

DTU



Simulated exercise on CRE outbreak – Klebsiella

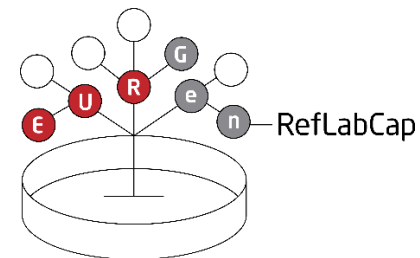
EURGen-RefLabCap

Virtual multidisciplinary training workshop

12 May 2023

Jette S. Kjeldgaard & Faisal Khan

(jetk@food.dtu.dk – fakh@food.dtu.dk)



Scenario

Anatomy of an extensively drug-resistant *Klebsiella pneumoniae* outbreak in Tuscany, Italy

Melissa J. Martin^a, Brendan W. Corey^a , Filomena Sannio^b, Lindsey R. Hall^a , Ulrike MacDonald^{c,d} ,
Brendan T. Jones^a, Emma G. Mills^a, Casey Harless^a, Jason Stam^a, Rosslyn Maybank^a, Yoon Kwak^a,
Katharina Schaufler^{e,f,g} , Karsten Becker^h , Nils-Olaf Hübnerⁱ, Stefania Cresti^{b,j}, Giacinta Tordini^{b,j},
Marcello Valassina^j, Maria Grazia Cusi^{b,j} , Jason W. Bennett^a, Thomas A. Russo^{c,d,k,l} , Patrick T. McGann^a ,
Francois Lebreton^{a,1,2} , and Jean-Denis Docquier^{b,m,1,2} 

- Based on the sequence data from Italian outbreak 2018-2021
- Dataset is available on NCBI under Project ID [PRJNA725484](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA725484)
 - <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA725484>
- The data for non-cluster outbreak from NCBI, including MinION data (Inject 4.1)

Scenario- ICU

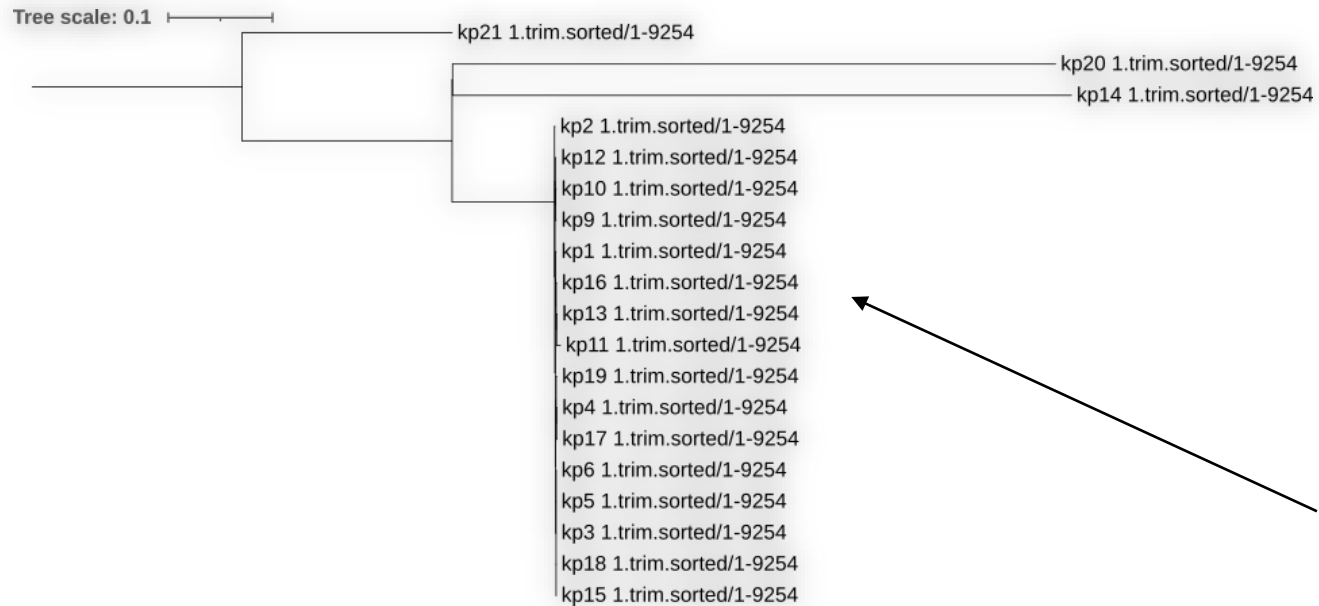
- Initially, 7 cases were identified in the ICU
- All isolates belonged to ST147 and were NDM-1 positive
- Kp7 and Kp8 are not part of the cluster

	kp1	kp2	kp3	kp4	kp5	kp6	kp7	kp8
Kp1	0	3	3	6	3	3	628	411
Kp2	3	0	4	7	4	4	629	412
Kp3	3	4	0	3	0	0	627	410
Kp4	6	7	3	0	3	3	630	413
Kp5	3	4	0	3	0	0	627	410
Kp6	3	4	0	3	0	0	627	410
Kp7	628	629	627	630	627	627	0	925
kp8	411	412	410	413	410	410	925	0



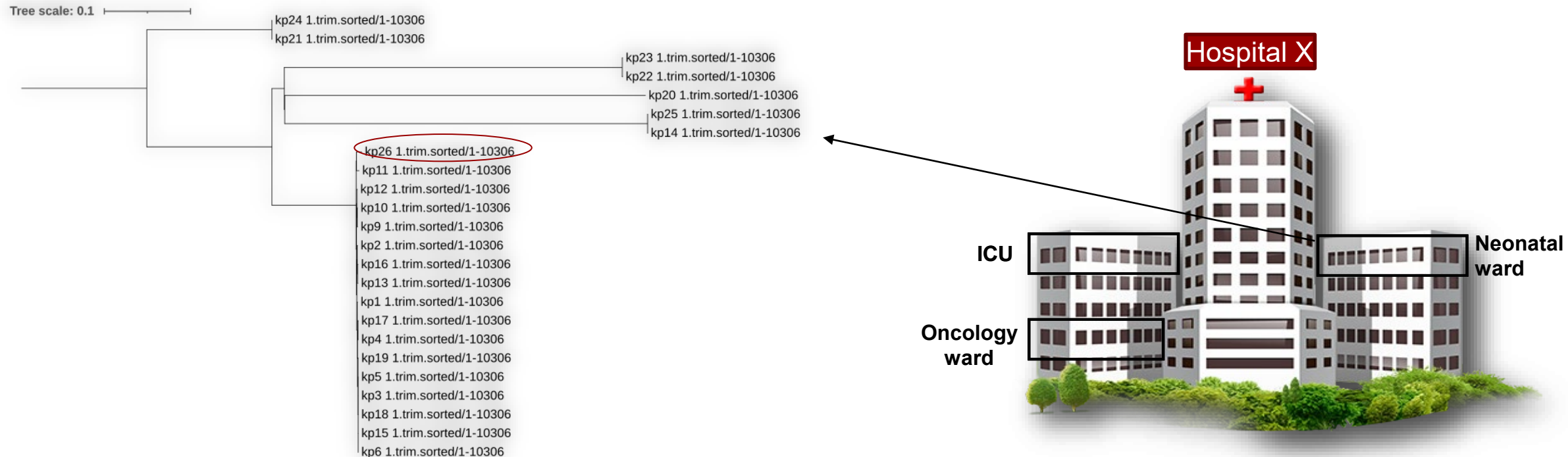
Inject 3.1 – Screening of oncology ward

- Which isolates are excluded from the cluster?
 - Kp14, Kp20, Kp21 are excluded
 - >3000 SNP difference

