ECDC NORMAL

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 T1.17.1

Report on the first annual inter-laboratory ring-trial of bioinformatics pipelines for *Salmonella* and *Campylobacter*

Version n°: 2

Date: 22-02-2023

Health and Digital Executive Agency

Third EU Health Programme

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, a new executive Agency with name HaDEA (Health and Digital Executive Agency) is taking 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-2920 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

TABLE OF CONTENTS

1. BACKGROUND

This report describes the first *in silico* inter-laboratory ring-trial of bioinformatics pipelines for prediction of AMR genes in antimicrobial-resistant *Salmonella* and *Campylobacter* (RingTrial1-WGS-AMR), the first out of three planned ring-trials, organised by Statens Serum Institut (SSI) in the FWD AMR-RefLabCap project in 2022-24.

Thirty-nine participants representing 37 public health laboratories (+ one veterinary institute) from 33 countries were invited. Even though 24 participants accepted the invitation, only 23 participants submitted the results. The participants could choose between receiving assembled sequences (fasta files) or links to short read sequences (fastq files) by email. Five chose to receive fasta files and 19 chose to receive links to fastq files. Finally, all 24 participants received personal links to the submission form created with the Enalyzer tool (www.enalyzer.com), where they could select to report their results for one or both pathogens.

Participants were encouraged to follow the analysis guidelines in the protocol (https://www.fwdamr-reflabcap.eu/resources/protocols-and-guidelines) developed in the FWD AMR-RefLabCap project. The aim of this and following ring-trials is to investigate the outcome of different databases, tools and bioinformatic pipelines used by the participants and enable them to compare their performance of antimicrobial resistance (AMR) gene and point mutation detection.

2. MATERIALS AND METHODS

2.1. Sequence selection

The four *Salmonella* and *Campylobacter* sequences were selected from the collection used for the eight External Quality Assessment on antimicrobial susceptibility testing (EQA8-AST) of *Salmonella* and *Campylobacter* in the Food- and Waterborne Diseases and Zoonoses Network, which was organised by SSI, as a part of a contract with the European Centre for Disease Prevention and Control (ECDC). Table 1 and Table 2 displays the sequences that were selected to represent genomes with different resistance genes and point mutations.

Table 1. Characteristics of the *Salmonella* sequences selected for the RingTrial1-WGS-AMR

Sequence	Serotype	ST	Genes*	Point mutations
TRING1S-1	Bredeney	505	*qnrB19*	*gyrA* D87G
TRING1S-2	Monophasic Typhimurium (O5-)	34	*aac(3)-IVa, aadA16, aph(3'')-Ib, aph(4)-Ia, aph(6)-Id, arr-3, blaTEM-1, catA2, dfrA27, floR, qacEdelta1, sul1, sul2, tet(D)*	None
TRING1S-3	Corvallis	1541	*aph(3'')-Ib, aph(6)-Id, qnrS1, sul2, tet(A)*	None
TRING1S-4	Emek	76	*sul1*	*gyrA* S83Y

*AMRFinderPlus output

Table 2. Characteristics of the *Campylobacter* sequences selected for the RingTrial1-WGS-AMR

Sequence	Species	ST	Genes*	Point mutations
TRING1C-1	*C. jejuni*	19	*blaOXA-193*	*gyrA* T86I
TRING1C-2	*C. jejuni*	464	*blaOXA, tet(O)*	50S L22 A103V, *gyrA* T86I
TRING1C-3	*C. coli*	8195	*blaOXA-193, tet(O)*	*gyrA* T86I
TRING1C-4	*C. coli*	832	*aad9, aadE, blaOXA-193, tet(O)*	50S L22 A103V, *gyrA* T86I

*AMRFinderPlus output

2.2. WGS analysis

A selection of candidate strains was sequenced using Illumina paired-end sequencing. The quality of sequences (genome size, N50, and total number of contigs) was checked with an in-house QC pipeline (https://github.com/ssi-dk/bifrost) for raw reads and an open source script (https://github.com/hcdenbakker/N50.sh) for assemblies.

Species identification was done using Kraken (https://github.com/DerrickWood/kraken). MLST calling was done with ARIBA (https://github.com/sanger-pathogens/ariba) using the typing schemes from the PubMLST database.
Salmonella serotypes were determined using Enterobase and SeqSero (https://github.com/denglab/SeqSero) as well as in-house developed scripts detecting the subspecies and *d*-Tartrate reaction.

The sequences were analysed for antibiotic resistance genes and point mutations (PMs) to generate 2 reference datasets:

a. RefRes: ResFinder (raw reads)
b. RefAMR: AMRFinderPlus (SPAdes assemblies)

a) RefRes: Antibiotic resistance genes were identified using raw reads that were run with KMA mapping (https://bitbucket.org/genomicepidemiology/kma/src/master/) to the ResFinder database. Point mutations were identified using KMA mapping to the PointFinder database.
b) RefAMR: Antibiotic resistance genes were identified using SPAdes (https://cab.spbu.ru/software/spades/) assemblies that were run through AMRFinderPlus. Point mutations were also identified using AMRFinderPlus.

2.3. Enalyzer survey

A reporting scheme was developed in Enalyzer survey tool. Participants received individual links where they could submit their results within 1 month from receiving the sequences.

The first part of the survey included questions about tools and databases used to identify the ST, AMR genes, point mutations, as well as serotype and species for *Salmonella* and *Campylobacter*, respectively. In the second part, it was possible, reporting one sequence at a time, to select the identified genes from a list. Furthermore, there was an option of entering additional genes in text boxes. For point mutations reporting, the participants were asked to type the detected mutations in text boxes as well.

2.4. Data analysis

Most of the laboratories reported results for both pathogens, but some countries submitted either for *Salmonella* or for *Campylobacter*. For this reason, there was a total of 23 participants, of which 17 submitted *Salmonella* data and 19 *Campylobacter* data. The participants were from Austria, Belgium, Bulgaria, Czech Republic, Estonia, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Malta, Norway, Poland, Portugal, Romania, Serbia, Slovenia and the Netherlands. Participants were assigned random codes RXX.

After collecting the submissions from 23 participants we analysed them by comparing the reported genes and point mutations to two reference data sets generated by the ring-trial provider with different tools and databases, as shown in Table 3.

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

Reference data set name	Tools applied	
	Gene detection	**Point mutation identification**
RefAMR	AMRFinderPlus on SPAdes assemblies	AMRFinderPlus on SPAdes assemblies
RefRes	KMA with ResFinder database	KMA with PointFinder database

The two reference data sets were generated based on two different databases for AMR gene detection and point mutation identification: ResFinder and AMRFinderPlus. This was in order to be able to compare whether similar tools and databases would generate similar results.

