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

EURL-*Salmonella* Proficiency Test Typing 2021

RIVM report 2022-0105
W.F. Jacobs-Reitsma et al.

EURL-*Salmonella* Proficiency Test Typing 2021

RIVM report 2022-0105

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-2022-0105

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* (2021) 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 2021

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 2021 quality control test on *Salmonella* typing. One laboratory was 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 2021, 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. Eight countries took part in 2021: the United Kingdom, the (potential) EU candidate countries Kosovo, North Macedonia, Serbia, 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 2021

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 2021 scoorden alle NRL's van de 27 EU-lidstaten goed bij deze kwaliteitscontrole op typering van *Salmonella*. Eén laboratorium had 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 2021 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 2021 waren dat er acht: het Verenigd Koninkrijk, de EU (potentiële) kandidaat-lidstaten Kosovo, Noord-Macedonië, Servië, 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 2021, the annual *Salmonella* typing Proficiency Test (PT) was organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands). The study's main objective was to evaluate whether the typing of *Salmonella* strains by the National Reference Laboratories for *Salmonella* (NRLs-*Salmonella*) in the European Union was carried out uniformly and whether comparable results were obtained.

A total of 35 laboratories participated in this study. These included the obligatory 27 NRLs-*Salmonella* in the 27 EU Member States. Eight additional NRLs participated voluntarily: the United Kingdom, the EU (potential) candidate countries Kosovo, North Macedonia, Serbia, and Türkiye and the EFTA countries Iceland, Norway and Switzerland.

All 35 laboratories performed serotyping. The EURL-*Salmonella* selected a total of twenty obligatory *Salmonella* strains plus one optional *Salmonella* 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). The laboratories were allowed to send strains for serotyping to another specialised laboratory in their country if this was part of their usual procedure.

Overall, 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 (Mooijman, 2007). Based on these criteria, the participants' performance was very good, including the two participants that submitted WGS-based results. All but one participant met the criteria for good performance in the first stage of the study. One participant had to participate in a follow-up study, including ten additional strains for serotyping. Ultimately, all 35 evaluated NRLs achieved good performance.

Nineteen NRLs also performed additional typing at the DNA level (MLVA and/or WGS) to investigate an additional set of ten *Salmonella* strains using cluster analysis. Similar to the PT Cluster Analysis 2020, the PT Cluster Analysis 2021 was mimicking an outbreak situation, with a *Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1 as the reference strain. Raw WGS data (compressed paired-end fastq files) of this reference strain were made available through a secure ftp server. Participants were asked to analyse the ten strains and to report per strain if a clustering match with the reference strain was found or not.

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.*
All five participants reported the MLVA-based cluster analysis results fully as expected. Fourteen of the 23 submissions (three participants with multiple submissions) reported the WGS-based cluster analysis

results fully as expected. All deviating results were related to one particular strain (21SCA08). The technical duplicates 21SCA06/21SCA09 were expected to be reported as (part of) one cluster and this was done in all 5 MLVA and all 23 WGS submissions.

1 Introduction

This report describes the 2021 Proficiency Test (PT) on the typing of *Salmonella* organised by the European Union Reference Laboratory for *Salmonella* (EURL-*Salmonella*, Bilthoven, the Netherlands) in November 2021.

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 35 laboratories participated in the PT Typing 2021. These included 27 NRLs-*Salmonella* in the 27 EU Member States and eight 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 three 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 2021 also included an optional part on cluster analysis. The cluster analysis involved ten *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 3-10-4-4-1 as the reference strain.
Raw WGS data (compressed paired-end fastq files) of this reference strain were made available through a secure ftp server. Participants were asked to analyse the ten strains and to report per strain whether a clustering match with the reference strain was found or not.

A total of nineteen NRLs participated in the cluster analysis: all nineteen performed WGS analysis and five participants also performed MLVA analysis.

2 Participants

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

[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: 1-35.