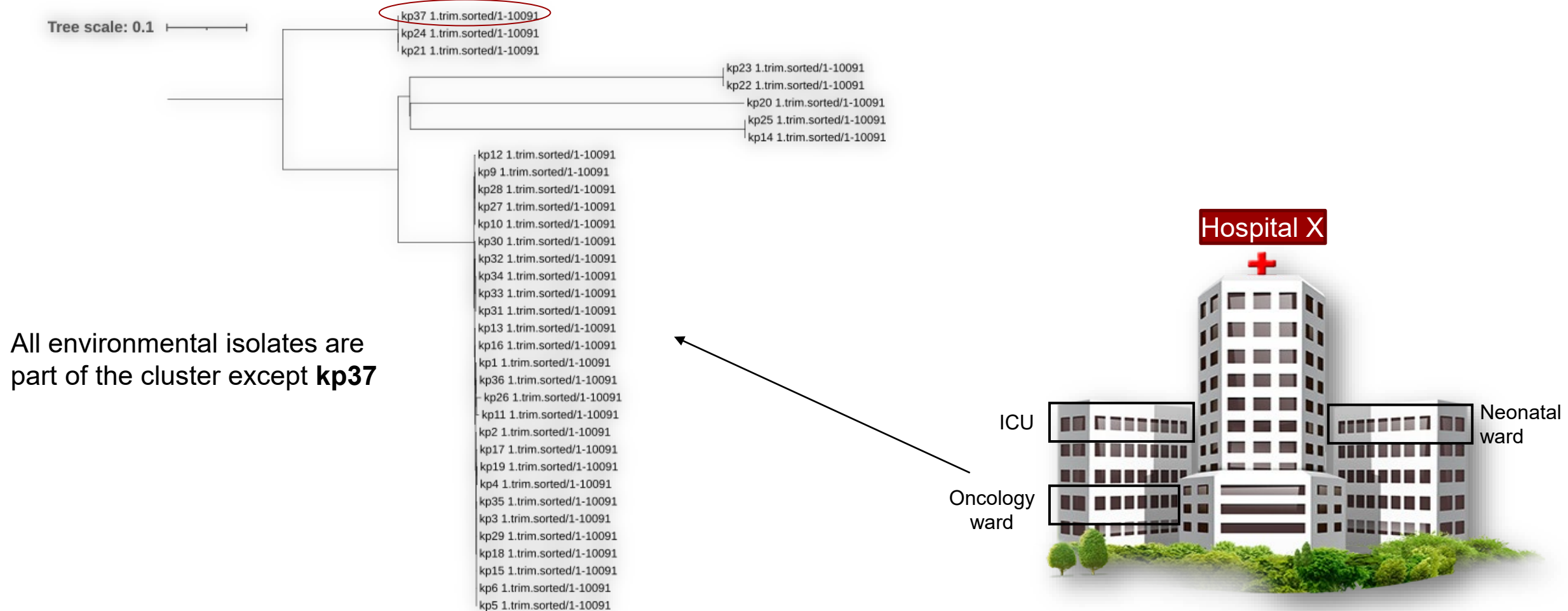


Inject 3.2 – Screening of Neonatal ward

- Which isolates are excluded from the cluster?
 - Kp14, Kp20, Kp21, Kp22, Kp23, Kp24, Kp25 = >3000 SNP difference
 - Kp26 = ~70 SNP difference. Is it NDM-1 positive, hvKp?



Inject 3.3 – Environmental screening



Inject 4- MinION sequences from Hospital Y

- Original minION data was not minION data 😊
- The MinION data that fits our case definition was limited

