
DTU

Results from the second External Quality Assessment (EQA) exercise

EURGen-RefLabCap

Faisal Ahmad Khan (fakh@food.dtu.dk)

Lauge Holm Sørensen (lahoso@food.dtu.dk)

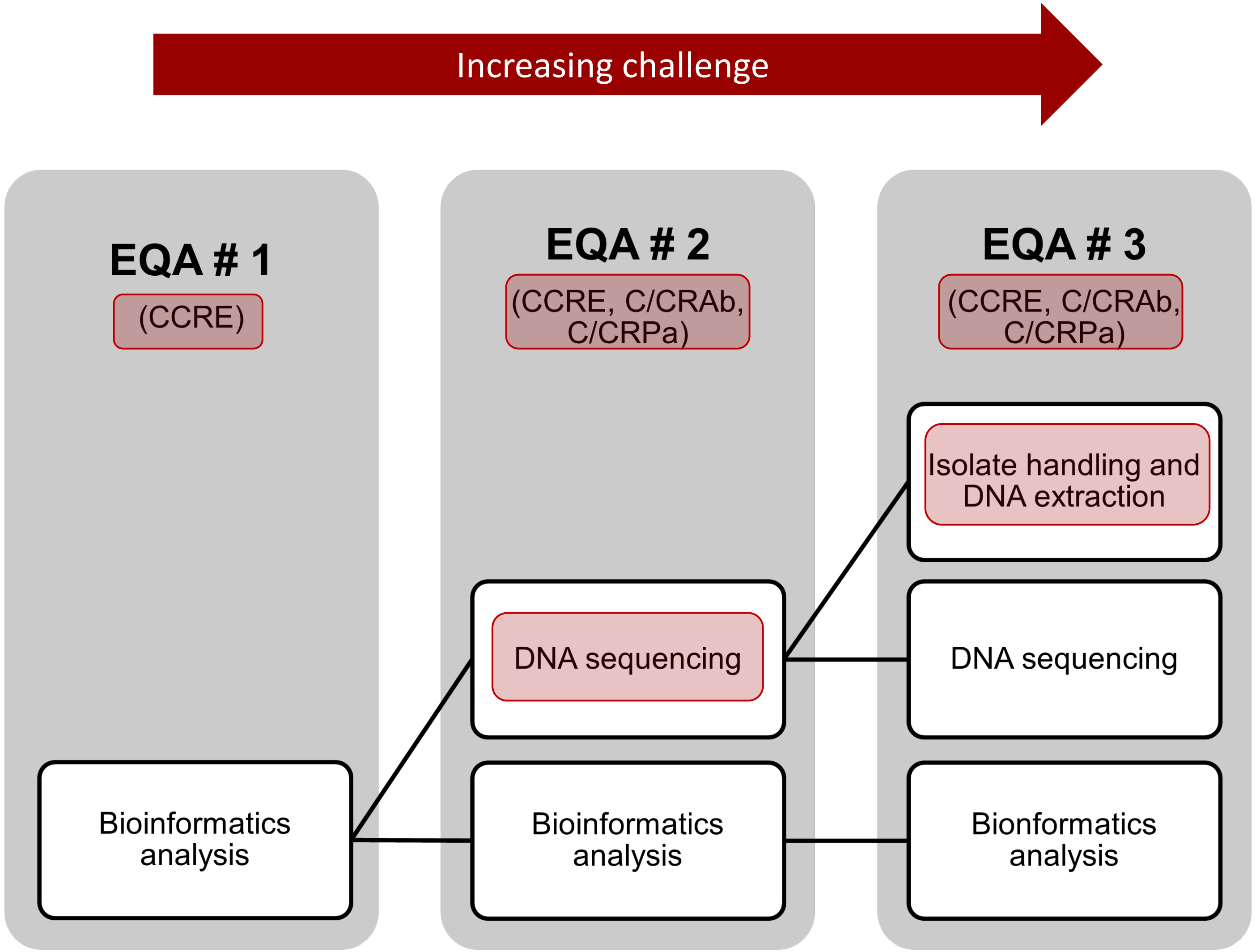
Funded by the European Union

INTRODUCTION

- OVERVIEW OF THE EURGen-RefLabCap EQAS
- DESIGN OF THE SECOND EQA
- PREPARATION OF EXPECTED RESULTS
- SCORING SYSTEM

Funded by the European Union

OVERVIEW OF ALL EURGen-RefLabCap EQAs



Increasing challenge

EQA # 1

(CCRE)

Bioinformatics analysis

EQA # 2

(CCRE, C/CRAb, C/CRPa)

DNA sequencing

Bioinformatics analysis

EQA # 3

(CCRE, C/CRAb, C/CRPa)

Isolate handling and DNA extraction

DNA sequencing

Bionformatics analysis

- Workstream 1 pathogens (WS1)
  - CRE/CCRE
- Workstream 2 pathogens (WS2)
  - C/CRPa
  - C/CRAb

Funded by the European Union

DESIGN OF SECOND EQA

Strains:
- EURGen-2023-01 (*Acinetobacter baumannii*)
- EURGen-2023-02 (*Escherichia coli*)
- EURGen-2023-03 (*Klebsiella pneumoniae*)
- EURGen-2023-04 (*Pseudomonas aeruginosa*)

Materials:
- Purified DNA
- (Raw and assembled reads from Illumina and nanopore sequencing technology)

Bioinformatics analyses included in EQA 2023:
1. Prediction of MLST
2. Prediction of plasmid replicon types
3. Detetcion of genes and chromosomal mutations mediating AMR
4. *In silico* prediction of AMR profiles

Additional Analyses

Quality control of sequences generated by participants

I. Short-read sequences
II. Long-read sequences

PREPARATION OF EXPECTED RESULTS

- Consensus results from TWO reference laboratories
  - Sequencing and bioinformatics analysis at DTU
  - Sequencing and bioinformatics analysis at SSI
- Bioinformatics tools used to prepare expected results
  - Mainly CGE tools
  - AMRFinder+
  - RGI (CARD database)
  - PathogenWatch
- Default thresholds (≥80% ID and ≥ 60% COV)

Final set of expected results

Categorical agreement

+

ID ≥ 90%

+

Coverage ≥ 90% (plasmids)/≥ 60% (ARGs)

Expected non mandatory results

- No consensus between reference labs
- Detetcion in only one type of dataset (lr or sr)
- Detection in only one tool

SCORING SYSTEM IN THE WEBTOOL

Analysis	**Submitted result**	**Score**
Prediction of MLST	Correct MLST	1
	Incorrect MLST	0
Detection of plasmid replicons, AMR genes and chromosomal mutations	Genetic determinant correctly identified	1
	Reporting a genetic determinant that was part of the expected results but not mandatory to report	blank
	Missing a genetic determinant	blank
	Reporting an unexpected genetic determinant	0
In-silico AMR profiles	AMR profile correctly reported for the antimicrobial	1
	Reporting an antimicrobial that was part of the expected results but not mandatory to report, or part of intrinsic resistance	blank
	Missing an antimicrobial	blank
	Reporting an AMR profile for an unexpected antimicrobial	0

SCORING SYSTEM

Maximum possible score of participants

Material and analysis	**EURGen-2023-01**	**EURGen-2023-02**	**EURGen-2023-03**	**EURGen-2023-04**	**Total**
Prediction of MLST	1	1	1	1	**4**
Detection of plasmid replicons	0	6	1	0	**7**
Detection of AMR genes and chromosomal PMs	6	17	4	6	**33**
In silico prediction of AMR profiles	10	18	5	10	**43**
Total	**17**	**41**	**12**	**17**	**87**

RESULTS AND DISCUSSION

- MATERIAL ANALYSED BY PARTICIPANTS
- OVERALL SCORES OF THE PARTICIPANTS
- PREDICTION OF MLST
- DETETCION OF PLASMID REPLIOCN GENES
- DETETCION OF AMR GENES AND MUTATIONS
- *IN SILICO* PREDICTION OF AMR PROFILES

MATERIAL ANALYSED BY PARTICIPANTS

Total submission: 30

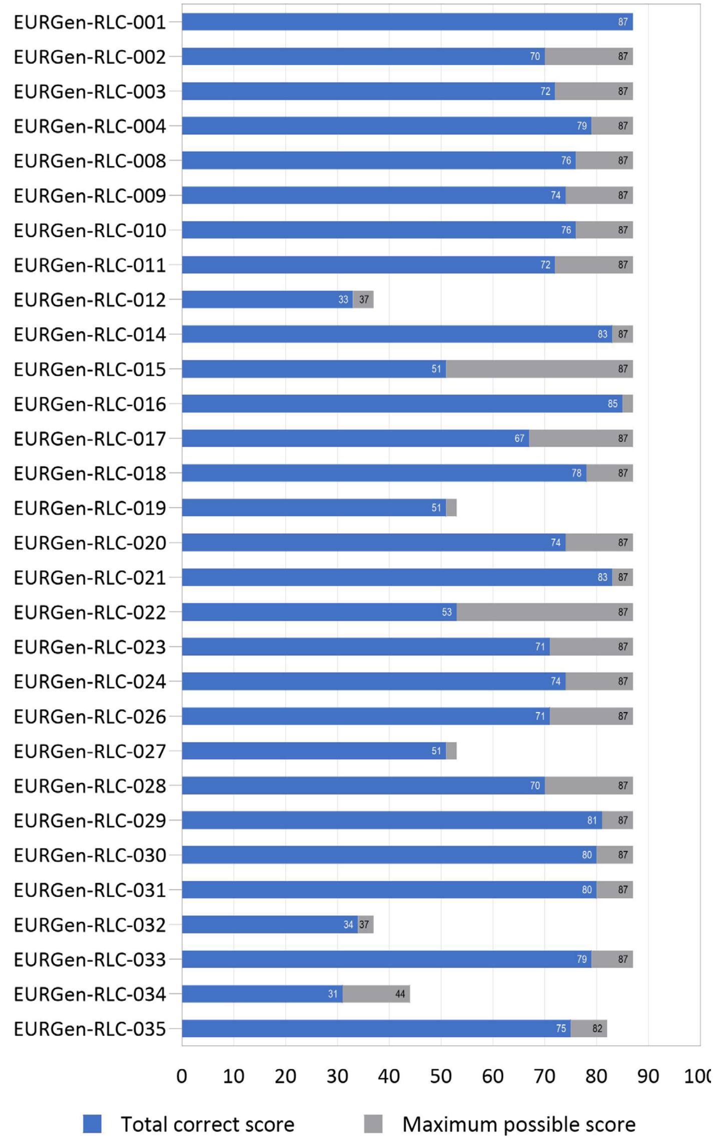
- WS1: 30 laboratories submitted results
- WS2: 28 laboratories submitted results

	EURGen-2023-01 ***(A. baumannii)***	**EURGen-2023-02** **(*E. coli*)**	**EURGen-2023-03** **(*K. pneumoniae*)**	**EURGen-2023-04** **(*P. aeruginosa*)**
Number of Laboratories	28	30	30	28

Test material used:

- DNA: 22 laboratories
- FASTQ: 5 laboratories
- FASTA: 2 laboratories
- DNA/FASTQ: 1 laboratory

OVERALL SCORES OF THE PARTICIPANTS



Averages of scores (%)

Prediction of MLST: **90%**
Detection of plasmid replicons: **81.9%**
Prediction of genetic AMR determinants: **86%**
Prediction of AMR profiles: **86.1%**

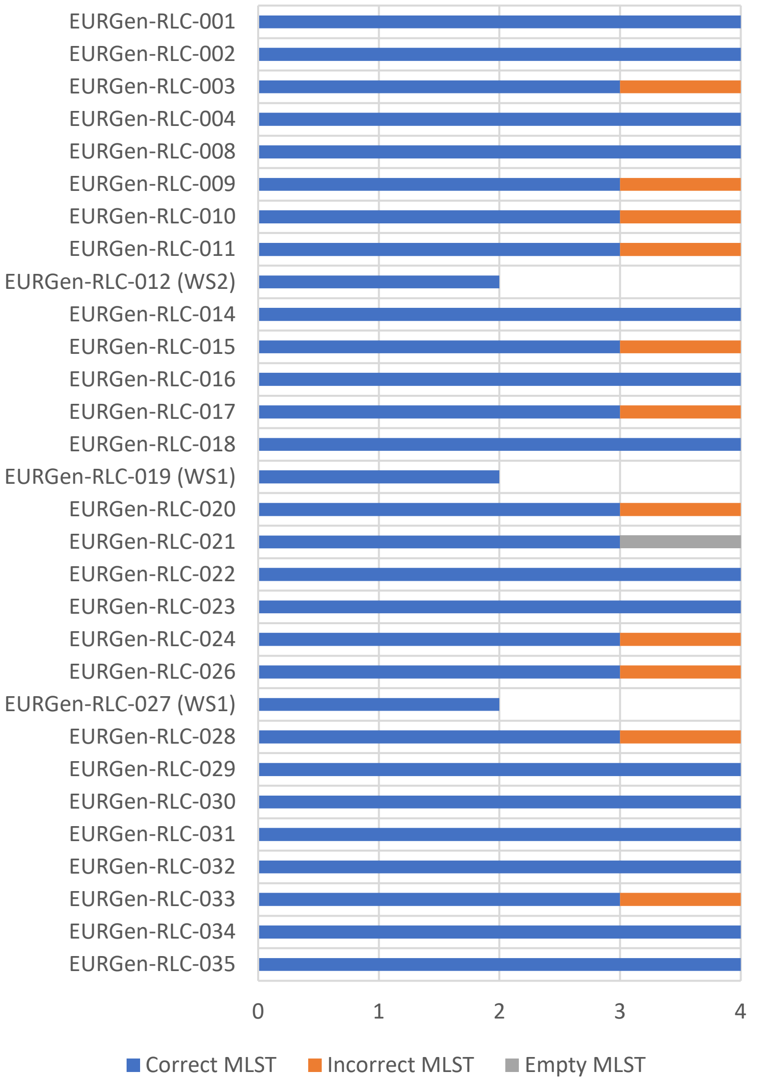
Total: 85.5%

PREDICTION OF MLST – EXPECTED RESULTS

Material	**MLST**	**Alleles assigned to each loci**						
EURGen-2023-01 (*A. baumannii*)	**136[a]**	*cpn60*	*gdhB*	*gltA*	*gpi*	*gyrB*	*recA*	*rpoD*
		2	3[a]	1	16	3	2	3
EURGen-2023-02 (*E. coli*)	**410**	*adk*	*fumC*	*gyrB*	*icd*	*mdh*	*purA*	*recA*
		6	4	12	1	20	18	7
EURGen-2023-03 (*K. pneumoniae*)	**4568**	*gapA*	*infB*	*mdh*	*pgi*	*phoE*	*rpoB*	*tonB*
		2	1	2	1	247	1	46
EURGen-2023-04 (*P. aeruginosa*)	**233**	*acsA*	*aroE*	*guaA*	*mutL*	*nuoD*	*ppsA*	*trpE*
		16	5	30	11	4	31	41

[a] The Oxford scheme reports two sequence types for EUGen-2023-01 due to presence of multicopy *gdhB* allele i.e., *gdhB*_189 and *gdhB*_3. The allele with lowest number was selected for the expected results.