3.1.2 *Protocol and test report*
Three weeks before the start of the PT, the NRLs received the protocol by email. Web-based result forms were used to report results. Instructions for completing these result forms and data entry were sent to the NRLs on 11 November 2021, in separate emails for serotyping and cluster analysis.
The protocol and blank result forms can be found on the EURL-*Salmonella* website:
https://www.eurlsalmonella.eu/proficiency-testing/typing-studies

3.1.3 *Transport*
The parcels containing the strains for serotyping and cluster analysis were sent by the EURL-*Salmonella* on 8 November 2021. 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 26th EURL-*Salmonella* Workshop (Mooijman, 2021), 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 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). However, participants were asked to report only the results as detected and on which the identification of serovar names was based. Three 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 2021*

Strain code	O-antigens	H-antigens		Serovar	Origin
		(phase 1)	(phase 2)		
S1	1,9,12	g,m	-	Enteritidis	Human
S2 [a)]	11	e,h	1,2	Chingola	Non-human
S3 [a)]	1,4,12,27	b	e,n,z15	Wagenia	Human
S4	4,[5],12	l,v	e,n,z15	Brandenburg	Human
S5	3,{10}{15}{15,34}	y	1,5	Orion	Human
S6	6,7,14	e,h	e,n,z15	Braenderup	Human
S7	{6,7,14}{54}	g,m,[p],s	[1,2,7]	Montevideo	Human
S8 [b)]	1,4,[5],12	i	-	4,5,12:i:-	Human
S9 [a)]	1,4,12,[27]	b	l,w	Wien	Human
S10	6,7,14	r	1,2	Virchow	Non-human
S11	1,9,12	l,z13	e,n,x	Napoli	Human
S12	8,20	z38	-	Apeyeme	Human
S13	6,7,14	r	1,5	Infantis	Human
S14	28	z10	e,n,x	Umbilo	Human
S15	6,8	z10	e,n,x	Hadar	Human
S16	4,[5],12	a	1,7	Arechavaleta	Human
S17	3,{10}{15}{15,34}	e,h	1,6	Anatum	Human
S18	6,8,20	r,[i]	1,5	Bovismorbificans	Human
S19	1,4,[5],12	i	1,2	Typhimurium	Human
S20	1,13,23	z29	-	Cubana	Human
S21[c)]	50	r	1,5,(7)	50:r:1,5 (IIIb)	Human

[a)] Represented in an EURL-*Salmonella* PT Serotyping for the first time.
[b)] Monophasic variant of Typhimurium as determined by PCR.
[c)] *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 (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 2021*

Strain code	O-antigens	H-antigens		Serovar	Origin
		(phase 1)	(phase 2)		
SF1	1,9,12	g,m	-	Enteritidis	Non-human
SF2	4,12	i	1,6	Agama	Non-human
SF3	8,20	r,[i]	z6	Altona	Human
SF4	1,4,[5],12	f,g,s	[1,2]	Agona	Rapeseed meal
SF5	1,4,[5],12	i	1,2	Typhimurium	Environmental
SF6[a)]	1,4,[5],12	i	-	1,4,[5],12:i:-	Human
SF7	6,7,14	z10	l,w	Jerusalem	Broiler
SF8	1,4,[5],12	i	1,5	Lagos	Human
SF9	3,{10}{15}{15,34}	e,h	l,w	Meleagridis	Meat and bone meal
SF10[b)]	4,[5],12	i	e,n,x	Farsta	Reptile

[a)] Monophasic variant of Typhimurium as determined by PCR.
[b)] Undated antigenic formula according to Issenhuth-Jeanjean et al., 2014.

3.3 Cluster analysis part of the PT

3.3.1 Salmonella *strains for cluster analysis*

A total of ten *Salmonella* strains (21SCA01 – 21SCA10) were included in the part on cluster analysis. Background information on the strains is given in Table 3.4.

