National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

EURL-*Salmonella* Proficiency Test Typing 2020

RIVM report 2021-0126
W.F. Jacobs-Reitsma et al.

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

EURL-*Salmonella* Proficiency Test Typing 2020

RIVM report 2021-0126

Colophon

© RIVM 2021
Parts of this publication may be reproduced provided acknowledgement is given to: National Institute for Public Health and the Environment, along with the title and year of publication.

RIVM attaches a great deal of importance to the accessibility of its products. However, it is at present not yet possible to provide this document in a completely accessible form. If a part is not accessible, it is mentioned as such. Also see www.rivm.nl/en/accessibility

DOI 10.21945/RIVM-2021-0126

W.F. Jacobs-Reitsma (author), RIVM
A. Verbruggen (author), RIVM
R.E. Diddens (author), RIVM
A.H.A.M. van Hoek (author), RIVM
K.A. Mooijman (author), RIVM

Contact:
W.F. Jacobs-Reitsma
Centre for Zoonoses and Environmental Microbiology
wilma.jacobs@rivm.nl

This investigation was performed by order and for the account of the European Commission, Directorate General for Health and Food Safety (DG-SANTE), within the framework of RIVM project number E/114506/20 European Union Reference Laboratory for *Salmonella*

Published by:
National Institute for Public Health and the Environment, RIVM
P.O. Box1 | 3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

Synopsis

EURL-*Salmonella* Proficiency Test Typing 2020

From 1992, National Reference Laboratories (NRLs) of European Union (EU) Member States have been obliged to participate in annual quality control 'Proficiency' Tests (PTs). NRLs from countries outside the EU occasionally participate in these tests on a voluntary basis. One of the PTs is on typing of *Salmonella* bacteria. The NRLs of all 27 EU Member States performed well in this 2020 quality control test on *Salmonella* typing. Overall, the participating laboratories were able to assign the correct name to 97% of the strains tested.

Laboratories are obliged to type *Salmonella* with the reference method (serotyping). In 2020, they could also perform additional typing at DNA level, for example by using Whole Genome Sequencing (WGS). More detailed DNA typing methods are sometimes needed to trace the source of a contamination.

Each Member State designates a specific laboratory within their national boundaries to be responsible for the detection and identification of *Salmonella* in animals and/or food products. These laboratories are referred to as the National Reference Laboratories (NRLs). The performance of these NRLs in *Salmonella* typing is assessed annually by testing their ability to identify 20 *Salmonella* strains.

The United Kingdom, the EU candidate countries Republic of North Macedonia and Serbia, as well as the European Free Trade Association (EFTA) countries Iceland, Norway and Switzerland took voluntary part in the 2020 assessment.

The annual Proficiency Test on *Salmonella* typing is organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*). The EURL-*Salmonella* is located at the National Institute for Public Health and the Environment (RIVM) in the Netherlands.

Keywords: EURL-*Salmonella*, *Salmonella*, serotyping, molecular typing, PFGE, MLVA, WGS, cluster analysis, Proficiency Test

Publiekssamenvatting

EURL-*Salmonella* ringonderzoek typering 2020

Sinds 1992 zijn de Nationale Referentie Laboratoria (NRL's) van de Europese lidstaten verplicht om elk jaar hun kwaliteit te laten toetsen met zogeheten ringonderzoeken. Soms doen NRL's van landen buiten de Europese Unie (EU) vrijwillig mee. Een van de ringonderzoeken is de typering van *Salmonella*-bacteriën. In 2020 scoorden alle NRL's van de 27 EU lidstaten goed bij deze kwaliteitscontrole op typering van *Salmonella*. Als groep konden de deelnemende laboratoria aan 97 procent van de geteste stammen de juiste naam geven.

De laboratoria zijn verplicht om *Salmonella* met een standaardmethode te typeren (serotypering). Daarnaast mochten zij in 2020 zelf aangeven of ze extra typeringen op DNA-niveau wilden doen, bijvoorbeeld met Whole Genome Sequencing (WGS). Deze preciezere typering kan soms nodig zijn om de bron van een besmetting op te sporen.

Voor de kwaliteitstoetsen wijst elke lidstaat een laboratorium aan, het Nationale Referentie Laboratorium (NRL). Dit NRL is namens dat land verantwoordelijk om *Salmonella* in monsters van levensmiddelen of dieren aan te tonen en te typeren. Om te controleren of de laboratoria hun werk goed doen, moeten zij onder andere twintig *Salmonella*-stammen de juiste naam kunnen geven.

In 2020 deden zes landen buiten de Europese Unie vrijwillig mee: het Verenigd Koninkrijk, de EU kandidaat lidstaten Republiek Noord-Macedonië en Servië, en de European Free Trade Association (EFTA) landen IJsland, Noorwegen en Zwitserland.

Het Europese Unie Referentie Laboratorium voor *Salmonella* (EURL-*Salmonella*) organiseert het jaarlijkse ringonderzoek *Salmonella*-typering. Dit laboratorium is gevestigd bij het RIVM in Nederland.

Kernwoorden: EURL-*Salmonella*, *Salmonella*, serotypering, moleculaire typering, PFGE, MLVA, WGS, cluster analyse, ringonderzoek

Contents

Summary

In November 2020, the annual *Salmonella* typing Proficiency Test (PT) was organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands). The study's main objective was to evaluate whether the typing of *Salmonella* strains by the National Reference Laboratories for *Salmonella* (NRLs-*Salmonella*) in the European Union was carried out uniformly, and whether comparable results were obtained.

A total of 37 laboratories participated in this study. These included 29 NRLs-*Salmonella* in the 27 EU Member States plus the United Kingdom, two NRLs of EU candidate countries Republic of North Macedonia and Serbia, three NRLs of EFTA countries Iceland, Norway and Switzerland, and three additional participants to compare with their WGS-based results.

All 37 laboratories performed serotyping. A total of twenty obligatory *Salmonella* strains plus one optional *Salmonella* strain were selected by the EURL-*Salmonella* for serotyping. The strains had to be typed according to the method routinely used in each laboratory, following the White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007). The laboratories were allowed to send strains for serotyping to another specialised laboratory in their country if this was part of their usual procedure.

Overall, 99% of the strains were typed correctly for the O-antigens, 98% of the strains were typed correctly for the H-antigens, and 97% of the strains were correctly named by the participants.
In 2007, criteria for 'good performance' with regard to serotyping were defined (Mooijman, 2007). Using these criteria, the performance of the participants was very good, including the performance of four participants that were submitting WGS-based results. All participants met the level of good performance at the first stage of this PT, and there was no need to organise a follow-up study.

Nineteen NRLs and two external partners also performed additional typing at DNA level (PFGE and/or MLVA and/or WGS) to investigate an additional set of ten *Salmonella* strains using cluster analysis. Based on the information gained from the first pilot in 2019, the second pilot PT Cluster Analysis 2020 was mimicking an outbreak situation, with a monophasic *Salmonella* Typhimurium ST34, MLVA type 3-14-13-NA-211 as the reference strain. Raw WGS data (compressed paired-end fastq files) of this reference strain were made available through a secure ftp server. Participants were asked to analyse the ten strains and to report per strain if a clustering match with the reference strain was found or not.

Evaluation of the participants' cluster analysis results was done by comparing the participants' results to the expected results in the outbreak investigation setting, as pre-defined by the EURL-*Salmonella.*

The two PFGE participants reported their PFGE-based cluster analysis results in complete agreement. Five out of the six participants reported the MLVA-based cluster analysis results completely as expected. All but one of the 23 submissions reported the WGS-based cluster analysis results completely as expected. The technical duplicate strains 20SCA06/20SCA08 were expected to be reported as (part of) one cluster. This was the case in 2/2 PFGE submissions, in 6/6 MLVA submissions, and in 22/23 WGS submissions.

1 Introduction

This report describes the 2020 Proficiency Test (PT) on typing of *Salmonella* organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands) in November 2020.

According to EC Regulation No. 2017/625 (EC, 2017), one of the tasks of the EURL-*Salmonella* is to organise PTs for the National Reference Laboratories for *Salmonella* (NRLs-*Salmonella*) in the European Union. The main objectives for PTs on typing of *Salmonella* are that the typing should be carried out uniformly in all Member States, and that comparable results should be obtained. The implementation of PTs on typing started in 1995.

A total of 37 laboratories participated in this study. These included 29 NRLs-*Salmonella* in the 27 EU Member States plus the United Kingdom, two NRLs of EU candidate countries Republic of North Macedonia and Serbia, three NRLs of EFTA countries Iceland, Norway and Switzerland, and three additional participants to compare with their WGS-based results. The main objective of this study was to check the performance of the EU NRLs in serotyping *Salmonella*. All NRLs performed serotyping of the 20 obligatory strains, and all but three of the participants serotyped the optional 21st strain. NRLs of EU Member States that do not achieve the defined level of good performance for serotyping have to participate in a follow-up study.

The typing study included a second pilot of an optional part on cluster analysis. The cluster analysis involved ten *Salmonella* strains, and could be performed up to the choice of the participant by PFGE and/or MLVA and/or WGS (or any combination of these methods), using their own routine procedures. Based on the information gained from the first pilot in 2019, the second pilot was mimicking an outbreak situation, with a monophasic *Salmonella* Typhimurium ST34, MLVA type 3-14-13-NA-211 as the reference strain. Raw WGS data (compressed paired-end fastq files) of this reference strain were made available through a secure ftp server. Participants were asked to analyse the ten strains and to report per strain if a clustering match with the reference strain was found or not.

A total of nineteen NRLs and two external partners participated in the cluster analysis, with two participants using PFGE analysis, six using MLVA analysis and 21 participants using WGS analysis.

2 Participants

Country	City	Institute
Austria	Graz	AGES
Belgium	Brussels	Sciensano
Bulgaria	Sofia	NDRVMI
Croatia	Zagreb	Croatian Veterinary Institute
Cyprus	Nicosia	Cyprus Veterinary Services
Czech Republic	Prague	State Veterinary Institute Prague
Denmark	Ringsted	Danish Veterinary and Food Administration (DVFA)
Estonia	Tartu	Veterinary and Food Laboratory
Finland	Kuopio	Finnish Food Authority
France	Maisons-Alfort	ANSES (Laboratoire de Sécurité des Aliments)
Germany	Berlin	German Federal Institute for Risk Assessment (BFR)
Greece	Chalkida	Veterinary Laboratory of Chalkis
Hungary	Budapest	National Food Chain Safety Office, Food Chain Safety Laboratory Directorate, Microbiological NRL
Iceland	Reykjavík	Landspítali University Hospital, Dept. of Clinical Microbiology
Ireland	Celbridge	Central Veterinary Research Laboratory
Italy	Parma	European Food Safety Authority
Italy	Legnaro	Istituto Zooprofilattico Sperimentale delle Venezie
Latvia	Riga	Institute of Food Safety, Animal Health and Environment (BIOR)
Lithuania	Vilnius	National Food and Veterinary Risk Assessment Institute
Luxembourg	Dudelange	Laboratoire National de Santé
Malta	Valletta	Malta Public Health Laboratory
Netherlands	Bilthoven	RIVM, Centre for Infectious Diseases Research, Diagnostics and Screening (IDS)
North Macedonia Republic of	Skopje	Faculty of Veterinary Medicine Food and feed microbiology laboratory
Norway	Oslo	Norwegian Veterinary Institute
Poland	Pulawy	National Veterinary Research Institute
Portugal	Oeiras	INIAV-Instituto Nacional de Investigação Agrária e Veterinária
Romania	Bucharest	Institute for Diagnosis and Animal Health
Serbia	Belgrade	NIVS Veterinary Institute of Serbia
Slovak Republic	Bratislava	State Veterinary and Food Institute
Slovenia	Ljubljana	UL, Veterinary Faculty, NVI
Spain	Algete-Madrid	Laboratorio Central de Veterinaria
Sweden	Uppsala	National Veterinary Institute (SVA)
Switzerland	Bern	Institute of Veterinary Bacteriology (ZOBA)
United Kingdom	Addlestone	Animal and Plant Health Agency (APHA)
United Kingdom	Belfast	AFBI – Northern Ireland
United Kingdom	London	Public Health England

3 Materials and methods

3.1 Design of the Proficiency Test (PT)

3.1.1 *Laboratory codes*

Each participant was randomly assigned a laboratory code: 1-34 for the NRLs, and 73, 91 and 96 for three additional (WGS) participants.

3.1.2 *Protocol and test report*

Three weeks before the start of the PT, the NRLs received the protocol by email. Web-based result forms were used to report results. Instructions for the completion of these result forms and data-entry were sent to the NRLs on 4 and 8 November 2020, in emails for serotyping and for the second pilot on cluster analysis, respectively. The protocol and blank result forms can be found on the EURL-*Salmonella* website: https://www.eurlsalmonella.eu/proficiency-testing/typing-studies

3.1.3 *Transport*

The parcels containing the strains for serotyping and cluster analysis were sent by the EURL-*Salmonella* on 2 November 2020. All samples were packed and transported as Biological Substance Category B (UN 3373) and transported by a door-to-door courier service.

3.2 Serotyping part of the PT

3.2.1 Salmonella *strains for serotyping*

A total of twenty *Salmonella* strains (coded S1–S20) had to be serotyped by the participants. As agreed at the 25th EURL-*Salmonella* Workshop (Mooijman, 2020), a less common strain (S21) was additionally included. Testing this strain was optional and results were not included in the evaluation. Laboratories were allowed to send strains for serotyping to another specialised laboratory in their country if this was part of their usual procedure.

The *Salmonella* strains used for the part on serotyping originated from the National *Salmonella* Centre collection in the Netherlands. The strains were verified by the Centre before distribution. The complete antigenic formulas of the 21 serovars, in accordance with the most recent White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007), are shown in Table 1. However, participants were asked to report only those results on which the identification of serovar names was based. Thirteen strains (Table 1) represented serovars included in the EURL-*Salmonella* serotyping PTs for the first time.

Table 1 Antigenic formulas of the 21 Salmonella *strains according to the White-Kauffmann-Le Minor scheme used in the EURL-*Salmonella *PT Serotyping 2020*

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Origin
S1 [a)]	13,23	i	e,n,z_{15}	Jukestown	Human
S2 [a)]	1,6,14,25	z_4,z_{23}	[e,n,z_{15}]	Bousso	Non-human
S3	6,8	z_{10}	e,n,x	Hadar	Human
S4 [a)]	1,4,12,27	z_{29}	-	Brancaster	Human
S5 [a)]	8	d	1,2	Virginia	Human
S6 [a)]	9,12	d	z_6	Zega	Chicken
S7	1,13,23	g,m,[s],[t]	-	Agbeni	Human
S8 [b)]	1,4,[5],12	i	-	1,4,[5],12:i:-	Human
S9 [a)]	30	k	e,n,[x],z_{15}	Odozi	Environment
S10 [a)]	1,4,12,[27]	l,[z_{13}],z_{28}	1,5	Tyresoe	Human
S11 [a)]	11	l,v	1,2	Stendal	Non-human
S12 [a)]	4,12,[27]	a	1,5	Hessarek	Chicken
S13	1,4,[5],12	i	1,2	Typhimurium	Human
S14 [a)]	6,7	e,h	1,2	Larochelle	Human
S15	6,7,14	r	1,2	Virchow	Chicken
S16	1,9,12	g,m	-	Enteritidis	Human
S17 [a)]	3,10	b	e,n,x	Benfica	Non-human
S18	6,7,14	r	1,5	Infantis	Chicken
S19 [a)]	4,12,[27]	b	1,6	Canada	Human
S20 [a)]	8,20	z_{38}	-	Apeyeme	Non-human
S21[c)]	50	r	1,5,(7)	50:r:1,5 (IIIb)	Human

[a)] Represented in an EURL-*Salmonella* PT Serotyping for the first time.
[b)] Typhimurium, monophasic variant as determined by PCR.
[c)] *Salmonella enterica* subspecies *diarizonae* (optional strain).

3.2.2 *Evaluation of the serotyping results*

The evaluation of the serotyping results is presented in Table 2.

Table 2 Evaluation of serotyping results

Results	Evaluation
Auto-agglutination or, Incomplete set of antisera (outside range of antisera)	Not typable
Partly typable due to incomplete set of antisera or, Part of the formula (for the name of the serovar) or, No serovar name	Partly correct
Wrong serovar or, Mixed sera formula	Incorrect

In 2007, the following criteria for 'good performance' in PTs on serotyping were defined (Mooijman, 2007).
Penalty points are given for the incorrect typing of strains, but a distinction is made between the five most important human health-related *Salmonella* serovars (as indicated in EU legislation, also sometimes referred to as 'top-5'), and all other strains:

- 4 penalty points: incorrect typing of *S*. Enteritidis, *S*. Typhimurium (including the monophasic variant), *S*. Hadar, *S*. Infantis or *S*. Virchow, or assigning the name of one of these five serovars to another strain;
- 1 penalty point: incorrect typing of all other *Salmonella* serovars.

The total number of penalty points is calculated for each NRL-*Salmonella*. The criterion for good performance is set at less than four penalty points. All EU Member State NRLs not meeting the criterion of good performance (four penalty points or more) have to participate in a follow-up study.

3.3 Cluster analysis part of the PT

3.3.1 Salmonella *strains for cluster analysis*

A total of ten *Salmonella* strains (shipped as SCA01–SCA10, but subsequently indicated as 20SCA01 – 20SCA10) were included in this second pilot on cluster analysis. Background information on the strains is given in Table 3.

Table 3 Background information on the Salmonella *strains used for cluster analysis in 2020*

Strain code	Serovar	ST	MLVA-profile	Origin
20SCA01 [a] (=19SCA09)	4,[5],12:i:-	34	3-13-9-NA-211	Human
20SCA02 [a]	4,5,12:i:-	34	3-14-9-NA-211	Human
20SCA03 [a]	4,5,12:i:-	34	3-15-9-NA-211	Human
20SCA04 [a] [c]	4,12:i:-	34	3-14-13-NA-211	Human
20SCA05 [a] [c]	4,5,12:i:-	34	3-14-13-NA-211	Human
20SCA06 [a] [b] [c] (=REF)	4,5,12:i:-	34	3-14-13-NA-211	Human
20SCA07 (=19SCA07)	Typhimurium	19	5-9-14-9-211	Human
20SCA08 [a] [b] [c]	4,5,12:i:-	34	3-14-13-NA-211	Human
20SCA09 [a]	4,12:i:-	34	3-11-8-NA-211	Human
20SCA10 (=19SCA03)	Typhimurium	19	3-16-7-17-311	Human

[a] Typhimurium, monophasic variant as determined by PCR.
[b] Technical duplicates (in bold).
[c] MLVA-based (in purple) clustering match with the REF strain.

