Highlights from EQA2-WGS-AMR and RingTrial2-WGS-AMR

Presented by Egle Kudirkiene and Małgorzata Ligowska-Marzęta
on behalf of EQA team

Foodborne Infections (FBI)
Statens Serum Institut

Outline

- EQA2 vs Ringtrial2 differences

- EQA

  - DNA and sequencing QC

  - Methods used (tools and databases)

  - Reporting of genes and point mutations

  - QC, methods and reporting effect on the NRLs results

- RingTrial

  - Methods and their effect on results

  - Reporting of genes and point mutations

- Conclusions and the plans for the next round

The aim of the exercises

EQA2-WGS-AMR

<u>Material to analyse:</u>
Dried DNA samples
(WGS required)

To evaluate and ensure the quality and comparability of the WGS-based data produced by the NRLs

Comparison of performance of the NRLs in AMR gene and point mutation detection **based on sequences produced by each NRL**
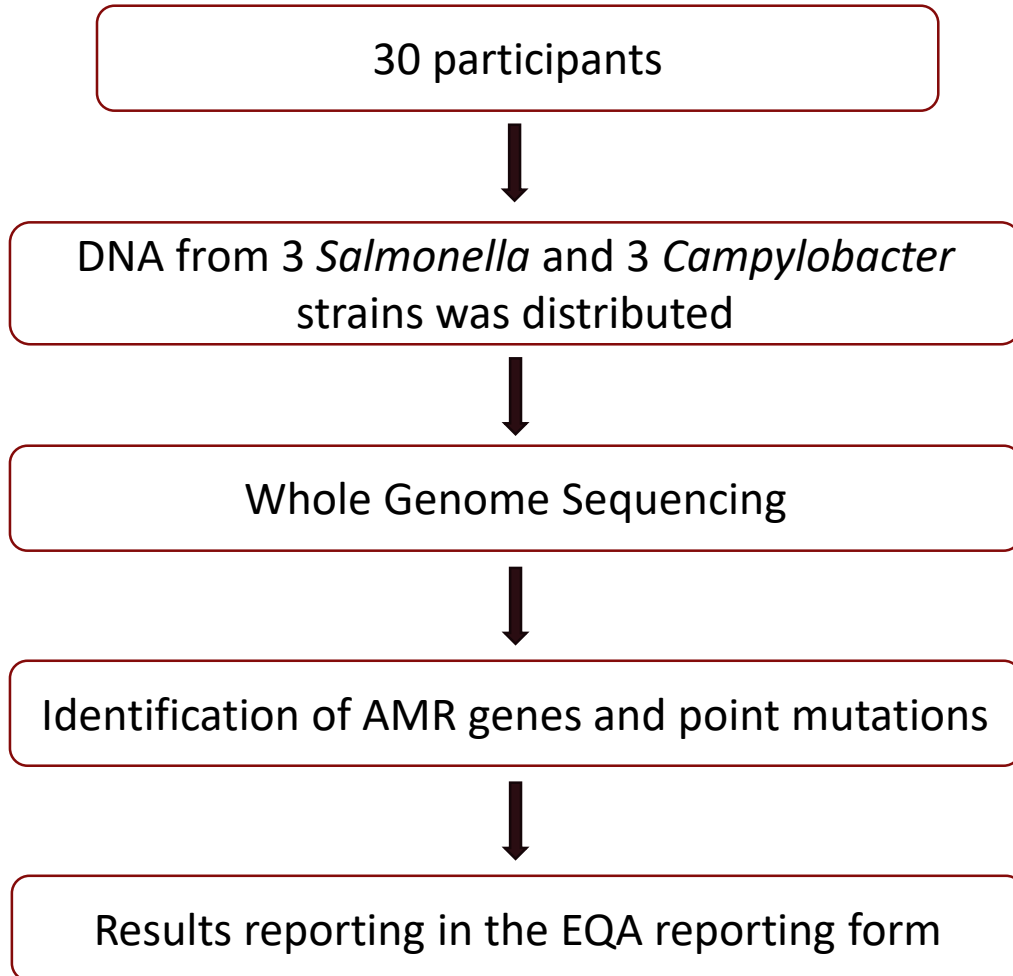
RingTrial2

<u>Material to analyse:</u>
Provided sequences
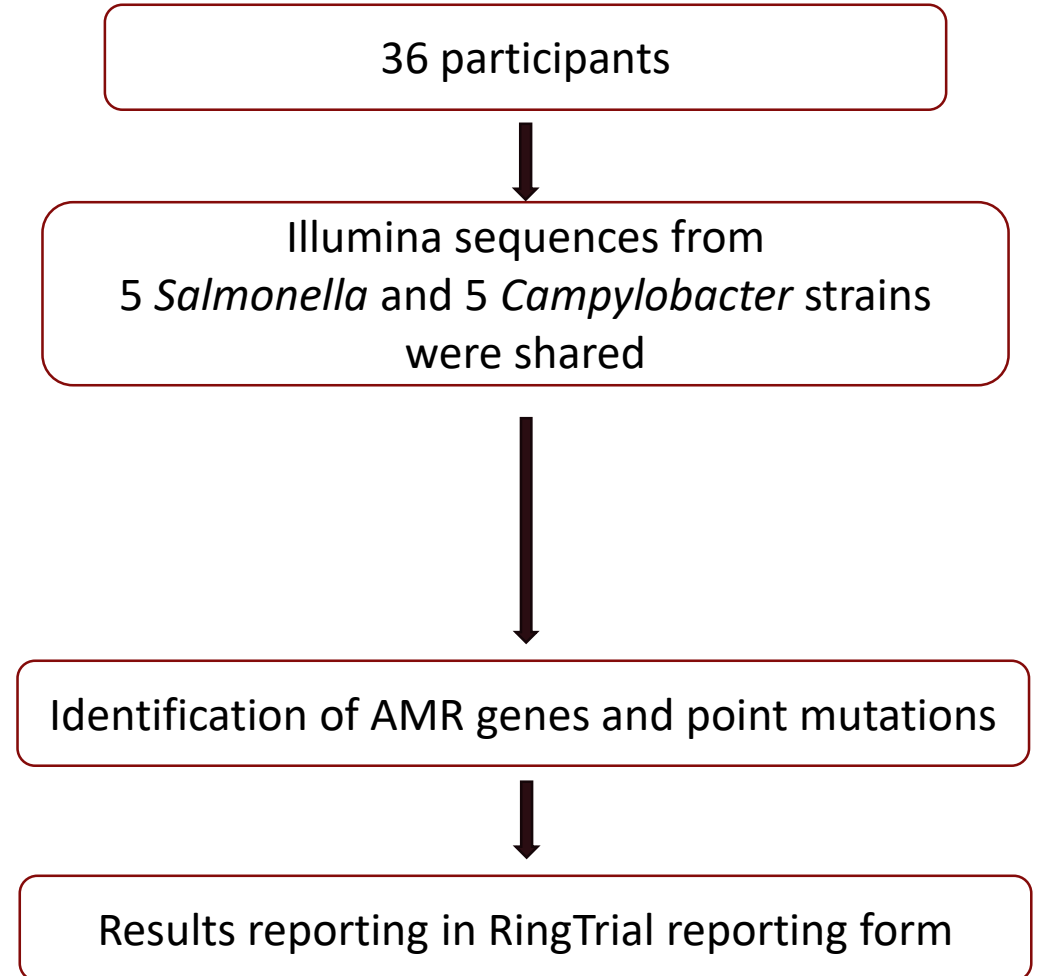(reads or assemblies)

Comparison of the outcomes of different databases, tools and bioinformatic pipelines used by NRLs

Comparison of performance of the NRLs in AMR gene and point mutation detection **based on provided sequences**

EQA: DNA and sequencing quality

DNA and sequencing QC evaluation parameters

DNA

- method
- concentration, ng/µl

Assembly

- method
- genome length, bp
- number of contigs (length trimming)
- N50, bp

Evaluated according to suggested thresholds for assembly QC:

- 4.4 Mb - 5.8 Mb (S), 1.5 Mb - 1.9 Mb (C)
- a N50 of > 30 000 bp
- contig number of <500 (S), <300 (C)

