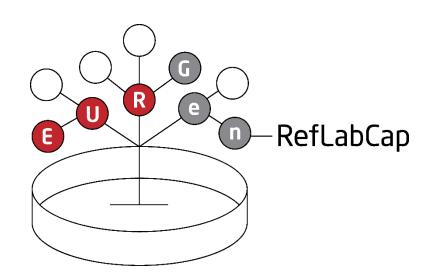




EURGen-RefLabCap

Results from the first external quality assessment (EQA) exercise

Ana Rita Rebelo (anrire@food.dtu.dk)







INTRODUCTION

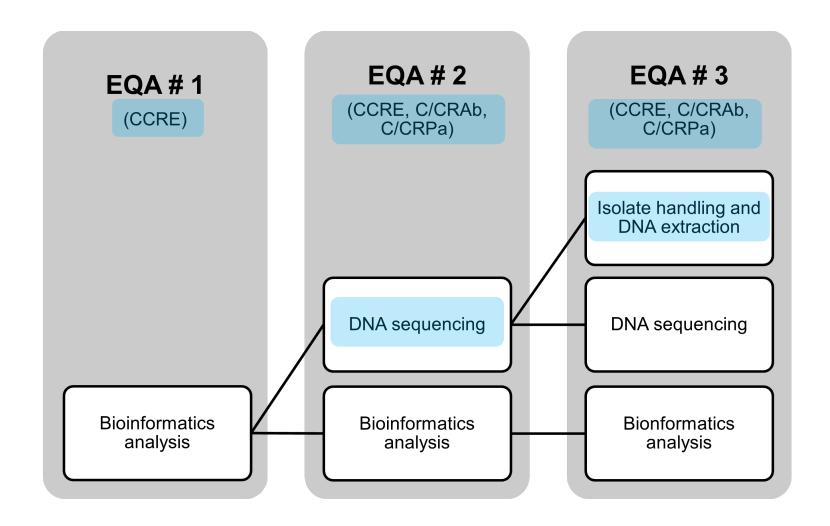
- OVERVIEW OF ALL EURGEN-REFLABCAP EQAS
- DESIGN OF THE FIRST EQA
- PREPARATION OF EXPECTED RESULTS
- SCORING SYSTEM

26. januar 2021 DTU Fødevareinstituttet



OVERVIEW OF ALL EURGEN-REFLABCAP EQAS







DESIGN OF THE FIRST EQA



Materials:

FASTA short-reads

FASTA long-reads

FASTQ short-reads

FASTQ long-reads

Strains:

EURGen-2022-01 E. coli

EURGen-2022-02 K. pneumoniae

EURGen-2022-03 K. pneumoniae

EURGen-2022-04 E. coli

Analyses:

- i) prediction of multi-locus sequence
- ii) detection of plasmid replicon types
- iii) detection of genes and chromosomal point mutations mediating AMR
- iv) in silico prediction of the AMR profiles





PREPARATION OF EXPECTED RESULTS



- Three external reference laboratories
 - DTU
 - SSI
 - Centre Hospitalier Universitaire de Caen Normandie
- Mainly tools from Center for Genomic Epidemiology (CGE)
- o Default thresholds (80% ID and 60% COV) or higher

Final set of expected results:

Categorical agreement

COV ≥ 90% (plasmids) / ≥ 60% (ARGs)





Analysis	Submitted result	Sco	re	
Prediction of MLST	Correct MLST	1		
Frediction of WLS1	Incorrect MLST	0		
Detection of plasmid	Genetic determinant correctly identified	1		
replicons, AMR genes and	Missing a genetic determinant	blar	ık	
chromosomal PMs	Reporting a genetic determinant that was not part of the expected results	0		
	Complete AMR profile correctly predicted	1		
In silico prediction of AMR profiles	Missing one or more antimicrobial in the complete AMR profile, or including antimicrobials that were not part of the expected profile	0		