Strains were pre-tested by the EURL-*Salmonella* to be suitable for cluster analysis using either MLVA or WGS. Initially, a set of eleven human surveillance strains, collected and sequenced in 2020 by the National *Salmonella* Centre at RIVM, was selected to be tested for potential use in the PT2020. Five strains from the PT2019 set were also included in the pre-testing. PT2019 strains 19SCA02, 19SCA03, 19SCA07, 19SCA09 and 19SCA10 were cultured from the -70°C stock, prepared from the transport tubes in November 2019. All test strains were freshly cultured on blood-agar plates and a single colony was selected to produce another blood-agar plate which was submitted for MLVA analysis. In addition, material from the same single colony was grown overnight in BHI broth. A cell pellet was made from 1,8 mL overnight culture and resuspended in DNA/RNA Shield (Zymo). This was submitted for WGS analysis on 25 September 2020. Approximately every other day, all test strains were sub-cultured, using alternately liquid (BPW) and solid (blood agar plates) media. The strains were

resubmitted for MLVA and WGS analysis (as described above) after ten times sub-culturing (8 October 2020).

Identical MLVA results were obtained before and after the ten times sub-culturing, and these results also completely matched with the November 2019 MLVA results for the five PT2019 strains.

WGS pre-test results are shown in Annex 5. Sequencing was performed externally, on an Illumina NovaSeq platform. Raw data were processed via an in-house developed pipeline (assembly_pipeline: https://github.com/Papos92), which includes the SPAdes 3.10.0 assembler. Cluster analysis was done in Ridom SeqSphere+, using the cgMLST Enterobase v2.0 scheme and visualised in a minimum spanning tree (MST, Figure A5).

Based on the pre-test results, eight stable strains were selected to be included for the PT Cluster Analysis 2020. In addition, the variable strain 19SCA03 from the PT2019 was added as strain 20SCA10, still showing its variability (Annex 5).
The tenth strain was a technical duplicate; strain 20SCA06 and strain 20SCA08 shipment tubes were both prepared from the same blood-agar plate containing strain 20SCA06. Figure 1 shows the WGS pre-test results as well as the EURL-*Salmonella* PT2020 results for the ten selected strains (Table 3).

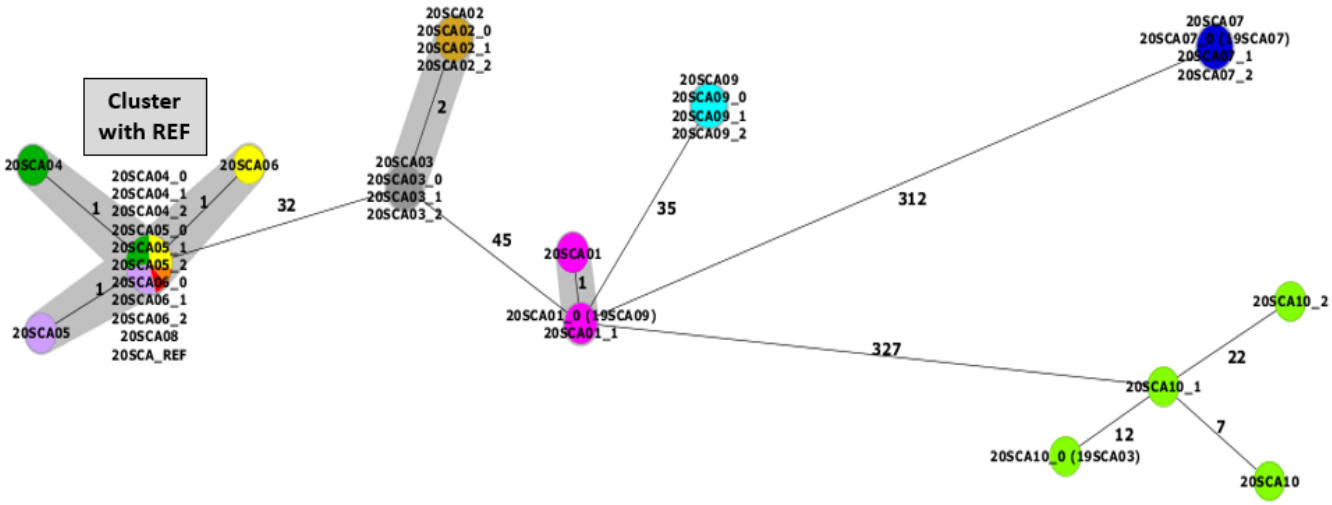


*Figure 1 MST of the EURL-*Salmonella *pre-test and PT2020 results, (RidomSeqSphere+,* S. enterica *MLST (7) and cgMLST (3002), pairwise ignoring missing values). Cluster Alert (in grey background) was set at six allelic differences.*
20SCA_0: Original WGS data from the stored strains (November 2019/Early 2020);
20SCA_1: WGS data from initial pre-testing (25 September 2020);
20SCA_2: WGS data after ten times sub-culturing (8 October 2020);
20SCA without underscore: PT2020 data (18 November 2020).

3.3.2 *Evaluation of the cluster analysis results in general*

Cluster analysis was performed up to the choice of the participant by PFGE and/or MLVA and/or WGS (or any combination of these methods), using their own routine procedures. However, the Protocol of the PT

Typing 2020 already indicated that PFGE is no longer performed at the EURL-*Salmonella* and evaluation of PFGE results would only be based on comparing the results as sent in by PFGE participants.
The pilot PT Cluster Analysis 2020 was mimicking an outbreak situation, with a monophasic *Salmonella* Typhimurium ST34, MLVA type 3-14-13-NA-211 as the reference strain. Raw WGS data of this strain (compressed paired-end fastq files) were made available through a secure ftp server. For this particular PT2020 situation, the cluster definition was set at maximum six allelic differences from the reference sequence (REF). For MLVA, the cluster definition was set at no loci with a different number of repeats.
Participants were asked to analyse the ten *Salmonella* strains and to report per strain if a clustering match with the reference strain was found or not. Details on the method(s) used and the outcome of the cluster analysis had to be reported in the electronic result form. Additionally, specific data for PFGE and WGS had to be sent by email or uploaded to a secure ftp server.
Evaluation (per methodology, see sections 3.3.3 – 3.3.5) of the participants' cluster analysis results was performed by comparing the participants' results to the expected results in the outbreak investigation setting, as pre-defined by the EURL-*Salmonella.*
No performance criteria were set for this second pilot PT on cluster analysis. As a minimum, it was expected that participants would report the technical duplicate strains 20SCA06 and 20SCA08 to be (part of) one cluster.

3.3.3 Evaluation of the cluster analysis results based on PFGE data

Data submission for PFGE results included:

- **Electronic result form**: protocol used, position of the lanes, potential cluster identification in case of an outbreak situation.
- **The PFGE gel image** had to be emailed as an uncompressed 8-bit grey scale TIFF file to the EURL-*Salmonella*. The laboratory code had to be included in the name of the .tif file, for example: Lab01_PFGE2020.tif.
- The ZIP export files were prepared from the analysis in BioNumerics, *including all test strains and reference strains*, *as well as the TIFF image*. **The BioNumerics analysis data** had to be emailed in a ZIP file to the EURL-*Salmonella*. The zip file had to include the laboratory code in the name, for example: Lab01_PFGE2020.zip.

Because PFGE is no longer performed at the EURL-*Salmonella*, evaluation of PFGE results could only be based on comparing the results as sent in by PFGE participants (see section 4.3.2).

3.3.4 Evaluation of the cluster analysis results based on MLVA data

Data submission for MLVA results included:

- **Electronic result form**: scheme/loci used, the allelic profile, cluster identification in case of an outbreak investigation.

Participants were asked to report per strain (Table 3) if a clustering match was found with the reference outbreak strain (REF) in the EURL-*Salmonella* PT Typing 2020: monophasic *Salmonella* Typhimurium, ST34, MLVA type 3-14-13-NA-211.

The MLVA cluster definition for the PT Typing 2020 was set at no loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 20SCA04, 20SCA05, SCA06 (reference strain) and 20SCA08 (technical duplicate of the reference strain) to be a clustering match with the REF outbreak strain.

3.3.5 Evaluation of the cluster analysis results based on WGS data

Data submission for WGS results included:

- **Electronic result form**: background information on the wet-lab and dry-lab methods used, cluster identification in case of an outbreak investigation (SNP-based and/or cgMLST/wgMLST-based).
- **Raw reads** (compressed fastq files) uploaded to the secure ftp server according to the instructions.
- **The distance matrix** emailed to the EURL-*Salmonella*.

Participants were asked to report per strain (Table 3) if a clustering match was found with the reference outbreak strain (REF) in the EURL-*Salmonella* PT Typing 2020: 20SCA_REF_R1.fq.gz and 20SCA_REF_R2.fq.gz (monophasic *Salmonella* Typhimurium, ST34, MLVA type 3-14-13-NA-211).

The WGS cluster definition for the PT Typing 2020 was set at maximum six allelic differences from the reference (REF). Based on this cluster definition, WGS-based results were expected to indicate strains 20SCA04, 20SCA05, SCA06 (reference strain) and 20SCA08 (technical duplicate of the reference strain) to be a clustering match with the provided REF outbreak strain (also see Figure 1).

4 Results and Discussion

4.1 Technical data

4.1.1 General

A total of 37 laboratories participated in this study (Chapter 2). These included 29 NRLs-*Salmonella* in the 27 EU Member States plus the United Kingdom, two NRLs of EU candidate countries, and three NRLs of EFTA countries. Data from three additional participants (Laboratory codes 73, 91, and 96) were included to compare with their WGS-based results.
The frequency of *Salmonella* serotyping at the participating laboratories and the number of strains serotyped in 2020 are summarised in Table 4.

Table 4 Frequency and number of Salmonella *strains serotyped in 2020*

Laboratory code	Serotyping frequency in 2020	No. of strains serotyped in 2020
1	Daily	130
9	Daily	160
10	Daily	315
18	Daily	328
21	Daily	330
14	Daily	400
26	Daily	400
30	Daily	450
5	Daily	500
11	Daily	550
19	Daily	600
28	Daily	600
13	Daily	800
32	Daily	850
24	Daily	900
31	Daily	1100
25	Daily	2000
23	Daily	2423
33	Daily	2500
12	Daily	3300
8	Daily	4400
2	Daily	4500
34	Daily	4500
17	Daily	5000
16	Once a week	20
22	Once a week	300
29	Once a week	1100
3	Once a week	2000
7	Twice a week	90
20	Twice a week	90
6	Twice a week	650
15	Twice a week	659

Laboratory code	Serotyping frequency in 2020	No. of strains serotyped in 2020
4	Thrice a week	73
27	Depends on programs	300
n=34		42318

4.1.2 Accreditation

Of the 34 participants, 33 are accredited for serotyping *Salmonella.* Thirty-one according to EN ISO/IEC 17025, and three (also) according to EN ISO 15189. One laboratory mentioned ISO 6579-1 only. The one non-EU laboratory not accredited for serotyping is known for this because of their relatively low numbers of serotyping strains.
All 33 laboratories stated that they are accredited for all *Salmonella* serovars.

4.1.3 Transport of samples

All but two participants received their package within two days after shipment on Monday 2 November 2020. One package was received in the laboratory on 6 November and the final one on 9 November 2020. All packages were received in good condition.

4.2 Serotyping results

4.2.1 General

The twenty obligatory strains were all tested by the NRLs-*Salmonella* in the participating countries, strain S20 was forwarded to their national typing centre by laboratory 19. Classical serology was used by 33 participants, six of them mentioned the combined use of classical serology and Luminex assays (3) or multiplex/real time PCR (3). One participant used Whole Genome Sequencing (WGS).
Additional data were obtained from participants with the Laboratory codes 73, 91, and 96, all using WGS in their routine serotyping.
Details on the number and the source of the sera used by the participants are summarised in Tables 5a and 5b.

Table 5a Number of laboratories using sera from various manufacturers

Manufacturer	Number of NRLs (n=33)
Biorad	16
Pro-Lab	5
Sifin	17
Statens Serum Institute (SSI)	29
Other	5
Own preparation	3

Table 5b Number of laboratories using sera from one or more manufacturers and/or in-house prepared sera

Number of manufacturers from which sera are obtained (including in-house preparations)	Number of NRLs (n=33)
1	6
2	13
3	13
4	1

4.2.2 *Biochemical testing*

Thirty participants indicated the use of biochemical tests. Details are given in Table 6, with specific attention for strain S12 (in purple), which was biochemically tested by 24 participants.

4.2.3 *Use of PCR for confirmation*

Sixteen laboratories used PCR to confirm strain S8, the monophasic variant of *S*. Typhimurium 1,4,[5],12:i:-, and ten of these (including WGS participant 73) also used PCR to confirm strain S13, *S*. Typhimurium. The majority of laboratories mentioned using the reference by Tennant et al., 2010.

4.2.4 *General comments on the PT 2020 serotyping evaluation*

Selection, preparation and shipment of the strains to the participants is always carried out with upmost care, and includes various quality control steps, including purity and typeability. This year, at least ten participants mentioned some or even many strains to be difficult to type, showing rough colonies, which needed additional passages (e.g. using U-tubes) before successful typing. Apart from strain S12, these problems could not be linked to specific strains, and no common cause could be determined.
Strain S12 (*S*. Hessarek) was often mentioned to show both smooth and rough colonies, and sometimes gave inconclusive results in the biochemical tests (Labs 10 and 28, with positive results for rhamnose, trehalose, gas from glucose but negative results for dulcitol, H_2S and Simmons citrate, also see section 4.2.2).
However, none of the twenty strains had to be excluded from the evaluation.

Table 6 Biochemical tests used by 30 participants on various strains and indicated by their number; strain 12 (S. Hessarek) in purple

Lab code	beta-glucuronidase	Dulcitol	Galacturonate	Gelatinase	Gas from Glucose	Glucose	H2S	Lactose	Lysine Decarboxylase	Malonate	Mucate	ONPG	Rhamnose	Salicine	Simmon's citrate	Sorbitol	Tartrate	Trehalose	TSI
1			12, P	12, P						12, P						12, P	12, P		
2							12			12, B			12		12				
3		12, F			12					F		F	12	F	12	F			
4		12, L				12	12			L			12						
5		K								K							K		
7																	4, 17		
8g)																			
9b)																			
10d)																			
11		12, D	12, D			12, D	12, D			12, D	12, D			12, D		12, D		12, D	
12		M, 12					12	21		M	12		12						
13		2, 12							12, A	12, D		12, A	12		12, A		2, D	12	
14	21	G						21		G		21		G					
15e)																			
17										N			12				12	12	
18c)																			
19		12, E	12, E							12, E	12, E	12, E				12, E			
21						12	12			17			12		12				
22a)																			
23		12, A		J						C									
24		H								H				H					
25		X									X							X	
26		12								12, D		21							
28d)																			
29										17							2, 12		
31		12								21		21							12
32		12			12		12			4			12		12			12	
33		12, K								12, K		12, K		12, K		12, K			
34f)																			
73		12			12								12		12				

a)	Strain S12 tested, but tests not stated	A	All strains S1-S21
b)	Strain S17 tested, but tests not stated	B	S2, S4, S5, S, S7, S9, S10, S17, S19, S21
c)	Strain S-21 tested, but tests not stated	C	S4, S6, S7, S9, S10, S17, S21
d)	Strains S12 and S21 tested, but tests not stated	D	S4, S9, S17, S19, S21
e)	Strains S12, S17 and S21 tested, but tests not stated	E	S4, S9, S17, S20, S21
f)	Strains S4, S12, S17 and S21 tested, but tests not stated	F	S2, S3, S17, S21
g)	All strains S1-S21 tested, but tests not stated	G	S2, S4, S17, S21
		H	S2, S8, S17, S21
X	tested, but strains not stated	J	S2, S9, S17, S21
		K	S4, S9, S17, S21
		L	S1, S4, S9
		M	S4, S7, S17
		N	S4, S9, S17
		P	S17, S21

With regard to strain S12, the White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007) states:
"Serovars Hessarek (4,12,27 :a :1,5) and Fulica (4,[5],12 :a :[1,5]), which formula could be similar, are not combined because they differ by biochemical characters. Rhamnose, gas production from glucose, dulcitol, trehalose, Simmons citrate, L(+) tartrate (=*d*-tartrate), mucate, H_2S, and tetrathionate-reductase are positive for Hessarek and negative for Fulica. This latter serovar is very rare."

The EURL-*Salmonella* extensively tested strain S12: on rhamnose, gas from glucose, dulcitol, Simmons citrate, mucate, and TSI (H_2S), all with a positive result.

4.2.5 *Serotyping results per laboratory*

The percentages of correct results per laboratory are shown in Figure 2. The evaluation of the type of errors for O- and H-antigens and for identification of the strains are shown in Figures 3, 4 and 5.
The O-antigens were completely typed correctly by 29 of the 37 participants (78%). This corresponds to 99% of the total number of strains. The H-antigens were completely typed correctly by 31 of the 37 participants (84%), corresponding to 98% of the total number of strains. As a result, 28 participants (76%) gave completely the correct serovar names, corresponding to 97% of all strains evaluated.