FWD AMR· RefLabCap

Sequencing QC – *Salmonella*   1/3

- **DNA concentration and method**
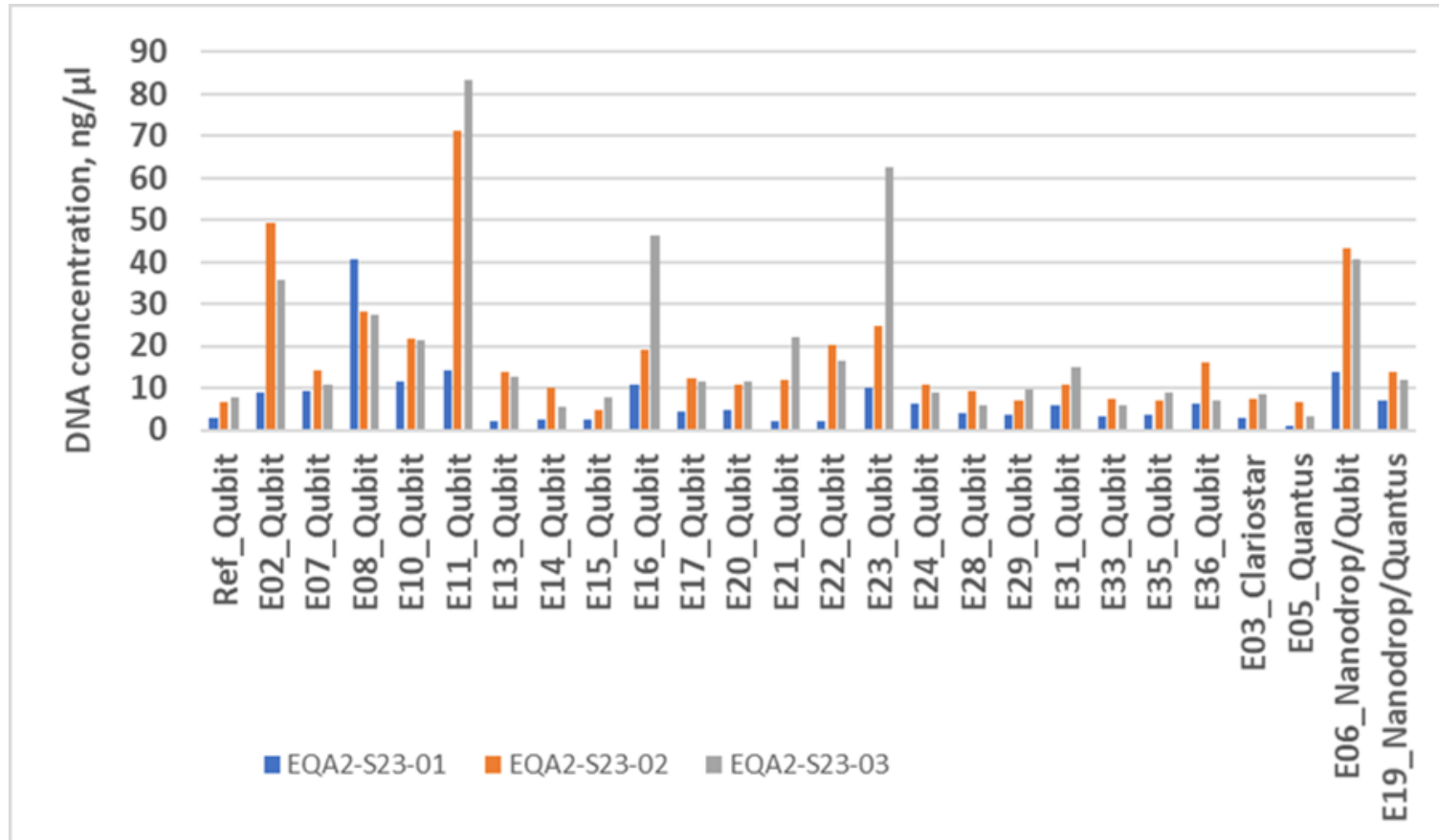
Figure 1. The distribution of Salmonella DNA concentrations among 25 participants with the indicated method(s) used.

EQA2-WGS-AMR

Sequencing QC – *Salmonella*   2/3

- ## Assembly

Figure 2. The distribution of number of contigs among 25 participants for three Salmonella samples. The horizontal axis labels indicate the ID of the participant, the tool used for genome assembly, and if contig length-based filtering was used it also indicates the filtering length. *Note: the figure does not include the results with the number of contigs of >500 (EQA2-S23-01: E05, E14, E15, E22, E35, and for EQA2-S23-02: E05)*

EQA2-WGS-AMR

Sequencing QC – *Salmonella*   3/3

QC failed – **not analysed further by the participants**

	Participant code	DNA		Assembly			
		Method	Concentration, ng/µl	Method*	Genome size, Mb	N50, bp	No. of contigs
EQA2-S23-01	E05	Quantus	1.1	SPAdes	1.02	454	2214
EQA2-S23-01	E15	Qubit	2.4	NR	2.70	351	7654
EQA2-S23-01	E23	Qubit	10.0	SPAdes	0.10	270	482

QC partly failed – **participants analysed and reported results**

Strain	Participant code	DNA		Assembly			
		Method	Concentration, ng/µl	Method*	Genome size, Mb	N50, bp	No.of contigs
EQA2-S23-01	E14	Qubit	2.6	SPAdes_200	4.70	1443	4687
EQA2-S23-01	E22	Qubit	2.0	SPAdes	5.03	6503	1813
EQA2-S23-01	E35	Qubit	3.7	SPAdes	4.62	27839	953
EQA2-S23-02	E05	Quantus	6.6	SPAdes	4.92	14321	728

