European Commission

Service contract for the provision of EU networking and support for public health reference laboratory functions for antimicrobial resistance in *Salmonella* species and *Campylobacter* species in human samples
SC 2019 74 09

Deliverable T1.16.2

Report on the second annual *in vitro* external quality assessment scheme for WGS-based resistome profiling of *Salmonella* and *Campylobacter*

EQA2-WGS-AMR

Version n°: 2

Written by Statens Serum Institut
January 2024

STATENS SERUM INSTITUT

DTU

FWD AMR-RefLabCap

This report was produced under the EU Third Health Programme 2014-2020 under a service contract with the Consumers, Health, Agriculture and Food Executive Agency (Chafea) acting under the mandate from the European Commission. From 1 April 2021, the European Health and Digital Executive Agency (HaDEA) took over all contractual obligations from Chafea. The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the Commission/Executive Agency. The Commission/Executive Agency do not guarantee the accuracy of the data included in this study. Neither the Commission/Executive Agency nor any person acting on the Commission's/Executive Agency's behalf may be held responsible for the use which may be made of the information contained therein.

EUROPEAN COMMISSION

Directorate-General Health and Food Safety (DG SANTE)

Directorate B — Public health, Cancer and Health security

Unit B2 — Health security

*L-2*920 Luxembourg

Email : SANTE-CONSULT-B2@ec.europa.eu

Health and Digital Executive Agency (HaDEA)

HaDEA COV2

Place Rogier, 16

B-1049 BRUXELLES

Belgium

Email : HaDEA-HP-TENDER@ec.europa.eu

Deliverable T1.16.2

Report on the second annual *in vitro* external quality assessment scheme for WGS-based resistome profiling of *Salmonella* and *Campylobacter*

EQA2-WGS-AMR

Authors (SSI): Małgorzata Ligowska-Marzęta, Egle Kudirkiene, Jeppe Boel, Eva Møller Nielsen

Contributors from DTU: Ana Rita Bastos Rebelo, Susanne Karlsmose Pedersen, Rene S. Hendriksen

Contributors from SSI: Karen Loaiza Conza, Anne Sophie Majgaard Uldall, Mia Torpdahl, Susanne Schjørring, Pernille Gymoese, Katrine Grimstrup Joensen

STATENS SERUM INSTITUT

DTU

FWD AMR· RefLabCap

Table of Contents

1. Background

This report presents the second External Quality Assessment for WGS-based resistome profiling in antimicrobial-resistant *Salmonella* and *Campylobacter* (EQA2-WGS-AMR). The EQA is second out of three planned EQAs (Deliverable T1.16), organized by Statens Serum Institut (SSI) in collaboration with the Technical University of Denmark (DTU) in the FWD AMR-RefLabCap project.

The aim of this EQA was to compare the participant's ability to identify genes and point mutations that confer antimicrobial resistance in *Salmonella* and *Campylobacter* using whole genome sequencing (WGS), based on provided DNA samples. The participants were recommended to follow the analytical guidelines described in the protocol (https://www.fwdamr-reflabcap.eu/resources/reflabcap-protocols-and-guidelines) that were developed in the FWD AMR-RefLabCap project. Participation in the EQA2-WGS-AMR will enable the laboratories to identify strengths and weaknesses in their technical and analytical setup and implement improvements, when needed.

DNA from three isolates of *Salmonella* and three isolates of *Campylobacter* was included in this EQA. Fourty public health laboratories from the FWD network and one veterinary institute were invited to participate. Thirty four laboratories accepted the invitation and 30 participants submitted results. The participants represented a total of 26 countries, including nine priority countries.

2. Materials and methods

2.1. Strain selection

The strains used in this EQA2-WGS-AMR represent a wide array of antimicrobial resistance markers and were selected from the SSI strain collection. The genotypic and phenotypic antimicrobial resistance features of each strain are shown in Table 1 and Table 2.

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)
Abbreviations used are based on EUCAST system :
https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/Disk_abbreviations/EUCAST_system_for_antimicrobial_abbreviations.pdf

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)
Abbreviations used are based on EUCAST system : https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/Disk_abbreviations/EUCAST_system_for_antimicrobial_abbreviations.pdf

2.2. Phenotypic testing by the EQA provider

The strains were phenotypically tested for antimicrobial susceptibility by determination of MIC values and subsequent classification as wild type (wt) or non-wild type (nwt) using epidomiological break point values (1). MIC determination was performed following the harmonised EU AST protocol using microbroth dilution method with EUVSEC and EUVSEC2 TREK panels from Thermo Scientific, Denmark for *Salmonella* and EUCAMP2 panels for *Campylobacter* (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020D1729&from).

The *Salmonella* panels included the following antimicrobials: Ampicillin, Azithromycin, Cefepime, Cefotaxime, Cefoxitin, Ceftazidime, Chloramphenicol, Ciprofloxacin, Colistin, Ertapenem, Gentamicin, Meropenem, Nalidixic acid, Sulfamethoxazole, Temocillin, Tetracycline, Trimethoprim and Trimethoprim-sulfamethoxazole. For *Campylobacter*, the panels included Ciprofloxacin, Erythromycin, Gentamicin and Tetracycline. The results are shown in Table 1 and Table 2. The selection of antimicrobials tested was based on the priority list of antimicrobial agents set in the harmonised EU AST protocol (1), recommended by the ECDC.

In most cases there is a correlation between the pheno- and genotypes and from the established phenotypes for the test strains it is possible to evaluate the phenotypic predictions for these antimicrobials. However, some strains included in this EQA harbour genes that confer resistance towards antimicrobials that the test strains have not been tested for and some strains have a phenotype for which there is no genetic determinant present. This is *e.g.* the case for colistin resistance phenotype in *Salmonella* and chloramphenicol and beta-lactam genes in *Campylobacter*.

2.3. Strain culturing, DNA extraction and distribution

The DNA samples were prepared by DTU. The isolates were cultured on blood agar plates. The *Salmonella* strains were incubated aerobically for 16-20 hours at 35±1°C and *Campylobacter* strains were incubated for 24 hours at 41.5±1°C in microaerophilic atmosphere.

After initial incubation, a lawn of bacteria was created on blood agar by streaking a suspension of the strains using a plate rotator and incubating overnight. Colony mass from a 10 µl loop was harvested from each strain and resuspended in 3 ml PBS. Tubes were centrifuged at 20 000 G for 5 minutes and after decantation, the pellet was resuspended in 600 µl PBS buffer.

For DNA isolation, the Easy-DNA Kit for genomic DNA isolation was used (Invitrogen). Protocol #3 from the kit was applied in triple volume and nuclease-free water with 2 mg/ml RNase was used to resuspend the precipitated pellet. The DNA concentration was measured using Qubit with 2 µl DNA sample. The DNA was diluted to approximately 50 ng/µl and stored at -20°C.

Twenty microliter of the dilution was aliquoted into 1.5 ml Eppendorf tubes and vaccum dried using a vacuum centrifuge (Eppendorf Concentator plus) for 30 minutes at temperature 25-30°C using the default settings. The tubes with dried samples were stored in bags at room temperature together with a silica bag. The pellet from one tube of each strain was resuspended in 50 ul nuclease free water and concentration was measured using Qubit™ dsDNA Quantification High Sensitivity (HS) kit. The measured concentrations are presented in Table 3.

Table 3. DNA concentrations measured using Qubit™ dsDNA HS kit.

Species	Strain	Concentration [ng]
Salmonella	EQA2-S23-01	720
	EQA2-S23-02	610
	EQA2-S23-03	1100
Campylobacter	EQA2-C23-01	750
	EQA2-C23-02	1155
	EQA2-C23-03	705

Fourty tubes with dried DNA from each strain were produced and delivered to SSI, packed in zip-lock bags together with silica bags placed in bubble envelopes. The packages with DNA samples were shipped at room temperature.

2.4. WGS analysis by the EQA provider

DNA from *Salmonella* and *Campylobacter* strains was sequenced using paired-end Illumina sequencing. The quality of the sequences (genome size, N50, total number of contigs) was checked with an in-house QC pipeline (https://github.com/ssi-dk/bifrost) for raw reads and BioNumerics for assemblies.

Salmonella serotypes were determined using Enterobase and SeqSero (https://github.com/denglab/SeqSero), as well as in-house developed scripts detecting the subspecies and genetic marker implicating the *d*-Tartrate reaction for distinguishing *S* Paratyphi B var. L(+) tartrate+ (var. Java) from *S* Paratyphi B.

For *Campylobacter* species identification, Kraken was used (https://github.com/DerrickWood/kraken). MLST calling was performed with ARIBA (https://github.com/sanger-pathogens/ariba) using the typing schemes from the PubMLST database.

The sequences were analysed by the EQA provider for the presence of antimicrobial resistance genes and point mutations by querying two different databases: ResFinder and AMRFinderPlus. The results obtained with the two approaches shown in Table 4, will be referred to as “reference datasets” in the report.

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)

In the result analysis, each reference dataset was compared to genes and point mutations reported by the participants using the same database or a combination of databases.

2.5. SurveyXact reporting scheme and collection of results

The reporting platform was developed in the SurveyXact survey tool (https://www.survey-xact.dk).

The reporting scheme consisted of three parts. The first part included questions about method and quality parameters for each strain, such as sequencing technology, method for DNA concentration measurement, as well as DNA concentrations, genome size, N50 and total number of contigs. The second part of the scheme focused on tools and databases used to identify the sequence type (ST), AMR genes, point mutations, as well as the serotype and species for *Salmonella* and *Campylobacter*, respectively. Furthermore, questions in this part included identity- and coverage cut-offs used for identifying genes and point mutations, as well as an additional question for participants that reported that they had used more than one database. The third part was for reporting AMR genes and point mutations (PMs). It was possible to select multiple genes from a list of genes in alphabetical order. It was also possible to report a gene in a free text field, in case it was not present on the default list. For reporting of point mutations, the participants were asked to type the detected mutations in text boxes.

All participants received individual links to the reporting form, where it was possible to report results for one or both pathogens. The time given for reporting of the results was two months from the sample shipping date.

Twenty laboratories reported results for both *Salmonella* and for *Campylobacter*, five laboratories for *Salmonella* only and five for *Campylobacter* only. The participating countries were Austria, Belgium, Czech Republic, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovenia, Spain and Sweden. The participating laboratories were randomly assigned codes and these codes were used for identification of laboratories in the EQA.

3. *Salmonella* results

3.1. Quality metrics for all *Salmonella* DNA samples

3.1.1. DNA concentration and its evaluation

All 25 participants reported DNA concentrations and the methods used for concentraton determination for all three *Salmonella* samples (Table S 1 – S 3).

The participants reported varying DNA concentrations for the three samples (Figure 1). The reported concentrations of sample EQA2-S23-01 were lower than the concentrations of the other two samples as illustrated in Figure S 1. The low concentration meant that several laboratories did not attempt to sequence the samples (see paragraph 3.1.2).

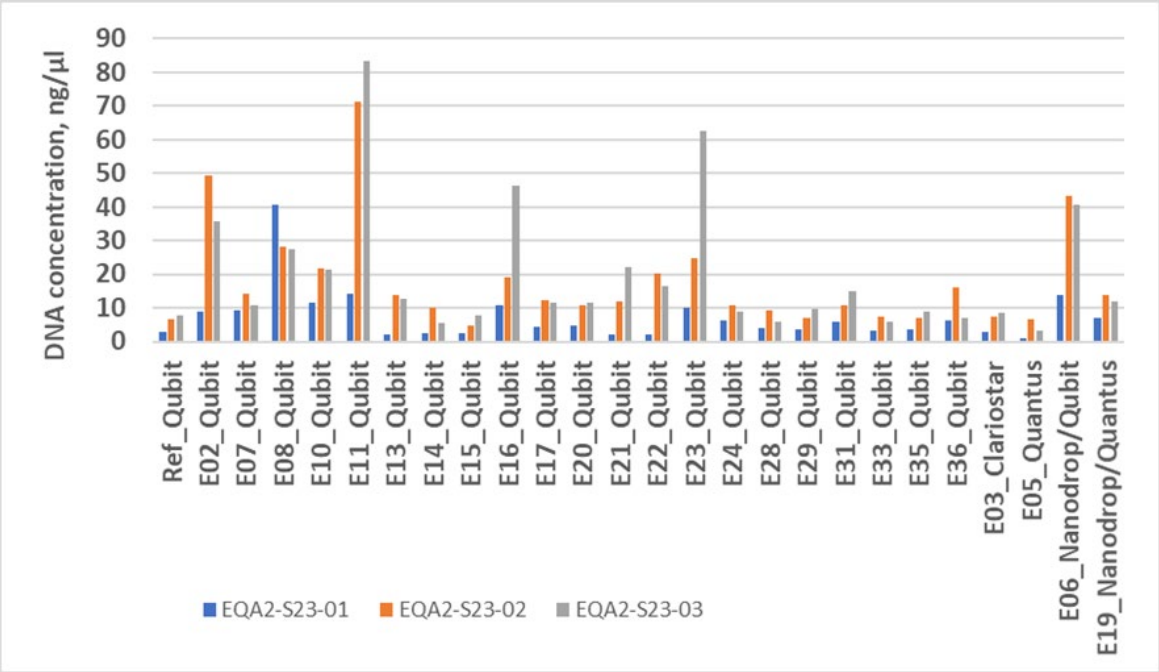

Figure 1. The distribution of Salmonella DNA concentrations among 25 participants with the indicated method(s) used.

3.1.2. Evaluation of genome assembly

Twenty-three participants used Illumina technology and two participants used Ion Torrent technology. The quality parameters of the genome assemblies were evaluated against the thresholds recommended in the suggested WGS protocol (https://www.fwdamr-reflabcap.eu/resources/reflabcap-protocols-and-guidelines) which are a genome size of 4.4 Mb-5.8 Mb, a N50 of > 30 000 bp, and a contig number of <500. The laboratories were expected to use their usual pipelines for evaluating the quality of assemblies. To assemble *Salmonella* genomes, most participants (n=19) used SPAdes alone with or without contig filtering where the cut-off varied from 200 to 500 bp (Figure 2).

