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

EURL-*Salmonella* Proficiency Test Typing 2022

RIVM report 2023-0339
W.F. Jacobs-Reitsma et al.

EURL-*Salmonella* Proficiency Test Typing 2022

RIVM report 2023-0339

Colophon

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-2023-0339

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 within the framework of RIVM project number E/114506/21 European Union Reference Laboratory for *Salmonella* (2022) and was co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the granting authority European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

Published by:
National Institute for Public Health and the Environment
P.O. Box 1 | 3720 BA Bilthoven
the Netherlands
www.rivm.nl/en

Co-funded by the European Union

Synopsis

EURL-*Salmonella* Proficiency Test Typing 2022

Since 1992, National Reference Laboratories (NRLs) of European Union (EU) Member States have been obliged to participate in annual quality control 'Proficiency' Tests (PTs). One of the PTs is on typing of *Salmonella* bacteria. The NRLs of all 27 EU Member States performed well in this 2022 quality control test on *Salmonella* typing. Two laboratories were found to require a follow-up study after the initial test. Overall, the participating laboratories were able to assign the correct name to 98% of the strains tested.

Laboratories are obliged to type *Salmonella* with the reference method (serotyping). In 2022, they could also perform additional typing at the 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 correctly identify 20 *Salmonella* strains.

NRLs from countries outside the EU occasionally participate in these tests on a voluntary basis. Seven countries took part in 2022: the United Kingdom, the (potential) EU candidate countries Kosovo, Moldova, and Türkiye as well as the European Free Trade Association (EFTA) countries Iceland, Norway and Switzerland.

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, MLVA, WGS, cluster analysis, Proficiency Test

Publiekssamenvatting

EURL-*Salmonella* ringonderzoek typering 2022

Sinds 1992 zijn de Nationale Referentie Laboratoria (NRL's) van de lidstaten van de Europese Unie verplicht om elk jaar hun kwaliteit te laten toetsen met zogeheten ringonderzoeken. Een van de ringonderzoeken is de typering van *Salmonella*-bacteriën. In 2022 scoorden alle NRL's van de 27 EU-lidstaten goed bij deze kwaliteitscontrole op typering van *Salmonella*. Twee laboratoria hadden hiervoor een herkansing nodig. Als groep konden de deelnemende laboratoria aan 98 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 2022 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.

Soms doen er ook NRL's van landen buiten de EU vrijwillig aan mee. In 2022 waren dat er zeven: het Verenigd Koninkrijk, de EU (potentiële) kandidaat-lidstaten Kosovo, Moldavië, en Turkije 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, MLVA, WGS, cluster analyse, ringonderzoek

Contents

Summary

In November 2022, the annual *Salmonella* typing Proficiency Test (PT) was organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands). The PT'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 34 laboratories participated in this PT. These included the obligatory 27 NRLs-*Salmonella* in the 27 EU Member States. Seven additional NRLs participated voluntarily: the United Kingdom, the EU (potential) candidate countries Kosovo, Moldova, and Türkiye as well as the European Free Trade Association (EFTA) countries Iceland, Norway and Switzerland.

All 34 laboratories performed serotyping. The EURL-*Salmonella* selected a total of twenty obligatory *Salmonella* strains plus one optional strain 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).

Overall, nearly 100% of the strains were typed correctly for the O-antigens, 98% of the strains were typed correctly for the H-antigens, and 98% of the strains were correctly named by the participants.
In 2007, criteria for 'good performance' concerning serotyping were defined. Based on these criteria, the participants' performance was very good, including the three participants that submitted Whole Genome Sequencing(WGS)-based serotyping results. All but two participants met the criteria for good performance in the first stage of the PT. Two participants had to participate in a follow-up study, including ten additional strains for serotyping. Ultimately, all 34 evaluated NRLs achieved good performance.

The PT Typing 2022 also included an optional part on cluster analysis. The cluster analysis involved six 'wet' *Salmonella* strains and allowed participants to choose either Multiple-Locus Variable number of tandem repeat Analysis (MLVA) and/or WGS.The PT was mimicking an outbreak situation, with a *Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1 as the reference strain.
Raw sequence data on this reference strain, as well as on another six *Salmonella* strains, were made available to the participants via a secure ftp server for 'dry' evaluation.
Participants were asked to analyse the six 'wet' *Salmonella* strains (MLVA/WGS) and the six 'dry' *Salmonella* strains (WGS only), and to report per strain whether a clustering match with the reference strain was found or not.
A total of 20 NRLs participated in the cluster analysis: all 20 performed WGS analysis and five participants also performed MLVA analysis.

The participants' cluster analysis results were evaluated by comparing their results to the expected results in the outbreak investigation setting, as pre-defined by the EURL-*Salmonella.*
Four of the five participants reported the MLVA-based cluster analysis results fully as expected. Nineteen of the 26 submissions (by fifteen of the twenty participants) reported the WGS-based cluster analysis results fully as expected. Three participants' deviations were mainly due to a misunderstanding or mistakes in reporting the data. Three participants did not exclude strain 22SCA13 from their cluster analysis, although this strain was expected not to pass the participants' quality control (QC), because it also contained numerous *E. coli* reads.

1 Introduction

This report describes the 2022 Proficiency Test (PT) on the typing of *Salmonella* organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands) in November 2022.

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 34 laboratories participated in the PT Typing 2022. These included 27 NRLs-*Salmonella* in the 27 EU Member States and seven NRLs from third countries (EU candidate or potential EU candidate Member States, members of the European Free Trade Association (EFTA), and the United Kingdom).
The main objective of this PT was to evaluate the performance of the EU NRLs in serotyping *Salmonella*. All NRLs performed serotyping of the 20 obligatory strains, and all but five participants serotyped the optional 21st strain. NRLs of EU Member States that would not achieve the defined level of good performance for serotyping had to participate in a follow-up study.

The PT Typing 2022 also included an optional part on cluster analysis. The cluster analysis involved six 'wet' *Salmonella* strains and allowed participants to choose either MLVA and/or WGS.
The PT was mimicking an outbreak situation, with a *Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1 as the reference strain.
Raw sequence data on this reference strain, as well as on another six *Salmonella* strains, were made available to the participants via a secure ftp server for 'dry' evaluation.
Participants were asked to analyse the six 'wet' *Salmonella* strains (MLVA/WGS) and the six 'dry' *Salmonella* strains (WGS only), and to report per strain whether a clustering match with the reference strain was found or not.
A total of 20 NRLs participated in the cluster analysis: all 20 performed WGS analysis and five participants also performed MLVA analysis.

2 Participants

Country	City	Institute
Austria	Graz	AGES
Belgium	Brussels	Sciensano
Bulgaria	Sofia	National Diagnostic Research Veterinary Medical Institute (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
Iceland	Reykjavík	Landspítali University Hospital, Dept. of Clinical Microbiology
Ireland[a)]	Celbridge	Central Veterinary Research Laboratory
Italy	Legnaro	Istituto Zooprofilattico Sperimentale delle Venezie
Kosovo	Prishtina	Kosovo Food and Veterinary Laboratory
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
Moldova	Chisinau	Republican Center for Veterinary Diagnostic
Netherlands	Bilthoven	RIVM, Centre for Infectious Diseases Research, Diagnostics and Screening (IDS)
Norway	Ås	Norwegian Veterinary Institute
Poland	Pulawy	National Veterinary Research Institute
Portugal	Oeiras	Instituto Nacional de Investigação Agrária e Veterinária (INIAV)
Romania	Bucharest	Institute for Diagnosis and Animal Health
Serbia	Belgrade	NIVS Veterinary Institute of Serbia
Slovak Republic	Dolný Kubín	Veterinary and Food Institute in Dolný Kubín
Slovenia	Ljubljana	UL, Veterinary Faculty, NVI
Spain	Algete-Madrid	Laboratorio Central de Veterinaria
Sweden	Uppsala	National Veterinary Institute (SVA)
Switzerland	Zurich	Institute for Food Safety and Hygiene
Türkiye	Ankara	Veterinary Control Central Research Institute
United Kingdom	Addlestone	Animal and Plant Health Agency (APHA) Weybridge

[a)] Also representing the NRL-*Salmonella*-Typing in Northern-Ireland.

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.

3.1.2 Protocol and test report

Three weeks before the start of the PT, the NRLs received the protocol by email. Participants used web-based forms to submit their results. Instructions for completing these result forms and data entry were sent to the NRLs on 8 November 2022, in separate emails for serotyping and cluster analysis.
The protocol and screenshots of the 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 7 November 2022. 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*

Participants had to serotype a total of twenty *Salmonella* strains (coded S1–S20). As agreed at the 27th EURL-*Salmonella* Workshop (Mooijman, 2022), an 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 serologically verified by the Centre before distribution. Table 3.1 presents the complete antigenic formulas of the 21 serovars in accordance with the most recent White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007) plus published supplements no. 47 (Guibourdenche et al., 2010) and no. 48 (Issenhuth-Jeanjean et al., 2014). However, participants were asked to report only the results as detected and on which the identification of serovar names was based. Eleven strains (Table 3.1) represented serovars included in the EURL-*Salmonella* serotyping PTs for the first time.

Table 3.1 Antigenic formulas of the 21 Salmonella *strains according to the White-Kauffmann-Le Minor scheme used in the EURL-*Salmonella *PT Serotyping 2022*

Strain code	O-antigens	H-antigens		Serovar	Origin
		(phase 1)	(phase 2)		
S1[a)]	6,7	k	e,n,x	Singapore	Human
S2[b)]	1,4,[5],12	i	-	1,4,[5],12:i:-	Human
S3[a)]	4,12	z_{10}	1,6	Tudu	Human
S4	1,4,[5],12	f,g,s	[1,2]	Agona	Human
S5	1,9,12	g,m	-	Enteritidis	Human
S6[a)]	6,7	l,z_{13}	e,n,x	Kenya	Human
S7	1,13,23	m,t	-	Kintambo	Human
S8 [a) c)]	4,[5],12	i	e,n,x	Farsta	Reptile
S9	6,8	z_{10}	e,n,x	Hadar	Human
S10 [a)]	16	b	1,2	Hull	Human
S11 [a)]	11	d	[e,n,x]	Chandans	Human
S12	6,7,14	r	1,2	Virchow	Human
S13 [a)]	1,4,[5],12	g,m,s	[1,2]	Hato	Human
S14	1,9,12	e,h	1,5	Eastbourne	Human
S15 [a)]	3,10	z_{35}	z_6	Cairina	Human
S16 [a)]	1,6,14,25	a	1,5	Garba	Human
S17 [a)]	1,13,23	d	1,5	Mishmarhaemek	Human
S18	1,4,[5],12	i	1,2	Typhimurium	Human
S19 [a)]	28	c	1,5	Hermannswerder	Human
S20	6,7,14	r	1,5	Infantis	Human
S21[d)]	47	k	z_{35}	47:k:z_{35} (IIIb)	Human

[a)] Represented in an EURL-*Salmonella* PT Serotyping for the first time.
[b)] Monophasic variant of *S.* Typhimurium based on genomic sequences. Phenotypic result: 4,5:i:-.
[c)] In accordance with Supplement no. 48 to the White-Kauffmann-Le Minor scheme.
[d)] *Salmonella enterica* subspecies *diarizonae* (optional strain).

3.2.2 *Evaluation of the serotyping results*

The evaluation of deviating serotyping results is presented in Table 3.2.

Table 3.2 Evaluation of deviating 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 (results with four penalty points or more) have to participate in a follow-up study.

3.2.3 *Follow-up study serotyping*

The follow-up study for serotyping consisted of typing an additional set of ten *Salmonella* strains. The strains selected for the follow-up study are shown in Table 3.3.

Table 3.3 Antigenic formulas of the ten Salmonella *strains according to the White-Kauffmann-Le Minor scheme used in the follow-up part of the EURL-*Salmonella *PT Serotyping 2022*

Strain code	O-antigens	H-antigens		Serovar	Origin
		(phase 1)	(phase 2)		
SF1	28	c	1,5	Hermannswerder	Human
SF2	3,{10}{15}{15,34}	e,h	1,6	Anatum	Human
SF3[a)]	4,[5],12	i	e,n,x	Farsta	Reptile
SF4[b)]	1,4,[5],12	i	-	1,4,[5],12:i:-	Human
SF5	16	c	l,w	Yoruba	Meat and bone meal
SF6	1,13,23	m,t	-	Kintambo	Human
SF7	3,{10},{15}	r	z_6	Weltevreden	Human
SF8	11	d	[e,n,x]	Chandans	Human
SF9	4,12	z_{10}	1,6	Tudu	Human
SF10	{6,7,14}{54}	g,m,[p],s	[1,2,7]	Montevideo	Human

[a)] In accordance with Supplement no. 48 to the White-Kauffmann-Le Minor scheme.
[b)] Monophasic variant of *S.* Typhimurium based on genomic sequences. Phenotypic result: 4,5:i:-.

3.3 Cluster analysis part of the PT

3.3.1 Salmonella *strains for cluster analysis*

A total of six *Salmonella* strains (22SCA01 – 22SCA06) in Heart Infusion (HI) agar transport tubes were sent to the participants in the part on cluster analysis. Background information on these 'wet' strains is given in Table 3.4A. In addition, raw sequence data (fastq.gz files, md5 checksums) on another seven *Salmonella* strains (22SCA11 – 22SCA16, plus the 22SCA-REF) were made available to the participants via a secure ftp server for 'dry' evaluation (WGS only). Background information on the 'dry' strains is given in Table 3.4B.

Table 3.4A Background information on the 'wet' Salmonella *strains used for cluster analysis in 2022*

Strain code	Serovar	ST	Origin	MLVA-profile
22SCA01 [a)]	Enteritidis	11	Human	3-10-5-3-1
22SCA02 [c)]	Enteritidis	11	Human	3-10-4-4-1
22SCA03	Enteritidis	11	Human	2-9-7-4-2
22SCA04 [b)] =REF	Enteritidis	11	Human	3-10-6-3-1
22SCA05 [a)]	Enteritidis	11	Human	3-10-5-3-1
22SCA06 [b)]	Enteritidis	11	Human	3-10-6-3-1

a) Technical duplicates.
b) Technical duplicates.
c) Biological duplicate strain 21SCA08.