- The in silico prediction of AMR profiles was evaluated as a <u>single answer:</u>
 - The antimicrobials included in the AMR profiles were not evaluated individually
 - All antimicrobials were <u>evaluated together</u> as one complete AMR profile:
 - > To score of "1" the participants had to correctly identify all antimicrobials that were part of the complete AMR profiles
 - Missing one antimicrobial, or including unexpected antimicrobials scored the answer as wrong with a value of "0"

o Example: EURGen-2022-01

Expected	AMC	AMP	CEP	СТА	CTZ	CTV	ERT	IMI	MER	PIT	Unexp.	Score
Lab 1	Х	х	х	х	х	х	Х	Х	Х	Х		1
Lab 2	х	-	х	х	х	х	х	х	Х	Х		0
Lab 3	х	Х	Х	Х	х	Х	Х	Х	Х	Х	(COL)	0







Maximum possible score for the participants

Material and analysis	EURGen- 2022-01	EURGen- 2022-02	EURGen- 2022-03	EURGen- 2022-04	Total
Prediction of MLST	1	1	1	1	4
Detection of plasmid replicons	2	7 ^a	3	5	17 ^a
Detection of AMR genes and chromosomal PMs	1	14	13	10	38
In silico prediction of AMR profiles	1	1	1	1	4
Total	5	23 ^a	18	17	63 ^a

^a If using data produced by long-read sequencing, the maximum possible score is n+1 due to the presence of one extra expected plasmid replicon in those data, when compared with short-read data







- Scoring system: all unexpected genetic determinants receive score "0"
- During validation of the submitted results: special situations were a score "0" was deemed inappropriate:
 - Reporting the genes bla_{SHV-64} or bla_{SHV-67} instead of the expected genes bla_{SHV-11} or bla_{SHV-12}
 - Reporting the genes bla_{SHV-1} or $bla_{SHV-106}$ instead of the expected gene bla_{SHV-28}
 - Reporting the gene *aac(3)-lle* instead of the expected gene *aac(3)-lla*
 - Reporting the gene aac(6')-lb instead of the expected gene aac(6')-llc
 - Reporting bla_{TEM} variants different from bla_{TEM-1}
 - Reporting the plasmid replicon IncHI1B(pNDM-MAR), when using short-read data
- The scoring was <u>manually adjusted</u> in these situations and the individual evaluation reports for each laboratory were updated: participants no longer receive a score "0" (which indicates an error), but instead will see a result of "blank" (indicating a discrepancy when compared with the expected results).







RESULTS AND DISCUSSION

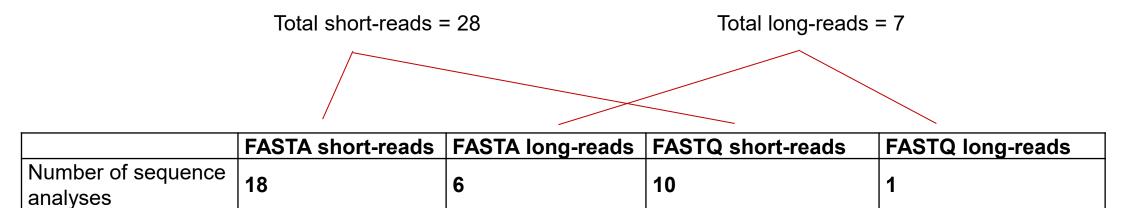
- ANALYSED MATERIALS
- OVERALL SCORES OF THE PARTICIPANTS
- PREDICTION OF MLST
- DETECTION OF PLASMID REPLICONS
- DETECTION OF GENES AND CHROMOSOMAL POINT MUTATIONS MEDIATING AMR
- IN SILICO PREDICTION OF AMR PROFILES