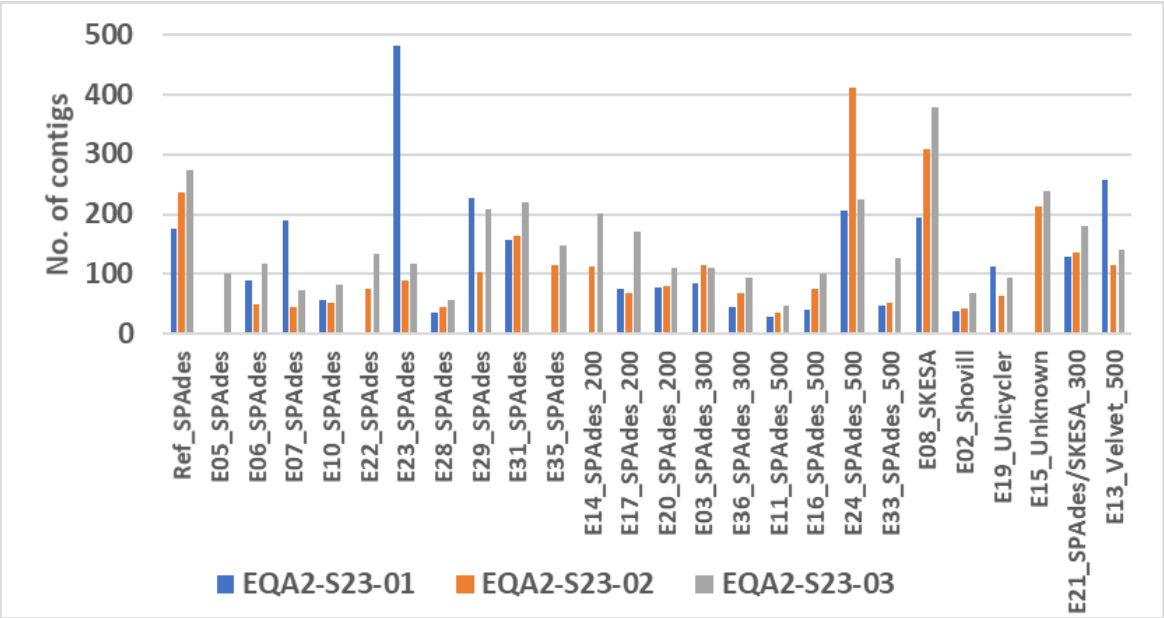

Figure 2. The distribution of number of contigs among 25 participants for three Salmonella samples. The horizontal axis labels indicate the ID of the participant, the tool used for genome assembly, and if contig length-based filtering was used it also indicates the filtering length. Note: the figure does not include the results with the number of contigs of >500 (EQA2-S23-01: E05, E14, E15, E22, E23, E35, and for EQA2-S23-02: E05)

All participants provided data on quality parameters of the assembled genomes: genome size, N50 and number of contigs (Table S 1 - S 3). The evaluation of these parameters are provided for each DNA sample separately below.

EQA2-S23-01

Genome sizes between 0.1 and 5.0 Mb (av.=4.5), N50 between 270 bp and 420 012 bp (av. 159 848), and numbers of contigs between 28 and 7 654 (av.=796) were reported for the DNA sample EQA2-S23-01. Nineteen out of 25 participants reported genomes sizes, N50 and contig number that complied with the recommended limits (Table S 1).

Six participants (E05, E14, E15, E22, E23, E35) reported results below the recommended thresholds for at least one of the three QC parameters. Laboratories E05, E15 and E23 reported all parameters below the recommended thresholds and, consequently, they did not analyse this sample further. Laboratories E14, E22 and E35 reported the expected genome size, but N50 of <30 000 bp and contig numbers of >500. Despite that, all three participants reported the AMR genes and point mutations in this strain. One participant, E22, commented on the poor QC metrics for the strain, whereas the two other laboratories did not comment on the issue.

EQA2-S23-02

The participants reported genome sizes between 4.9 and 5.0 Mb (av.=5.0), N50 between 14 321 bp and 437 353 bp (av. 235 076), and the number of contigs between 36 and 728 (av.=135).

All, except one laboratory, reported the genome sizes, N50 and contig number within the recommended limits for *Salmonella* (Table S 2). Participant E05 reported the expected genome size but with N50 of <30 000 bp and >500 contigs, and submitted results for the sample.

EQA2-S23-03

For this *Salmonella* DNA sample, the participants reported genome sizes between 4.8 and 5.0 Mb (av.=4.9), N50 between 32 046 bp to 254 048 bp (av. 156 386), and the number

of contigs between 48 and 378 (av.=142). All reported genome sizes, N50 and contig number were within the recommended limits (Table S 3).

3.2. AMR gene and PMs detection methods used

3.2.1. Tools and databases used for AMR gene detection

All 25 participants reported the applied tools, databases, types of files used as inputs, thresholds for sequence coverage and sequence identity for AMR gene detection, as well as how they reported the AMR genes. Overall, 22 unique combinations of tools/databases/inputs/thresholds/gene reporting strategies were used by 25 participants (Table S 7).

The most commonly used tool was ResFinder, followed by AMRFinderPlus. ResFinder was used by 21 participants, AMRFinderPlus by 10 participants, and RGI by five participants. AbriTAMR 1.0.13, ARIBA, Blast, and Abricate were used by one participant each (Figure 3).

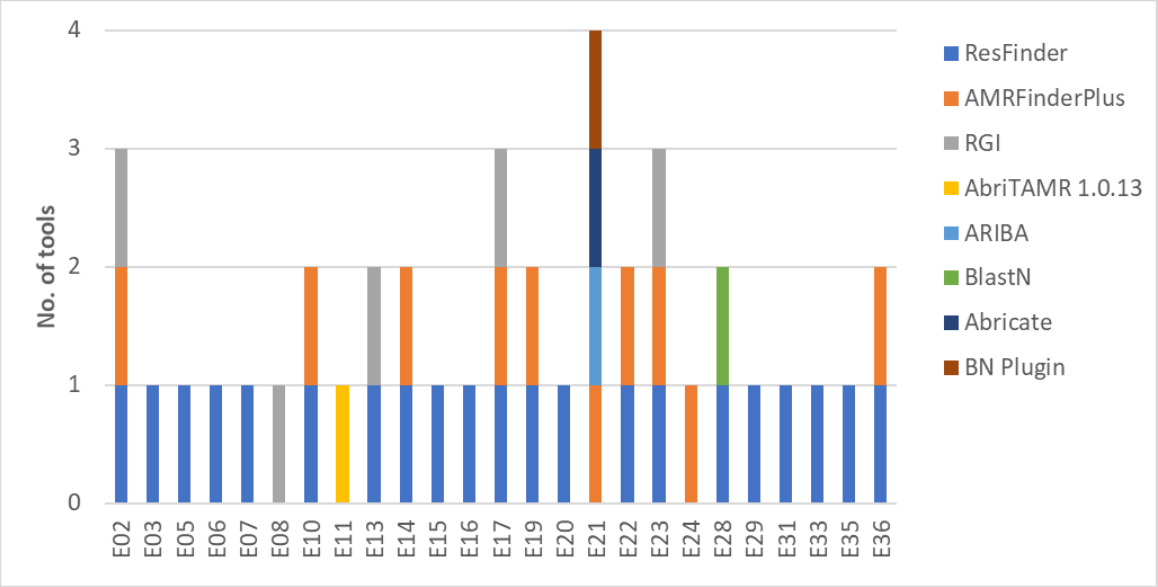

Figure 3. An overview of tools used by 25 participants for AMR gene detection in Salmonella

The ResFinder database was used by 22 participants and the AMRFinderPlus database was used by 10 participants. The CARD database was used by three participants, and the AbriTAMR 1.0.13 was used by one participant. It is assumed by the provider that the participant who reported the AbriTAMR 1.0.13 tool as the database used in fact the AMRFinderPlus database, which is the default for this tool. The participants who used more than one database also indicated how they reported AMR genes. Four participants reported a consensus list of genes (common genes present in all databases used), four participants reported a subset of genes based on experience/knowledge/literature, and two participants reported all genes from all databases. Additionally, participant E02 indicated that they reported AMR genes present in at least one of the three databases and participant E13 that they reported the consensus of AMR genes based on a literature review (Figure 4).

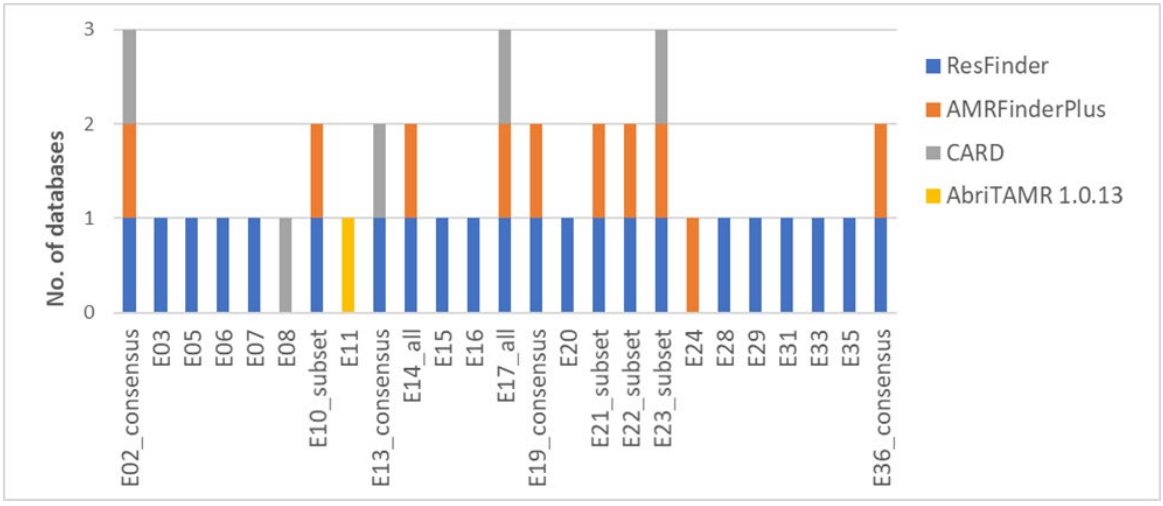

Figure 4. An overview of databases used by 25 participants for AMR gene detection in Salmonella. The horizontal labels indicate the participant ID and how they reported AMR genes in cases when more than one database was used: consensus – common genes present in all databases used, subset – based on experience/knowledge/literature, all – genes from all databases. It is assumed by the provider that participant E11, who reported the tool AbriTAMR 1.0.13 as the database, used the AMRFinderPlus database, which is the default for this tool.

3.2.2. Tools and databases used for point mutations detection

All participants reported the tools, the databases, and the inputs that they used for point mutations detection as well as the approach applied in cases where more than one database was used. Overall, 14 unique combinations of tools/databases/inputs/reporting strategies were used by the 25 participants (Table S 8).

PointFinder was the preferred tool, being used by 22 participants, either alone or in combination with another tool. AMRFinderPlus was the second most common tool and was used by 10 participants (Figure 5).

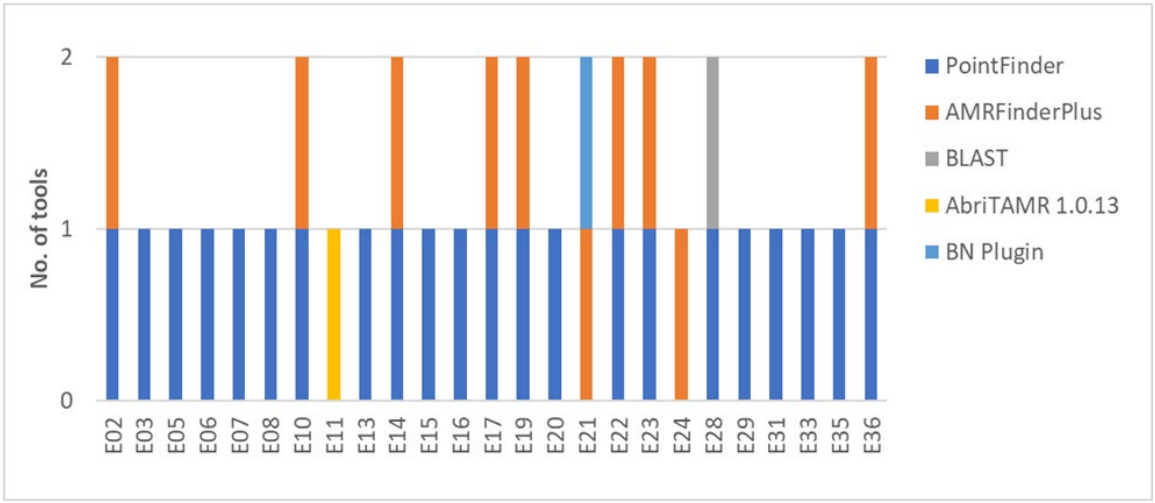

Figure 5. An overview of tools used by 25 participants for point mutations detection in Salmonella

Seventeen participants used only one database and eight participants used two databases (Figure 6). The eight participants which used two databases, used ResFinder and AMRFinderPlus. The participants that used more than one database also indicated how they reported the point mutations (Figure 6).

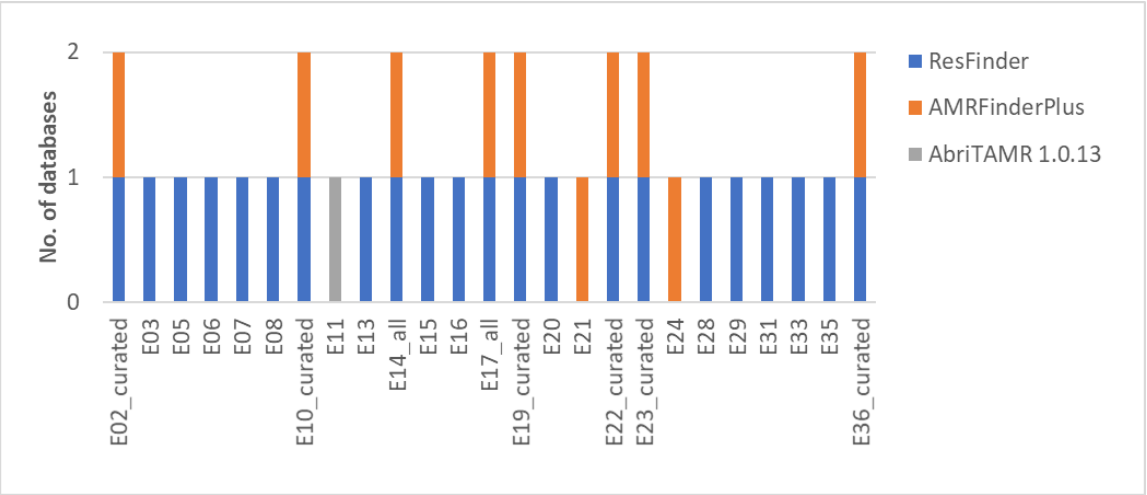

Figure 6. An overview of databases used by 25 participants for point mutations detection in Salmonella. The horizontal labels indicate the participant ID and how they reported AMR genes in cases when more than one database was used: "all" - reported all point mutations from both databases without curation, "curated" – reported the curated results the point mutations from all databases for duplicates. It is assumed by the provider that participant E11, who reported the tool AbriTAMR 1.0.13 as the database, used the AMRFinderPlus database, which is the default for this tool.