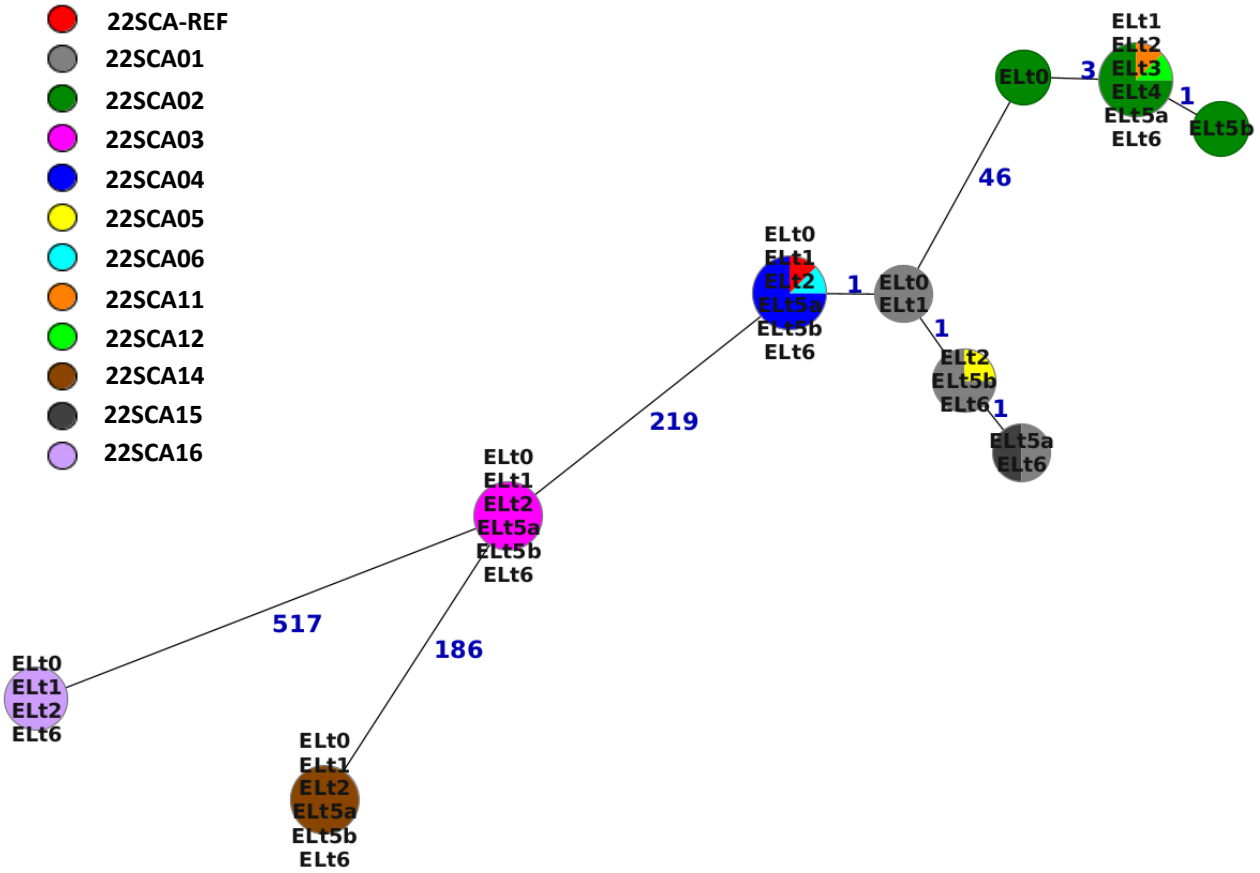
Table 3.4B Background information on the 'dry' Salmonella *strains used for cluster analysis in 2022 (WGS only)*

Strain code	Serovar	ST	Origin
22SCA11 [c)]	Enteritidis	11	Human
22SCA12 [c)]	Enteritidis	11	Human
22SCA13 *	Enteritidis	n.a.	Unknown
22SCA14	Enteritidis	11	Human
22SCA15 [d)]	Enteritidis	11	Human
22SCA16	Enteritidis	11	Human
22SCA-REF [e)]	Enteritidis	11	Human

c) Strain 21SCA08, raw data PT 2021 from 2 different participants.
d) Biological duplicate strain 22SCA01.
e) EURL-*Salmonella* raw data strain 22SCA04 (after the one-year storage at minus 70°C, ELt5a, Figure 3.1).
* *S.* Enteritidis contaminated with *E. coli* reads.
n.a. not applicable (QC not passed).

Strains were selected by the EURL-*Salmonella* to be suitable for analysis using either MLVA ('wet' strains only) or WGS. In preparation of the PT 2021 on cluster analysis, a set of 15 human surveillance *Salmonella* strains were re-cultured from storage (2019) on blood-agar plates and submitted for MLVA and WGS analysis both directly and after sub-culturing for ten times alternately in Buffered Peptone Water (BPW) and on blood-agar plates. Re-cultured strains were stored both at minus 70°C and in HI agar transport tubes. Strains were re-cultured again on blood-agar plates from both types of storage in the summer of 2022 and submitted for WGS analysis. Subsequently, six 'wet' strains and seven 'dry' strains (including the reference) were selected for inclusion in the PT 2022 (also see Figure 3.1). Two sets of 'wet' technical duplicates were included: strain 22SCA01 and strain 22SCA05 shipment tubes were both prepared from the same blood-agar plate containing strain 22SCA01; strain 22SCA04 and strain 22SCA06 shipment tubes were both prepared from the same blood-agar plate containing strain 22SCA04.
Figure 3.1 shows the WGS pre-test results as well as the EURL-*Salmonella* PT 2022 results for the twelve selected strains, and also includes the reference strain (Tables 3.4A and 3.4B).
Sequencing was performed in-house, on an Illumina NextSeq platform. Raw data were processed via an in-house developed Juno-assembly

pipeline (https://rivm-bioinformatics.github.io/ids_bacteriology_man/juno-assembly.html), which includes the SPAdes 3.15.3 assembler. Cluster analysis was done in Ridom SeqSphere+, using the cgMLST Enterobase v2.0 scheme and visualised in a minimum spanning tree (MST, Figure 3.1).
Stable and consistent cgMLST analysis results were obtained for both the minus 70°C-stored and the HI-tubes-stored strains (Figure 3.1, ELt5a and ELt5b). Subsequently, the 'wet' strains selected to be included for the PT 2022 (Figure 3.1, ELt6) were freshly prepared from the minus 70°C stocks (2021).



ELt0: Original WGS data from the stored human surveillance Salmonella strains (2019);
ELt1: WGS data from initial pre-testing for PT 2021 (8 July 2021);
ELt2: WGS data after ten times sub-culturing (blood-agar/BPW) for PT 2021 (17 August 2021);
ELt3: PT 2021 data at the start of the PT (November 2021);
ELt4: PT 2021 data at the end of the PT (February 2022);
ELt5a: WGS data after one-year storage at minus 70°C (September 2022);
ELt5b: WGS data after one-year storage in HI agar transport tubes (September 2022);
ELt6: PT 2022 data at the start of the PT (November 2022).

Figure 3.1 MST of the EURL-Salmonella (EL) pre-tests and PT 2022 results, (RidomSeqSphere+, cgMLST (3002), pairwise ignoring missing values).

3.3.2 *Evaluation of the cluster analysis results in general*

Cluster analysis could be performed up to the choice of the participant by MLVA and/or WGS, and using their own routine method(s).

Like in the previous two years, the PT Cluster Analysis 2022 was mimicking an outbreak situation, with a *Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1 as the reference strain (22SCA-REF).

However, the number of strains to be shipped to the participants for 'wet' analysis (MLVA/WGS) was reduced from ten to six. In addition, raw WGS data of six strains plus the reference strain were made available through a secure ftp server for 'dry' analysis (WGS only), reducing the workload and costs on the 'wet' lab part for these strains.

Participants were asked to analyse the six 'wet' *Salmonella* strains (MLVA/WGS) and the six 'dry' *Salmonella* strains (WGS only) and to report per strain whether a clustering match with the given 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 WGS had to be sent by email or uploaded to the secure ftp server.

Evaluation (per methodology, see sections 3.3.3 and 3.3.4) 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* (Protocol EURL-*Salmonella* PT Typing 2022).
https://www.eurlsalmonella.eu/documenten/protocol-eurl-salmonella-pt-typing-2022

No specific performance criteria were set for this PT on cluster analysis. As a minimum, it was expected that participants would report any technical duplicate strains to be (part of) one cluster.

3.3.3 *Evaluation of the cluster analysis results based on MLVA data*

Data submission for MLVA results included:

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

Participants were asked to report per strain (Tables 3.4A/B) whether or not they found a clustering match with the given reference outbreak strain (22SCA-REF) in the EURL-*Salmonella* PT 2022: *Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1.
The MLVA cluster definition for the PT Typing 2022 was set at zero loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 22SCA04 (reference strain) and 22SCA06 (technical duplicate of the reference strain) to be a clustering match with the reference outbreak strain.

3.3.4 *Evaluation of the cluster analysis results based on WGS data*

Data submission for WGS results included:

- **Result form**: background information on the wet-lab and dry-lab methods used, copy/paste of the md5sum output, 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 (Tables 3.4A and 3.4B):
- whether the data passed their quality control (QC) criteria or not;
- whether or not a clustering match was found with the provided reference strain in the EURL-*Salmonella* PT Typing 2022: 22SCA-REF (*Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1).

Apart from the reference cluster, any further clusters could be reported optionally.

Strain 22SCA13 was expected not to pass the QC of the participants, because the data files of this strain also contained numerous *E. coli* reads. The PT Typing 2022 Protocol indicated to exclude strains from the cluster analysis if the data did not pass the QC.

The WGS cluster definition for this particular PT Typing 2022 situation was set at maximum six allelic differences from the reference sequence. Based on this (cgMLST-based) cluster definition, WGS-based results were expected to indicate the 'wet' strains 22SCA04 (reference strain), 22SCA06 (technical duplicate of the reference strain), 22SCA01 (clustering with the reference strain), 22SCA05 (technical duplicate of strain 22SCA01) and the 'dry' strain 22SCA15 (ELt5a data of strain 22SCA01) to be a clustering match with the provided reference outbreak strain 22SCA-REF data (also see Figure 3.1).

4 Results and Discussion

4.1 Technical data

4.1.1 General

A total of 34 laboratories participated in this PT (Chapter 2). These included 27 NRLs-*Salmonella* in the 27 EU Member States and seven NRLs from third countries (EU candidate or potential EU candidate Member States, members of the European Free Trade Association (EFTA), and the United Kingdom).
The frequency of *Salmonella* serotyping at the participating laboratories and the number of strains (approximately) serotyped in 2022 are summarised in Table 4.1.

Table 4.1 Frequency and number of Salmonella *strains serotyped in 2022*

Laboratory code	Serotyping frequency in 2022	No. of strains serotyped in 2022
32	Daily	170
7	Daily	217
15	Daily	330
20	Daily	350
26	Daily	350
31	Daily	500
28	Daily	650
10	Daily	800
24	Daily	900
9	Daily	1100
19	Daily	1800
18	Daily	2268
8	Daily	2500
11	Daily	2500
27	Daily	3000
33	Daily	3000
1	Daily	3200
17	Daily	4000
29	Daily	7500
4	Thrice a week	20
30	Thrice a week	480
3	Thrice a week	700
2	Thrice a week	1800
13	Twice a week	80
5	Twice a week	94
16	Twice a week	110
34	Twice a week	150
23	Twice a week	250
14	Twice a week	383
25	Once a week	100
22	Once a week	510
12	Monthly	10

Laboratory code	Serotyping frequency in 2022	No. of strains serotyped in 2022
21	Monthly	10
6	Monthly	58
n=34		39 890

4.1.2 *Accreditation*

Of the 34 participants, 32 are accredited for serotyping *Salmonella.* Thirty according to EN ISO/IEC 17025, two of them combined with EN ISO 15189. One laboratory mentioned the combination of EN ISO 15189 and EN ISO 16140-6.
The one non-EU laboratory not accredited for serotyping is known for this because of its relatively low numbers of serotyping strains. The one EU NRL currently not accredited for serotyping indicated to plan the re-accreditation for 2023/2024.
All 32 laboratories stated that they are accredited for all *Salmonella* serovars.

4.1.3 *Transport of samples*

All but seven participants received their package within two days after shipment on Monday 7 November 2022. Six packages were received by laboratories on 10 November, and the final one arrived on 11 November 2022. All laboratories received the packages 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. A total of 33 participants used classical serology. Eight of them mentioned the combined use of classical serology and Luminex assays (2), multiplex/real-time PCR (4), or WGS (2). One participant used Whole Genome Sequencing (WGS), supplemented with traditional agglutination using O:6 and O:8 (strain S9) or O:22 and O:23 (strains S7 and S17).
Details on the number and the source of the antisera used by the participants are summarised in Tables 4.2 and 4.3.

Table 4.2 Number of laboratories using antisera from various manufacturers

Manufacturer	Number of NRLs (n=32*)
Bio-Rad	14
Pro-Lab	5
Sifin	19
Statens Serum Institut (SSI)	27
Other	5
Own preparation	3

*Missing data from two participants.

Table 4.3 Number of laboratories using antisera from one or more manufacturers and/or in-house prepared sera

Number of manufacturers from which antisera are obtained (including in-house preparations)	Number of NRLs (n=32*)
1	11
2	6
3	11
4	3
5	1

*Missing data from two participants.

4.2.2 *Biochemical testing*

Twenty participants indicated the use of (a variety of) biochemical tests on all or on a selected number of strains. Laboratories 13 and 15 routinely tested all 21 strains using MALDI-TOF.

4.2.3 *Use of PCR for confirmation*

Fourteen laboratories used PCR to confirm strain S2, the monophasic variant of *S*. Typhimurium 4,5:i:-, and seven of them also used PCR to confirm strain S18, *S*. Typhimurium. Most laboratories mentioned for this the reference 'Tennant et al., 2010'.

4.2.4 *Serotyping results per laboratory*

The percentages of correct results per laboratory are shown in Figure 4.1. The evaluation of the type of errors for O- and H-antigens and the identification of the strains are shown in Figures 4.2, 4.3 and 4.4.
The O-antigens were all typed correctly by 31 of the 34 participants (91%). This corresponds to nearly 100% of the total number of strains. The H-antigens were completely typed correctly by 27 of the 34 participants (79%), corresponding to 98% of the total number of strains. As a result, 25 participants (74%) reported all serovar names correctly, which corresponds to 98% of all strains evaluated.