Table 3.4 Background information on the Salmonella *strains used for cluster analysis in 2021*

Strain code	Serovar	ST	MLVA-profile	Origin
21SCA01	Enteritidis	11	2-9-9-4-2	Human
21SCA02	Enteritidis	183	2-11-9-3-1	Human
21SCA03	Enteritidis	183	2-11-9-3-1	Human
21SCA04	Enteritidis	11	3-10-4-4-1	Human
21SCA05	Enteritidis	1925	3-10-5-4-1	Human
21SCA06 [a)]	Enteritidis	11	3-10-4-4-1	Human
21SCA07	Enteritidis	3406	2-14-NA-7-NA	Human
21SCA08	Enteritidis	11	3-10-4-4-1	Human
21SCA09 [a)]	Enteritidis	11	3-10-4-4-1	Human
21SCA10	Enteritidis	11	1-10-7-3-2	Human

[a)] Technical duplicates (in bold).

Strains were pre-tested by the EURL-*Salmonella* to be suitable for cluster analysis using either MLVA or WGS. Initially, a set of fifteen human surveillance strains, collected and sequenced in 2019 by the National *Salmonella* Centre at RIVM, was selected to be tested for potential use in the PT2021. All test strains were freshly cultured on blood-agar plates and a single colony was selected to produce two blood-agar plates which were submitted for MLVA and WGS analysis respectively (8 July 2021). Approximately every other day, all test strains were subcultured, using alternately liquid (BPW) and solid (blood-agar plates) media. The strains were resubmitted for MLVA and WGS analysis (as described above) after ten times sub-culturing (17 August 2021).

Identical MLVA results were obtained before and after the ten times sub-culturing, and these results also fully matched the original MLVA results in 2019.
WGS pre-test results are shown in Figure 3.1. 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).

Based on the pre-test results, nine stable strains were selected to be included in the PT Cluster Analysis 2021. The tenth strain was a technical duplicate; strain 21SCA06 and strain 21SCA09 shipment tubes were both prepared from the same blood-agar plate containing strain 21SCA06.
Figure 3.1 shows the WGS pre-test results as well as the EURL-*Salmonella* PT2021 results for the ten selected strains (Table 3.4).
The original WGS data from the human surveillance strains in 2019 are indicated with ELt0, the WGS data for initial testing are indicated with ELt1, and the WGS data after ten times sub-culturing are indicated with ELt2. The PT 2021 set of strains was additionally tested both at the start

of the PT (November 2021: ELt3) and the end of the data submission period (February 2022: ELt4).

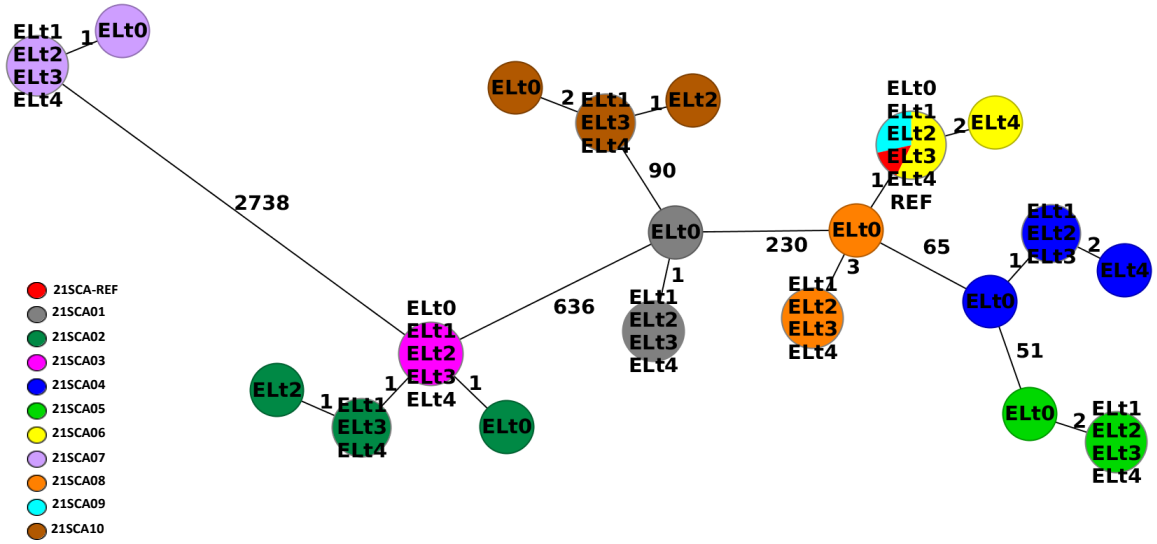


*Figure 3.1 MST of the EURL-*Salmonella *pre-test and PT 2021 results, (Ridom SeqSphere⁺, cgMLST (3002), pairwise ignoring missing values).*