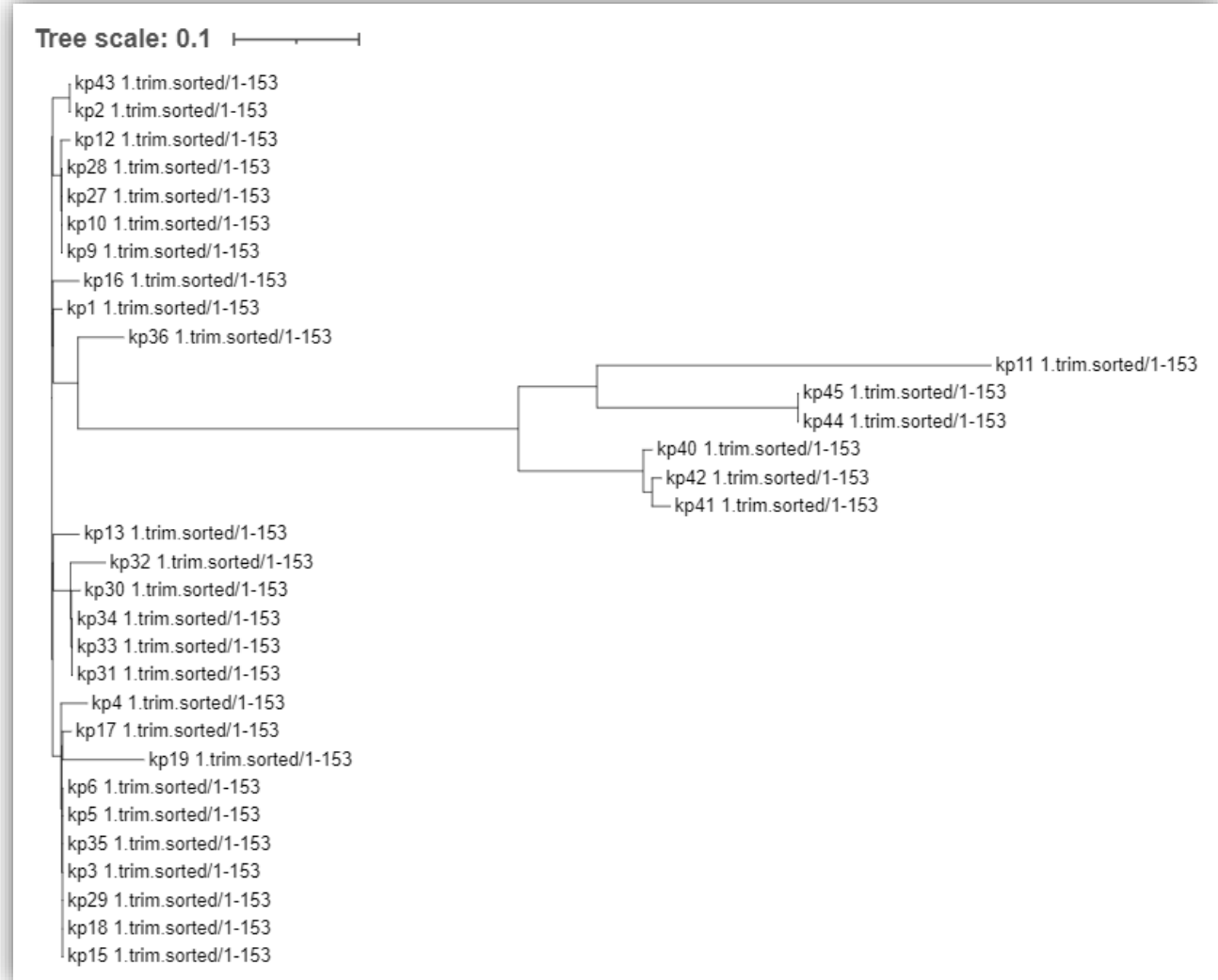


	kp38_trim.fsa	kp39_trim.fsa	kp1_1.fsa	kp2_1.fsa	kp3_1.fsa
kp38_trim.fsa					
kp39_trim.fsa	248				
kp1_1.fsa	18	258			
kp2_1.fsa	18	258	0		
kp3_1.fsa	17	257	1	1	

- **MinTyper: Percentage coverage was only 36%**
- **Problem with the MinION data**

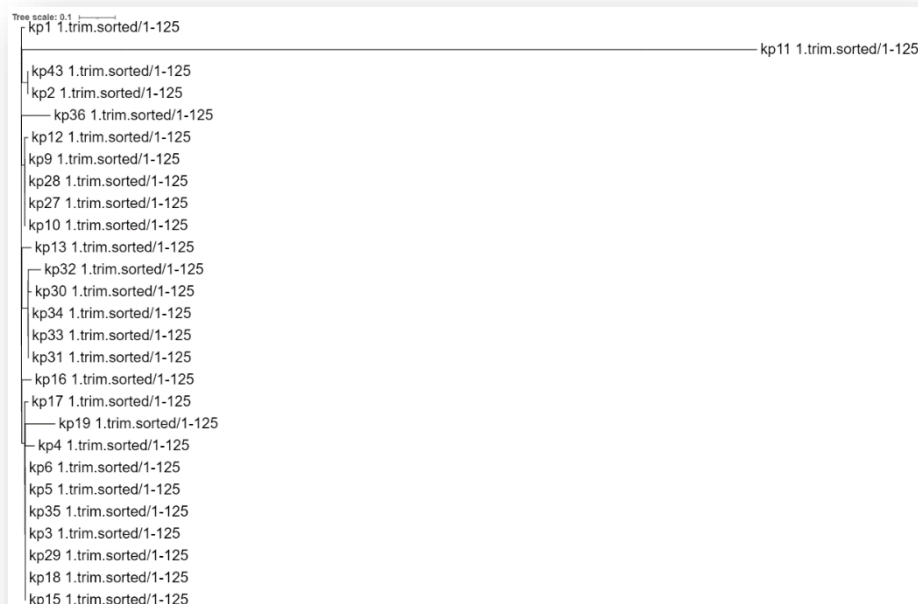
Inject 5 – Archived samples

- Family 1- travel to India (Kp1 and Kp43)
 - Kp43 introduced the outbreak strain to hospital X
 - Rest of the patients are not part of cluster and do not have travel history



The Final Outbreak Cluster

- Not all outbreak isolates have NDM-1 and hypervirulence
- No all ST147 are part of the outbreak cluster?
 - Kp11 is a different ST and is not part of the cluster.



Isolate ID	Inject	NDM-1	Part of the MLST cluster	Hypervirulent
kp1	1,2	Yes	147 Yes	Yes
kp2	1,2	Yes	147 Yes	Yes
kp3	1,2	Yes	147 Yes	Yes
kp4	2	Yes	147 Yes	Yes
kp5	2	Yes	147 Yes	Yes
kp6	2	Yes	147 Yes	Yes
kp9	3,1	Yes	147 Yes	No
kp10	3,1	Yes	147 Yes	Yes
kp11	3,1	Yes	4843 Yes	Yes
kp12	3,1	Yes	147 Yes	Yes
kp13	3,1	Yes	147 Yes	No
kp15	3,1	Yes	147 Yes	Yes
kp16	3,1	Yes	147 Yes	Yes
kp17	3,1	Yes	147 Yes	No
kp18	3,1	Yes	147 Yes	No
kp19	3,1	Yes	147 Yes	No
kp27	3,3	Yes	147 Yes	No
kp28	3,3	Yes	147 Yes	No
kp29	3,3	Yes	147 Yes	Yes
kp30	3,3	Yes	147 Yes	Yes
kp31	3,3	Yes	147 Yes	Yes
kp32	3,3	Yes	147 Yes	Yes
kp33	3,3	Yes	147 Yes	Yes
kp34	3,3	Yes	147 Yes	Yes
kp35	3,3	No	147 Yes	Yes
kp36	3,3	No	147 Yes	Yes
kp43	5	Yes	147 Yes	Yes

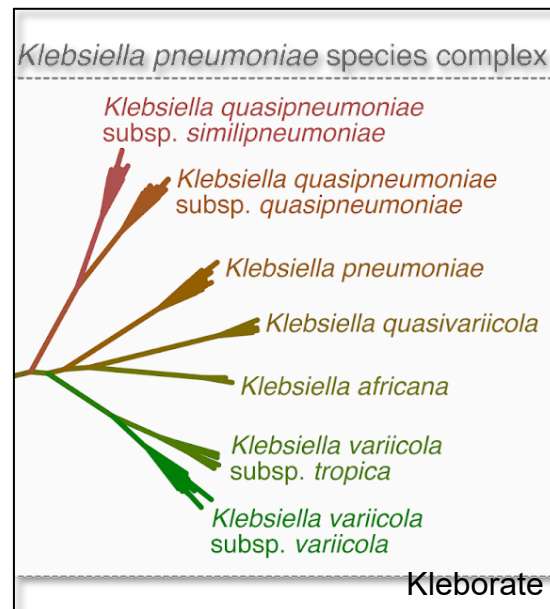
Establishing a case definition-feedback

- How do you make a case definition?
- How specific it should be?
- Your Opinion?