PREDICTION OF MLST - SUBMITTED RESULTS



WS1: Only submitted results for workstream 1 pathogens; WS2: Only submitted results for workstream 2 pathogens.

- Average score : 90%

- All MLST were correct for EURGen-2023-02, EURGen-2023-03, and EURGen-2023-04

- 11 Incorrect MLST
  - All for EURGen-2023-01

PREDICTION OF MLST - DISCUSSION

11 Incorrect MLST results for *A. baumannii* (EURGen-2023-01)

- Due to MLST scheme used
  - Participants used Pasteur scheme (ST 02)
  - Oxford scheme used for expected results (ST 136)

- Due to presence of multicopy *gdhB* allele i.e., *gdhB_3* and *gdhB_189* (Oxford scheme)
  - Resulting in two sequence types i.e., ST 136 and ST 1851
  - Also detected in expected results

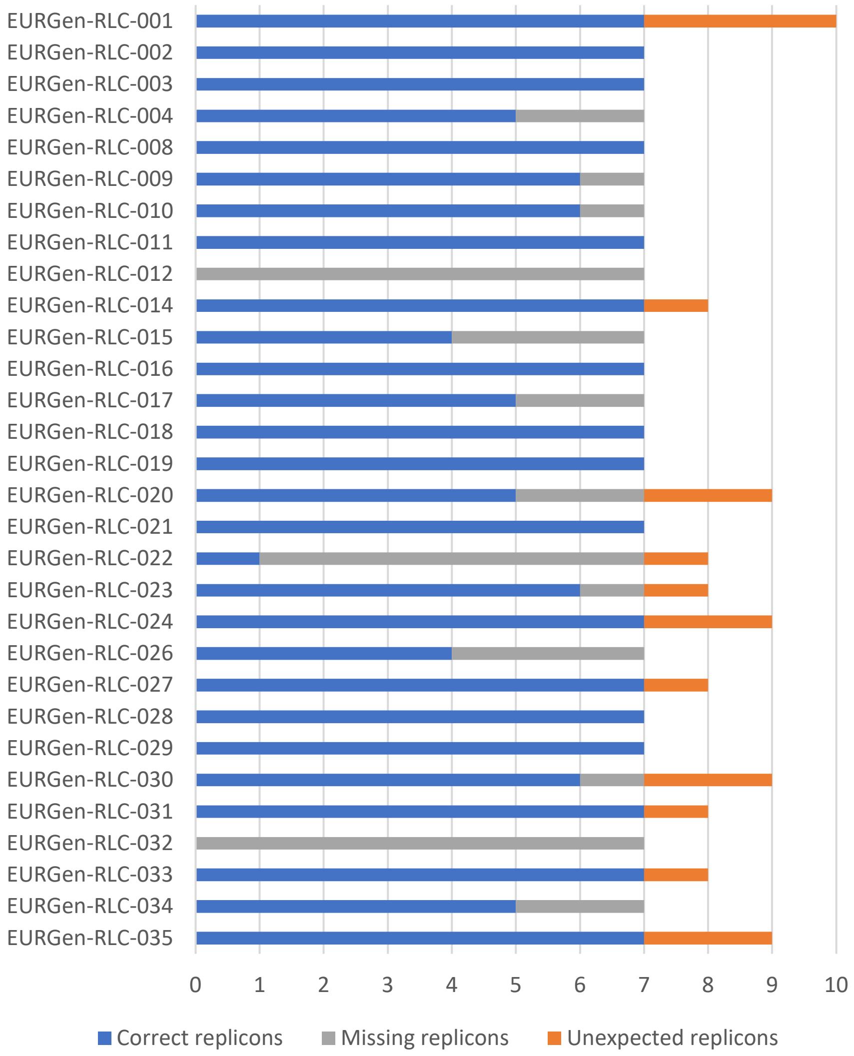
- For the self-evaluation, it should be considered that these discrepancies do not represent a flaw in the bioinformatics analysis performed
- It is important to understand that the bioinformatics capacity and knowledge required for using either MLST scheme is the same

DETECTION OF PLASMID REPLICONS – EXPECTED RESULTS

Material	Plasmid replicons	Nr.
EURGen-2023-01 (*A. baumannii*)	No plasmid replicon detected	0
EURGen-2023-02 (*E. coli*)	**Expected:** Col(BS512), ColKP3, IncFIA, IncFIB(AP001918), IncFII(pAMA1167-NDM-5), IncX3	6
	Expected but non-mandatory: IncQ1, Col(pHAD28)	2
EURGen-2023-03 (*K. pneumoniae*)	**Expected**: repB(R1701)	1
	Expected but non-mandatory: Col(pHAD28)	1
EURGen-2023-04 (*P. aeruginosa*)	No plasmid replicon detected	0

Funded by the European Union

DETECTION OF PLASMID REPLICONS - SUBMITTED RESULTS



Total submissions: 60

- 50% submissions were fully correct (n=30)
- 33.3% had missing replicons (n=20)
- 23.3% had unexpected replicons (n=14)

6.7% Simultaneous (n=4)

Average score: 81.9%

- 17 participants achieved 100% of max. possible score

DETECTION OF PLASMID REPLICONS - SUBMITTED RESULTS

Strain	Missing expected replicons	Unexpected replicons
EURGen-2023-02	Col(pHAD28)* (n=24)	ColpVC (n=3)
	IncQ1* (n=17)	IncFII(pRSB107) (n=2)
	ColKP3 (n=8)	IncFIA(HI1) (n=1)
	IncFIB(AP001918)(n=7)	
	IncFII(pAMA1167-NDM-5) (n=5)	
	IncX3 (n=3)	
	Col(BS512) (n=3)	
EURGen-2023-03	Col(pHAD28)* (n=24)	Col(MG828) (n=8)
	repB(R1701) (n=8)	Col440I (n=3)

* Expected but non-mandatory

DETECTION OF PLASMID REPLICONS - DISCUSSION

- The non-mandatory replicons were missing in most results
  - Col(pHAD28) (n=43) and IncQ1 (n=17)
  - IncQ1 was only detected in long-reads sequencing 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 (?)

Different approaches according to purpose:

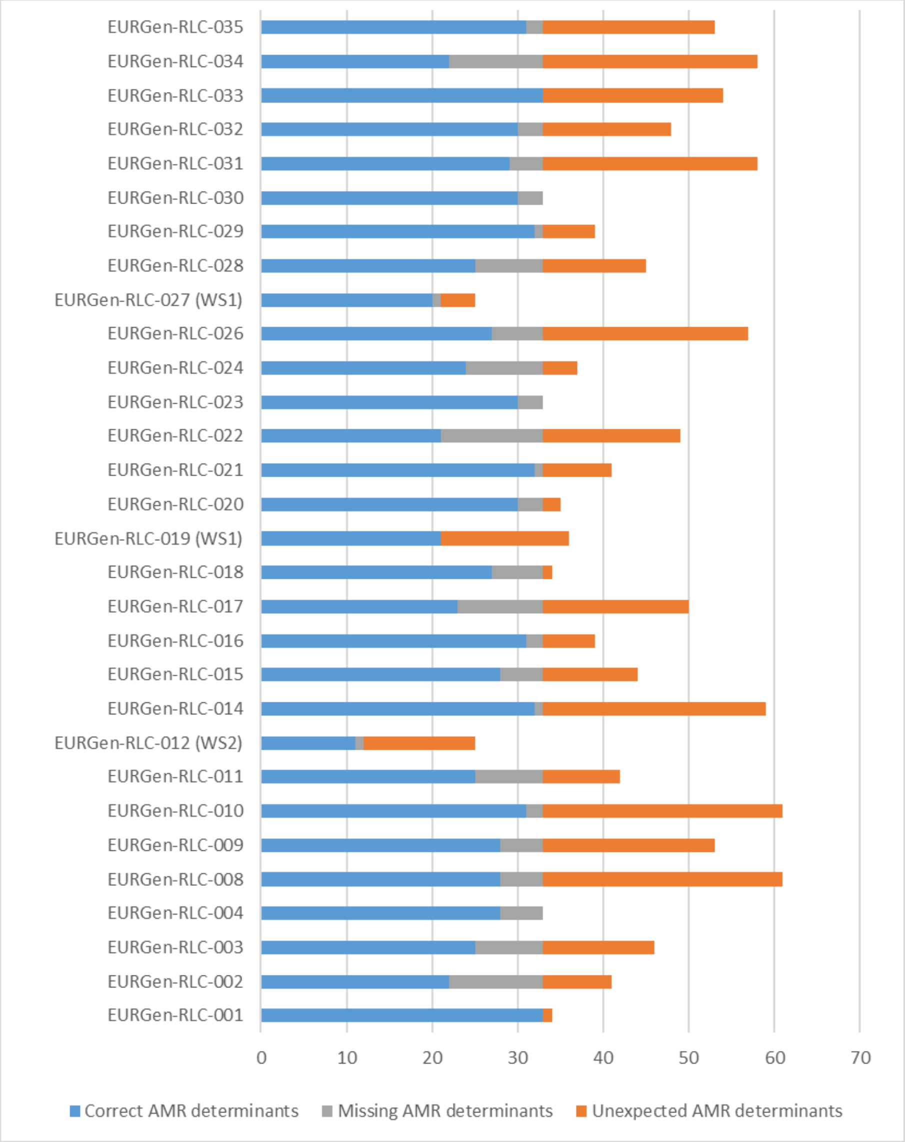
- Thresholds can be adjusted for different analyses
- Perhaps better to be less strict and manually evaluate results

DETECTION OF GENES AND MUTATIONS MEDIATING AMR – EXPECTED RESULTS

Material	**AMR genes and chromosomal mutations**
EURGen-2023-01 (*A. baumannii*)	**Expected**: *aph(3')-VI*[a], *armA, bla*$_{\text{NDM-1,}}$ *bla*$_{\text{OXA-23,}}$ *gyrA* S81L, *parC* S84L
	Expected but non-mandatory: *ftsI* A515V, *parC* V104I, *parC* D105E
EURGen-2023-02 (*E. coli*)	**Expected:** *aac(3)-IId*, *aac(6')-Ib-cr*[b], *bla*$_{\text{CMY-2}}$[c], *bla*$_{\text{OXA-181}}$, *bla*$_{\text{NDM-5}}$, *bla*$_{\text{TEM-1}}$[d], *bla*$_{\text{CTX-M-15}}$, *qnrS1*, *sul1*, *sul2*, *dfrA12*, *dfrA17, glpT* E448K, *gyrA* D87N, *gyrA* S83L, *parE* S458A, *parC* S80I
	Expected but non-mandatory: *bla*$_{\text{OXA-1,}}$ *ftsI* N337NYRIN, *pmrB* Y358N
EURGen-2023-03 (*K. pneumoniae*)	**Expected:** *bla*$_{\text{SHV-1}}$[e], *bla*$_{\text{TEM-1}}$[d], *bla*$_{\text{CTX-M-3}}$, *qnrS1*
	Expected but non-mandatory: *mgrB*::IS*1*
EURGen-2023-04 (*P. aeruginosa*)	**Expected:** aac(6')-Il, aac(3)-Id, *bla*$_{\text{VIM-2}}$, *bla*$_{\text{OXA-4,}}$ *crpP, gyrA* T83I
	Expected but non-mandatory: *parC* S87L

[a] Either *aph(3')-VIa* or *aph(3')-VI*
[b] Either *aac(6')-Ib-cr5* or *aac(6')-Ib-cr*
[c] Either the *bla*$_{\text{CMY-2}}$ or *bla*$_{\text{CMY-59}}$
[d] Either *bla*$_{\text{TEM-1}}$, *bla*$_{\text{TEM-1A}}$, *bla*$_{\text{TEM-1B}}$, *bla*$_{\text{TEM-1C}}$ or *bla*$_{\text{TEM-1D}}$
[e] Either *bla*$_{\text{SHV-1}}$, *bla*$_{\text{SHV-185}}$ or *bla*$_{\text{SHV-187}}$

DETECTION OF GENES AND MUTATIONS – SUBMITTED RESULTS



Total submissions: 114

- 9.6% submissions were fully correct (n=11)
- 61.4% had missing AMR determinants (n=70)
- 74.6% had unexpected determinants (n=85)

45.6% Simultaneous (n=52)

Average score: 86%

- Three participants achieved 100% of their max. possible score

DETECTION OF GENES AND MUTATIONS – SUBMITTED RESULTS

Strain	Total expected	Correct	Missing	Unexpected
EURGen-2023-01	168	132 (78.6%)	36 (21.4%)	73 (↑ 43.6%)
EURGen-2023-02	493	436 (88.4%)	57 (11.6%)	76 (↑15.4%)
EURGen-2023-03	116	108 (93.1%)	8 (6.9%)	104 (↑89.7%)
EURGen-2023-04	168	134 (79.8%)	34 (20.2%)	125 (↑74.4%)

Funded by the European Union

DETECTION OF GENES AND MUTATIONS – DISCUSSION

Strain	Examples of problems
01	Missing *gyrA* S81L (n=16) or *parC* S84L (n=15)
	Unexpected $bla_{\text{OXA-66}}$ (n=15) or $bla_{\text{ADC-25}}$ (n=12)
	Unexpected *aph(3'')-Ib* (n=12), *aph(6)-Id* (n=12), *tet(B)* (n=12)
02	Missing *glpT* E448K (n=19)
	Unexpected *aph(3'')-Ib* (n=14), *aph(6)-Id* (n=13), *aadA2* (n=12), *aadA5* (n=12), *tet(B)* (n=12)
03	Missing *mgrB*::IS*1* (n=29)
	Unexpected *fosA (*n=19*)*, *oqxA/oqxB* (n=16)
	Unexpected *acrR* mutations (n=42)
04	Missing *gyrA* T83I (n=12)
	Unexpected bla_{PAO} (n=21), $bla_{\text{OXA-486}}$ (n=12), *fosA* (n=20)
	Unexpected *aadA2* (n=11), *aph(3')-IIb* (n=12), *sul1* (n=12), *dfrB5* (n=14)

Missing mutations due to lacking database

- PointFinder can't detect PMs in *A. baumannii* and *P. aeruginosa* – No database!
- *glpT* mutations not present in PointFinder database (ResFinder)
- mgrB::IS*1,* was not possible to report in webtool

⇩

Multiple tools and database can be used:

- AMRFinder+
- CARD
- ResFinder
- Other tools (PathogenWatch?)