3. SALMONELLA RESULTS

3.1. AMR genes and PMs detection methods used

The methods used by the participants varied a lot. The majority of participants used the ResFinder database, either on assemblies, as chosen by 8 participants, or by using raw reads as nine participants chose. Different tools were used to query or map to the database. The second most popular database used by 9 participants was AMRFinderPlus; however, no participant used AMRFinderPlus as the only database for gene detection. The input for analysis in AMRFinderPlus was assemblies primarily made with SPAdes. SPAdes was used by 6 participants and 3 participants used SKESA, Velvet and an unknown assembler.

Figure 1 summarizes different tools reported as used by the participants for antimicrobial resistance gene detection.

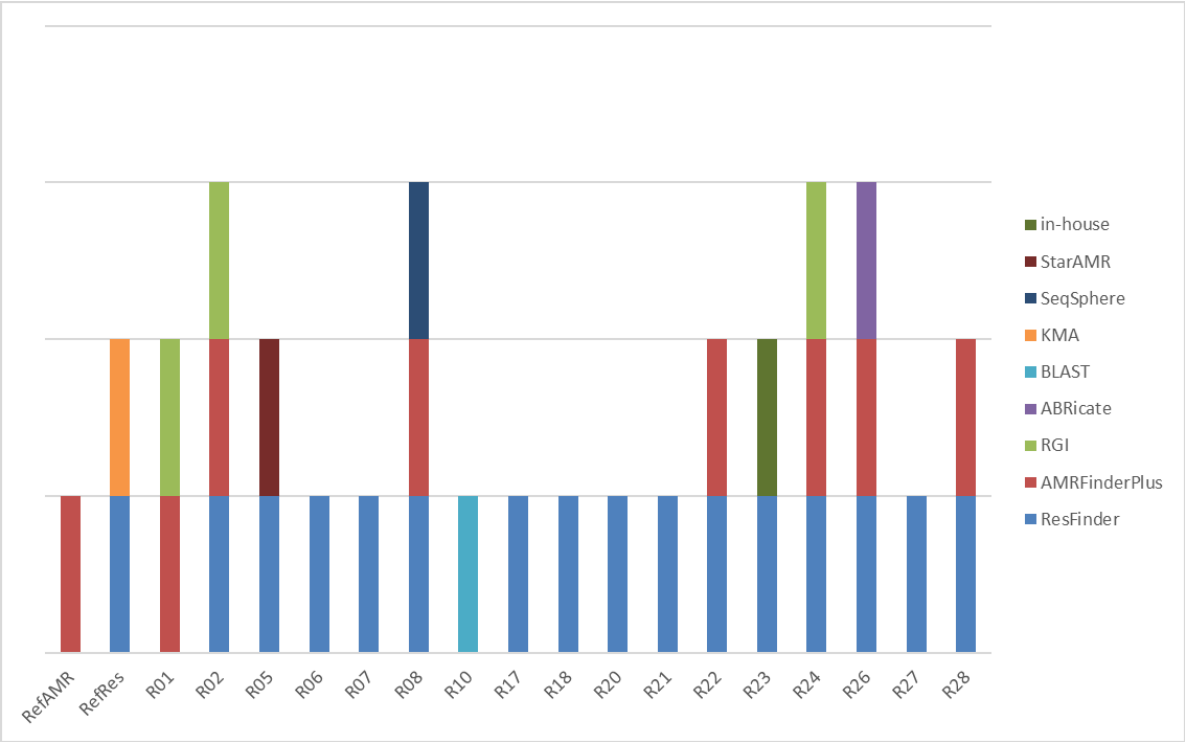

Figure 1. Tools used by the participants for AMR gene detection for all *Salmonella* sequences.

Taken together, ResFinder and AMRFinderPlus were the two most commonly used tools. Seven participants used ResFinder alone and eight in combination with other tools. AMRFinderPlus was used by seven laboratories in combination with other tools.

The tools used for point mutations detection for all participants are summarized in Figure 2.

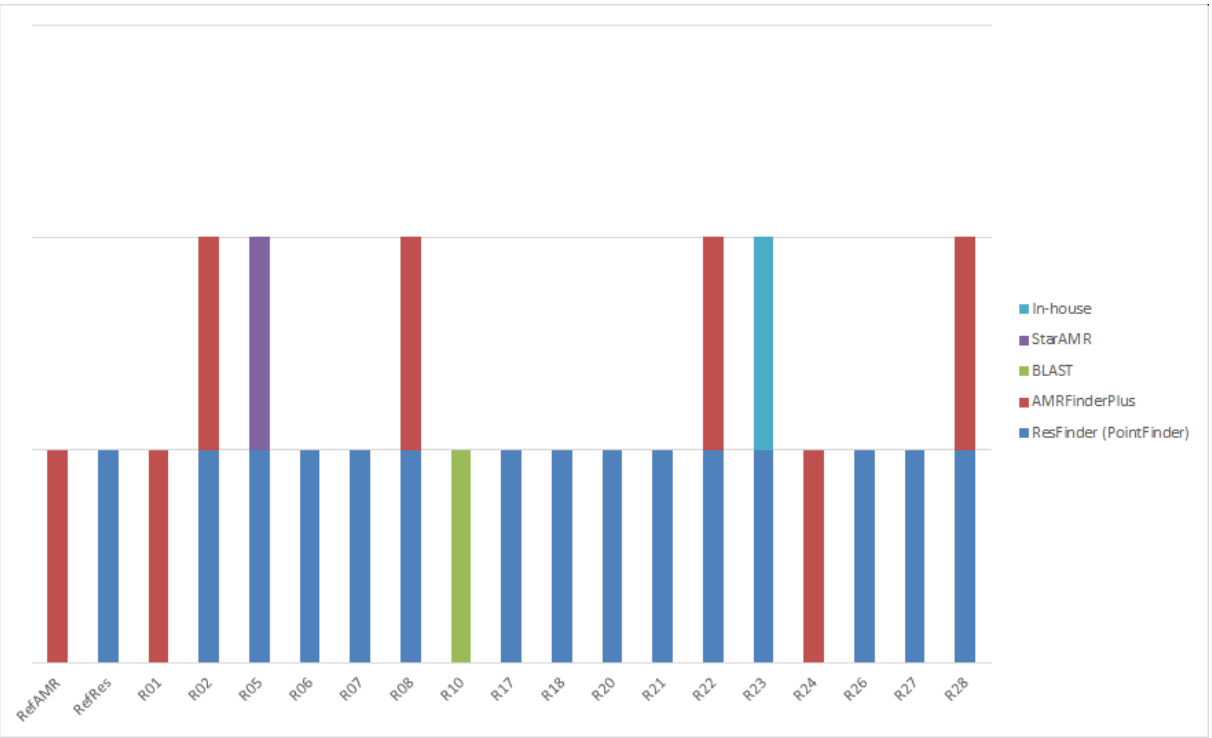

Figure 2. Overview of tools used for point mutation detection by all participants for all *Salmonella* sequences.