26. januar 2021 DTU Fødevareinstituttet



ANALYSED MATERIALS





Total 24 FASTA	Total 11 FASTQ
	·

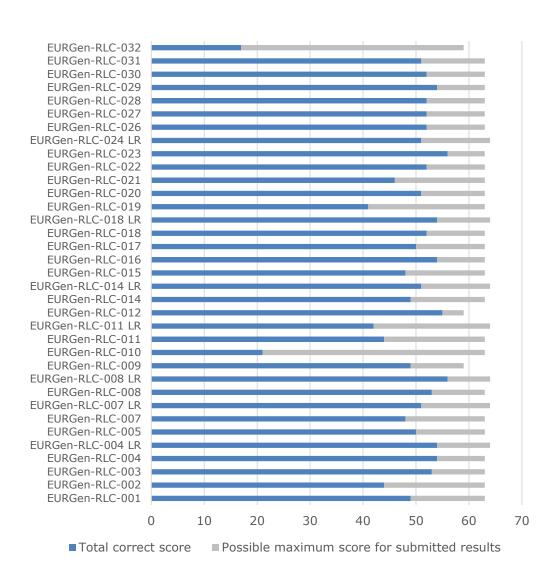
Total 35 allalyses

Most common bioinformatics approaches: CGE tools



OVERALL SCORES OF THE PARTICIPANTS





Averages of scores (%):

Prediction of MSLT: 96.4%

Detection of plasmid replicons: **74.0%**

Prediction of genetic AMR determinants: 83.3%

Prediction of AMR profiles: 14.8% *

Total: 77.6%

* Analysis of individual AMR profiles: 87.6%





PREDICTION OF MLST — EXPECTED RESULTS



Material	MLST	Alleles as	ssigned to	each loci	, from the	scheme <i>E</i>	. coli #1	
Waterial	IVILSI	adk	fumC	gyrB	icd	mdh	purA	recA
EURGen-2022-01	399	6	4	1	95	69	8	20
EURGen-2022-04	635	6	107	1	95	69	8	7

Material	MLST	Alleles a	ssigned to	each loci	, from the	scheme <i>K</i>	. pneumoi	niae
Wateriai	MILSI	gapA	infB	mdh	pgi	phoE	rpoB	tonB
EURGen-2022-02	147 ^a	3 ^a	4	6	1	7	4	38
EURGen-2022-03	307	4	1	2	52	1	1	7

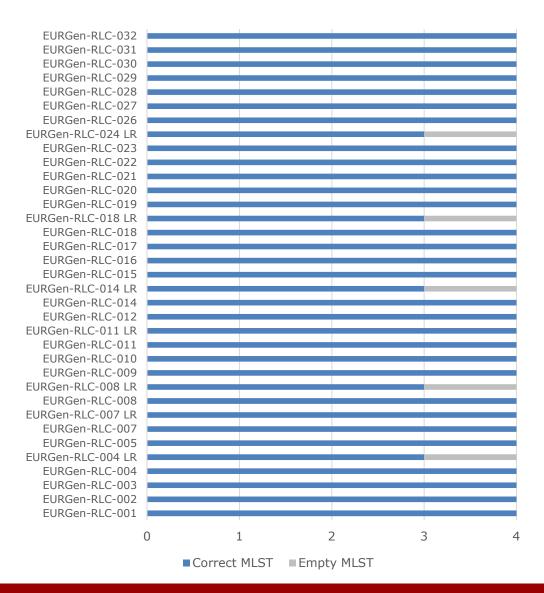


^a The assembled file produced with short-read sequencing yielded a perfect hit for the *gapA* locus, but the file produced with long-read sequencing did not generate a perfect hit.



PREDICTION OF MLST - SUBMITTED RESULTS





96.4% correct (n=135)

3.6% empty (n=5)

EURGen-2022-02 Long-reads



PREDICTION OF MLST - DISCUSSION



The five empty MLST results:

- Absence of a perfect hit for the gapA loci
- Also seen in the expected results



Discrepancy between short- and long-read data:

- there are differences between these technologies
- generally, short-read sequencing data have higher accuracy for analyses that are sensitive to SNPs

Recommendations:

- During routine laboratory work report the closest MLST
- Report the imperfect allele to the respective database





DETECTION OF PLASMID REPLICONS — EXPECTED RESULTS



Material	Plasmid replicons	Nr.
EURGen-2022-01	IncN2; IncFIB(K)	2
EURGen-2022-02	IncFII; IncL; IncFII(Yp); IncFIB(K)(pCAV1099-114); Col156; IncFIB(pB171); Col(pHAD28); IncHI1B(pNDM-MAR) ^a	7 / 8
EURGen-2022-03	IncFIB(pQiI); IncFIB(K); IncFII(K)	3
EURGen-2022-04	IncHI2A; IncHI2; IncFIB; IncFIB(K); Col(pHAD28)	5