3.3. Serotypes and STs reported

3.3.1. Serotyping methods and serotypes

Sixteen participants used only one tool/software for *Salmonella* serotyping, and nine participants used a combination of two or three tools/softwares. The most commonly used tool was SeqSero2, used by 20 participants (Figure 7).

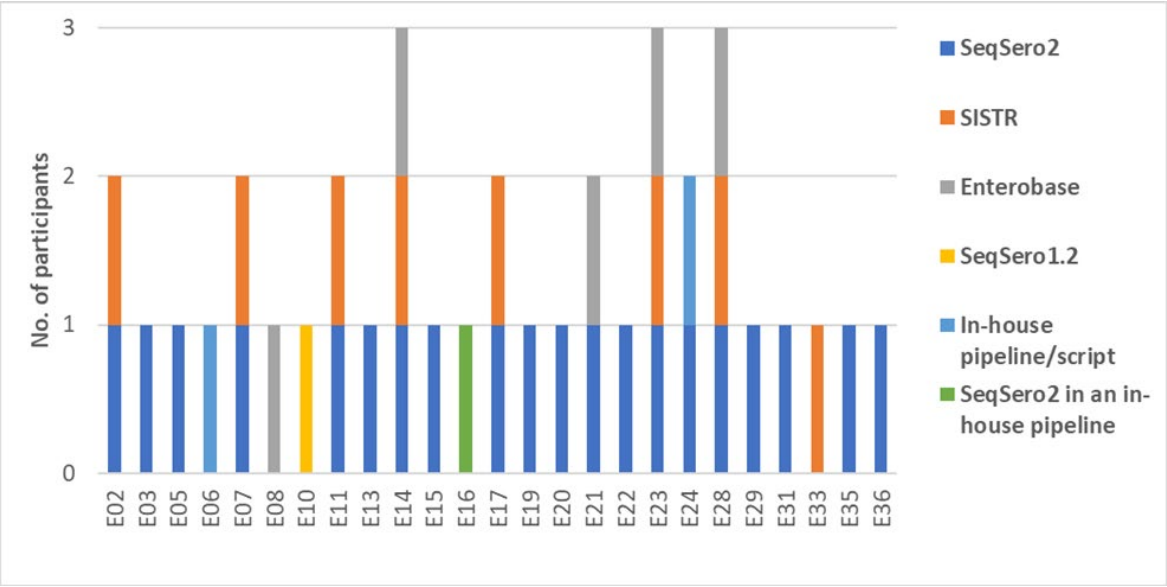

Figure 7. An overview of the tools/softwares used by 25 participants for Salmonella serotyping.

All 25 participants reported serotypes for the DNA samples EQA2-S23-02 and EQA2-S23-03, and 22 participants provided serotypes for the sample EQA2-S23-01. The lacking serotypes for EQA2-S23-01 were due to quality issues (see paragraph 3.1.2 for details).

An overview of concordance of the reported serotypes for the three test samples is presented in Table 5. In general, the participants reported serotypes that were in accordance with the serotypes established by the EQA provider. In total, six non-concording

results were reported, of which three were reported by the same participant (Table S 12). The reason for reporting a non-concording result for sample EQA2-S23-01 by participant E05 is likely related to insufficient sequence quality of this sample.

Table 5. Reported Salmonella serotypes and concordance with EQA provider's results

Sample	**EQA2-S23-01**	**EQA2-S23-02**	**EQA2-S23-03**
Serotype	Dublin	Stanley	Rissen
Concording results	19	24	24
Non-concording results	3	1	1
No serotype reported	3		

3.3.2. MLST methods and STs

Eleven participants used the MLST2.0 scheme that is available in CGE tools (https://cge.food.dtu.dk/services/MLST/), five participants used the Enterobase MLST scheme, and four participants used the Tsemann MLST scheme for 7 gene MLST typing. The remaining five participants indicated that they used other schemes/tools: Ridom SeqSphere+, senterica_achtman_2 scheme, used PubMLST, and in-house Bifrost using Enterobase scheme.

For strain EQA2-S23-01, 21 participants correctly reported the strain as ST10 and four participants did not report the ST. For strain EQA2-S23-02, 24 participants correctly reported ST29, whereas one participant reported the strain as ST3241. Twenty-four participants correctly reported the ST469 for strain EQA2-S23-03 and one participant did not report ST for this strain (Table S 11).

3.4. AMR genes and PMs reported for *Salmonella* DNA samples

The genes identified by the EQA provider using two different tools and databases, namely Res_Ref and AMR_Ref, and the genes identified by the participants are presented for each strain in the following paragraphs. The letter "X" indicates the detection of a specific gene.

The results from the participants' and the EQA provider's reference datasets are divided into three categories based on which database was used. The green (ResFinder) category indicates laboratories that only used the ResFinder database (2)(3). Participants that used the AMRFinderPlus database (4), either alone or with ResFinder, are grouped in the blue category (AMRFinderPlus +/- ResFinder). When relevant, the third, yellow, category is applied, grouping the laboratories that used CARD (5), either alone or in combination with any other databases. In cases where a participant used a database different from ResFinder, AMRFinderPlus or CARD, it is marked with an asterisk in the tables and the database name is stated in the table footer.

For each gene and PM table, the concordance of the reported results among the participants was calculated. This number is expressed as a percentage of the total number of laboratories that reported the same genes or point mutations for a given DNA sample. That number varies between DNA samples from 18 to 25, as not all participants reported point mutations and certain participants were unable to report genes for *Salmonella* strain

EQA2-S23-01. When possible, explanations of observed discrepancies between reference datasets and participant's results or between databases are explained.

3.4.1. Strain EQA2-S23-01

When applying the two reference datasets (Res_Ref and AMR_Ref) on the same sequences, certain systematic differences are observed (Table 6). The ResFinder database contains the *bla* gene variant *blaTEM-1B*, in contrast to the AMRFinderPlus database that will report the same variant as *blaTEM-1*. Additionally, the ResFinder database reports the gene *aac(6')-Iaa*, which is endogenous to *Salmonella* but is considered cryptic, as it does not contribute to aminoglycoside resistance (6). For this reason, it is not present in the AMRFinderPlus database. Three efflux pump genes, *emrD*, *mdsA* and *mdsB* are reported in the AMRFinderPlus dataset, but not in the Res_Ref dataset. These efflux genes are not present in ResFinder database.

Table 6. Genes reported in Salmonella strain EQA2-S23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Participants E05, E15 and E23 did not report any genes for this strain. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder											AMRFinderPlus +/- ResFinder									CARD +/- other				
	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	% concordance
ResFinder															*										
AMRFinderPlus															*										
CARD																									
aac(6')-Iaa	X	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	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	X	X		X	X	X	95

Three participants, E05, E15 and E23, did not report any genes for this strain, due to insufficient DNA concentration of this strain.
* Laboratory E21 used also BioNumerics plugin which has its own database

Depending on the applied database, eighteen out of twenty-two participants reported the gene *blaTEM-1B* (from ResFinder) or *blaTEM-1* (from AMRFinderPlus). Twenty-one out of twenty-two participants reported the presence of *sul2* and *tet(A)* genes.

The participants who queried more than one database were asked to give more details about their reporting strategy. Participants E17 and E14 stated that they reported all genes from all databases used. However, they did not report the *blaTEM-1* gene, which was likely given as output from the AMRFinderPlus database. Participant E02 used three databases: AMRFinderPlus, ResFinder and CARD and reported both *blaTEM-1* and *blaTEM-1B*, suggesting that all genes from all databases were reported. This participant noted that when using three databases, genes present in at least 2 databases were reported in a voting system, suggesting a custom approach.

Table 7. Point mutations (PMs) reported in Salmonella strain EQA2-S23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – PointFinder, Blue – AMRFinderPlus with or without PointFinder. Participants E05, E15 and E23 did not report any PMs for this strain. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder													AMRFinderPlus +/- PointFinder										
	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		
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

% concordance
32
100

Three participants, E05, E15 and E23, did not report any PMs for this strain, due to insufficient DNA concentration of this strain.
* Laboratory E21 used also BioNumerics plugin which has its own database

The labs that used the AMR_Ref dataset reported two point mutations, *ramR* T18P and *acrB* R717L, whereas only the *acrB* mutation was reported when the Res_Ref dataset was used (Table 7). As no point mutations in the *ramR* are present in the PointFinder database, this PM will not be called by this database alone.

The point mutation in *ramR* was reported by all except two participants that used AMRFinderPlus as the only database, by one participant (E21) using AMRFinderPlus as well as the BioNumerics plugin, and by four out of seven participants that used both AMRFinderPlus and ResFinder as databases.

3.4.2. Strain EQA2-S23-02

Similarly to strain EQA2-S23-01, several database-related differences between the reference datasets, Res_Ref and AMR_Ref, were observed for genes *aac(6')-Iaa*, *blaTEM-1/blaTEM-1B, emrD, mdsA* and *mdsB* (see paragraph 3.4.1 for details). Additionally, the genes *aadA1* and *aadA2* were reported in the reference set AMR_Ref, but in Res_Ref, only *aadA2* was reported (Table 8). The gene *ant(3")-Ia* was only reported in the dataset using ResFinder. Gene names *ant(3")-Ia* and *aadA2* can be used interchangeably (7).

Table 8. Genes reported in Salmonella strain EQA2-S23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder													AMRFinderPlus +/- ResFinder									CARD +/- other					
	Res_Ref	E31	E15	E06	E33	E35	E05	E03	E16	E07	E20	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02	E23	% concordance
ResFinder																	*											
AMRFinderPlus																	*											
CARD																												
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	88
aac(6')-Iaa	X	X	X			X	X		X	X	X	X	X								X	X			X	X		52
aadA1						X		X					X	X	X	X	X	X	X	X	X		X		X	X	X	56
aadA2	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	
ant(3'')-Ia	X	X	X	X					X	X	X	X						X			X	X						96
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	84
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	X	X	88
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	X	X	X	96
blaTEM-1												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		100
catA2	X	X	X	X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	88
dfrA12	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
emrD														X														0
floR	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	X	84
mdsA														X		X												4
mdsB														X		X												4
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	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	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	96
sul3	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	96
tet(M)	X	X	X	X	X	X		X	X	X	X	X	X	X		X	X	X	X	X	X	X		X	X	X	X	88

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

All 25 participants reported the genes *dfrA12*, *mph(A)* and *qnrS1*. Twenty-four out of 25 participants reported genes: *aadA2*, *aph(6)-Id*, *sul1* and *sul3*. Four out of 25 laboratories did not report the *floR* gene and three out of 25 laboratories failed to identify the *catA2* gene. Possible reasons for the absence of gene detection in some of the laboratories are discussed below based on two examples.

Two of the four laboratories mentioned above used thresholds for gene detection that were higher than the default. Laboratory E03 used thresholds of 99% identity and 100% coverage in ResFinder, where the defaults are 90% and 60%. In the EQA provider's reference dataset, Res_Ref, the percentage identity for the gene *floR* was 93% and for *catA2*, 96%, with coverage being 94% and 100%, respectively. This could explain why laboratory E03 missed the *catA2* gene when applying the high thresholds. It is worth noting that the percentage identity for these two genes was the same regardless of whether reads or assemblies were used in the ResFinder tool. Laboratory E08 used the CARD database with "Perfect" algorithm, which detects proteins with a 100% match to a reference sequence in CARD (5). When the EQA provider performed analysis using RGI in CARD for this DNA sample (data not shown), the gene "*catII* from Escherichia coli K-12" was detected with a 98% identity and 100% coverage (under "Strict" algorithm). This could potentially explain how the gene could have been missed by laboratory E08 if the laboratory only reported the "Perfect" hits.

Table 9. Point mutations (PMs) reported in Salmonella strain EQA2-S23-02. Reference dataset, Res_Ref is shaded grey. Participants are grouped based on database(s) used : Green – PointFinder, Blue – AMRFinderPlus with or without PointFinder. Participants E11, E10, E23, E05, E22, E24, E36 did not report any PMs for this strain. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder														AMRFinderPlus + Pointfinder					
	Res_Ref	E31	E15	E06	E08	E33	E13	E35	E03	E16	E07	E20	E28	E29	E21	E17	E02	E14	E19	% concordance
PointFinder																				
AMRFinderPlus															*					
parC T57S	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100

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

The point mutation *parC* T57S was reported only by the Res_Ref reference data set because this mutation is absent from the AMRFinderPlus database. This mutation was reported by 13 participants using PointFinder (Table 9). It was also reported by the participant E21, that used both AMRFinderPlus and the BioNumerics plugin, that queries its own database. There is currently no consensus whether this mutation contributes to quinolone resistance in *Salmonella* spp. (8).

3.4.3. Strain EQA2-S23-03

For strain EQA2-S23-03 there were also several database-related differences between the reference datasets, Res_Ref and AMR_Ref for the genes *aac(6')-Iaa, emrD, mdsA* and *mdsB* (as described in paragaph 3.4.1).

Table 10. Genes reported in Salmonella strain EQA2-S23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder														AMRFinderPlus +/- ResFinder										CARD +/- other					
	Res_Ref	E31	E15	E06	E33	E35	E05	E03	E16	E07	E20	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02	E23	% concordance
ResFinder																	*											
AMRFinderPlus																	*											
CARD																												
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	100
aac(6')-Iaa	X	X	X			X	X		X	X	X	X	X								X	X			X	X		52
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	X	X	X	92
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	X	X	92
blaCTX-M-55	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
emrD														X														0
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	92
mdsA														X		X												4
mdsB														X		X												4
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	96
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	100
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	X	X	96

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

All 25 participants reported the genes *aac(3)-IId*, *blaCTX-M-55* and *sul2* (Table 10). All but one participants reported the *qnrS1* and *tet(A)* gene. Two participants missed the *aph(3")-Ib*, *aph(6)-Id* and *floR* genes.