For point mutation detection, PointFinder was by far the preferred tool, being used by nine participants as the only tool and by six in combination with other tools. AMRFinderPlus was the second most popular tool, being used by six participants together with another tool or used exclusively.

3.2. Serotypes and STs reported

Participants were asked to report the serotype and ST of all four *Salmonella* sequences. Sixteen participants reported the STs (Table 4) and only one participant did not report the correct STs. It seems to be a reporting issue since the right STs were reported but not for the corresponding sequences.

Table 4. STs reported by the 16 participants.

Lab_ID	TRING1S-1	TRING1S-2	TRING1S-3	TRING1S-4
R01	505	34	1541	76
R02	505	34	1541	76
R05	505	34	1541	76
R06	505	34	1541	76
R08	505	34	1541	76
R10	505	34	1541	76
R17	505	34	1541	76
R18	76	34	505	1541
R20	505	34	1541	76
R21	505	34	1541	76
R22	505	34	1541	76
R23	505	34	1541	76
R24	505	34	1541	76
R26	505	34	1541	76
R27	505	34	1541	76
R28	505	34	1541	76

Serotypes were reported by 17 participants as seen in Table 5 and overall the serotypes were correctly predicted. The majority of participants reported strain TRING1S-2 as Monophasic Typhimurium with a missing O5, however the reporting as 'Monophasic Typhimurium' and 'Typhimurium' was regarded concordant. One participant did not correctly predict the *S*. Bredeney, *S*. Corvallis or *S*. Emek, again a reporting issue might be the case here. Furthermore, six participants could not differentiate sequence TRING1S-3 between *S*. Corvallis and *S*. Chailey and eight participants were not able to distinguish whether the sequence TRING1S-4 was *S*. Chincol or *S*. Emek. Almonst all participants who could not differentiate used SeqSero as the only method. A look-up of the ST in Enterobase would help differentiate between these serotypes.

Table 5. Salmonella serotypes reported by the participants

Lab_ID	TRING1S-1	TRING1S-2	TRING1S-3	TRING1S-4
R01	Bredeney	Monophasic Typhimurium (O5-)	Corvallis or Chailey	Emek
R02	Bredeney	Monophasic Typhimurium	Corvallis	Emek
R05	Bredeney	Monophasic Typhimurium	Corvallis	Emek
R06	Bredeney	Monophasic Typhimurium (O5-)	Corvallis or Chailey	Chincol or Emek
R07	Bredeney	Typhimurium	Corvallis	Chincol
R08	Bredeney	Typhimurium (O5-)	Corvallis	Chincol or Emek
R10	Bredeney	Monophasic Typhimurium	Corvallis	Chincol or Emek
R17	Bredeney	Monophasic Typhimurium (O5-)	Corvallis	Emek
R18	Emek	Monophasic Typhimurium	Bredeney	Corvallis
R20	Bredeney	Monophasic Typhimurium (O5-)	Corvallis or Chailey	Chincol or Emek
R21	Bredeney	Monophasic Typhimurium (O5-)	Corvallis	Emek
R22	Bredeney	Monophasic Typhimurium	Corvallis or Chailey	Chincol or Emek
R23	Bredeney	Monophasic Typhimurium (O5-)	Corvallis	Chincol or Emek
R24	Bredeney	Monophasic Typhimurium (O5-)	Corvallis	Emek
R26	Bredeney	Monophasic Typhimurium (O5-)	Corvallis or Chailey	Chincol or Emek
R27	Bredeney	Monophasic Typhimurium (O5-)	Corvallis or Chailey	Chincol or Emek
R28	Bredeney	Monophasic Typhimurium	Corvallis	Emek

3.3. AMR genes and PMs reported for Salmonella strains

In general, most participants demonstrated a qualified detection of AMR genes and point mutations in the four *Salmonella* sequences used in this ring-trial. Only a few participants had problems in detecting a few genes and point mutations. Furthermore, we strongly suspect that participant R18 has mixed up the reporting of the AMR genes and point mutations of the sequences. We can see, as with the serotypes and STs, that the reporting would have been correct, if reported for a different sequence than the one in question. Therefore, the result from participant R18 has been reported in the following sequence-specific tables (marked in light grey), but disregarded in the discussion of results.

The gene *aac(6')-Iaa* was reported by the majority of the participants. This gene is endogenous to the *Salmonella* genus, however, it is considered a cryptic gene that does not contribute to aminoglycoside resistance (Magnet et al., 1999), hence, it is not reported in the AMRFinderPlus database and it might not be reported in certain other databases as well (Bharat et al., 2022).

In the following sequence-specific tables, the expected antibiotic resistance genes for each sequence are marked with an "X" in the columns RefRes and RefAMR, referring to the two reference datasets as explained in Table 3. The participant laboratories are grouped into three categories. The categories are based on the different databases used to query for antibiotic resistance genes: ResFinder only (Green) and AMRFinderPlus and ResFinder (Blue), together with the corresponding reference datasets (RefRes and RefAMR). Participants that have used the CARD database in addition to other databases are grouped

in the Mix (Yellow) category. For an overview of the tools used by each participant to query the databases for antibiotic gene detection, see Figure 1.

The expected point mutations are also marked with an "X" in separate sequence-specific tables in columns RefAMR and RefRes. The participants are grouped into the same categories as for the antibiotic gene detection. For an overview of which tools different participants used for point mutation detection, see Figure 2.

3.3.1. Sequence TRING1S-1

Sequence TRING1S-1 is a *Salmonella* Bredeney, ST 505.

Table 6. Genes found in sequence TRING1S-1, Green – ResFinder, Blue – AMRFinder and ResFinder, Yellow – a mix of databases.

Lab #	RefRes	R05	R06	R07	R17	R18	R20	R21	R23	R27	RefAMR	R08	R10	R22	R26	R28	R01	R02	R24
	ResFinder										ResAMR						Mix		
ResFinder_db																			
AMR_Finder_db																			
CARD_db																			
aac(6')-Iaa	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X		X	X
qnrB19	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
qnrB5																			X
qnrB81																			X
sul1						X													

All but one participant reported gene *aac*(6')-*Iaa*, but as previously mentioned this gene does not confer resistance in *Salmonella* and is therefore not present in all databases (Table 6). All participants reported the *qnrB19* gene and participant R24 also reported additional *qnr* genes.