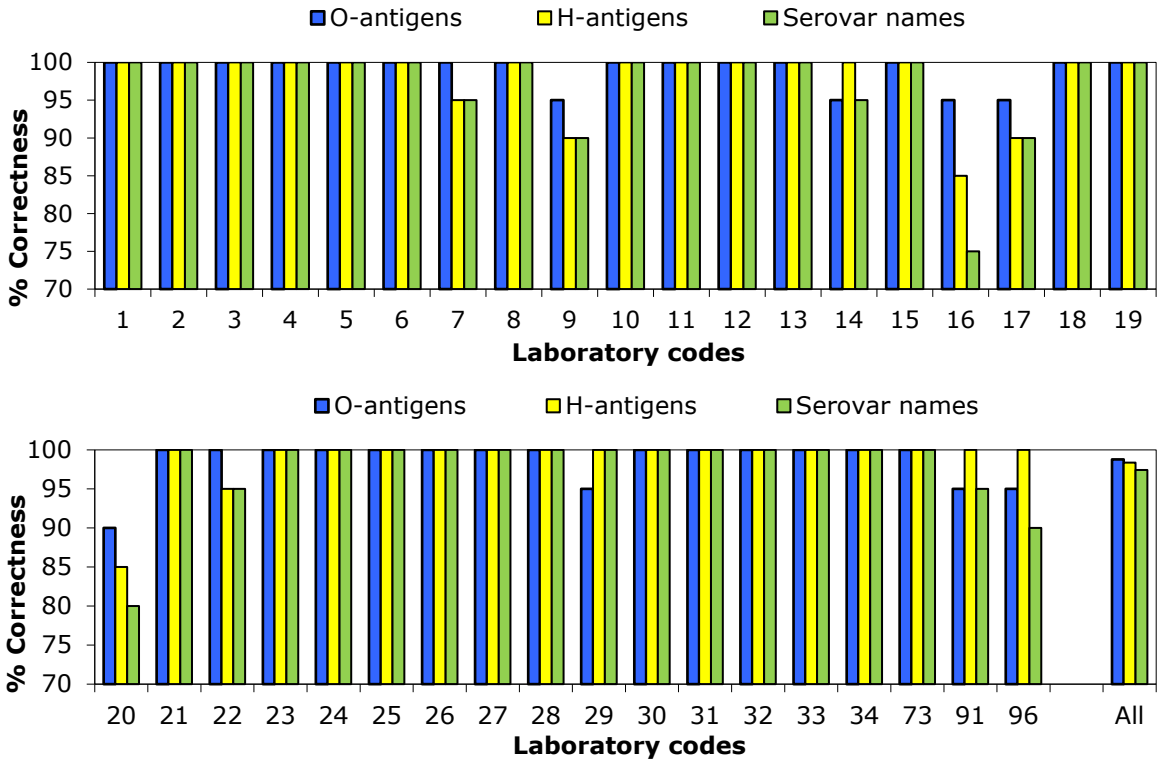


Figure 2 Percentages of correct serotyping results, per participant

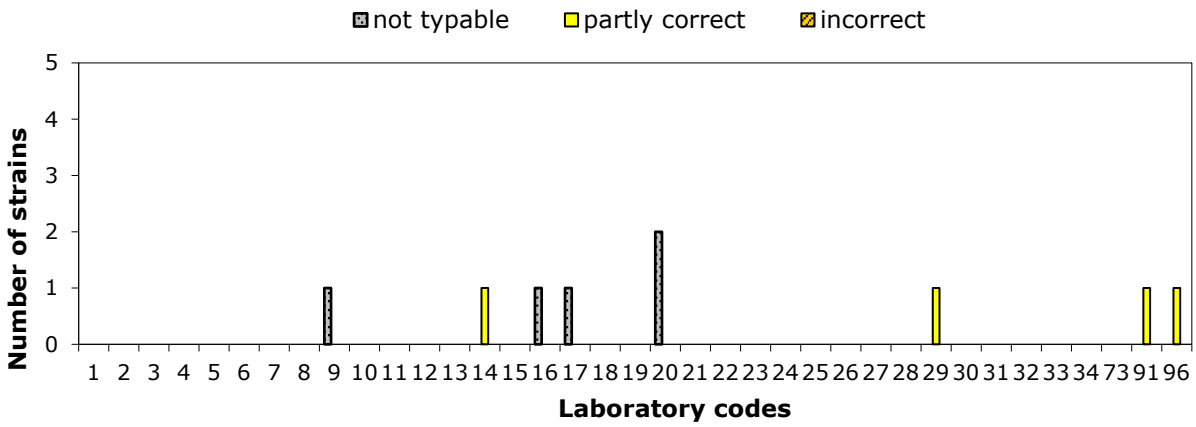


Figure 3 Evaluation of type of errors for O-antigens, per participant

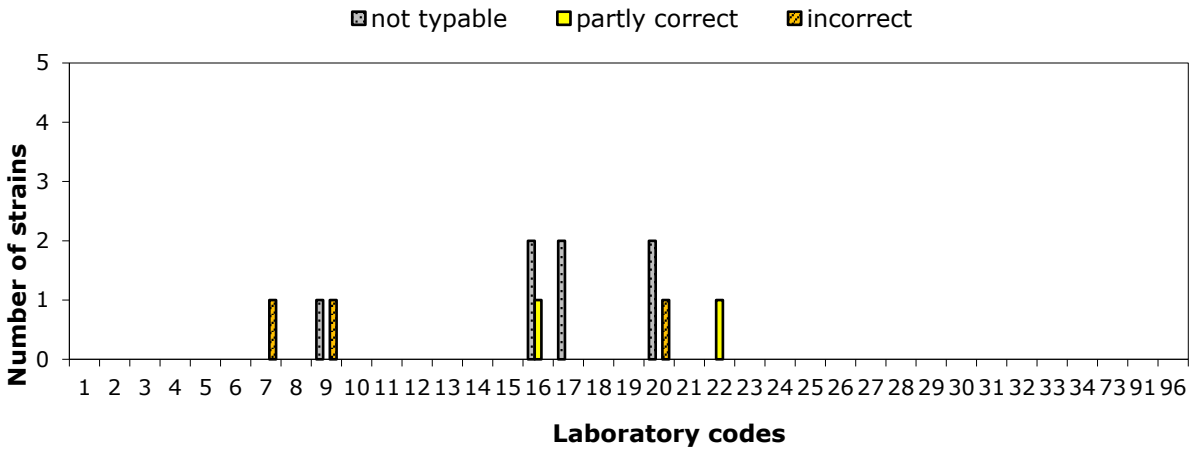


Figure 4 Evaluation of type of errors for H-antigens, per participant

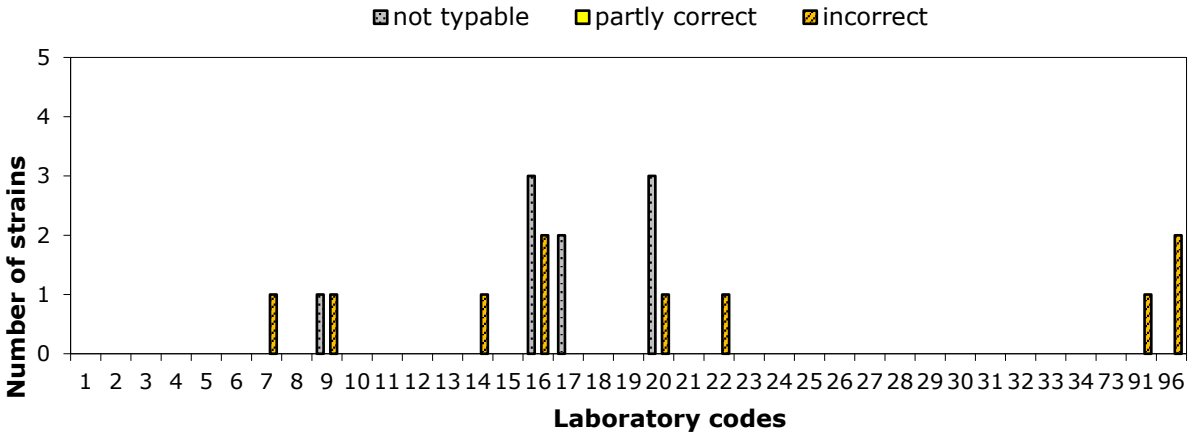


Figure 5 Evaluation of the type of errors in the identification of the serovar names, per participant

4.2.6 *Performance of the participants*

The number of penalty points was determined for each NRL using the guidelines described in Section 3.2.2. Table 7 shows the number of penalty points for each NRL and indicates whether the level of good performance was achieved (yes or no).

Overall, the performance of the NRLs in the PT Serotyping 2020 was very good, including the performance of 4 participants that were submitting WGS-based results. All participants met the level of good performance at the first stage of this PT, and there was no need to organize a follow-up study.

All participants received their individual laboratory evaluation report as well as the interim summary report on serotyping on 11 March 2021. An example of an individual laboratory evaluation report on serotyping results is given in Annex 1. The interim summary report is available on the EURL-*Salmonella* website: www.eurlsalmonella.eu/publications/proficiency-test-reports.

Table 7 Evaluation of serotyping results per NRL

Laboratory code	Penalty points	Good performance	Laboratory code	Penalty points	Good performance
1	0	yes	20	1	yes
2	0	yes	21	0	yes
3	0	yes	22	1	yes
4	0	yes	23	0	yes
5	0	yes	24	0	yes
6	0	yes	25	0	yes
7	1	yes	26	0	yes
8	0	yes	27	0	yes
9	1	yes	28	0	yes
10	0	yes	29	0	yes
11	0	yes	30	0	yes
12	0	yes	31	0	yes
13	0	yes	32	0	yes
14	1	yes	33	0	yes
15	0	yes	34	0	yes
16	2	yes	73	0	yes
17	0	yes	91	1	yes
18	0	yes	96	2	yes
19	0	yes			

4.2.7 *Serotyping results per strain*

The final naming results reported per strain (S1 – S20) and per laboratory are given in Annex 2. A completely correct identification was obtained for nine *Salmonella* serovars: Bousso (S2), Hadar (S3), Zega (S6), Typhimurium (S13), Larochelle (S14), Virchow (S15), Enteritidis (S16), Benfica (S17), and Infantis (S18).
The reported serovar names for strain 1,4,[5],12:i:- (S8) are also shown in Annex 2. Sixteen participants (including WGS-participant 73) used a PCR method to confirm this strain to be a monophasic Typhimurium strain.

Details on the strains that caused problems or inconsistencies in serotyping are shown in Annex 3. Interestingly, some inconsistencies were seen in the submitted results for strains S3 (Hadar) and S5 (Muenchen), especially by the four participants that were using WGS (laboratory codes 29, 73, 91, and 96). Both serovars belong to the pairs of serovars in *Salmonella* serogroup C_2 which differ only by the minor antigen $O:6_1$ and that may show variable expression (also described as "colonial form variation", Hendriksen et al., 2009; Mikoleit et al., 2012). Laboratory 73 reported to confirm separately for presence of O:6. Retrospectively, also laboratory 29 reported to have checked for the presence of both O:6 and O:8 by agglutination and both antigens were found. Therefore the O-antigens should have been reported as 6,8 and not just 8 as mistakenly done initially. The other two laboratories may not have this particular option of additional testing in their routine WGS pipelines/protocols.

Details on the additional and optional strain S21 are given in Annex 4. All but three participants tried to serotype strain S21, a *Salmonella enterica* subsp. *diarizonae* (IIIb). Some laboratories did not have access to the required antisera to finalise this (50:r:1,5).

4.2.8 Trend analysis of the serotyping results of the EU NRLs

Historical data for all participants of the EURL-*Salmonella* PTs on the serotyping of *Salmonella* can be found on the EURL-*Salmonella* website: www.eurlsalmonella.eu/.

The historical data on the EU NRLs only are visualised in Figure 6, showing the percentages of correctly typed strains, and in Figure 7, showing the number of penalty points and non-good performance. The percentages of correctly typed strains are stable over time, usually showing a better performance for the O-antigens than for the H-antigens.

The number of penalty points has clearly declined, from 35 points when this system started in 2007, to three points in the 2020 study. The rise as seen for the 2018 study was mainly caused by the relatively large number of seven EU NRLs that made a mistake in typing a *S.* Cannstatt strain. Moreover, the number of EU NRLs with a non-good performance is low: two in the period 2010 – 2013, one in the 2014, 2015 and 2018 studies, and none in the 2016, 2017, 2019 and 2020 studies.

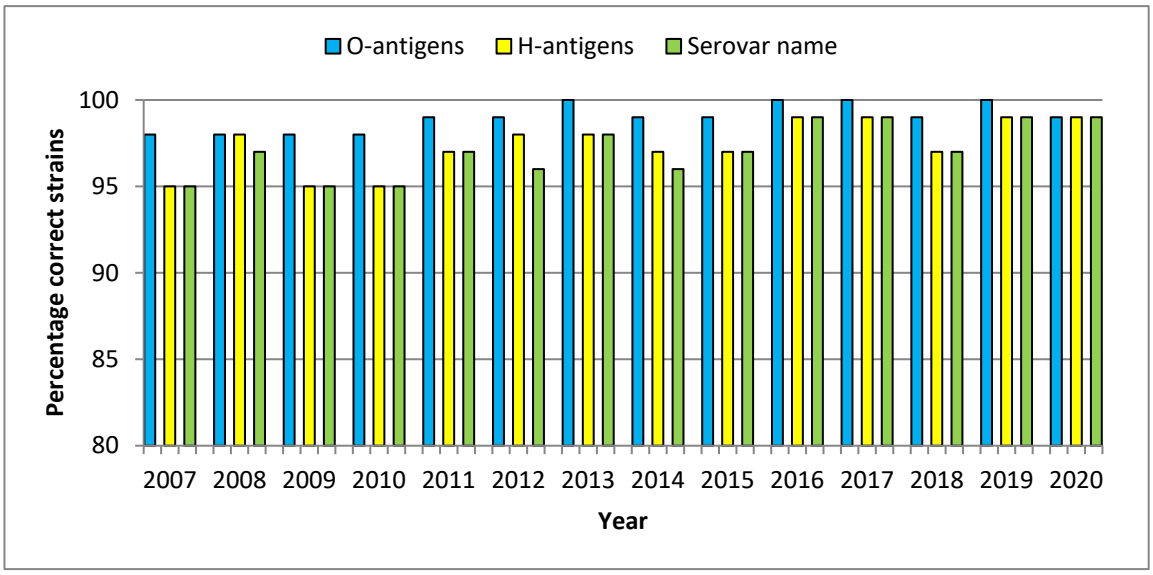


Figure 6 Serotyping results of the EU NRLs, based on the percentages of correctly typed strains

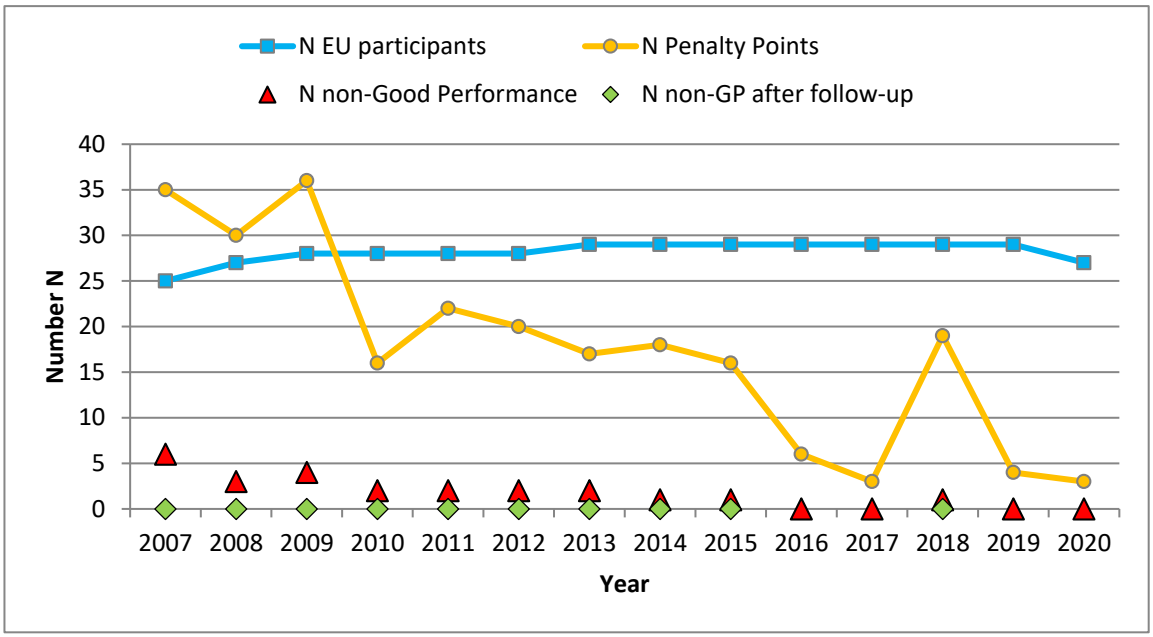


*Figure 7 Serotyping results of the EU NRLs-*Salmonella*, based on the number (N) of Penalty Points and non-Good Performance (non-GP)*

4.3 Cluster analysis results

4.3.1 General

Cluster analysis was performed up to the choice of the participant by PFGE and/or MLVA and/or WGS (or any combination of these methods), using their own routine procedures.
A total of nineteen NRLs and two external partners participated in the cluster analysis; two participants used PFGE analysis, six used MLVA analysis and 21 used WGS analysis (Table 8).
All participants received their individual laboratory evaluation report of the second pilot on cluster analysis on 27 May 2021, together with the interim summary report on the overall results. An example of an individual laboratory evaluation report on cluster analysis results is given in Annex 6. The interim summary report is available on the EURL-*Salmonella* website:
https://www.eurlsalmonella.eu/publications/proficiency-test-reports.

As a general question, the participants were asked if they serotyped the ten strains. Fifteen participants indicated to have serotyped the strains. These serotyping results are given in Annex 7, for information purposes only.

Table 8 Participation in PT Cluster Analysis in 2020, per method or combination of methods used

Method used:			Number of participants	Laboratory codes
		WGS	15	1, 2, 3, 6, 12, 14, 18, 19, 21, 24, 25, 32, 34, 91, 96
	MLVA	WGS	4	8, 11, 28, 31
PFGE	MLVA	WGS	2	17, 33
Total PFGE:	Total MLVA:	Total WGS:	Total overall:	
2	6	21	21	

4.3.2 *Results cluster analysis based on PFGE data*

Only two participants (Laboratory codes 17 and 33) submitted results based on PFGE data and were using BioNumerics for the cluster analysis. The combined data sets in BioNumerics are shown in Figures 8 and 9. Initially, similarity was calculated as recommended by EFSA (Jacobs et al., 2014) using the Dice coefficient, with both tolerance and optimisation at 1,5% (Figure 8). However, the optimal setting for this specific analysis appeared to be an adjusted setting with tolerance and optimisation at 1% (Figure 9), as was remarked by participant 33. By using these adjusted settings, the PFGE clustering would also match with both the MLVA-based and the WGS-based clustering. Clarification by participant 17 revealed that they used their standard tolerance and optimisation at 1%, which explains the result as reported by this participant (Table 9).

Based on Figure 9, both PFGE participants' results confirm a clustering match of the REF strain 20SCA06 with strains 20SCA04, 20SCA05, and 20SCA08 (the technical duplicate).

The technical duplicates SCA06/SCA08 were expected to be reported as (part of) one cluster and this was done by both participants (Table 9).