DETECTION OF GENES AND MUTATIONS – DISCUSSION

Strain	Examples of problems
01	Missing *gyrA* S81L (n=16) or *parC* S84L (n=15)
	Unexpected *bla*$_{\text{OXA-66}}$ (n=15) or *bla*$_{\text{ADC-25}}$ (n=12)
	Unexpected *aph(3'')-Ib* (n=12), *aph(6)-Id* (n=12), *tet(B)* (n=12)
02	Missing *glpT* E448K (n=19)
	Unexpected *aph(3'')-Ib* (n=14), *aph(6)-Id* (n=13), *aadA2* (n=12), *aadA5* (n=12), *tet(B)* (n=12)
03	Missing *mgrB*::IS*1* (n=29)
	Unexpected *fosA (n=19), oqxA/oqxB* (n=16)
	Unexpected *acrR* mutations (n=42)
04	Missing *gyrA* T83I (n=12)
	Unexpected *bla*$_{\text{PAO}}$ (n=21), *bla*$_{\text{OXA-486}}$ (n=12), *fosA* (n=20)
	Unexpected *aadA2* (n=11), *aph(3')-IIb* (n=12), *sul1* (n=12), *dfrB5* (n=14)

Reporting intrinsic resistance genes

- Present in the strains but do no contribute to the elevated resistance in non-WT phenotype

⇩

Results must be evaluated critically:

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

DETECTION OF GENES AND MUTATIONS – DISCUSSION

Strain	Examples of problems
01	Missing *gyrA* S81L (n=16) or *parC* S84L (n=15)
	Unexpected *bla*$_{\text{OXA-66}}$ (n=15) or *bla*$_{\text{ADC-25}}$ (n=12)
	Unexpected *aph(3'')-Ib* (n=12), *aph(6)-Id* (n=12), *tet(B)* (n=12)
02	Missing *glpT* E448K (n=19)
	Unexpected *aph(3'')-Ib* (n=14), *aph(6)-Id* (n=13), *aadA2* (n=12), *aadA5* (n=12), *tet(B)* (n=12)
03	Missing mgrB::IS*1* (n=29)
	Unexpected *fosA* (n=19), *oqxA/oqxB* (n=16)
	Unexpected *acrR* mutations (n=42)
04	Missing *gyrA* T83I (n=12)
	Unexpected *bla*$_{\text{PAO}}$ (n=21), *bla*$_{\text{OXA-486}}$ (n=12), *fosA* (n=20)
	Unexpected *aadA2* (n=11), *aph(3')-IIb* (n=12), *sul1* (n=12), *dfrB5* (n=14)

Reporting AMR genes for antimicrobials not included in EQA

- These genes are present in the strains but they confer resistance to antimicrobials not relevant for the species
- Tetracycline (*tet(B), tet(G)*), Streptomycin (*aadA2, aadA5, aph(6)-Id, aph(3'')-Ib*) and Kanamycin (*aph(3')-IIb*).

⇩

Results must be evaluated critically:

- Results from the bioinformatics tools should not be reported without critical evaluation
- Carefully report AMR determinants

DETECTION OF GENES AND MUTATIONS – DISCUSSION

Strain	Examples of problems
01	Missing *gyrA* S81L (n=16) or *parC* S84L (n=15)
	Unexpected $bla_{\text{OXA-66}}$ (n=15) or $bla_{\text{ADC-25}}$ (n=12)
	Unexpected *aph(3'')-Ib* (n=12), *aph(6)-Id* (n=12), *tet(B)* (n=12)
02	Missing *glpT* E448K (n=19)
	Unexpected *aph(3'')-Ib* (n=14), *aph(6)-Id* (n=13), *aadA2* (n=12), *aadA5* (n=12), *tet(B)* (n=12)
03	Missing *mgrB*::IS*1* (n=29)
	Unexpected *fosA* (n=19), *oqxA/oqxB* (n=16)
	Unexpected *acrR* mutations (n=42)
04	Missing *gyrA* T83I (n=12)
	Unexpected bla_{PAO} (n=21), $bla_{\text{OXA-486}}$ (n=12), *fosA* (n=20)
	Unexpected *aadA2* (n=11), *aph(3')-IIb* (n=12), *sul1* (n=12), *dfrB5* (n=14)

Reporting PMs (and genes) with unconfirmed impact on the AMR profiles

- PMs in *acrR* and *ramR*

⇩

Results must be evaluated critically:

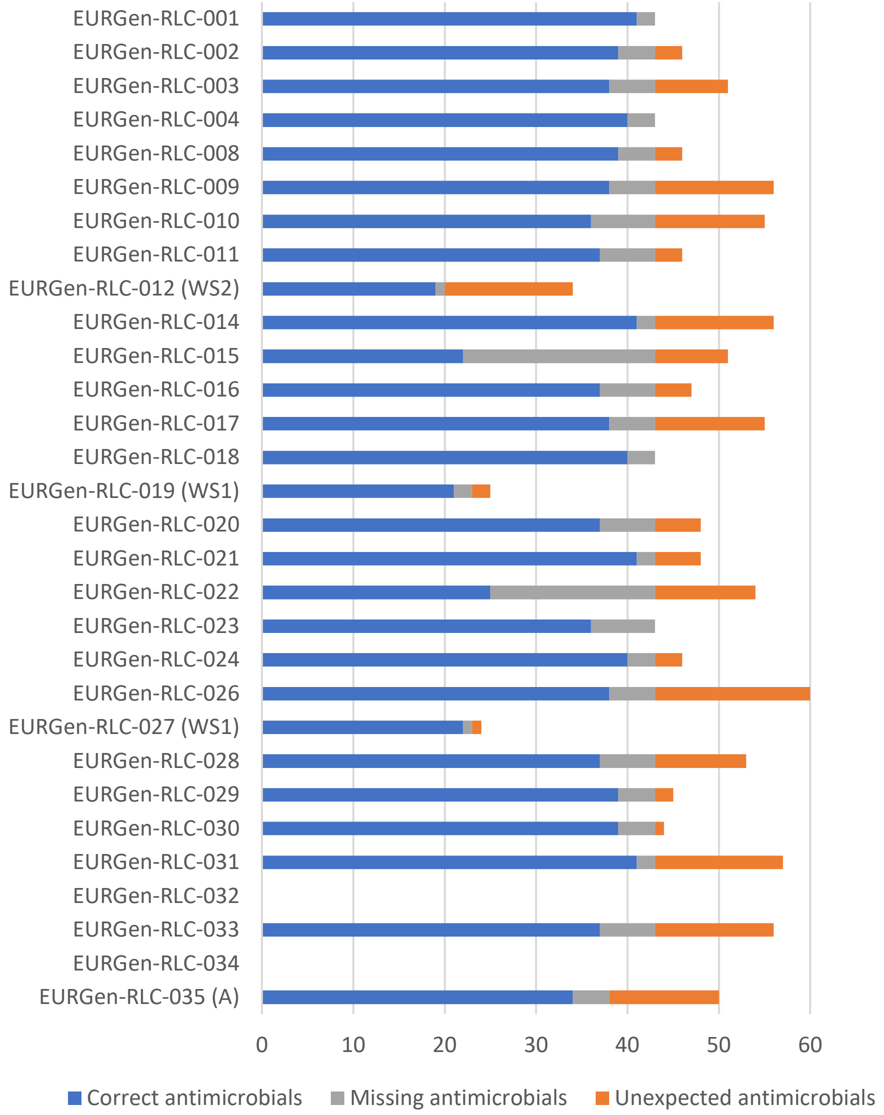
- Might highlight the insufficient knowledge of genetic AMR mechanisms

IN SILICO PREDICTION OF AMR PROFILES – EXPECTED REULTS

Material	**Associated prediction of AMR profiles**
EURGen-2023-01 (*A. baumannii*)	**Expected**: Amikacin, ciprofloxacin, cefepime, ceftazidime, ceftazidime-avibactam, gentamicin, imipenem, meropenem, piperacillin-tazobactam, tobramycin
	Intrinsic: Aztreonam, fosfomycin
EURGen-2023-02 (*E. coli*)	**Expected:** Amikacin, amoxicillin-clavulanic acid, ampicillin, aztreonam, cefepime, cefotaxime, ceftazidime, ceftazidime-avibactam, ciprofloxacin, ertapenem, <u>fosfomycin</u>, gentamicin, imipenem, meropenem, piperacillin-tazobactam, sulfamethoxazole, tobramycin, trimethoprim
	Expected non-mandatory: Colistin[a]
EURGen-2023-03 (*K. pneumoniae*)	**Expected:** Aztreonam, cefepime, cefotaxime, ceftazidime, ciprofloxacin
	Expected non-mandatory: Colistin[b]
	Intrinsic: Ampicillin
EURGen-2023-04 (*P. aeruginosa*)	**Expected:** Amikacin, cefepime, ceftazidime, ceftazidime-avibactam, ciprofloxacin, gentamicin, imipenem, meropenem, piperacillin-tazobactam, tobramycin

Funded by the European Union

IN SILICO PREDICTION OF AMR PROFILES – EXPECTED REULTS



Total submissions: 105

- 10.5% submissions were fully correct (n=11)
- 63.8% had missing antimicrobials (n=67)
- 52.2% had unexpected antimicrobials (n=58)

29.5% Simultaneous (n=31)

Average score: 87.2%

- None of participants achieved 100% of their max. possible score

IN SILICO PREDICTION OF AMR PROFILES – SUBMITTED REULTS

Strain	Total expected	Correct	Missing	Unexpected
EURGen-2023-01	260	214 (82.3%)	46 (17.7%)	54 (↑ 20.8%)
EURGen-2023-02	486	430 (88.5%)	56 (11.5%)	2 (↑ 0.4%)
EURGen-2023-03	130	120 (92.3%)	10 (7.7%)	43 (↑ 33.1%)
EURGen-2023-04	260	228 (87.7%)	32 (12.3%)	90 (↑ 34.6%)

IN SILICO PREDICTION OF AMR PROFILES – SUBMITTED REULTS

Strain EURGen-2023-01 (*A. baumannii*)

	Expected										Unexpected												
Laboratories	Amikacin	Cefepime	Ceftazidime	Ceftazidime-avibactam	Ciprofloxacin	Gentamicin	Imipenem	Meropenem	Piperacillin-tazobactam	Tobramycin	Ampicillin	Amoxicillin-clavulanic acid	Aztreonam[b]	Cefotaxime	Colistin	Ertapenem	Fosfomycin[b]	Sulfamethoxazole	Tigecycline	Trimethoprim	**Correct (nr.)**	**Missing (nr.)**	**Unexpected (nr.)**
EURGen-RLC-001	x	x	x	x	x	x	x	x	x	x											10	0	0
EURGen-RLC-002	x	x	x	-	-	x	x	x	x	x											8	2	0
EURGen-RLC-003	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-004	x	x	x	x	x	x	x	x	x	x											10	0	0
EURGen-RLC-008	x	x	x	-	-	x	x	x	x	x											8	2	0
EURGen-RLC-009	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-010	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-011	x	x	x	x	-	x	x	x	x	x											9	1	0
EURGen-RLC-012	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		x	10	0	7
EURGen-RLC-014	x	x	x	x	-	x	x	x	x	x	x	x		x		x					9	1	4
EURGen-RLC-015	x	-	-	-	-	x	-	x	-	-	x	x									3	7	2
EURGen-RLC-016	x	x	x	x	x	x	x	x	x	x			x				x				10	0	0
EURGen-RLC-017	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-018	x	x	x	x	-	x	x	x	x	x											9	1	0
EURGen-RLC-020	x	x	x	-	x	x	x	x	-	-			x								7	3	0
EURGen-RLC-021	x	x	x	x	x	x	x	x	x	x			x								10	0	0
EURGen-RLC-022	x	-	-	-	-	x	x	x	x	x	x	x		x		x					6	4	4
EURGen-RLC-023	x	-	-	-	-	x	x	x	-	x											5	5	0
EURGen-RLC-024	x	x	x	x	-	x	x	x	x	x											9	1	0
EURGen-RLC-026	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-028	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-029	x	x	x	-	-	x	x	x	x	x											8	2	0
EURGen-RLC-030	x	x	x	x	-	x	x	x	x	x											9	1	0
EURGen-RLC-031	x	x	x	x	-	x	x	x	x	x	x	x	x	x		x			x		9	1	5
EURGen-RLC-033	x	x	x	-	-	x	x	x	x	x	x	x		x		x					8	2	4
EURGen-RLC-035	x	x	x	-	x	x	x	x	x	x	x	x	x	x		x					9	1	4
EURGen-RLC-019																					NA	NA	NA
EURGen-RLC-027																					NA	NA	NA
EURGen-RLC-032																					NA	NA	NA
EURGen-RLC-034																					NA	NA	NA
Correct (nr.)	26	23	23	11	7	26	25	26	23	24	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Total		
Missing or UN (nr.)	0	3	3	15	19	0	1	0	3	2	13	13	6	12	1	12	1	1	1	1	214	46	54