ELt0: Original WGS data from the stored strains (2019);
ELt1: WGS data from initial pre-testing (8 July 2021);
ELt2: WGS data after ten times sub-culturing (17 August 2021);
ELt3: PT2021 data at the start of the PT (November 2021);
ELt4: PT2021 data at the end of the PT (February 2022).

3.3.2 *Evaluation of the cluster analysis results in general*

Cluster analysis was performed up to the choice of the participant by MLVA and/or WGS, and using their own routine method(s).
The PT Cluster Analysis 2021 was mimicking an outbreak situation, with a *Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1 as the reference strain (21SCA-REF). Raw WGS data of this strain (compressed paired-end fastq files) were made available through a secure ftp server.
For this particular PT2021 situation, the cluster definition was set at a maximum of seven allelic differences from the reference sequence. For MLVA, the cluster definition was set at loci with a different number of repeats.
Participants were asked to analyse the ten *Salmonella* strains and to report per strain whether a clustering match with the reference strain was found or not. Details on the method(s) used and the outcome of the cluster analysis had to be reported in the electronic result form. Additionally, specific data for WGS had to be sent by email or uploaded to a 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.*
No specific performance criteria were set for this PT on cluster analysis. As a minimum, it was expected that participants would report the

technical duplicate strains 21SCA06 and 21SCA09 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:

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

Participants were asked to report per strain (Table 3.4) whether they found a clustering match with the reference outbreak strain (21SCA-REF) in the EURL-*Salmonella* PT Typing 2021:
Salmonella Enteritidis ST11, MLVA type 3-10-4-4-1.
The MLVA cluster definition for the PT Typing 2021 was set at no loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 21SCA04, 21SCA06 (reference strain), 21SCA08 and 21SCA09 (technical duplicate of the reference strain) to be a clustering match with the reference outbreak strain as detailed in the PT Typing 2021.

3.3.4 Evaluation of the cluster analysis results based on WGS data

Data submission for WGS results included:

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

Participants were asked to report per strain (Table 3.4) whether a clustering match was found with the reference outbreak strain in the EURL-*Salmonella* PT Typing 2021: 21SCA-REF (*Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1).
The WGS cluster definition for the PT Typing 2021 was set at a maximum of seven allelic differences from the reference.
Based on this (cgMLST-)cluster definition, WGS-based results were expected to indicate strains 21SCA06 (reference strain), 21SCA08, and 21SCA09 (technical duplicate of the reference strain) to be a clustering match with the provided reference outbreak strain 21SCA-REF data as detailed in the PT Typing 2021 (also see Figure 3.1).

4 Results and Discussion

4.1 Technical data

4.1.1 General

A total of 35 laboratories participated in this PT (Chapter 2). These included 27 NRLs-*Salmonella* in the 27 EU Member States and 8 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 2021 are summarised in Table 4.1.

Table 4.1 Frequency and number of Salmonella *strains serotyped in 2021*

Laboratory code	Serotyping frequency in 2021	No. of strains serotyped in 2021
20	Daily	90
14	Daily	150
17	Daily	330
12	Daily	347
13	Daily	350
1	Daily	500
11	Daily	500
18	Daily	500
26	Daily	600
6	Daily	650
32	Daily	1100
21	Daily	1200
29	Daily	2400
23	Daily	2800
19	Daily	3000
34	Daily	3000
7	Daily	3500
10	Daily	3500
35	Daily	4000
27	Daily	4100
30	Twice a week	120
9	Twice a week	150
16	Twice a week	350
3	Twice a week	500
24	Twice a week	1000
22	Twice a week	2000
28	Thrice a week	158
2	Thrice a week	520
31	Once a week	45
15	Once a week	55
25	Once a week	60
8	Once a week	100

Laboratory code	Serotyping frequency in 2021	No. of strains serotyped in 2021
33	Once a week	100
5	Once a week	1100
4	As necessary	20
n=35		38895

4.1.2 Accreditation

Of the 35 participants, 33 are accredited for serotyping *Salmonella.* Thirty-one 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, and another laboratory mentioned EN ISO 6579-1 only.
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 2022.
32 laboratories stated that they are accredited for all *Salmonella* serovars, and one laboratory indicated to be accredited for serotyping *S. enterica* subsp. *enterica*.