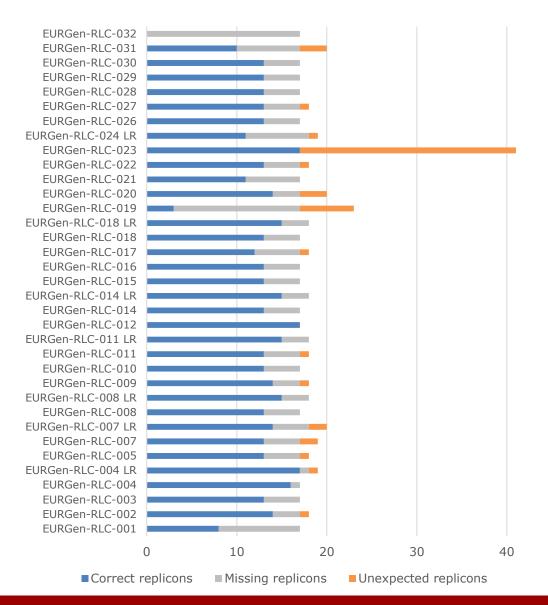


^a The replicon was only expected in data produced with long-read sequencing technologies. Participants that requested to analyse FASTQ files, including those using files produced through short-read sequencing, could also potentially detect the replicon. However, the replicon was not included nor scored as part of expected results for short-read data since its presence was not uniform between FASTQ and FASTA datasets.





DETECTION OF PLASMID REPLICONS — SUBMITTED RESULTS



48.6% fully correct (n=68)

47.9% missing replicons (n=67)

18.6% with unexpected replicons (n=26)

15% simultaneous (n=21)



DETECTION OF PLASMID REPLICONS — SUBMITTED RESULTS



Missing replicons in strain 01	Type of data
IncFIB(K) (n=3)	SR
IncN2 (n=2)	SR

Missing replicons in strain 02	Type of data
Col(pHAD28) (n=25)	SR
Col(pHAD28) (n=6)	LR
Col156 (n=22)	SR
Col156 (n=5)	LR
IncFIB(pB171) (n=22)	SR
IncFIB(pB171) (n=5)	LR
IncFIB(K)(pCAV1099-114) (n=7)	SR
IncFIB(K)(pCAV1099-114) (n=2)	LR
IncFII(Yp) (n=5)	SR
IncL (n=2)	SR
IncL (n=2)	LR
IncFII (n=1)	SR
IncFII (n=1)	LR

Missing replicons in strain 03	Type of data
IncFII(K) (n=1)	LR
IncFII(K) (n=5)	SR
IncFIB(K) (n=4)	SR
IncFIB(pQil) (n=2)	SR

Missing replicons in strain 04	Type of data
Col(pHAD28) (n=2)	LR
Col(pHAD28) (n=23)	SR
IncFIB(K) (n=3)	SR
IncFII (n=3)	SR
IncHI2 (n=2)	SR
IncHI2A (n=1)	SR



DETECTION OF PLASMID REPLICONS — DISCUSSION



The difference between expected results:

 IncHI1B(pNDM-MAR) in strain 02 was only part of the expected results for long-read data



Discrepancy between short- and long-read data:

- long-read sequencing is overall more adequate for detection of plasmids
- the assembly process might fail to properly capture sequences that were present in raw data

The missing plasmid replicons:

- Choice of different thresholds (?)
- o Other (?)