EQA2-WGS-AMR

- **DNA concentration and method**
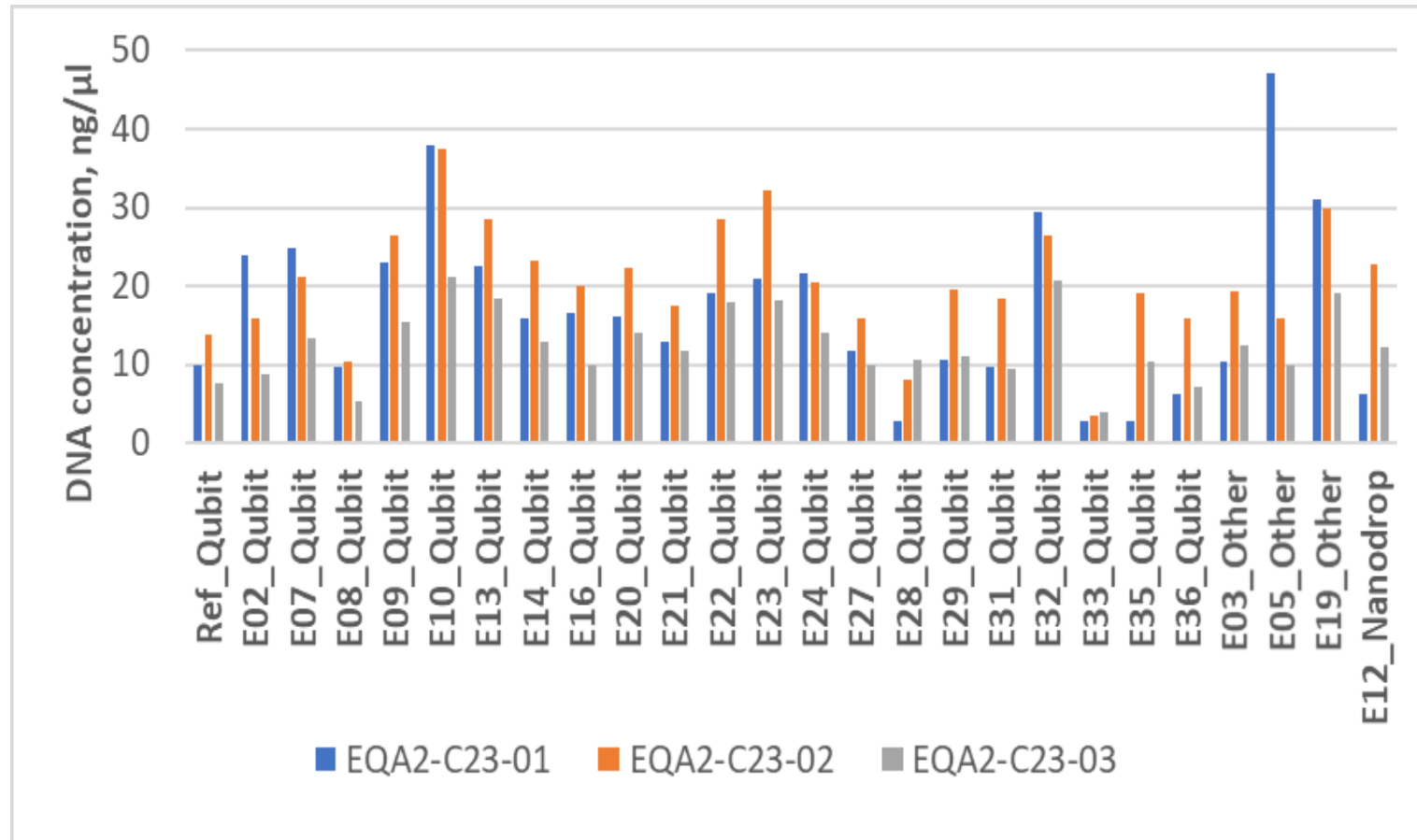
Figure 8. The distribution of Campylobacter DNA concentrations among 25 participants with the indicated methods used.
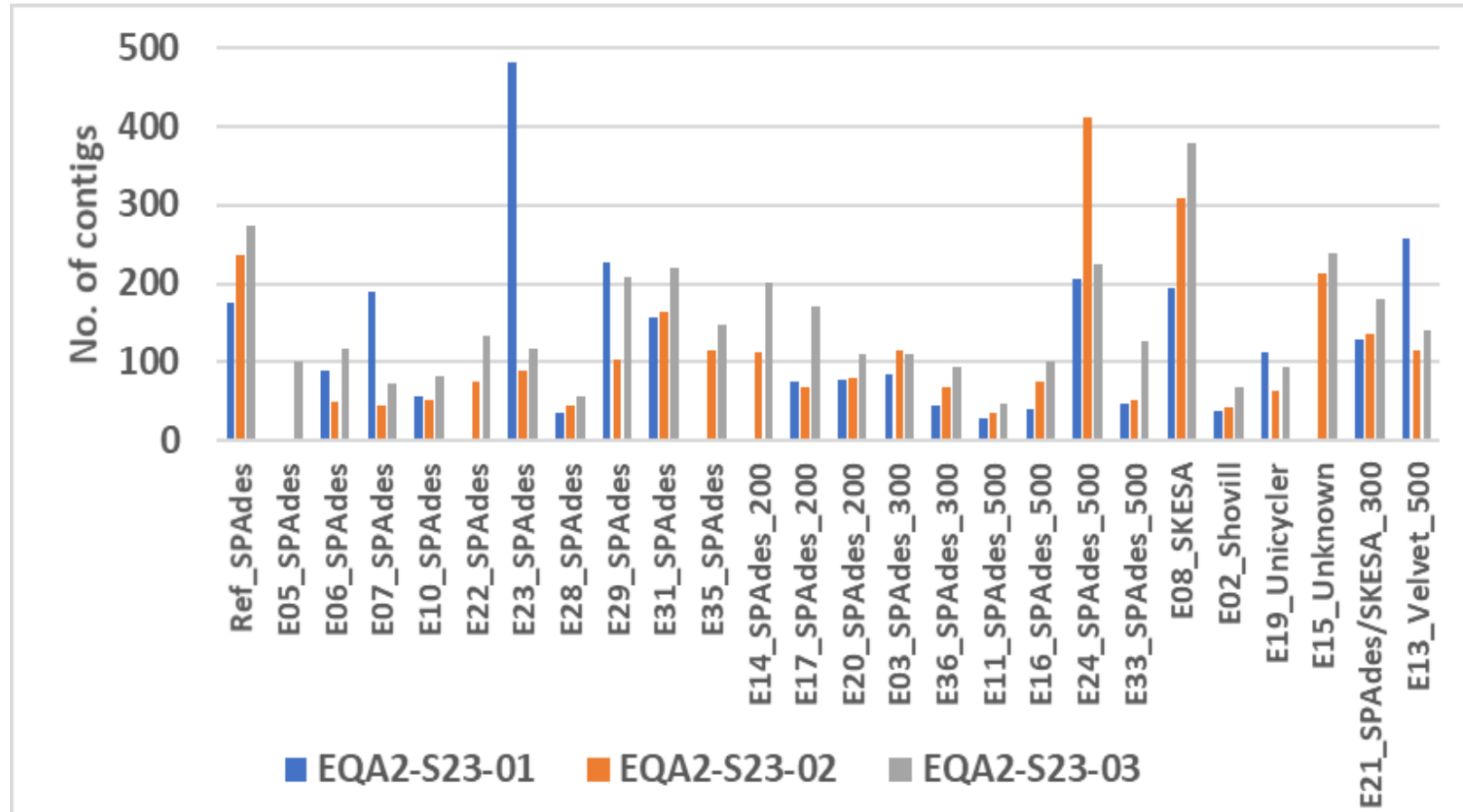
EQA2-WGS-AMR
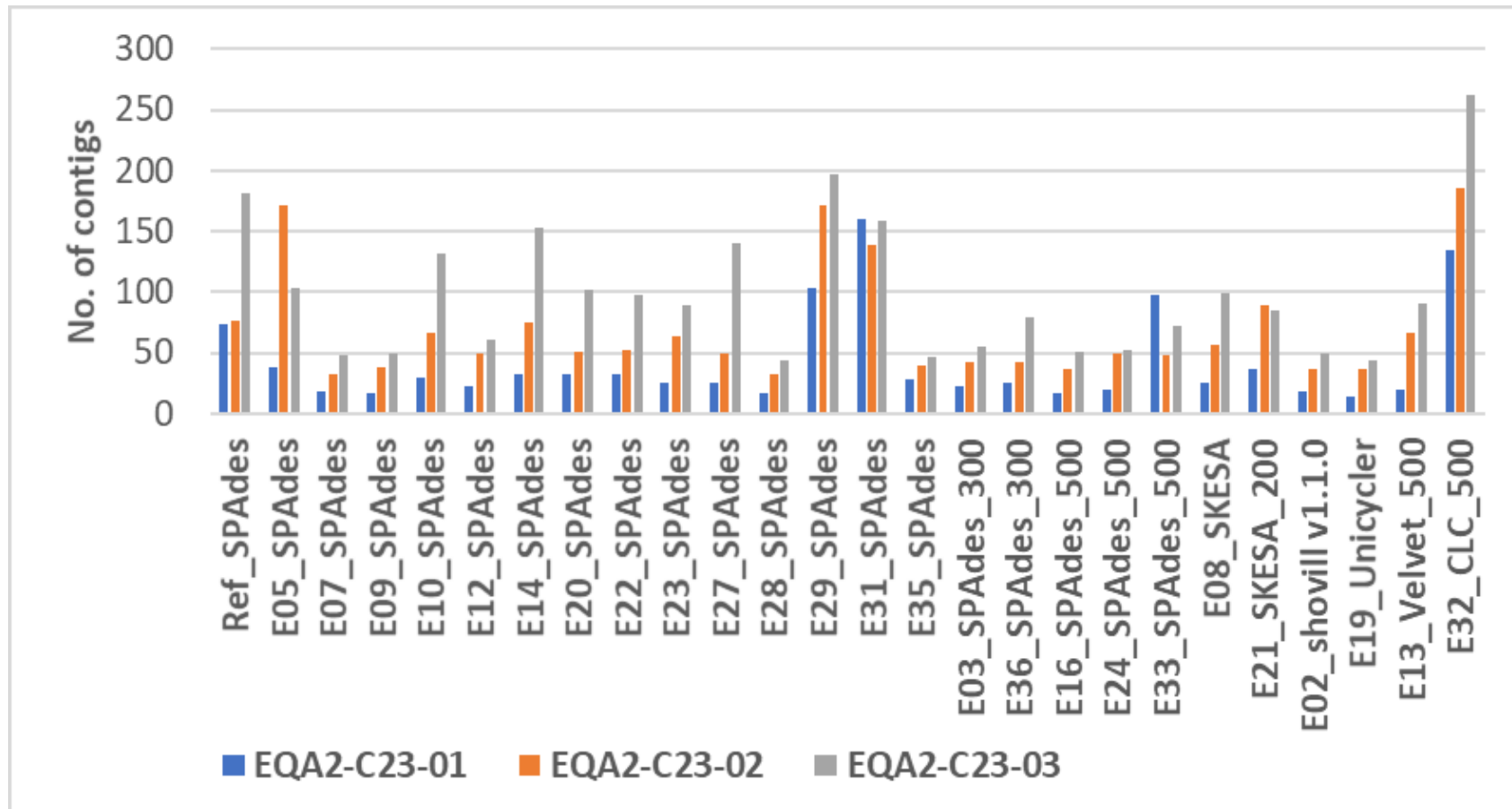
- **Assembly**

Figure 9. The distribution of number of contigs among 25 participants for three Campylobacter samples. The horizontal axis labels indicate the ID of the participant, the tool used for genome assembly, and if contig length-based filtering was used it also indicates the length used.

EQA2-WGS-AMR

EQA: Methods for gene and point mutation detection and reporting

Methods used for gene and point mutation detection and reporting

- In the survey we asked about:
  - ✓ Tools + input
  - ✓ Databases + Input
  - ✓ Thresholds for sequence length and identity
  - ✓ Result reporting when >1 database was used

Overview of reported tools, databases, inputs and reporting strategies

FWD AMR· RefLabCap

Supplementary Table 7. Detection and reporting of AMR genes in *Salmonella*

Unique combination	Tools/Inputs[A]	Databases/Inputs[A]	No. of participants	Partici-pants ID[B]	Identity (%)	Cove-rage (%)
	1 tool,1 input	1 database,1 input				
1	ResFinder_N	ResFinder_N	4	E03	99	100
2				E06	30	20
3				E29	90	60
3				E33	90	60
4	ResFinder_R	ResFinder_R	4	E15	90	60
5				E16	85	60
4				E20	90	60
6				E35	80	60
7	RGI_N	CARD_N	1	E08	perfect	perfect
8	AMRFinderPlus_N	AMRFinderPlus_N	1	E24	97	97
9	AbriTAMR 1.0.13	AbriTAMR 1.0.13	1	E11	default	default
	1 tool, >1 input	1 database, >1 input				
10	ResFinder_N_R	ResFinder_N_R	2	E07	90	60
10				E31	90	60
11	ResFinder_N_P_R	ResFinder_N_P_R	1	E05	90	60
	2 tools, 1 input	1 database, >1 input				
12	BlastN/ResFinder_R	ResFinder_N_R	1	E28	90	60
	2 tools >1 approach	2 databases				
13	ResFinder_N_R/RGI_N	ResFinder_N_R/CARD_N	1	E13[C]	90	60
	2 tools, 1 input	2 databases, 1 input				
14	AMRFinderPlus_N/ResFinder_N	AMRFinderPlus_N/ResFinder_N	2	E10	98	60
15				E36	90	60
16	AMRFinderPlus_N/ResFinder_R	AMRFinderPlus_N/ResFinder_R	1	E19	90	60
17	AMRFinderPlus_R/ResFinder_R	AMRFinderPlus_R/ResFinder_R	1	E14	>90	
	2 tools, >1 input	2 databases, >1 input				
18	AMRFinderPlus_N/ResFinder_N_R	AMRFinderPlus_N/ResFinder_N_R	1	E22	90	40
	3 tools, 1 input	3 databases, >1 input				
19	AMRFinderPlus_N/ResFinder_N/RGI_N	AMRFinderPlus_N/ResFinder_N/CARD_N	1	E17	90	60
20	AMRFinderPlus_N/ResFinder_R/RGI_N	AMRFinderPlus_N/ResFinder_R/CARD_N	1	E23	at least 99[D]	at least 99[D]
	3 tools, >1 input	3 databases, 1 input				
21	AMRFinderPlus_N/ResFinder_N_R/RGI_N	AMRFinderPlus_N/ResFinder_N_R/CARD_N	1	E02[E]	default	90
	4 tools, 1 input	2 databases, 1 input				
22	ARIBA_R/Abricate_R/AMRFinderPlus_N/BN Plugin	AMRFinderPlus_N/ResFinder_R	1	E21	90/85/90[F]	90/85/90[F]