4.1.3 Transport of samples

All but four participants received their package within two days after shipment on Monday 8 November 2021. One package was received by the laboratory on 12 November, two on 15 November and the final one on 21 November 2021. 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. Laboratory 26 forwarded strains S12 and S14 to their national typing centre. A total of 34 participants used classical serology. Seven of them mentioned the combined use of classical serology and Luminex assays (2), multiplex/real-time PCR (4), or WGS (1). One participant used Whole Genome Sequencing (WGS), supplemented with traditional agglutination using O:6 and O:8 (strains S12, S15, S18) or O:22 and O:23 (strain S20).
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=34)
Biorad	15
Pro-Lab	4
Sifin	20
Statens Serum Institute (SSI)	29
Other	6
Own preparation	3

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=34)
1	10
2	8
3	14
4	1
5	1

4.2.2 *Biochemical testing*

Seventeen participants indicated the use of (a variety of) biochemical tests on all or a limited number of strains. All seventeen participants biochemically tested strain S20 (*S*. Cubana). Laboratories 17 and 24 routinely tested all 21 strains using MALDI-TOF.

4.2.3 *Use of PCR for confirmation*

Fifteen laboratories used PCR to confirm strain S8, the monophasic variant of *S*. Typhimurium 1,4,[5],12:i:-, and five of these also used PCR to confirm strain S19, *S*. Typhimurium. Most laboratories mentioned using the reference by 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 completely typed correctly by 32 of the 35 participants (91%). This corresponds to nearly 100% of the total number of strains. The H-antigens were completely typed correctly by 26 of the 35 participants (74%), corresponding to 98% of the total number of strains. As a result, 26 participants (74%) reported all serovar names correctly, which corresponds to 98% of all strains evaluated.