Klebsiella Typing tools



- *K. pneumoniae* species complex (KpSC)
 - Seven species
 - Identified as *K. pneumoniae* in clinical microbiology lab
- Further characterization of KpSC using Kleborate



Species identification

MLST

Virulence genotyping

AMR genotyping

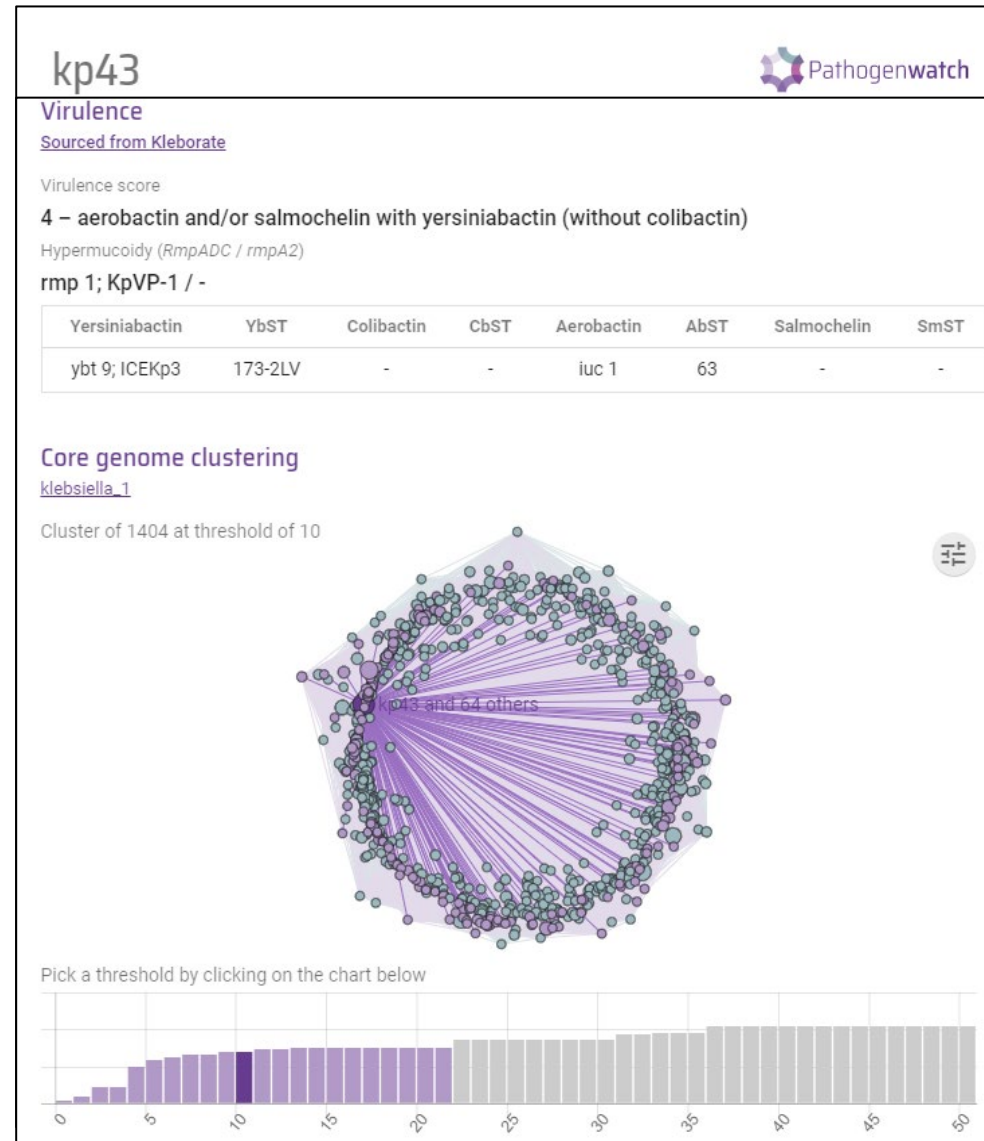
Capsule (K), LPS (O) locus

Assembly QC/stats

Kleborate

Kleborate and Pathogenwatch

- Kleborate is a command-line tool
- Can also be used via Pathogenwatch
 - <https://pathogen.watch/upload>
- Can also do cgMLST for *K. pneumoniae*
- Additional Typing
 - K (capsule) and O (LPS) serotyping
 - Virulence typing
 - Yersiniabctin and aerobactin typing
- Cluster analysis with global isolates



Pathogenwatch and Kleborate Tutorials

- Pathogenwatch Tutorial
 - https://docs.google.com/document/d/1D8IsODA-6_0_g0Bg2b2ZvYDPRRng8QEAgJnnYeHjWw/edit
 - Kleborate Tutorial (Command-line)
 - https://docs.google.com/document/d/19iz_bLGBj2yE3xAOTp_VIIJu0jMlzlAJIPzyfLMDYGo/edit#heading=h.7e1hzu2asyop
 - Other resources
 - <https://klebnet.org/training/>

Survey results

- Two surveys to catch up and see how you interpret the data
- 25 countries has signed up for the workshop
 - 21 countries replied to survey 1 (inject 1-2)
 - 20 countries replied to survey 2 (inject 3-4)
- Some points and questions are raised

Survey results – survey 1

- Characterisation of kp1:

	Responses
Can you confirm the NDM-gene - with which gene variant?	Yes, NDM-1 (100 %)
Which MLST is this isolate?	ST147 (100 %)
Is the isolate classified as hypervirulent? Specify the genes found	Yes - rmpA, rmpA2, iucA, peg344 (100 %) Additionally iroB (33 %) ybt 9 (ICEKp3) (5 %)
For Kp2 and Kp3 - did you find the same characteristics?	Yes (100 %)

Survey results – survey 1

- Tools used:

Tools used	Number of participants (%)
Kmerfinder	14 (66 %)
Resfinder	19 (91 %)
MLST 2.0 (CGE)	16 (76 %)
K-PAM	20 (95 %)
AMRFinder	4 (2 %)
PubMLST	3 (1.4 %)
rMLST	4 (2 %)

- Additional tools used:
 - Bifrost pipeline, Kleborate, Kraken2, Aquamis, Shovill, PlasmidFinder

Survey results – survey 1

Additional results

Question:	Responses
MLST of kp4, kp5, kp6, kp7, kp8	ST147 (100 %)
How many of the eight isolates are part of an outbreak cluster?	6 (81 %) 4/5/7 (~5 % each)
Minimal number of SNP differences in this outbreak cluster	0 SNPs (62 %) 1 SNP (10 %); 3 SNPs (14 %) 5 SNPs (5 %)
Maximal number of SNP differences in this outbreak cluster	7 SNPs (62 %) 3 (5%); 10 (10 %) 413 SNPs (5 %)

Cluster A - Kp3, Kp4, Kp5, Kp6. There is a smaller cluster of 2 isolates Kp1, Kp2 (cluster B)