Inputs:
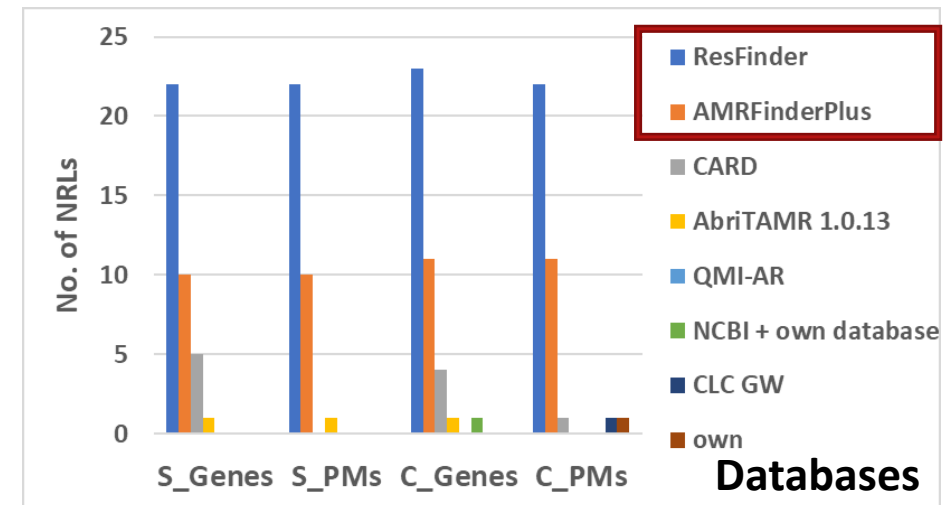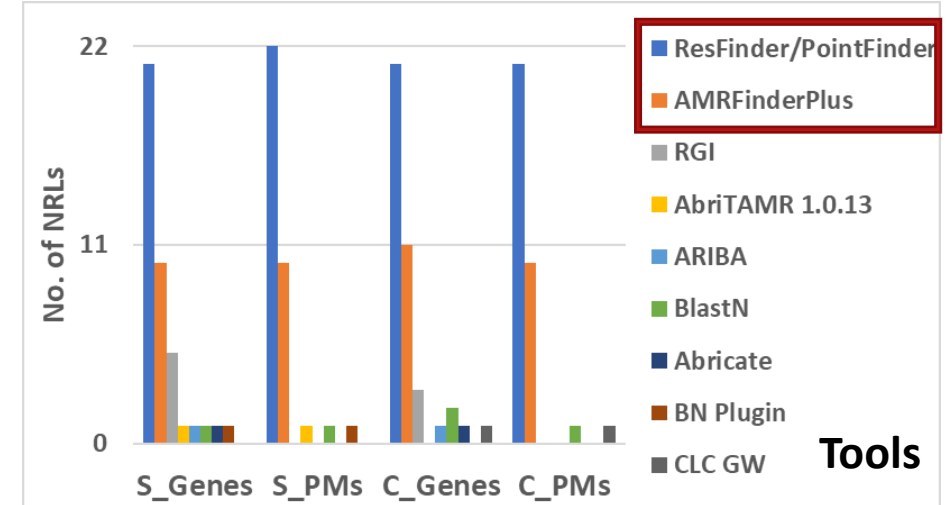**N** - DNA fasta,
P - protein fasta,
R - raw reads.
>1 input, if different inputs were used for at least one of the tools/databases

Reporting strategy:
light yellow - reported all genes from all databases,
light red - reported a subset of genes based on experience/knowledge/literature,
light green - reported a consensus list of genes (common genes present in all databases used)

Overview of the tools, databases and inputs used by 25 NRLs

Combinations of Tools, Inputs and databases used	*Salmonella*		*Campylobacter*	
	Genes	PMs	Genes	PMs
1 tool, 1 input/1 database	11	12	9	11
1 tool, >1 input/1 database	3	3	4	5
1 tool,1 input/4 databases			1	1
2 tools, 1 input/1 database	1	2		
2 tools, 1 input/2 databases	4	6	3	3
2 tools, >1 input/2 databases	2	2	4	5
3 tools, 1 input/3 databases	2		1	
3 tools, >1 input/2 databases			2	
3 tools, >1 input/3 databases	1		1	
4 tools, 1 input/2 databases	1			
No. of unique combinations	22	14	21	13
No. of shared unique combinations	3	6	2	6

Tools
- ResFinder/PointFinder
- AMRFinderPlus
- RGI
- AbriTAMR 1.0.13
- ARIBA
- BlastN
- Abricate
- BN Plugin
- CLC GW

Databases
- ResFinder
- AMRFinderPlus
- CARD
- AbriTAMR 1.0.13
- QMI-AR
- NCBI + own database
- CLC GW
- own

EQA2-WGS-AMR

Reporting of the input

Left table

Unique combination	Tools/Inputs[A]	Databases/Inputs[A]	No. of participants	Participants ID[B]	Identity (%)	Coverage (%)
	1 tool,1 input	1 database,1 input				
1	ResFinder_N	ResFinder_N	4	E03	99	100
2				E06	30	20
3				E29	90	60
3				E33	90	60
4	ResFinder_R	ResFinder_R	4	E15	90	60
5				E16	85	60
4				E20	90	60
6				E35	80	60
7	RGI_N	CARD_N	1	E08	perfect	perfect
8	AMRFinderPlus_N	AMRFinderPlus_N	1	E24	97	97
9	AbriTAMR 1.0.13	AbriTAMR 1.0.13	1	E11	default	default
	1 tool, >1 input	1 database, >1 input				
10	ResFinder_N_R	ResFinder_N_R	2	E07	90	60
10				E31	90	60
11	ResFinder_N [P] R	ResFinder_N [P] R	1	E05	90	60
	2 tools, 1 input	1 database, >1 input				
12	BlastN/ResFinder_R	ResFinder_N_R	1	E28	90	60
	2 tools >1 approach	2 databases				
13	ResFinder_N_R/ RGI_N	ResFinder_N_R/ CARD_N	1	E13[C]	90	60
	2 tools, 1 input	2 databases, 1 input				
14	AMRFinderPlus_N/ ResFinder_N	AMRFinderPlus_N/ ResFinder_N	2	E10	98	60
15				E36	90	60
16	AMRFinderPlus_N/ ResFinder_R	AMRFinderPlus_N/ ResFinder_R	1	E19	90	60
17	AMRFinderPlus_R/ ResFinder_R	AMRFinderPlus_R/ ResFinder_R	1	E14	>90	
	2 tools, >1 input	2 databases, >1 input				
18	AMRFinderPlus_N/ ResFinder_N_R	AMRFinderPlus_N/ ResFinder_N_R	1	E22	90	40
	3 tools, 1 input	3 databases, >1 input				
19	AMRFinderPlus_N/ ResFinder_N/RGI_N	AMRFinderPlus_N/ ResFinder_N/CARD_N	1	E17	90	60
20	AMRFinderPlus_N/ ResFinder_R/RGI_N	AMRFinderPlus_N/ ResFinder_R/CARD_N	1	E23	at least 99[D]	at least 99[D]
	3 tools, >1 input	3 databases, 1 input				
21	AMRFinderPlus_N/ ResFinder_N_R/RGI_N	AMRFinderPlus_N/ ResFinder_N_R/CARD_N	1	E02[E]	default	90
	4 tools, 1 input	2 databases, 1 input				
22	ARIBA_R/Abricate_R/ N/B	AMRFinderPlus_N/ ResFinder_R	1	E21	90/85/90[F]	90/85/90[F]

EQA2-WGS-AMR

Right table

Unique combinations	Tools/Inputs[A]	Databases/Inputs[A]	No. of participants	Participants ID[B]
	1 tool, 1 input	1 database, 1 input		
1	PointFinder_N	ResFinder_N	6	E03
1				E06
1				E08
1				E29
1				E33
2	PointFinder_R	ResFinder_R	4	E15
2				E16
2				E20
2				E35
3	AbriTAMR 1.0.13	AbriTAMR 1.0.13	1	E11
4	AMRFinderPlus_N	AMRFinderPlus_N	1	E24
	1 tool, >1 inputs	1 database, >1 input		
5	PointFinder_N_R	ResFinder_N_R	2	E13
5				E31
6	PointFinder_N [P] R	ResFinder_N [P] R	1	E05
	1 tool, 1 input	1 database, >1 input		
7	PointFinder_N	ResFinder_N_R	1	E07
	2 tools, 1 input	1 database, 1 input		
8	AMRFinderPlus_N/ BN Plugin[C]	AMRFinderPlus_N	1	E21
	2 tools, 1 input	1 database, >1 input		
9	PointFinder_R/BLAST_N	ResFinder_N_R	1	E28
	2 tools, 1 input	2 databases, 1 input		
10	PointFinder_N/ AMRFinderPlus_N	ResFinder_N/ AMRFinderPlus_N	3	E10
11				E17
10				E36
12	PointFinder_R/ AMRFinderPlus [R]	ResFinder_R/ AMRFinderPlus [R]	1	E14
13	PointFinder_R/ AMRFinderPlus_N	ResFinder_R/ AMRFinderPlus_N	2	E19
13				E23
	2 tools, >1 input	2 databases, >1 input		
14	PointFinder_N_R/ AMRFinderPlus_N	ResFinder_N_R/ AMRFinderPlus_N	2	E02
14				E22

Errors in reporting

How to report tools, databases and inputs?