There were two point mutations in sequence TRING1S-1, *gyrA* D87G and *parC* T57S, where the latter is considered non-informative by AMRFinderPlus and thus not reported by participants using this database exclusively. All participants and both reference datasets reported *gyrA* D87G (Table 7).

Table 7. Point mutation found in sequence TRING1S-1, Green – PointFinder, Blue – AMRFinder and Yellow – PointFinder and AMRFinderPlus databases.

Lab #	RefRes	R05	R06	R07	R10	R17	R18	R20	R21	R23	R26	R27	RefAMR	R01	R24	R02	R08	R22	R28
	ResFinder												AMRF			Mix			
ResFinder (PointFinder)																			
AMRFinderPlus																			
gyrA D87G	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
gyrA S83Y							X												
parC T57S	X	X	X		X	X		X	X	X	X	X				X	X	X	

For an overview of which tools different participants used for detecting point mutations, see Figure 2.

3.3.2. Sequence TRING1S-2

Sequence TRING1S-2 is a *Salmonella* monophasic Typhimurium (O5-), ST 34.

Table 8. Genes found in sequence TRING1S-2, Green – ResFinder, Blue – AMRFinder and ResFinder, Yellow – a mix of databases.

Lab #	RefRes	R05	R06	R07	R17	R18	R20	R21	R23	R27	RefAMR	R08	R10	R22	R26	R28	R01	R02	R24
	ResFinder										ResAMR						Mix		
ResFinder_db																			
AMRFinderPlus_db																			
CARD_db																			
aac(3)-IV	X	X	X	X	X		X	X	X	X		X	X		X			X	X
aac(3)-IVa											X	X		X		X			X
aac(3)-VIa																	X		
aac(6')-Iaa	X		X	X	X	X	X	X	X	X		X	X	X	X			X	X
aadA16	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
aph(3'')-Ib		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
aph(4)-Ia	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
aph(6)-Id	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
arr-3	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
blaTEM-1											X	X	X	X		X	X		X
blaTEM-1B	X	X	X	X	X		X	X	X	X		X			X			X	X
catA2	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
dfrA27	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
floR	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
mcr-9	X	X	X	X	X		X	X	X	X		X	X	X	X	X	X	X	X
qacE	X		X						X										
qacEdelta1											X				X	X		X	X
qnrS1						X													
sitABCD						X													
sul1	X	X	X		X		X	X	X	X	X	X	X	X	X	X	X	X	X
sul2	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X		X	
tet(A)						X													
tet(D)	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X

For this sequence, several differences were observed between the two reference datasets, RefRes and RefAMR. Many of these differences are caused by the different databases queried.

Among the participants, there was an unanimous detection of genes *aph(6)-Id*, *aadA16*, *aph(4)-Ia*, *arr-3*, *dfrA27*, *floR* and *tetD*. All participants except one reported genes *sul1* and *catA2*.

The *mcr-9* gene was not reported by the AMRFinderPlus reference dataset. A recent phenotype study by National Antimicrobial Resistance Monitoring System (NARMS) found that *mcr-9* did not confer resistance to colistin in over 100 natural *mcr-9*+ isolates (Feldgarden et al., 2022), and therefore the gene is not reported in the AMRFinderPlus database as conferring resistance.

No point mutation was detected or reported in sequence TRING1S-2.

3.3.3. Sequence TRING1S-3

Sequence TRING1S-3 is a *Salmonella* Corvallis, ST 1541.

Table 9. Genes found in sequence TRING1S-3, Green – ResFinder, Blue – AMRFinder and ResFinder, Yellow – a mix of databases.

Lab #	RefRes	R05	R06	R07	R17	R18	R20	R21	R23	R27	RefAMR	R08	R10	R22	R26	R28	R01	R02	R24
Database used	ResFinder										ResAMR						Mix		
ResFinder_db																			
AMRFinderPlus_db																			
CARD_db																			
aac(3)-IV						X													
aac(6')-Iaa	X	X	X	X	X	X	X	X	X	X		X		X	X	X		X	X
aadA16						X													
aph(3'')-Ib	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
aph(4)-Ia						X													
aph(6)-Id	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
blaTEM-1B						X													
catA2						X													
dfrA27						X													
floR						X													
mcr-9						X													
qacE						X													
sitABCD						X													
qnrS1	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
sul1						X													
sul2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
tet(A)	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
tet(D)						X													

In this sequence, only one difference between the two reference datasets, RefRes and RefAMR, was observed. The cryptic gene *aac(6')-Iaa* was detected by RefRes, but not by RefAMR. All participants reported the presence of *aph(6)-Id, qnrS1, sul2* and *tet(A)* genes. All participants, except R07, reported the *aph(3'')-Ib* gene.

One point mutation was detected in TRING1S-3. The *parC* T57S substitution, classified as non-informative by AMRFinderPlus, was reported by elleven of seventeen participants.

Table 10. Point mutation found in sequence TRING1S-3, Green – PointFinder, Blue – AMRFinder, Yellow – PointFinder and AMRFinderPlus databases.

Lab #	RefRes	R05	R06	R07	R10	R17	R18	R20	R21	R23	R26	R27	RefAMR	R01	R24	R02	R08	R22	R28
	ResFinder												AMRF			Mix			
ResFinder (PointFinder)																			
AMRFinderPlus																			
parC p.T57S	X	X	X		X	X		X	X	X	X	X				X		X	

3.3.4. Sequence TRING1S-4

Sequence TRING1S-4 is a *Salmonella* Emek, ST 76.

Table 11. Genes found in sequence TRING1S-4, Green – ResFinder, Blue – AMRFinder and ResFinder, Yellow – a mix of databases.

Lab #	RefRes	R05	R06	R07	R17	R18	R20	R21	R23	R27	RefAMR	R08	R10	R22	R26	R28	R01	R02	R24
	ResFinder										ResAMR						Mix		
ResFinder_db																			
AMRFinderPlus_db																			
CARD_db																			
aac(6')-Iaa	X		X	X	X	X	X	X	X	X		X	X	X	X	X		X	X
qnrB19						X													
sitABCD						X													
sul1	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X

Fifteen of seventeen participants reported gene *aac*(6')-*Iaa*, but as previously mentioned this gene does not confer resistance in *Salmonella* and is therefore not present in all databases. All participants reported the *sul1* gene.

There was one point mutation in sequence TRING1S-4, *gyrA* S83Y, and all participants and both reference datasets reported this point mutation (Table 12).

Table 12. Point mutation found in sequence TRING1S-4, Green – PointFinder, Blue – AMRFinder and Yellow – PointFinder and AMRFinderPlus databases.