Kp7 and Kp8 can be also part of the outbreak based on plasmid analysis. Kp7 and Kp8 differs by 1 or 2 plasmids, respectively. These isolates could be from the same source but with plasmid recombination. Further analysis of plasmid SNPs carrying NDM-1 gene would be useful

Understanding of the SNP differences –it's a relative measure!

	kp1	kp2	kp3	kp4	kp5	kp6	kp7	kp8		kp1_1.trim	kp2_1.trim	kp3_1.trim	kp4_1.trim	kp5_1.trim	
kp1_1.trim	0	3	3	6	3	3	2383	1127		.sorted/1-9	.sorted/1-9	.sorted/1-9	.sorted/1-9	.sorted/1-9	kp6_1.trim.sorted/1-9
kp2_1.trim	3	0	4	7	4	4	2384	1128	kp1_1.trim	0	4	3	6	3	3
kp3_1.trim	3	4	0	3	0	0	2382	1126	kp2_1.trim	4	0	5	8	5	5
kp4_1.trim	6	7	3	0	3	3	2385	1129	kp3_1.trim	3	5	0	3	0	0
kp5_1.trim	3	4	0	3	0	0	2382	1126	kp4_1.trim	6	8	3	0	3	3
kp6_1.trim	3	4	0	3	0	0	2382	1126	kp5_1.trim	3	5	0	3	0	0
kp7_1.trim	2383	2384	2382	2385	2382	2382	0	3135	kp6_1.trim	3	5	0	3	0	0
kp8_1.trim	1127	1128	1126	1129	1126	1126	3135	0	min: 0 max: 8						
min: 0 max: 3135															

Percentage of reference genome covered by all isolates: 90.6273963945683

5274461 positions was found in all analyzed genomes.

Size of reference genome: 5819941

Below is listed the number of positions that are shared and trusted between

File	Valid positions	Pct. of reference
kp7_1.trim.ignored_snps	5465853	93.9159520689299
kp6_1.trim.ignored_snps	5814419	99.9051193130652
kp3_1.trim.ignored_snps	5816912	99.9479547988545
kp5_1.trim.ignored_snps	5813062	99.8818029254936
kp8_1.trim.ignored_snps	5576639	95.819510885076
kp4_1.trim.ignored_snps	5803576	99.7188115824542
kp1_1.trim.ignored_snps	5818756	99.9796389688486
kp2_1.trim.ignored_snps	5815425	99.9224047116629

Percentage of reference genome covered by all isolates: 98.876345997322

5754545 positions was found in all analyzed genomes.

Size of reference genome: 5819941

Below is listed the number of positions that are shared and trusted between each isolate and the reference genome.

File	Valid positions	Pct. of reference
kp6_1.trim.ignored_snps	5814419	99.9051193130652
kp5_1.trim.ignored_snps	5813062	99.8818029254936
kp3_1.trim.ignored_snps	5816912	99.9479547988545
kp1_1.trim.ignored_snps	5818756	99.9796389688486
kp2_1.trim.ignored_snps	5815425	99.9224047116629
kp4_1.trim.ignored_snps	5803576	99.7188115824542

Survey 2 results – clustering

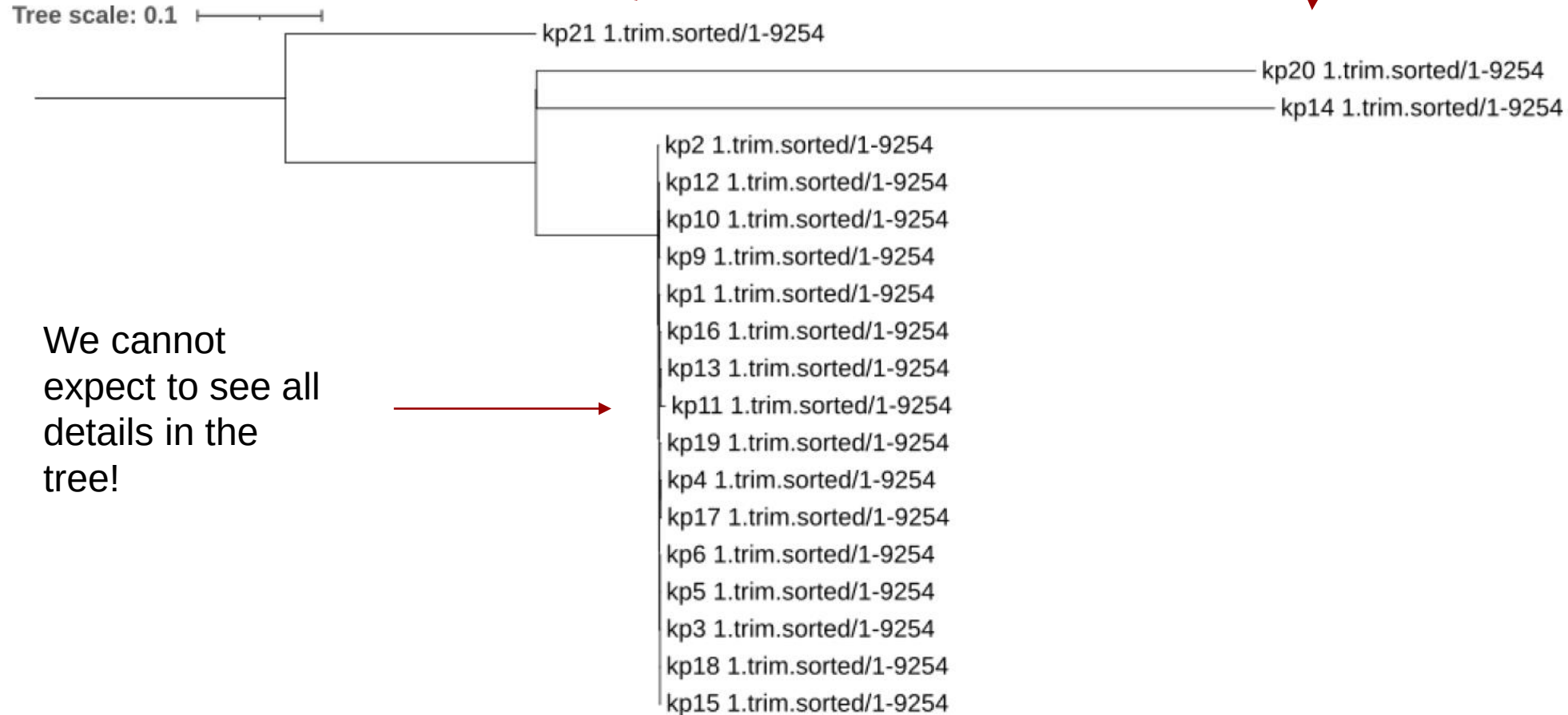
- In inject 3 we get 3 datasets
 - Screening of patients and environment in oncology and neonatal wards
 - Determine if they are part of the outbreak

Inject	# of samples to exclude (% of responses)	Samples to exclude
3.1	3 (30 %) 4 (70 %)	kp14, kp20, kp21 Kp11 , kp14, p20, kp21
3.2	All 5 (45 %) 4 (55 %)	Kp22, kp23, kp24, kp25, kp26 Kp22, kp23, kp24, kp25
3.3	1 (100 %)	Kp37

Inject 3.1

To be excluded

Custer analysis results



We cannot expect to see all details in the tree!