Different approaches according to purpose:

- Thresholds can be adjusted for different analyses
- Perhaps better to be less strict and manually evaluate results





DETECTION OF PLASMID REPLICONS — DISCUSSION



Reported unexpected results:

One participant with 24 unexpected replicons



- Too much noise can hide important information
- Results can / should be confirmed with other tools







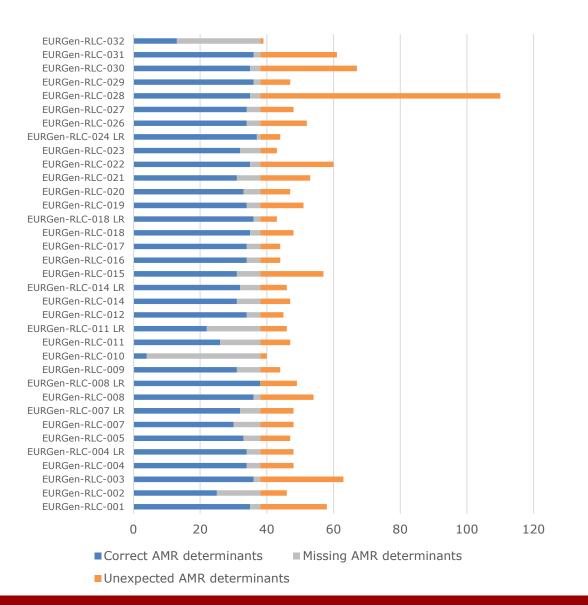
DETECTION OF GENES AND MUTATIONS MEDIATING AMR AND PREDICTION OF AMR PROFILES — EXPECTED RESULTS

Material	AMR genes and chromosomal PMs	Nr.	Associated prediction of AMR profiles	Nr.
EURGen-2022-01	bla _{NDM-1}	1	Amoxicillin-clavulanic acid, ampicillin, cefepime, cefotaxime, ceftazidime, ceftazidime-avibactam, ertapenem, imipenem, meropenem, piperacillin-tazobactam	10
EURGen-2022-02	bla _{CTX-M-15} , bla _{NDM-1} , bla _{OXA-1} , bla _{OXA-48} , bla _{SHV-11} , bla _{SHV-12} , bla _{TEM-1} , rmtC, aac(6')- lb-cr, aac(3)-lia, aac(6')-lb-cr, gyrA S83I, parC S80I, dfrA17, sul1	14	Amoxicillin-clavulanic acid, ampicillin, aztreonam, cefepime, cefotaxime, ceftazidime, ceftazidime-avibactam, ertapenem, imipenem, meropenem, piperacillin-tazobactam, Amikacin, gentamicin, tobramycin, Ciprofloxacin, Trimethoprim, Sulfamethoxazole	17
EURGen-2022-03	bla _{KPC-3} , bla _{CTX-M-15} , bla _{OXA-1} , bla _{OXA-9} , bla _{TEM-1} , bla _{SHV-28} , aac(3)-lla, aac(6')-lb-cr, qnrB1, aac(6')-lb-cr, gyrA S83I, parC S80I, dfrA14, sul2, mgrB::IS1	13	Amoxicillin-clavulanic acid, ampicillin, aztreonam, cefepime, cefotaxime, ceftazidime, ertapenem, imipenem, meropenem, piperacillin-tazobactam, Amikacin, gentamicin, tobramycin, Ciprofloxacin, Trimethoprim, Sulfamethoxazole, Colistin	16 (+1)
EURGen-2022-04	bla _{OXA-10} , bla _{OXA-436} , bla _{SHV-12} , bla _{TEM-1} , aac(6')-llc, gyrA S83I, qnrA1, dfrA19, sul1, sul2	10	Amoxicillin-clavulanic acid, ampicillin, aztreonam, cefepime, cefotaxime, ceftazidime, ertapenem, imipenem, meropenem, piperacillin-tazobactam, Gentamicin, tobramycin, Ciprofloxacin, Trimethoprim, Sulfamethoxazole	15









24.3% fully correct (n=34)

58.6% missing determinants (n=82)

72.1% with unexpected determinants (n=26)

55% simultaneous (n=77)



DETECTION OF GENES AND MUTATIONS — SUBMITTED RESULTS



Strain	Total expected	Correct	Missing	Unexpected
01	35	35 (100%)	0	7 († 20%)
02	490	405 (82.7%)	85 (17.3%)	176 († 36%)
03	455	369 (81.1%)	86 (18.9%)	224 († 49%)
04	350	299 (85.4%)	51 (14.6%)	45 († 13%)

Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-IIa (n=15)
02	Unexpected fosA (n=22)
02	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-IIe (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)