Table 9 Number of clusters, and their identification reported by the two PFGE participants

Laboratory code	# Clusters reported	Cluster 1
17	1	SCA04;SCA05;SCA06;SC08 [a)]
33	1	SCA03-SCA04-SCA05-SCA06-SCA08 [b)]

a) After clarification: tolerance and optimisation were set at 1%

b) Remark by laboratory 33: adjusting the tolerance and optimisation to 1%, strain 20SCA03 would not be included in the cluster anymore

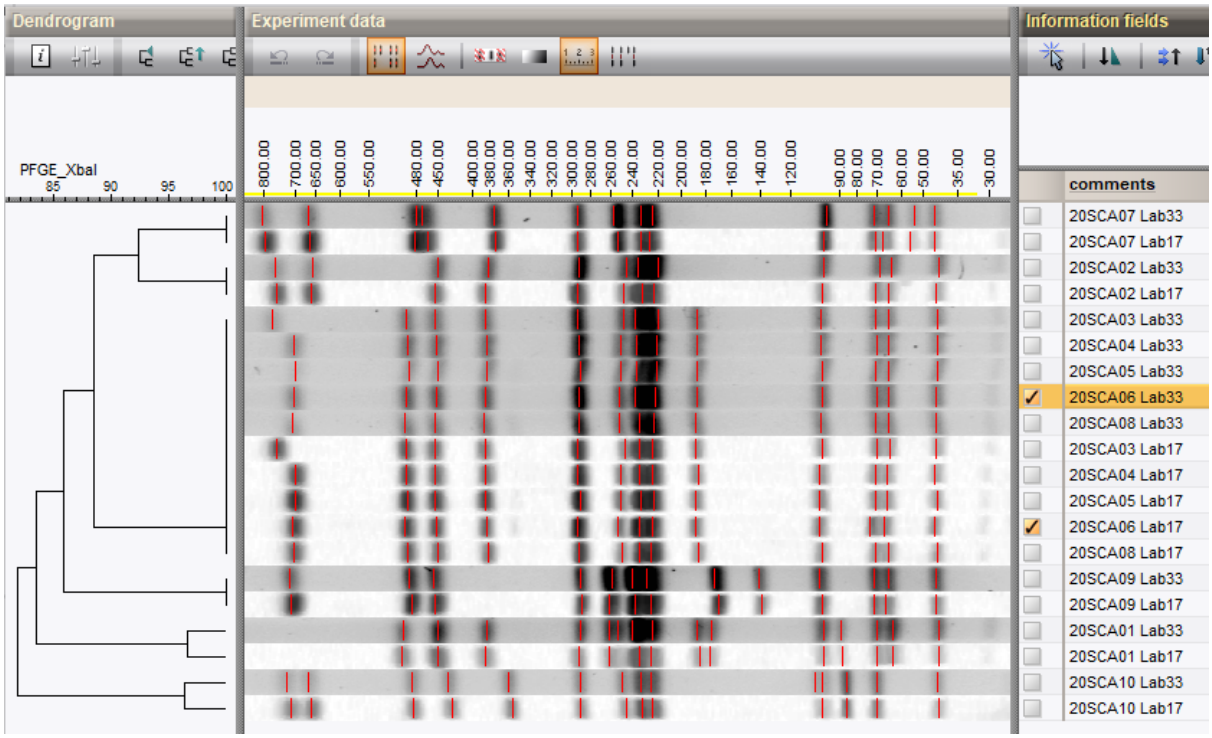


Figure 8 Cluster analysis based on PFGE data from Laboratory codes 17 and 33, using the Dice coefficient, with both tolerance and optimisation ***at 1,5%***

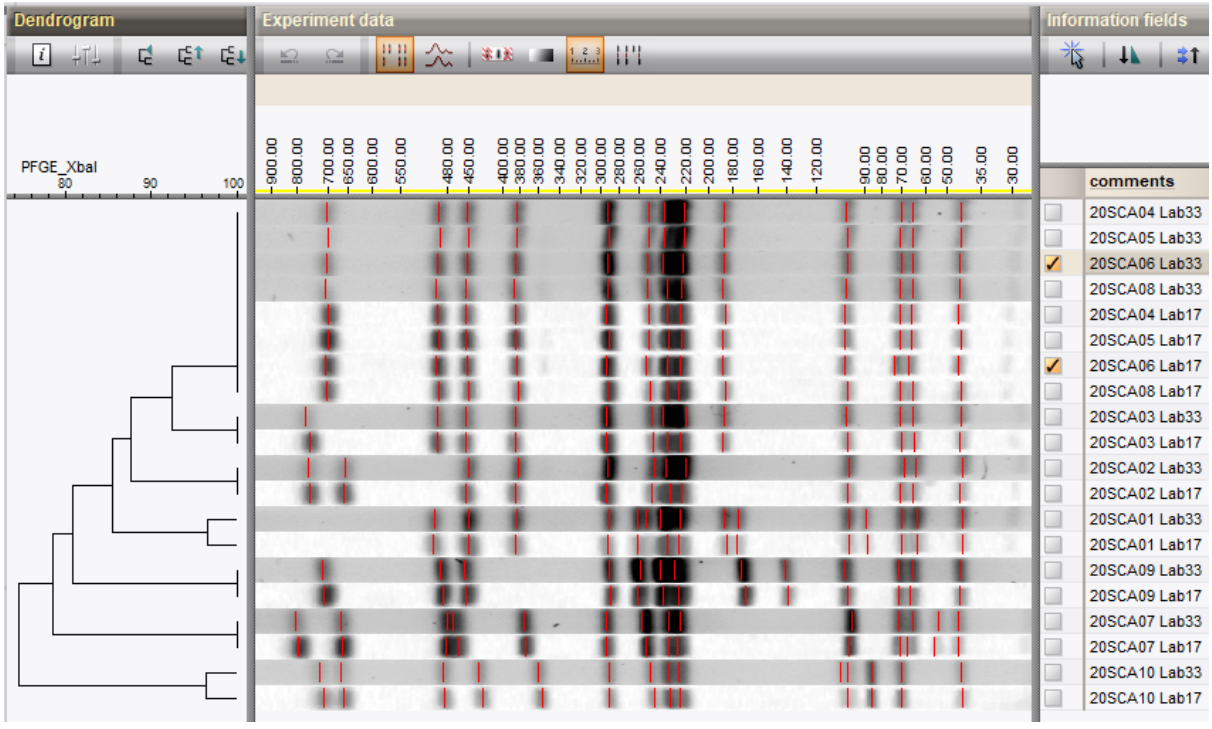


Figure 9 Cluster analysis based on PFGE data from Laboratory codes 17 and 33, using the Dice coefficient, with both tolerance and optimisation ***at 1,0%***

4.3.3 *Results cluster analysis based on MLVA data*

Six participants (Laboratory codes 8, 11, 17, 28, 31, and 33) submitted cluster analysis results based on MLVA data.
The allelic profiles submitted by the participants are given in Annex 8.

Participants were asked to report per strain if (yes or no) a clustering match was found with the reference outbreak strain (REF) in the EURL-*Salmonella* PT Typing 2020: monophasic *Salmonella* Typhimurium, ST34, MLVA type 3-14-13-NA-211.
The MLVA cluster definition for the PT Typing 2020 was set at no loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 20SCA04, 20SCA05, SCA06 (reference strain) and 20SCA08 (technical duplicate of the reference strain) to be a clustering match with the REF outbreak strain.

Five participants (Laboratory codes 8, 17, 28, 31, and 33) out of the six submissions reported the MLVA-based cluster analysis results completely as expected (Table 10).
Laboratory 11 reported incorrect results for strains 20SCA05 and 20SCA07, which was retrospectively clarified by a mistake in the identification of the samples (also see Annex 8).

The technical duplicates SCA06/SCA08 were expected to be reported as (part of) one cluster and this was done by all six participants (Table 10).

Table 10 Expected cluster analysis results and the cluster analysis results reported by the six MLVA participants

Lab code	20 SCA01	20 SCA02	20 SCA03	20 SCA04	20 SCA05	20 SCA06	20 SCA07	20 SCA08	20 SCA09	20 SCA10
Expected	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**	**Yes**	**No**	**No**
8	No	No	No	Yes	Yes	Yes	No	Yes	No	No
11	No	No	No	Yes	No	Yes	Yes	Yes	No	No
17	No	No	No	Yes	Yes	Yes	No	Yes	No	No
28	No	No	No	Yes	Yes	Yes	No	Yes	No	No
31	No	No	No	Yes	Yes	Yes	No	Yes	No	No
33	No	No	No	Yes	Yes	Yes	No	Yes	No	No

Deviation from the expected result

4.3.4 *Results cluster analysis based on WGS data*

Twenty-one participants (Table 8) submitted cluster analysis results based on WGS data; two participants (two laboratory codes each: 2/82 and 6/86) submitted both cgMLST-based and SNP-based data.

General details on the wet-lab and dry-lab protocols performed by the participants and the EURL-*Salmonella* (EL) are given in Annex 9. All participants performed DNA extraction, library preparation and sequencing in-house, except for participants 14, 19 and 2/82 (library preparation and sequencing outsourced) and participants 11, 18, 96 and the EURL-*Salmonella* (all outsourced). The Illumina MiSeq platform was used most often (12x), followed by the Illumina NovaSeq or NextSeq

(4x each), and Illumina MiniSeq or HiSeq (1x each). Including the EURL-*Salmonella*, 16 participants used cgMLST for data analysis and 8 participants used SNP-based analysis (6x reference-based and 2x assembly-based).
Tools used for this analysis varied from in-house (chewBBaca-based) pipelines to commercial ones, most often Ridom SeqSphere (7x). Minimum Spanning Tree (MST, 14x), followed by both Maximum likelihood (ML, 4x) and Neighbor joining (NJ, 3x) were commonly used for the cluster analysis.

All participants' Quality Criteria (QC) parameters reported for the evaluation of their data are listed in Annex 10. A variety in naming these QC parameters, and in the thresholds used, was observed. An overview of the most widely used (names of) parameters is given in Table 11.

Twelve participants reported the md5 checksum for the compressed paired-end fastq files of the REF and these were correct for ten participants, indicating that the transfer of data from the secure ftp site went alright.

20SCA_REF_R1.fq.gz: 257eece96dfe3169c2e1f00e797c1dca
20SCA_REF_R2.fq.gz: 29adc5a5b60e3e96ca69fdf31cfc1022

Two participants reported md5 checksums that deviated from the expected ones. The md5 checksums reported by one of the two participants matched with the ones of the uncompressed fastq files:
20SCA_REF_R1.fq: 455b0fcf8dfa37edc5b19c2fa7050900
20SCA_REF_R2.fq: e990142ae84307676c89fade41b87e76

Table 11 Participants' most widely used QC parameters

Criteria indicated	**Thresholds (# participants)**
Genome size (sometimes also indicated as e.g. Assembly length, Total length, Contamination)	4.6-5.2 Mbases (2); ~5MBases (2); 4,5 - 5,2 Mb (2); 4,5 - 5,5 Mb (1); ; 4,8 - 5,6 Mbp (1); [3.6 Mb, 6.0 Mb] (1); 4 - 5,8 Mbp (1); Deviation <0,5 million bp from the expected genome size (1); +/- 20% (1); Length of contigs assembled<ref genome + 10% (1)
Assembly contamination	Completeness > 99.0 Contamination < 2.0 (1)
Contamination check	No threshold given (3); <4% (2); > 5% contaminating species = fail (1); Rejected if there is >10% contamination (1); Around 10% (1); > 75% Salmonella (1); Identity ≥ 0,95 (1);
Contamination of genomic sequences	Pure bacterial culture (1)
Fraction of reads uniquely assigned to Salmonella enterica	> 0.90 (1)
Genome fraction %	>90% (1)
Contamination (on fastq)	TrueCoverage_absente_genes < 2; TrueCoverage_multiple_alleles <1 Confindr_Genus "Salmonella" Confindr_NumContamSNVs < 30 (1)

Coverage	>30x (8); >50 (2); >20 (1); Minimum 25x (2); ≥10 (1); Minimum 20-30x (1); min 30x, max 100x (1); 90x (1)
Avg. coverage (assembled)	50x but if it's less, the % of good targets should be >95% (1)
Coverage (after mapping)	Avg cov > 25 (1)
Breath coverage	Min coverage: 80% (1)
Coverage cgMLST	≥90% (1)
Number of contigs (>1000)	<100 (1)
Number of contigs (>200)	< 250 (1)
Number of contigs	<500 (4); <300 (3); No threshold (yet) (2); <115 (1); ≤400 (1)
N50	>10 000 (3); >15 000 (minimum) (3); >30 000 bp (2); >20 Kb (1), >48 230 bp (1), >50 000 (1), >55 000 (1), >80 000 (1), >200 000 (1)

All but one of the participants' raw data (fastq files) were successfully processed through the in-house assembly pipeline as discussed in section 3.3.5. Raw data from participant 28 were processed using a Unicycler assembly pipeline (Galaxy Version 0.4.8.0), because this concerned single-end fastq files which cannot be analysed by the in-house assembly pipeline. All *de novo* assembled genomes (fasta files) were analysed in Ridom SeqSphere⁺, using the cgMLST Enterobase v2.0 and visualised in a MST (Figure 10). Data per strain are given in Annex 11.

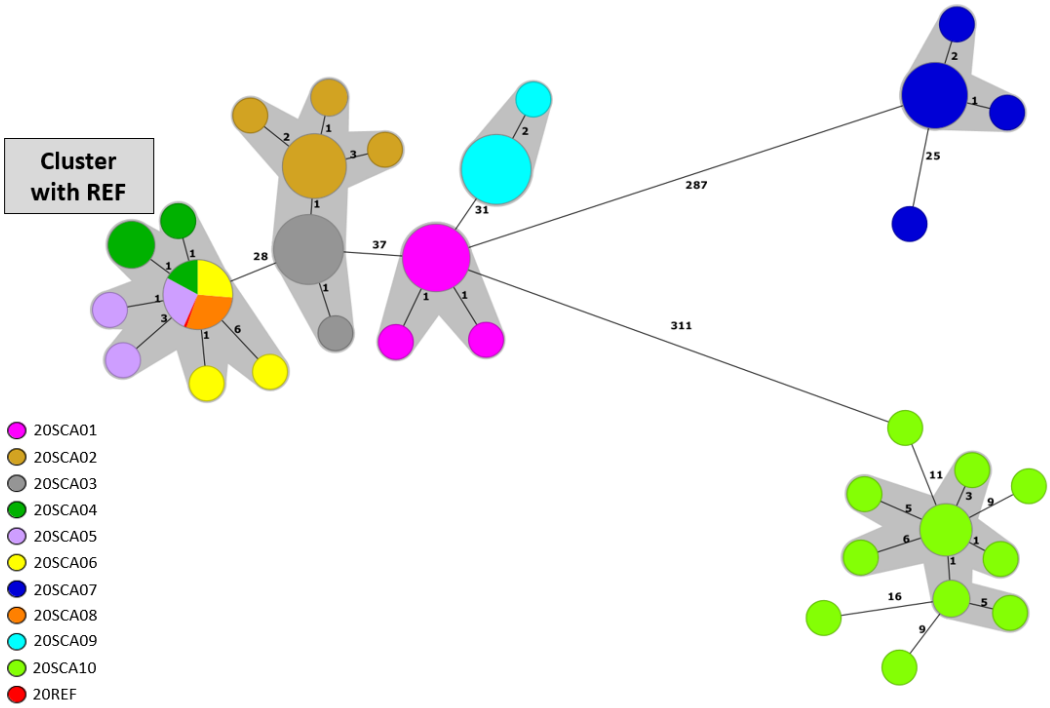


Figure 10 MST of all strains from all participants' processed raw data (Ridom SeqSphere⁺, S. enterica *MLST (7) and cgMLST (3002), pairwise ignoring missing values)*

An overview of the main QC parameters results on all in-house *de novo* assembled genomes (fasta files) is given in Table 12, summarised per participant and in Table 13, summarised per strain. Detailed data per participant are given in Annex 12.

Table 12 Results QC parameters on the de novo assembled genomes, average per participant

Laboratory code	Average # contigs	Average Largest contig	Average Total length	Average N50	Average Coverage
Lab01	80	648058	4950288	235324	53
Lab02	189	459373	4957966	87684	962
Lab03	75	764556	4951969	243991	283
Lab06	65	806224	4962002	294076	116
Lab08	85	700378	4970199	245535	106
Lab11	91	574166	4966062	204893	49
Lab12	80	579582	4943425	253092	88
Lab14	78	632938	4950235	247145	219
Lab17	145	836515	5020191	300054	63
Lab18	181	428233	4903240	141743	63
Lab19	80	523877	4963096	215384	194
Lab21	80	507649	4950226	218237	39
Lab24	70	814426	4967022	304022	131
Lab25	109	806206	4986081	297366	216
Lab31	91	616943	4963903	215376	119
Lab32	76	773030	4967867	280425	64
Lab33	122	762146	5003337	264416	171
Lab34	221	278147	4965411	93662	102
Lab91	79	558815	4949754	242570	112
Lab96	82	543597	4941781	232017	205
EL PT (18-11-2020)	78	492864	4948864	214803	127
EL_1 (25-9-2020)	75	758593	4949126	215117	237
EL_2 (8-10-2020)	72	745121	4945305	212222	226

Table 13 Results QC parameters on the de novo assembled genomes, average per strain

Strain number	Average # contigs	Average Largest contig	Average Total length	Average N50	Average Coverage
20SCA01	100	702539	4990000	238841	167
20SCA02	114	518682	4907016	232538	171
20SCA03	99	566993	4924267	238779	177
20SCA04	104	668308	4991667	238762	166
20SCA05	106	695052	4993815	236536	169
20SCA06	99	672200	4983799	240603	171
20SCA07	94	630567	4883317	241791	166
20SCA08	98	657744	4990408	234701	155
20SCA09	102	662473	4978520	201360	179
20SCA10	89	568203	4961737	185663	183

Participants were asked to report per strain if (yes or no) a clustering match was found with the reference outbreak strain (REF) in the EURL-*Salmonella* PT Typing 2020: 20SCA_REF_R1.fq.gz and 20SCA_REF_R2.fq.gz (monophasic *Salmonella* Typhimurium, ST34, MLVA type 3-14-13-NA-211).

The WGS cluster definition for the PT Typing 2020 was set at maximum six allelic differences from the reference (REF). Based on this cluster definition, WGS-based results were expected to indicate strains 20SCA04, 20SCA05, SCA06 (reference strain) and 20SCA08 (technical duplicate of the reference strain) to be a clustering match with the provided REF outbreak strain (also see Figures 1 and 10).

All but one of the 23 submissions (two participants with both a SNP-based and a cgMLST-based submission) reported the WGS-based cluster analysis results completely as expected (Table 14). Laboratory 32 reported strain 20SCA08 not to be clustering with the reference strain, but remarked that "I would from this analysis without any metadata also suggest strain SCA08 to possibly be part of the cluster due to 9 SNP differences". Notably, the cgMLST-based analysis on all participants' data showed no allelic differences for clustering strain 20SCA08 at all (Annex 11).

Although this was not a specific question in the result form, two participants commented that strains 20SCA02 and 20SCA03 would fall into the definition of a second WGS-based cluster (Figure 1 and Figure 10). Note that this was not the case when using the PFGE-based or MLVA-based cluster definitions (Figure 9 and Table 3).

The technical duplicates SCA06/SCA08 were expected to be reported as (part of) one cluster and this was done in 22 of the 23 submissions (Table 14).