Inject 3.1

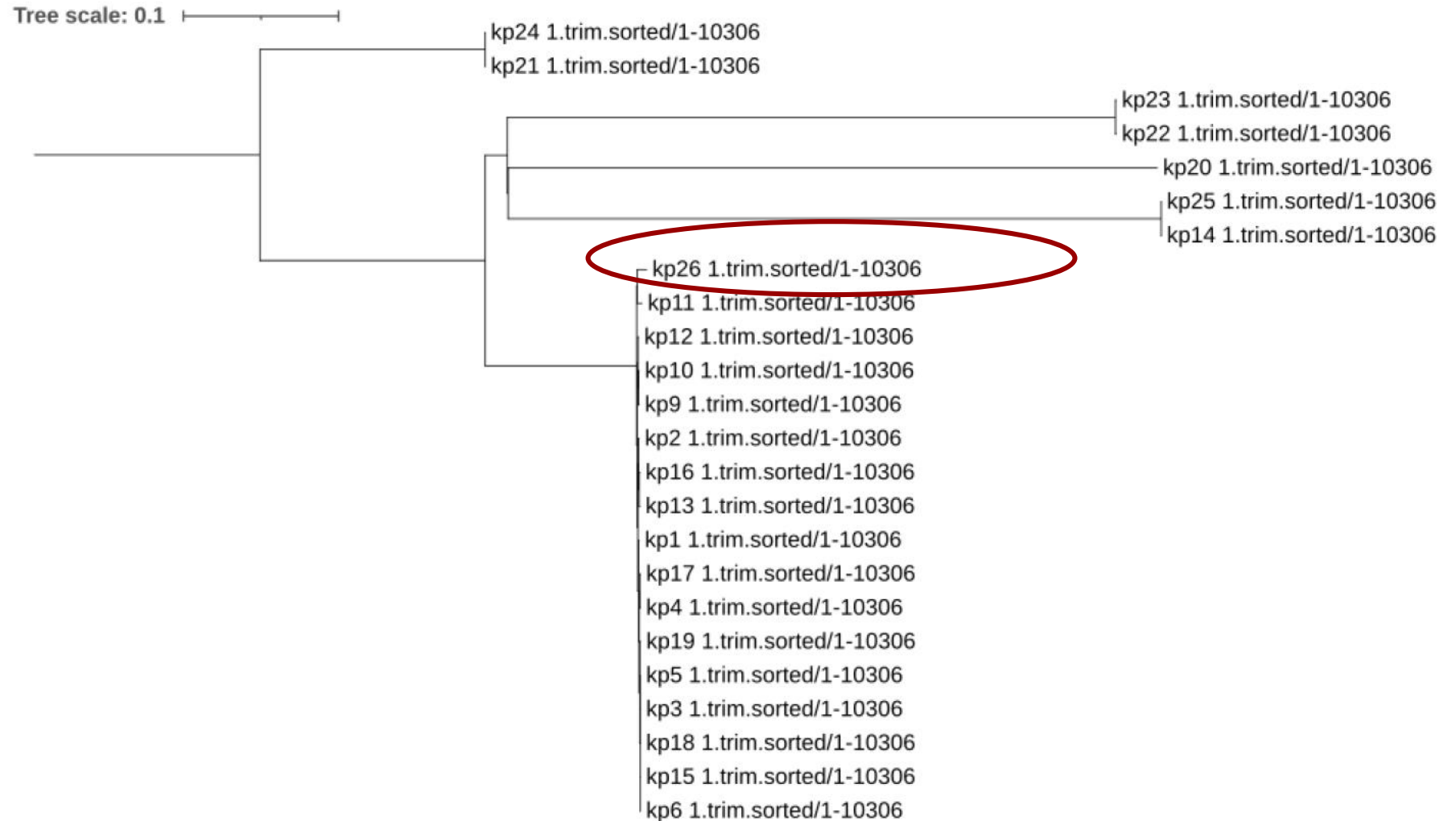
	kp10_1.trim.sorted/1-9254	kp11_1.trim.sorted/1-9254	kp1_1.trim.sorted/1-9254	kp12_1.trim.sorted/1-9254	kp13_1.trim.sorted/1-9254	kp14_1.trim.sorted/1-9254
kp10_1.trim.sorted/1-9254	0	43	1	1	4	4363
kp11_1.trim.sorted/1-9254	43	0	42	44	45	4391
kp1_1.trim.sorted/1-9254	1	42	0	2	3	4362
kp12_1.trim.sorted/1-9254	1	44	2	0	5	4364
kp13_1.trim.sorted/1-9254	4	45	3	5	0	4365
kp14_1.trim.sorted/1-9254	4363	4391	4362	4364	4365	0
kp15_1.trim.sorted/1-9254	2	43	1	3	4	4363
kp16_1.trim.sorted/1-9254	4	45	3	5	6	4365
kp17_1.trim.sorted/1-9254	3	44	2	4	5	4364
kp18_1.trim.sorted/1-9254	2	43	1	3	4	4363
kp19_1.trim.sorted/1-9254	8	49	7	9	10	4369
kp20_1.trim.sorted/1-9254	4349	4378	4348	4350	4351	5744
kp21_1.trim.sorted/1-9254	3541	3570	3540	3542	3543	5516
kp2_1.trim.sorted/1-9254	3	44	2	4	5	4364
kp3_1.trim.sorted/1-9254	2	43	1	3	4	4363
kp4_1.trim.sorted/1-9254	5	46	4	6	7	4366
kp5_1.trim.sorted/1-9254	2	43	1	3	4	4363
kp6_1.trim.sorted/1-9254	2	43	1	3	4	4363
kp9_1.trim.sorted/1-9254	1	44	2	2	5	4364
min: 0 max: 5744						

Is isolate with 43 SNPs still in the cluster?

	kp21_1.trim.sorted/1-9254	kp2_1.trim.sorted/1-9254	kp3_1.trim.sorted/1-9254	kp4_1.trim.sorted/1-9254	kp5_1.trim.sorted/1-9254	kp6_1.trim.sorted/1-9254	kp9_1.trim.sorted/1-9254
49	3541	3	2	5	2	2	1
78	3570	44	43	46	43	43	44
48	3540	2	1	4	1	1	2
50	3542	4	3	6	3	3	2
51	3543	5	4	7	4	4	5
44	5516	4364	4363	4366	4363	4363	4364
49	3541	3	0	3	0	0	3
51	3543	5	4	7	4	4	5
50	3542	4	1	4	1	1	4
49	3541	3	0	3	0	0	3
55	3547	9	6	9	6	6	9
0	5474	4350	4349	4352	4349	4349	4350
74	0	3542	3541	3544	3541	3541	3542
50	3542	0	3	6	3	3	4
49	3541	3	0	3	0	0	3
52	3544	6	3	0	3	3	6
49	3541	3	0	3	0	0	3
49	3541	3	0	3	0	0	3
50	3542	4	3	6	3	3	0