Strain	Examples of problems		
	Missing blaSHV-11 (n=25)		
	Missing gyrA or parC PM (n=18)		
	Missing aac(3)-IIa (n=15)		
00	Unexpected fosA (n=22)		
02	Unexpected oqxA/oqxB (n=20)		
	Unexpected aac(3)-IIe (n=10)		
	Unexpected PM in acrR (n=58)		
	Unexpected PM in ramR (n=11)		
03	Unexpected blaTEM (n=47)		
04	Unexpected mcr-9 (n=31)		

Several variants from the same gene present simultaneously:

Participants failed to report one of the variants



Results must be evaluated critically:

 The different variants are not in the same genomic location and are present simultaneously

Different approaches according to purpose:

Results can / should be confirmed with other tools







Strain	Examples of problems	
	Missing blaSHV-11 (n=25)	
	Missing gyrA or parC PM (n=18)	
	Missing aac(3)-IIa (n=15)	
02	Unexpected fosA (n=22)	
02	Unexpected oqxA/oqxB (n=20)	
	Unexpected aac(3)-IIe (n=10)	
	Unexpected PM in acrR (n=58)	
	Unexpected PM in ramR (n=11)	
03	Unexpected blaTEM (n=47)	
04	Unexpected mcr-9 (n=31)	

Reporting an incomplete PM profile:

Impact on expected phenotype (cumulative effect)



Results must be evaluated critically:

Results can / should be confirmed with other tools







Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-IIa (n=15)
02	Unexpected fosA (n=22)
	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-IIe (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)

Differences between bioinformatics databases:

 The same genetic sequence is annotated differently in ResFinder and AMRFinderPlus



- Important to understand the different databases supporting the tools
- Results can / should be confirmed with other tools







Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-Ila (n=15)
02	Unexpected fosA (n=22)
	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-Ile (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)

Reporting intrinsic genes:

Present in strain but with no impact in WT vs. non-WT phenotypes



- Too much noise can hide important information
- Insufficient knowledge regarding genetic mechanisms of AMR might lead to incorrect reporting of R-profiles





Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-Ila (n=15)
00	Unexpected fosA (n=22)
02	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-Ile (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)

Reporting PM (or genes) not relevant for the EQA

 Including mutations without proved association with Rprofiles



- Too much noise can hide important information
- Might point to insufficient knowledge of genetic AMR mechanisms







Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-IIa (n=15)
02	Unexpected fosA (n=22)
02	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-IIe (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)

Reported unexpected results:

 Nine participants reporting >2 variants of blaTEM from the same genomic location



- Too much noise can hide important information
- Results can / should be confirmed with other tools
- Different genetic variants can have different phenotypic implications







Strain	Examples of problems
	Missing blaSHV-11 (n=25)
	Missing gyrA or parC PM (n=18)
	Missing aac(3)-IIa (n=15)
02	Unexpected fosA (n=22)
02	Unexpected oqxA/oqxB (n=20)
	Unexpected aac(3)-IIe (n=10)
	Unexpected PM in acrR (n=58)
	Unexpected PM in ramR (n=11)
03	Unexpected blaTEM (n=47)
04	Unexpected mcr-9 (n=31)

Reporting different variants that occupy the same position in the genome:

Present in the strain, but does not confer phenotypic resistance



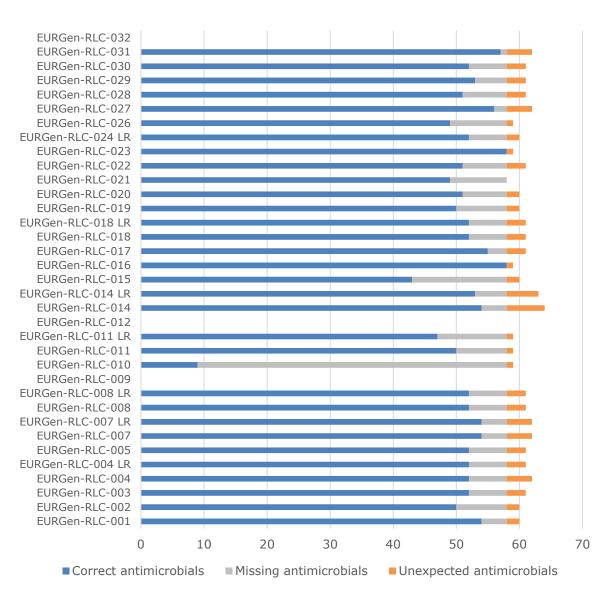
- Too much noise can hide important information
- Points to insufficient knowledge of genetic AMR mechanisms
- Might lead to incorrect reporting of R-profiles