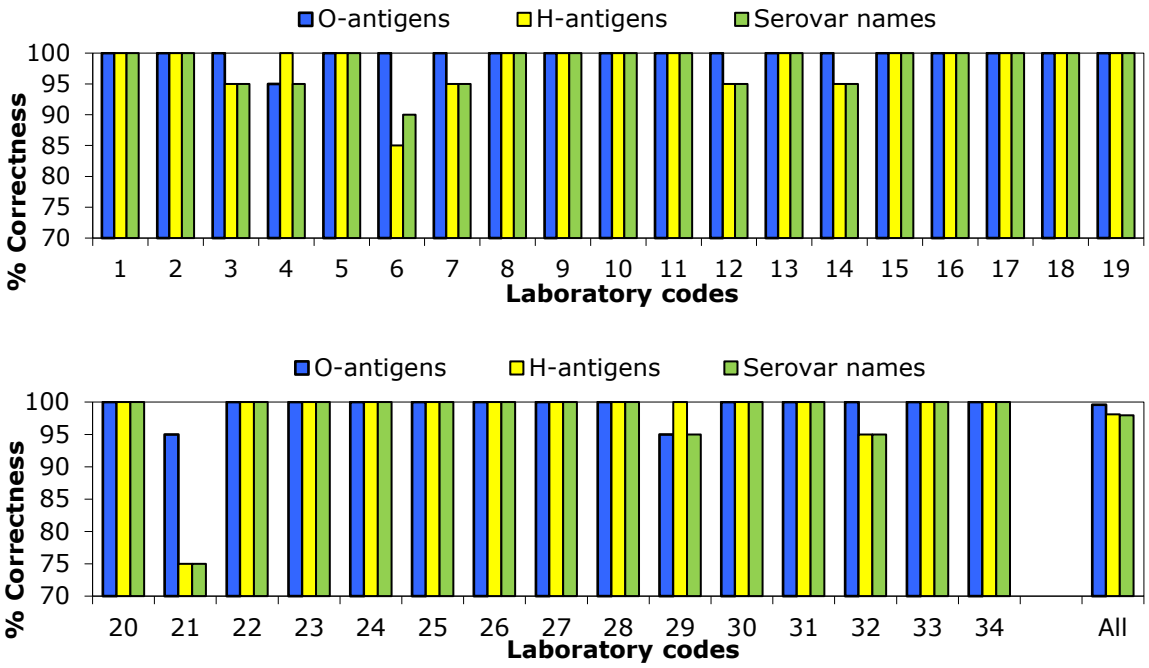


Figure 4.1 Percentages of correct serotyping results, per participant

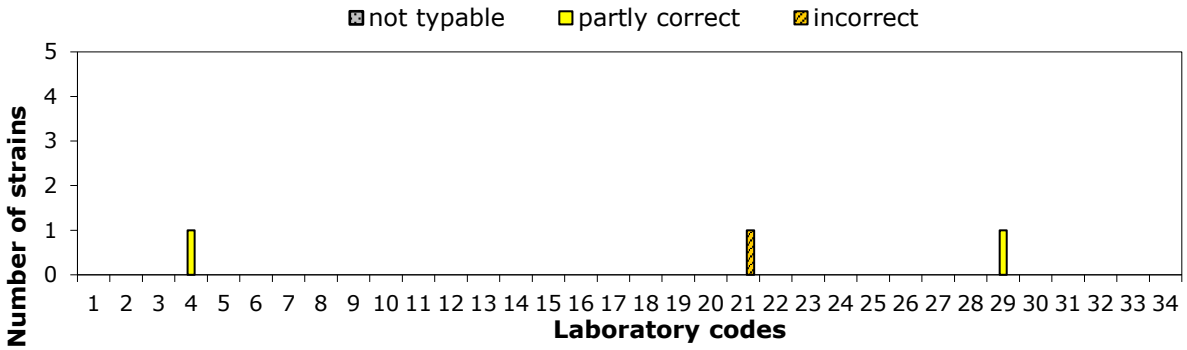


Figure 4.2 Evaluation of type of errors for O-antigens, per participant

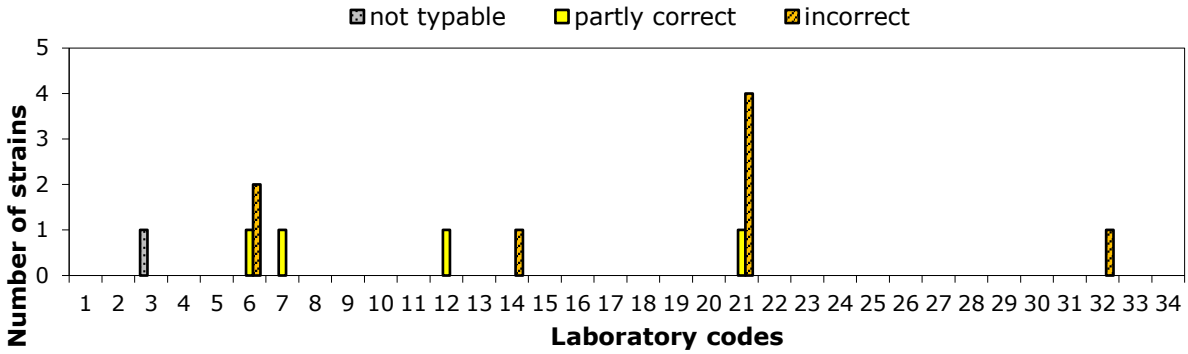


Figure 4.3 Evaluation of type of errors for H-antigens, per participant

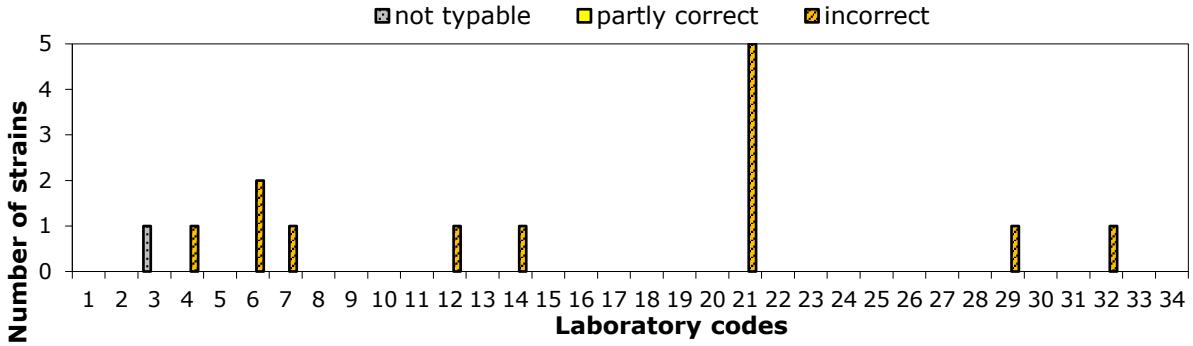


Figure 4.4 Evaluation of the type of errors in the identification of the serovar names, per participant

4.2.5 *Performance of the participants*

The number of penalty points was determined for each NRL using the guidelines described in Section 3.2.2. Table 4.4 shows the number of penalty points for each NRL and indicates whether the level of good performance was achieved (yes or no).

Overall, the participants' performance in the PT Serotyping 2022 was very good. Two EU Member State NRLs (Laboratories 14 and 21) did not meet the level of good performance at the first stage of the PT. Laboratory 21 discussed the deviating results (mistakes in five different strains, also see Annex 2 and 3) with the technicians, but no clear explanation could be identified. A follow-up study for both laboratories was organised in March/April 2023 (see 4.2.7).

All participants received both their individual laboratory evaluation report and the interim summary report on serotyping on 1 March 2023. Annex 1 shows an example of an individual laboratory evaluation report on serotyping. The interim summary report is available on the EURL-*Salmonella* website:
https://www.eurlsalmonella.eu/documenten/interim-summary-report-eurl-salmonella-pt-serotyping-2022

Table 4.4 Evaluation of serotyping results per NRL

Laboratory code	Penalty points	Good performance
1	0	yes
2	0	yes
3	0	yes
4	1	yes
5	0	yes
6	2	yes
7	1	yes
8	0	yes
9	0	yes
10	0	yes
11	0	yes
12	1	yes
13	0	yes
14	4	**NO**

Laboratory code	Penalty points	Good performance
15	0	yes
16	0	yes
17	0	yes
18	0	yes
19	0	yes
20	0	yes
21	5	**NO**
22	0	yes
23	0	yes
24	0	yes
25	0	yes
26	0	yes
27	0	yes
28	0	yes
29	1	yes
30	0	yes
31	0	yes
32	1	yes
33	0	yes
34	0	yes

4.2.6 *Serotyping results per strain*

Annex 2 displays the final naming results reported per strain (S1 – S20) and per laboratory (1-34).
A completely correct identification was obtained for ten *Salmonella* serovars: Singapore (S1), Agona (S4), Enteritidis (S5), Kenya (S6), Hadar (S9), Hull (S10), Virchow (S12), Hato (S13), Mishmarhaemek (S17), and Infantis (S20). Annex 2 also shows the reported serovar names for strain 1,4,[5],12:i:- (S2). Fourteen participants used a PCR method to confirm this strain to be a monophasic *S.* Typhimurium strain.
Strain S8 was characterised with antigenic formula 4,5,12:i:e,n,x, and in accordance with Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor scheme this new variant of the previously described serovar Farsta (4,12:i:e,n,x) is now recognised with the updated antigenic formula: 4,[5],12:i:e,n,x (Issenhuth-Jeanjean et al., 2014).
Most problems occurred with the serovar Kintambo (S7). Four laboratories had difficulties assigning the correct serovar name to this strain, due to problems with completing the designation of the (phase 1) H-antigens. Annex 3 includes all details on the strains that caused problems in serotyping.
Annex 4 describes details on the additional and optional strain S21.
All but five participants tried to serotype strain S21, a *Salmonella enterica* subsp. *diarizonae* (IIIb). A few laboratories did not have access to all required antisera to finalise the serotyping of this strain (47:k:z_{35}).

4.2.7 *Results follow-up study*

Two EU NRLs did not achieve the level of good performance in the first part of the PT (Table 4.4) and participated in a follow-up study. Both NRLs received ten additional strains for serotyping in week 13, 2023. For the follow-up study, the number of penalty points was also determined using the guidelines described in Section 3.2.2. Table 4.5 shows the results of the follow-up study: both participants achieved the level of good performance.

Table 4.5 Evaluation of serotyping results per NRL in the follow-up study

Laboratory code	Penalty points	Good performance
14	0	Yes
21	0	Yes

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: http://www.eurlsalmonella.eu/publications/proficiency-test-reports
The historical data on the EU NRLs-*Salmonella* only are visualised in Figure 4.5, showing the percentages of correctly typed strains. Figure 4.6 shows the number of penalty points and non-good performance.
The percentages of correctly typed strains are stable over time, usually showing better performance for the O-antigens than for the H-antigens (Figure 4.5).
The number of penalty points has clearly declined, from 35 points when this system started in 2007 to three points in the 2020 PT. The rise seen in the 2018 PT was mainly caused by the seven EU NRLs that made a mistake in typing a *S.* Cannstatt strain. The total numbers of penalty points are strongly affected by the system of four penalty points for one mistake in the 'top-5' *Salmonella* serovars, as is seen in the PT 2021 as well as the PT 2022 results (Figure 4.6).
However, the number of EU NRLs with an initial non-good performance is low: two in the period 2010 – 2013 plus in the current PT 2022, one in the 2014, 2015, 2018, and 2021 PTs, and none in the 2016, 2017, 2019 and 2020 PTs. All follow-up studies organised for these EU NRLs, only occasionally the same ones, resulted in a good performance after all.

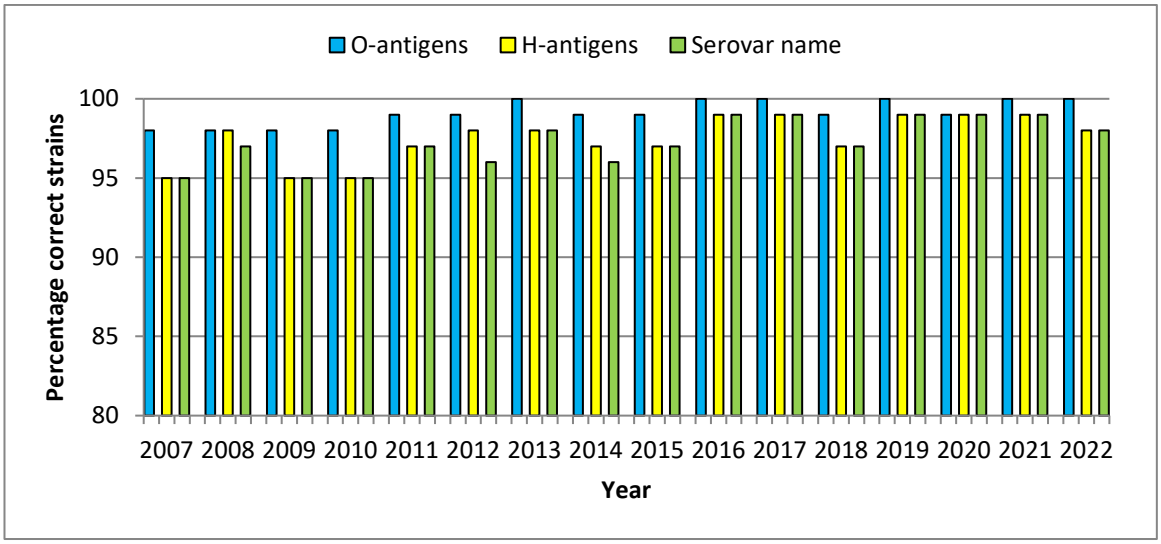


*Figure 4.5 Serotyping results of the EU NRLs-*Salmonella*, based on the percentages of correctly typed strains*

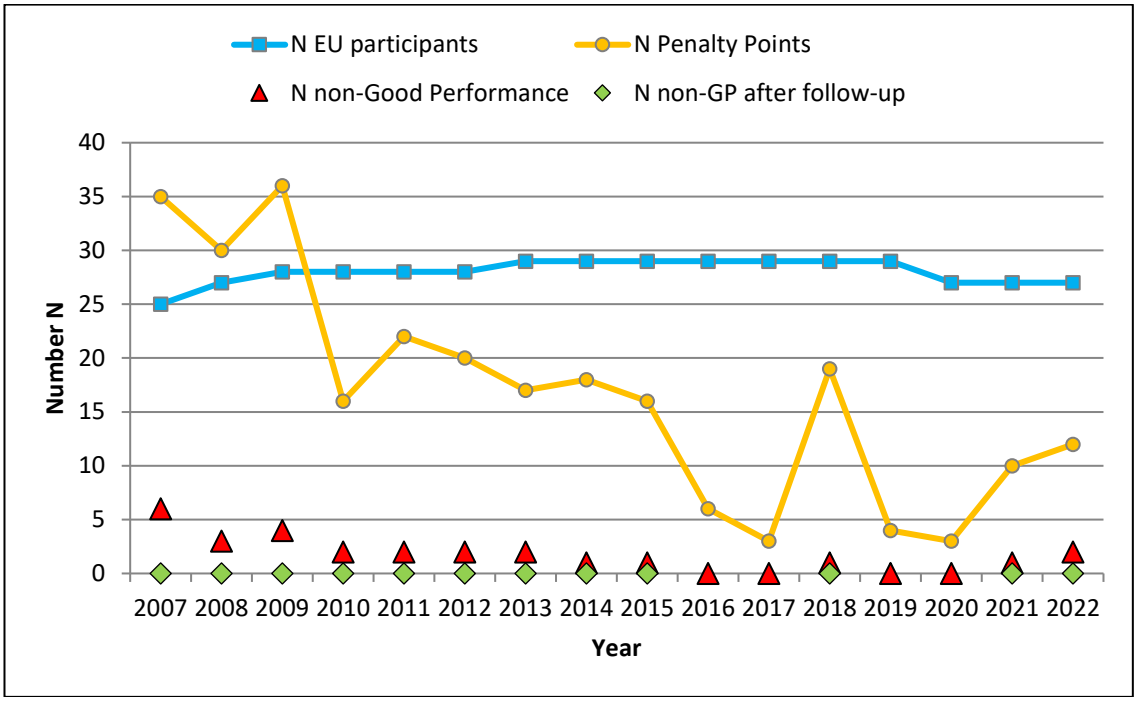


*Figure 4.6 Serotyping results of the EU NRLs-*Salmonella*, based on the number (N) of Penalty Points and non-Good Performance (non-GP)*

4.3 Cluster analysis results

4.3.1 General

Participants could choose to use either MLVA and/or WGS to perform the cluster analysis, using their own routine procedures.
A total of twenty NRLs participated in the cluster analysis. All twenty performed WGS analysis. Five participants additionally performed MLVA analysis.
All participants received their individual laboratory evaluation report on 19 May 2023. Annex 5 gives an example of an individual laboratory evaluation report on cluster analysis results.

Overall results were presented at the online EURL-*Salmonella* Workshop on 22 May 2023 (https://www.eurlsalmonella.eu/workshop-2023).
The interim summary report on overall results is also available on the EURL-*Salmonella* website: https://www.eurlsalmonella.eu/media/3611

As a general question, the participants were asked if and how they serotyped the six 'wet' and the six 'dry' strains. Sixteen participants indicated to have serotyped the 'wet' strains, thirteen also serotyped the 'dry' strains. Annex 6 shows these serotyping results, for information purposes only.

4.3.2 *Results cluster analysis based on MLVA data*

Five participants (Laboratory codes 1, 17, 19, 28, 33) submitted cluster analysis results based on MLVA data.
Annex 7 shows the allelic profiles submitted by the participants.
Laboratory 19 did not report the results in the expected format and therefore these results are regarded as deviating.

Participants were asked to report per strain (Tables 3.4A/B) whether or not a clustering match was found with the reference outbreak strain in the EURL-*Salmonella* PT 2022, being:
Salmonella Enteritidis ST11, MLVA type 3-10-6-3-1.
The MLVA cluster definition for the PT 2022 Typing was set at no loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 22SCA04 (reference strain) and 22SCA06 (technical duplicate of the reference strain) to be a clustering match with the reference outbreak strain.
Four participants reported the MLVA-based cluster analysis results completely as expected (Table 4.6). This included the expected reporting of the technical duplicates 22SCA04/22SCA06 as (part of) one cluster (Table 4.6).
The fifth participant reported the allelic profiles in a deviating format; therefore, the evaluation of their results cannot be done according to the PT Typing 2022 Protocol.