Karen Loaiza Conza

- Certain inputs are used for gene and PM detection depending on the tool in use

- Other tools/softwares may use ResFinder, AMRFinder and RGI databases but only with one possible input

ResFinder					
genes			**point mutations**		
input	**method**	**db**	**input**	**method**	**db**
DNA fastq	KMA	nucleotide	DNA fastq	KMA	nucleotide
DNA fasta	BLASTN		DNA fasta	BLASTN	

AMRFinderPlus					
genes			**point mutations**		
input	**method**	**db**	**input**	**method**	**db**
DNA fasta	BLASTX	protein	DNA fasta	BLASTN	nucleotide
Protein fasta	BLASTP		Protein fasta	-	-

RGI					
genes			**point mutations**		
input	**method**	**db**	**input**	**method**	**db**
DNA fasta	PRODIGAL + BLAST	protein	DNA fasta	BLASTN	nucleotide
Protein fasta	BLASTP		Protein fasta	BLASTP	protein

Abricate						Ariba					
genes			**point mutations**			**genes**			**point mutations**		
input	**method**	**db**	**input**	**method**	**db**	**input**	**method**	**db**	**input**	**method**	**db**
-	-	-	-	-	-	DNA reads	Bowtie2	nucleotide	-	-	-
DNA fasta	BLASTN	nucleotide	-	-	-	-	-	-	-	-	-

Reporting genes and point mutations

Salmonella and *Campylobacter* DNA samples characteristics

Table 1. Genotypic and phenotypic characteristics of the Salmonella strains selected for the EQA2-WGS-AMR

Strain	EQA2-S23-01	EQA2-S23-02	EQA2-S23-03
Serotype	Dublin	Stanley	Rissen
ST	10	29	469
Genes[A]	blaTEM-1, emrD, mdsA, mdsB, sul2, tetA	aac(3)-IId, aadA1, aadA2, aph(3')-Ia, aph(3'')-Ib, aph(6)-Id, blaTEM-1, catA2, dfrA12, emrD, floR, mdsA, mdsB, mphA, qnrS1, sul1, sul3, tetM	aac(3)-IId, aph(3'')-Ib, aph(6)-Id, blaCTX-M-55, emrD, floR, mdsA, mdsB, qnrS1, sul2, tetA
PMs[A]	ramR T18P, acrB R717L	None	gyrA D87N
NWT Phenotypes[B]	AMP, AZI, COL, SME, TET	AMP, AZI, CHL, CIP, GEN, SME, TRI	AMP, CEP, CTA, CTZ, CHL, CIP, GEN, NAL, SME, TEM, TET

A According to AMRFinderPlus
B Abbreviations of antimicrobials: AMP (Ampicillin), AZI (Azithromycin), CEP (Cefepime), Cefotaxime (CTA), Ceftazidime (CTZ), CHL (Chloramphenicol), CIP (Ciprofloxacin), COL (Colistin), GEN (Gentamicin), NAL (Nalidixic acid), SME (Sulfamethoxazole), TEM (Temocilin), TET (Tetracycline), TRI (Trimethoprim)

Table 2. Genotypic and phenotypic characteristics of the Campylobacter strains selected for the EQA2-WGS-AMR

Strain	EQA2-C23-01	EQA2-C23-02	EQA2-C23-03
Species	C. coli	C. coli	C. coli
ST	888	1586	872
Genes[A]	aadE-Cc, tet(O)	aac(6')-Ie/aph(2'')-Ia, aad9, aadE, aph(2'')-If, aph(3')-IIIa, blaOXA-193, catA13, ermB, tet(O/M/O)	aadE-Cc, blaOXA-489, tetO
PMs[A]	gyrA T86I, 50S L22 A103V	gyrA T86I	23S A2075G, gyrA T86I
NWT Phenotypes[B]	CIP, ERY, GEN, TET	CIP, ERY	CIP, ERY, TET

A According to AMRFinderPlus
B Abbreviations of antimicrobials: CIP (Ciprofloxacin), ERY (Erythromycin), GEN (Gentamicin), TET (Tetracycline)

EQA2-WGS-AMR

19

Gene and point mutation reporting

* denotes extra tools used by participants, in this case laboratory E21 used BioNumerics (separate database)

Concordance defined as number of laboratories that reported the same genes or point mutations for a given DNA sample (expressed in percentage) – taking into consideration different nomenclature

Common differences between reference datasets

Table 4. Tools and databases used in provider's reference datasets, Res_Ref and AMR_Ref, for Salmonella and Campylobacter

Reference dataset		Res_Ref	AMR_Ref
AMR gene detection	Database	ResFinder	AMRFinderPlus
	Tool	ResFinder (CGE server)	AMRFinderPlus
	Input	Raw reads (fastq)	SPAdes assembly (fasta protein)
	Cutoffs	90% identity, 60% coverage	90% identity, 50% coverage
Point mutation identification	Database	ResFinder	AMRFinderPlus
	Tool	PointFinder (CGE server)	AMRFinderPlus
	Input	Raw reads (fastq)	SPAdes assembly (fasta nucleotide)

Salmonella

Res_Ref	AMR_Ref
blaTEM-1B	blaTEM-1
aac(6')-Iaa	-
-	Efflux genes: emrD, mdsA, mdsB
aadA2 / ant(3'')-Ia	aadA2

Campylobacter

Res_Ref	AMR_Ref
-	50S L22 A103V
cat	catA13
aac(6')-aph(2'')	aac(6')-Ie/aph(2'')-Ia
ant(6)-Ia	aadE

Salmonella examples: EQA2-S23-02 - thresholds

* Laboratory E21 used also BioNumerics plugin which has its own database

Thresholds in the provider reference values:

Gene	Pct identity	Coverage	Tool / database
catA2	96%	100%	ResFinder
	98%	100%	RGI (CARD)
floR	93%	94%	ResFinder
	99%	100%	RGI (CARD)

Thresholds in some reported values:

Gene	Pct identity	Coverage	Tool / database
E03	99%	100%	ResFinder
E08	100%	100%	RGI (CARD)

"Perfect" algorithm was applied

Campylobacter example: EQA2-C23-02 - nomenclature

aac(6')-aph(2'') in ResFinder
(option not given in the survey)

aadE and **ant(6)-Ia** are synonyms

cat and **catA13**

Sporadic reporting of *tet* genes
NO phenotypic resistance to tetracycline

	Res_Ref	E31	E08	E33	E35	E09	E05	E03	E16	E07	E20	E28	E29	AMR_Ref	E24	E27	E21	E10	E36	E22	E14	E19	E12	E13	E02	E23	E32	% concordance
							ResFinder							AMR_Ref		AMRFinder +/- ResFinder							CARD +/- others					
ResFinder						*																						
AMRFinderPlus																												
CARD																												
Other database																							**				***	
aac(6')-Ie/aph(2'')-Ia			X		X	X	X			X		X	X	X	X	X	X	X	X	X	X	X	X	X	X		
aac(6')-aph(2'')	X	X		X				X	X	X		X															96
aad9													X	X	X	X	X	X	X		X		X		32
aadE													X	X	X	X		X		X		X		80
ant(6)-Ia	X	X	X	X	X		X	X	X		X	X			X	X	X								80
aph(2'')-If	X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	92
aph(3')-III	X	X			X		X	X	X	X		X													
aph(3')-IIIa		X	X		X				X		X	X	X	X	X	X	X	X	X	X	X	X	X	100
blaOXA-193	X	X			X	X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	72
cat	X	X	X	X	X	X	X	X	X	X	X	X							X				X	
catA13													X	X	X	X	X	X			X	X	100
erm(B)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
tet(O)		X				X					X			X	X	X		X		X	
tet(O/M/O)													X		X		X	X			X	48

* BioNumerics 8.1
** NCBI + an home-made database
*** QMI-AR Peptide Marker Database (2021-08).

AMRFinderPlus Ref: 99% identity and 68% coverage for *tet(O/M/O)*
ResFinder Ref: No *tet* gene detected
CARD: 99.5% identity, 68% coverage for *tet(O/M/O)*

EQA2-WGS-AMR

QC, methods and reporting effect on the NRLs results

Additional comparison...not in the report

- **Between the participants** - when <u>same unique combination</u> of tools, databases, inputs, thresholds and reporting strategy is in use

- **With EQA provider** results – when <u>low sequence QC</u> was reported

EQA2-WGS-AMR

Salmonella genes – same unique combination, different results

FWD AMR· RefLabCap

EQA2-S23-02

Unique combination	Tools/Inputs[A]	Databases/Inputs[A]	No. of participants	Partici-pants ID[B]	Identity (%)	Cove-rage (%)
	1 tool,1 input	1 database,1 input				
1	ResFinder_N	ResFinder_N	4	E03	99	100
2				E06	30	20
3				E29	90	60
3				E33	90	60
4	ResFinder_R	ResFinder_R	4	E15	90	60
5				E16	85	60
4				E20	90	60
6				E35	80	60
7	RGI_N	CARD_N	1	E08	perfect	perfect
8	AMRFinderPlus_N	AMRFinderPlus_N	1	E24	97	97

The QC was good for both NRLs

Different reporting? Different versions of tools/databases?

Reporting errors?

- *aac(6')-Iaa* – cryptic

- *catA2* - ?

- *aadA1* - ?