Table 14 Expected cluster analysis results and the cluster analysis results reported by the 21 WGS participants

Lab code	20 SCA01	20 SCA02	20 SCA03	20 SCA04	20 SCA05	20 SCA06	20 SCA07	20 SCA08	20 SCA09	20 SCA10
Expected	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**	**Yes**	**No**	**No**
1	No	No	No	Yes	Yes	Yes	No	Yes	No	No
2-cgMLST	No	No	No	Yes	Yes	Yes	No	Yes	No	No
2-SNP	No	No	No	Yes	Yes	Yes	No	Yes	No	No
3	No	No	No	Yes	Yes	Yes	No	Yes	No	No
6-cgMLST	No	No	No	Yes	Yes	Yes	No	Yes	No	No
6-SNP	No	No	No	Yes	Yes	Yes	No	Yes	No	No
8	No	No	No	Yes	Yes	Yes	No	Yes	No	No
11	No	No	No	Yes	Yes	Yes	No	Yes	No	No
12	No	No	No	Yes	Yes	Yes	No	Yes	No	No
14	No	No	No	Yes	Yes	Yes	No	Yes	No	No
17	No	No	No	Yes	Yes	Yes	No	Yes	No	No
18	No	No	No	Yes	Yes	Yes	No	Yes	No	No
19	No	No	No	Yes	Yes	Yes	No	Yes	No	No
21	No	No	No	Yes	Yes	Yes	No	Yes	No	No
24	No	No	No	Yes	Yes	Yes	No	Yes	No	No
25	No	No	No	Yes	Yes	Yes	No	Yes	No	No
28	No	No	No	Yes	Yes	Yes	No	Yes	No	No
31	No	No	No	Yes	Yes	Yes	No	Yes	No	No
32	No	No	No	Yes	Yes	Yes	No	No	No	No
33	No	No	No	Yes	Yes	Yes	No	Yes	No	No
34	No	No	No	Yes	Yes	Yes	No	Yes	No	No
91	No	No	No	Yes	Yes	Yes	No	Yes	No	No
96	No	No	No	Yes	Yes	Yes	No	Yes	No	No

5 Conclusions

5.1 Serotyping

- Overall results for the 37 evaluated participants are:
  - 99% of the strains were typed correctly for the O-antigens.
  - 98% of the strains were typed correctly for the H-antigens.
  - 97% of the strains were correctly named.
- All 29 NRLs-*Salmonella* in the 27 EU Member States plus the United Kingdom, the five non-EU NRLs, plus the additional three WGS participants met the level of good performance at the first stage of this PT, and there was no need to organise a follow-up study.

5.2 Cluster analysis

- The second pilot on optional cluster analysis was based on the simulation of an outbreak-related request to the NRL-network from the EURL-*Salmonella* (EFSA/ECDC), including a description of the cluster definition.
- Selection of suitable PT strains was improved by including extended pre-testing of the strains by the EURL-*Salmonella*, based on MLVA and WGS.
- A total of 21 participants performed the second cluster analysis pilot, with two participants using PFGE analysis, six using MLVA analysis and 21 participants using WGS analysis.
- The two PFGE participants reported their PFGE-based cluster analysis results in complete agreement.
- Five out of the six participants reported the MLVA-based cluster analysis results completely as expected.
- All but one of the 23 submissions (two participants with both a SNP-based and a cgMLST-based submission) reported the WGS-based cluster analysis results completely as expected.
- The technical duplicate strains 20SCA06/20SCA08 were expected to be reported as (part of) one cluster. This was the case in 2/2 PFGE submissions, in 6/6 MLVA submissions, and in 22/23 WGS submissions.

List of abbreviations

BN	BioNumerics
BPW	Buffered Peptone Water
cgMLST	core genome Multilocus Sequence Typing
DG-SANTE	Directorate General for Health and Food Safety
EC	European Commision
ECDC	European Centre for Disease prevention and Control
EFSA	European Food Safety Authority
EFTA	European Free Trade Association
EL	EURL-*Salmonella* Laboratory
EU	European Union
EURL-*Salmonella*	European Union Reference Laboratory for *Salmonella*
ftp	file transfer protocol
ISO	International Organization for Standardization
MLVA	Multiple-Locus Variable number of tandem repeat Analysis
MST	Minimum Spanning Tree
n.a.	not applicable
NRL-*Salmonella*	National Reference Laboratory for *Salmonella*
PCR	Polymerase Chain Reaction
PFGE	Pulsed Field Gel Electrophoresis
PT	Proficiency Test
QC	Quality Control
REF	Reference
RIVM	National Institute for Public Health and the Environment (Bilthoven, The Netherlands)
SNP	Single Nucleotide Polymorphism
SSI	Statens Serum Institut (Copenhagen, Denmark)
ST	Sequence Type
TIFF	Tagged Image File Format
wgMLST	whole genome Multilocus Sequence Typing
WGS	Whole Genome Sequencing

References

EC, 2017. Regulation (EU) 2017/625 of the European Parliament and of the Council of 15 March 2017 on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products. Official Journal of the European Union L95: 7 April 2017. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0625&rid=3 (accessed 28-9-2021).

ECDC (European Centre for Disease Prevention and Control), 2011. ECDC Technical Report, ECDC 2011-06, version 1.2. Laboratory standard operating procedure for MLVA of *Salmonella enterica* serotype Typhimurium. https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/1109_SOP_Salmonella_Typhimurium_MLVA.pdf (accessed 28-9-2021).

Grimont, P.A.D. and F.-X. Weill, 2007. Antigenic formulae of the *Salmonella* serovars, 9th ed. WHO Collaborating Centre for Reference and Research on *Salmonella*. Institute Pasteur, Paris, France. https://www.pasteur.fr/sites/default/files/veng_0.pdf (accessed 28-9-2021).

EN ISO 6579-1. Microbiology of the food chain – Horizontal method for the detection, enumeration and serotyping of *Salmonella* – Part 1: Detection of *Salmonella* spp. International Organization for Standardization, Geneva, Switzerland.

EN ISO 15189. Medical laboratories - Requirements for quality and competence. International Organization for Standardization, Geneva, Switzerland.

EN ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories. International Organization for Standardization, Geneva, Switzerland.

Hendriksen, R.S., et al., 2009. WHO Global Salm-Surv External Quality Assurance System for Serotyping of *Salmonella* Isolates from 2000 to 2007. J Clin Microbiol 2009(47): 2729-2736. doi:10.1128/JCM.02437-08

Jacobs, W., S. Kuiling, K. van der Zwaluw, 2014. Molecular typing of *Salmonella* strains isolated from food, feed and animals: state of play and standard operating procedures for pulsed field gel electrophoresis (PFGE) and Multiple-Locus Variable number tandem repeat Analysis (MLVA) typing, profiles interpretation and curation. EFSA supporting publication 2014:EN-703, 74 pp. https://www.efsa.europa.eu/en/supporting/pub/en-703 (accessed 28-9-2021).

Mikoleit, M., M.S. Van Duyne, J. Halpin, B. McGlinchey, and P.I. Fields, 2012. Variable Expression of O:61 in *Salmonella* Group C2. J Clin Microbiol 2012(50): 4098-4099. doi:10.1128/JCM.01676-12

Mooijman, K.A., 2007. The twelfth CRL-*Salmonella* Workshop; 7 and 8 May 2007, Bilthoven, the Netherlands. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report no.: 330604006. http://www.eurlsalmonella.eu/Publications/Workshop_Reports (accessed 28-9-2021).

Mooijman, K.A., 2020. The 25th EURL-*Salmonella* workshop; 18-19 September 2020, Online. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report no.: 2020-0202. http://www.eurlsalmonella.eu/Publications/Workshop_Reports (accessed 28-9-2021).

Tennant, S.M., S. Diallo, H. Levy, S. Livio, S.O. Sow, M. Tapia, P.I. Fields, M. Mikoleit, B. Tamboura, K.L. Kotloff, J.P. Nataro, J.E. Galen and M.M. Levine, 2010. Identification by PCR of non-typhoidal *Salmonella enterica* serovars associated with invasive infections among febrile patients in Mali. PLoS. Negl. Trop. Dis 4(3): 621.

Wick, R.R., L.M. Judd, C.L. Gorrie, and K.E. Holt (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13(6): e1005595. Available online: https://doi.org/10.1371/journal.pcbi.1005595 (accessed 28-9-2021).

Annex 1 Example of an individual laboratory evaluation report on serotyping results

Results
EURL-*Salmonella* PT Serotyping 2020

Number of penalty points: 0 **Evaluation: Good Performance**

	Reference Results				**Results NRL labcode:**			**1**
Strain	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar
S1	13,23	i	e,n,z15	Jukestown	13,23	i	e,n,z15	Jukestown
S2	1,6,14,25	z4,z23	[e,n,z15]	Bousso	6,14,25	z4,z23	-	Bousso
S3	6,8	z10	e,n,x	Hadar	6,8	z10	e,n,x	Hadar
S4	1,4,12,27	z29	-	Brancaster	4,12,27	z29	-	Brancaster
S5	8	d	1,2	Virginia	8	d	1,2	Virginia
S6	9,12	d	z6	Zega	9,12	d	z6	Zega
S7	1,13,23	g,m,[s],[t]	-	Agbeni	13,23	g,m	-	Agbeni
S8[a)]	1,4,[5],12	i	-	1,4,[5],12:i:-	4,12	i	-	4,12:i:-
S9	30	k	e,n,[x],z15	Odozi	30	k	e,n,z15	Odozi
S10	1,4,12,[27]	l,[z13],z28	1,5	Tyresoe	4,12	l,z28	1,5	Tyresoe
S11	11	l,v	1,2	Stendal	11	l,v	1,2	Stendal
S12	4,12,[27]	a	1,5	Hessarek	4,12	a	1,5	Hessarek
S13	1,4,[5],12	i	1,2	Typhimurium	4,5,12	i	1,2	Typhimurium
S14	6,7	e,h	1,2	Larochelle	6,7	e,h	1,2	Larochelle
S15	6,7,14	r	1,2	Virchow	6,7	r	1,2	Virchow
S16	1,9,12	g,m	-	Enteritidis	9,12	g,m	-	Enteritidis
S17	3,10	b	e,n,x	Benfica	3,10	b	e,n,x	Benfica
S18	6,7,14	r	1,5	Infantis	6,7	r	1,5	Infantis
S19	4,12,[27]	b	1,6	Canada	4,12	b	1,6	Canada
S20	8,20	z38	-	Apeyeme	8,20	z38	-	Apeyeme
S21[b)]	50	r	1,5,(7)	50:r:1,5 (IIIb)	50	r	1,5,7	50:r:1,5,7

a) Typhimurium, monophasic variant as determined by PCR.
b) *Salmonella enterica* subspecies *diarizonae*

Lab 1: S19: inconsistent agglutination with various antisera

Results
EURL-*Salmonella* PT Serotyping 2020

For back-ground information, reference results are given completely according to the White-Kauffmann-le Minor scheme (2007).
Participants were asked to report only those results, on which the identification of serovar names was based.

Colour coding:

	remark (e.g. spelling error, or deviations in the results of optional strain S21)
	not typable (e.g. antisera not available, rough strain)
	partly correct; the naming: no penalty points
	incorrect; in the naming: 1 penalty point
	incorrect; in the naming: 4 penalty points

As decided at the 25th EURL-*Salmonella* Workshop (Online, 2020), Strain S-21 was an additional strain to the study.
Testing of this strain was optional and results were not included in the evaluation (remarks in blue or grey only).
The evaluation of the serotyping results was performed as indicated in Table 1 of the Protocol as sent to the participants.
In addition to that, Good Performance was evaluated on the basis of penalty points as indicated below.

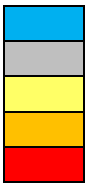
4 penalty points: Incorrect typing of *S*. Enteritidis, *S*. Typhimurium (including monophasic variant), *S*. Hadar, *S*. Infantis or *S*. Virchow or assigning the name of one of these 5 serovars to another serovar.
1 penalty point: Incorrect typing of all other *Salmonella* serovars.
(no penalty points are given in case a strain was non-typable due to auto-agglutination)

Good Performance is defined as < 4 penalty points.

Annex 2 Serotyping results per strain and per laboratory

Lab: REF	S1 Jukestown	S2 Bousso	S3[a] Hadar	S4 Brancaster	S5[b] Virginia	S6 Zega	S7 Agbeni	S8 1,4,[5],12:i:-	S9 Odozi	S10 Tyresoe
1	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
2	Jukestown	Bousso	Istanbul	Brancaster	Muenchen	Zega	Agbeni	1,4,12:i:-	Odozi	Tyresoe
3	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
4	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	1,4,12;i;-	Odozi	Tyresoe
5	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
6	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
7	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
8	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,5,12:i:-	Odozi	Tyresoe
9	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	monophasic Typhimurium	Odozi	Tyresoe
10	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	Typhimurium Monophasic	Odozi	Tyresoe
11	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	monophasic Typhimurium	Odozi	Tyresoe
12	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
13	Jukestown	Bousso	Hadar	Brancaster	Muenchen	Zega	Agbeni	1,4,5,12:i:-	Odozi	Tyresoe
14	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,5,12:i:-	Odozi	Tyresoe
15	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
16	Jukestown	Bousso	Hadar	4,12:HME:-	Virginia	Zega	Agbeni	Typhimurium	OMC:k:e,n,z15	Azteca
17	I:13,23:i:-	Bousso	Hadar	Brancaster	Virginia	Zega	-:gm:-	I:4:i:- (monophasic TM)	Odozi	Tyresoe
18	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
19	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	Typhimurium monophasic variant	Odozi	Tyresoe
20	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,5,12:i:-	?	Tyresoe
21	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	1,4,12; i; -	Odozi	Tyresoe
22	jukestown	bousso	hadar	brancaster	virginia	zega	agbeni	Monophasic Salmonella typhimurium	odozi	tyresoe
23	Jukestown	Bousso	Hadar	Brancaster	Muenchen	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
24	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4 : i : -	Odozi	Tyresoe
25	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	1,4,12:i:- (mST)	Odozi	Tyresoe
26	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
27	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4, 5, 12: i: -	Odozi	Tyresoe
28	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,5,12:i:-	Odozi	Tyresoe
29	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,(5),12:i:-	Odozi	Tyresoe
30	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	Monophasic Typhimurium	Odozi	Tyresoe
31	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4:i:-	Obdozi	Tyresoe
32	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	Typhimurium, monophasic 4,12 : i -	Odozi	Tyresoe
33	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4:i:-	Odozi	Tyresoe
34	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,12:i:-	Odozi	Tyresoe
73	Jukestown	Bousso	Hadar	Brancaster	Virginia	Zega	Agbeni	4,[5],12:i:-	Odozi	Tyresoe
91	Jukestown	Bousso	Hadar	Brancaster	Muenchen	Zega	Agbeni	Typhimurium - monophasic	Odozi	Tyresoe
96	Juketown	Bousso	Hadar	Brancaster	Muenchen	Zega	Agbeni	I 4,[5],12:i:-	Angoda	Tyresoe
X	0	0	0	0	2	0	0	1	1	1

S11 Stendal	S12 Hessarek	S13 Typhimurium	S14 Larochelle	S15 Virchow	S16 Enteritidis	S17 Benfica	S18 Infantis	S19 Canada	S20 Apeyeme	Lab: REF
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	1
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	2
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	3
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	4
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	5
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	6
Stendal	Lagos	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	7
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	8
Stendal	-	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Uppsala	Apeyeme	9
Stendal	Fulica / Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	10
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	11
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	12
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	13
Stendal	Paratyphi A	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	14
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	15
Stendal	4,12:a:1,5	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	8,20:HME:-	16
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	17
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	18
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	19
Stendal	?	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Uppsala	?	20
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	21
tours	hessarek	typhimurium	larochelle	virchow	enteritidis	benfica	infantis	canada	apeyeme	22
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	23
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	24
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benefica	Infantis	Canada	Apeyeme	25
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	26
Stendal	Fulica	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	27
Stendal	Fulica, Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apayeme	28
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	29
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	30
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	31
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	32
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	33
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	34
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	73
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	91
Stendal	Hessarek	Typhimurium	Larochelle	Virchow	Enteritidis	Benfica	Infantis	Canada	Apeyeme	96
1	2	0	0	0	0	0	0	2	0	**X**



remark (e.g., spelling error)
not typable (e.g., antisera not available, rough strain)
partly correct, in the naming: no penalty points
incorrect; in the naming: 1 penalty point
incorrect; in the naming: 4 penalty points

X = number of deviating laboratories (by penalty points) per strain.

Results for Strain S21 are given in Annex 4.

a) Remark on Strain S3: According to the protocol of this PT, an 8:z_{10}:e,n,z typed strain should have been reported as “Istanbul” (Laboratory code 2). An “Hadar” named strain would have been expected to show 6,8 for the O-antigen result, therefore an 8 result for the O-antigen is (for this PT) considered as “partly correct” (Laboratory codes 29, 91, and 96).
b) Remark on Strain S5: According to the protocol of this PT, an 8:d:1,2 typed strain should have been reported as "Virginia" and a 6,8:d:1,2 typed strain should have been reported as "Muenchen". Therefore, the 8:d:1,2 results named Muenchen are (for this PT) considered as "incorrect" (Laboratory codes 91 and 96).