Table 4.6 Expected cluster analysis results and the cluster analysis results reported by the five MLVA participants

	Strain code					
Lab code	**22SCA01**	**22SCA02**	**22SCA03**	**22SCA04**	**22SCA05**	**22SCA06**
Expected	**No**	**No**	**No**	**Yes**	**No**	**Yes**
1	No	No	No	Yes	No	Yes
17	No	No	No	Yes	No	Yes
19*	Yes	Yes	Yes	Yes*	Yes	Yes*
28	No	No	No	Yes	No	Yes
33	No	No	No	Yes	No	Yes

*The allelic profiles were not reported in the expected format, therefore the evaluation of the cluster analysis results cannot be done according to the PT Typing 2022 Protocol.
In blue: Deviation from the expected result.

4.3.3 *Results cluster analysis based on WGS data*

Twenty participants (Table 4.8) submitted a total of 26 cluster analysis results based on WGS data; four participants submitted both cgMLST-based and SNP-based data results, and one participant submitted two cgMLST-based and one SNP-based data analyses.

Annex 8 shows the general details of the wet-lab and dry-lab protocols performed by the participants and the EURL-*Salmonella* (EL). All participants and the EL performed DNA extraction, library preparation and sequencing in-house, except for Lab 8 (library preparation and sequencing outsourced) and Lab 10 (sequencing outsourced). Most participants used the Illumina MiSeq platform(14x), followed by the Illumina NextSeq (5x), and the Illumina MiniSeq (2x) or NovaSeq (1x). Including the EL, 17 submissions were based on cgMLST for data analysis and ten submissions were based on SNP-based analysis (8x reference-based and 2x assembly-based).
Tools used for this analysis varied from in-house pipelines (using the Enterobase scheme) to commercial ones, most often Ridom SeqSphere^{+} (9x). The most commonly used methods for cluster analysis were Minimum Spanning Tree (MST, 12x) and Maximum Likelihood (ML, 9x).

Annex 9 lists all participants' QC criteria reported for evaluating their data. A variety in naming these QC criteria and in the used thresholds was observed, similar to the previous PTs on cluster analysis (Jacobs-Reitsma et al., 2020, Jacobs-Reitsma et al., 2021, Jacobs-Reitsma et al., 2023). Contamination, coverage, GC%, N50, total number of contigs, and total length of assembly were the most commonly referred parameters.

Fourteen compressed paired-end fastq files (strains 22SCA11 – 22SCA16 plus 22SCA-REF) had to be downloaded for analysis from the secure ftp server. The md5 checksums for these files were available on the server as well (Annex 10). Participants were asked whether they checked the md5sum values after downloading, and 16 participants (80%) indicated that they did this.
Participants were also asked to copy/paste in the result form 'your md5 output for all your strains'. Regrettably, this question was not clear for everyone. But 8 of 16 participants did enter the md5 checksums for the sequence files they had generated and that they had to upload to the secure ftp server (for strains 22SCA01 – 22SCA06), which was inclined with this question. After downloading the raw data files from the participants at the EURL-*Salmonella*, this was checked to be correct, indicating that also the transfer of data via the secure ftp site went alright.

All participants' raw data (compressed fastq files) for the six 'wet' strains (22SCA01 – 22SCA06) were successfully processed through the Juno-assembly pipeline as discussed in section 3.3.1. The *de novo* assembled genomes (fasta files) were analysed in Ridom SeqSphere+, using the cgMLST Enterobase v2.0 and visualised in a MST, which also includes the 'dry' strain data (22SCA-REF, 22SCA11 – 22SCA16, but excluding contaminated strain 22SCA13) (Figure 4.7). Annex 11 shows the data per 'wet' strain. Results for Laboratory 26 indicate a swap for strains 22SCA02, 22SCA03, and 22SCA04.

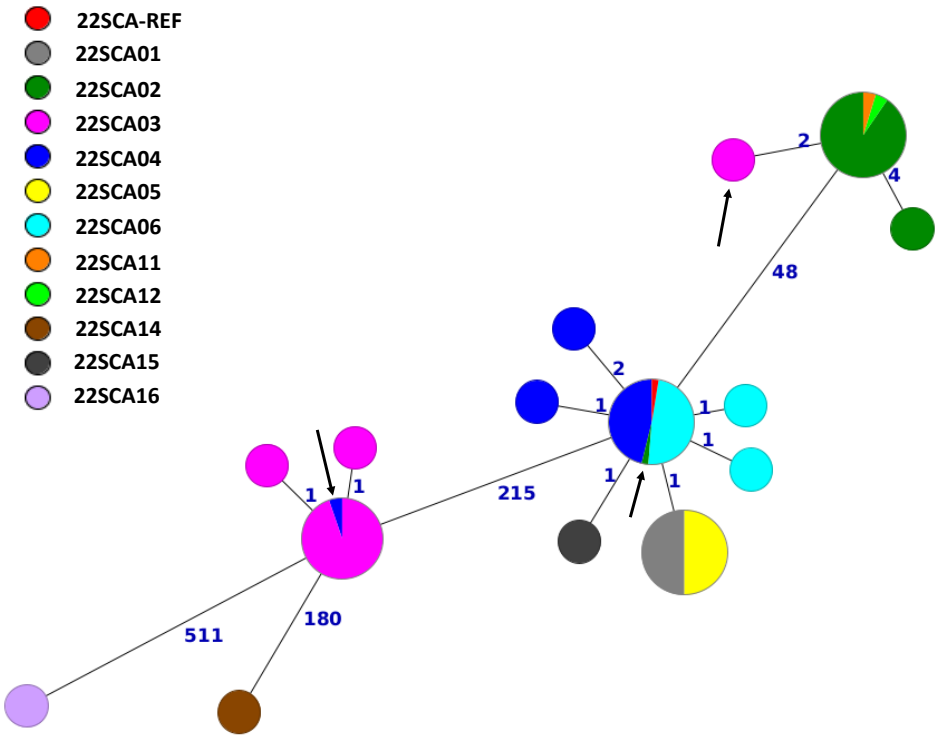


*Figure 4.7 MST of the strains from the participants' processed raw data plus the 'dry' strain data (22SCA-REF, 22SCA11 – 22SCA16, excluding contaminated strain 22SCA13) (Ridom SeqSphere+, cgMLST (3002), pairwise ignoring missing values)**

*Three arrows are indicating the swap of Laboratory 26 for strains 22SCA02, 22SCA03, and 22SCA04.

An overview of the main QC results on all in-house *de novo* assembled genomes (fasta files) is given in Table 4.7. Annex 12 shows detailed data per participant.

Table 4.7 QC results of the in-house de novo *assembled genomes (22SCA01 - 22SCA06), average per participant*

Laboratory code	Illumina Platform	Average contamination	Average # contigs	Average Largest contig	Average Total length	Average N50	Average Coverage
Lab01	MiSeq	0,52	29	995839	4699908	406060	134
Lab02	NextSeq	0,58	72	462377	4704060	153824	110
Lab03	MiSeq	0,52	26	1449562	4701632	452333	176
Lab07	MiSeq	0,53	29	1330154	4702350	453875	121
Lab08	NovaSeq	0,55	31	1321631	4701247	443868	57
Lab09	MiSeq	0,63	24	1542632	4701490	488170	727
Lab10	NextSeq	0,53	30	1489631	4703327	488647	109
Lab14	MiSeq	0,52	38	1166966	4693599	336342	142
Lab16	MiSeq	0,54	26	1624795	4701612	500376	71
Lab17	MiSeq	0,52	26	1542912	4702056	490979	93
Lab19	MiniSeq	0,99	51	1455860	4718226	488772	75
Lab23	NextSeq	0,52	27	1515335	4700311	495667	51
Lab24	MiSeq	0,56	26	1543332	4701450	491848	662
Lab26	MiSeq	1,09	67	785893	4709479	241894	112
Lab27	NextSeq	0,53	77	331746	4698863	145285	98
Lab28	MiSeq	0,52	30	1110971	4700529	380104	97
Lab29	MiSeq	0,52	30	1074276	4701296	371301	25
Lab30	MiSeq	0,52	93	373056	4697609	156370	101
Lab32	MiniSeq	0,52	27	1400734	4701828	452767	69
Lab33	MiSeq	0,52	26	1488435	4700645	523105	80
EL	NextSeq	0,52	52	870862	4701618	214836	136

Participants were asked to report per strain:

- whether the data passed their QC criteria or not;
- whether a clustering match with the reference strain in the EURL-*Salmonella* PT Typing 2022 (22SCA-REF) was found or not.

Apart from the reference cluster, any further clusters could be reported optionally.

Strain 22SCA13 was expected not to pass the QC of the participants, because the data files of this strain also contained numerous *E. coli* reads. The PT Typing 2022 Protocol indicated to exclude strains from the cluster analysis if the data did not pass the QC.
Strain 22SCA13 was reported to be excluded from cluster analysis by 17 of the 20 participants, although it was still included in the distance matrix of one of their submissions. Reasons for (not) excluding strain 22SCA13 are given in Annex 13.
Annex 14 shows per submission the participants' distance matrix data for their comparison of the 22SCA-REF with the final 11 strains (strain 22SCA13 expected to be excluded from the cluster analysis).

The cluster definition for this particular PT Typing 2022 situation was set at maximum six allelic differences from the reference sequence. Based on this (cgMLST-based) cluster definition, WGS-based results were expected to indicate the 'wet' strains 22SCA04 (reference strain), 22SCA06 (technical duplicate of the reference strain), 22SCA01 (clustering with the reference strain), 22SCA05 (technical duplicate of strain 22SCA01) and the 'dry' strain 22SCA15 (ELt5a data of strain 22SCA01) to be a clustering match with the provided reference outbreak strain 22SCA-REF data (also see Figure 3.1).
Nineteen of the 26 submissions (five participants with multiple submissions) reported the WGS-based cluster analysis results completely as expected (Table 4.8).
Technical duplicates 22SCA01 and 22SCA05 were reported within one cluster in all 26 submissions. Technical duplicates 22SCA04 and 22SCA06 were reported within one cluster in all but one of the submissions (Table 4.8).

Some observations on the interpretation of Table 4.8 are given below:
Laboratory 14: there may have been a misunderstanding in the way to report the results of the cluster analysis. Based on the submitted distance matrix, and the analyses shown in Figure 4.7/Annex 11, data are in line with the expected results except for strain 22SCA13 (Annex 13 and 14).
Laboratory 19: this deviation may have been a mistake in filling the result form, because this answer is not supported by the distance matrix that was submitted, nor by Figure 4.7/Annex 11.
Laboratory 26: there may have been a swap of strain numberings: Data of strain 22SCA02 reported as strain 22SCA04; Data of strain 22SCA03 reported as strain 22SCA04; Data of strain 22SCA04 reported as strain 22SCA02. Based on the submitted distance matrix (with the wrong numbering), and the analysis shown in Figure 4.7/Annex 11, data seem to be in line with the expected results.
Laboratory 28 and laboratory 29: strain 22SCA13 was expected to be excluded from the cluster analysis (also see Annex 13 and 14).

Apart from the cluster with the reference strain, a second cluster was optionally to be identified: 22SCA02, 22SCA11, and 22SCA12 (Figure 3.1). A second cluster was reported in 21 of the 26 submissions, three of these were deviating from the expected results. Laboratory 26 reported 22SCA11 and 22SCA12 to be a second cluster, without 22SCA02. Laboratory 29 reported the second cluster correctly, but also considered the four remaining strains as a third cluster, 'although quite divergent'. Laboratory 33 reported the second cluster correctly, but reported strains 22SCA04 and 22SCA14 as a third cluster.

Table 4.8 Expected cluster analysis results and the cluster analysis results reported per data analysis method by the 20 WGS participants

	Strain code											
Lab code - method	**22 SCA01**	**22 SCA02**	**22 SCA03**	**22 SCA04**	**22 SCA05**	**22 SCA06**	**22 SCA11**	**22 SCA12**	**22 SCA13**	**22 SCA14**	**22 SCA15**	**22 SCA16**
Expected	**Yes**	**No**	**No**	**Yes**	**Yes**	**Yes**	**No**	**No**	**n.a.**	**No**	**Yes**	**No**
1-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
2-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
3-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
7-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
8-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
8-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
9-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
10-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
14-cgMLST	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14-SNPr	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
16-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
17-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
19-cgMLST	Yes	No	Yes	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
23-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
24-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
26-SNPa	Yes	Yes	No	No	Yes	Yes	No	No	n.a.	No	Yes	No
27-cgMLST1	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
27-cgMLST2	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
27-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
28-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	No	No	Yes	No
28-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	No	No	Yes	No

	Strain code											
Lab code - method	**22 SCA01**	**22 SCA02**	**22 SCA03**	**22 SCA04**	**22 SCA05**	**22 SCA06**	**22 SCA11**	**22 SCA12**	**22 SCA13**	**22 SCA14**	**22 SCA15**	**22 SCA16**
29-SNPr	Yes	No	No	Yes	Yes	Yes	No	No	No	No	Yes	No
30-SNPa	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
30-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
32-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No
33-cgMLST	Yes	No	No	Yes	Yes	Yes	No	No	n.a.	No	Yes	No

Yes/No: Whether or not a clustering match with the reference strain in the EURL-*Salmonella* PT Typing 2022 (22SCA-REF) was found.
n.a.: Not applicable (QC not passed).
In blue: Deviation from the expected result.

5 Conclusions

5.1 Serotyping

- The overall results for the 34 participants are:
  - They typed nearly 100% of the strains correctly for the O-antigens.
  - They typed 98% of the strains correctly for the H-antigens.
  - They named 98% of the strains correctly.
- Two EU NRLs-*Salmonella* initially did not achieve the defined level of good performance and participated in a follow-up study, typing an additional set of ten strains.
- Ultimately, all 27 EU NRLs and the seven non-EU NRLs achieved the defined level of good performance.

5.2 Cluster analysis

- The 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 included pre-testing the strains by the EURL-*Salmonella*, based on WGS.
  - Six strains were shipped to the participants for 'wet' analysis (MLVA/WGS).
  - Raw WGS data of six strains plus the reference strain were made available through a secure ftp server for 'dry' analysis (WGS only).
- A total of twenty participants performed cluster analysis; five using MLVA analysis and all twenty using WGS analysis.
- Four of the five participants reported the MLVA-based cluster analysis results fully as expected.
- Nineteen of the 26 submissions (fifteen of the 20 participants) reported the WGS-based cluster analysis results fully as expected.
- Three participants did not exclude strain 22SCA13 from their cluster analysis, although this strain was expected not to pass the participants' QC because it also contained numerous *E. coli* reads.

Acknowledgements

The authors would like to thank the IDS-bioinformatics team (RIVM) for their valuable help with the Juno-assembly pipeline, and in particular Maaike van den Beld for kindly providing strains for the cluster analysis. Also, the technical assistance by Wendy van Overbeek (RIVM) in the preparation of all sample materials is highly appreciated.

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 26-7-2023).

ECDC (European Centre for Disease Prevention and Control), 2016. Laboratory standard operating procedure for multiple-locus variable-number tandem repeat analysis of *Salmonella enterica* serotype Enteritidis. doi 10.2900/973540 https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Salmonella-Enteritidis-Laboratory-standard-operating-procedure.pdf (accessed 26-7-2023).

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

EN ISO 16140-6. Microbiology of the food chain — Method validation — Part 6: Protocol for the validation of alternative (proprietary) methods for microbiological confirmation and typing procedures. 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.

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 26-7-2023).