EQA2-WGS-AMR

Salmonella – low QC, different unique combinations

EQA2-S23-01 - genes

	ResFinder											AMRFinderPlus +/- ResFinder									CARD +/- other				% concordance
	Res_Ref	E31	E06	E33	E35	E03	E16	E07	E20	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02	
ResFinder															*										
AMRFinderPlus															*										
CARD																									
aac(6')-Iaa	X	X			X		X	X	X	X	X								⊗	X		X	X	X	55
blaTEM-1													X	X	X			X			X			X	
blaTEM-1B	X	X	X	X	X	X	X	X	X	X	X					X	X	X	X	X		X	X	X	100
emrD													X												
mdsA													X		X										5
mdsB													X		X										5
sul2	X	X	X	X	X	X	X	X	X	X	X					X	X	X	X	X		X	X	X	95
tet(A)	X	X		X	X	X	X	X	X	X	X					X	X	X	X	X		X	X	X	95

EQA2-S23-01 – point mutations

	PointFinder											AMRFinderPlus +/- PointFinder									% concordance
	Res_Ref	E31	E06	E08	E33	E13	E35	E03	E16	E07	E20	E28	E29	AMR_Ref	E11	E21	E24	E17	E10	E02	E36	E22	E14	E19	
PointFinder																									
AMRFinderPlus																*									
ramR T18P														X	X	X	X		X	X	X	X	⊗		32
acrB R717L	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100

E35
ResFinder_R

E22
ResFinder_N_R/
AMRFinderPlus_N
Curated for
duplicates

E14
ResFinder_R/
AMRFinderPlus_R
All mutations were
reported

E35
ResFinder_R
80/60

N50 = 27839
Contig No. = 953

E22
ResFinder_N_R/
AMRFinderPlus_N
90/40
A subset was
Reported

N50 = 6503
Contig No. = 1813

E14
ResFinder_R/
AMRFinderPlus_R
>90/?
All genes
were reported

N50 = 1443
Contig No. = 4687

Similar reporting to EQA providers results, except single cases.
Differences due to:
- Low QC?
- Reporting errors?

EQA2-WGS-AMR

Possible reasons of differences in reporting

Content of the different databases and tools
- Versions ?

AMR gene and PM reporting strategies
- When only one databases was used ?

Sequencing quality
- Participants with the same method in use ?

Mistakes in reporting

Ring-trial methods and their effect on the results

Overview of the tools, databases and inputs used by 29 NRLs

Combinations of Tools, Inputs and databases	Salmonella		Campylobacter	
	Genes	PMs	Genes	PMs
1 tool, 1 input/1 database	12	16	13	16
1 tool, >1 input/1 database	4	3	3	2
1 tool, 1 input/2 databases	1			
1 tool, >1 input/2 databases	1			
1 tool, 1 input/4 databases			1	1
2 tools, 1 input/1 database	1			
2 tools, 1 input/2 databases	2	3	3	4
2 tools, >1 input/2 databases	3	6	4	6
3 tools, 1 input/1 database		1		
3 tools, >1 input/2 databases	2		1	
3 tools, >1 input/3 databases	3		4	
No. of unique combinations	25	15	24	16
No. of shared unique combinations	2	7	2	5
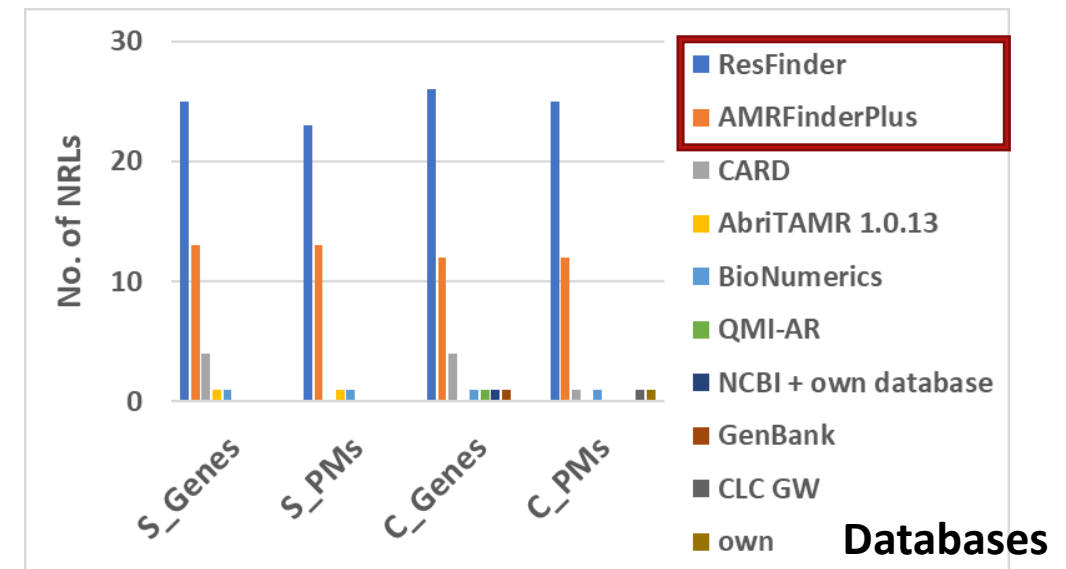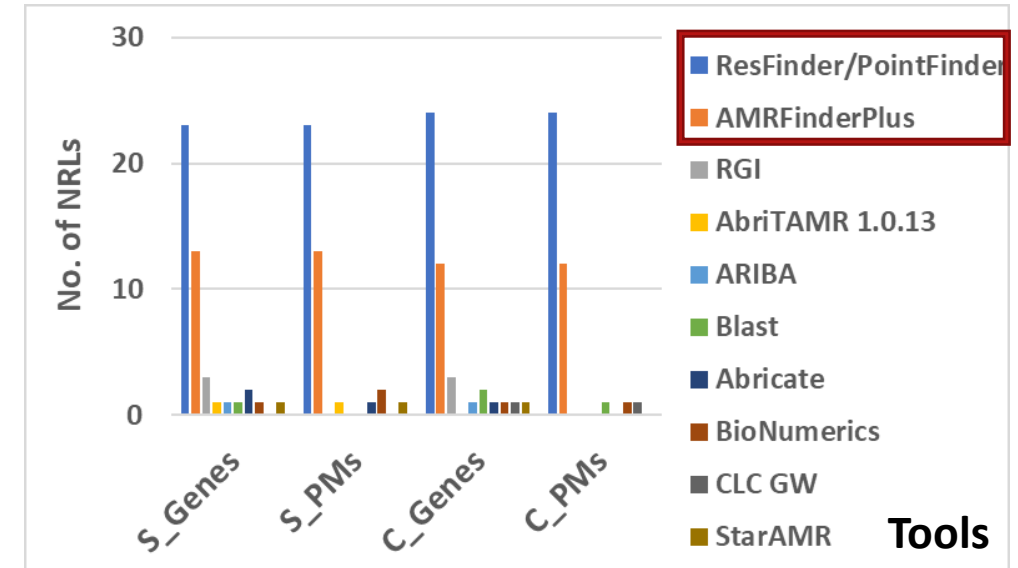
RingTrial2-WGS-AMR

Salmonella and *Campylobacter* DNA samples characteristics

FWD AMR· RefLabCap

Table 1. Genotypic and phenotypic characteristics of Salmonella strains selected for the RingTrial2

Strain	TRING2S-1	TRING2S-2	TRING2S-4	TRING2S-7	TRING2S-10
Serotype	Saintpaul	Meleagridis	Typhimurium	Newport	Monophasic Typhimurium
ST	50	463	19	132	34
Genes[A]	aac(3)-IId, aadA2, aph(3')-Ia, aph(3'')-Ib, aph(6)-Id, arr-2, blaTEM-1, dfrA14, floR, lnu(F), mph(A), qnrS1, sul2, tet(A)	aac(3)-IId, aac(6')-Ib-cr5, aadA16, aph(3'')-Ib, aph(6)-Id, arr-3, blaTEM-1, catA2, dfrA27, floR, fosA7.4, mph(A), qnrB6, sul1, sul2, tet(A)	aadA2, ant(2'')-Ia, blaCTX-M-9, catA1, dfrA16, qnrA1, sul1, tet(A)	aadA2, blaCARB-2, dfrA1, floR, mph(A), qnrA1, sul1, tet(A)	aac(3)-IIg, aac(6')-Ib3, aac(6')-IIc, aadA2, aph(3')-Ia, aph(3'')-Ib, aph(6)-Id, arr, blaSHV-12, blaTEM-1, dfrA19, ere(A), qnrB2, sul1, sul2, tet(B), tet(D)
PMs[A]	gyrA S83Y	None	None	None	None
NWT Phenotypes[B]	AMP, AZI, CHL, CIP, GEN, NAL, SME, TET, TIG, TRI	AMP, AZI, CHL, CIP, SME, TET, TRI	AMP, CTA, CHL, CIP, GEN, SME, TET, TRI	AMP, AZI, CHL, CIP, SME, TET, TRI	AMP, CTA, CTZ, CIP, GEN, SME, TET, TRI

A According to AMRFinderPlus
B Abbreviations of antimicrobials: AMP (Ampicillin), AZI (Azithromycin), Cefotaxim (Ciprofloxacin), GEN (Gentamicin), NAL (Nalidixic acid), SME (Sulfamethoxazole), (Trimethoprim) . Abbreviations used are based on EUCAST system :

Table 2. Genotypic and phenotypic characteristics of Campylobacter strains selected for the RingTrial2