Lab #	RefRes	R05	R06	R07	R10	R17	R18	R20	R21	R23	R26	R27	RefAMR	R01	R24	R02	R08	R22	R28
	ResFinder												AMRF			Mix			
ResFinder (PointFinder)																			
AMRFinderPlus																			
gyrA D87G							X												
gyrA S83Y	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
parC T57S							X												

4. CAMPYLOBACTER RESULTS

4.1. AMR genes and PMs detection methods used

The methods used by the participants for gene detection and point mutation identification varied, however, the majority of participants (18 out of 19) used the ResFinder database, either on assemblies or using raw reads. Other databases used included AMRFinderPlus, used by eight participants, and CARD, used by five participants.

Different tools were used to query or map to the databases. ResFinder as a tool was used by eight participants alone and by ten participants in combination with another tool, as shown in Figure 3. Six participants used AMRFinderPlus as a tool in combination with other tools, such as ResFinder, RGI, SeqSphere, ABRicate, BioNumerics or an in-house pipeline.

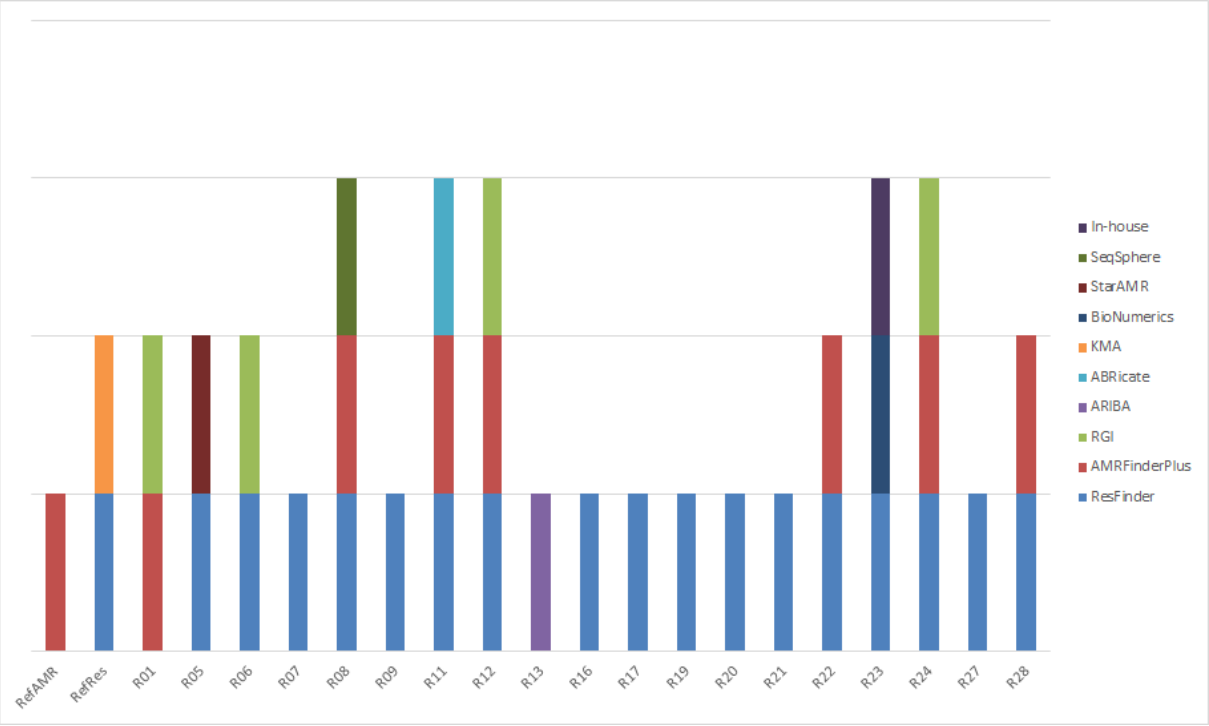

Figure 3. Tools used by the participants for AMR gene detection for all *Campylobacter* sequences.

Eleven out of eighteen participants used assemblies as input for detection of AMR genes and the most commonly used assembler was SPAdes. One participant reported using Velvet and one SKESA. As mentioned in section 3.2, the SAUTE assembler could be considered for assemblies instead of SKESA.

The tools used for point mutations detection for all participants are summarized in Figure 4.

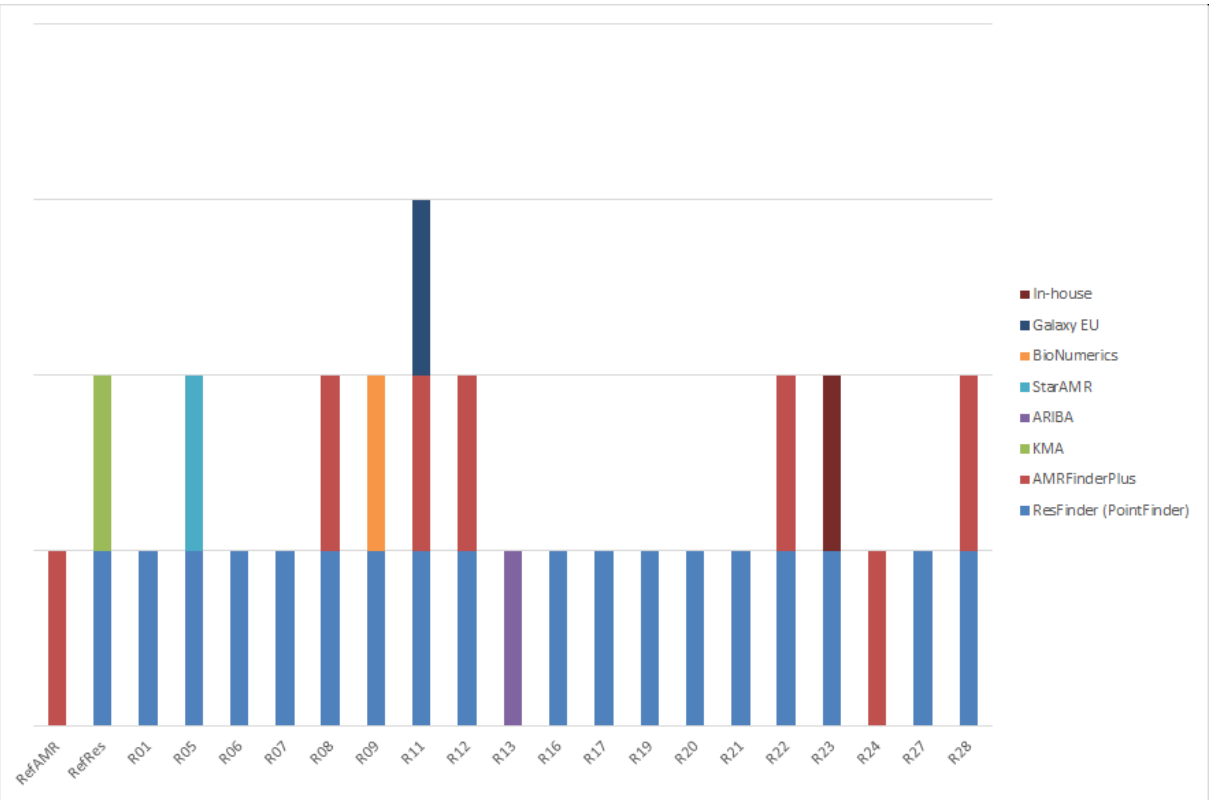

Figure 4. Overview of tools used for point mutation detection by all participants for all *Campylobacter* sequences.