Guibourdenche, M., et al., 2010. Supplement 2003–2007 (no. 47) to the White–Kauffmann–Le Minor scheme. Research in Microbiology, 2010, 161, pp. 26–29. https://doi.org/10.1016/j.resmic.2009.10.002 (accessed 28-08-2023)

Issenhuth-Jeanjean, S., et al., 2014. Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor scheme. Research in Microbiology 2014 (165): 526-530. https://doi.org/10.1016/j.resmic.2014.07.004 (accessed 28-08-2023)

Jacobs-Reitsma, W.F., A. Verbruggen, R.E. Diddens, A.H.A.M. van Hoek, and K.A. Mooijman, 2020. EURL-*Salmonella* Proficiency Test Typing 2019. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report 2020-0084. https://www.rivm.nl/bibliotheek/rapporten/2020-0084.pdf (accessed 26-7-2023).

Jacobs-Reitsma, W.F., A. Verbruggen, R.E. Diddens, A.H.A.M. van Hoek, and K.A. Mooijman, 2021. EURL-*Salmonella* Proficiency Test Typing 2020. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report 2021-0126. https://www.rivm.nl/bibliotheek/rapporten/2021-0126.pdf (accessed 26-7-2023).

Jacobs-Reitsma, W.F., A. Verbruggen, R.E. Diddens, A.H.A.M. van Hoek, and K.A. Mooijman, 2023. EURL-*Salmonella* Proficiency Test Typing 2021. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report 2022-0105. https://www.rivm.nl/bibliotheek/rapporten/2022-0105.pdf (accessed 26-7-2023).

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. https://www.rivm.nl/bibliotheek/rapporten/330604006.pdf (accessed 26-7-2023).

Mooijman, K.A., 2022. The 27th EURL-*Salmonella* workshop; 23 and 24 May 2022, Online. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report no.: 2022-0107. https://www.rivm.nl/bibliotheek/rapporten/2022-0107.pdf (accessed 26-7-2023).

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.

List of abbreviations

BPW	Buffered Peptone Water
CA	Cluster Analysis
cgMLST	core genome Multilocus Sequence Typing
DG-SANTE	Directorate General for Health and Food Safety
EC	European Commission
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
HI agar	Hearth Infusion agar (in transport tubes)
ISO	International Organization for Standardization
MALDI-TOF	Matrix Assisted Laser Desorption Ionization Time-Of-Flight mass spectrometry
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
PT	Proficiency Test
QC	Quality Control
REF	Reference
RIVM	National Institute for Public Health and the Environment (Bilthoven, The Netherlands)
SNP	Single Nucleotide Polymorphism
SNPa	assembly-based SNP analysis
SNPr	reference-based SNP analysis
SSI	Statens Serum Institut (Copenhagen, Denmark)
ST	Sequence Type
wgMLST	whole genome Multilocus Sequence Typing
WGS	Whole Genome Sequencing

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

Results

EURL-*Salmonella* PT Serotyping 2022

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

	Reference Results				**Results NRL lab code:**			**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	6,7	k	e,n,x	Singapore	6,7	k	e,n,x	Singapore
S2[a)]	1,4,[5],12	i	-	1,4,[5],12:i:-	4	i	-	4:i:-
S3	4,12	z10	1,6	Tudu	4	z10	6	Tudu
S4	1,4,[5],12	f,g,s	[1,2]	Agona	4	f,g,s	-	Agona
S5	1,9,12	g,m	-	Enteritidis	9,12	g,m	-	Enteritidis
S6	6,7	l,z13	e,n,x	Kenya	6,7	z13	e,n,x	Kenya
S7	1,13,23	m,t	-	Kintambo	13,23	m,t	-	Kintambo
S8[b)]	4,[5],12	i	e,n,x	Farsta	4,5	i	e,n,x	Farsta
S9	6,8	z10	e,n,x	Hadar	6,8	z10	e,n,x	Hadar
S10	16	b	1,2	Hull	16	b	2	Hull
S11	11	d	[e,n,x]	Chandans	11	d	e,n,x	Chandas
S12	6,7,14	r	1,2	Virchow	6,7	r	2	Virchow
S13	1,4,[5],12	g,m,s	[1,2]	Hato	4	g,m,s	-	Hato
S14	1,9,12	e,h	1,5	Eastbourne	9,12	h	5	Eastbourne
S15	3,10	z35	z6	Cairina	10	z35	z6	Cairina
S16	1,6,14,25	a	1,5	Garba	6,14,25	a	5	Garba
S17	1,13,23	d	1,5	Mishmarhaemek	13,23	d	5	Mishmarhaemek
S18	1,4,[5],12	i	1,2	Typhimurium	4,5	i	2	Typhimurium
S19	28	c	1,5	Hermannswerder	28	c	5	Hermannswerder
S20	6,7,14	r	1,5	Infantis	6,7	r	5	Infantis
S21[c)]	47	k	z35	47:k:z35 (IIIb)	47	k	z35	S. IIIb (Salmonella enterica subsp. diarizonae) 47:k:z35

[a)] monophasic variant of *S.* Typhimurium based on genomic sequences.
[b)] in accordance with Supplement 2008-2010 (no. 48) to the White-Kauffmann-Le Minor scheme.
[c)] *Salmonella enterica* subspecies *diarizonae.*

Results

EURL-*Salmonella* PT Serotyping 2022

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 27th EURL-*Salmonella* Workshop (23 May 2022, online), Strain S21 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 Singapore	S2 1,4,[5],12:i:-	S3 Tudu	S4 Agona	S5 Enteritidis	S6 Kenya	S7 Kintambo	S8 Farsta	S9 Hadar	S10 Hull
1	Singapore	4:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
2	Singapore	I 4:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
3	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
4	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
5	Singapore	1,4,5,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
6	Singapore	Typhimurium monofaza	Tudu	Agona	Enteritidis	Kenya	enterica II	Farsta	Hadar	Hull
7	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Agbeni	Farsta	Hadar	Hull
8	Singapore	1,4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
9	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
10	Singapore	monophasic Typhimurium	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
11	Singapore	4,12:i:- (mST)	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
12	Singapore	4,5,12:i:-	Tudu	Agona	Enteritidis	Kenya	Agbeni	Farsta	Hadar	Hull
13	Singapore	monophasic Typhimurium	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
14	Singapore	Tumodi	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
15	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
16	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
17	Singapore	Monophasic Typhimurium 4:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
18	Singapore	4,5,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
19	Singapore	Sub I 4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
20	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
21	Singapore	Monophasic Salmonella Typhimurium	Lexington	Agona	Enteritidis	Kenya	Agbeni	Chester	Hadar	Hall
22	Singapore	4,5,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
23	Singapore	4,12:i: -	Tudu	Agona	Enteritidis	Kenya	Kintambo	Fasta	Hadar	Hull
24	Singapore	4:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	4,5:i:e,n,x	Hadar	Hull
25	Singapore	4:i:- (monophasic Typhimurium)	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
26	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
27	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
28	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
29	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
30	Singapore	Typhimurium, monophasic (4,12:i:-)	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
31	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
32	Singapore	4,12:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
33	Singapore	4,5:i:-	Tudu	Agona	Enteritidis	Kenya	Kintambo	Farsta	Hadar	Hull
34	Singapore	4,[5],12:i:- (monophasic Typhimurium)	Tudu	Agona	Enteritidis	Kenya	Kimtambo	Farsta	Hadar	Hull
X	0	1	1	0	0	0	4	1	0	0

S11 Chandans	S12 Virchow	S13 Hato	S14 Eastbourne	S15 Cairina	S16 Garba	S17 Mishmarhaemek	S18 Typhimurium	S19 Hermannswerder	S20 Infantis	Lab: REF
Chandas	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	1
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	2
Chandans	Virchow	Hato	Eastbourne	3,10:-:z6	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	3
Chandans	Virchow	Hato	Waedenswil	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	4
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	5
Chandans	Virchov	Hato	Eastbourne	enterica II	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	6
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	7
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	8
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	9
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	10
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	11
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	12
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	13
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	14
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	15
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	16
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	17
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	18
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	19
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	20
Findorff	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Vanier	Infantis	21
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	22
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	23
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	24
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	25
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	26
Chandans	Vichow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	27
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	28
Chandans	Virchow	Hato	Eastbourne	Cairina	Sanjuan	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	29
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	30
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	31
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	4,5,12:i:-	Hermannswerder	Infantis	32
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	33
Chandans	Virchow	Hato	Eastbourne	Cairina	Garba	Mishmarhaemek	Typhimurium	Hermannswerder	Infantis	34
1	0	0	1	1	1	0	1	1	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.

NOTE: The Table reflects the raw data submitted by the participants. However, the electronic result form does not allow input in special fonts like ‘italic’ or ‘subscript’.

Results for strain S21 are given in Annex 4.

Annex 3 Details per strain that caused problems in serotyping

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-2	**1,4,[5],12**	**i**	**-**	**1,4,[5],12:i:-**	**REF**
S-2	1,4,12	i	z6	Tumodi	14
S-3	**4,12**	**z10**	**1,6**	**Tudu**	**REF**
S-3	3	z10	1,5	Lexington	21
S-7	**1,13,23**	**m,t**	**-**	**Kintambo**	**REF**
S-7	1,13,23	g,m,t	1,5	enterica II	6
S-7	13,23	g,m,t	-	Agbeni	7
S-7	13,23	g,m	-	Agbeni	12
S-7	13,23	g,m,t	-	Agbeni	21
S-7	13,23	m,t	-	Kimtambo	34
S-8	**4,[5],12**	**i**	**e,n,x**	**Farsta**	**REF**
S-8	4,5,12	e,h	e,n,x	Chester	21
S-8	4,5,12	i	e,n,x	Fasta	23
S-8	4,5	i	e,n,x	4,5:i:e,n,x	24
S-10	**16**	**b**	**1,2**	**Hull**	**REF**
S-10	16	b	1,2	Hall	21
S-11	**11**	**d**	**[e,n,x]**	**Chandans**	**REF**
S-11	11	d	e,n,x	Chandas	1
S-11	11	d	z6	Findorff	21
S-12	**6,7,14**	**r**	**1,2**	**Virchow**	**REF**
S-12	6,7	r	1,2	Virchov	6
S-12	6,7	r	1,2	Vichow	27
S-14	**1,9,12**	**e,h**	**1,5**	**Eastbourne**	**REF**
S-14	9,46	e,h	1,5	Waedenswil	4
S-15	**3,10**	**z35**	**z6**	**Cairina**	**REF**
S-15	3,10	-	z6	3,10:-:z6	3
S-15	3,10	z35	e,n,x,z15	enterica II	6
S-15	3,1	z35	z6	Cairina	13
S-16	**1,6,14,25**	**a**	**1,5**	**Garba**	**REF**
S-16	6,7,14	a	1,5	Sanjuan	29
S-18	**1,4,[5],12**	**i**	**1,2**	**Typhimurium**	**REF**
S-18	4,5,12	i	-	4,5,12:i:-	32
S-19	**28**	**c**	**1,5**	**Hermannswerder**	**REF**
S-19	28	z	1,5	Vanier	21

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

Annex 4 Details of serotyping results for strain S21

Strain code	O-antigens	H-antigens		Serovar	Lab code
		(phase 1)	(phase 2)		
S-21	**47**	**k**	**z35**	**47:k:z35**	**REF**
S-21	47	k	z35	S. IIIb (Salmonella enterica subsp. diarizonae) 47:k:z35	1
S-21	47	k	z35	IIIb 47:k:z35	2
S-21	47	k	-	47:k:-	3
S-21	47	k	z35	47:k:z35	4
S-21	47	k	z35	Salmonella enterica subspecies diarizonae 47:k:z35 (IIIb)	5
S-21	-	-	-	-	6
S-21	47	k	-	enterica subsp diarizonae	7
S-21	47	k	z35	IIIb 47:k:z35	8
S-21	47	k	z35	47:5:z35	9
S-21		k	z35	S.enterica subsp. diarizonae IIIb	10
S-21	47	k	z35	47:k:z35 (IIIb)	11
S-21					12
S-21					13
S-21	47	k	z35	III b	14
S-21	47	k	z35	IIIb 47:k:z35	15
S-21					16
S-21	47	k	z35	IIIb:47:k:z35	17
S-21	47	k	z35	47:k:z35	18
S-21	47	k	z35	Sub IIIb 47:k:z35 (diarizonae)	19
S-21	47	k	z35	47:k:z35 (IIIb)	20
S-21	47	k	z35	Lyon III b	21
S-21	47	k	z35	IIIb 47:k:z35	22
S-21	47	k	z35	47:k:z35	23
S-21	OME+	k	z35	OME+:k:z35	24
S-21	-	-	-	-	25
S-21	47	k	z35	47:k:z35	26
S-21	47	k	z35	47:k:z35	27
S-21	47	k	z35	47:k:z35 (IIIb)	28
S-21	47	k	z35	47:z:z35 sg IIIb	29
S-21	47	k	z35	Salmonella enterica subsp. diarizonae serovar 47:k:z35	30
S-21	47	k	z35	47:k:z35 (IIIb)	31
S-21	47	k	z35	47:k:z35	32
S-21	47	k	z35	S.enterica subsp.diarizonae (Group O:X)	33
S-21	47	k	z35	IIIb	34

Reference strain
remark (e.g. spelling error)
not typable (e.g. antisera not available, rough strain)

NOTE: The Table reflects the raw data submitted by the participants. However, the electronic result form does not allow input in special fonts like 'italic' or 'subscript'.

Annex 5 Example of an individual laboratory evaluation report on cluster analysis results

Evaluation

EURL-Salmonella PT Cluster Analysis 2022

Laboratory code: 01

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 2022).

As a minimum, it was expected to have any technical duplicate strains reported as (part of) one cluster.
No specific performance criteria were set for this 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 2022 (www.eurlsalmonella.eu/publications/proficiency-test-reports).

Did you serotype the 6 'wet' strains that were shipped to the participants: Yes
Methodology used: Classical serology/SeqSero

Strain:	**22SCA01[a)]**	**22SCA02[c)]**	**22SCA03**	**22SCA04[b)] (REF)**	**22SCA05[a)]**	**22SCA06[b)]**
Expected results:	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
Reported results:	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis

a) Technical duplicates b) Technical duplicates c) Biological duplicate strain 2021SCA08

Did you serotype the 6 'dry' strains (bioinformatic data only): Yes
Bioinformatic tool(s) used: SeqSero

Strain:	**22SCA11[c)]**	**22SCA12 [c)]**	**22SCA13***	**22SCA14**	**22SCA15 [d)]**	**22SCA16**
Expected results:	Enteritidis	Enteritidis	Enteritidis*	Enteritidis	Enteritidis	Enteritidis
Reported results:	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis

c) Strain 2021SCA08, raw data from 2 different participants PT 2021 d) Biological duplicate strain 22SCA01 * *S.* Enteritidis was contaminated with *E. coli*

Submission of **MLVA** results: Yes

The allelic profile was asked to be reported in the format SENTR7-SENTR5-SENTR6-SENTR4-SE-3.

Strain:	**22SCA01[a)]**	**22SCA02 [c)]**	**22SCA03**	**22SCA04[b)] (REF)**	**22SCA05[a)]**	**22SCA06[b)]**
Expected results:	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1
Reported results:	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1

a) Technical duplicates b) Technical duplicates c) Biological duplicate strain 2021SCA08

MLVA-based cluster identification in the PT Typing 2022 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 2022:

***Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1**

The cluster definition for MLVA was set at zero loci with a different number of repeats.