- Missing from most submitted results
  - Ciprofloxacin (n=19)
    - Most participants used ResFinder which lacks database of AMR PMs for *A. baumannii*
  - Ceftazidime-avibactam (n=15)
    - Missing from *bla*$_{\text{NDM-1}}$ resistance profile in ResFinder database (but present for other genes)
- Reporting antimicrobials not included in the EQA
  - Ampicillin (n=13), Amoxicillin-clavulanic acid (n=13), Cefotaxime (n=12), Etrapenem (n=12)

Results must be evaluated critically:

- Results can / should be confirmed with other tools
- Points to insufficient knowledge of genetic AMR mechanisms

IN SILICO PREDICTION OF AMR PROFILES – SUBMITTED REULTS

Strain EURGen-2023-02 (*E. coli*)

	Expected																							
Laboratories	Amikacin	Amoxicillin-clavulanic acid	Ampicillin	Aztreonam	Cefepime	Cefotaxime	Ceftazidime	Ceftazidime-avibactam	Ciprofloxacin	Colistin[a]	Ertapenem	Fosfomycin	Gentamicin	Imipenem	Meropenem	Piperacillin-tazobactam	Sulfamethoxazole	Tobramycin	Trimethoprim	Tigecycline	**Correct (nr.)**	**Missing (nr.)**	**Unexpected (nr.)**	**Expected non-mandatory (nr.)**
EURGen-RLC-001	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		18	0	0	1
EURGen-RLC-002	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-003	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-004	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x		18	0	0	0
EURGen-RLC-008	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-009	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-010	-	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		15	3	0	0
EURGen-RLC-011	x	x	x	x	x	x	x	-	x	-	x	-	-	x	x	x	x	x	x		15	3	0	0
EURGen-RLC-014	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x		17	1	0	0
EURGen-RLC-015	x	x	x	-	-	-	-	-	x	-	-	-	x	-	x	-	x	x	x		9	9	0	0
EURGen-RLC-016	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x		18	0	0	0
EURGen-RLC-017	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-018	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x		17	1	0	0
EURGen-RLC-019	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-020	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x	x	x	x	17	1	1	1
EURGen-RLC-021	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		18	0	0	1
EURGen-RLC-022	-	x	x	-	-	x	x	-	x	-	x	-	-	x	x	x	-	-	x		10	8	0	0
EURGen-RLC-023	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x		17	1	0	0
EURGen-RLC-024	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x		17	1	0	0
EURGen-RLC-026	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		16	2	0	0
EURGen-RLC-027	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x		17	1	0	0
EURGen-RLC-028	x	x	x	-	x	x	x	-	x	-	x	-	-	x	x	x	x	x	x		14	4	0	0
EURGen-RLC-029	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x	x	x	x	x		17	1	0	1
EURGen-RLC-030	x	x	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x	x		17	1	0	1
EURGen-RLC-031	x	x	x	x	x	x	x	x	x	-	x	-	x	x	x	x	x	x	x	x	17	1	1	0
EURGen-RLC-033	x	x	x	-	x	x	x	-	x	-	x	-	x	x	x	x	x	x	x		15	3	0	0
EURGen-RLC-035	x	x	x	x	x	x	x	-	x	-	x	-	x	x	x	-	x	x	x		15	3	0	0
EURGen-RLC-012																					NA	NA	NA	NA
EURGen-RLC-032																					NA	NA	NA	NA
EURGen-RLC-034																					NA	NA	NA	NA
Correct (nr.)	25	27	27	23	25	26	26	11	27	5	26	6	24	26	27	25	26	26	27	NA	Total			
Missing or UN (nr.)	2	0	0	4	2	1	1	16	0	22	1	21	3	1	0	2	1	1	0	2	430	56	2	5

- Missing from most submitted results
  - Ceftazidime-avibactam (n=16)
    - Missing from *bla*$_{\text{NDM-5}}$ R-profile in ResFinder database (but present for other genes)
  - Fosfomycin (n=21)
    - *glpT* E448K missing in ResFinder

Results must be evaluated critically:

- Results can / should be confirmed with other tools

Funded by the European Union

IN SILICO PREDICTION OF AMR PROFILES – SUBMITTED REULTS

Strain EURGen-2023-03 (*K.pneumoniae*)

	Expected						Unexpected												
	Aztreonam	Cefepime	Cefotaxime	Ceftazidime	Ciprofloxacin	Colistin[a]	Ampicillin[b]	Amoxicillin-clavulanic acid	Ertapenem	Fosfomycin	Gentamicin	Imipenem	Meropenem	Piperacillin-tazobactam	Trimethoprim	**Correct (nr.)**	**Missing (nr.)**	**Unexpected (nr.)**	**Expected non-mandatory (nr.)**
EURGen-RLC-001	x	x	x	x	x	-	x									5	0	0	0
EURGen-RLC-002	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-003	x	x	x	x	x	-				x					x	5	0	2	0
EURGen-RLC-004	x	x	x	-	x	-	x									4	1	0	0
EURGen-RLC-008	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-009	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-010	x	x	x	x	x	-	x			x						5	0	1	0
EURGen-RLC-011	x	x	x	x	-	-	x			x					x	4	1	2	0
EURGen-RLC-014	x	x	x	x	x	-	x			x	x				x	5	0	3	0
EURGen-RLC-015	-	-	-	-	x	-	x			x						1	4	1	0
EURGen-RLC-016	x	x	x	x	x	-	x	x	x					x		5	0	3	0
EURGen-RLC-017	-	x	x	x	x	-	x			x						4	1	1	0
EURGen-RLC-018	x	x	x	x	x	-	x									5	0	0	0
EURGen-RLC-019	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-020	x	x	x	x	x	-	x	x						x		5	0	2	0
EURGen-RLC-021	-	x	x	x	-	-	x	x		x				x	x	3	2	4	0
EURGen-RLC-022	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-023	-	x	x	x	x	x	x									4	1	0	1
EURGen-RLC-024	x	x	x	x	x	-	x			x						5	0	1	0
EURGen-RLC-026	x	x	x	x	x	-			x	x		x	x	x	x	5	0	6	0
EURGen-RLC-027	x	x	x	x	x	x	x	x								5	0	1	1
EURGen-RLC-028	x	x	x	x	x	-	x			x						5	0	1	0
EURGen-RLC-029	x	x	x	x	x	-	x			x						5	0	1	0
EURGen-RLC-030	x	x	x	x	x	-	x	x								5	0	1	0
EURGen-RLC-031	x	x	x	x	x	-	x			x						5	0	1	0
EURGen-RLC-033	x	x	x	x	x	-	x			x					x	5	0	2	0
EURGen-RLC-012																NA	NA	NA	NA
EURGen-RLC-032																NA	NA	NA	NA
EURGen-RLC-034																NA	NA	NA	NA
EURGen-RLC-035																NA	NA	NA	NA
Correct (nr.)	22	25	25	24	24	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	Total			
Missing or UN (nr.)	4	1	1	2	2	24	24	5	2	18	1	1	1	4	11	120	10	43	2

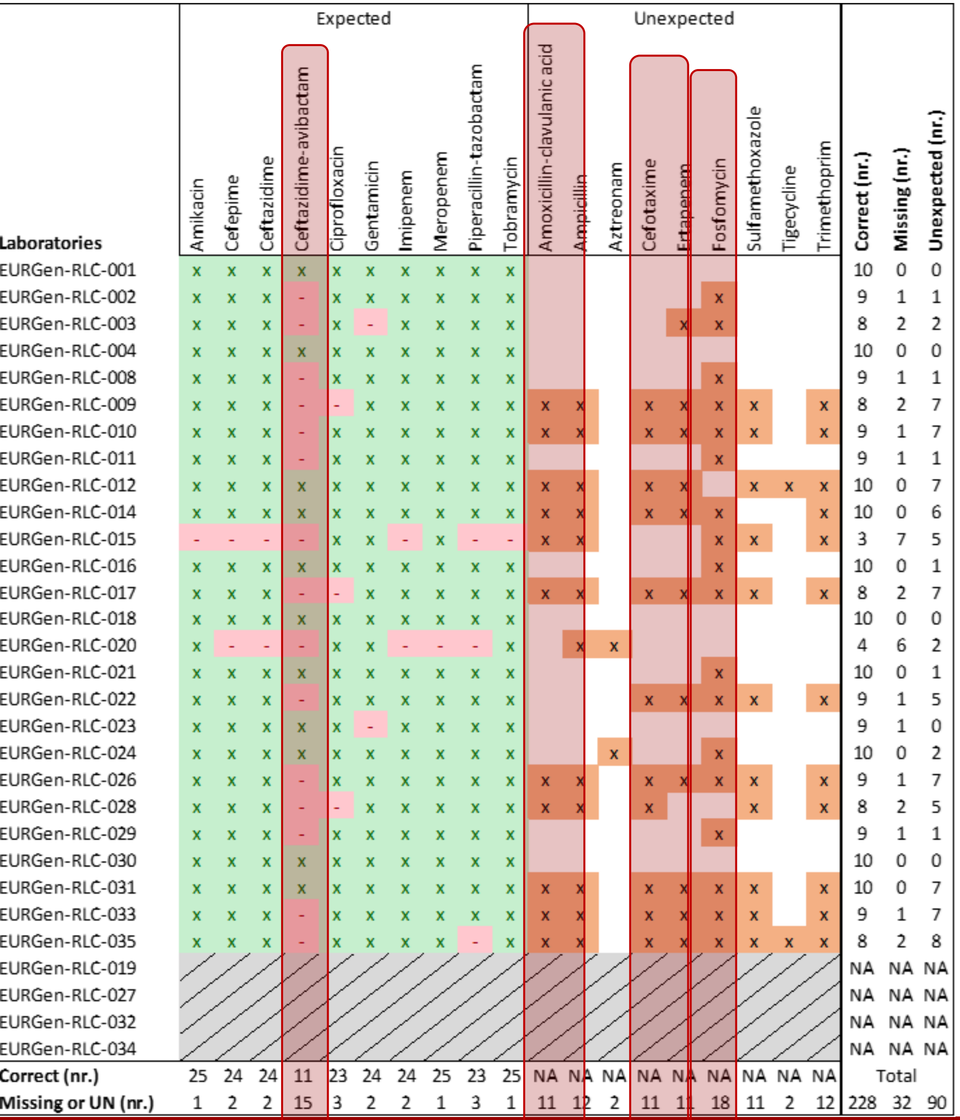
- Missing from most submitted results
  - Colistin (n=24)
    - Due to missing *mgrB*::IS*1* mutation
- Reporting of Intrinsic resistance
  - Ampicillin (n=24)
  - Fosfomycin (n=18)

Results must be evaluated critically:

- Results can / should be confirmed with other tools
- Points to insufficient knowledge of genetic mechanisms of AMR

IN SILICO PREDICTION OF AMR PROFILES – SUBMITTED REULTS

Strain EURGen-2023-04 (*P. aeruginosa*)



	Expected										Unexpected											
Laboratories	Amikacin	Cefepime	Ceftazidime	Ceftazidime-avibactam	Ciprofloxacin	Gentamicin	Imipenem	Meropenem	Piperacillin-tazobactam	Tobramycin	Amoxicillin-clavulanic acid	Ampicillin	Aztreonam	Cefotaxime	Ertapenem	Fosfomycin	Sulfamethoxazole	Tigecycline	Trimethoprim	**Correct (nr.)**	**Missing (nr.)**	**Unexpected (nr.)**
EURGen-RLC-001	x	x	x	x	x	x	x	x	x	x										10	0	0
EURGen-RLC-002	x	x	x	-	x	x	x	x	x	x						x				9	1	1
EURGen-RLC-003	x	x	x	-	x	-	x	x	x	x					x	x				8	2	2
EURGen-RLC-004	x	x	x	x	x	x	x	x	x	x										10	0	0
EURGen-RLC-008	x	x	x	-	x	x	x	x	x	x						x				9	1	1
EURGen-RLC-009	x	x	x	-	-	x	x	x	x	x	x	x		x	x	x	x		x	8	2	7
EURGen-RLC-010	x	x	x	-	x	x	x	x	x	x	x	x		x	x	x	x		x	9	1	7
EURGen-RLC-011	x	x	x	-	x	x	x	x	x	x						x				9	1	1
EURGen-RLC-012	x	x	x	x	x	x	x	x	x	x	x	x		x	x		x	x	x	10	0	7
EURGen-RLC-014	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x			x	10	0	6
EURGen-RLC-015	-	-	-	-	x	x	-	x	-	-	x	x				x	x		x	3	7	5
EURGen-RLC-016	x	x	x	x	x	x	x	x	x	x						x				10	0	1
EURGen-RLC-017	x	x	x	-	-	x	x	x	x	x	x	x		x	x	x	x		x	8	2	7
EURGen-RLC-018	x	x	x	x	x	x	x	x	x	x										10	0	0
EURGen-RLC-020	x	-	-	-	x	x	-	-	-	x		x	x							4	6	2
EURGen-RLC-021	x	x	x	x	x	x	x	x	x	x						x				10	0	1
EURGen-RLC-022	x	x	x	-	x	x	x	x	x	x				x	x	x	x		x	9	1	5
EURGen-RLC-023	x	x	x	x	x	-	x	x	x	x										9	1	0
EURGen-RLC-024	x	x	x	x	x	x	x	x	x	x			x			x				10	0	2
EURGen-RLC-026	x	x	x	-	x	x	x	x	x	x	x	x		x	x	x	x		x	9	1	7
EURGen-RLC-028	x	x	x	-	-	x	x	x	x	x	x	x		x			x		x	8	2	5
EURGen-RLC-029	x	x	x	-	x	x	x	x	x	x						x				9	1	1
EURGen-RLC-030	x	x	x	x	x	x	x	x	x	x										10	0	0
EURGen-RLC-031	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x		x	10	0	7
EURGen-RLC-033	x	x	x	-	x	x	x	x	x	x	x	x		x	x	x	x		x	9	1	7
EURGen-RLC-035	x	x	x	-	x	x	x	x	-	x	x	x		x	x	x	x	x	x	8	2	8
EURGen-RLC-019																				NA	NA	NA
EURGen-RLC-027																				NA	NA	NA
EURGen-RLC-032																				NA	NA	NA
EURGen-RLC-034																				NA	NA	NA
Correct (nr.)	25	24	24	11	23	24	24	25	23	25	NA	NA	NA	NA	NA	NA	NA	NA	NA	Total		
Missing or UN (nr.)	1	2	2	15	3	2	2	1	3	1	11	12	2	11	11	18	11	2	12	228	32	90

- Missing from most submitted results
  - – Ceftazidime-avibactam (n=15)
    - Missing from $bla_{\text{NDM-5}}$ R-profile in ResFinder database (but present for other genes)

- Reporting antimicrobials not included in the EQA
  - – Ampicillin (n=12), Amoxicillin-clavulanic acid (n=11), Cefotaxime (n=11), Etrapenem (n=11)

- Reporting incorrect resistance
  - – Fosfomycin (n=18)
  - – Not intrinsic in *P. aruginosa* (although carries *fosA)*

Results must be evaluated critically:

- Results can / should be confirmed with other tools
- Points to insufficient knowledge of genetic mechanisms of AMR

GENERAL RECOMMENDATIONS

- For discrepancies due to misinterpretation of the EQA protocol and/or insufficient knowledge about certain genetic elements:
  - Laboratories should ensure sufficient knowledge about the genetic mechanisms mediating AMR and other important genetic elements;
  - Laboratories should ensure the use of multiple bioinformatics tools and databases for the detection of genetic determinants since bioinformatics tools and databases can be limited to the analysis of only a few bacterial species which contributes to the false-negative results;
  - Laboratories should be familiar with the bioinformatics tools they use, and the contents of the respective databases;
  - Laboratories should analyse their results critically and, when needed, perform confirmatory testing, to ensure that the information being reported is accurate.