PointFinder was the preferred tool, as nine participants used it as the only tool and nine participants in combination with other methods. AMRFinderPlus was the second most commonly used tool, being used by six participants in total, one of which used it as the only tool. Other tools included StarAMR, BioNumerics, KMA, ARIBA, Galaxy EU and one participant reported using an in-house pipeline, where ResFinder is incorporated.

Eighteen out of nineteen participants applied the PointFinder database for point mutation identification. Seven participants used the AMRFinderPlus database (five of which used it in combination with the PointFinder database). Other participants reported using Galaxy EU and an in-house database.

4.2. Species and STs reported

Two participants did not complete the reporting of species and ST of the four *Campylobacter* sequences. The remaining seventeen participants correctly reported the species of all *Campylobacter* sequences. All but one participant reported the correct STs. Participant R07 reported the wrong ST for all four sequences. For the correct species and ST of the sequences, see Table 2.

4.3. AMR genes and PMs reported for Campylobacter sequences

In this section, we used two reference data sets for comparison, as explained in Table 3. In the following sequence-specific tables, the expected antibiotic resistance genes for each sequence are marked with an "X" in the columns RefRes and RefAMR. The participants are grouped into three categories, based on the databases used to identify the antibiotic resistance genes: ResFinder (Green – ResFinder database only), ResAMR (Blue - ResFinder and AMRFinderPlus database) and Mix (Yellow – CARD database in combination with other databases).

The expected point mutations are also marked with an "X" in the sequence-specific tables in columns RefRes and RefAMR. The participants are grouped into similar categories as for the antibiotic gene detection, except that the AMRF category (blue) represents participants that used AMRFinderPlus as the only database and the Mix category (yellow) groups participants that have used more than one database.

4.3.1. Sequence TRING1C-1

Sequence TRING1C-1 is a *C. jejuni*, ST 19.

Table 13. Genes found in sequence TRING1C-1, Green – ResFinder, Blue – AMRFinder + ResFinder, Yellow – CARD and other databases.

Lab #	RefRes	R05	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R08	R11	R22	R28	R06	R13*	R01	R12	R24
	ResFinder											ResAMR					Mix				
ResFinder																					
AMRFinderPlus																					
CARD																					
blaOXA					X																
blaOXA-193	X	X		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
blaOXA-450				X																	
blaOXA-451				X																	
blaOXA-452				X																	
blaOXA-453				X																	
blaOXA-489		X		X																	X
blaOXA-61		X	X	X														X		X	X
cmeABC+R																		X			
tet(O)																			X		

* used an in-house database

In sequence TRING1C-1, the *blaOXA-193* gene was detected in both reference datasets used, RefRes and RefAMR. This gene was also reported by the majority of participants. Six participants reported *blaOXA-61* gene.

One point mutation, *gyrA* T86I, was detected in both reference datasets, RefRes and RefAMR, as well as almost all participants.

Table 14. Point mutations found in sequence TRING1C-1, Green – ResFinder, Blue – AMRFinder, Yellow – Different databases.

Lab #	RefRes	R01	R05	R06	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R24	R08	R11	R12	R13 *	R22	R28
	ResFinder													AMRF		Mix					
ResFinder (PointFinder)																					
AMRFinderPlus																					
23S						X	X														
gyrA T86I	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X

* used an in-house database

Participant R09 reported one additional point mutation in gyrA, as well as six point mutations in the *23S* gene (Table S 1).

4.3.2. Sequence TRING1C-2

Sequence TRING1C-2 is a *Campylobacter jejuni*, ST 464.

Table 15. Genes found in sequence TRING1C-2, Green – ResFinder, Blue – AMRFinder + ResFinder, Yellow – CARD and other databases.

Lab #	RefRes	R05	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R08	R11	R22	R28	R06	R13 *	R01	R12	R24
	ResFinder											ResAMR					Mix				
ResFinder																					
AMRFinderPlus																					
CARD																					
blaOXA												X			X						
blaOXA-193	X					X	X	X	X	X	X		X			X					
blaOXA-461																		X			
OXA-660																					X
tet(O)				X								X	X					X			X
tet(O/32/O)	X	X	X	X	X	X	X	X	X	X	X			X	X	X	X			X	X
tet(O/M/O)																		X			

* used an in-house database

Both reference datasets, RefRes and RefAMR identified the same genes, *blaOXA* and *tet(O)*, however, ResFinder identified the variants *blaOXA-193* and *tet(O/32/O)*. Most participants identified the *tet(O/32/O)* gene variant and one participant reported *tet(O/M/O)* gene instead. AMRFinderPlus only determines the *blaOXA* gene. This is caused by an imperfect match to the underlying sequences e.g. *blaOXA-193*, thus AMRFinderPlus does not report an imperfect match to a variant as it is potentially a novel variant (Feldgarden et al., 2019).

Table 16. Point mutations found in sequence TRING1C-2, Green – ResFinder, Blue – AMRFinder, Yellow – Different databases.

Lab #	RefRes	R01	R05	R06	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R24	R08	R11	R12	R13 *	R22	R28
	ResFinder													AMRF		Mix					
ResFinder (PointFinder)																					
AMRFinderPlus																					
23S						X	X		X												
50S L22 A103V		X												X	X		X	X		X	X
cmeR						X	X		X												
gyrA							X		X												
gyrA 2 T86I				X																	
gyrA T86I	X					X		X		X	X	X	X	X	X	X	X	X	X	X	X

* used an in-house database

The gyrA T86I point mutation was identified in both reference data sets, RefRes and RefAMR. However, the mutation 50S L22 A103V was reported only in the RefAMR dataset.