Inject 3.2

- SNP matrix shows 71 SNPs between kp26 and kp1
- There are >10,000 SNPs in total in this analysis
- High variation leads to a lower coverage (percentage of bases) being included in the analysis



Survey 2 results - characterisation

Inject		Samples to exclude
3.1	MLST of non-cluster isolates (Outbreak cluster ST147)	Kp11: ST4843 Kp14: ST15 Kp20: ST17 Kp21: ST395
3.2	NDM gene in neonatal isolates	Kp22: NDM-1 Kp23: NDM-1 Kp24: negative Kp25: NDM-1 Kp26: NDM-1
3.3	Carba-genes in the selected isolates (“low MIC”)	Kp26: NDM-1; OXA-48 Kp27: NDM-1 Kp35: none Kp36: none Kp37: NDM-1; NDM-5; OXA-232

MLST

- Based on 7 'housekeeping' genes
- Each gene variant is assigned a variant number (allele)
- The combination of allele numbers -> MLST number
- Just one nucleotide change in one gene -> new MLST number
- One allele variant can still be closely related

MLST-2.0 Server - Results

mlst Profile: *kpneumoniae*

Organism: *Klebsiella pneumoniae*

Sequence Type: 147 Input Files: *kp1_1.fa*

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

Sequence Type: 4843

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_273
tonB	100	100	414	414	0	tonB_38

extended output

Input Files: *kp11_1.fa*

MLST – one allele difference

- In this case, there is only one SNP between the rpoB alleles
- One allele variant can still be closely related!

```
# Length: 501
# Identity: 500/501 (99.8%)
# Similarity: 500/501 (99.8%)
# Gaps: 0/501 ( 0.0%)
# Score: 2496.0
#
#
#=====
```

```
1 ACCCACTACGGTCGCGTATGTCCGATCGAAACGCCTGAAGGTCCGAACAT 50
1 ACCCACTACGGTCGCGTATGTCCGATCGAAACGCCTGAAGGTCCGAACAT 50

51 CGGTCTGATTAACCTCCCTGTCCGTGTACGCGCAGACCAACGAATATGGCT 100
51 CGGTCTGATTAACCTCCCTGTCCGTGTACGCGCAGACCAACGAATATGGCT 100

101 TCCTTGAGACGCCGTATCGTAAAGTGACCGACGGTGTGGTTACTGACGAA 150
101 TCCTTGAGACGCCGTATCGTAAAGTGACCGACGGTGTGGTTACTGACGAA 150

151 ATTCACCTACCTGTCTGCTATCGAAGAAGGCAACTACGTTATCGCTCAGGC 200
151 ATTCACCTACCTGTCTGCTATCGAAGAAGGCAACTACGTTATCGCTCAGGC 200

201 GAACTCCAACCTGGATGAAAACGGCCACTTCGTAGAAGATCTGGTTACCT 250
201 GAACTCCAACCTGGATGAAAACGGCCACTTCGTAGAAGATCTGGTTACCT 250

251 GCCGTAGCAAAGGCGAATCCAGCTTGTTTCAGCCGCGACCAAGTTGACTAC 300
251 GCCGTAGCAAAGGCGAATCCAGCTTGTTTCAGCCGCGACCAAGTTGACTAC 300

301 ATGGACGTATCCACCCAGCAGGCGTATCCGTCGGTGCGTCCCTGATCCC 350
301 ATGGACGTATCCACCCAGCAGGCGTATCCGTCGGTGCGTCCCTGATCCC 350

351 GTTCCTGGAACACGATGACGCCAACCGTGCAATTGATGGGTGCGAACATGC 400
351 GTTCCTGGAACACGATGACGCCAACCGTGCAATTGATGGGTGCGAACATGC 400

401 AACGTCAGGCGGTTCCGACTCTGCGCGCTGATAAGCCGCTGGTTGGTACC 450
401 AACGTCAGGCGGTTCCGACTCTGCGCGCTGATAAGCCGCTGGTTGGTACC 450

451 GGTATGGAACGTGCTGTTGCCGTTGACTCCGGTGTTACTGCCGTGGCTAA 500
451 GGTATGGAACGTGCTGTTGCCGTTGACTCCGGTGTTACTGCCGTGGCTAA 500

501 A 501
|
501 A 501
```

What to include/exclude

Would you consider removing any isolates from the phylogeny provided by the exercise coordinators for Inject 3.3? If yes, please mention the isolates you would like to remove and why:

Resistance genotype

- 4 isolates (kp30, kp34, kp35, kp37) should be removed from the phylogeny, because these isolates do not have the same carbapenem resistance genes as cluster isolates (blaNDM-1)
- We remove Kp26 - it has OXA-48 in addition of NDM-1
- We would remove kp35 and 36 - no NDM1 or other carbapenem resistance genes

Subtyping/ MLST

- kp37 (different ST)

Sequence quality

- Kp28 (high contig count, low N50, but good enough for resistance and outbreak detection - no significant discrepancies when comparing to the rest of cluster but fragmented virulence genes)

of SNPs (and ST)

- kp11, kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp26, kp37 - reason: >18 SNPs / difference in ST
- kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp37 - > 1000 SNP difference from cluster

What to include/exclude?

Would you consider removing any isolates from the phylogeny provided by the exercise coordinators for Inject 3.3? If yes, please mention the isolates you would like to remove and why:

Resistance genotype

- 4 isolates (kp30, kp34, kp35, kp37) should be removed from the phylogeny, because these isolates do not have the same carbapenem resistance genes as cluster isolates (blaNDM-1)
- We remove Kp26 - it has OXA-48 in addition of NDM-1
- We would remove kp35 and 36 - no NDM1 or other carbapenem resistance genes

Subtyping/ MLST

- kp37 (different ST)

Sequence quality

- Kp28 (high contig count, low N50, but good enough for resistance and outbreak detection - no significant discrepancies when comparing to the rest of cluster but fragmented virulence genes)

of SNPs (and ST)

- kp11, kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp26, kp37 - reason: >18 SNPs / difference in ST
- kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp37 - > 1000 SNP difference from cluster

What to include/exclude - continued

of SNPs
(and ST)