Table 11. Point mutations (PMs) reported in Salmonella strain EQA2-S23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – PointFinder, Blue – AMRFinderPlus with or without PointFinder. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder															AMRFinderPlus +/- Pointfinder												
	Res_Ref	E31	E15	E06	E08	E33	E13	E35	E05	E03	E16	E07	E20	E28	E29	AMR_Ref	E11	E21	E24	E17	E10	E02	E36	E23	E22	E14	E19	% concordance
PointFinder																												
AMRFinderPlus																		*										
gyrA D87N	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
parC T57S	X	X	X	X		X	X	X		X		X	X	X	X			X		X	X	X				X	X	68

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

The point mutation *gyrA* D87N was reported for strain EQA2-S23-03 in both reference datasets, whereas the mutation *parC* T57S was reported only in ResFinder dataset, due to the database differences mentioned in paragraph 3.4.2.

All 25 participants reported the substitution in *gyrA*, whereas 11 out of 14 participants using ResFinder reported the substitution in *parC* (Table 11). The same was true for five out of eight participants using both ResFinder and AMRFinderPlus as databases.

4. *Campylobacter* results

4.1. Quality metrics for all *Campylobacter* strains

4.1.1. DNA concentration and it's evaluation

In total, 25 participants reported the DNA concentrations including the applied methods for all three *Campylobacter* samples (Table S 4-S 6). The participants provided varying DNA concentration values for the three samples (Figure 8, Figure S 2).

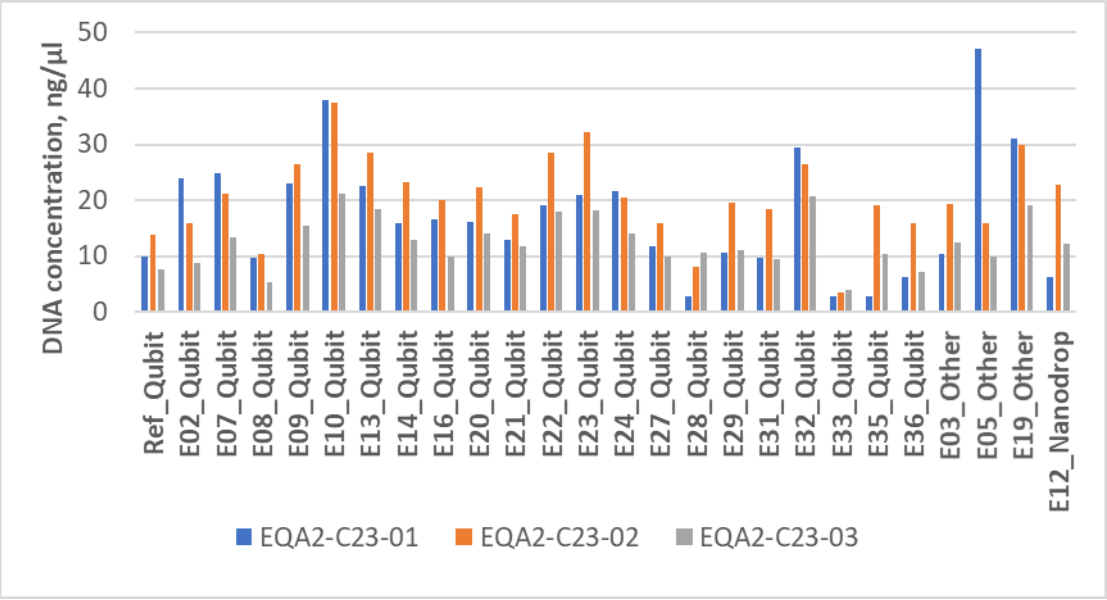

Figure 8. The distribution of Campylobacter DNA concentrations among 25 participants with the indicated methods used.

4.1.2. Evaluation of genome assembly

Twenty-four laboratories sequenced *Campylobacter* using Illumina technology and one laboratory used Ion Torrent technology. The quality parameters of genome assemblies generated by participants using assembly tools of their choice were evaluated. The parameters were evaluated according to the thresholds recommended in the suggested WGS protocol (https://www.fwdamr-reflabcap.eu/resources/reflabcap-protocols-and-guidelines): genome size of 1.5 Mb -1.9 Mb, N50 of >30 000 bp, and contig number of <500 contigs. The laboratories were expected to use their usual pipelines for evaluating the quality of assemblies.

To assemble *Campylobacter* genomes, most participants used SPAdes (n=19) with or without contig filtering which varied between 200 to 500 bp (Figure 9).

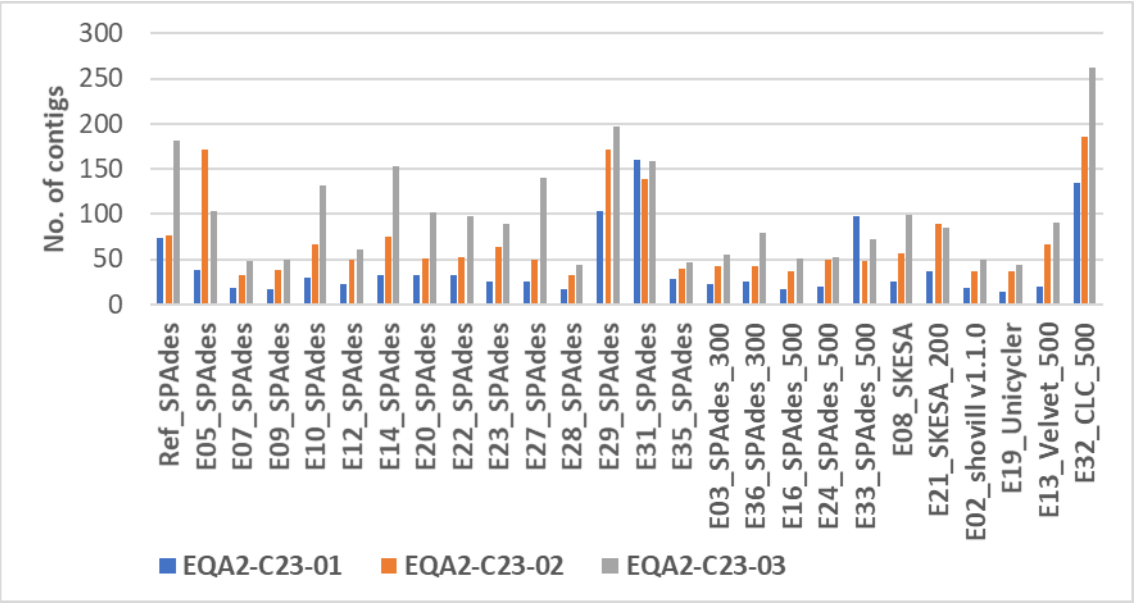

Figure 9. The distribution of number of contigs among 25 participants for three Campylobacter samples. The horizontal axis labels indicate the ID of the participant, the tool used for genome assembly, and if contig length-based filtering was used it also indicates the length used.

All the laboratories reported the genomes size, N50 and contig number within the recommended limits for *Campylobacter* for all three samples (Table S 4-S 6). The overview of these parameters for each sample is provided below.

EQA2-C23-01

The participants reported genome size between 1.7 and 1.7 Mb (av.=1.7), N50 between 30 103 bp and 263 055 bp (av. 192 285), and contigs numbers between 14 and 160 (av.=40).

EQA2-C23-02

For this strain the participants reported genome size between 1.7 to 1.8 Mb (av.=1.8), N50 between 18 850 bp to 132 930 bp (av. 104 738), and contig numbers between 33 and 186 (av.=69).

EQA2-C23-03

For the third strain, the participants reported genome size between 1.7 and 1.8 Mb (av.=1.7), N50 between 16 188 bp and 162 605 bp (av. 128 777), and contig numbers between 33 and 186 (av.=69).

4.2. AMR gene and PMs detection methods used

4.2.1. Tools and databases used for gene detection

All 25 participants reported the tools, the databases, types of files used as inputs, the thresholds for sequence coverage, sequence identity and AMR gene reporting strategy when more than one database was used. Overall, 21 unique combinations of tools/databases/inputs/thresholds/gene reporting strategies were used by 25 participants (Table S 9).

The ResFinder tool was used by 21 participants, AMRFinderPlus was used by 11 participants, RGI was used by three participants, and BLAST by two participants. ARIBA, ABRicate, and CLC Genomics Workbench were used by one participant each (Figure 10).

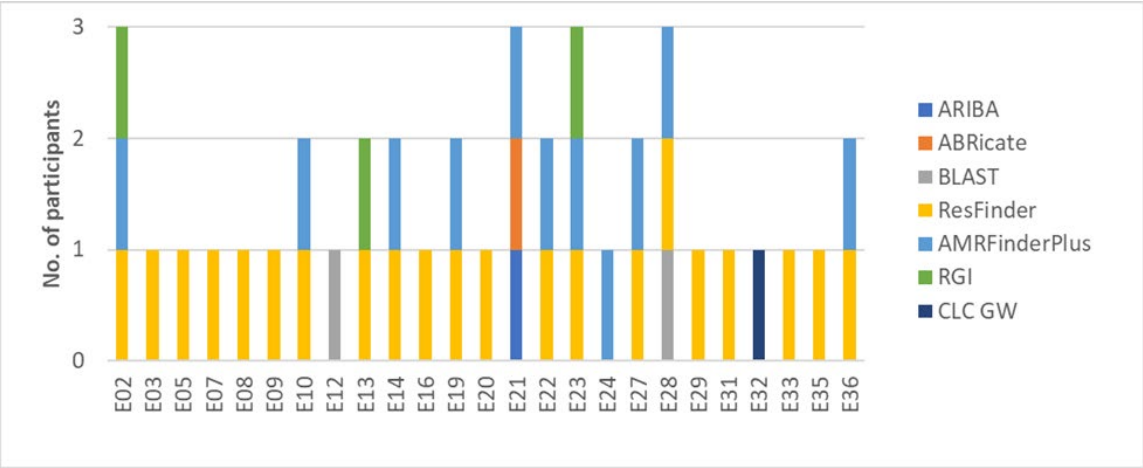

Figure 10. An overview of tools used by 25 participants for AMR gene detection in Campylobacter.

The ResFinder database was used by 23 participants, AMRFinderPlus database by 11 participants, CARD database by four participants, and QIAGEN Microbial Insight – Antimicrobial Resistance (QMI-AR) and a combination of NCBI and own database by one participant each. The participants which used more than one database also indicated how they reported AMR genes. Four participants reported a subset of genes based on experience/knowledge/literature, three other participants reported a consensus list of genes (common genes present in all databases used), and two participants reported all genes from all databases. Additionally, participant E02 indicated that they used a voting system, participant E13 that they also included a literature review, and participant E27 that they did not report genes which do not confer resistance (Figure 11).

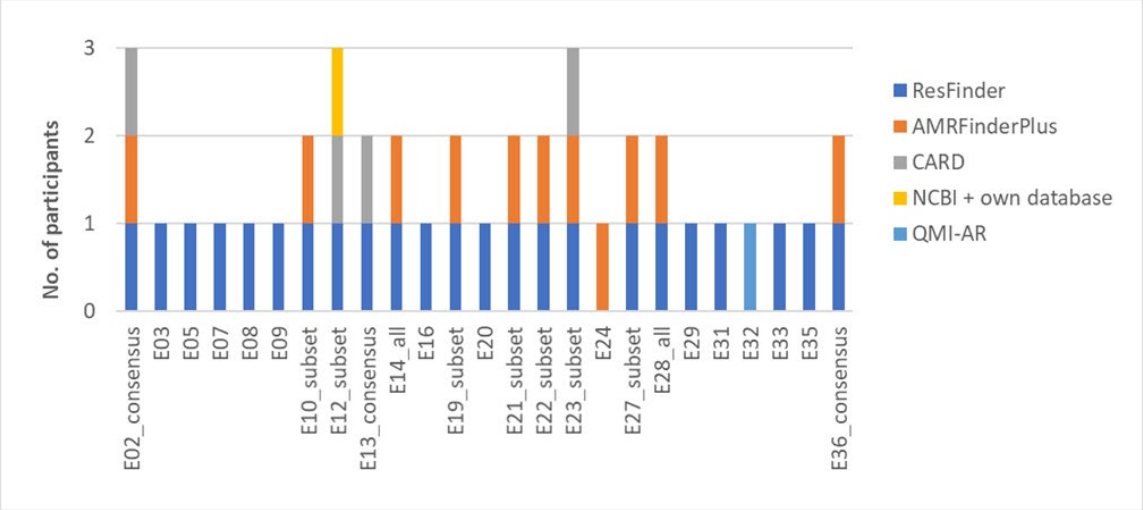

Figure 11. An overview of databases used by 25 participants for AMR gene detection in Campylobacter. The horizontal labels indicate the participant ID and how they reported AMR genes in cases when more than one database was used: "consensus" – common genes present in all databases used, "subset" – based on experience/knowledge/literature, "all" – genes from all databases.

4.2.2. Tools and databases used for point mutations detection

All 25 participants reported the tools, the databases, and the inputs that they used for point mutations detection in *Campylobacter* and what approach they used for reporting when more than one database was used. Overall, 13 unique combinations of tools/databases/inputs/reporting strategies were used by 25 participants (Table S 10).

PointFinder was the most commonly used tool, applied by 21 participants, AMRFinderPlus was used by 10 participants, whereas Blast and CLC Genomics Workbench (CLC GW) was used by one participant each (Figure 12).