Annex 3 Details per strain that caused problems or inconsistencies in serotyping

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-1	13,23	i	e,n,z15	Jukestown	REF
S-1	13,23	i	-	I:13,23:i:-	17
S-1	13	i	e,n,z15	Juketown	96
S-3[a)]	6,8	z10	e,n,x	Hadar	REF
S-3	8	z10	e,n,x	Istanbul	2
S-3	8	z10	e,n,x	Hadar	29
S-3	8 (O6 confirmation: +)	z10	e,n,x	Hadar	73
S-3	8	z10	e,n,x	Hadar	91
S-3	8	z10	e,n,x	Hadar	96
S-4	1,4,12,27	z29	-	Brancaster	REF
S-4	4,12	HME	-	4,12:HME:-	16
S-5[b)]	8	d	1,2	Virginia	REF
S-5	6,8	d	1,2	Muenchen	2
S-5	6,8	d	1,2	Muenchen	13
S-5	6,8	d	1,2	Muenchen	23
S-5	8	d	1,2	Virginia	29
S-5	8 (O6 confirmation: -)	d	1,2	Virginia	73
S-5	8	d	1,2	Muenchen	91
S-5	8	d	1,2	Muenchen	96
S-7	1,13,23	g,m,[s],[t]	-	Agbeni	REF
S-7	-	g,m	-	-:gm:-	17
S-8	1,4,[5],12	i	-	1,4,[5],12:i:-	REF
S-8	4,12	i	-	1,4,12;i;-	4
S-8	4,12	i	-	Typhimurium	16
S-9	30	k	e,n,[x],z15	Odozi	REF
S-9	OMC	k	e,n,z15	OMC:k:e,n,z15	16
S-9	OMC	k	e,n,z15	?	20
S-9	30	k	e,n, z15	Obdozi	31
S-9	30	k	e,n,z15	Angoda	96
S-10	1,4,12,[27]	l,[z13],z28	1,5	Tyresoe	REF
S-10	4,5,12,27	l,v	1,5	Azteca	16
S-11	11	l,v	1,2	Stendal	REF
S-11	11	l,z13	1,2	tours	22
S-12	4,12,[27]	a	1,5	Hessarek	REF
S-12	4,12	i	1,5	Lagos	7
S-12	-	-	-	-	9
S-12	4,12	a	1,5	Fulica / Hessarek	10

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-1	13,23	i	e,n,z15	Jukestown	REF
S-12	2,12	a	5	Paratyphi A	14
S-12	4,12	a	1,5	4,12:a:1,5	16
S-12	?	?	?	?	20
S-12	4, 12	a	-	Fulica	27
S-12	4,12	a	5	Fulica, Hessarek	28
S-17	3,10	b	e,n,x	Benfica	REF
S-17	3,10	b	e,n,x	Benefica	25
S-19	4,12,[27]	b	1,6	Canada	REF
S-19	4	b	1,7	Uppsala	9
S-19	4,12	b	1,7	Uppsala	20
S-20	8,20	z38	-	Apeyeme	REF
S-20	8,20	HME	-	8,20:HME:-	16
S-20	8,20	HMD	-	?	20

Reference strain
remark (e.g. spelling error)
not typable (e.g. antisera not available, rough strain)
partly correct; in the naming: no penalty points
incorrect; in the naming: 1 penalty point
incorrect; in the naming: 4 penalty points

a) Remark on Strain S3: According to the protocol of this PT, an 8:z_{10}:e,n,z typed strain should have been reported as "Istanbul" (Laboratory code 2). An "Hadar" named strain would have been expected to show 6,8 for the O-antigen result, therefore an 8 result for the O-antigen is (for this PT) considered as "partly correct" (Laboratory codes 29, 91, and 96).
   Retrospectively, Lab 29 reported: "Strains S3: We checked for the presence of both O:6 and O:8 using agglutination, and we found both O-antigens. Therefore the O-antigens should have been 6,8 and not just 8 as I have reported. This also corresponds with the serovar name of Hadar."
b) Remark on Strain S5: According to the protocol of this PT, an 8:d:1,2 typed strain should have been reported as "Virginia" and a 6,8:d:1,2 typed strain should have been reported as "Muenchen". Therefore, the 8:d:1,2 results named Muenchen are (for this PT) considered as "incorrect" (Laboratory codes 91 and 96).

Annex 4 Details of serotyping results for strain S21

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-21	**50**	**r**	**1,5,(7)**	**IIIb 50:r:1,5**	**REF**
S-21	50	r	1,5,7	50:r:1,5,7	1
S-21	50	r	1,5,7	IIIb 50:r:1,5,7	2
S-21	50	r	1,5	SIIIb 50:r:1,5	3
S-21	50	r	1,5	Salmonella enterica subspecies diarizonae 50:r:1,5	4
S-21	50	r	1,5,7	50:r:1,5,7	5
S-21	50	-	1,5	50:-:1,5	6
S-21					7
S-21	50	r	5	Salmonella enterica subsp. diarizonae 50:r:1,5(7)	8
S-21	50	r	1,5	IIIb (diarizonae)	9
S-21	61	r	1,5,7		10
S-21	OME	r	1,5,7	OME : r : 1,5,7 (IIIb)	11
S-21	50	r	1,5,7	S. IIIb 50:r:1,5,7	12
S-21	50	r	1,5,7	50:r:1,5,7	13
S-21	50	r	5	IIIb 50:r:1,5,(7)	14
S-21	50	r	1,5	50:r:1,5	15
S-21	-	-	-	-:-:-	16
S-21	50	r	1,5	IIIb:50:r:1,5	17
S-21	50	r	1,5	50:r:1,5	18
S-21	50	r	1,5	IIIb 50 : r : 1,5,(7)	19
S-21					20
S-21	50	r	1,5,7	50; r; 1,5,7	21
S-21	11	r	1,5	senegal	22
S-21	50	r	1,5	(IIIb) 50:r:1,5	23
S-21	?	r	5	OME + : r : 5	24
S-21	50	r	1,5,7	50:r:1,5,7	25
S-21	50	r	1,5	IIIb (diarizonae) - 50:r:1,5	26
S-21					27
S-21	61	r	5	IIIa arizonae	28
S-21	50	r	1,5	Subspecies IIIb	29
S-21	61	r	1,5,7	Diarizonae	30
S-21	50	?	?	Subspec III**	31
S-21	50	r	1,5,7	Salmonella enterica subsp. diarizonae serovar 50 : r ; 1,5,7	32
S-21	50	r	1,5,7	IIIb 50:r:1,5,7	33
S-21	50	r	1,5,7	sg IIIb 50:r:1,5,7	34
S-21	50	r	1,5,7	IIIb 50:r:1,5,(7)	73
S-21	50	r	1,5,7	IIIb 50:r:1,5,(7)	91
S-21	IIIa 50	r	1,5,7	IIIa 50:r:1,5,7	96

Annex 5 Minimum Spanning Tree of EURL-*Salmonella* pre-tested strains in the cluster analysis

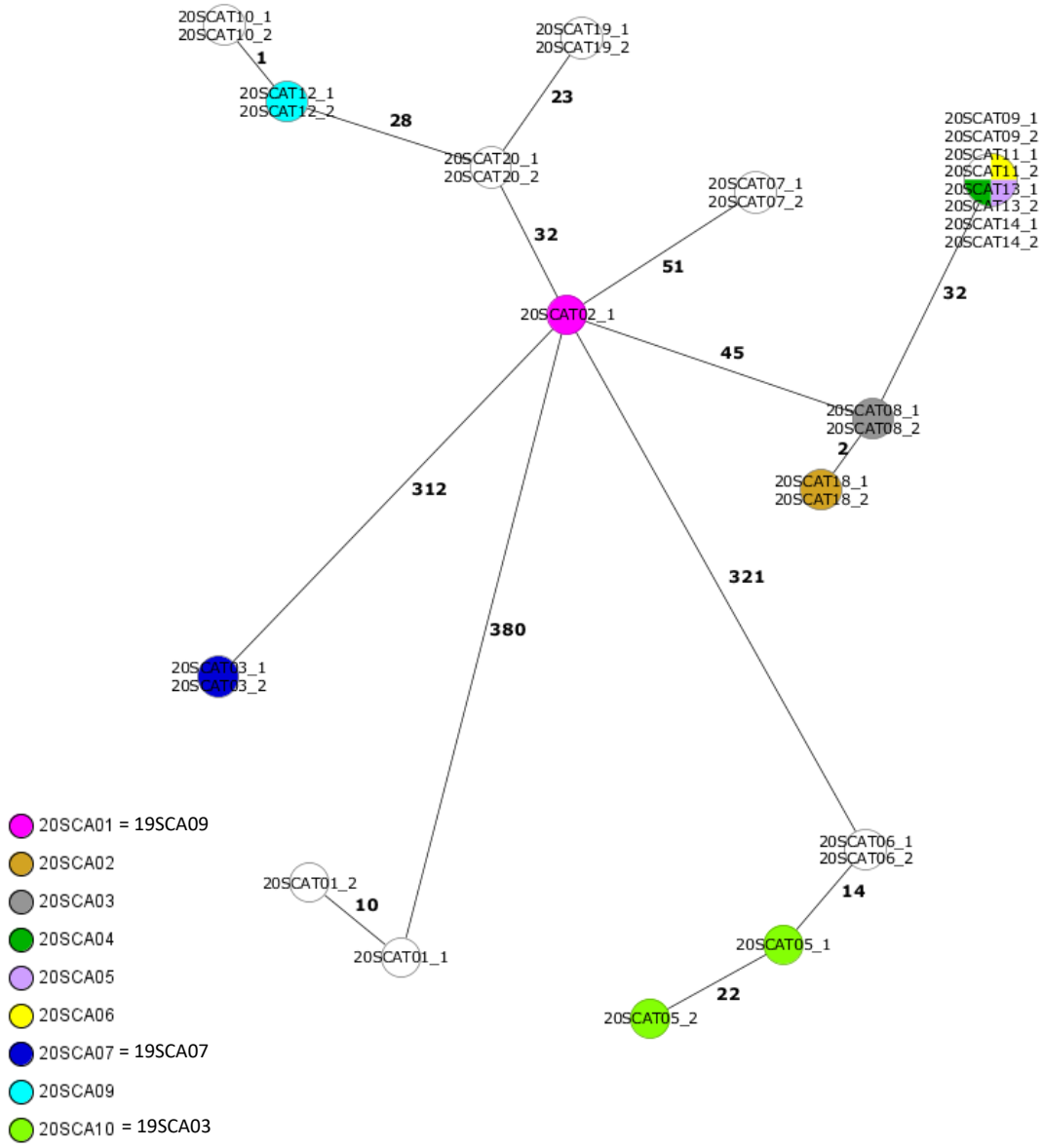


*Figure A5 MST of 16 20SCA**T** test strains, before (_1) and after (_2) ten times sub-culturing.*

The 20SCAT02_2 data are not available.
20SCAT01 = 19SCA10, 20SCAT06 = 19SCA02.
The legend shows which (coloured) strains were selected and renamed for the PT2020 cluster analysis (20SCA01-20SCA10).

Annex 6 Example of an individual laboratory evaluation report on cluster analysis results

Evaluation
EURL-*Salmonella* PT Cluster Analysis 2020

Laboratory code: 33
Evaluation (per methodology) of the participants' cluster analysis results was done by comparing the participants' results to the expected results in an outbreak situation setting, as pre-defined by the EURL-*Salmonella* (Protocol PT Typing 2020).

As a minimum, it was expected to have any technical duplicate strains reported as (part) of one cluster.
No performance criteria were set for this second pilot PT on cluster analysis.

In general, deviations (of any kind) from the expected (REF) results are indicated in blue:

Background details and overall results can be found in the interim summary report EURL-*Salmonella* PT Cluster Analysis 2020 (www.eurlsalmonella.eu)

Did you serotype the strains: Yes
Methodology used: Classical serology; xMAP Salmonella Serotyping Assay

Strain:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**
Expected results:	4,[5],12:i:-	4,5,12:i:-	4,5,12:i:-	4,12:i:-	4,5,12:i:-
Reported results:	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-

Strain:	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
Expected results:	4,5,12:i:-	Typhimurium	4,5,12:i:-	4,12:i:-	Typhimurium
Reported results:	4,5,12:i:-	Typhimurium	4,5,12:i:-	4,5,12:i:-	Typhimurium

Submission of PFGE results: Yes

Number of reported clusters detected by PFGE data analysis:	1
Expected result*:	1

	PFGE Cluster 1
Expected result*:	SCA03-SCA04-SCA05-SCA06-SCA08
Reported IDs for the strains per cluster:	SCA03-SCA04-SCA05-SCA06-SCA08

*based on the results by the 2 PFGE participants (Labs 17 and 33), Dice coefficient, both tolerance and optimization at 1,5% (EFSA recommendation). Also see the interim summary report for details.

PFGE comment by Lab 33:
The strains included in the cluster 1 (SCA03-SCA04-SCA05-SCA06-SCA08) were obtained by using optimization and tolerance indices equal to 1.5, according to the protocol defined by EFSA (2014). Anyway, if we adjust both the indices to 1.0, strain SCA03 is not included in cluster 1 anymore. Thus, according to the modified parameters the final cluster 1 includes the strains SCA04-SCA05-SCA06-SCA08.

PFGE-based cluster identification as expected:	Yes
Technical duplicates 20SCA06 and 20SCA08 reported within one cluster:	Yes

Submission of MLVA results: Yes

Strain:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**
Expected results:	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
Reported results:	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211

Strain:	**20SCA06 (REF)**	**20SCA07**	**20SCA08 (ref)**	**20SCA09**	**20SCA10**
Expected results:	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
Reported results:	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311

MLVA-based cluster identification in the PT Typing 2020 setting included:

Report per strain if [yes or no] a clustering match was found with the **Reference outbreak strain (REF)** in the EURL-*Salmonella* PT Typing 2020:

monophasic *Salmonella* Typhimurium ST34, MLVA type 3-14-13-NA-211

The cluster definition for MLVA is set at no loci with a different number of repeats.

Strain:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
Expected results:	No	No	No	Yes	Yes	Yes	No	Yes	No	No
Reported results:	No	No	No	Yes	Yes	Yes	No	Yes	No	No

MLVA-based cluster identification as expected: Yes
Technical duplicates 20SCA06 and 20SCA08 reported within one cluster: Yes

Submission of WGS results: Yes
WGS platform used: Illumina Mi-Seq
Analysis used for WGS data: cgMLST-based
Tool used for analysis: chewBBaca
Method used or phylogenetic analysis: Minimum Spanning Tree (MST)

	20SCA_REF_R1.fq.gz	20SCA_REF_R2.fq.gz
Expected md5 checksum:	257eece96dfe3169c2e1f00e797c1dca	29adc5a5b60e3e96ca69fdf31cfc1022
Reported md5 checksum:		

WGS-based cluster identification in the PT Typing 2020 setting included:

Report per strain if [yes or no] a clustering match was found with the **Reference outbreak strain (REF)** in the EURL-*Salmonella* PT Typing 2020:
20SCA_REF_R1.fq.gz and **20SCA_REF_R2.fq.gz**
(monophasic *Salmonella* Typhimurium ST34, MLVA type 3-14-13-NA-211)
The cluster definition for WGS is set at maximum 6 allele differences from the reference sequence.

Strain:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
Expected results:	No	No	No	Yes	Yes	Yes	No	Yes	No	No
Reported results:	No	No	No	Yes	Yes	Yes	No	Yes	No	No

WGS-based cluster identification as expected: Yes
Technical duplicates 20SCA06 and 20SCA08 reported within one cluster: Yes

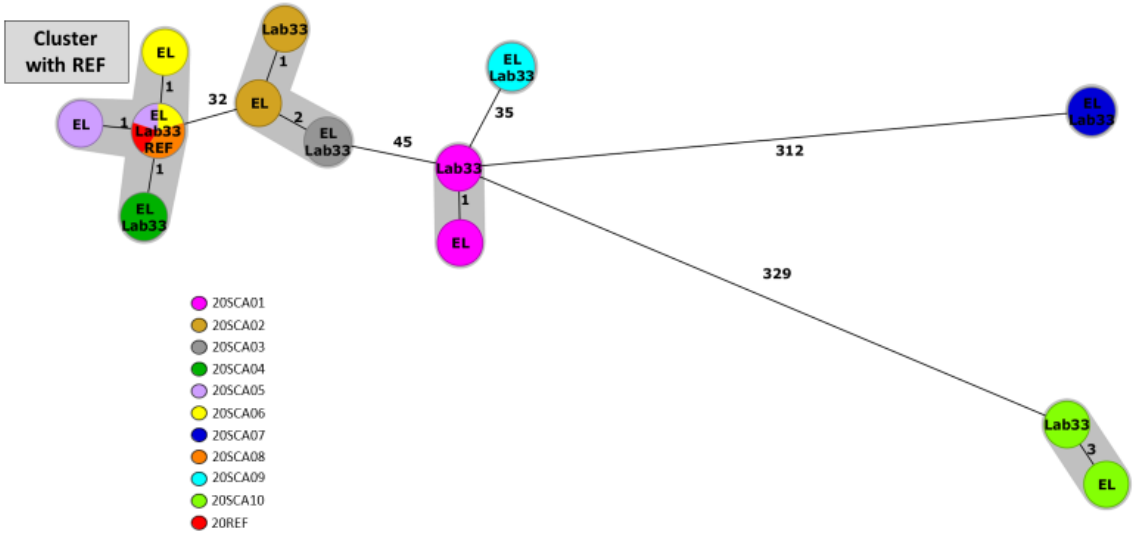


*Figure A6 Minimum Spanning Tree of the participants' results and the EURL-*Salmonella *(EL) results, analysed in Ridom SeqSphere+, (assembly_pipeline: https://github.com/Papos92), S. enterica cgMLST (3002), pairwise ignoring missing values*

Annex 7 Serotyping results cluster analysis part

Lab code	Serotyping method used	20SCA01	20SCA02	20SCA03	20SCA04	20SCA05	20SCA06 (REF)	20SCA07	20SCA08 (REF)	20SCA09	20SCA10
REF	Classical serology/PCR	4,[5],12:i:-	4,5,12:i:-	4,5,12:i:-	4,12:i:-	4,5,12:i:-	4,5,12:i:-	Typhimurium	4,5,12:i:-	4,12:i:-	Typhimurium
6	BioNumerics 8.0 Salmonella plugin	Monophasic ST	Monophasic ST	Monophasic ST	Monophasic ST	Monophasic ST	Monophasic ST	S. Typhimurium	Monophasic ST	Monophasic ST	S. Typhimurium
8	Classical serology	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. Typhimurium	S. 4,5,12 : i : -	S. 4,5,12 : i : -	S. 4,5,12 : i : -
11	Classical serology Tennant et al.,2010	monophasic Typhimurium	monophasic Typhimurium	monophasic Typhimurium	monophasic Typhimurium	monophasic Typhimurium	monophasic Typhimurium	Typhimurium	monophasic Typhimurium	monophasic Typhimurium	Typhimurium
12	Classical serology	S. Typhimurium O5-, monophasic	S. Typhimurium, monophasic	S. Typhimurium, monophasic	S. Typhimurium, monophasic	S. Typhimurium, monophasic	S. Typhimurium, monophasic	S. Typhimurium	S. Typhimurium, monophasic	S. Typhimurium O5-, monophasic	S. Typhimurium
14	WGS (SeqSero2 v1.1.0)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	Typhimurium	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	1,4,[5],12:i:- (monophasic Salmonella Typhimurium)	Typhimurium
19	http://www.denglab.info/SeqSero/	potential monophasic variant of Typhimurium(O5-)	potential monophasic variant of Typhimurium(O5-)	potential monophasic variant of Typhimurium	potential monophasic variant of Typhimurium	potential monophasic variant of Typhimurium	potential monophasic variant of Typhimurium	Typhimurium	potential monophasic variant of Typhimurium	potential monophasic variant of Typhimurium	Typhimurium
24	Classical serology	4:i:-, monophasic STM by PCR	4,5:i:-, monophasic STM by PCR	4,5:i:-, monophasic STM by PCR	4,5:i:-, monophasic STM by PCR	4,5:i:-, monophasic STM by PCR	4,5:i:-, monophasic STM by PCR	Typhimurium	4,5:i:-, monophasic STM by PCR	4:i:-, monophasic STM by PCR	Typhimurium
25	sistr, seqsero2	Typhimurium	1,4,[5],12:i:-	1,4,[5],12:i:-	1,4,[5],12:i:-	1,4,[5],12:i:-	1,4,[5],12:i:-	Typhimurium	1,4,[5],12:i:-	1,4,[5],12:i:-	Typhimurium
28	https://cge.cbs.dtu.dk/services/SeqSero/	monophasic 4,12:i:-	monophasic 4,12:i:-	monophasic 4,12:i:-	monophasic 4,12:i:-	monophasic 4,12:i:-	monophasic 4,12:i:-	S. Typhimurium	monophasic 4,12:i:-	monophasic 4,12:i:-	S.Typhimurium
31	SeqSero	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:1,2 Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:- monophasic variant of Typhimurium	4:i:1,2 Typhimurium
32	WGS, SeqSero vers. 1.2 (CGE web tool)	4:i:-	4:i:-	4:i:-	4:i:-	4:i:-	4:i:-	4:-:1,2	4:i:-	4:i:-	4:-:1,2
33	Classical serology xMAP Salmonella Serotyping Assay	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	4,5,12:i:-	Typhimurium	4,5,12:i:-	4,5,12:i:-	Typhimurium
34	Most, SeqSero and Sistr	Typhimurium	Monophasic Typhimurium	Monophasic Typhimurium	Monophasic Typhimurium	Typhimurium	Monophasic Typhimurium	Typhimurium	Monophasic Typhimurium	Monophasic Typhimurium	Typhimurium
91	WGS - ST & SeqSero	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium	Salmonella Typhimurium - monophasic	Salmonella Typhimurium - monophasic	Salmonella Typhimurium
96	seqsero2 and SISTR	Typhimurium/I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	Typhimurium	I 1,4,[5],12:i:-	I 1,4,[5],12:i:-	Typhimurium