Funded by the European Union

GENERAL RECOMMENDATIONS

- For discrepancies due to variations between the type of data and the chosen bioinformatics tools and databases:
  - 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.
- For curators of bioinformatics tools for AMR surveillance:
  - Curators of widely-used bioinformatics tools and databases should try to improve the databases and regularly update the databases
  - Curators of bioinformatics tools and databases should engage in ongoing, active dialogue to ensure conformity between approaches;

Funded by the European Union

QUALITY CONTROL ANALYSIS OF SUBMITTED SEQUENCES

- METRICS INCLUDED IN QUALITY CONTROL ANALYSIS
- TOTAL ACHIEVED ILLUMINA SCORES
- RESULTS FROM QUALITY METRICS- ILLUMINA SEQUENCES
- NOTE ON OXFORD NANOPORE SEQUENCING

Funded by the European Union

QC METRICS

- Applied thresholds are metric specific
- Submitted genomes which deviate more than 10% from expected genomic size or does not have at least 95% of cgMLST alleles identified are expected to fail outright
- Coverage, Q-scores and proportion of mapping reads - are to some extent platform dependant
- Correlation groups 2 and 3 are non-correlating groups, submission failing across both are indicative of large QC issue

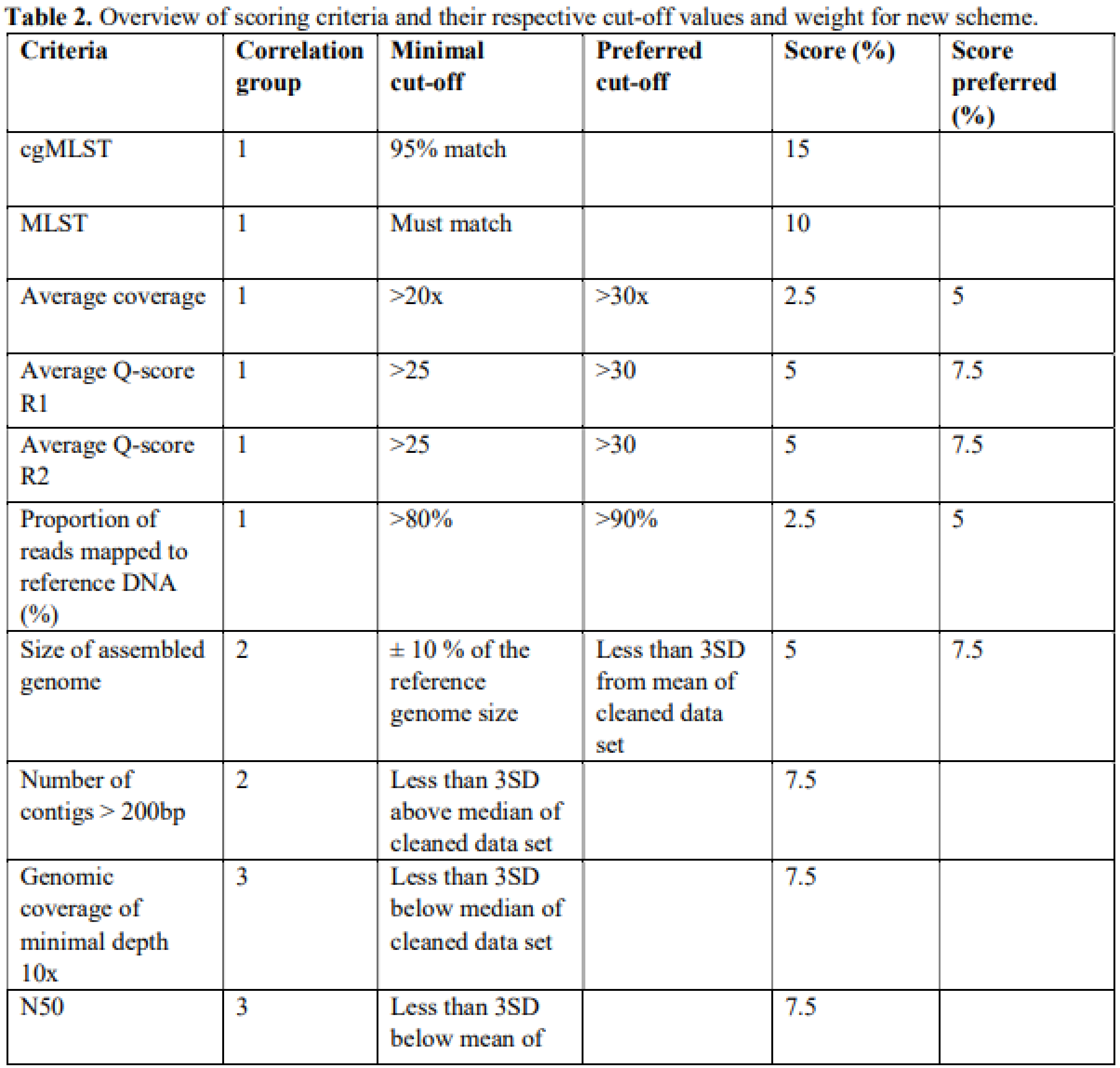


Table 2. Overview of scoring criteria and their respective cut-off values and weight for new scheme.

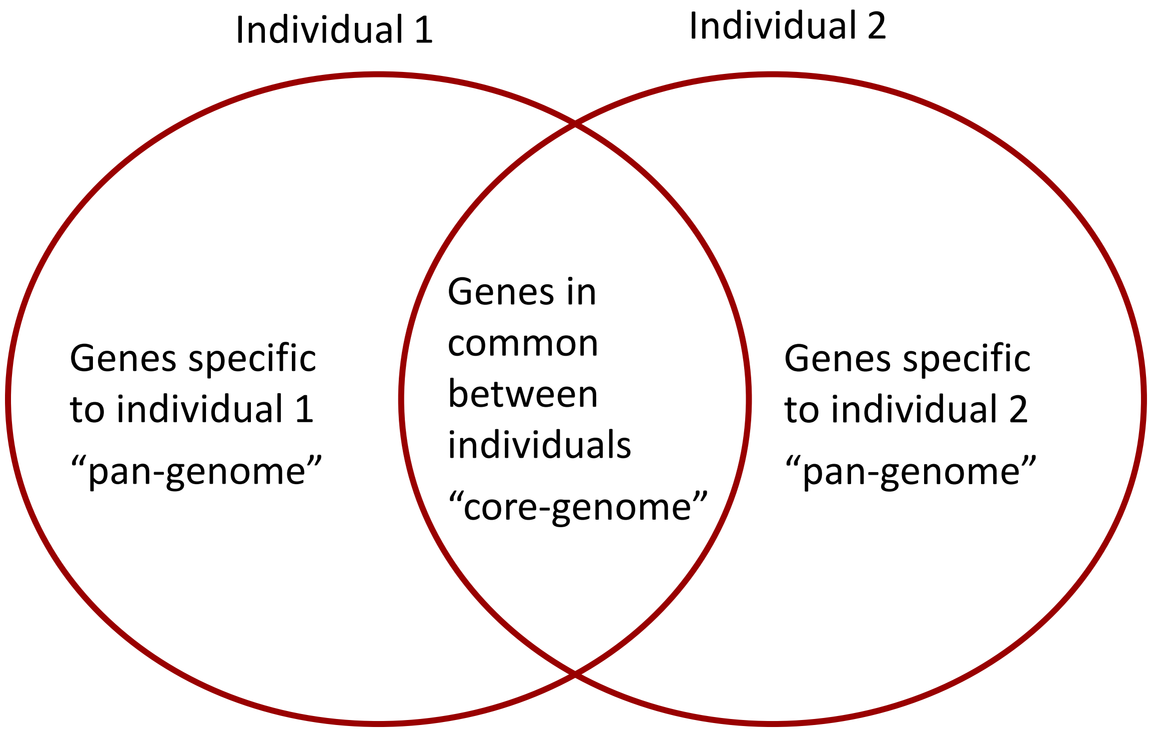
Criteria	Correlation group	Minimal cut-off	Preferred cut-off	Score (%)	Score preferred (%)
cgMLST	1	95% match		15	
MLST	1	Must match		10	
Average coverage	1	>20x	>30x	2.5	5
Average Q-score R1	1	>25	>30	5	7.5
Average Q-score R2	1	>25	>30	5	7.5
Proportion of reads mapped to reference DNA (%)	1	>80%	>90%	2.5	5
Size of assembled genome	2	± 10 % of the reference genome size	Less than 3SD from mean of cleaned data set	5	7.5
Number of contigs > 200bp	2	Less than 3SD above median of cleaned data set		7.5	
Genomic coverage of minimal depth 10x	3	Less than 3SD below median of cleaned data set		7.5	
N50	3	Less than 3SD below mean of		7.5	

Table 2 from the Readme file for Illumina sequencing

METRICS SHORT REFRESHER - MLST

- MLST
  - Scheme of 7 specific core genes in isolates
  - Each gene has a specific DNA sequence
  - Any genetic variation is classified as a new allele
  - The combination of these 7 unique alleles defines a sequence type
  - Perfect match expected

- cgMLST
  - The concept above is applied to a predefined set of core genes expected to be present in every individual member of the species
  - 95% match expected



Individual 1

Individual 2

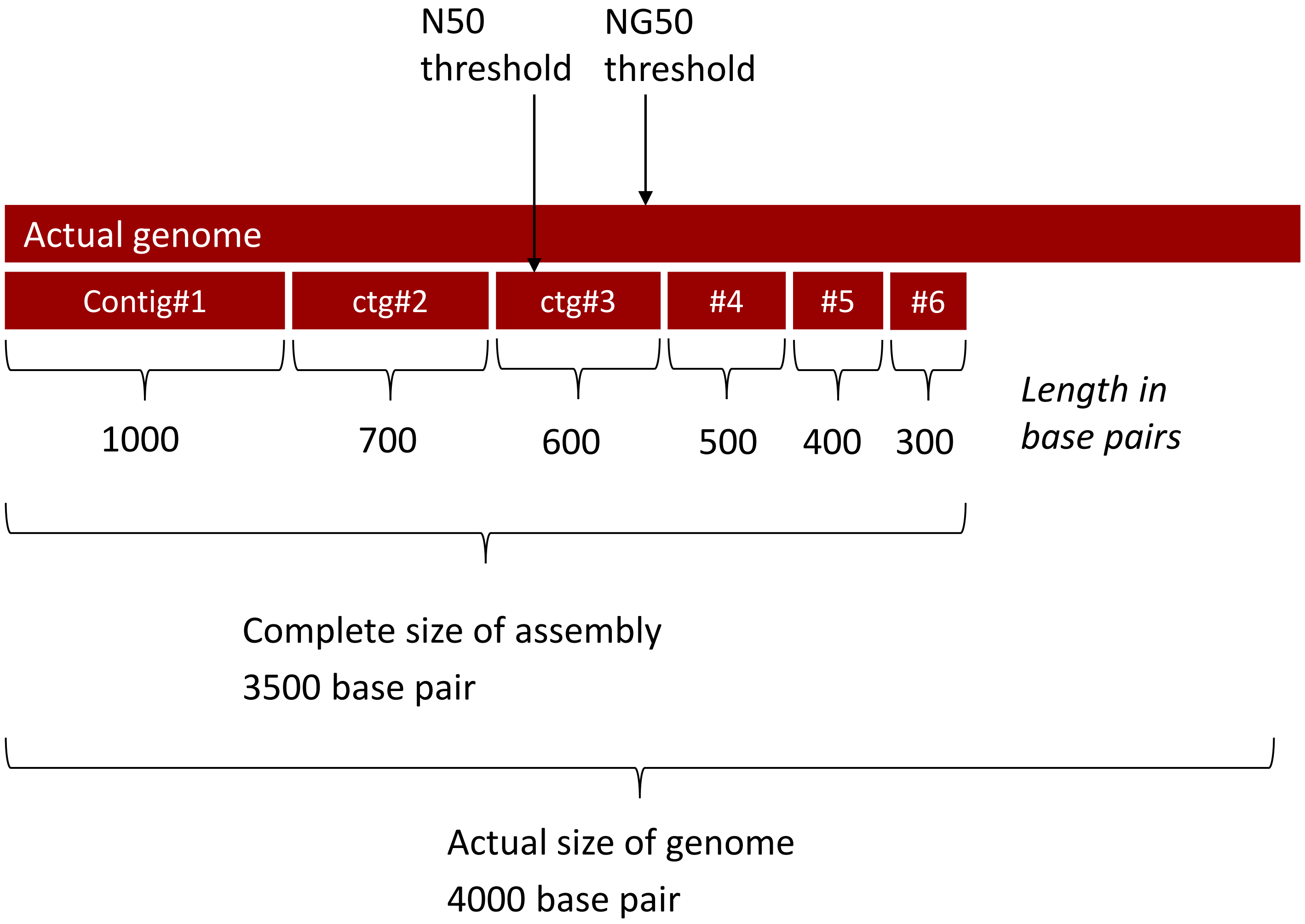
Genes specific to individual 1 “pan-genome”