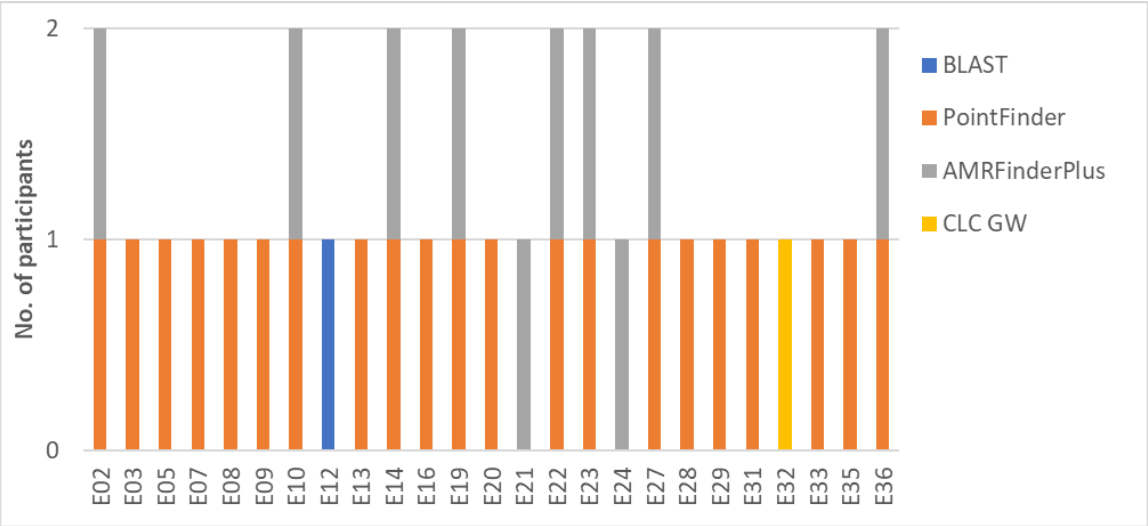

Figure 12. An overview of tools used by 25 participants for point mutations detection in Campylobacter.

The most used databases were ResFinder and AMRFinderPlus, that were queried by 22 and 11 participants, respectively. The participants which used more than one database also indicated how they reported point mutations (Figure 13).

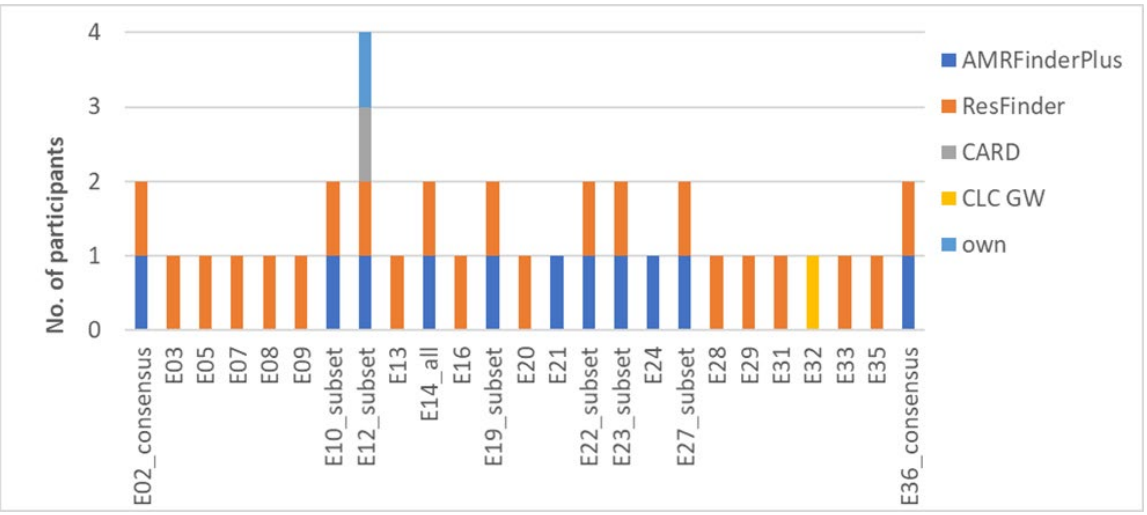

Figure 13. An overview of databases used by 25 participants for point mutations detection in Campylobacter. The horizontal labels indicate the participant ID and how they reported point mutations in cases when more than one database was used : "all" - all point mutations from all databases, "subset" - a subset of point mutations based on experience/knowledge/literature, "consensus" - a consensus list of point mutations (common point mutations present in all databases used).

4.3. Species and STs reported

4.3.1. Methods used for species identification and results

Nineteen participants used only one tool/software for *Campylobacter* species detection, five participants used two, and one participant three. The most commonly used tools were KmerFinder and Kraken, used by 13 and five participants, respectively. Remaining tools/softwares were used by 1-2 participants (Figure 14).

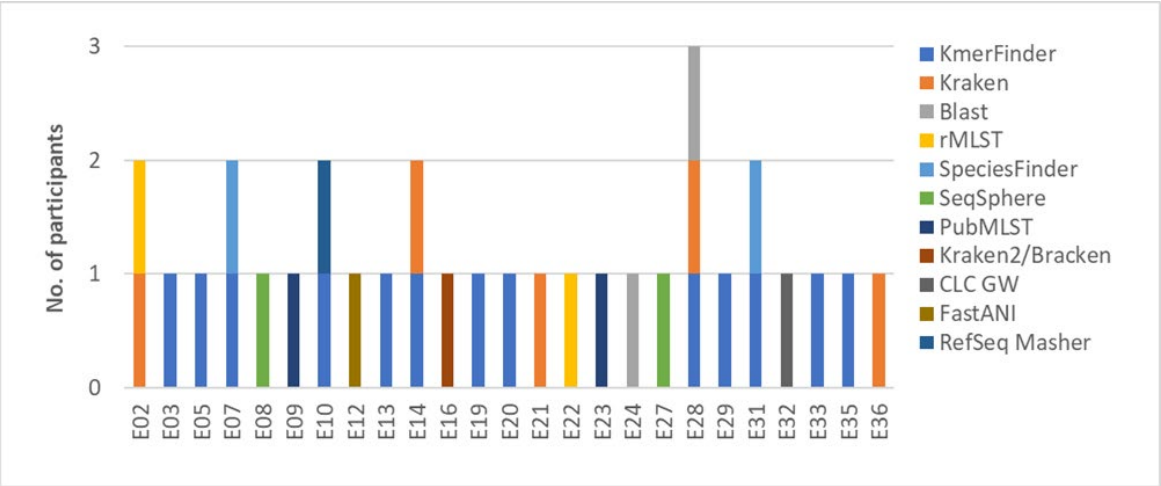

Figure 14. An overview of the tools/softwares used by 25 participants for Campylobacter species identification.

All 25 participants reported the *Campylobacter* species correctly for all three strains (Table S 14).

4.3.2. Methods used for MLST detection and STs

Twelve participants used MLST2.0 scheme available as CGE tools, four participants used tsemann MLST scheme, four participants used PubMLST, three participants used Ridom SeqSphere+, one CLC genomics Workbench and one BioNumerics for 7 gene MLST typing of *Campylobacter*. All 25 participants reported the ST for *Campylobacter* EQA2-C23-01 and EQA2-WGS-02 correctly. The same was true for isolate EQA2-C23-03, with the exception of participant E05, who did not report the ST (Table S 13).

4.4. AMR genes and PMs reported for *Campylobacter* strains

In this section, strain-specific tables are used to demonstrate the genes identified by EQA provider using two different tools and databases, namely Res_Ref and AMR_Ref, as well as by the participants. The letter “X” indicates gene detection.

Results from the participants and the reference datasets are divided into three categories based on which database was used. The green (ResFinder) category contains laboratories that only used the ResFinder database and the PointFinder for point mutations. Participants that used the AMRFinderPlus database, either alone or with ResFinder, are grouped in the blue category (AMRFinderPlus +/- ResFinder). When relevant, the third, yellow, category is applied, grouping laboratories that used CARD, either alone or in combination with any other databases. In cases where a participant used a database different from ResFinder, AMRFinderPlus or CARD, it is marked with an asterisk in the tables and the database name is stated in the table footer.

For each gene and PM table, we calculated the concordance of the reported results among the participants. This number is expressed in percentage of total number of laboratories that reported genes or point mutations for a given strain. That number varies between strains, from 18 to 25, as not all participants wished to report point mutations.

4.4.1. Strain EQA2-C23-01

The *aadE*-Cc gene was identified in both reference datasets, but the *tetO* gene was not detected in the Res_Ref dataset (Table 12). When the EQA provider compared an

output from ResFinder based on paired Illumina reads with an output based on SPADes assemblies, the *tetO* gene was only detected when an assembly was used.

Table 12. Genes reported in Campylobacter strain EQA2-C23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder													AMRFinder +/- ResFinder									CARD +/- others					
	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						*																						
AMRFinderPlus																												
CARD																												
Other database																							**				***	
aadE-Cc	X	X	X	X	X	X	X		X		X	X	X	X	X	X	X	X		X	X	X		X	X	X		80
tet(O)		X	X	X	X	X		X					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	80

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

Among the participants, the *tetO* gene was reported by all 8 participants using AMRFinderPlus as one of the databases, as well as by all 5 participants using CARD or other databases, but only by 7 out of 12 participants using ResFinder. In the case of laboratories E05, E16, E07, E20 and E28 that did not report the gene, all used reads as input, with two labs, E05 and E07 using assemblies in addition to reads. Six out of seven laboratories that did report the *tetO* gene, used assemblies as input and out of those, laboratories E03 and E31 used reads additionally. Laboratory E35 reported this gene despite using reads only.

The *aadE-Cc* gene was not reported by 2 out of 12 laboratories using ResFinder, by 1 out of 8 laboratories using AMRFinderPlus and 2 out of 5 laboratories using CARD or other databases.

Table 13. Point mutations (PMs) reported in Campylobacter strain EQA2-C23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – PointFinder, Blue – AMRFinderPlus with or without PointFinder. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder															AMRFinderPlus +/- PointFinder											O	
	Res_Ref	E28	E31	E08	E33	E13	E35	E09	E05	E03	E16	E07	E20	E29	E32	AMR_Ref	E21	E24	E27	E10	E02	E36	E23	E22	E14	E19	E12	% concordance
PointFinder								*																				
AMRFinderPlus																												
Other database																											**	
23S A2075G	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				X	X	X	X	X	X	X	X	X	92
gyrA T86I										X			X			X	X	X		X	X	X	X		X		X	
gyrA_2 T86I	X	X	X	X	X	X	X	X	X			X		X					X					X		X		92
50S L22 A103V																X	X	X	X	X		X	X	X	X			32

BioNumerics 8.1
** ResFinder, Card, NCBI, home-made database

In point mutations reporting (Table 13), differences between reference datasets were observed. The mutation *23S* A2075G was reported in the Res_Ref dataset, but not in the AMR_Ref. Conversely, mutation *50S* L22 A103V was reported in AMR_Ref dataset, but not in Res_Ref. The *gyrA* T86I substitution was reported in both datasets but either as *gyrA_2* (Res_Ref) or *gyrA* (AMRFinderPlus). These differences are related to the presence and absence of these point mutations in the respective databases: the mutation in 23S is only present in the ResFinder database and the mutation in 50S is only present in the AMRFinderPlus database. The latter mutation is equally common among resistant and sensitive isolates and, for this reason, not necessarily resulting in a phenotype (9).

The mutation in *23S* gene was reported by 23 out of 25 participants. It was reported by all 14 participants using solely ResFinder and by all 8 participants that used both AMRFinderPlus and ResFinder. Two participants using AMRFinderPlus alone, did not report this mutation, which is in accordance with the AMR_Ref dataset. None of the participants using ResFinder reported the *50S* L22 A103V mutation but it was reported by 8 out of 10 participants using AMRFinderPlus with or without ResFinder. One participant using a combination of databases such as ResFinder, CARD, NCBI and a home-made database did not report this mutation, but reported the other two. This participant's approach to multiple databases was to report a subset of genes based on experience, knowledge or literature.

4.4.2. Strain EQA2-C23-02

In strain EQA2-C23-02, quite a few differences were observed between the two reference datasets (Table 14). Only three genes were reported in the same way in both reference datasets: *aph(2'')-If*, *blaOXA-193* and *ermB*. Certain genes were reported by both reference datasets but in the form of different variants, for example *cat / catA13*, *aac(6')-aph(2'') / aac(6')-Ie/aph(2'')-Ia* or, in one case, in the form of synonyms : *ant(6)-Ia* and *aadE*. The genes *aad9* and *tet(O/M/O)* gene was reported only in the AMR_Ref dataset.

Table 14. Genes reported in Campylobacter strain EQA2-C23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder													AMRFinder +/- ResFinder									CARD +/- others					
	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						*																						
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		X			X		
ant(6)-Ia	X	X	X	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		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	X	X	100
blaOXA-193	X	X			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		X	X	X		X	
catA13														X	X	X	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	X	X	X	100
tet(O)		X				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).

Similar differences as those observed for the different reference datasets, were also seen among participants, depending on which databases were queried. The gene *aadE* was reported by 6 out of 8 participants using AMRFinderPlus database (alone or with another database). All 8 participants using AMRFinderPlus with or without ResFinder, as well as all participants using CARD alone or in combination with other databases reported the gene *aph(3')-IIIa*. The same gene was reported by all Resfinder users, either in the subvariant form *aph(3')-IIIa* (5 out of 12 participants) or in the variant form *aph(3')-III* (7 out of 12 participants).

For the cluster gene *aac(6')-aph(2'')* some reporting bias can be observed. All 8 AMRFinderPlus users reported the gene in the variant form *aac(6')-Ie/aph(2'')-Ia*, just as it is present in the AMRFinderPlus database. The same gene is present only as *aac(6')-aph(2'')* in ResFinder database and yet it was reported as *aac(6')-Ie/aph(2'')-Ia* by 6 out of

12 participants that used ResFinder as the only database. The provider's guess is that this is due to the way the reporting scheme has been set up, where only the gene *aac(6')-Ie/aph(2")-Ia* was present as a choice on the gene list. All the participants reported the chloramphenicol resistance gene, *catA*, either as *catA* (ResFinder) or *catA13* (AMRFinderPlus and other databases).

None of the participants using ResFinder reported the mosaic *tet(O/M/O)* gene and only 3 participants using the AMRFinderPlus reported it. From the AMR_Ref dataset analysis, it is apparent that this gene is detected with 68% coverage and 99% identity.