Strain	TRING2C-1	TRING2C-4	TRING2C-7	TRING2C-9	TRING2C-10
Species	C. coli	C. jejuni	C. jejuni	C. jejuni	C. coli
ST	9263	257	7433	572	12073
Genes[A]	aad9, aadE, aadE-Cc, blaOXA-193, lnuC, tetO	blaOXA-461	aad9, aph(2'')-If, aph(3')-IIIa, blaOXA-193, catA13, tetO	aadE, aph(3')-IIIa, blaOXA-193, sat4	aad9, aadE, aph(3')-IIIa, blaOXA-193, catA, ermB, sat4, tetO
PMs[A]	gyrA T86I	None	50S L22 A103V, gyrA T86I	gyrA T86I	gyrA T86I, rpsL K43R
NWT Phenotypes[B]	CIP, ERY, NAL, STR, TET	None	CHL, CIP, ERT, GEN, TET	CIP, TET	CHL, CIP, ERT, ERY, TET

A According to AMRFinderPlus
B Abbreviations of antimicrobials: CHL (Chloramphenicol), CIP (Ciprofloxacin), ERT (Ertapenem), ERY (Erythromycin), GEN (Gentamicin), NAL (Nalidixic acid), STR (Streptomycin), TET (Tetracycline). Abbreviations used are based on EUCAST system :

RingTrial2-WGS-AMR

Common differences between reference datasets

Table 3. Tools and databases used in provider's reference data sets for Salmonella and Campylobacter

	AMR gene detection			Point mutation identification		
	Database	**Tool**	**Input**	**Database**	**Tool**	**Input**
Ref_Res	ResFinder	ResFinder	Raw reads (fastq)	ResFinder	PointFinder	Raw reads (fastq)
Ref_AMR	AMRFinderPlus	AMRFinderPlus	SPAdes assembly (fasta protein)	AMRFinderPlus	AMRFinderPlus	SPAdes assembly (fasta nucleotide)

Salmonella

Ref_Res	Ref_AMR
blaTEM-1B	blaTEM-1
aac(6')-Iaa	-
aac(6')-Ib-cr	-

Campylobacter

Ref_Res	Ref_AMR
cat	catA13
aph(3')-III	aph(3')-IIIa
ant(6)-Ia	aadE

Salmonella example: TRING2S-1 – reads vs assemblies

FWD AMR· RefLabCap

7 out of 13 participants that used ResFinder only, reported this gene – **all used reads** (but 2 of them also used assemblies)

The other 6 participants that used ResFinder only, and did not report this gene – **all used assemblies** (but 3 of them also used reads)

	ResFinder													AMRFinderPlus +/- ResFinder											CARD +/- others						% concordance
	Res_Ref	R15	R23	R25	R17	R30	R31	R27	R07	R33	R05	R18	R20	R38	AMR_Ref	R08	R35	R01	R10	R14	R04	R28	R21	R34	R39	R06	R02	R24	R32	R22	R40	
ResFinder																																
AMRFinderPlus																					**											
CARD																																
Other database																														*	**	
aac(3)-IId	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	93
aac(6')-Iaa	X	X	X	X	X	X		X	X	X		X	X	X							X	X			X		X	X			X	55
aadA2		X													X	X	X	X	X	X	X	X		X		X	X	X	X			
aadA17	X	X	X		X	X	X	X	X	X	X	X	X								X				X						83
aph(3')-Ia	X	X	X	X		X		X	X	X	X	X	X	X	X						X	X		X		X	X	X	X	X	X	90
aph(3")-Ib	X	X	X		X			X	X	X	X			X						X	X	X	X		X	X	X	X	X			62
aph(6)-Id	X	X	X	X	X		X		X	X	X	X	X	X						X	X	X	X	X	X	X	X	X	X	X		90
arr-2	X	X	X	X	X		X	X	X	X	X	X	X	X						X	X	X	X	X	X	X	X	X	X	X		93
	X	Ⓧ	Ⓧ		Ⓧ	Ⓧ		Ⓧ		Ⓧ			Ⓧ								X					X						31
blaTEM-1															X	X	X	X	X	X	X	X				X	X				
blaTEM-1B	X	X	X	X	X	X	X	X	X	X	X	X	X	X					X			X	X	X	X	X				X	X	97
dfrA14	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	97
floR	X	X	X		X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	90
lnu(F)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	93
mph(A)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
qnrS1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	97
sul2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	97
tet(A)	X	X	X	X	X	X		X		X		X	X	X	X					X	X	X	X	X	X	X	X	X	X	X	X	76

* GenBank
** BioNumerics 8.1

Gene is not detected when assemblies are used – split into contigs?

Salmonella example: TRING2S-10

arr gene not detected in ResFinder or CARD (provider's data), even though there are 10 variants of *arr* gene in ResFinder

AMRFinderPlus has many alleles of this gene, it was detected with 100% identity and coverage in AMR_Ref

Rifamycin resistance – not tested

	ResFinder														AMRFinderPlus +/- ResFinder											CARD +/- others						% concordance
	Res_Ref	R15	R23	R25	R17	R30	R31	R27	R07	R33	R05	R18	R20	R38	AMR_Ref	R08	R35	R01	R10	R14	R04	R28	R21	R34	R39	R06	R02	R24	R32	R22	R40	
ResFinder																					**											
AMRFinderPlus																					**											
CARD																																
Other database																														*	**	
aac(3)-IIg															X	X	X	X	X	X	X	X	X	X		X		X	X			38
aac(6')-Ib	X				X		X		X																							
aac(6')-Ib3	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
aac(6')-Ib-cr	X	X	X		X	X			X	X	X	X	X	X								X	X	X		X						100
aac(6')-IIc	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
aac(6')-Iaa	X	X	X	X	X	X		X				X	X	X					X	X	X	X	X		X		X		X		X	62
aadA2	X	X	X		X	X		X		X	X		X	X	X	X	X		X		X		X			X	X	X	X			
aadA2b				X					X				X																			90
aph(3')-Ia	X	X	X	X	X	X		X	X	X	X		X	X	X	X	X		X		X	X	X			X	X	X	X			90
aph(3'')-Ib	X	X	X	X	X	X	X	X		X	X	X	X	X	X						X	X	X		X	X	X		X		X	83
aph(6)-Id	X	X	X	X	X	X	X	X		X	X	X	X	X	X				X		X	X	X		X	X	X		X		X	83
arr															X	X	X	X	X		X		X		X		X		X	X		38
blaSHV-12	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
blaTEM-1															X	X	X	X	X	X	X	X	X	X			X	X				
blaTEM-1B	X	X	X	X	X	X	X	X	X	X	X	X	X	X					X		X	X	X	X	X	X	X		X	X	X	100
dfrA19	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	97
ere(A)	X	X	X	X	X	X		X				X	X	X	X											X	X	X	X	X	X	66
mcr-9	X	X	X			X	X	X	X	X	X	X	X	X			X			X		X	X	X	X	X	X				X	79
qnrB2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
sul1	X	X	X	X	X	X		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	97
sul2	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X		X	X	X	93
tet(B)	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	97
tet(D)	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	97

* GenBank
** BioNumerics 8.1

tet genes in *Campylobacter* - always some issues!

TRING2C-1: 99.9% identity, 65% coverage   91% identity, 65% coverage

TRING2C-7: tet(O) not detected in ResFinder in April 2023, but detected in November 2023

TRING2C-9: both *tet(O)* (>90% identity and coverage) and *tet(O/32/O)* (100% identity and coverage)
    Only when reads were used (provider and participants)

TRING2C-10: 99.8% identity, 100% coverage   93% identity, 100% coverage

RingTrial2-WGS-AMR

tet genes in *Campylobacter* - always some issues!

TRING2C-1: 99.9% identity, 65% coverage   91% identity, 65% coverage

tet(O)														X												X		X	
tet(O/32/O)	X	X	X	X	X	X	X	X	X		X	X	X				X		X	X	X	X	X	X		X	X	X	79

TRING2C-7:   ...mber 2023

tet(O)						X	X	X	X	X	
tet(O/32/O)											83

TRING2C-9: both *tet(O)...*   ...ge)
Only when...

tet(O)												
tet(O/32/O)							X					
tet(O/M/O)								X	X	X		62

General remarks

- *tet* gene usually reported as *tet(O)* with AMRFinderPlus and as *tet(O/32/O)* by ResFinder, with some exceptions

- Often issues when assemblies are used – likely split into contigs

- Previous study (Dahl et al. 2021) identified *tet(O/32/O)* gene in 27 strains by mapping but not by assembly-based methods

TRING2C-10: 99.8% identity, 100% coverage   93% identity, 100% coverage

tet(O)															X			X	X						X	X					X			
tet(O/32/O)	X	X	X	X	X	X	X	X	X	X	X	X	X	X				X	X		X	X	X	X	X		X	X	X		90

Salmonella genes – same unique combination, different results

FWD AMR·RefLabCap

	Tools/Inputs[A]	Databases/Inputs[A]	No. of participants	Participant ID[B]	Identity (%)	Coverage (%)
	1 tool, 1 input	1 database, 1 input				
1	ResFinder_N	ResFinder_N	2	R18	30	20
2				R20	90	60
3	ResFinder_R	ResFinder_R	6	R15	90	60
3				R17	90	60
3				R30	90	60
3				R33	90	60
4				R23	85	60
5				R38	80	60

TRING2S-1:
- R17 and R33 did not report *aph(3'')-Ib*
- Same reporting for the remaining *18 genes*

TRING2S-2:
- R33 did not report *aph(3'')-Ib* and *aph(6)-Id*
- Same reporting for the remaining 18 genes

TRING2S-4:
- Same reporting for all 11 genes

TRING2S-7:
- Same reporting for all 10 genes

TRING2S-10:
- R15 and R17 did not report *aac(6')-Ib*
- R33 did not report *ere(A)*