Genes in common between individuals “core-genome”

Genes specific to individual 2 “pan-genome”

ASSEMBLY METRICS

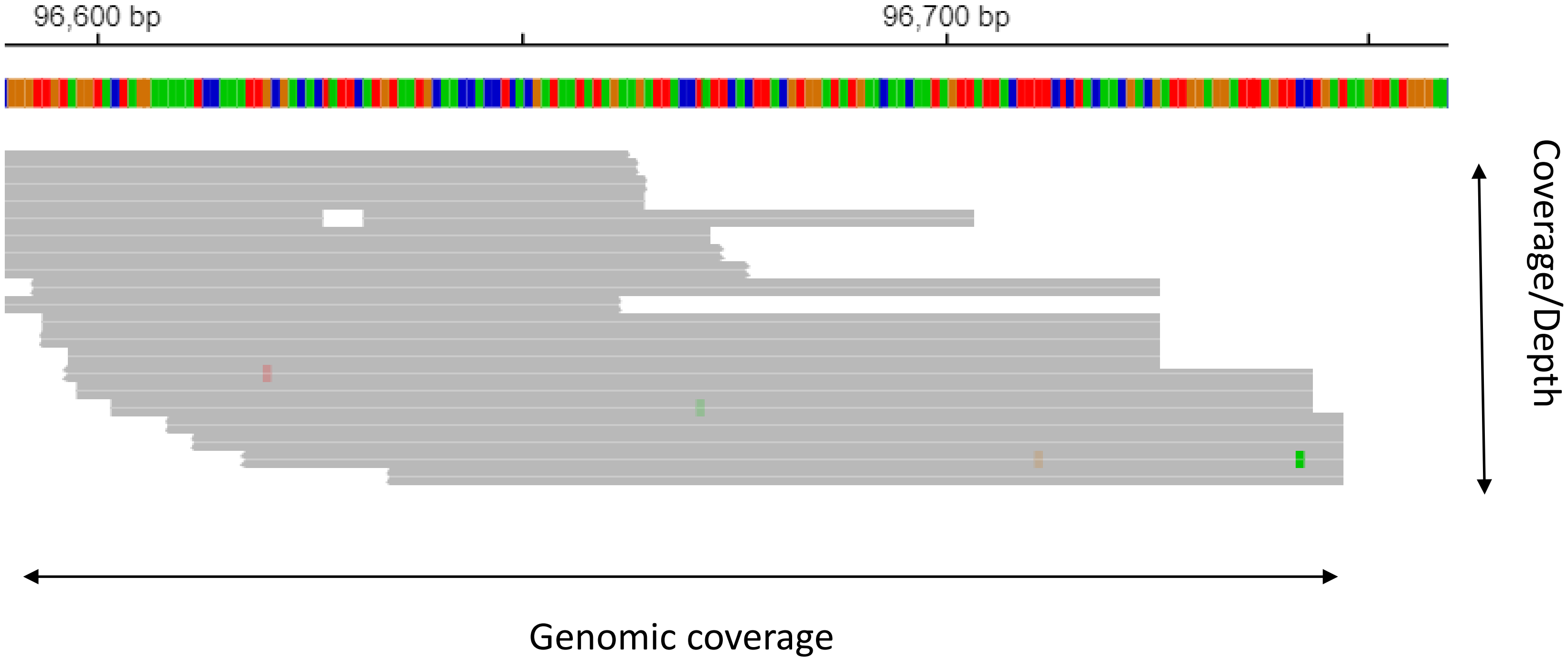
- Commonly used metrics
  - Number of contigs more than 200 bp
    - Here 6
  - N50
    - Adding lengths of contigs, going from longest to shortest, what is the minimal length included to reach 50% of the complete assembly size
  - Size of assembly compared to reference genome
    - Indicative of contamination or sequence quality
    - Here 87.5%
- Adjusted quality thresholds used for evaluation



COVERAGE METRICS

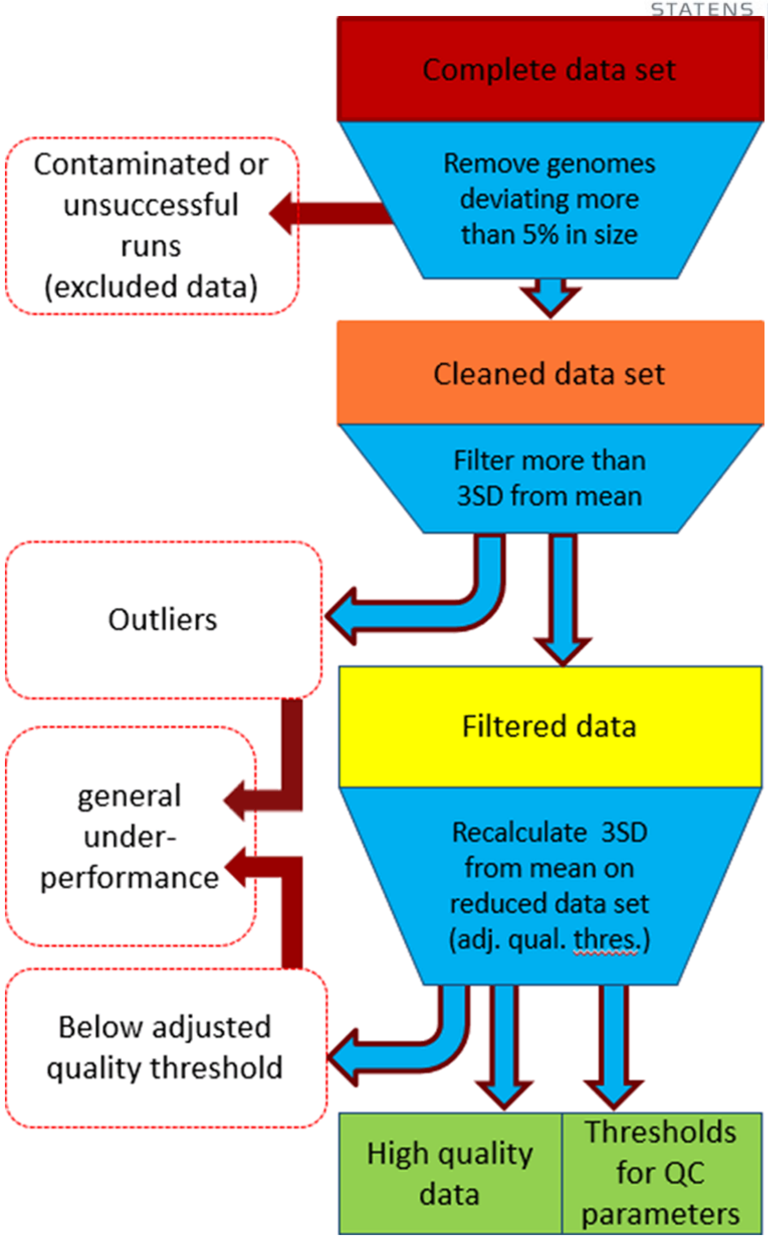
- Average coverage
  - On average, how many times is the genome covered by reads
  - Completely dependent on:
    - Sequencing yield
    - Size of the genome
  - Minimum of 20x, 30x prefered

- Genomic coverage of minimal depth 10x
  - The breadth of the coverage of the reference genome
  - Low depth leads to greater uncertainty of the base call
  - Assembly is complicated by low coverage regions
  - Adjusted quality threshold



DEFINITION OF QC THRESHOLDS

- Data was evaluated in three steps:
  - Initial exclusion
    - 10% > deviance from reference size
    - <95% cgMLST alleles
  - Outlier detection
    - More than 3 standard deviations from the mode
    - Mode is estimated either by mean or by median of the submitted genome
  - Setting adjusted quality threshold
    - Detected outliers were removed and standard deviations and mode was recalculated
    - This is done to remove data points which deviate by a large margin from the main body of data, as such outliers have a very large impact on the estimation of standard deviation and mode of the distribution



Complete data set

Remove genomes deviating more than 5% in size

Contaminated or unsuccessful runs (excluded data)

Cleaned data set

Filter more than 3SD from mean

Outliers

Filtered data

Recalculate 3SD from mean on reduced data set (adj. qual. thres.)

general under-performance

Below adjusted quality threshold

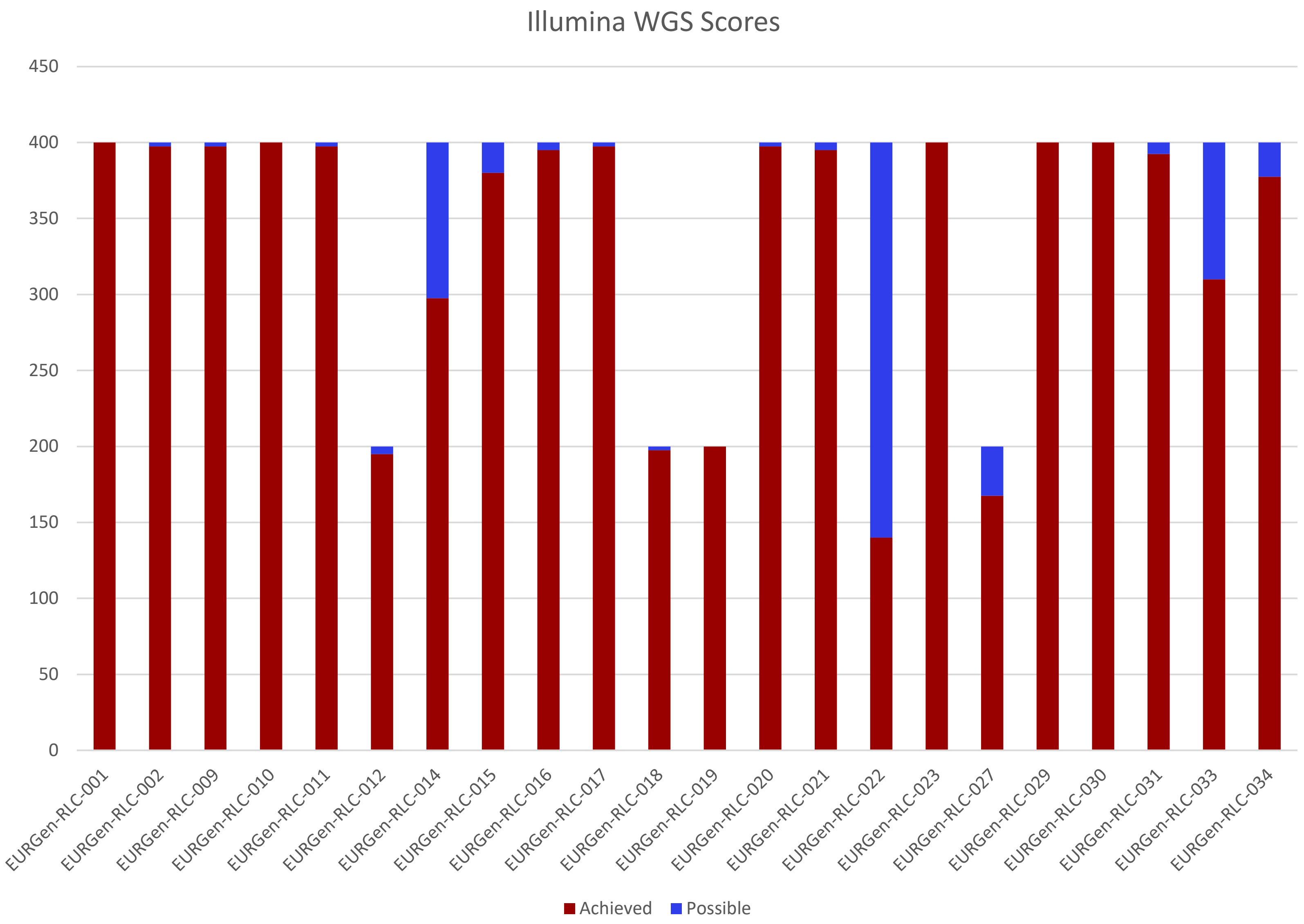
High quality data

Thresholds for QC parameters

Ref: Kristensen T. and Sørensen L.H. et al, Microb Genom. 2023 Aug;9(8):mgen001076. doi: 10.1099/mgen.0.001076. PMID: 37526643