14.8% fully correct (n=19)

63.3% missing antimicrobials (n=81)

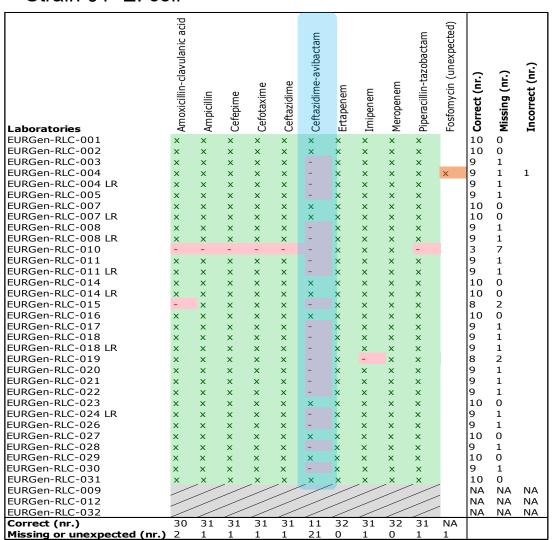
60.9% with unexpected antimicrobials (n=78)

39.1% simultaneous (n=50)





Strain 01 E. coli



Missing from most submitted results:

 Antimicrobial missing from blaNDM-1 R-profile in ResFinder database (but present for other genes)

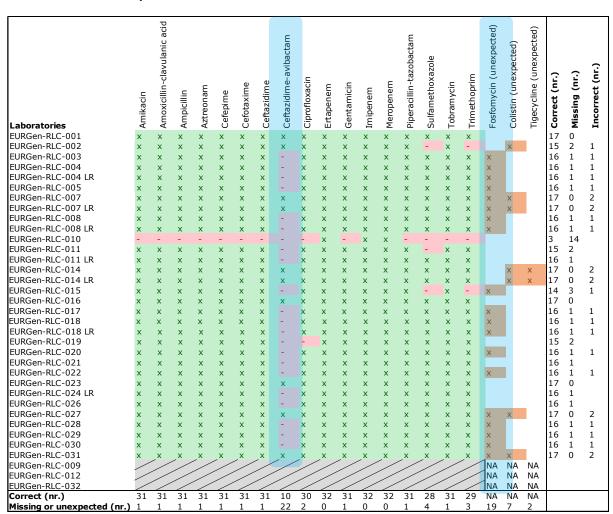


- Results can / should be confirmed with other tools
- Points to insufficient knowledge of genetic AMR mechanisms
- Important to understand the different databases supporting the tools





Strain 02 K. pneumoniae



Reporting unexpected results:

Due to do the reporting of intrinsic genes



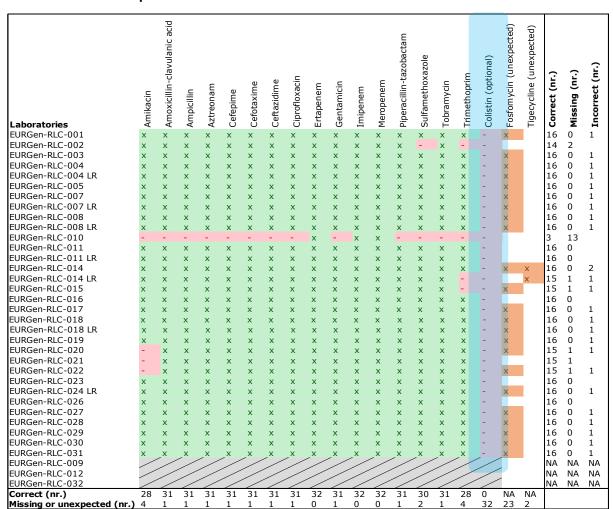
- Too much noise can hide important information
- Points to insufficient knowledge of genetic AMR mechanisms
- Might lead to incorrect reporting of R-profiles