- kp11, kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp26, kp37 - reason: >18 SNPs / difference in ST
- kp14, kp20, kp21, kp22, kp23, kp24, kp25, kp37 - > 1000 SNP difference from cluster
- To analyse more precisely the outbreak, I would remove all non-related isolates to go deeper into ST-147 and ST-147 like isolates (ST4843) : Kp37, Kp24, Kp21, Kp23, Kp22, Kp20, Kp25, Kp14

Additional comments

- Small SNP differences between isolates with hypervirulent and classical pathotype
 - (for example 6 SNP difference between classical kp13 vs hypervirulent kp16 or 17)
 - Kp11 and Kp26 are in the same cluster as Kp1 in your tree, but we refer on our tree
 - cgMLST Seqsphere
 - Based on the PlasmidFinder results and SNP analysis, we believe that kp35 and kp36 are part of the outbreak, however lost NDM-1 through plasmid loss (maybe IncFIB(pQil))
- Further, it seems like other clusters could be identifies:
 - Neonatal cluster kp22 (pt21) and kp23 (pt22) (0 allelic differences on SS+)
 - kp14 and kp25 (pt13 and pt24) with same profiles, but admitted to different wards. No clear epidemiological link
 - kp21 and kp24 (pt20 and pt23) and kp37 from the environmental assesment

Inject 4

- Characterisation of the isolates kp38 and 39

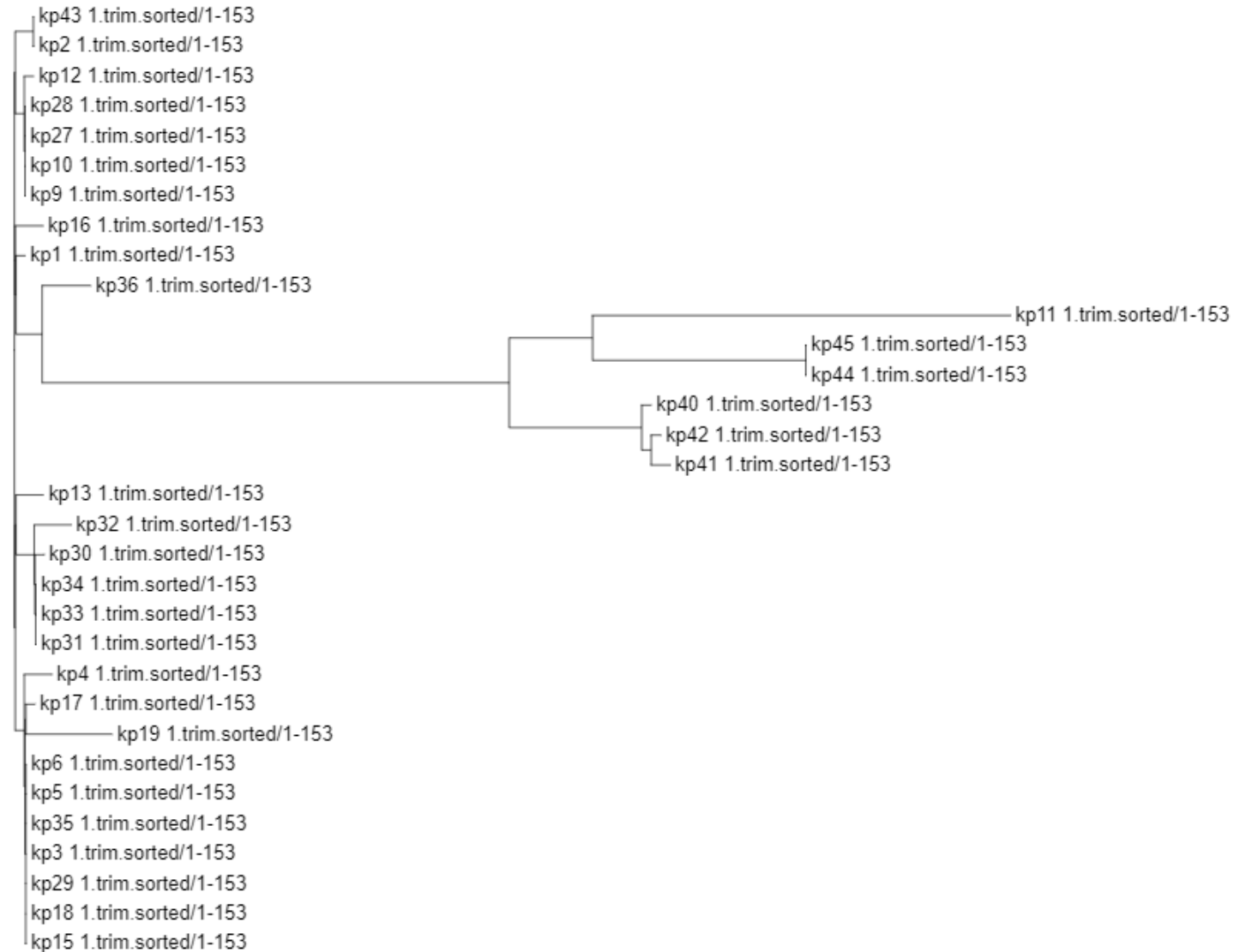
	MLST	NDM-1	Other NDM	Hypervirulent
kp38	ST147	NDM-1	-	-
kp39	Unknown	NDM-1	Yes VIM-producer	-

- Comments for kp38:
 - Kp38 might be in the cluster because it has less SNPs with kp1, kp2 and kp3, but the coverage is very low and we can not be sure.
- Comments for kp39:
 - the isolate MLST has few alleles matching the ST147, indicating possible relation, but still uncertain due to low sequence quality.
 - NDM might be missing due to sequence quality.
 - Were not able to determine hypervirulence.

Inject 5

- Comparison with ‘older isolates’
 - Kp43 is matching
 - Kp11 out of cluster
 - Now with >80 SNPs in a collection of more related isolates (max SNP difference is ca 150)
- Outbreak isolates SNP difference
 - 1-16 SNPs

Tree scale: 0.1



Sum-up

- You cannot (always) exclude isolates based on
 - MLST
 - Not all isolates of same ST cluster
 - One allele variant can still be closely related
 - Resistance genotype
 - Isolates can lose or gain plasmids with resistance genes between patients or samples
 - Or during sampling
 - SNP difference
 - How many SNPs allows inclusion in cluster?
 - An iterative process
 - Run analysis, exclude most distant, repeat

Data problems

- Inject 3.1
 - Fastq reads of kp14, kp25 and kp37 does not contain quality scores
 - Fastq files of kp15 are mixed with kp14
 - appear identical in CSIphylogeny, but program recognises f/r files as separate isolates
 - Fasta files are right – showing different isolates
- Inject 4
 - MinION data of questionable quality
 - Failing to be run correct in MinTyper

Evaluation survey

- Please fill it out
- **We would love your feedback!**
- https://ec.europa.eu/eusurvey/runner/EURGen_MD_WS_evaluation_may23
- Link will be on Slack too
- Thank you for taking the time to participate!

Thank you

Questions?