TOTAL ACHIEVED ILLUMINA SCORES

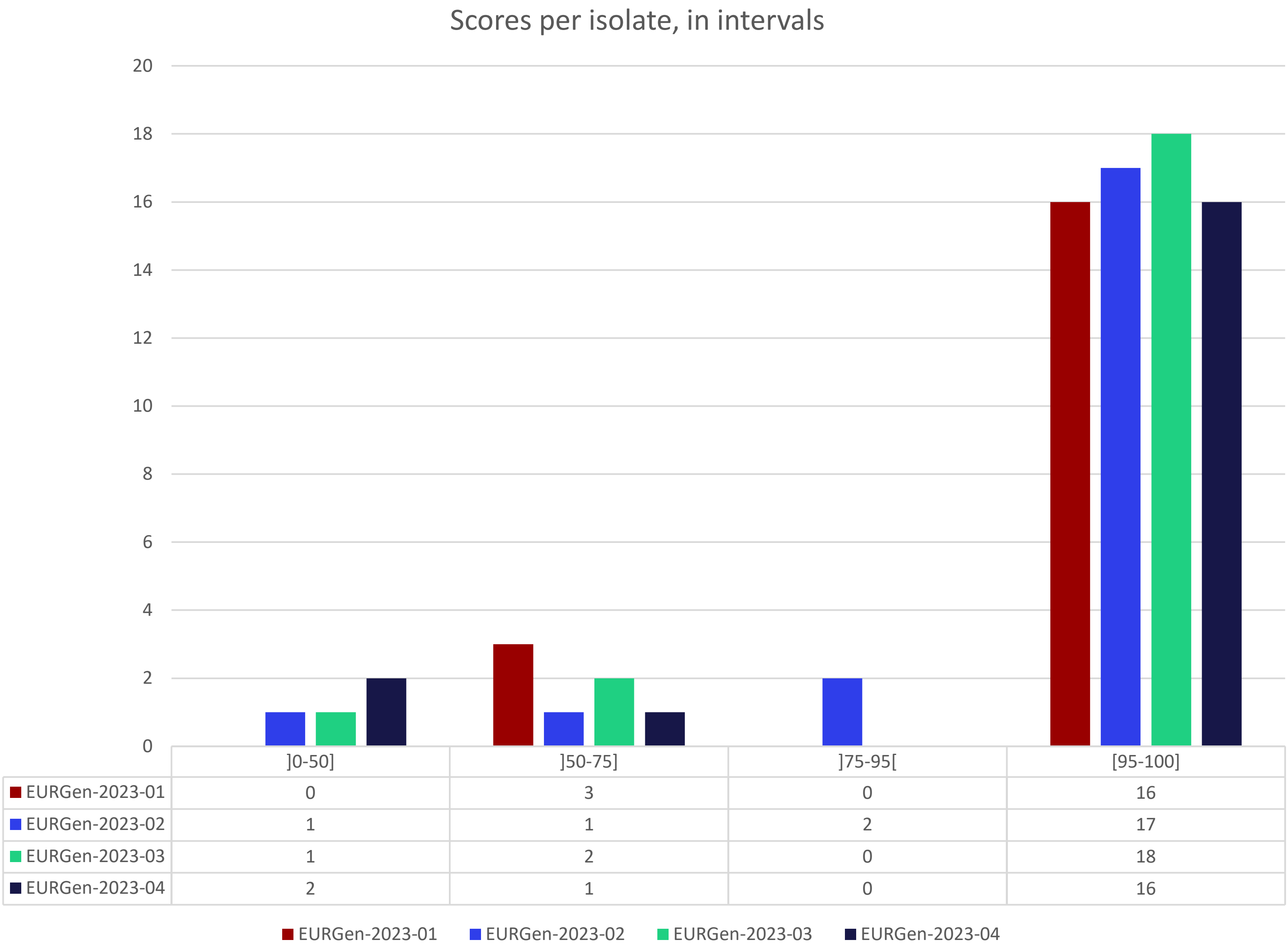
- We received Illumina WGS data from 22 laboratories
- Four participants submitted WGs for only two species, remaining submitted for all four available
- Three were less successful, mainly due to high variance in genome size compared to reference and lack of coverage



TOTAL ACHIEVED ILLUMINA SCORES - SUMMARY

- For each species we received:
  - 19 *A. baumannii*
  - 21 *E. coli*
  - 21 *K. pneumonia*
  - 19 *P. aeruginosa*

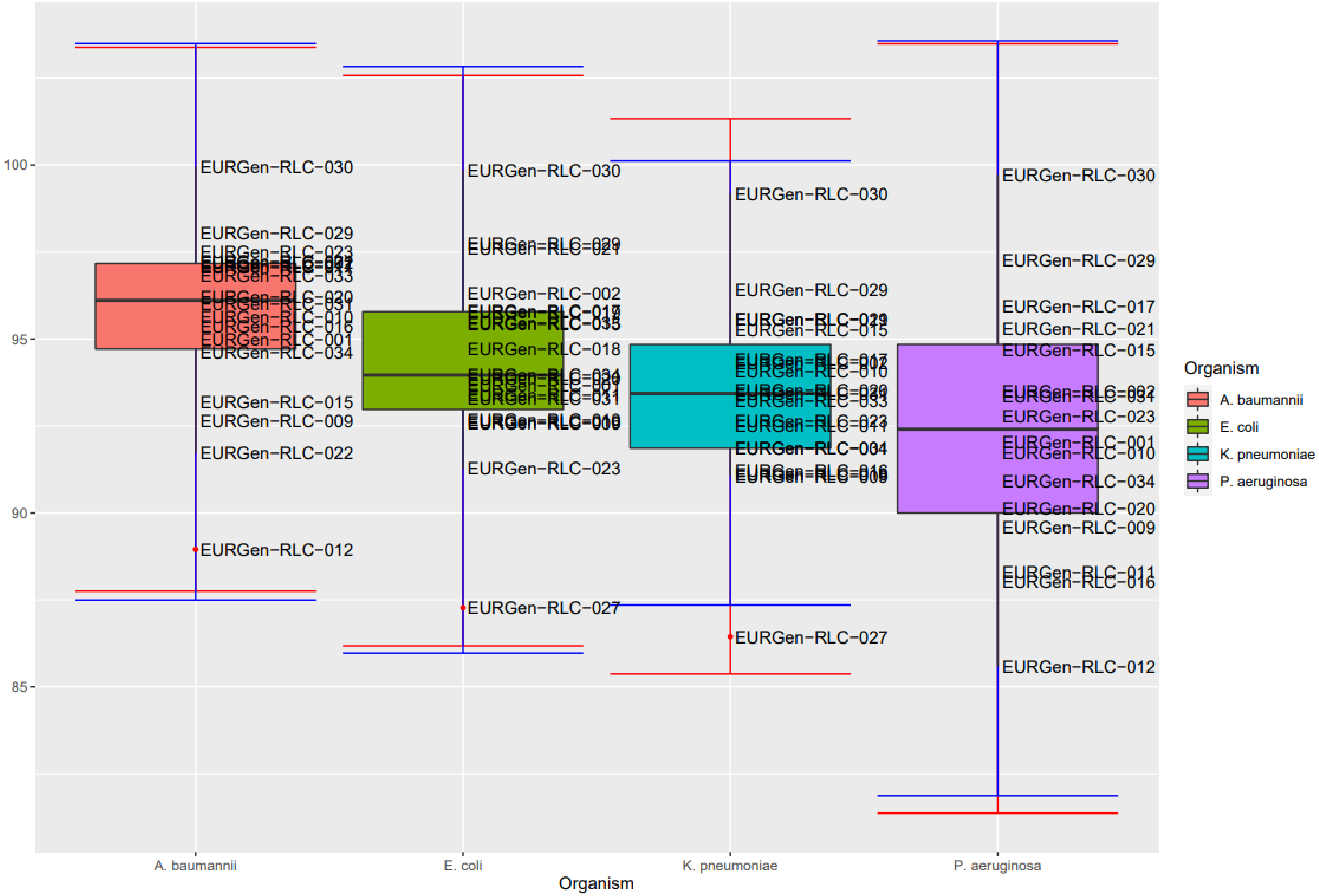
- Issues were evenly found in all species, though *A. baumannii* and *P. aeruginosa* showed widest distributions in QC parameters



	]0-50]	]50-75]	]75-95[	[95-100]
🟥 EURGen-2023-01	0	3	0	16
🟦 EURGen-2023-02	1	1	2	17
🟩 EURGen-2023-03	1	2	0	18
■ EURGen-2023-04	2	1	0	16

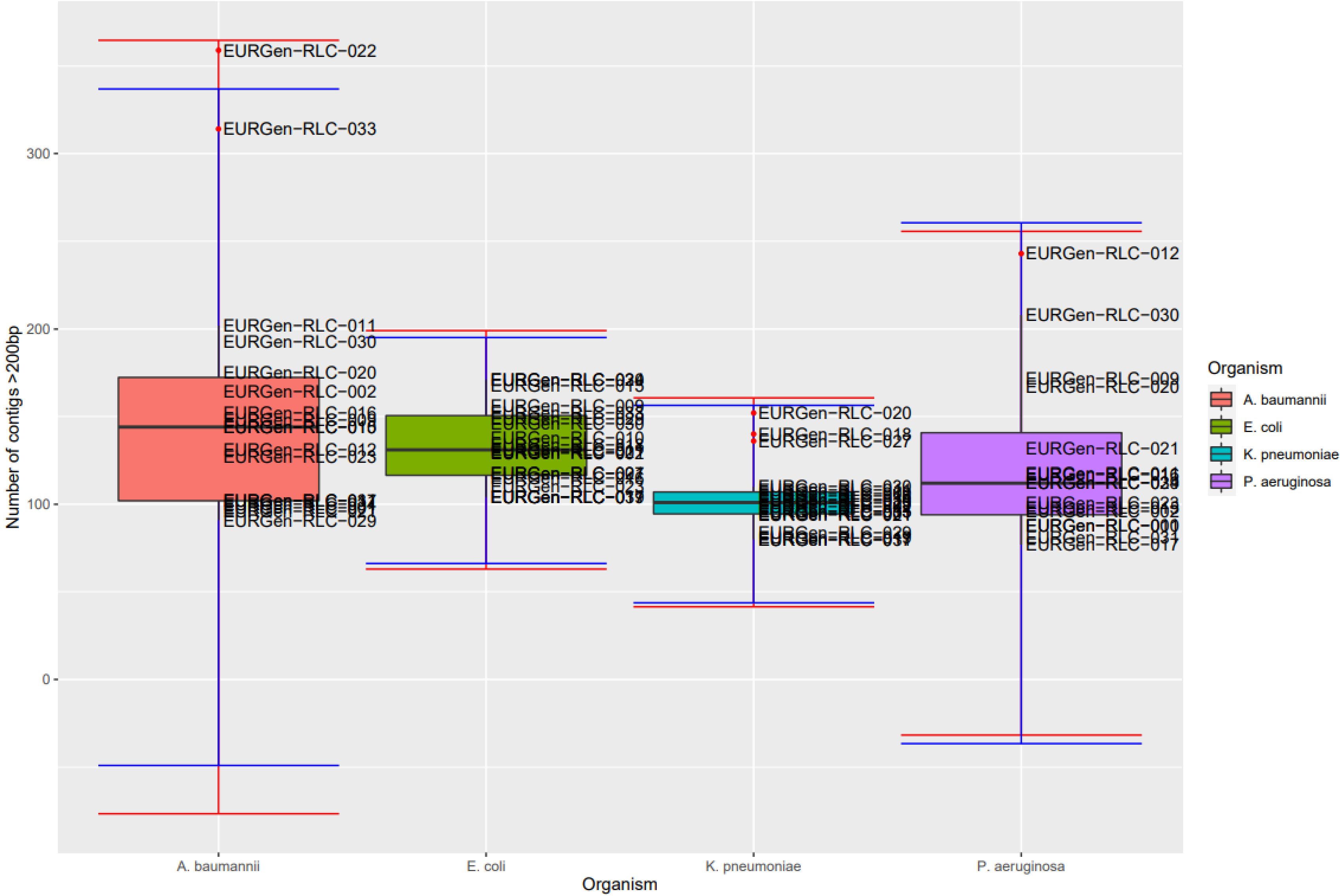
PROPORTION OF READS MAPPING TO REFERENCE

- Disregarding genomes removed in initial exclusion, all submission were above accepted threshold of 80%
- Most are above preferred threshold of 90%
- This metric is to some extent dependent on the sequencing platform



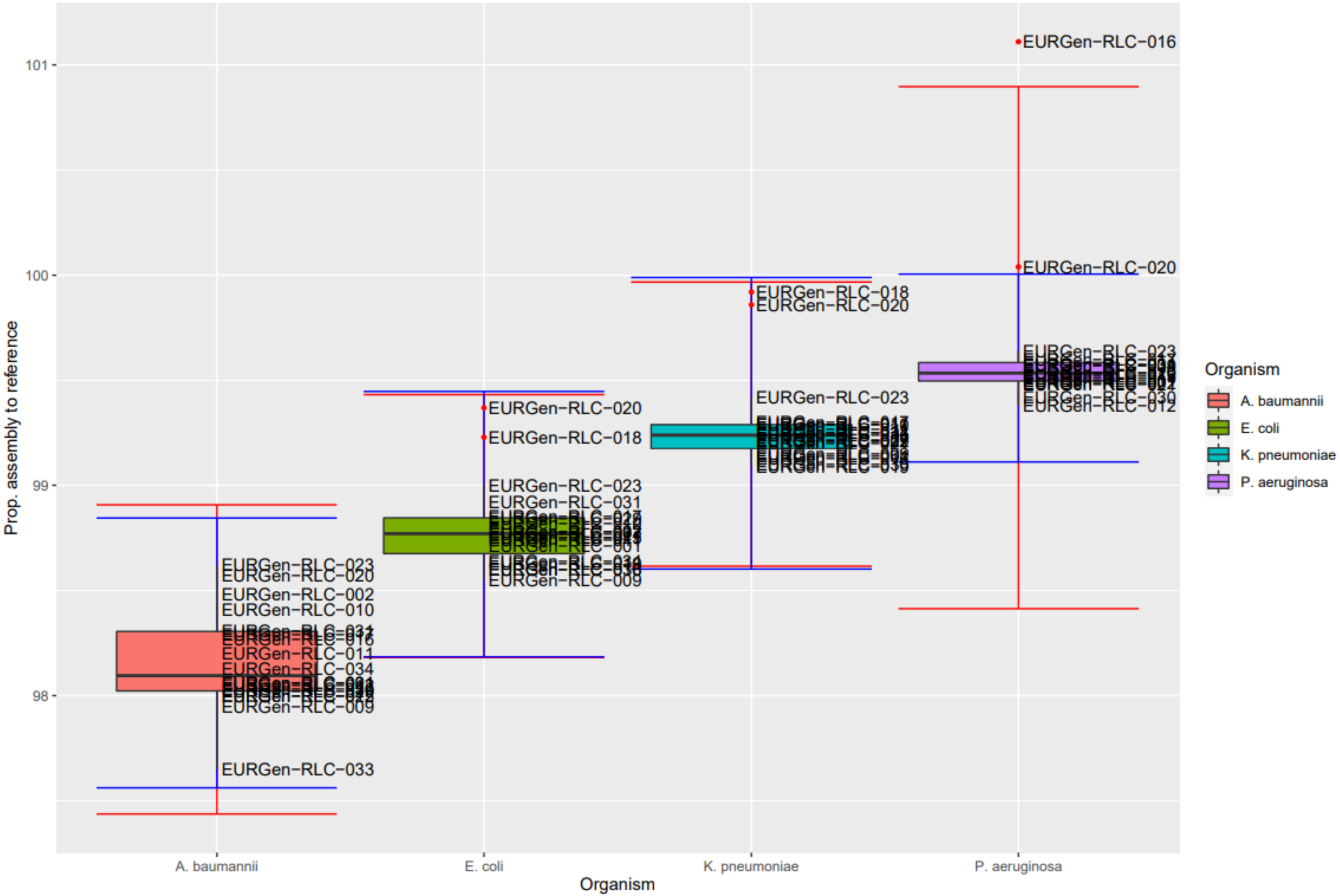
NUMBER OF CONTIGS >200 bp

- Close distribution with some variance in *P. aeruginosa*
- Red whiskers are outlier threshold
- Blue whiskers are adj. qual. thresholds
- A couple of genomes with high counts in *A. baumannii* one of which is outside the adjusted quality threshold