Strain 03 K. pneumoniae



None of the participants reported colistin resistance

 Because none of the participants found the underlying genetic mechanism of resistance (chromosomal PM mgrB::IS1)



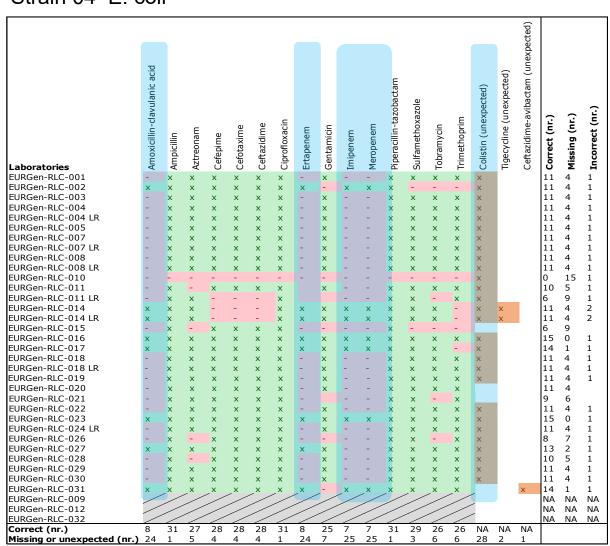
- Results can / should be confirmed with other tools
- Very Major Errors can have important clinical implications







Strain 04 E. coli



Very Major Errors of high significance:

- Almost all missed resistance toward carbapenems
- blaOXA-436 has an incomplete phenotype in ResFinder



- Results can / should be confirmed with other tools
- Very Major Errors can have important clinical implications
- Points to insufficient knowledge of genetic AMR mechanisms



OTHER GENERAL RECOMMENDATIONS



For discrepancies due to variations between the type of data and the chosen bioinformatics tools and databases:

- Laboratories planning to implement or in the process of implementing WGS-based analysis in their settings should aim at using harmonized protocols such as the one created during the EURGen-RefLabCap project;
- Laboratories currently using WGS could consider aligning their own protocols with other harmonized protocols;
- Laboratories should implement verification steps such as using multiple bioinformatics approaches to confirm the obtained results;
- Laboratories should communicate their suggestions, strange observations and potential problems to the curators of bioinformatics tools and databases;
- Laboratories should be aware of differences between short- and long-read sequencing data and select the most adequate approach depending on their aims.





OTHER GENERAL RECOMMENDATIONS



For discrepancies due to misinterpretation of the EQA protocol and/or insufficient knowledge about certain genetic elements:

- Laboratories should ensure at least basic, and ideally extensive, knowledge about the genetic mechanisms mediating AMR and other important genetic elements;
- Laboratories should be familiar with the mode of action of the bioinformatics tools they use, and the contents of the respective databases;
- Laboratories should analyse their data with the understanding that, currently, there is no "fit-forall" approach and some types of data and some suites of bioinformatics tools are more adequate for certain purposes than others;
- Laboratories should analyse their results critically and, when needed, perform confirmatory testing, to ensure that the information being reported is accurate and actionable.





OTHER GENERAL RECOMMENDATIONS



For everyone:

- Continue to participate in genomic EQAs
- Use well-defined QC parameters and respective thresholds
- Use benchmarking datasets to validate the bioinformatics approaches

- Repeat the analysis of the EQA materials and update the bioinformatics pipelines to ensure detection of all genetic determinants
- Contact the EQA organizers for troubleshooting





Upcoming task

Email consultation for the report

Document will be distributed in late June including updated individual evaluation reports

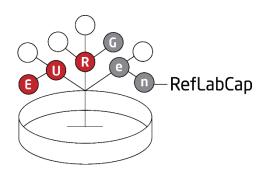
2 weeks of email consultation period

Changes will be implemented and the updated document will be shared by email





Thank you on behalf of the EURGen-RefLabCap team



EURGen-RefLabCap@food.dtu.dk