Strain:	**22SCA01[a)]**	**22SCA02 [c)]**	**22SCA03**	**22SCA04[b)] (REF)**	**22SCA05[a)]**	**22SCA06[b)]**
Expected results:	No	No	No	Yes	No	Yes
Reported results:	No	No	No	Yes	No	Yes

a) Technical duplicates b) Technical duplicates c) Biological duplicate strain 2021SCA08

MLVA-based cluster identification as expected: Yes
Technical duplicates 22SCA04 (REF) and 22SCA06 reported within one cluster: Yes

Submission of **WGS** results:	Yes
WGS platform used:	Illumina MiSeq
Analysis used for WGS data:	cgMLST-based
Tool used for analysis:	Ridom SeqSphere
Method used for cluster analysis:	Minimum Spanning Tree (MST)
Did you check the md5sum values for all 14 compressed fastq files that you downloaded from the sftp server?	Yes

Strain 22SCA13 was expected to be reported for not passing the Quality Control (QC), this *S*. Enteritidis strain was contaminated with *E. coli*.
Strains not passing the QC had to be excluded from the cluster analysis (Protocol EURL-*Salmonella* PT Typing 2022).

Strains reported for not passing the Quality Control (QC):	22SCA13
Reason(s) for not passing the QC:	Contamination Check Result: Potential contamination by second species above 10% detected: *Escherichia coli*; genome size too big (12.7 MB); strain excluded for MST and matrix
Excluded from the cluster analysis (distance matrix):	Yes

WGS-based cluster identification in the PT Typing 2022 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 2022:
22SCA_REF_R1.fq.gz and **22SCA_REF_R2.fq.gz** (*Salmonella* Enteritidis ST11, MLVA type 3-10-6-3-1).
The cgMLST-based cluster definition was set at maximum 6 allelic differences from the reference sequence.

Strain:	**22SCA01[a)]**	**22SCA02[c)]**	**22SCA03**	**22SCA04[b)]**	**22SCA05[a)]**	**22SCA06[b)]**
Expected results:	Yes	No	No	Yes	Yes	Yes
Reported results:	Yes	No	No	Yes	Yes	Yes

Strain:	**22SCA11[c)]**	**22SCA12[c)]**	**22SCA13***	**22SCA14**	**22SCA15[d)]**	**22SCA16**
Expected results:	No	No	n.a.	No	Yes	No
Reported results:	No	No	n.a.	No	Yes	No

a) Technical duplicates
b) Technical duplicates (REF)
c) 22SCA02: Biological duplicate strain 2021SCA08; 22SCA11 and 22SCA12: Strain 2021SCA08, raw data from 2 different Laboratories
d) Biological duplicate strain 22SCA01
* *S.* Enteritidis strain contaminated with *E. coli*
n.a.: not applicable (QC not passed)

WGS-based cluster identification with the reference strain as expected:	Yes
Technical duplicates 22SCA04 (REF) and 22SCA06 reported within one cluster:	Yes
Technical duplicates 22SCA01 and 22SCA05 reported within one cluster:	Yes
Apart from the cluster with the reference strain, a second cluster was optionally to be identified:	22SCA02, 22SCA11, and 22SCA12
Optionally, any further cluster(s) reported:	22SCA02, 22SCA11, 22SCA12
Second cluster identified as expected:	Yes

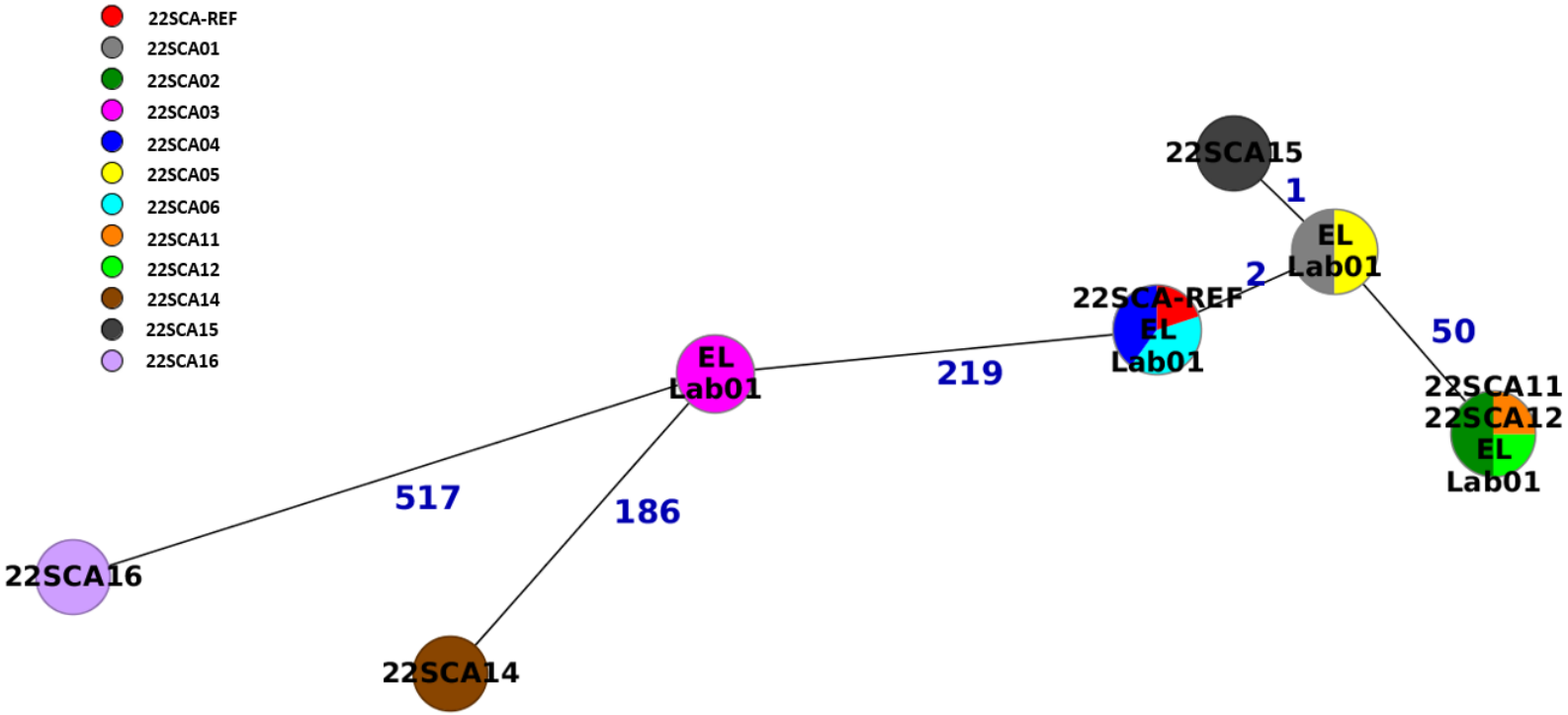


Figure 1 Minimum Spanning Tree of the participant's results, the EL results and the 'dry' strains (Ridom SeqSphere+, cgMLST (3002), pairwise ignoring missing values).

Annex 6 Serotyping results cluster analysis part

Table A6.1 Reported serotyping results 'wet' strains 22SCA01 – 22SCA06

Lab code	Serotyping method(s) used	22SCA01	22SCA02	22SCA03	22SCA04	22SCA05	22SCA06
REF	Luminex/In-house Juno pipeline (SeqSero2)	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
1	Classical serology/SeqSero	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis
2	WGS Illumina short reads, in-house pipeline, typing based on SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
3	Bionumerics Seqsero	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
8/78	SeqSero2+SISTR	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
9	Sistr	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
10	https://cge.cbs.dtu.dk/services/SeqSero/	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
14/74	WGS	S.enteritidis, O-9:g,m	S.enteritidis, O-9:g,m	S.enteritidis, O-9:g,m	S.enteritidis, O-9:g,m	S.enteritidis, O-9:g,m	S.enteritidis, O-9:g,m
16	SeqSero 1.2 (https://cge.food.dtu.dk/services/SeqSero/)	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-
23	SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
24	Classical serology	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
26	WGS-SeqSero 1.2 Server	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-
27/67/77	Classical serology	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis
28/68	Sistr	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
29	Classical serology	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis
30/70	WGS, SISTR Pipeline in Irida platform (VIGAS-P)	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
33	Classical serology/SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis

Table A6.2 Reported serotyping results 'dry' strains 22SCA11 – 22SCA16

Lab code	Serotyping method(s) used	22SCA11	22SCA12	22SCA13*	22SCA14	22SCA15	22SCA16
REF	Luminex/In-house Juno pipeline (SeqSero2)	Enteritidis	Enteritidis	n.a.	Enteritidis	Enteritidis	Enteritidis
1	SeqSero	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis
2	WGS Illumina short reads, in-house pipeline, typing based on SeqSero2	Enteritidis	Enteritidis	WGS quality insufficient	Enteritidis	Enteritidis	Enteritidis
3	Bionumerics Seqsero	Enteritidis	Enteritidis	Contamination detected	Enteritidis	Enteritidis	Enteritidis
8/78	SeqSero2+SISTR	Enteritidis	Enteritidis	NA	Enteritidis	Enteritidis	Enteritidis
9	Sistr	Enteritidis	Enteritidis		Enteritidis	Enteritidis	Enteritidis
14/74	Sciensano Galaxy, SeqSero2 1.2.1	Enteritidis[1)]	Enteritidis[1)]	Enteritidis[1)]	Enteritidis[1)]	Enteritidis[1)]	Enteritidis[1)]
16	SeqSero 1.2 (https://cge.food.dtu.dk/services/SeqSero/)	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-
23	SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
26	SeqSero 1.2 Server	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-
27/67/77	in-house Pipeline (Bakcharak Version 3.0.3) implemented tool SISTR	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis	S.Enteritidis
28/68	Sistr	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
30/70	SISTR (in IRIDA platform, VIGAS-P)	Enteritidis	Enteritidis	Not pass quality control	Enteritidis	Enteritidis	Enteritidis
33	SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis

* The data files of this *S.* Enteritidis strain also contained numerous *E. coli* reads (not passing QC).

[1)] original result entered as: *Salmonella enterica* subsp. *enterica* serovar Enteritidis.

In blue: Deviation from the expected result. Strain 22SCA13 was expected not to be included in the cluster analysis, due to not passing QC. This would also be applicable to serotyping the data of this strain.

Annex 7 Expected and reported MLVA results for all five participants, cluster analysis part

Loci were asked to be reported in the order: SENTR7-SENTR5-SENTR6-SENTR4-SE-3.

	Strain code					
Lab code	**22SCA01**	**22SCA02**	**22SCA03**	**22SCA04**	**22SCA05**	**22SCA06**
Expected	**3-10-5-3-1**	**3-10-4-4-1**	**2-9-7-4-2**	**3-10-6-3-1**	**3-10-5-3-1**	**3-10-6-3-1**
1	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1
17	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1
19	5-2-3-7-6-6-11	5-2-3-7-6-6-11	5-2-3-7-6-6-11	5-2-3-7-6-6-11	5-2-3-7-6-6-11	5-2-3-7-6-6-11
28	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1
33	3-10-5-3-1	3-10-4-4-1	2-9-7-4-2	3-10-6-3-1	3-10-5-3-1	3-10-6-3-1

In blue: Deviation from the expected result.

Annex 8 WGS results cluster analysis part, methods used by the participants

Lab code	Wet lab [a)]	WGS platform [b)]	Data analysis	Tool for analysis	Method for cluster analysis [c)]
1	In-In-In	MiSeq	cgMLST-based	Ridom SeqSphere	MST
2	In-In-In	NextSeq	cgMLST-based	Ridom SeqSphere	MST
3	In-In-In	MiSeq	cgMLST-based	BioNumerics	MST
7	In-In-In	MiSeq	cgMLST-based	linux command line	NJ
8-cgMLST	In-Out-Out	NovaSeq	cgMLST-based	Ridom SeqSphere	MST
8-SNPr	In-Out-Out	NovaSeq	SNP-based - reference-based	SNIPPY	ML
9	In-In-In	MiSeq	SNP-based - reference-based	Python script	MST
10	In-In-Out	NextSeq	SNP-based - reference-based	https://cge.food.dtu.dk/services/CSIPhylogeny/	ML
14-cgMLST	In-In-In	MiSeq	cgMLST-based	galaxy.sciensano	ML
14-SNPr	In-In-In	MiSeq	SNP-based - reference-based	Galaxy Sciensano	ML
16	In-In-In	MiSeq	SNP-based - reference-based	CSIPhylogeny (https://cge.food.dtu.dk/services/CSIPhylogeny/)	ML
17	In-In-In	MiSeq	cgMLST-based	in-house galaxy	MSTreeV2
19	In-In-In	Miniseq	cgMLST-based	Ridom SeqSphere	MST
23	In-In-In	NextSeq 2000	cgMLST-based	Ridom SeqSphere	MST
24	In-In-In	MiSeq	cgMLST-based	Ridom SeqSphere	MST
26	In-In-In	MiSeq	SNP-based - assembly-based	CSIPhylogeny 1.4	ML
27-cgMLST1	In-In-In	NextSeq	cgMLST-based	inhouse chewieSnake pipeline (Enterobase scheme)	single linkage hierarchical clustering
27-cgMLST2	In-In-In	NextSeq	cgMLST-based	Ridom SeqSphere+Enterobase scheme	single linkage hierarchical clustering

Lab code	Wet lab [a)]	WGS platform [b)]	Data analysis	Tool for analysis	Method for cluster analysis [c)]
27-SNPr	In-In-In	NextSeq	SNP-based - reference-based	SNP-analysis using SnippySnake pipeline	single linkage hierarchical clustering
28-cgMLST	In-In-In	MiSeq	cgMLST-based	PyMLST v1	MST
28-SNPr	In-In-In	MiSeq	SNP-based - reference-based	BWA, bcftools, RAxML	ML
29	In-In-In	MiSeq	SNP-based - reference-based	Snippy, Snapper DB, Gubbins, RAxML, iToL	ML and SNP address analysis
30-cgMLST	In-In-In	MiSeq	cgMLST-based	chewBBACA using the scheme from Enterobase	Calculated AD based on output chewBBACA
30-SNPa	In-In-In	MiSeq	SNP-based - assembly-based	In house pipeline[d)]	ML
32	In-In-In	MiniSeq	cgMLST-based	Ridom SeqSphere	Distance matrix only
33	In-In-In	MiSeq	cgMLST-based	ChewBBaCa	MST
EL	In-In-In	NextSeq	cgMLST-based	Ridom SeqSphere	MST

a) Wet lab preparations: DNA extraction, Library preparation, sequencing. IN: In-house, Out: Outsourced.
b) All Illumina platforms.
c) ML: Maximum Likelihood, MST: Minimum Spanning Tree, NJ: Neighbor Joining.
d) Based on parSNP, Gubbins, creating a ML tree in IQTree, creating a SNP distance matrix with snp-dists.