Annex 8 MLVA results cluster analysis part

Lab code	20SCA01	20SCA02	20SCA03	20SCA04	20SCA05
Expected	**3-13-9-NA-211**	**3-14-9-NA-211**	**3-15-9-NA-211**	**3-14-13-NA-211**	**3-14-13-NA-211**
8	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
11-a*	3-13-9-NA-211	3-16-7-17-311	3-11-8-NA-211	3-14-13-NA-211	5-9-14-9-211
11-b*	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
17	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
28	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
31	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211
33	3-13-9-NA-211	3-14-9-NA-211	3-15-9-NA-211	3-14-13-NA-211	3-14-13-NA-211

Lab code	20SCA06 (REF1)	20SCA07	20SCA08 (REF2)	20SCA09	20SCA10
Expected	**3-14-13-NA-211**	**5-9-14-9-211**	**3-14-13-NA-211**	**3-11-8-NA-211**	**3-16-7-17-311**
8	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
11-a*	3-14-13-NA-211	3-14-13-NA-211	3-14-13-NA-211	3-15-9-NA-211	3-14-9-NA-211
11-b*	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
17	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
28	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
31	3-14-13-NA-211	3-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311
33	3-14-13-NA-211	5-9-14-9-211	3-14-13-NA-211	3-11-8-NA-211	3-16-7-17-311

Loci reported in the order: STTR9, STTR5, STTR6, STTR10, STTR3

In blue: Deviation from the expected result.

*Laboratory 11 investigated the potential swap of strains and concluded that they made a mistake in the identification of the samples: initially, samples were tested/reported in the order 10 – 2 (11-a), instead of in the order 2 – 10 (11-b)

Annex 9 WGS results cluster analysis part, methods used by the participants

Lab code	DNA extraction, library preparation and sequencing	WGS platform	Data analysis	Tool used for analysis	Method used for cluster analysis
6-cgMLST	In-house	Illumina MiSeq	cgMLST-based	BioNumerics	Minimum Spanning Tree (MST)
14	Combined[a)]	Illumina NovaSeq	cgMLST-based	BioNumerics	Minimum Spanning Tree (MST)
33	In-house	Illumina MiSeq	cgMLST-based	chewBBaca	Minimum Spanning Tree (MST)
25	In-house	Illumina MiSeq	cgMLST-based	chewBBACA, https://github.com/B-UMMI/chewBBACA	Minimum Spanning Tree (MST)
EL	Outsourced	Illumina NovaSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
1	In-house	MiniSeq Illumina	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
2-cgMLST	Combined[a)]	Illumina NovaSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
3	In-house	Illumina NextSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
8	In-house	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
21	In-house	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
24	In-house	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
96	Outsourced	Illumina NextSeq	cgMLST-based	https://chewbbaca.online/species/4; https://github.com/B-UMMI/chewBBACA	MSTtree V2 GrapeTree https://github.com/achtman-lab/GrapeTree
19	Combined[a)]	NovaSeq6000	cgMLST-based	chewbbaca, used Salmonella.cgMLSTv2 from Enterobase	Neighbor joining (NJ)
17	In-house	Illumina MiSeq	cgMLST-based	in-house Galaxy	Neighbor joining (NJ)
11	Outsourced	Illumina MiSeq	cgMLST-based	in house automated CHEWBBACA based pipeline	single linkage hierarchical clustering
12	In-house	Illumina NextSeq	cgMLST-based	inhouse automated CHEWBACCA based Pipeline	single linkage hierarchical clustering
32	In-house	Illumina MiSeq	SNP-based - A[b)]	In house pipeline[c)]	Maximum likelihood (ML)
18	Outsourced	Illumina MiSeq	SNP-based - A[b)]		Neighbor joining (NJ)
28	In-house	Illumina NextSeq	SNP-based - R[b)]	CSI Phylogeny 1.4; https://cge.cbs.dtu.dk/services/CSIPhylogeny/	Maximum likelihood (ML)
2-SNP	Combined[a)]	Illumina NovaSeq	SNP-based - R[b)]	in-house : iVARCall2	Maximum likelihood (ML)
34	In-house	Illumina MiSeq	SNP-based - R[b)]	Snippy, Gubbins, Raxml, iTol	Maximum likelihood (ML)
6-SNP	In-house	Illumina MiSeq	SNP-based - R[b)]	BioNumerics	Minimum Spanning Tree (MST)
31	In-house	Illumina MiSeq	SNP-based - R[b)]	In-house pipeline	Minimum Spanning Tree (MST)
91	In-house	Illumina HiSeq	SNP-based - R[b)]	SNapper DB	Variant Call Format

[a)] Combined: DNA extraction in-house, library preparation and sequencing outsourced
[b)] A: assembly-based, R: reference-based
[c)] In house pipeline based on parSNP, Gubbins, creating a ML tree in IQTree, creating a SNP distance matrix with snp-dists (https://github.com/NorwegianVeterinaryInstitute/ALPPACA/wiki/Pipeline-and-program-descriptions)

Data sorted by 'Data analysis', 'Method used for cluster analysis', and 'Tool used for analysis'.

Annex 10 WGS cluster analysis part, QC criteria as listed by the participants

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
1	Allele calling	cgMLST CT7 Enterobase & cgMLST Statistics	cgMLST alleles found and called > 95%
1	Avg. coverage (assembled)	Assembly Statistics in SeqSphere	50x but if it's less, the % of good targets should be >95%
1	Contamination check	Mash Screen - SeqSphere	Identity >= 0.95
1	Genome size	Assembly Statistics in SeqSphere	length of contigs assembled<ref genome + 10%
2/82	Breath coverage	Python	min coverage : 80%
2/82	Contamination	Confindr	around 10% (appreciation)
2/82	Coverage	BBtool	min 30X, max 100X
2/82	De novo assembly	Spades	
2/82	Gap Closing	GapCloser	
2/82	Genome Assembly Evaluation	Quast	
2/82	N50	Quast	Appreciation (no threshold)
2/82	Number of contigs	Quast	Appreciation (no threshold)
2/82	Scaffolding	MeDuSa	Delete scaffolds <200b
2/82	Trimming	Trimmomatic	Min lenght : 50pb, Phred score < 20
3	average coverage		>=10
3	contamination	CheckM	<4%
3	coverage cgMLST	SeqSphere	>=90%
3	GC%		51.6-52.3
3	genome completeness		>96%
3	N50		>10000

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
3	number of contigs		<300
3	phred score		>30
3	Total length		4.54-5.21 Mb
6/86	Core	BioNumerics 8.0	98%
6/86	Coverage	BioNumerics 8.0	>30
6/86	Genome size (contamination)	BioNumerics 8.0	4.6 - 5.2 million bp
6/86	N50	BioNumerics 8.0	>15,000
6/86	Number of Contigs	BioNumerics 8.0	< or equal to 400
8	allele calling result - percentage of good targets	Ridom SeqSphere	98%
8	assembly lenght	Ridom SeqSphere	~5MBases for Salmonella
8	coverage	Ridom SeqSphere	minimum 20-30x
8	No. of. contigs	Ridom SeqSphere	200 bases (contigs shorter than 200 have to be ignored)
11	Confirmation of genus	K-merFinder-3,2	
11	Contamination check	K-merFinder-3,2	
11	Coverage (depth)	FASTQC	>25x
11	GC content	FASTQC	%similar between strains
11	GC%	QUAST	>51 and <53
11	N50	QUAST	>55000
11	Number of contigs	QUAST	<500
11	Percent mactching targets in S.enterica cgMLST scheme	chewBBACA	more than 95%
11	Serotyping	SeqSeroV2	

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
12	confirmation of serotyping	SISTR	confirmed serotype
12	coverage depth	shovill - quast	>30
12	duplicated orthologs	shovill - quast	almost no duplicated orthologs
12	fraction of reads uniquely assigned to Salmonella enterica	KRAKEN	> 0.90
12	number of contigs	shovill - quast	>200
12	predicted species	mash	Salmonella species
12	Q30 base fraction	fastp	> 0.80
12	single copy orthologs (genome completeness)	shovill - quast	nearly all single copy orthologs
12	total length	shovill - quast	4.5-5.5 Mb for Salmonella
14	De novo assembly: contigs	BioNumerics	<115
14	De novo assembly: Sequence length	BioNumerics	[3.6 Mb, 6.0 Mb]
14	N50	BioNumerics	>48230 bp
14	Raw data statistics: expected coverage	BioNumerics	>30x
14	Summary calls: % core present	BioNumerics	>80%
17	Contamination check	Kraken2	> 5% contaminating species = fail
17	fastQC	fastQC	
17	Median coverage	bowtie2 map 2.3.0 - samtools depth 1.9	>20
17	N50	Quast	
17	Total length	Quast	around 5 Mbp
17	total number of contigs	Quast	< 500
18	#contigs	quast	<300

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
18	coverage mean	qualimap	>30
18	Genome fraction %	quast	>90%
18	insert size	Qualimap	350-400
18	N50	Quast	15000
18	Total Lenght	quast	4,8-5*10ˇ6 bp
19	basic statistics fastqc (for the reads)	fastqc	pass
19	coverage	fastqc, quast	>50
19	N50	quast	>50000
19	number of contigs (>1000)	quast	<100
21	Contamination	kraken	-
21	Coverage	FastQC	minimum 25x
21	N50	FastQC	minimum 10000
24	Coverage	Ridom SeqSphere	50
24	N50	Ridom SeqSphere	80 000
24	Perc. Good cgMLST Targets	Ridom SeqSphere	> 99 %
24	Quality score	Ridom SeqSphere	30
25	contamination	kraken2/centrifuge	
25	Coverage	fastqc	30x
25	N50	quast, http://bioinf.spbau.ru/quast	>20kb
25	number of bases	quast, http://bioinf.spbau.ru/quast	3.7 Mbp – 6.4 Mbp
28	Average read length	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Should be similar to the expected read length from the sequencing platform.
28	Contamination of genomic sequences	KmerFinder tool; https://cge.cbs.dtu.dk/services/KmerFinder	Pure bacterial culture

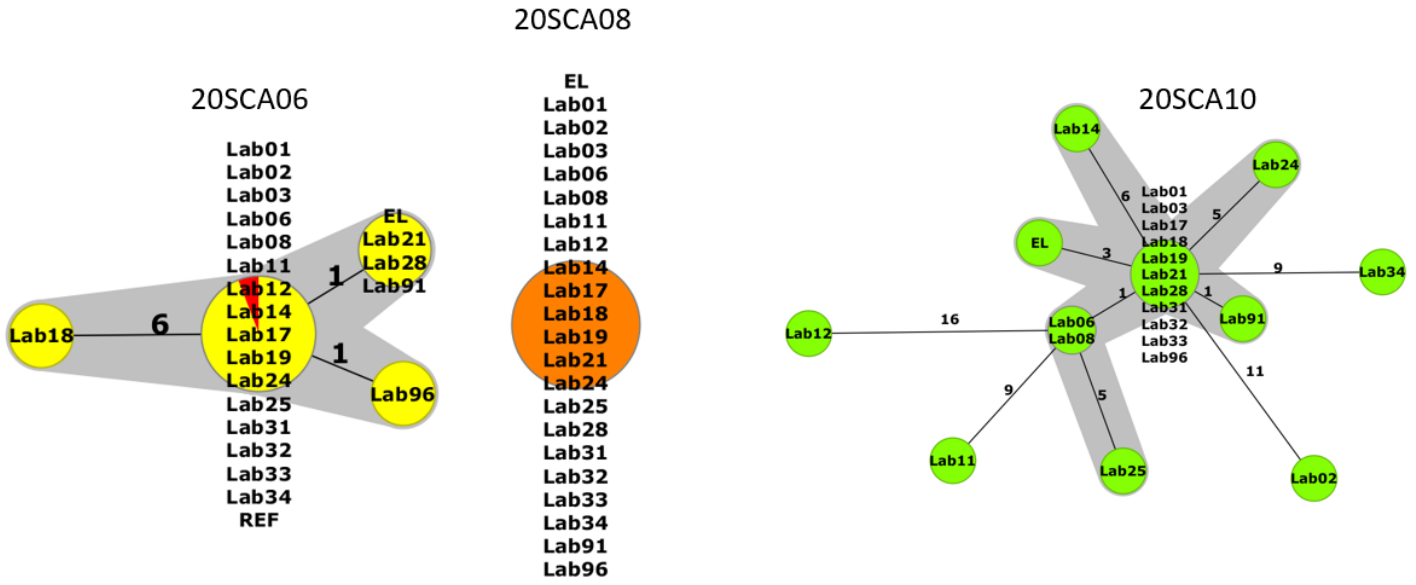
Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
28	Depth of coverage	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Coverage >30x
28	N50	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	>30 000 bp
28	Number of reads	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	The number of reads refers to the sequence yield, how much was sequenced. (No criteria established).
28	Size of assembled genome	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Deviation <0,5 million bp from the expected genome size.
28	Total number of contigs (after assembly)	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	<500 contigs
31	Coverage (after mapping)	QualiMap	Avg cov > 25
31	Nr of reads	FastQC	~0.6 M reads for 2x250 or ~1.0 M reads for 2x150.
31	Read length	Trimmomatic	36 bases
31	Sequence quality	FastQC, Trimmomatic	Sliding window trimming of bases with avg. qual < 20 in 4 bp window.
32	%GC	Data from multiQC	Not an exact threshold, but will give you an idea if you have sequenced the right species, so more like an indicator
32	Genome coverage	Data exported from fastQC/multiQC and calculated manually in Excel	Usually about 30X coverage to aim for
32	N50	Quast	Still no absolute threshold for this, but at least 15.000 bp (would probably be a bit sceptical to a dataset with lower than 50.000 bp, but will probably depend on species sequenced)
32	Number of contigs	Quast	We have still no exact threshold for this. We see that number of contigs might be species specific. But for now we lean towards suggestions from EU-RL AMR less than 500 contigs. But will

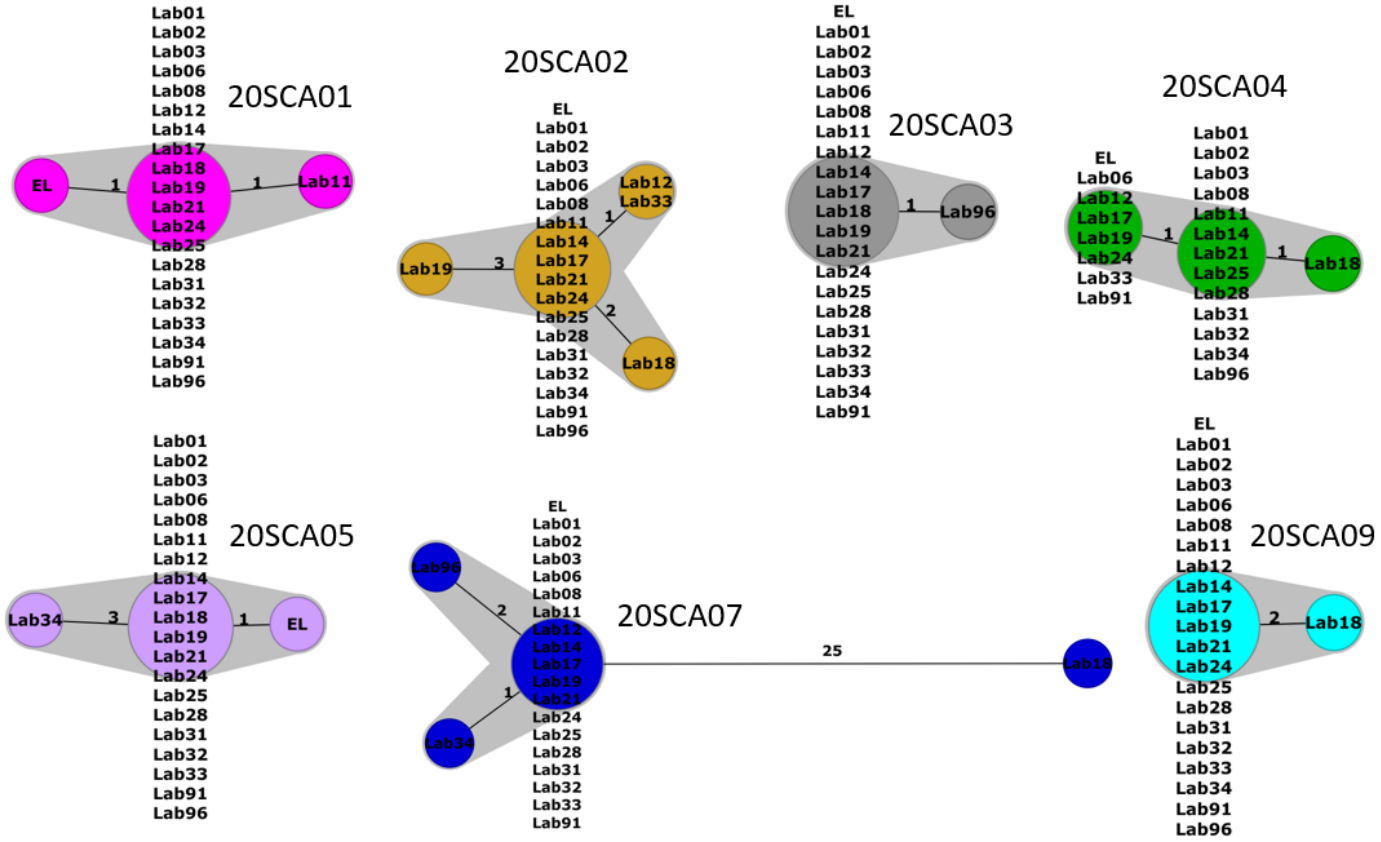
Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
			probably look into it if it's very different from what we use to see for a specific species.
32	Quality of raw reads	FastQC/MultiQC	Not a real threshold on this, also depending on read length etcbut will be evaluated
32	total length of assembly	Quast	If this differs too much from what to expect. We do not have an exact threshold for this, but lean towards suggested from ER-RL AMR +/- 20%
33	assembly Length	in house python script	4.6-5.2 Mbases
33	Coverage	in house python script	90X
33	minimum quality of reads	Trimmomatic	Q min al 3' > 20
33	N50	in house python script	>200000
33	num of contigs Longer than 200 bp	in house python script	< 250
34	Check contaminations in the sample	Kmerid	>75% Salmonella
34	Check contaminations in the sample	Shovill assembled gemones	4-5.8 Mbp
34	Check sequencing quaility	Qualimap	Mean coverage >30
34	Check sequencing quality	Quast	N50 value >30Kb
34	Check sequencing quality	Quast	number of contigs < 500
91	Assessment of bacterial contamination	KMER-look at similarity and reference genome,	it is rejected if there is >10% contamination
91	Minimum Read Count	in house	>10,000
91	Minimum Read Length	in house	>50 after trimming with trimmomatic
96	assembly contamination	https://github.com/Ecogenomics/CheckM	Completeness > 99.0 Contamination < 2.0