The gyrA T86I substitution was reported by 13 out of 21 participants in strain TRING1C-2. Two participants reported an unspecified mutation in *gyrA* gene and one participant reported the mutation in *gyrA_2* variant of the gene, present in PointFinder database since June 2022. The 50S L22 A103V mutation was reported by six out of 21 participants. It is worth noting that the presence of this mutation is not necessarily related to the phenotype, as it is equally common among resistant and sensitive isolates in a set of 516 *Campylobacter* isolates tested in a recent study (Dahl et al., 2021). Three participants reported mutations in *cmeR* and *23S* genes.

Participant R09, similarly as for the previous sequence, reported many nucleotide substitutions, listed in Table S 2.

4.3.3. Sequence TRING1C-3

Sequence TRING1C-3 is a *Campylobacter coli*, ST 8195.

Table 17. Genes found in sequence TRING1C-3, Green – ResFinder, Blue – AMRFinder + ResFinder, Yellow – CARD and other databases.

Lab #	RefRes	R05	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R08	R11	R22	R28	R06	R13 *	R01	R12	R24
	ResFinder											ResAMR					Mix				
ResFinder																					
AMRFinderPlus																					
CARD																					
blaOXA					X																
blaOXA-193	X	X		X		X		X	X	X	X	X	X	X	X	X	X	X		X	X
blaOXA-489		X		X																	
blaOXA-450				X																	
blaOXA-451				X																	
blaOXA-453				X																	
blaOXA-61		X	X	X														X		X	X
tet(O)				X								X	X					X			X
tet(O/32/O)	X	X	X	X	X	X		X	X	X	X			X	X	X	X	X		X	X
tet(O/M/O)																		X			

* used an in-house database

In sequence TRING1C-3, the *blaOXA-193* gene was identified in both reference sets, RefRes and RefAMR. The *tet(O)* gene was also dientified by both datasets, but in RefRes the gene was reported as *tet(O/32/O)*. These two genes were identified by most of the participants. Additional *blaOXA* genes were also reported by a few participants.

Mutation *gyrA* T86I was identified in the RefAMR reference dataset, but not in the RefRes dataset.

Table 18. Point mutations found in sequence TRING1C-3, Green – ResFinder, Blue – AMRFinder, Yellow – Different databases.

Lab #	RefRes	R01	R05	R06	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R24	R08	R11	R12	R13 *	R22	R28
	ResFinder													AMRF		Mix					
ResFinder (PointFinder)																					
AMRFinderPlus																					
23S						X															
gyrA: T86I						X							X	X	X	X	X	X	X	X	X
gyrA_2 p.T86I			X	X	X		X	X		X	X										

* used an in-house database

The majority of participants (16 out of 19) identified the gyrA T86I substitution, with many PoinFinder users reporting it on the *gyrA_2* variant. Participant R09 reported additional 7 point mutations in gene 23S (shown in
Table S 3).

4.3.4. Sequence TRING1C-4

Sequence TRING1C-4 is a *Campylobacter coli*, ST 832.

Table 19. Genes found in sequence TRING1C-4, Green – ResFinder, Blue – AMRFinder + ResFinder, Yellow – CARD and other databases.

Lab #	RefRes	R05	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R08	R11	R22	R28	R06	R13 *	R01	R12	R24
	ResFinder											ResAMR					Mix				
ResFinder																					
AMRFinderPlus																					
CARD																					
aad9												X			X	X					X
aadE												X	X		X	X			X		X
ant(6)-Ia	X	X		X	X	X		X	X	X	X			X		X	X	X		X	X
blaOXA-193	X	X		X		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
blaOXA-450				X																	
blaOXA-451				X																	
blaOXA-452				X																	
blaOXA-453				X																	
blaOXA-489		X		X																	
blaOXA-61		X	X	X														X		X	X
tet(O)		X		X								X			X	X	X	X			X
tet(O/32/O)	X		X							X											
tet(O/M/O)																		X			

* used an in-house database

Genes *aad9* and *aadE* were identified in the RefAMR reference dataset, whereas in the RefRes dataset, *ant(6)-Ia* was reported. This is not contradictory, as genes of aminoglycoside nucleotidyltransferase subfamily (ANT(6)-I) are also known as aminoglycosides adenyltransferases of the AADE family (Hormeño et al., 2018). This difference is also seen among participants, where 8 out of 10 participants using ResFinder reported this gene as *ant(6)-Ia* (one participant did not report any of these genes), whereas 3 out of 4 participants using AMRFinderPlus in addition to ResFinder, reported *aadE*, *aad9*, or both. Four out of five participants using the CARD database in addition to other tools, reported the *ant(6)-Ia* gene and one participant reported both these genes.

The *blaOXA-193* gene was identified by both reference datasets, as well as majority of the participants (16 out of 19). The *tet(O)* gene or its variant *tet(O/32/O)* was reported by both reference datasets and by only 9 out of 19 participants. One participant, reported this gene as both *tet(O)* and *tet(O/M/O)* based on their in-house database for *Campylobacter* resistance.

Table 20. Point mutations found in sequence TRING1C-4, Green – ResFinder, Blue – AMRFinder, Yellow – Different databases.

Lab #	RefRes	R01	R05	R06	R07	R09	R16	R17	R19	R20	R21	R23	R27	RefAMR	R24	R08	R11	R12	R13 *	R22	R28
	ResFinder													AMRF		Mix					
ResFinder (PointFinder)																					
AMRFinderPlus																					
23S							X														
50S L22 A103V		X												X	X		X	X		X	X
gyrA 2							X														
gyrA 2 T86I			X	X	X			X		X	X										
gyrA T86I						X							X	X	X	X	X	X	X	X	X
rpsL						X															

* used an in-house database

In strain TRING1C-4, mutations in genes *gyrA* and *50S L22* were identified only in one of the two reference datasets, RefAMR. The gyrA T86I substitution was reported by all participants that used AMRFinderPlus database, as well as a combination of AMRFinderPlus and ResFinder databases, but by 8 out of 12 participants that only used ResFinder database.

5. CONCLUSIONS

This RingTrial1-WGS-AMR, organised by Statens Serum Institut (SSI), is the first exercise in the FWDAMR-RefLabCap project, in a series of three, spanning over 3 years. The aim of this and following ring-trials is to investigate the outcome of different databases, tools and analytic pipelines used and enable participants to compare their performance of AMR gene and point mutation detection.

Twenty-four participants accepted the invitation and 23 participants submitted the results.