Annex 9 WGS results cluster analysis part, QC criteria as listed by the participants

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
1	Contamination	Mash (integrated in SeqSphere+)	second species above 10%
1	Contig size	In House Pipeline (via SeqSphere+)	200 bases (contigs shorter than 200 bases were ignored)
1	Coverage	In House Pipeline (via SeqSphere+)	50x
1	Genome size (~5 MB for Salmonella)	In House Pipeline (via SeqSphere+)	4.6-5.3 MB
1	Percentage of good cgMLST targets	In House Pipeline (via SeqSphere+)	~98% minimum
2	average coverage	bbtools	>30
2	Completeness	CheckM	>96
2	contamination	CheckM	<4
2	GC%	QUAST	51.6 - 52.3
2	N50	QUAST	>30000
2	phred score	FastQC	>30
2	Total length of assembly	QUAST	4.4 - 5.8 Mb
2	Total number of contigs	QUAST	<300
3	Contamination	Kmer finder	N/A
3	Core Percent	Bionumerics	Greater than/equal to 95
3	Coverage	Bionumerics	Greater than/equal to 30X
3	N50	Bionumerics	Greater than/equal to 15000
3	Total length of assembly	Bionumerics	4500000-5400000 bp
3	Total number of contigs	Bionumerics	Less than/equal to 400
7	Contamination	Confindr	

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
7	Coverage	Qualimap	>300
7	Insert size median	Qualimap	340-400
7	N50	Quast	>15000
7	Total length of assembly	Quast	4,7-5,3*10E6
7	Total number of contigs	Quast	<300
8/78	Contamination	Confindr	5% - intra contaminations
8/78	Contamination	Kraken2	appreciation - inter contaminations
8/78	Coverage	BBmap	50X
8/78	GC%	Quast	
8/78	N50	Quast	
8/78	Q30	Fastp	80%
8/78	Recombinations	Gubbins	NA
8/78	Total lengh of assembly	Quast	
8/78	Total number of contigs	Quast	max 100 after assembly
9	Contamination	Min 80% Salmonella sp.	Kraken2
9	Coverage	Python script	Min 10x
9	Total length of assembly	Python script with SeqIO package	Min 4250000 bases
10	CG %	FastQC	parameter pass
10	Contamination	KmerFinder tool; https://cge.cbs.dtu.dk/services/KmerFinder	Pure bacterial culture
10	Coverage	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Coverage >30x
10	Expected genome size	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Deviation <0,5 million bp from the expected genome size.
10	N50	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	>30 000 pb

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
10	Total length of assembly	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Should be similar to the expected read length from the sequencing platform.
10	Total number of contigs	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	<500 contigs
14	Average quality score	galaxy.sciensano	30
14	Contamination	galaxy.sciensano	1.00%
14	Coverage against the assembled contigs		20.00x
14	GC%	galaxy.sciensano	2.00%
74	Contamination	kraken2 2.0.7	1.00%
74	Coverage	Galaxy Sciensano Salmonella pipeline	20.00x
16	Contamination	Quast + KmerFinder (https://cge.food.dtu.dk/services/KmerFinder/)	Contamination suggested by assembly criteria, KmerFinder used to confirm
16	Coverage	Quast + clc	>24
16	GC%	Quast	52,0-52,3
16	Total length of assembly	Quast	4,4-5,2 Mbps
16	Total number of contigs	Quast	<500
17	cgMLST loci detected	BLASTn	95% warning; <90% fail
17	Contamination	kraken2 2.0.7	1% warning; 5% fail
17	Coverage	Quast	20x warning; 10x fail
17	N50	Quast	>20 000
17	Total length of assembly	Quast	4.7-5.3 Mb
17	Total number of contigs	Quast	<100
19	poor cgMLST result	Ridom Seqsphere software	less than 90% coverage in cgMLST analysis

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
19	read quality	FastQC	Read sets with quality scores below 20 at any position were filtered
19	Total length of assembly	seqkit stats	too long or too short assembly; more than 20% deviation of average length
23	Contamination	Kraken2	The majority of taxonomically classifed reads should be assigned to the target species
23	Coverage	ead mapping by bwa, processing by samtools and coverage calculation by QualiMap	>=30X
23	N50	QUAST	>=10 kb
23	Total length of assembly	QUAST	+/-10% of the median genome size for species in NCBI Genome database
24	% of good targets in cgMLST	Ridom seqsphere	> 98 %
24	Coverage	Ridom seqsphere	> 20
26	Contamination	KmerFinder 3.2; SPAdes; QUAST	Species identification; Total length not exceeding 20% genome size
26	N. reads and percentage after filtering	Trimmomatic	
26	N50	SPAdes; QUAST	N50>15000
26	Total length of assembly	SPAdes; QUAST	Total length not exceeding 20% genome size
26	Total number of contigs	SPAdes; QUAST	N.contigs<500
27/67/77	Contamination Parameter: Read Fraction Majority Genus	Aquamis in-house Pipeline v 1.3.11	x > 0.95 (PASS), x </= 0.95 (Fail)
27/67/77	Contamination Parameter: NumContamSNVs (ConFindr 0.7.4)	Aquamis v 1.3.11	x </= 6 (PASS), 6 < x </=; 7 (Warning), x > 7 (Fail)

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
27/67/77	Coverage	Aquamis v 1.3.11	x > 40 (PASS), 30 < x </= 40 (Warning), x </= 30 (Fail)
27/67/77	Duplication ratio		x </= 1.002 (PASS), x > 1.002 (Warning)
27/67/77	GC%	Aquamis v 1.3.11	51.8 < x </= 52.2975 (PASS), x </= 51.8 x > 52.2975 (Warning)
27/67/77	N50	Aquamis v 1.3.11	x > 53027 (PASS), x </= 53027 (Warning)
27/67/77	Parameter: Single copy Orthologs	Aquamis v 1.3.11	x > 0.95 (PASS), x </= 0.95 (Fail)
27/67/77	Total length of assembly	Aquamis v 1.3.11	4627000 < x </= 5006000 (PASS), 4351000 < x </= 4627000, 5006000 < x </= 5326000 (Warning), x </= 4351000, x > 5326000 (Fail)
27/67/77	Total number of contigs (>= 1000 bp)	Aquamis v 1.3.11	x </= 167.5 (PASS), x > 167.5 (Fail)
28/68	Contamination	Kraken	Less than 90% of the reads belonging to another taxon than Salmonella.
28/68	Coverage	SPAdes, Pilon, Samtools Depth	15X
28/68	GC%	Quast	About 52%
28/68	N50	Quast	100000 pb
28/68	Total length of assembly	Quast	>4500000 pb and <5500000 pb
28/68	Total number of contigs	Quast	Less than 200 contigs
29	Average read Depth	Samtools and bash	x30
29	Contamination	Burrows-Wheeler Aligner (BWA)	N/A
29	Read quality (length and quality score)	Trim Galore	Min quality score: 30, min length: 50

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
29	Total length of assembly	Python	4Mb
30/70	Contamination	Kraken2/Bracken for species confirmation/contamination-check. In house pipeline, VIGAS-P platform (built on IRIDA)	Still no absolute value, but aim for 95% of the reads identified as the species of interest
30/70	Coverage	Automatically calculated and listed in VIGAS-P (own platform)	Usually aim for minimum 30X coverage
30/70	GC%	Data from multiQC and from Quast	Not an exact threshold, but will give you an idea if you have sequenced the right species, so more like an indicator of contamination
30/70	N50	Quast (built in as a tool in the assembly pipeline in VIGAS-P)	Not a real threshold on this, also depending on read length etc but will be evaluated
30/70	Total length of assembly	Quast (built in as a tool in the assembly pipeline in VIGAS-P)	If this differs too much from what to expect. We do not have an exact threshold for this, but lean towards suggestion from EU-RL AMR +/- 20% of average Salmonella genome or whether the length of one assembly is very different from the rest.
30/70	Total number of contigs	Quast (built in as a tool in the assembly pipeline in VIGAS-P)	We have 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 probably look into it if it's very different from what we use to see for a specific species.
32	% of cgMLST alleles found and called	Statistics implemented in SeqSphere	>95%
32	Contamination	Mash Screen implemented in SeqSphere	+/- 5%
32	Coverage	statistics implemented in SeqSphere	50X but if less, % of good targets should be >98%

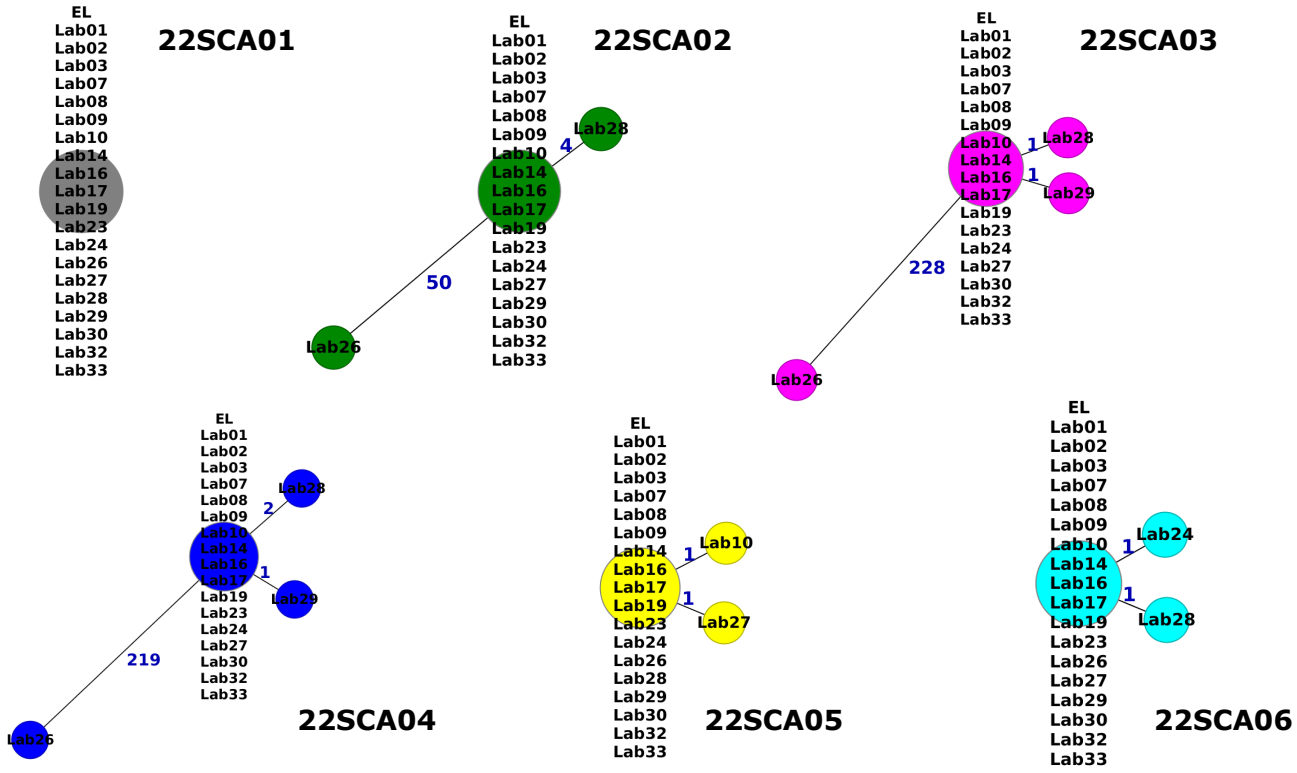
Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
32	Total length of assembly	statistics implemented in SeqSphere	length assembled < ref genome + 10%
32	Total number of contigs	statistics implemented in SeqSphere	<500
33	Contamination	CheckM	contamination < 2%
33	Contamination	Confindr	contamination status = False
33	N50	CheckM	>12500
33	Total length of assembly	CheckM	4200000-5200000 bpairs
33	Total number of contigs	CheckM	<300
EL	% good targets cgMLST	Ridom SeqSphere	>95%
EL	Contamination	CheckM	<4%
EL	Contamination	Kraken (MultiQC)	
EL	Coverage	Formula: (total reads * length of read)/length of genome sequenced	>30
EL	Total length of assembly	Ridom SeqSphere	4.5 - 5.2 Mb
EL	Total number of contigs	QUAST	<300

Annex 10 Md5 checksums of the 14 files that had to be downloaded from the sftp server for further analysis

ef67c2f8a661568dd6ea6b416b31c935	22SCA11_R1.fastq.gz
a1faf3e3d910d3ffa7626e7f2133d657	22SCA11_R2.fastq.gz
7df616cd89c6ec530eee347c812950cf	22SCA12_R1.fastq.gz
77dc74c4e5b45f55f58d7032b24bc8b0	22SCA12_R2.fastq.gz
0ea86d67a119bf13acbf67d75d46ebd8	22SCA13_R1.fastq.gz
f8fdd64d6d8bdfef56545007f03105f6	22SCA13_R2.fastq.gz
e61775a2192d0fcf5e51256c56a6ac90	22SCA14_R1.fastq.gz
9b6758df047758026294e018e0a6c139	22SCA14_R2.fastq.gz
e93b28a1abe7099909851421b657e0d7	22SCA15_R1.fastq.gz
e798974e3fd4ab4b63cc98b5562a3a71	22SCA15_R2.fastq.gz
04d9e7263606fe78ec625cce8c2c536f	22SCA16_R1.fastq.gz
4f6339282180b90db737e551a60dde49	22SCA16_R2.fastq.gz
65f572c91b90478c144d599e3035e432	22SCA-REF_R1.fastq.gz
3004180d62c0bf76115a61129e858119	22SCA-REF_R2.fastq.gz

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

MST for each 'wet' strain, using all participants' raw data, processed with the in-house developed Juno-assembly pipeline (Ridom SeqSphere+, cgMLST (3002), pairwise ignoring missing values). Results for Laboratory 26 indicate a swap between their results for strains 22SCA02, 22SCA03, and 22SCA04.



Annex 12 WGS results cluster analysis part, Results QC parameters on the in-house *de novo* assembled genomes, per participant

All statistics are based on contigs of size ≥ 500 bp.

	Laboratory code: 01		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,71	83	653667	4715215	52,2	136555	1451096	300	92,3
22SCA02	99,61	0,52	70	402068	4707142	52,1	135056	1499004	300	95,5
22SCA03	99,61	0,49	41	812126	4696973	52,1	267179	1922938	300	122,8
22SCA04	99,61	0,52	80	318580	4704542	52,1	132049	1655678	300	105,6
22SCA05	99,61	0,71	112	258089	4698902	52,2	84359	1594626	300	101,8
22SCA06	99,61	0,54	46	329729	4701584	52,1	167744	2236780	300	142,7

	Laboratory code: 02		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	24	1549270	4702094	52,1	421588	4525804	150	144,4
22SCA02	99,61	0,52	27	1549024	4701812	52,1	489949	6315842	150	201,5
22SCA03	99,61	0,49	23	1508817	4697928	52,1	490091	6230634	150	198,9
22SCA04	99,61	0,52	24	1549095	4702335	52,1	489948	5490754	150	175,1
22SCA05	99,61	0,52	29	991723	4702738	52,1	400838	3697564	150	117,9
22SCA06	99,61	0,52	26	1549444	4702882	52,1	421586	6788904	150	216,5

	Laboratory code: 03		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	28	1549444	4703986	52,1	489948	2680394	300	170,9
22SCA02	99,61	0,52	34	1050228	4702601	52,1	284466	1765938	300	112,7
22SCA03	99,61	0,49	31	1278622	4698836	52,1	478990	1428022	300	91,2
22SCA04	99,61	0,52	27	1549444	4702770	52,1	489948	1955196	300	124,7
22SCA05	99,61	0,6	28	1281349	4703106	52,1	489948	1908934	300	121,8
22SCA06	99,61	0,52	27	1271836	4702799	52,1	489948	1688440	300	107,7

	Laboratory code: 07		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,71	28	1550284	4704307	52,1	491898	857710	300	54,7
22SCA02	99,61	0,52	25	1549198	4701607	52,1	489949	919326	300	58,7
22SCA03	99,61	0,52	26	1508817	4699165	52,1	490379	925424	300	59,1
22SCA04	99,61	0,52	23	1549444	4702547	52,1	489948	959526	300	61,2
22SCA05	99,61	0,52	57	488493	4697828	52,2	211083	862448	300	55,1
22SCA06	99,61	0,52	26	1283547	4702026	52,1	489948	820656	300	52,4

	Laboratory code: 08		Platform used: NovaSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,57	24	1549444	4702720	52,1	489948	21511962	150	686,2
22SCA02	99,61	0,64	24	1549198	4701409	52,1	489949	24261104	150	774,1
22SCA03	99,61	0,72	25	1508817	4697948	52,1	478990	25058374	150	800,1
22SCA04	99,61	0,63	24	1549444	4702338	52,1	489948	22744924	150	725,5
22SCA05	99,61	0,52	25	1549444	4702316	52,1	489948	23410020	150	746,8
22SCA06	99,61	0,67	24	1549444	4702207	52,1	490239	19700806	150	628,5

