, to close the gaps between plasmid contigs

Task 3 - Formulate hypotheses of transmission pathways

- Q2: While analysing strain cpe004, you observed multiple replicon types, i.e., IncHI1B and IncFIB. How might these replicon types influence the outcomes in plasmid reconstruction and subsequent bioinformatics analyses?
 - **Both replicon types are part of the same plasmid, suggesting I will obtain a single reconstructed plasmid containing both IncHI1B and IncFIB (10)**
 - **The presence of multiple replicons may complicate the identification and annotation of plasmids, or cause ambiguous phylogenetic analysis and evolutionary conclusions (9)**
 - **The presence of multiple replicons suggests a fusion event, where two or more plasmids merged into a single structure (7)**

Task 4: Confirm the routes of transmission

For each of the horizontal transmission hypotheses, indicate with an X whether the plasmids involved are conjugative, mobilizable, or non-mobilizable (MOB-typer results):

Hypothesis ID	Conjugative	Mobilizable	Non-mobilizable	Strains
HT1 - ColKP3		X		cpe027, cpe052, cpe047, cpe060, cpe085, cpe086, cpe091, cpe092
HT2 - IncFII/IncFIA/Inc FIB	X			cpe023, cpe024, cpe026, cpe082, cpe081
HT3 IncA/C2	X			cpe069, cpe070, cpe072, cpe073
HT4 IncH/Inc1B/IncFIB	X			cpe001, cpe004, cpe033, cpe035, cpe058, cpe061, (cpe097)*

* cpe097 plasmid is different based on DCJ distance due to several rearrangement events but based on the containment distance, it is very similar to *K. pneumoniae* plasmids

Task 4: Confirm the routes of transmission

- Q3B. How does the type of plasmid (conjugative vs. mobilizable) support each hypothesis of horizontal transmission? Would the presence of non-mobilizable plasmids change your interpretation of the results?

Q4. Using the Pling results, select all statements that you would report to hospital X

- The analysis confirmed the presence of an OXA-232-carrying ColKP3 plasmid conferring resistance to carbapenems in *E. coli* and *K. pneumoniae* strains isolated between 2017-2019, indicating **a plasmid outbreak (10)**
- The analysis supports **direct (clonal) transmission** of *K. pneumoniae* with an NDM-1-carrying IncA/C2 plasmid between the two patients (P_084 and P_086) **(9)**
- The analysis confirmed the circulation of an **NDM-1-carrying IncA/C2 plasmid** among two patients (P_084 and P_086) admitted at hospital X at the same time. Each patient carried *E. coli* and *K. pneumoniae* with this plasmid **(8)**
- The analysis confirmed **horizontal transmission** of an NDM-5-carrying IncFII/IncFIA plasmid among *E. coli* isolates with different genetic backgrounds (chromosome) isolated between 2017-2018, indicating a plasmid outbreak **(7)**
- The analysis confirms that an **NDM-1-carrying plasmid** was acquired by *K. quasipneumoniae* from a *K. pneumoniae* strain repeatedly detected in CPE-positive screening and clinical samples between 2017-2018 **(4)**

Sum up of hypotheses for outbreaks

- Clonal transmission:
 - *K. pneumoniae* with an NDM-1-carrying IncA/C2 plasmid
 - *E. coli* with an NDM-5-carrying IncFII/IncFIA plasmid
- Plasmid transmission:
 - ColKP3 plasmid carrying OXA-232 in both *E. coli* and *K. pneumoniae*
 - IncA/C2 plasmid carrying NDM-1 in *K. pneumoniae*
 - IncFII/IncFIA plasmid carrying NDM-5 in *E. coli*
 - IncHI1B/IncFIB plasmid carrying NDM-1 in *K. pneumoniae* and *K. quasipneumoniae* (Based on CD, but not when considering DCJ distance)

Questions/Comments?