Table 15. Point mutations (PMs) reported in Campylobacter strain EQA2-C23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without PointFinder. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder															AMRFinderPlus +/- PointFinder											O	% concordance
	Res_Ref	E28	E31	E08	E33	E13	E35	E09	E05	E03	E16	E07	E20	E29	E32	AMR_Ref	E21	E24	E27	E10	E02	E36	E23	E22	E14	E19	E12	
PointFinder								*																				
AMRFinderPlus																												
Other database																											**	
gyrA T86I										X			X			X	X	X		X	X	X	X		X		X	
gyrA_2 T86I	X	X	X	X	X	X	X	X	X			X		X					X					X		X		92

* BioNumerics 8.1
** Card, Ncbi, home-made database

Strain EQA2-C23-02 has one point mutation responsible for ciprofloxacin resistance, *gyrA* T86I (Table 15). This mutation was reported mostly as *gyrA_2* by the ResFinder users (as well as in the reference dataset) and as *gyrA* among AMRFinderPlus users. This is related to how the genes are present in the databases. The gene variant *gyrA_2* has been present in the PointFinder database since June 2022.

4.4.3. Strain EQA2-C23-03

For strain EQA2-C23-03 (Table 16), there was a complete agreement between the two reference datasets. Most participants reported the three antibiotic resistance genes: *aadE-Cc*, *blaOXA-489* and *tet(O)*, with the exception of three laboratories that did not report *aadE-Cc* and one laboratory that did not report the *blaOXA-489*.

Table 16. Genes reported in Campylobacter strain EQA2-C23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	ResFinder														AMRFinder +/- ResFinder									CARD +/- others					% concordance
	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	
ResFinder						*																						
AMRFinderPlus																												
CARD																												
Other database																							**				***	
aadE-Cc	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		X	X	X		88
blaOXA-489	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		96
tet(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	X	X	X	100

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

Two point mutations are present in this strain: *23S* A2075G, responsible for erythromycin resistance and *gyrA* T86I, responsible for ciprofloxacin resistance (Table 17). The first mutation was detected by all but one of the 25 laboratories. The mutation in *gyrA* was missed by 2 out of 25 laboratories.

Table 17. Point mutations (PMs) reported in Campylobacter strain EQA2-C23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder. Percentage concordance is based on following scale : darkest orange colour : 100% concordance among participants, lighter orange colour : 90-99% concordance, lightest orange colour : 80-89% concordance. Concordance lower than 80% is without colour.

	PointFinder															AMRFinderPlus +/- PointFinder											O	
	Res_Ref	E28	E31	E08	E33	E13	E35	E09	E05	E03	E16	E07	E20	E29	E32	AMR_Ref	E21	E24	E27	E10	E02	E36	E23	E22	E14	E19	E12	% concordance
PointFinder								*																				
AMRFinderPlus																												
Other database																											**	
23S A2075G	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	96
gyrA T86I										X			X			X	X	X		X	X	X	X		X		X	
gyrA_2 T86I	X	X	X	X	X	X	X	X	X			X		X					X					X		X		92

* BioNumerics 8.1
** Card, Ncbi, home-made database

5. Conclusions

The EQA2-WGS-AMR, organised by Statens Serum Institut in collaboration with the Technical University of Denmark, is the second exercise in the FWDAMR-RefLabCap project in series of three, spanning over three years. The purpose of this and the following third EQA is to support the implementation and further development of the standard protocol for AMR gene detection and clone identification (https://www.fwdamr-reflabcap.eu/resources/reflabcap-protocols-and-guidelines).

Thirty-four laboratories accepted the invitation to participate in the EQA. Three laboratories did not submit results due to problems with Illumina machine or lack of laboratory capacity to perform sequencing. Two laboratories from one country (one laboratory for each pathogen) submitted their results jointly, resulting in total of 30 participants that submitted results. This is an increase compared to the last year, where 25 participants submitted the results in EQA1-WGS-AMR.

In the EQA2-WGS-AMR, dried DNA samples were shipped to the participants. Two participants reported too low concentration for *Salmonella* strain EQA2-S23-01 and one participant reported measuring too low concentrations for all three *Salmonella* samples. Therefore, new vials with dried DNA were sent to those three laboratories per request. Two laboratories could subsequently perform sequencing and submit the results, but one laboratory reported unsatisfactory quality parameters on strain EQA2-S23-01 again and was unable to submit results for this strain. Additionally, two other laboratories that measured too low concentrations did not ask for an extra vial and did not submit results for this isolate. As noted in section 3.1.2, three participants reported poor DNA quality metrics for *Salmonella* strain EQA2-S23-01, but they still reported the genes that were in concordance with the reference dataset.

This round of EQA was the first one where dried DNA samples were distributed. Participants received a detailed instruction together with the DNA samples on how to properly rehydrate the DNA. One of the crucial steps of rehydration was to incubate the dried DNA with nuclease-free water for 15 minutes to allow sufficient rehydration, as well as pipetting the water up and down, while flushing the sides of the Eppendorf tube in order to facilitate the rehydration of any DNA that might be attached to the side of the Eppendorf tube. Conducting these steps without proper caution could have an effect on the amount of the DNA in the final solution. The instruction for rehydration of DNA samples will be highlighted even more in the final EQA round in order to ensure a uniform concentration of the DNA samples.

Similarly to what was observed in the EQA1-WGS-AMR, the discrepancies in the genes detected by the participants were partially due to presence or absence of different variants or different nomenclature of the same gene, in different databases. This is of importance when comparing outputs from different databases.

The identity and coverage thresholds applied for gene identification varied among the participants. Four participants applied thresholds that were higher than default settings for *Salmonella* (see laboratories E03, E08, E24 and E23 in Table S 7). The participants E03 and E08 did not report two genes, *floR* and *catA2* in the DNA sample EQA2-S23-02 (paragraph 3.4.2) and this was likely due to their high thresholds used.

The phenotypic testing results made available by the EQA provider (Table 1 and Table 2) will enable the participants to correlate their genotypic results (predictions) with the provided phenotypes for the tested antimicrobials. However, due to limitations in the number of antimicrobials tested, this correlation cannot be established for all reported genotypes. In some cases, the genetic determinant to which a phenotype can be attributed cannot be identified. This was the case with strain EQA2-S23-01, where colistin phenotype was observed by the EQA provider despite the absence of *mcr* genes in this strain. The mechanism of colistin resistance of this strain is not clear but it might involve, for example, the efflux pumps (10) or overexpression of a two-component system regulator, PmrD (11). It is also worth noting that strain EQA2-S23-01 is of serotype Dublin, which has been shown to be less susceptible to colistin than the other serotypes (12).

Campylobacter strain EQA2-C23-02 harbours the *cat* gene but the phenotypic expression of this gene was not investigated by the EQA provider as the strain was not tested for chloramphenicol susceptibility (see paragraph 2.2). Strains EQA2-C23-02 and EQA2-C23-03 both harbour a *blaOXA* gene. That is, however, not reflected in phenotype either, due to absence of beta-lactams among the antimicrobials tested. In general, correlating the presence of *blaOXA* genes in *Campylobacter* to phenotypic resistance is complicated. It has been shown that the presence of a G to T mutation in the promoter of the *blaOXA-61* gene conferred resistance to ampicillin and not the presence of the gene alone (13, 14).

Participants in EQA2-WGS-AMR applied 22 and 21 unique combinations of tools, databases, inputs and thresholds for gene detection in *Salmonella* and *Campylobacter*, respectively (Table S 7 and Table S 9). For point mutation detection, 14 and 13 unique combinations were used (Table S 8 and Table S 10) for *Salmonella* and *Campylobacter*, respectively. In general, even when comparing genes and point mutations reported by the participants using identical combinations of tools and databases, we found the participants did not report the same genes. The reason for this could be the manual curation applied to the results obtained from the different tools.

The majority of participants used one database for both *Salmonella* and *Campylobacter* and both for gene detection and point mutation identification. Those participants that used more than one database, applied a variety of curation methods to the obtained results. Regardless of the approach used, the genes and point mutations reported were generally concordant and only minor differences were reported by the participants.

In general, this round of EQA was successful and provided comparable results from the participants using a variety of approaches to identify genes and point mutations. The provider of this EQA will use the experiences gathered to improve the next year's EQA, for example by including the possibility to report the predicted phenotype for those antimicrobials that were phenotypically tested by the EQA provider.

6. References

1. European Centre for Disease Prevention and Control. 2016. EU protocol for harmonised monitoring of antimicrobial resistance in human Salmonella and Campylobacter isolates - June 2016. Stockholm.

2. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644.

3. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. 2017. PointFinder: A novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J Antimicrob Chemother 72:2764–2768.

4. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, Hoffmann M, Pettengill JB, Prasad AB, Tillman GE, Tyson GH, Klimke W. 2021. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci Rep 11:1–9.

5. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

6. Magnet S, Courvalin P, Lambert T. 1999. Activation of the cryptic aac(6')-Iy aminoglycoside resistance gene of Salmonella by a chromosomal deletion generating a transcriptional fusion. J Bacteriol 181:6650–6655.

7. Frye JG, Jackson CR. 2013. Genetic mechanisms of antimicrobial resistance identified in Salmonella enterica, Escherichia coli, and Enteroccocus spp. isolated from U.S. food animals. Front Microbiol 4:1–22.

8. Chang MX, Zhang JF, Sun YH, Li RS, Lin XL, Yang L, Webber MA, Jiang HX. 2021. Contribution of Different Mechanisms to Ciprofloxacin Resistance in Salmonella spp. Front Microbiol 12.

9. Dahl LG, Joensen KG, Østerlund MT, Kiil K, Nielsen EM. 2021. Prediction of antimicrobial resistance in clinical Campylobacter jejuni isolates from whole-genome sequencing data. Eur J Clin Microbiol Infect Dis 40:673–682.

10. Baron S, Hadjadj L, Rolain JM, Olaitan AO. 2016. Molecular mechanisms of polymyxin resistance: knowns and unknowns. Int J Antimicrob Agents 48:583–591.

11. Hjort K, Nicoloff H, Andersson DI. 2016. Unstable tandem gene amplification generates heteroresistance (variation in resistance within a population) to colistin in Salmonella enterica. Mol Microbiol 102:274–289.

12. Agersø Y, Torpdahl M, Zachariasen C, Seyfarth A, Hammerum AM, Nielsen EM. 2012. Tentative colistin epidemiological cut-off value for Salmonella spp. Foodborne Pathog Dis 9:367–369.

13. Zeng X, Brown S, Gillespie B, Lin J. 2014. A single nucleotide in the promoter region modulates the expression of the β-lactamase OXA-61 in campylobacter jejuni. J Antimicrob Chemother 69:1215–1223.

14. Ocejo M, Oporto B, Lavín JL, Hurtado A. 2021. Whole genome-based characterisation of antimicrobial resistance and genetic diversity in Campylobacter jejuni and Campylobacter coli from ruminants. Sci Rep 11:8998.

Annex A

6.1. Supplementary material on QC metrics for both organisms

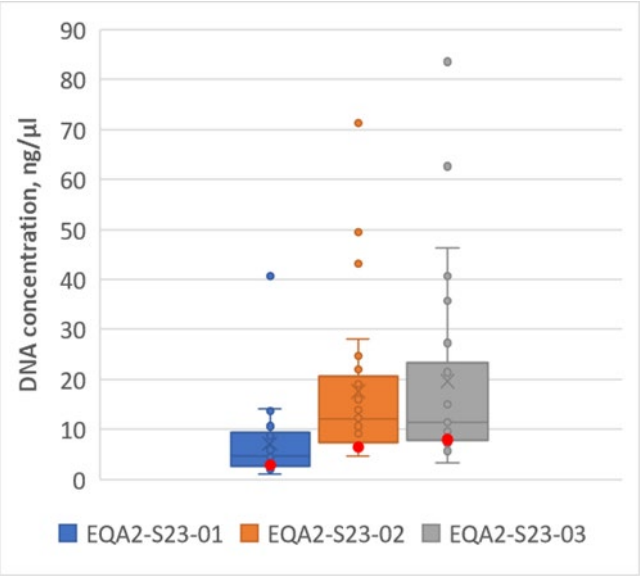

Figure S 1. Distribution of Salmonella DNA concentration among 25 participating laboratories. EQA provider's values for each sample are marked as a red dot.

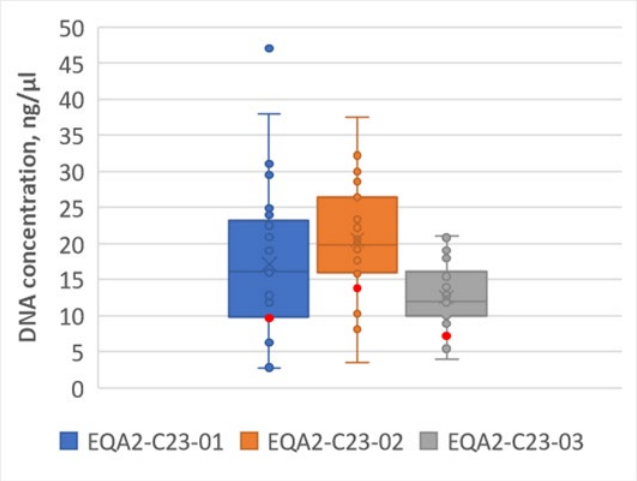

Figure S 2. Distribution of Campylobacter DNA concentration among 25 participating laboratories. EQA provider's values for each sample are marked as a red dot.