RingTrial2-WGS-AMR

Differences in reporting?
Different versions of tools/databases?
Reporting erorrs?
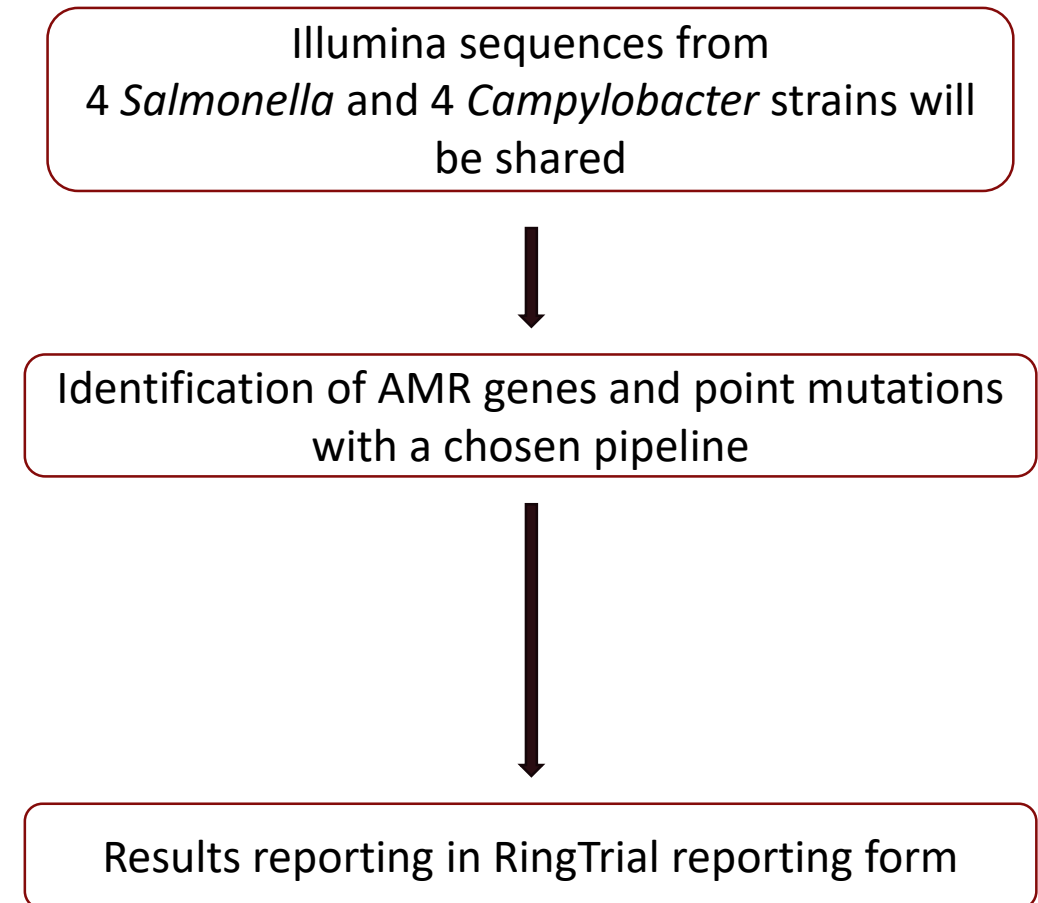
Conclusions – both EQA2 and RingTrial2

- Majority of participants produced sequences of acceptable quality

- High variation of combinations of tools/software/thresholds/reporting strategies used by participants – difficult to compare results

- In general good concordance with some exceptions

- Thresholds, reads vs assemblies, nomenclature can have effect on reporting

- Cannot always find explanations for everything
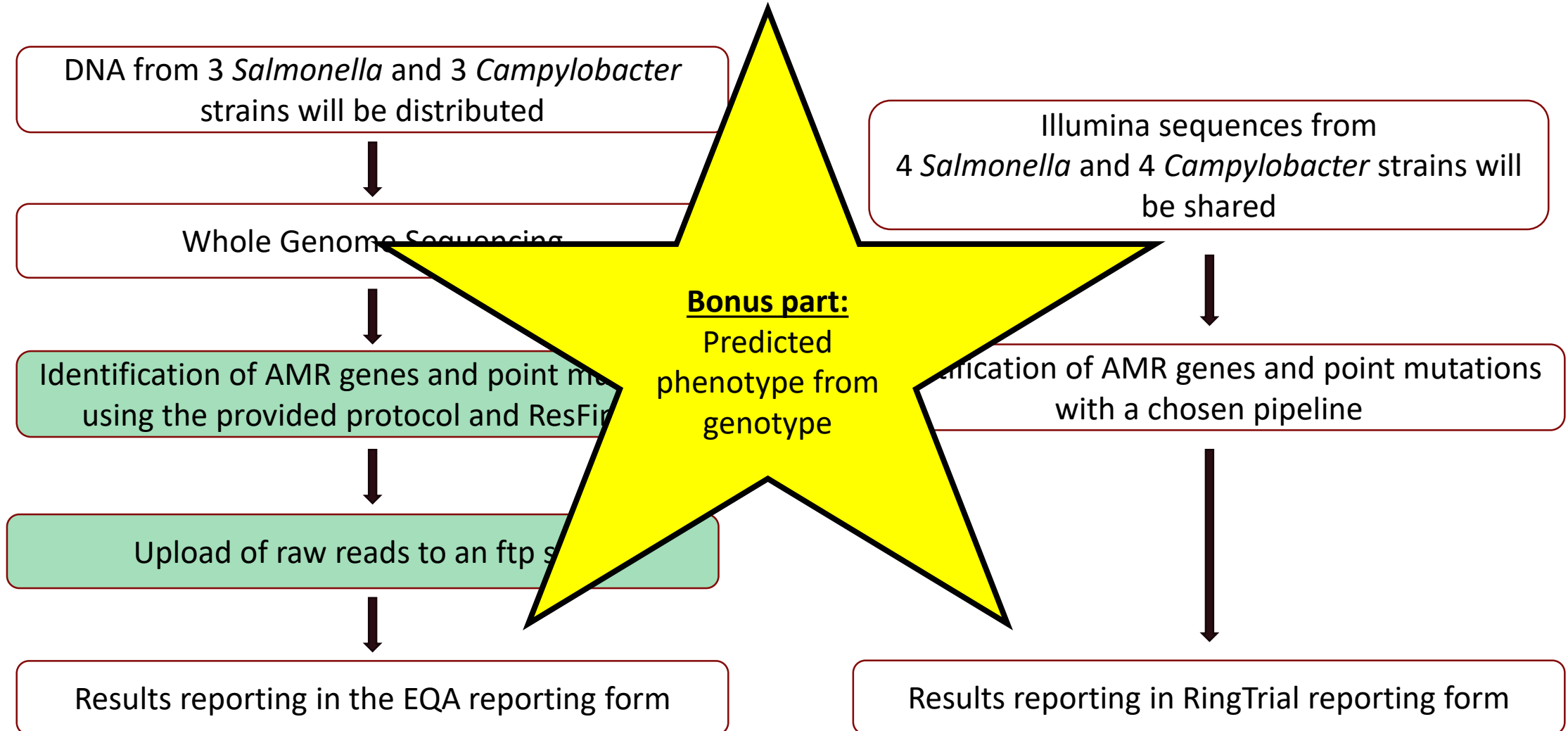
Plans for the next round

EQA3-WGS-AMR

DNA from 3 *Salmonella* and 3 *Campylobacter* strains will be distributed

↓

Whole Genome Sequencing

↓

Identification of AMR genes and point mutations using the provided protocol and ResFinder

↓

Upload of raw reads to an ftp site

↓

Results reporting in the EQA reporting form

RingTrial3-WGS-AMR

Illumina sequences from 4 *Salmonella* and 4 *Campylobacter* strains will be shared

↓

Identification of AMR genes and point mutations with a chosen pipeline

↓

Results reporting in RingTrial reporting form

Plans for the next round

FWD AMR·RefLabCap

EQA3-WGS-AMR

DNA from 3 *Salmonella* and 3 *Campylobacter* strains will be distributed

⬇

Whole Genome Sequencing

⬇

Identification of AMR genes and point mutations using the provided protocol and ResFinder

⬇

Upload of raw reads to an ftp server

⬇

Results reporting in the EQA reporting form

RingTrial3-WGS-AMR

Illumina sequences from 4 *Salmonella* and 4 *Campylobacter* strains will be shared

⬇

Identification of AMR genes and point mutations with a chosen pipeline

⬇

Results reporting in RingTrial reporting form

Bonus part:
Predicted phenotype from genotype

Eva Møller Nielsen

Karen Loaiza Conza

FWD AMR·
RefLabCap

Thank you for your attention and big thanks to these colleagues
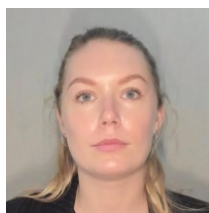
Anne Sophie Majgaard
Uldall
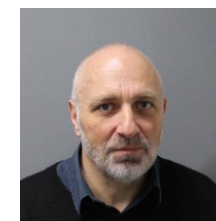
Jeppe Boel
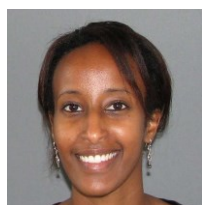
Ibado Mahad

Mia Torpdahl

Colleagues from DTU
René Hendriksen
Susanne Karlsmose Pedersen
Ana Rita Bastos Rebelo
Elif Seyda Tosun
Sarah Marvig Johansson