RingTrial1-WGS-AMR participants received sequences from the provider, either as raw data (fastq files) or assemblies (fasta files), and performed sequence analysis to identify antimicrobial resistance genes and point mutations. Overall, the participants performed well and identified the expected targets. This ring-trial highlighted that small differences in the results are to be expected among participants likely due to using different input data types (raw reads or assemblies), tools used to perform assemblies and, finally, tools used to detect genes and point mutations. Different tools can give different results. Whether using mapping of raw reads or blasting of assemblies, there are different points of attention to consider. If there is a gene in a genomic region that is difficult to assemble, it could potentially be missed in an assembly based approach. If there are many closely related variants in the database used, it might complicate determination of the exact variant through mapping.

Difference in input data can also have an effect on results. In *Campylobacter* genome TRING1C-4 only 9 of 19 participants detected the *tet(O)* gene or variants hereof. Very surprisingly, when using reads as input in ResFinder as opposed to assemblies, the gene is not detected with the default settings (90% sequence identity). When using reads as input the gene only has an 88% sequence identity. This explains why many participants analysing reads in ResFinder did not report this gene.

Reference gene databases also reports different results. Some genes have different nomenclature in different databases, some databases are based on amino acids and some on nucleotides. Finally, it is always important to ensure that the latest version of the database is being used, as well as to ensure that the database is regularly updated. Find more information on databases and tools in the protocol (https://www.fwdamr-reflabcap.eu/resources/protocols-and-guidelines).

In this ring-trial, an example of how presence or absence of certain genes in some databases can affect the results was the *mcr-9* gene found in *Salmonella*, present in ResFinder database, but removed from AMRFinderPlus database. Another example is the *blaOXA* gene in *Campylobacter* detected only to gene level by AMRFinderPlus, but ResFinder predicts the closest match to a variant, even if it is not a perfect match.

Several participants reported results based on more than one database. If output from two or more databases are merged without any critical assessment, too many genes will potentially be reported due to differences in nomenclature, especially different criteria for designating variants of the genes. For example, some participants reported several variants of the same gene (e.g. *blaOXA* in *Campylobacter* and *qnr* in *Salmonella*), which is likely due to using tools that allow less than 100% match to the complete gene. None of the participants used the AMRFinderPlus as the only database. For example, this database do not include the cryptic gene *aac(6')-Iaa* and the gene *mcr-9* which does not confer phenotypic resistance, but a majority of participants that also used ResFinder reported these. For routine resistome profiling, it may be less demanding to use the output from only one well-curated database as this will require less special knowledge on gene function, variants, etc. The use of more tools and databases might be useful for participants that have the capacity to investigate the genetic contents of the genomes in more detail in relation to phenotype, as well as the capacity to maintain a high level of special knowledge on AMR genes and functions. Such specialised knowledge should contribute to the curation of the databases.

For the next round of ring-trial, RingTrial2-WGS-AMR in 2023, the provider is planning to redesign the questionnaire, modify questions and structure used for gene reporting and add additional questions about tools and databases and how they are used.

6. REFERENCES

Bharat, A., Petkau, A., Avery, B. P., Chen, J., Folster, J., Carson, C. A., Kearney, A., Nadon, C., Mabon, P., Thiessen, J., Alexander, D. C., Allen, V., El Bailey, S., Bekal, S., German, G. J., Haldane, D., Hoang, L., Chui, L., Minion, J., … Mulvey, M. R. (2022). Correlation between Phenotypic and In Silico Detection of Antimicrobial Resistance in Salmonella enterica in Canada Using Staramr. *Microorganisms*, *10*(2), 1–10. https://doi.org/10.3390/microorganisms10020292

Dahl, L. G., Joensen, K. G., Østerlund, M. T., Kiil, K., & Nielsen, E. M. (2021). Prediction of antimicrobial resistance in clinical Campylobacter jejuni isolates from whole-genome sequencing data. *European Journal of Clinical Microbiology and Infectious Diseases*, *40*(4), 673–682. https://doi.org/10.1007/s10096-020-04043-y

Feldgarden, M., Brover, V., Fedorov, B., Haft, D. H., Prasad, A. B., & Klimke, W. (2022). Curation of the AMRFinderPlus databases: Applications, functionality and impact. *Microbial Genomics*, *8*(6), 1–10. https://doi.org/10.1099/mgen.0.000832

Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., Tyson, G. H., Zhao, S., Hsu, C. H., McDermott, P. F., Tadesse, D. A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J. P., & Klimke, W. (2019). Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrobial Agents and Chemotherapy*, *63*(11). https://doi.org/10.1128/AAC.00483-19

Hormeño, L., Ugarte-Ruiz, M., Palomo, G., Borge, C., Florez-Cuadrado, D., Vadillo, S., Píriz, S., Domínguez, L., Campos, M. J., & Quesada, A. (2018). Ant(6)-I genes encoding aminoglycoside O-nucleotidyltransferases are widely spread Among streptomycin resistant strains of Campylobacter jejuni and Campylobacter coli. *Frontiers in Microbiology*, *9*(OCT), 1–8. https://doi.org/10.3389/fmicb.2018.02515

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. *Journal of Bacteriology*, *181*(21), 6650–6655. https://doi.org/10.1128/jb.181.21.6650-6655.1999

7. ANNEX

Table S 1. Unique point mutations reported by participant R09 in *Campylobacter* sequence TRING1C-1

Laboratory	Gene	Point mutations reported
R09	*gyrA*	R285K agg -> aag
	23S	327G>A g -> a, 643G>R g -> r, 554A>C a -> c, 298G>A g -> a, 571T>G t -> g, 1027A>G

Table S 2. Unique point mutations reported by participant R09 in *Campylobacter* sequence TRING1C-2

Laboratory	Gene	Point mutations reported
R09	*gyrA*	Q863* caa -> taa, R285K agg -> aag,
	23S	298G>A g -> a, 364G>C g -> c
	cmeR	I115V atc -> gtt, E84K gaa -> aaa, G86S ggc -> agt

Table S 3. Unique point mutations reported by participant R09 in *Campylobacter* sequence TRING1C-3

Laboratory	Gene	Point mutations reported
R09	*23S*	364G>C g -> c, 1735T>C t -> c, 2113C>T c -> t 1730A>C a -> c, 554A>C a -> c, 571T>G t -> g 416T>G t -> g

More information on the European Union is available on the internet (http://europa.eu).
Luxembourg: Publications Office of the European Union, 2021
ISBN
doi:

Printed in
PRINTED ON ELEMENTAL CHLORINE-FREE BLEACHED PAPER (ECF)
PRINTED ON TOTALLY CHLORINE-FREE BLEACHED PAPER (TCF)
PRINTED ON RECYCLED PAPER
PRINTED ON PROCESS CHLORINE-FREE RECYCLED PAPER (PCF)

Health and Digital Executive Agency

Third EU Health Programme