Table S 1. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-S23-01. In red, results that did not pass the recommended threshold according to agreed WGS protocol. NR – results were not reported.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method*	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	2.7	SPAdes	4.97	84878	175
E02	Qubit	8.8	shovill v1.1.0	4.97	299397	37
E03	Clariostar	2.9	SPAdes_300	4.97	101678	85
E05	Quantus	1.1	SPAdes	1.02	454	2214
E06	Qubit&Nanodrop	13.7	SPAdes	4.97	116997	89
E07	Qubit	9.3	SPAdes	4.96	47033	190
E08	Qubit	40.6	SKESA	4.90	61410	195
E10	Qubit	11.5	SPAdes	4.90	221706	56
E11	Qubit	14.2	SPAdes_500	4.97	405889	28
E13	Qubit	1.9	Velvet_500	5.00	68935	258
E14	Qubit	2.6	SPAdes_200	4.70	1443	4687
E15	Qubit	2.4	NR	2.70	351	7654
E16	Qubit	10.6	SPAdes_500	4.96	420012	41
E17	Qubit	4.5	SPAdes_200	4.98	284311	75
E19	Quantus	7.1	Unicycler	4.90	71250	113
E20	Qubit	4.6	SPAdes_200	4.97	179990	77
E21	Qubit	2.2	SKESA/SPADes_300	4.90	107532	130
E22	Qubit	2.0	SPAdes	5.03	6503	1813
E23	Qubit	10.0	SPAdes	0.10	270	482
E24	Qubit	6.2	SPAdes_500	4.99	308630	207
E28	Qubit	4.1	SPAdes	4.96	332347	35
E29	Qubit	3.6	SPAdes	5.05	223582	228
E31	Qubit	5.9	SPAdes	4.99	220197	156
E33	Qubit	3.2	SPAdes_500	4.97	260404	48
E35	Qubit	3.7	SPAdes	4.62	27839	953
E36	Qubit	6.3	SPAdes_300	4.97	228045	46

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

Table S 2. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-S23-02. In red, results that did not pass the recommended threshold according to agreed WGS protocol. NR – results were not reported.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method*	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	6.5	SPAdes	5.02	63808	236
E02	Qubit	49.4	shovill v1.1.0	5.02	287106	43
E03	Clariostar	7.4	SPAdes_300	5.02	90467	115
E05	Quantus	6.6	SPAdes	4.92	14321	728
E06	Qubit&Nanodrop	43.2	SPAdes	5.02	413753	50
E07	Qubit	14.2	SPAdes	5.02	398149	46
E08	Qubit	28.1	SKESA	4.90	31780	308
E10	Qubit	21.9	SPAdes	5.00	308094	53
E11	Qubit	71.4	SPAdes_500	5.02	437353	36
E13	Qubit	13.8	Velvet_500	5.00	223205	116
E14	Qubit	10.0	SPAdes_200	5.00	437012	112
E15	Qubit	4.7	NR	5.00	68105	214
E16	Qubit	19.0	SPAdes_500	5.01	143369	75
E17	Qubit	12.3	SPAdes_200	5.03	223715	68
E19	Quantus	14.0	Unicycler	5.00	287116	63
E20	Qubit	10.7	SPAdes_200	5.01	169321	79
E21	Qubit	11.9	SKESA/SPADes_300	5.00	85108	136
E22	Qubit	20.2	SPAdes	5.02	287387	76
E23	Qubit	24.7	SPAdes	5.00	113077	90
E24	Qubit	10.6	SPAdes_500	5.02	227359	411
E28	Qubit	9.2	SPAdes	5.01	326172	45
E29	Qubit	7.0	SPAdes	5.04	414039	104
E31	Qubit	10.6	SPAdes	5.01	280766	165
E33	Qubit	7.2	SPAdes_500	5.02	223254	52
E35	Qubit	6.9	SPAdes	4.99	196896	115
E36	Qubit	16.0	SPAdes_300	5.03	189983	69

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

Table S 3. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-S23-03. In red, results that did not pass the recommended threshold according to agreed WGS protocol. NR – results were not reported.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method*	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	7.9	SPAdes	4.94	97830	274
E02	Qubit	35.7	shovill v1.1.0	4.92	232475	69
E03	Clariostar	8.5	SPAdes_300	4.93	105303	111
E05	Quantus	3.4	SPAdes	4.93	134490	101
E06	Qubit&Nanodrop	40.7	SPAdes	4.92	107178	118
E07	Qubit	10.8	SPAdes	4.93	247483	73
E08	Qubit	27.3	SKESA	4.80	32046	378
E10	Qubit	21.4	SPAdes	4.90	190632	83
E11	Qubit	83.6	SPAdes_500	4.92	254048	48
E13	Qubit	12.7	Velvet_500	4.90	181493	140
E14	Qubit	5.7	SPAdes_200	4.90	190632	201
E15	Qubit	7.8	NR	4.90	63348	238
E16	Qubit	46.4	SPAdes_500	4.92	144255	100
E17	Qubit	11.6	SPAdes_200	4.94	121558	171
E19	Quantus	12.0	Unicycler	4.90	222290	94
E20	Qubit	11.4	SPAdes_200	4.92	118831	111
E21	Qubit	22.1	SKESA/SPADes_300	4.90	79341	181
E22	Qubit	16.3	SPAdes	4.94	181476	133
E23	Qubit	62.6	SPAdes	4.90	148816	118
E24	Qubit	8.9	SPAdes_500	4.95	157826	224
E28	Qubit	6.0	SPAdes	4.92	247225	56
E29	Qubit	9.6	SPAdes	4.96	109458	209
E31	Qubit	15.0	SPAdes	4.95	151292	219
E33	Qubit	5.8	SPAdes_500	4.92	100308	126
E35	Qubit	8.7	SPAdes	4.89	148448	147
E36	Qubit	7.1	SPAdes_300	4.94	239409	94

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

Table S 4. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-C23-01.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	9.8	SPAdes	1.68	144979	73
E02	Qubit	23.9	shovill v1.1.0	1.68	225045	18
E03	Other	10.4	SPADes_300	1.68	171101	23
E05	Other	47.0	SPADes	1.67	109952	38
E07	Qubit	24.9	SPADes	1.68	259160	18
E08	Qubit	9.8	SKESA	1.70	173958	25
E09	Qubit	23.0	SPADes	1.67	263054	16
E10	Qubit	38.0	SPADes	1.68	262847	29
E12	Nanodrop	6.3	SPADes	1.68	72942	23
E13	Qubit	22.5	Velvet_500	1.70	263000	20
E14	Qubit	15.9	SPADes	1.70	176284	33
E16	Qubit	16.5	SPADes_500	1.68	225253	17
E19	Other	31.0	Unicycler	1.70	262807	14
E20	Qubit	16.2	SPADes	1.67	166329	32
E21	Qubit	12.8	SKESA_200	1.67	113524	37
E22	Qubit	19.0	SPADes	1.68	224693	32
E23	Qubit	20.9	SPADes	1.70	173958	25
E24	Qubit	21.7	SPADes_500	1.68	263055	19
E27	Qubit	11.8	SPADes	1.68	262948	25
E28	Qubit	2.8	SPADes	1.67	174851	17
E29	Qubit	10.5	SPADes	1.72	263054	104
E31	Qubit	9.6	SPADes	1.71	224655	160
E32	Qubit	29.5	CLC_500	1.68	42952	134
E33	Qubit	2.7	SPADes_500	1.67	30103	98
E35	Qubit	2.8	SPADes	1.67	225209	28
E36	Qubit	6.3	SPADes_300	1.68	176384	25

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

Table S 5. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-C23-02.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	13.8	SPAdes	1.76	114301	77
E02	Qubit	15.8	shovill v1.1.0	1.76	107402	37
E03	Other	19.4	SPADes_300	1.76	106370	42
E05	Other	16.0	SPADes	1.73	18850	172
E07	Qubit	21.1	SPADes	1.76	126864	33
E08	Qubit	10.3	SKESA	1.70	124873	56
E09	Qubit	26.4	SPADes	1.76	126678	38
E10	Qubit	37.5	SPADes	1.76	126284	66
E12	Nanodrop	22.8	SPADes	1.76	35870	49
E13	Qubit	28.6	Velvet_500	1.80	126831	67
E14	Qubit	23.3	SPADes	1.80	129756	75
E16	Qubit	20.0	SPADes_500	1.75	129796	37
E19	Other	30.0	Unicycler	1.80	129105	36
E20	Qubit	22.2	SPADes	1.76	119498	51
E21	Qubit	17.6	SKESA_200	1.70	56663	89
E22	Qubit	28.6	SPADes	1.76	132930	52
E23	Qubit	32.2	SPADes	1.70	107218	63
E24	Qubit	20.5	SPADes_500	1.76	97625	50
E27	Qubit	15.9	SPADes	1.76	126678	49
E28	Qubit	8.1	SPADes	1.75	126422	33
E29	Qubit	19.6	SPADes	1.80	107845	171
E31	Qubit	18.5	SPADes	1.78	126284	139
E32	Qubit	26.5	CLC_500	1.75	29295	186
E33	Qubit	3.5	SPADes_500	1.75	72318	48
E35	Qubit	19.2	SPADes	1.75	107745	39
E36	Qubit	16.0	SPADes_300	1.76	119249	42

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

Table S 6. Distribution of DNA concentrations, assembly methods and assembly quality parameters among 25 laboratories and EQA provider for sample EQA2-C23-03.

Participant code	DNA		Assembly			
	Method	Concentration, ng/µl	Method	Genome size, Mb	N50, bp	No. of contigs
Ref	Qubit	7.7	SPAdes	1.74	126384	182
E02	Qubit	8.8	shovill v1.1.0	1.72	140284	50
E03	Other	12.4	SPADes_300	1.72	162605	55
E05	Other	10.0	SPADes	1.70	38167	104
E07	Qubit	13.3	SPADes	1.71	162505	48
E08	Qubit	5.4	SKESA	1.70	140100	99
E09	Qubit	15.4	SPADes	1.73	162605	49
E10	Qubit	21.1	SPADes	1.73	162505	131
E12	Nanodrop	12.2	SPADes	1.72	28207	61
E13	Qubit	18.5	Velvet_500	1.70	162455	91
E14	Qubit	12.8	SPADes	1.70	162505	153
E16	Qubit	10.0	SPADes_500	1.71	162549	51
E19	Other	19.0	Unicycler	1.70	143633	43
E20	Qubit	14.0	SPADes	1.72	99657	102
E21	Qubit	11.8	SKESA_200	1.70	103500	85
E22	Qubit	17.9	SPADes	1.73	141674	98
E23	Qubit	18.1	SPADes	1.70	140100	89
E24	Qubit	14.0	SPADes_500	1.72	162605	52
E27	Qubit	10.0	SPADes	1.74	140356	140
E28	Qubit	10.5	SPADes	1.70	140100	43
E29	Qubit	11.1	SPADes	1.78	140356	197
E31	Qubit	9.6	SPADes	1.75	141214	159
E32	Qubit	20.8	CLC_500	1.70	16188	263
E33	Qubit	3.9	SPADes_500	1.72	89744	72
E35	Qubit	10.4	SPADes	1.71	135460	47
E36	Qubit	7.1	SPADes_300	1.73	140356	79

*filtering based on the contig size. No number means that filtering was not applied. Indicated number means that filtering of a certain length in base pairs was applied.

7. Annex B

7.1. Supplementary materials, methods for gene and point mutation detection

Table S 7. An overview of the tools, databases, inputs, thresholds for sequence coverage and identity used by 25 participants for the detection and reporting of AMR genes in Salmonella. Same number in the first column indicates that these participants used same tools and databases with the same inputs, identity, coverage and same strategy of reporting.

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/CAR D_N	1	E02[E]	default	90
	4 tools, 1 input	2 databases, 1 input				
22	ARIBA_R/Abricate_ R/ AMRFinderPlus_N/B N Plugin	AMRFinderPlus_N/ ResFinder_R	1	E21	90/85/ 90[F]	90/85/ 90[F]

[A]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

[B]light yellow indicated that theses participants reported all genes from all databases, light red - that participants reported a subset of genes based on experience/knowledge/literature, light green - that participants reported a consensus list of genes (common genes present in all databases used)
[C]in my report I have also used the literature where I found it necessary
[D]at least 99% on one of the three databases used
[E]since we use 3 databases we report genes that are present in at least 2 databases, a sort of a voting system
[F]AMRFinder (90), BN plugin (85), ResFinder (90)

Table S 8. An overview of tools, databases and inputs used by 25 participants for the detection and reporting of point mutations in Salmonella. Same number in the first column indicates that these participants used same tools and databases with the same inputs, and same strategy of reporting.

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

[A]Inputs : N - DNA fasta, P - protein fasta, R - raw reads. >1 input, if different inputs were used for at least one of the databases/tools
[B]light yellow indicated that theses participants reported point mutations from all databases without curating, light red - that participants curated the point mutations from all databases for duplicates
[C]Bionumerics Plugin (AMRFinder)

Table S 9. An overview of tools, databases and inputs used by 25 participants for the detection and reporting of AMR genes in Campylobacter. Same number in the first column indicates that these participants used same tools and databases with the same inputs, identity, coverage and same strategy of reporting.

Unique combinations	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	E08	90	60
1				E09[C]	90	60
1				E29	90	60
1				E33	90	60
2	ResFinder_R	ResFinder_R	3	E16	85	60
3				E20	90	60
4				E35	80	60
5	AMRFinderPlus_N	AMRFinderPlus_N	1	E24	97	97
6	CLC GW	QMI-AR	1	E32	98	60
	1 tool, 1 input	4 databases, 1 input				
7	BLAST_N	ResFinder_N/ CARD_N/NCBI/own	1	E12	90	90
	1 tool, >1 input	1 database, >1 input				
8	ResFinder_N_R	ResFinder_N_R	3	E03	99	100
9				E07	90	60
9				E31	90	60
10	ResFinder_N_P_R	ResFinder_N_P_R	1	E05	90	60
	2 tools, 1 input	2 databases, 1 input				
11	ResFinder_N/ AMRFinderPlus_N	ResFinder_N/ AMRFinderPlus_N	2	E10	80	50
12				E36	90	60
13	ResFinder_R/ AMRFinderPlus_R	ResFinder_R/ AMRFinderPlus_R	1	E14	>90	
	2 tools, >1 input	2 databases, >1 input				
14	ResFinder_N_R/ AMRFinderPlus_N	ResFinder_N_R/ AMRFinderPlus_N	2	E22	90	40
15				E27[D]	98	100
16	ResFinder_R/ AMRFinderPlus_N	ResFinder_R/ AMRFinderPlus_N	1	E19	90	60
17	ResFinder_N_R/ RGI_N	ResFinder_N_R/ CARD_N	1	E13[E]	90	60
	3 tools, 1 input	3 databases, 1 input				
18	ResFinder_N/ AMRFinderPlus_N/ RGI_N	ResFinder_N/ AMRFinderPlus_N/ CARD_N	1	E02	default	90
	3 tools, >1 input	2 databases, >1 input				
19	AMRFinderPlus_N/ Ariba_R/ABRicate_R	ResFinder_R/ AMRFinderPlus_N	1	E21	90	90
20	ResFinder_R/ AMRFinderPlus_R/ BLAST_N	ResFinder_R/ AMRFinderPlus_R	1	E28	90	60
	3 tools, >1 input	3 databases, >1 input				
21	ResFinder_R/ AMRFinderPlus_N/ RGI_N	ResFinder_R/ AMRFinderPlus_N/ CARD_N	1	E23	99[F]	99[F]

[A]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
[B]light yellow indicated that theses participants reported all genes from all databases, light red - that participants reported a subset of genes based on experience/knowledge/literature, light green - that participants reported a consensus list of genes (common genes present in all databases used)
[C]We use DTU'S CGE Resfinder within Bionumerics 8.1

[D]genes which do not confer resistance are not reported.
[E]In my report I have also used the literature where I found it necessary
[F]at least 99% on one of the three databases used

Table S 10. An overview of tools, databases and inputs used by 25 participants for the detection and reporting of point mutations in Campylobacter. Same number in the first column indicates that these participants used same tools and databases with the same inputs, and same strategy of reporting.