	Laboratory code: 09		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	26	1549444	4703386	52,1	490239	2137478	250	113,6
22SCA02	99,61	0,52	28	1228501	4702475	52,1	490376	2231734	250	118,6
22SCA03	99,61	0,55	31	1508817	4697777	52,1	478989	1819684	250	96,8
22SCA04	99,61	0,52	28	1549444	4702987	52,1	490239	1764100	250	93,8
22SCA05	99,61	0,52	36	1550284	4710429	52,1	491799	2333308	250	123,8
22SCA06	99,61	0,52	30	1551298	4702910	52,1	490239	2049582	250	109,0

	Laboratory code: 10		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	41	867290	4695786	52,1	350646	8609480	75	137,5
22SCA02	99,61	0,52	31	1469793	4694208	52,1	349540	8301998	75	132,6
22SCA03	99,61	0,49	40	771635	4687791	52,1	268816	9281860	75	148,5
22SCA04	99,61	0,52	41	1228351	4692822	52,1	349594	9312060	75	148,8
22SCA05	99,61	0,52	38	1390507	4697002	52,1	349729	8933014	75	142,6
22SCA06	99,61	0,52	35	1274218	4693985	52,1	349728	9015734	75	144,1

	Laboratory code: 14		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,65	25	1549444	4703126	52,1	489948	1655214	150	52,8
22SCA02	99,61	0,52	27	1549198	4702057	52,1	401355	2255354	150	71,9
22SCA03	99,61	0,49	26	2001793	4697097	52,1	729634	2339430	150	74,7
22SCA04	99,61	0,52	27	1549444	4702471	52,1	490239	2654732	150	84,7
22SCA05	99,61	0,52	26	1549444	4702966	52,1	490239	2278496	150	72,7
22SCA06	99,61	0,52	25	1549444	4701954	52,1	400838	2116838	150	67,5

	Laboratory code: 16		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	25	1550284	4702696	52,1	491898	1108672	300	70,7
22SCA02	99,61	0,52	28	1549198	4702804	52,1	491509	928728	300	59,2
22SCA03	99,61	0,49	28	1508817	4698264	52,1	490379	1681828	300	107,4
22SCA04	99,61	0,52	26	1550284	4702846	52,1	491608	1416022	300	90,3
22SCA05	99,61	0,52	25	1549444	4703074	52,1	490239	1245814	300	79,5
22SCA06	99,61	0,52	26	1549444	4702654	52,1	490239	2318224	300	147,9

	Laboratory code: 17		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	1,48	78	1027519	4735505	52,2	489948	1059748	250	55,9
22SCA02	99,61	0,6	38	1549653	4712865	52,1	491608	1249550	250	66,3
22SCA03	99,61	0,65	47	1508817	4707466	52,1	478990	1703696	250	90,5
22SCA04	99,61	0,64	28	1549444	4706282	52,1	490239	1655268	250	87,9
22SCA05	99,61	1,43	57	1550284	4722045	52,2	491898	1576296	250	83,5
22SCA06	99,61	1,14	56	1549444	4725191	52,2	489948	1194330	250	63,2

	Laboratory code: 19		Platform used: MiniSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	26	1549270	4701439	52,1	421896	1874784	150	59,8
22SCA02	99,61	0,52	27	1548745	4700545	52,1	478897	1537402	150	49,1
22SCA03	99,61	0,51	29	1508458	4695731	52,1	682217	1224444	150	39,1
22SCA04	99,61	0,52	28	1386609	4701658	52,1	490121	1574948	150	50,2
22SCA05	99,61	0,52	27	1548816	4701825	52,1	421478	1374674	150	43,9
22SCA06	99,61	0,52	26	1550110	4700669	52,1	479391	1916042	150	61,1

	Laboratory code: 23		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,72	25	1550284	4701889	52,1	491898	22219158	150	708,8
22SCA02	99,61	0,54	27	1549198	4701434	52,1	491608	15785742	150	503,6
22SCA03	99,61	0,49	26	1509657	4697614	52,1	491888	22386646	150	714,8
22SCA04	99,61	0,59	25	1550284	4702071	52,1	491898	29821926	150	951,3
22SCA05	99,61	0,52	25	1550284	4702162	52,1	491898	15666678	150	499,8
22SCA06	99,61	0,52	27	1550284	4703527	52,1	491898	18650474	150	594,8

	Laboratory code: 24		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,65	1,17	66	461547	4709371	52,1	139508	1940410	250	103,0
22SCA02	99,62	1,25	102	397459	4715440	52,1	102021	2155624	250	114,3
22SCA03	99,59	2,11	129	262967	4716920	52,1	88972	2134730	250	113,1
22SCA04	99,61	0,55	42	492817	4704903	52,1	227935	2304878	250	122,5
22SCA05	99,61	0,93	30	1550284	4704973	52,1	491898	2450678	250	130,2
22SCA06	99,61	0,52	31	1550284	4705268	52,1	401028	1689752	250	89,8

	Laboratory code: 26		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	63	356621	4699635	52,1	166834	3515358	150	112,2
22SCA02	99,61	0,52	57	356621	4699992	52,1	173666	3973228	150	126,8
22SCA03	99,52	0,6	151	187417	4698056	52,1	62774	1185102	150	37,8
22SCA04	99,61	0,49	56	429578	4697220	52,1	174279	3750800	150	119,8
22SCA05	99,61	0,52	68	371866	4698845	52,1	152221	2654866	150	84,8
22SCA06	99,61	0,52	68	288372	4699431	52,1	141935	3256582	150	103,9

	Laboratory code: 27		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	30	1222363	4700831	52,1	323233	2687898	150	85,8
22SCA02	99,61	0,52	31	1386306	4701720	52,1	421590	3558108	150	113,5
22SCA03	99,61	0,49	30	951040	4696353	52,1	406454	2631280	150	84,0
22SCA04	99,61	0,52	32	650535	4700769	52,1	395575	3434892	150	109,6
22SCA05	99,61	0,52	29	1221452	4702257	52,1	327634	3257408	150	103,9
22SCA06	99,61	0,52	29	1234130	4701245	52,1	406138	2679328	150	85,5

	Laboratory code: 28		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	25	1549444	4702913	52,1	400838	878928	150	28,0
22SCA02	99,61	0,53	39	479389	4701756	52,1	275971	551588	150	17,6
22SCA03	99,61	0,49	32	1278622	4696332	52,1	400759	882840	150	28,2
22SCA04	99,61	0,52	27	1549270	4702688	52,1	257745	832766	150	26,6
22SCA05	99,61	0,52	28	872086	4701943	52,1	491846	764450	150	24,4
22SCA06	99,61	0,52	29	716846	4702143	52,1	400649	772176	150	24,6

	Laboratory code: 29		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	43	421585	4701154	52,1	253111	3099440	150	98,9
22SCA02	99,61	0,52	39	863170	4699362	52,1	375907	5657520	150	180,6
22SCA03	99,61	0,49	148	221041	4691830	52,2	61468	2462454	150	78,7
22SCA04	99,61	0,56	125	228179	4696581	52,2	74500	2263174	150	72,3
22SCA05	99,61	0,52	83	232524	4698644	52,1	106151	3183174	150	101,6
22SCA06	99,61	0,52	121	271836	4698080	52,1	67085	2271450	150	72,5

	Laboratory code: 30		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	25	1549444	4703075	52,1	490239	1573528	300	100,4
22SCA02	99,61	0,52	26	1549120	4702495	52,1	489950	1007748	300	64,3
22SCA03	99,61	0,49	26	1508817	4696847	52,1	490379	978842	300	62,5
22SCA04	99,61	0,55	24	1549444	4702993	52,1	489948	942652	300	60,1
22SCA05	99,61	0,52	25	1550284	4702344	52,1	491898	738092	300	47,1
22SCA06	99,61	0,52	36	697296	4703215	52,1	264187	1203752	300	76,8

	Laboratory code: 32		Platform used: MiniSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	27	1386764	4702287	52,1	491898	3495104	150	111,5
22SCA02	99,61	0,52	27	1549198	4701599	52,1	489839	3169170	150	101,1
22SCA03	99,61	0,49	26	1508643	4697027	52,1	479000	2640514	150	84,3
22SCA04	99,61	0,52	26	1385879	4700336	52,1	694319	1551012	150	49,5
22SCA05	99,61	0,52	24	1550017	4700977	52,1	491788	1754034	150	56,0
22SCA06	99,61	0,52	26	1550110	4701643	52,1	491788	2385798	150	76,1

	Laboratory code: 33		Platform used: MiSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,53	85	549501	4702024	52,1	109130	1217364	300	77,7
22SCA02	99,61	0,52	32	1279658	4702536	52,1	284466	1976320	300	126,1
22SCA03	99,61	0,49	39	1215803	4697705	52,1	261776	2457226	300	156,9
22SCA04	99,61	0,52	72	472602	4701345	52,1	140717	3152258	300	201,2
22SCA05	99,61	0,52	47	681733	4703164	52,1	244735	1896272	300	121,0
22SCA06	99,61	0,52	38	1025875	4702935	52,1	248193	2089472	300	133,3

	Laboratory code: EL		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA01	99,61	0,52	31	977938	4700830	52,1	406009	3794370	150	121,1
22SCA02	99,61	0,52	31	1025281	4700255	52,1	406510	4901484	150	156,4
22SCA03	99,61	0,49	25	864024	4693923	52,1	421566	3892224	150	124,4
22SCA04	99,61	0,52	27	1386331	4701978	52,1	477278	4187256	150	133,6
22SCA05	99,61	0,52	29	951041	4700855	52,1	401824	4118072	150	131,4
22SCA06	99,61	0,52	32	770421	4701605	52,1	323171	4339682	150	138,5

	Laboratory code: EL		Platform used: NextSeq							
Strain	**Completeness**	**Contamination**	**# contigs**	**Largest contig**	**Total length**	**GC (%)**	**N50**	**Input read pairs**	**Read Length**	**Coverage**
22SCA11	99,61	0,52	27	1386344	4702037	52,1	489949	3482630	150	111,1
22SCA12	99,61	0,52	30	921650	4700086	52,1	410690	2255252	150	72,0
22SCA13	99,61	176.4	3857	1385207	12444142	51,1	11982	7141008	150	86,1
22SCA14	99,61	0,49	27	1509036	4715706	52,1	477416	6934580	150	220,1
22SCA15	99,61	0,52	25	1550284	4701883	52,1	491898	7449404	150	237,7
22SCA16	99,61	0,53	28	1482359	4780641	52,2	432463	4971470	150	156,0
22SCA-REF	99,61	0,56	25	1550109	4701633	52,1	491898	8313470	150	265,2

Annex 13 Reasons for (not) excluding strain 22SCA13 from cluster analysis

Lab code	Strain 22SCA13 excluded from cluster analysis	Reason(s) not passing QC
1	Yes	Contamination Check Result: Potential contamination by second species above 10% detected: Escherichia coli; genome size too big (12.7 MB); strain excluded for MST and matrix
2	Yes	total length too high, # contigs too high, GC% too low, contamination too high: mostly E. coli reads
3	Yes	Failed on assembly size, contig number and N50 values. Contamination confirmed using Kmer Finder.
7	Yes	Contamination with E.coli, number of contigs over 3400, total length is oversize
8	Yes	Contamination by Escherichia 42.26%: coli(39.96%) - Salmonella 37.83%: enterica(37.57%)
9	Yes	Contamination and failed assembly
10	Yes	purity of culture; CG%; No. contigs; genome size
16	Yes	Total length 10,7Mbp, GC% 50,9, 3490 contigs, contamination confirmed by KmerFinder
17	Yes	contaminated with E. coli, low N50, total length 2x, number of contigs too high, <90% MLST loci detected
19	Yes	final assembly length too large
23	Yes	Contaminated with E. coli, assembly size too big for Salmonella
24	Yes	Contamination with E. coli
26	Yes	N. contigs >500; Total length higher than expected; N50<15000
27	Yes	Fail: Total length 12,718,480 bp; Read Fraction Majority Genus 0.488; Contam SNVs 2508 (inter and intra contamination); Warning: # Contigs 5,602; N50 11,563; Single copy orthologs 0.500; Duplication Rate 1.415; GC 51.07
30	Yes	Contamination with other species (E. coli)
32	Yes	Potential contamination by second species above 10% detected: Escherichia coli
33	Yes	contamination status = True
14	No	
28	No*	Sample 22SCA13 was contaminated (only about 35% of the reads were classified as belonging to the Salmonella taxon). We select those reads removing that way the contamination. Thus, we continue the analysis just with the reads classified as Salmonella.
29	No	

*The PT Typing 2022 Protocol indicated to exclude strains from the cluster analysis if the data did not pass the QC, therefore the approach by Laboratory 28 was considered as deviating.

In blue: Deviation from the expected result.

Annex 14 Per submission, the participants' distance matrix data for their comparison to the reference strain 22SCA-REF with the test strains

	Strain code												
Lab code-method	**22 SCA-REF**	**22 SCA01**	**22 SCA02**	**22 SCA03**	**22 SCA04**	**22 SCA05**	**22 SCA06**	**22 SCA11**	**22 SCA12**	**22 SCA13***	**22 SCA14**	**22 SCA15**	**22 SCA16**
1-cgMLST	0	2	50	218	0	2	0	50	50		249	2	544
2-cgMLST	0	2	50	220	0	2	0	50	50		251	3	546
3-cgMLST	0	1	51	225	0	1	0	51	51		259	1	561
7-cgMLST	0	4	54	223	1	4	1	54	54		251	4	552
8-cgMLST	0	2	50	218	0	2	0	50	50		250	2	544
14-cgMLST	0	6	103	411	0	6	0	103	103	414	472	6	1220
17-cgMLST	0	6	58	229	1	6	1	58	58		262	5	567
19-cgMLST	0	3	51	219	1	3	1	51	52		250	3	545
23-cgMLST	0	2	50	220	0	2	0	50	51		251	2	546
24-cgMLST	0	2	49	218	0	2	0	50	50		249	2	543
27-cgMLST1	0	4	53	220	1	4	1	53	54	213[a)]	250	4	544
27-cgMLST2	0	2	51	220	0	3	0	51	51		251	2	546
28-cgMLST	0	3	53	212	2	5	1	48	49	232[b)]	246	3	537
30-cgMLST		4			1	4	1		54			5	
32-cgMLST	0	2	50	219	0	2	0	50	50		248	2	545
33-cgMLST	0	4	64	246	0	4	0	64	65		270	4	606
EL-Salm-cgMLST	0	2	50	220	0	2	0	50	50		251	3	546
26-SNPa	0	7	1	126	455	9	3	108	110		521	9	1413
30-SNPa	0	10	111	453	4	10	4	111	109		516	11	1220
8-SNPr	0	6	123	501	0	6	0	117	118		596	6	1321
9-SNPr	0	6	98	447	0	6	0	98	98		500	6	1223
10-SNPr	0	6	101	443	0	6	0	101	101		506	6	1447
14-SNPr	0	6	109	449	0	6	0	109	109	468	504	6	1327
16-SNPr	0	9	111	451	3	9	2	109	112	403[c)]	520	8	1420
27-SNPr	0	6	108	479	0	6	0	108	108		531	6	1683
28-SNPr	0	7	112	495	2	7	1	114	112	516[b)]	636	6	1765
29-SNPr	0	6	102	418	1	6	1	102	102	418	461	6	1127

* Strain 22SCA13 was expected not to be included in the cluster analysis (due to not passing QC).
[a)] 22SCA13 QC failed and will not be included for reporting. However we checked the allelic differences for own interests.
[b)] See Annex 13.
[c)] Reported to be excluded from the cluster analysis.
Empty cells: no data reported.
In blue: Deviation from the expected result.

Published by:

National Institute for Public Health and the Enviroment, RIVM
P.O. Box 1 | 3720 BA Bilthoven
www.rivm.nl/en
The Netherlands

November 2023

Committed to
health and sustainability