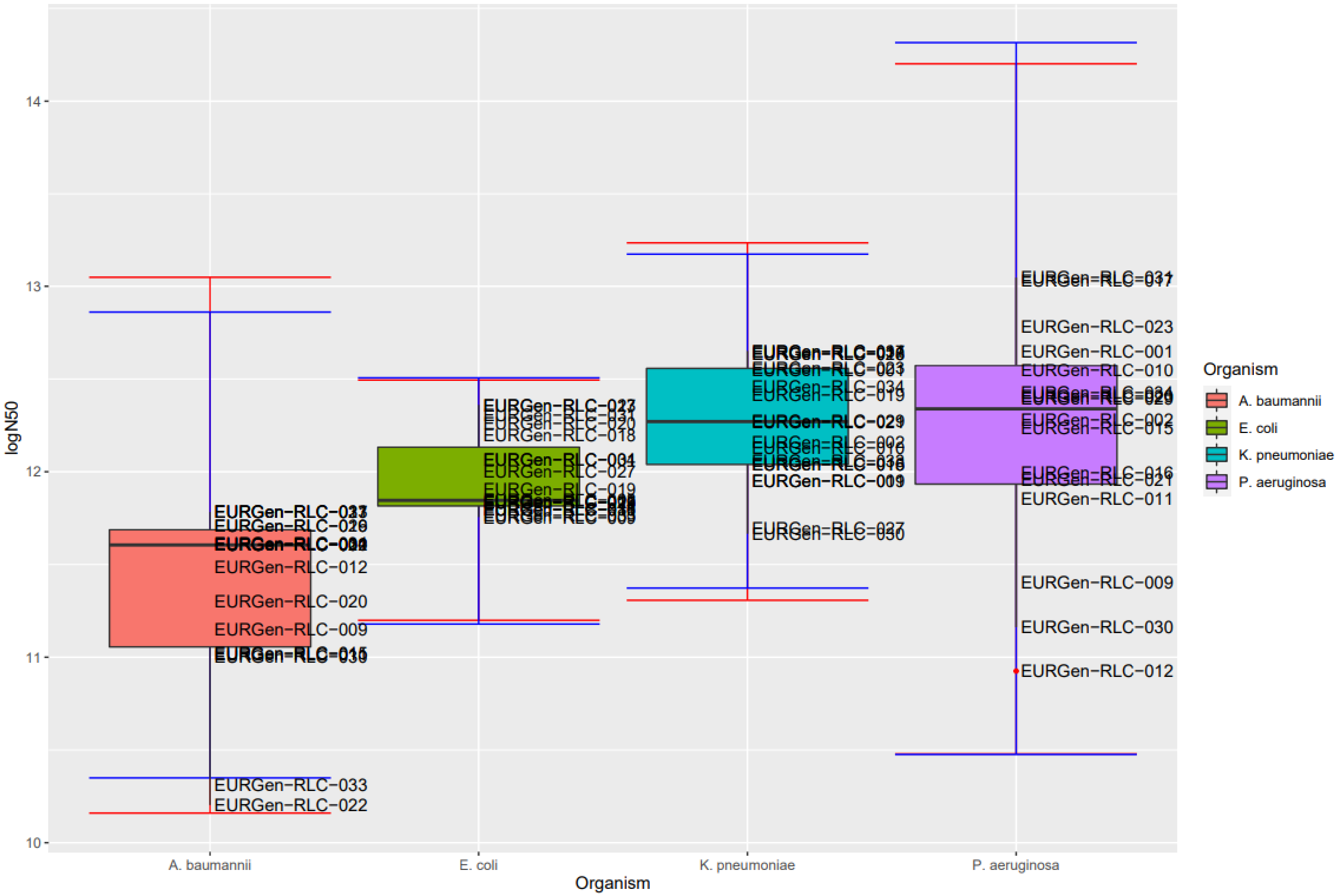
SIZE OF ASSEMBLY COMPARED TO REFERENCE

- Most assembly sizes are with expected thresholds
- Two submission for *P. aeruginosa* above general size of submissions, but notably only deviates 0.5-1.5%
- This metric is generally expected to be below 100% as this is not mapping, but the relative assembly size



N50

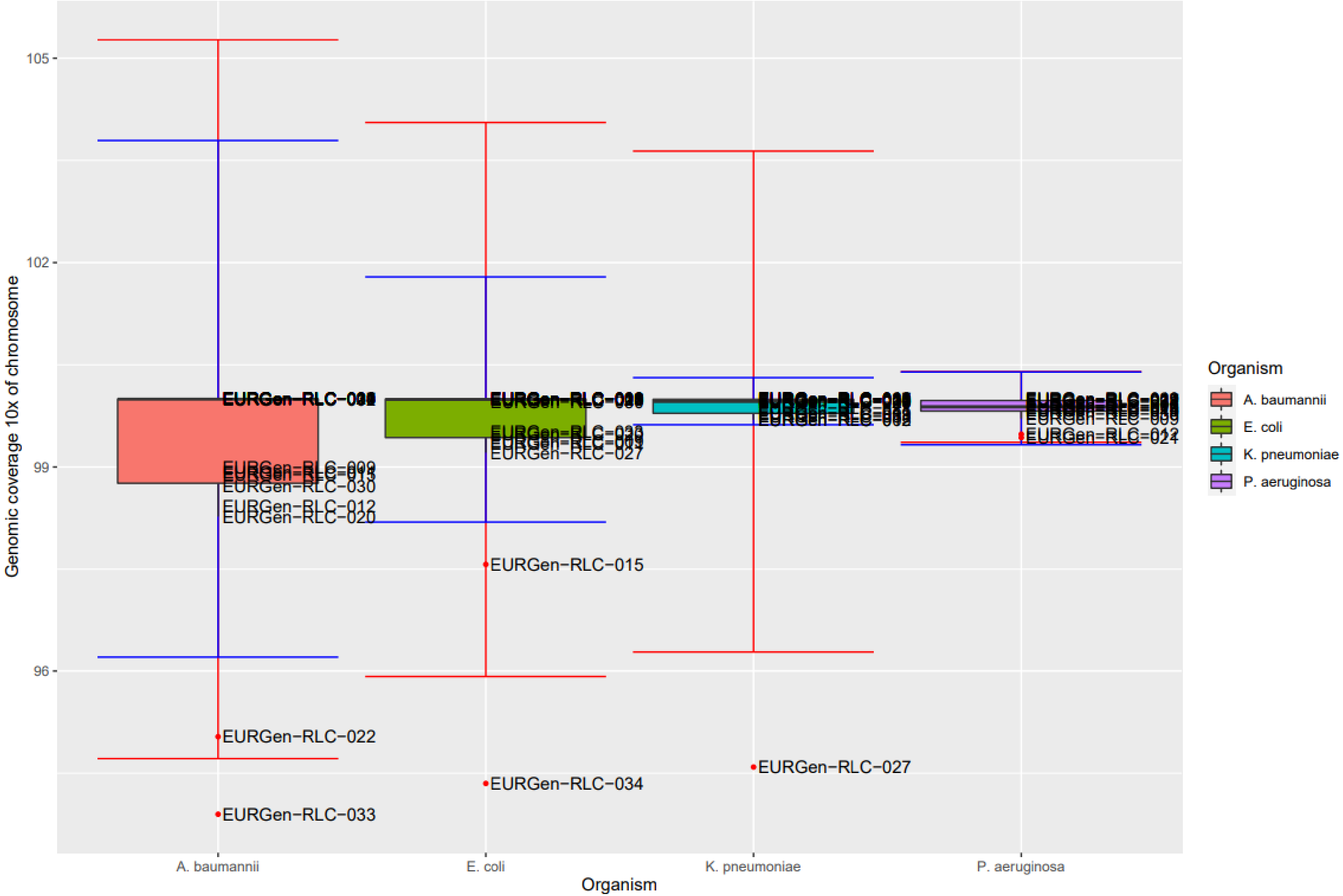
- Natural logarithmic transformation applied
- N50 generally shows comparable quality among participants
- In *A. baumannii,* two submission found outside adj. qual. threshold
- *P. aeruginosa* show a wide distribution comparable to other isolates



Funded by the European Union

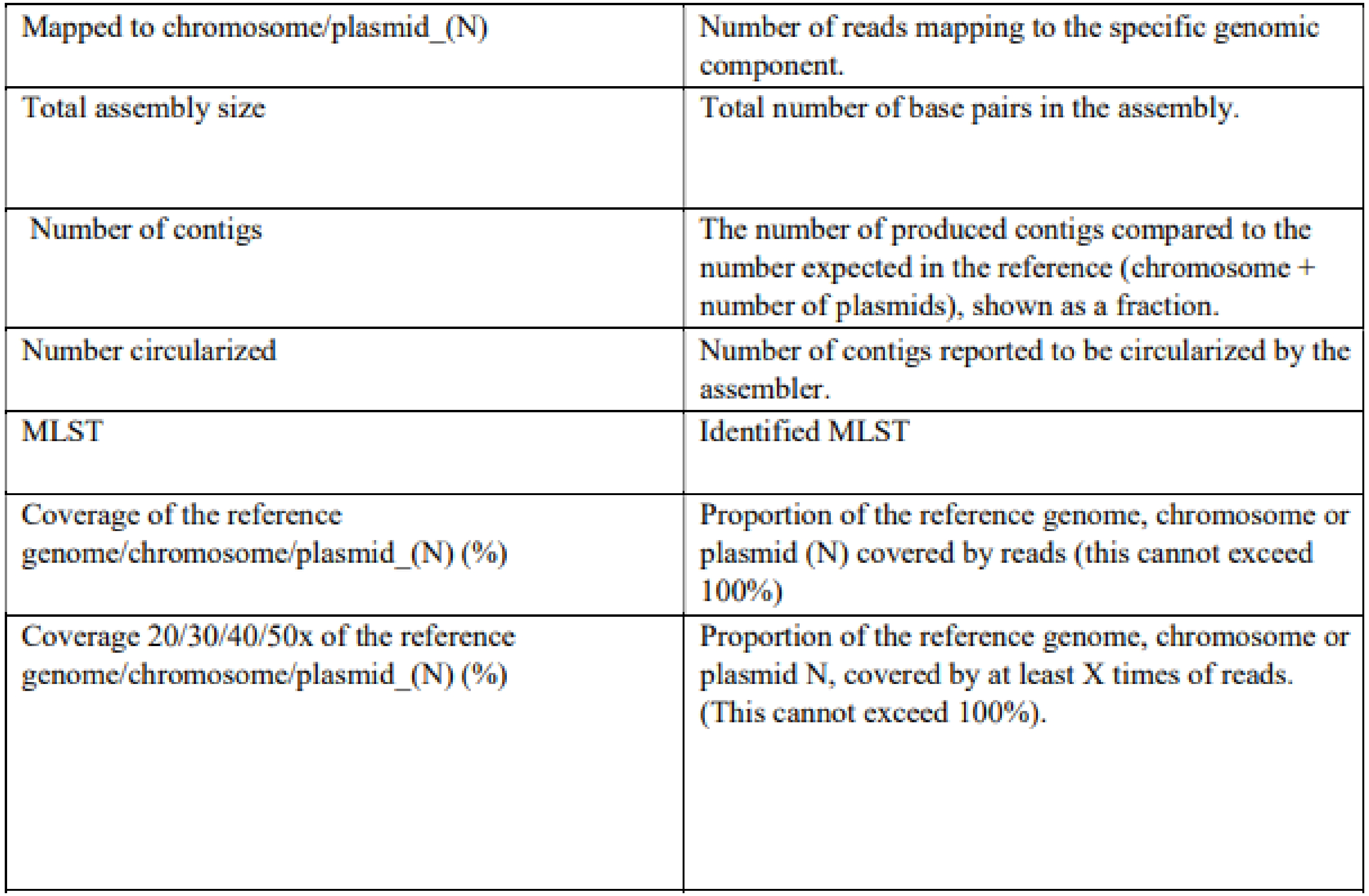
GENOMIC COVERAGE OF MINIMAL 10X DEPTH

- Most submitted genomes are above 99%, with notable exception of *A. baumannii*
- Three submission are outliers, two are below adj. qual. thresholds
- Having a high coverage of 10x is likely mainly influenced by the yield of the sequencing, but could also be due to bias in the workflow



NOTE ON OXFORD NANOPORE SEQUENCING

- Filtered on length and quality before assembly
- Evaluation more individualized as quality control of ONT is less well understood
- Issues identified:
  - Achieving high enough yield for recommend coverage of 30x
  - Coverage of plasmids



Mapped to chromosome/plasmid_(N)	Number of reads mapping to the specific genomic component.
Total assembly size	Total number of base pairs in the assembly.
Number of contigs	The number of produced contigs compared to the number expected in the reference (chromosome + number of plasmids), shown as a fraction.
Number circularized	Number of contigs reported to be circularized by the assembler.
MLST	Identified MLST
Coverage of the reference genome/chromosome/plasmid_(N) (%)	Proportion of the reference genome, chromosome or plasmid (N) covered by reads (this cannot exceed 100%)
Coverage 20/30/40/50x of the reference genome/chromosome/plasmid_(N) (%)	Proportion of the reference genome, chromosome or plasmid N, covered by at least X times of reads. (This cannot exceed 100%).

Subset of Table 1 from the ONT readme file

FEEDBACK SURVEY

- Questions
  - How useful was this EQA to your laboratory? (scale:1-10)
  - Was the preliminary individual EQA evaluation report you received clear and useful? (Yes/No)
  - Did you take any corrective action(s)? (Yes/No)
  - Comment section for suggestions

- So far only 3 respondents

- The survey is still open (deadline= Nov 30)

- Via this link (https://ec.europa.eu/eusurvey/runner/EURGen-RefLabCap_EQA_2023_feedback_survey_October_2023_3e60685e-87a1-35eb-3bc9-486d41fd38a6)

Thank you on behalf of the EURGen-RefLabCap team

EURGen-RefLabCap@food.dtu.dk

Funded by the European Union