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	4	E03
1				E29
1				E08
1				E09[C]
2	PointFinder_R	ResFinder_R	4	E16
2				E20
2				E28
2				E35
3	AMRFinderPlus_N	AMRFinderPlus_N	2	E21
3				E24
4	CLC GW	CLC GW[D]	1	E32
	1 tool, >1 input	1 database, >1 input		
5	PointFinder_N_R	ResFinder_N_R	4	E07
5				E13
5				E31
5				E33
6	PointFinder_N_P_R	ResFinder_N_P_R	1	E05
	1 tool, 1 input	4 databases, 1 input		
7	BLAST_N	AMRFinderPlus/ResFinder_N/ CARD/own	1	E12
	2 tools, 1 input	2 databases, 1 input		
8	PointFinder_N/ AMRFinderPlus_N	ResFinder_N/ AMRFinderPlus_N	2	E10
9				E36
10	PointFinder_R/ AMRFinderPlus_R	ResFinder_R/ AMRFinderPlus_R	1	E14
	2 tools, >1 input	2 databases, >1 input		
11	PointFinder_R/ AMRFinderPlus_N	ResFinder_R/ AMRFinderPlus_N	3	E02
12				E19
12				E23
13	PointFinder_N_R/ AMRFinderPlus_N	ResFinder_N_R/ AMRFinderPlus_N	2	E27
13				E22

[A]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
[B]light yellow indicated that theses participants reported all point mutations from all databases, light red - that participants reported a subset of point mutations based on experience/knowledge/literature, light green - that participants reported a consensus list of point mutations (common point mutations present in all databases used).
[C]We use DTU'S CGE resfinder within Bionumerics 8.1
[D]module PointFinder database for *Campylobacter* (2019-08)

7.2. Supplementary materials, serotype/species and ST identification

Table S 11. Salmonella ST and methods used for identification by the participants

Lab code	MLST method	EQA2-S23-01	EQA2-S23-02	EQA2-S23-03
E02	MLST (tsemann)	10	29	469
E03	MLST (tsemann)	10	29	469
E05	MLST2.0 (CGE tools)		3241	469
E06	Enterobase	10	29	469
E07	MLST2.0 (CGE tools)	10	29	469
E08	Ridom SeqSphere+	10	29	469
E10	senterica_achtman_2	10	29	469
E11	MLST2.0 (CGE tools)	10	29	469
E13	MLST2.0 (CGE tools)	10	29	469
E14	Enterobase		29	469
E15	PubMLST		29	469
E16	MLST2.0 (CGE tools)	10	29	469
E17	MLST2.0 (CGE tools)	10	29	469
E19	Enterobase	10	29	469
E20	MLST2.0 (CGE tools)	10	29	469
E21	In-house Bifrost using Enterobase scheme	10	29	469
E22	Ridom SeqSphere+	10	29	469
E23	Enterobase		29	469
E24	Enterobase	10	29	469
E28	MLST (tsemann)	10	29	469
E29	MLST2.0 (CGE tools)	10	29	469
E31	MLST2.0 (CGE tools)	10	29	469
E33	MLST2.0 (CGE tools)	10	29	
E35	MLST2.0 (CGE tools)	10	29	469
E36	MLST (tsemann)	10	29	469

Table S 12. Salmonella serotypes and methods used for identification by the participants

Lab code	SeqSero2	SISTR	Enterobase	Other	EQA2-S23-01	EQA2-S23-02	EQA2-S23-03
E02	X	X			Dublin	Stanley	Rissen
E03	X				Dublin	Stanley	Rissen
E05	X				O -9 ; H1 g,m ; H2 -	O -4 ; H1 d ; H2 1,2	O -7 ; H1 f, g : H2 -
E06				In-house script (unpublished)	Dublin	Stanley	Rissen
E07	X	X			Dublin	Stanley	Rissen
E08			X		S. Adelaide	S. Agona	S. Meleagridis
E10				Seq sero cge-dtu	Dublin	stanley	Rissen
E11	X	X			Dublin	Stanley	Rissen
E13	X				Salmonella Dublin (9:g,p)	Stanley (4:d:1,2)	Rissen (7:f,g)
E14	X	X	X			Stanley	Rissen
E15	X					Stanley (4:d:1,2)	Rissen (7:f,g:-)
E16				SeqSero2 incorporated in in-house pipeline	Dublin	Stanley	Rissen
E17	X	X			Dublin	Stanley	Rissen
E19	X				Dublin 9:g,p:-	Stanley 4:d:1,2	Rissen 7:f,g:-
E20	X				Dublin	Stanley	Rissen
E21	X		X		Dublin	Stanley	Rissen
E22	X				Dublin	Stanley	Rissen
E23	X	X	X			Stanley	Rissen
E24	X			an in-house database of STs and corresponding serovars in combination with SeqSero	Dublin	Stanley	Rissen
E28	X	X	X		dublin 9:g,p:-	4:d:1,2; Stanley	7:f,g:-; Rissen
E29	X				Dublin	Stanley	Rissen
E31	X				Dublin	Stanley	Rissen
E33		X			Dublin	Stanley	Rissen
E35	X				O-9, H1:g,p, H2:-	O:4, H1:1,2, H2:d	O:7, H1:f,g, H2:-
E36	X				4:d:1,2	4:d:1,2	7:f,g:-

Table S 13. Campylobacter ST and methods used for identification by the participants

Lab code	MLST method	EQA2-C23-01	EQA2-C23-02	EQA2-C23-03
E02	MLST (tsemann)	888	1586	872
E03	MLST (tsemann)	888	1586	872
E05	MLST2.0 (CGE tools)	888	1586	
E07	MLST2.0 (CGE tools)	888	1586	872
E08	Ridom SeqSphere+	888	1586	872
E09	Bionumerics 8.1	888	1586	872
E10	PubMLST	888	1586	872
E12	PubMLST	888	1586	872
E13	MLST2.0 (CGE tools)	888	1586	872
E14	MLST2.0 (CGE tools)	888	1586	872
E16	MLST2.0 (CGE tools)	888	1586	872
E19	MLST2.0 (CGE tools)	888	1586	872
E20	MLST2.0 (CGE tools)	888	1586	872
E21	in house Bifrost using pubMLST database	888	1586	872
E22	Ridom SeqSphere+	888	1586	872
E23	MLST2.0 (CGE tools)	888	1586	872
E24	PubMLST	888	1586	872
E27	SeqSphere	888	1586	872
E28	MLST (tsemann)	888	1586	872
E29	MLST2.0 (CGE tools)	888	1586	872
E31	MLST2.0 (CGE tools)	888	1586	872
E32	CLC Genomics workbench, Type a known species	888	1586	872
E33	MLST2.0 (CGE tools)	888	1586	872
E35	MLST2.0 (CGE tools)	888	1586	872
E36	MLST (tsemann)	888	1586	872

Table S 14. Campylobacter species and methods used for identification by the participants

Lab code	KmerFinder	Blast	Kraken	Unknown	Other	EQA2-C23-01	EQA2-C23-02	EQA2-C23-03
E02			X		rMLST	*C. coli*	*C. coli*	*C. coli*
E03	X					*C. coli*	*C. coli*	*C. coli*
E05	X					*C. coli*	*C. coli*	*C. coli*
E07	X				SpeciesFinder, ResFinder	*C. coli*	*C. coli*	*C. coli*
E08					Ridom SeqSphere+	*C. coli*	*C. coli*	*C. coli*
E09					PubMLST	*C. coli*	*C. coli*	*C. coli*
E10	X				RefSeq Masher	*C. coli*	*C. coli*	*C. coli*
E12					FastANI	*C. coli*	*C. coli*	*C. coli*
E13	X					*C. coli*	*C. coli*	*C. coli*
E14	X		X			*C. coli*	*C. coli*	*C. coli*
E16					Kraken2/Bracken included in in-house pipeline	*C. coli*	*C. coli*	*C. coli*
E18					PCR In gel	*C. coli*	*C. coli*	*C. coli*
E19	X					*C. coli*	*C. coli*	*C. coli*
E20	X					*C. coli*	*C. coli*	*C. coli*
E21			X		Kraken as part of Bifrost	*C. coli*	*C. coli*	*C. coli*
E22					rMLST	*C. coli*	*C. coli*	*C. coli*
E23					pubMLST	*C. coli*	*C. coli*	*C. coli*
E24		X				*C. coli*	*C. coli*	*C. coli*
E27					SeqSphere	*C. coli*	*C. coli*	*C. coli*
E28	X	X	X			*C. coli*	*C. coli*	*C. coli*
E29	X					*C. coli*	*C. coli*	*C. coli*
E31	X				SpeciesFinder	*C. coli*	*C. coli*	*C. coli*
E32					CLC Genomics Workbench, Find best match using k-mer spectra	*C. coli*	*C. coli*	*C. coli*
E33	X					*C. coli*	*C. coli*	*C. coli*
E35	X					*C. coli*	*C. coli*	*C. coli*
E36			X			*C. coli*	*C. coli*	*C. coli*

8. Annex C

8.1. Supplementary gene tables for both organisms

This section contains tables with genes that were reported by some participants but not reported by the reference datasets.

Table S 15. Additional genes reported in Salmonella strain EQA2-S23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database. Participants E05, E15 and E23 did not report any genes for this strain.

	ResFinder											AMRFinderPlus +/- ResFinder									CARD +/- other			
	Res_Ref	E31	E06	E33	E35	E03	E16	E07	E20_1	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02
ResFinder															*									
AMRFinderPlus															*									
CARD																								
aph(3')-Ia														X										
blaTEM-182										X														
golS																						X		
sitABCD			X																					
sul1																	X							
tet(B)									X															
tet(G)																								X

Table S 16. Additional genes reported in Salmonella strain EQA2-S23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database.

	ResFinder													AMRFinderPlus +/- ResFinder								CARD +/- other				
	Res_Ref	E31	E15	E06	E33	E35	E05	E03	E16	E07	E20_1	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02	E23
ResFinder																	*										
AMRFinderPlus																	*										
CARD																											
aac(3)-IIa			X			X															X						
aac(6')-Ib-cr											X																
aadA7											X																
ant(3")-Ib						X																					
aph(3')-IIa										X									X								
dfrA14											X																
golS																							X				
mrx																							X				
qacE		X				X			X																		
qacEdelta1																	X	X		X			X	X	X	X	
qacL		X	X														X	X		X					X	X	

Table S 17. Additional genes reported in Salmonella strain EQA2-S23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database.

	ResFinder													AMRFinderPlus +/- ResFinder								CARD +/- other				
	Res_Ref	E31	E15	E06	E33	E35	E05	E03	E16	E07	E20_1	E28	E29	AMR_Ref	E24	E11	E21	E10	E36	E22	E14	E19	E08	E13	E17	E02	E23
ResFinder																	*										
AMRFinderPlus																	*										
CARD																											
aac(3)-IIa																					X						
golS																							X	X			

Table S 18. Additional genes reported in Campylobacter strain EQA2-C23-01. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database.

	ResFinder													AMRFinder +/- ResFinder									CARD +/- others				
	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
ResFinder																											
AMRFinderPlus																											
CARD																											
Other database																							*				**
aadE										X									X								
ANT(6)-Ig																							X				
tet(O/32/O)									X																		

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

Table S 19. Additional genes reported in Campylobacter strain EQA2-C23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used: Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database.

	ResFinder													AMRFinder +/- ResFinder									CARD +/- others				
	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
ResFinder																											
AMRFinderPlus																											
CARD																											
Other database																							*				**
aph(3')-Ia					X																						
blaOXA-460																											X
blaOXA-489						X																					
blaOXA-61			X	X		X		X			X		X														
sat4																	X										
blaOXA-452						X																					
blaOXA-451						X																					
tet(O/32/O)									X																		
blaOXA-453						X																					
blaOXA-450						X																					

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

Table S 20. Additional genes reported in Campylobacter strain EQA2-C23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder, Yellow – CARD with or without any other database.

	ResFinder													AMRFinder +/- ResFinder								CARD +/- others				
	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
ResFinder																											
AMRFinderPlus																											
CARD																											
Other database																							*				**
aadE										X																	
blaOXA-460																											X
ANT(6)-Ig																							X				

8.2. Supplementary point mutation tables for both organisms

This section contains tables with point mutations that were reported by some participants but not reported by the reference datasets.

Table S 21. Additional point mutations (PMs) reported in Campylobacter strain EQA2-C23-02. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder.

	ResFinder															AMRFinderPlus +/- Resfinder											O
	Res_Ref	E28	E31	E08	E33	E13	E35	E09	E05	E03	E16	E07	E20	E29	E32	AMR_Ref	E21	E24	E27	E10	E02	E36	E23	E22	E14	E19	E12
ResFinder																											
AMRFinderPlus																											
Other database																											*
23S C2113T				X																							

* Card, Ncbi, home-made database

Table S 22. Additional point mutations (PMs) reported in Campylobacter strain EQA2-C23-03. Reference datasets, Res_Ref and AMR_Ref, are shaded grey. Participants are grouped based on database(s) used : Green – ResFinder, Blue – AMRFinderPlus with or without ResFinder.

	ResFinder															AMRFinderPlus +/- Resfinder											O
	Res_Ref	E28	E31	E08	E33	E13	E35	E09	E05	E03	E16	E07	E20	E29	E32	AMR_Ref	E21	E24	E27	E10	E02	E36	E23	E22	E14	E19	E12
ResFinder																											
AMRFinderPlus																											
Other database																											*
Promoting region of blaoxa G57T																											X
23S A1753G				X																							
rpsL A126T				X																							

* Card, Ncbi, home-made database

[Catalogue number]

Publications Office
of the European Union

ISBN [number]