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
96	assembly statistics	https://github.com/Ecogenomics/CheckM	Max 5% variation of the Ref 4.8 Mb genome size
96	Contamination (on fastq)	https://github.com/B-UMMI/INNUca/blob/master/modules/trueCoverage_rematch.py; https://github.com/OLC-Bioinformatics/ConFindr	TrueCoverage_absente_genes < 2; TrueCoverage_multiple_alleles <1 Confindr_Genus "Salmonella" Confindr_NumContamSNVs < 30
96	Integrity and coverage of fastq	https://github.com/assemblerflow/flowcraft/blob/master/flowcraft/templates/integrity_coverage.py	integrity; raw coverage on 4.8Mb reference > 25x
96	Per sequence quality scores; Per sequence GC content	FASTQC	PASS
96	Percentage of missing loci in cgMLST	cgMLST loci list https://zenodo.org/record/1323684	< 2%
96	Serotype prediction	https://github.com/denglab/SeqSero2	No multiple serovar detected
96	species confirmation	mash screen; https://github.com/marbl/Mash; assess_mash_screen.py from https://zenodo.org/record/2541486#.YAgpk-hKiUk	single species equal to "Salmonella enterica"
EL	Contamination	KrakenBracken	<4%
EL	Coverage	Formula: (total reads * length of read)/length of genome sequenced	>30
EL	GC%	FastQC	51.6-52.3
EL	N50	FastQC, Seqsphere	>10000
EL	number of contigs	FastQC, Seqsphere	<300
EL	Total length	FastQC, Seqsphere	4,5-5,2 Mbases

Annex 11 WGS results cluster analysis part, Minimum Spanning Tree per strain

MST for each strain, using all participants' processed raw data (Ridom SeqSphere+, *S. enterica* MLST (7) and cgMLST (3002), pairwise ignoring missing values).




20SCA01
EL
Lab11
1
1
20SCA02
Lab19
Lab12
Lab33
Lab18
3
1
2
20SCA03
Lab96
1
20SCA04
EL
Lab06
Lab12
Lab17
Lab19
Lab24
Lab33
Lab91
Lab18
1
1
20SCA05
Lab34
EL
3
1
20SCA07
Lab96
Lab34
2
1
25
Lab18
20SCA09
Lab18
2

Annex 12 Results QC parameters on the *de novo* assembled genomes, per participant

All statistics are based on contigs of size ≥ 500 bp, except Total reads and Coverage.

	Laboratory code: 01			Platform used: MiniSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	79	84	85	79	82	81	58	81	84	83
Largest contig	633697	436204	436204	825354	825354	825451	435777	825354	632210	604974
Total length	4975941	4877563	4912989	4981644	4982213	4981687	4883657	4982460	4971348	4953379
N50	267185	223164	224133	271057	223164	267192	247375	223809	201417	204744
Total reads	1915912	1336590	1504464	2786594	2307042	1052890	1236534	1607938	1913330	1898560
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	58	41	46	84	69	32	38	48	58	57

	Laboratory code: 02			Platform used: NovaSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	307	158	194	135	125	170	102	218	222	257
Largest contig	428740	342046	342046	435151	723309	389138	436145	436388	812326	248438
Total length	4996449	4886932	4920958	4985506	4987352	4981331	4888111	4993555	4974943	4964527
N50	43554	107215	66583	156067	113354	105386	112602	60110	58341	53625
Total reads	29763088	35031750	32818088	36567666	34600968	34559616	34474118	32297922	38853002	37193164
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	814	995	912	1016	956	936	1030	873	1075	1015

	Laboratory code: 03			Platform used: NextSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	74	77	79	76	76	79	57	77	84	74
Largest contig	868089	601225	632924	825354	825354	825354	825089	825354	812037	604784
Total length	4978709	4880334	4912461	4982932	4981578	4986962	4885402	4983110	4973037	4955169
N50	271050	271057	231741	223915	271057	271057	247375	271057	201337	180265
Total reads	11165414	9259254	9440380	6604510	9744502	10454486	6864030	9800132	11242546	9404488
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	335	284	288	198	291	313	210	293	337	282

	Laboratory code: 06			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	68	68	67	66	68	65	45	64	70	66
Largest contig	873560	602734	634712	907857	825655	907857	907776	907857	812135	682099
Total length	4989684	4891657	4924958	4992669	4991611	4991697	4894400	4993031	4985285	4965032
N50	270584	283111	282993	316084	282875	316045	376819	316084	270472	225691
Total reads	2196198	2085008	1981692	1920564	1538362	1611346	1641132	2174824	1999992	2137042
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	132	128	121	115	92	97	100	130	120	129

	Laboratory code: 08			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	80	83	90	90	90	95	67	93	76	83
Largest contig	873928	602734	634712	634440	377001	825471	732399	908225	811951	602921
Total length	4996301	4897926	4939642	4997424	4998153	5002327	4903367	5002841	4988551	4975458
N50	270584	270591	270591	251294	202946	241951	225813	270591	270472	180514
Total reads	2137988	1663830	2035134	1766494	1796722	1840508	1157162	1790594	1722514	1779322
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	128	102	123	106	108	110	71	107	103	107

	Laboratory code: 11			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	93	81	87	84	79	70	60	166	114	76
Largest contig	732357	602734	436528	494770	634440	907857	679673	194200	377001	682102
Total length	4993846	4901526	4925556	4993745	4994090	4997318	4899820	4984210	4995182	4975330
N50	239495	270591	165103	184319	190164	275676	320291	75212	123427	204652
Total reads	782370	771942	573790	667386	682848	832466	704882	869056	708724	1518832
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	47	47	35	40	41	50	43	52	42	91

	Laboratory code: 12			Platform used: NextSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	83	80	81	80	79	80	66	82	85	85
Largest contig	401568	600666	632522	632312	632406	632652	602642	632203	632084	396768
Total length	4968848	4872509	4904509	4974572	4974955	4973618	4881966	4974222	4964821	4944225
N50	253597	270447	276852	276852	282312	267148	247256	276793	201103	178557
Total reads	2750382	3049900	3227604	2853982	2861532	3139026	3058192	3072376	2770184	2334568
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	83	94	98	86	86	94	94	92	83	71

	Laboratory code: 14			Platform used: NovaSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	80	83	82	78	79	79	58	78	83	79
Largest contig	633767	600753	632993	825170	632997	632650	605785	528368	631912	604980
Total length	4976670	4879237	4912202	4982871	4981683	4979635	4884804	4980224	4971926	4953095
N50	267185	223164	223164	271057	271057	267192	247375	267192	229042	205017
Total reads	7293620	6948068	7894752	7138042	7532896	7161558	7565994	7734070	7124454	6122834
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	219	213	241	215	226	215	232	232	215	185

	Laboratory code: 17			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	189	165	116	189	151	100	151	140	135	112
Largest contig	950592	602734	858732	908041	907857	907857	907776	907857	811951	601748
Total length	5082753	4966413	4964499	5076330	5051610	5017118	4969396	5038951	5036939	4997901
N50	282867	316045	299823	316045	316486	282875	376819	316045	270472	223067
Total reads	1322978	1192560	1156244	1585194	1195648	1186780	1286240	1562994	1278106	1295966
Read length	250	250	250	250	250	250	250	250	250	250
Coverage	63	58	57	75	57	57	63	74	61	63

	Laboratory code: 18			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	78	171	70	105	92	350	545	83	246	72
Largest contig	873560	272843	634712	323666	552004	97013	104515	634426	107489	682099
Total length	4994479	4853455	4925177	4978145	4986086	4856764	4590032	4992099	4887291	4968871
N50	270584	78133	282993	129702	161519	33021	19030	176086	40669	225691
Total reads	1001066	1075324	1070942	1003896	1046702	1090012	842242	1159258	1058266	1047584
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	60	66	65	60	63	67	55	69	65	63

	Laboratory code: 19			Platform used: NovaSeq 6000						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	79	99	83	82	80	80	63	77	81	80
Largest contig	478970	436388	528582	436204	632882	436388	435961	436204	812037	605157
Total length	4977885	5006715	4911544	4979531	4979784	4981069	4885367	4980444	4974506	4954117
N50	229042	224022	229042	223809	223164	223809	224099	223815	174707	178330
Total reads	2758610	7934798	7522138	5613840	5255196	6240626	6229086	5172510	6752252	10735352
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	83	237	229	169	158	187	191	155	203	324

	Laboratory code: 21			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	82	82	81	74	81	78	62	90	87	83
Largest contig	634885	436204	436204	825538	633345	436204	436145	436204	424233	377524
Total length	4977074	4879098	4911712	4980897	4980839	4981276	4884807	4980877	4973368	4952310
N50	239141	224027	223164	223164	223164	223164	224994	223164	201417	176966
Total reads	1495098	1536604	1198572	1109826	1385024	918452	1518660	1099988	1405140	1246144
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	45	47	36	33	42	28	47	33	42	38

	Laboratory code: 24			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	112	65	68	68	65	66	46	66	73	67
Largest contig	873560	602734	634712	907857	907856	907857	907776	907857	811951	682100
Total length	5026536	4893283	4926864	4993759	4994461	4992700	4895426	4994550	4987105	4965539
N50	270584	316045	316355	316045	316084	316045	376819	316084	270472	225691
Total reads	2646776	2371816	2634926	2647254	2832844	2659882	2498948	2559338	3444030	1884078
Read length	250	250	250	250	250	250	250	250	250	250
Coverage	131	121	133	132	141	133	127	128	172	95

	Laboratory code: 25			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	74	76	128	131	103	137	104	132	142	67
Largest contig	873560	602734	634712	907857	907857	908041	907776	825471	811951	682099
Total length	4993262	4898396	4964167	5009037	5013007	5034352	4925174	5027075	5029119	4967223
N50	270584	316045	283092	316084	316045	316045	376819	282782	270472	225691
Total reads	5406048	4193758	3410166	3894938	3790014	3672242	2808040	3198368	2877692	2854506
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	324	256	204	231	225	217	169	189	170	172

	Laboratory code: 28			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	94	81	83	83	122	77	80	92	84	101
Largest contig	325733	303636	316961	303419	278390	280062	278376	235488	351114	277712
Total length	4921254	4779587	4835881	4883418	4920373	4807902	4841977	4855873	4903686	4843506
N50	129883	130058	112525	130057	124266	134933	134235	130055	101685	87518
Total reads	5948101	6115329	6014150	6494541	34600896	6074060	6690831	7048012	5759235	6981082
Read length	75	75	75	75	75	75	75	75	75	75
Coverage	90	96	93	100	523	95	104	109	88	108

	Laboratory code: 31			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	71	79	81	90	108	100	76	97	97	113
Largest contig	873293	601834	634712	464132	822584	373806	533499	635176	757981	472409
Total length	4987426	4893786	4925526	4990262	5000795	4992896	4897755	4991573	4988540	4970474
N50	270584	270591	282782	204144	282782	169688	170065	192009	211974	99139
Total reads	2335522	2550094	2693994	2310236	1846586	2579844	2853784	2396840	1841978	2277724
Read length	250	250	250	250	250	250	250	250	250	250
Coverage	117	130	136	115	91	129	145	119	92	114

	Laboratory code: 32			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	76	72	79	79	71	71	102	69	74	69
Largest contig	791174	601629	633607	825471	825655	907857	825390	825471	811951	682099
Total length	4989443	4891889	4931230	4996418	4992107	4991906	4943244	4991953	4985180	4965300
N50	253813	282900	316200	282875	282875	316045	290501	282875	270472	225691
Total reads	1148746	817978	1258666	1240460	1119506	724812	1110606	984250	1014766	1256576
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	69	50	76	74	67	43	67	59	61	76

	Laboratory code: 33			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	79	504	98	72	80	91	51	72	86	89
Largest contig	873560	601629	450142	907857	828270	825229	815064	825655	811951	682099
Total length	4999619	5209875	4939822	4999166	5002342	5008845	4901805	4998742	4994198	4978955
N50	253813	282875	270591	316045	282875	282875	261960	282875	213335	196920
Total reads	2779548	2768060	2840554	2603542	2838348	3007394	2027022	2893346	3351570	3528604
Read length	300	300	300	300	300	300	300	300	300	300
Coverage	166	156	172	156	170	179	124	173	201	212

	Laboratory code: 34			Platform used: MiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	161	190	343	420	534	99	139	149	91	83
Largest contig	329910	179518	146539	129488	108512	632892	248728	220720	407647	377513
Total length	4977965	4884617	4943990	5022300	5064029	4985677	4874138	4975885	4972250	4953257
N50	82197	74226	34893	28039	24193	186541	72306	82666	174590	176966
Total reads	3600000	3600000	3600000	3600000	3600000	3291752	4245246	3578478	2414834	4071026
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	107	106	101	84	93	97	130	107	71	123

	Laboratory code: 91			Platform used: HiSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	85	84	81	78	79	79	63	74	85	81
Largest contig	633912	464239	528434	632676	632849	528238	464515	633954	464111	605218
Total length	4974078	4878836	4911114	4978150	4980954	4979556	4888105	4980730	4971412	4954606
N50	267185	186541	228810	282450	282450	267192	224994	284002	225110	176966
Total reads	3315112	4526834	3276108	4254940	4106738	3532012	4539482	2942342	2256424	4115514
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	100	139	100	128	123	106	139	88	68	124

	Laboratory code: 96			Platform used: NextSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	84	85	80	83	86	82	65	81	87	85
Largest contig	475848	496267	632709	433494	464407	464407	602633	632580	631813	601809
Total length	4966783	4870938	4904272	4973146	4973244	4972439	4879409	4972792	4964114	4940669
N50	267117	222996	231697	233707	222996	233707	247324	282312	201395	176922
Total reads	6776282	5886728	7656788	7108676	7347760	8020248	7084090	3569892	6914108	7136036
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	204	181	234	214	221	242	217	107	209	216

	Laboratory code: EL*			Platform used: NovaSeq						
Parameters:	**20SCA01**	**20SCA02**	**20SCA03**	**20SCA04**	**20SCA05**	**20SCA06**	**20SCA07**	**20SCA08**	**20SCA09**	**20SCA10**
# contigs	82	83	64	81	80	80	61	79	86	85
Largest contig	478602	436388	635370	436388	633240	436388	436145	633091	424417	378610
Total length	4977096	4880625	4916282	4972443	4979726	4972730	4885701	4979244	4972086	4952708
N50	174626	186541	304127	220732	220732	223164	224997	223164	201337	168606
Total reads	3979002	3557272	6057630	3473132	4065556	4906268	3415146	4375690	4426384	3770568
Read length	150	150	150	150	150	150	150	150	150	150
Coverage	120	109	185	105	122	148	105	132	133	114

EL*: EURL-*Salmonella*, 18-11-2020 (PT Typing 2020)

	Laboratory code: EL_1*			Platform used: NovaSeq						
	20SCAT02	20SCAT18	20SCAT08	20SCAT14	20SCAT11	20SCAT09	20SCAT03	20SCAT09	20SCAT12	20SCAT05
Parameters:	**=20SCA01**	**=20SCA02**	**=20SCA03**	**=20SCA04**	**=20SCA05**	**=20SCA06**	**=20SCA07**	**=20SCA08**	**=20SCA09**	**=20SCA10**
# contigs	75	77	78	75	76	75	58	na	81	77
Largest contig	868736	602177	633778	826185	826185	826185	825920	na	812868	605300
Total length	4979155	4882562	4913747	4983836	4983060	4982262	4887336	na	4974301	4955873
N50	239141	224022	223164	224022	224022	224022	222779	na	174620	180265
Total reads	8915196	6410576	8691864	6626758	8156888	8194566	6497230	na	8429066	8585364
Read length	150	150	150	150	150	150	150	na	150	150
Coverage	268	197	265	199	245	246	199	na	254	259

EL_1*: EURL-*Salmonella*, 25-9-2020 (first pre-test of the strains)

	Laboratory code: EL_2*			Platform used: NovaSeq						
	20SCAT02	20SCAT18	20SCAT08	20SCAT14	20SCAT11	20SCAT09	20SCAT03	20SCAT09	20SCAT12	20SCAT05
Parameters:	**=20SCA01**	**=20SCA02**	**=20SCA03**	**=20SCA04**	**=20SCA05**	**=20SCA06**	**=20SCA07**	**=20SCA08**	**=20SCA09**	**=20SCA10**
# contigs	na	73	73	72	70	73	59	na	81	77
Largest contig	na	603282	635254	825817	826185	825817	825920	na	812868	605821
Total length	na	4883207	4914924	4983554	4984072	4983211	4887068	na	4970468	4955935
N50	na	224022	224022	224022	224022	224022	222779	na	174620	180265
Total reads	na	5996302	7231842	6408318	6742540	7139956	7472790	na	9381824	9355584
Read length	na	150	150	150	150	150	150	na	150	150
Coverage	na	184	220	192	202	214	229	na	282	282

EL_2*: EURL-*Salmonella*, 8-10-2020 (second pre-test of the strains, after ten times sub-culturing

RIVM
Committed to *health and sustainability* -