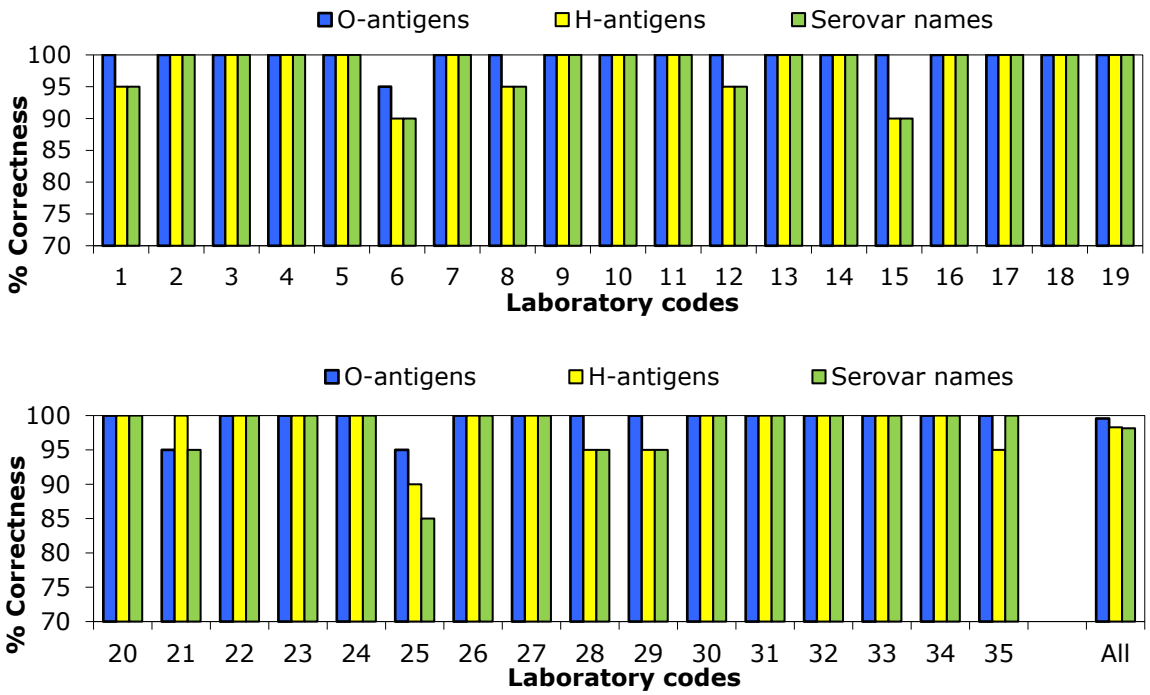


Figure 4.1 Percentages of correct serotyping results, per participant

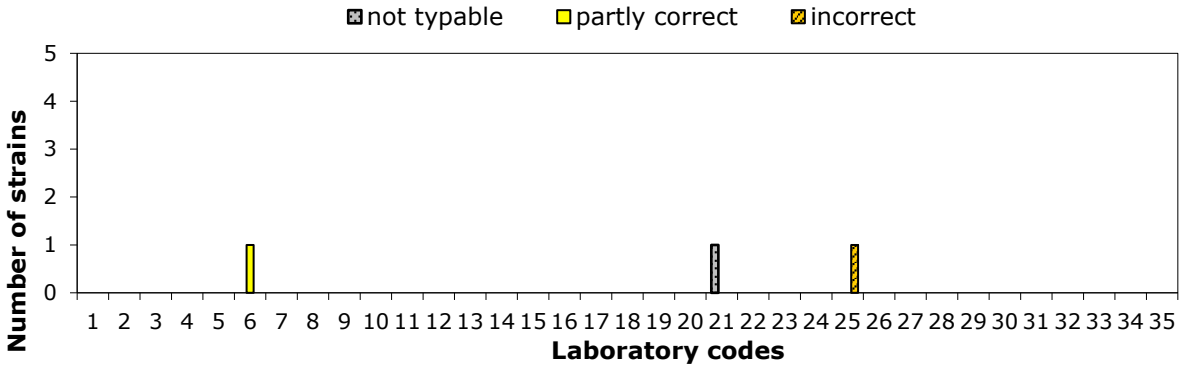


Figure 4.2 Evaluation of type of errors for O-antigens, per participant

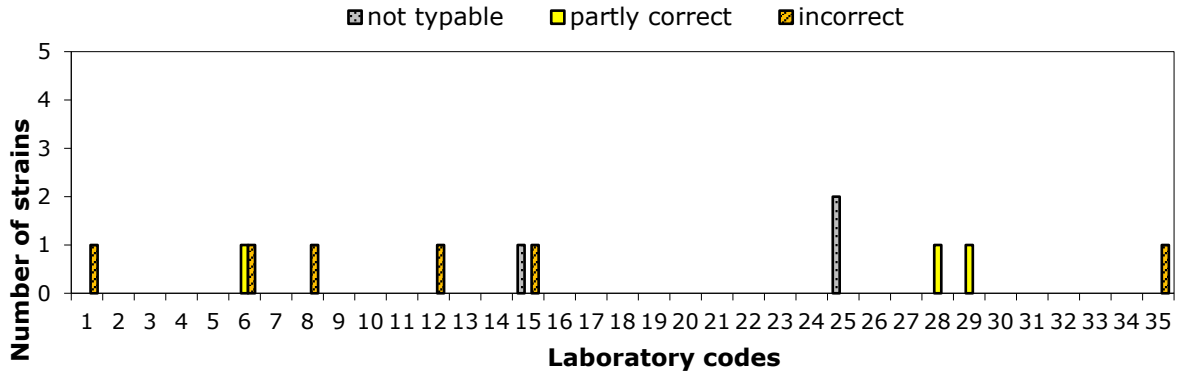


Figure 4.3 Evaluation of type of errors for H-antigens, per participant

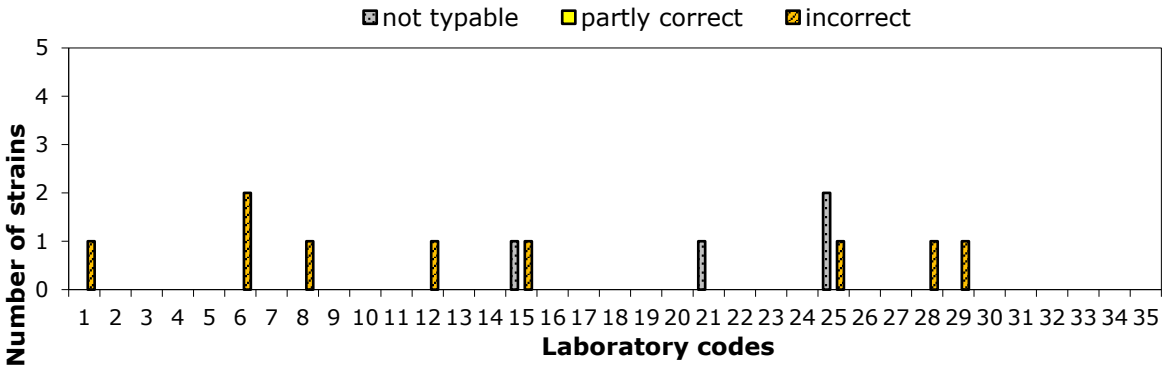


Figure 4.4 Evaluation of the type of errors in the identification of the serovar names, per participant

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 2021 was very good. One EU Member State NRL (Lab 12) did not meet the level of good performance at the first stage of the study. Therefore, a follow-up study for this laboratory was organised in March/April 2022.

All participants received both their individual laboratory evaluation report and the interim summary report on serotyping on 23 February 2022.

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-2021

Table 4.4 Evaluation of serotyping results per NRL

Laboratory code	Penalty points	Good performance
1	1	yes
2	0	yes
3	0	yes
4	0	yes
5	0	yes
6	2	yes
7	0	yes
8	1	yes
9	0	yes
10	0	yes
11	0	yes
12	4	NO
13	0	yes
14	0	yes
15	1	yes
16	0	yes

Laboratory code	Penalty points	Good performance
17	0	yes
18	0	yes
19	0	yes
20	0	yes
21	0	yes
22	0	yes
23	0	yes
24	0	yes
25	1	yes
26	0	yes
27	0	yes
28	1	yes
29	1	yes
30	0	yes
31	0	yes
32	0	yes
33	0	yes
34	0	yes
35	0	yes

4.2.6 Serotyping results per strain

Annex 2 displays the final naming results reported per strain (S1 – S20) and per laboratory (1-35).
A completely correct identification was obtained for ten *Salmonella* serovars: Enteritidis (S1), Chingola (S2), Braenderup (S6), Montevideo (S7), Wien (S9), Virchow (S10), Infantis (S13), Hadar (S15), Anatum (S17), and Typhimurium (S19).
Annex 2 also shows the reported serovar names for strain 1,4,[5],12:i:- (S8). Fifteen participants used a PCR method to confirm this strain to be a monophasic Typhimurium strain.
Annex 3 includes the details on the strains that caused problems in serotyping.
Annex 4 describes details on the additional and optional strain S21.
All but three participants tried to serotype strain S21, a *Salmonella enterica* subsp. *diarizonae* (IIIb). A few laboratories did not have access to the required antisera to finalise the serotyping of this strain (50:r:1,5).

4.2.7 Results follow-up study

One EU NRL 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. This NRL received ten additional strains for serotyping in week 12, 2022.
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: this participant 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
12	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 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 study. The rise seen in the 2018 study was mainly caused by the seven EU NRLs that made a mistake in typing an *S.* Cannstatt strain. The decreased 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 results (Figure 4.6).
However, the number of EU NRLs with a non-good performance is low: two in the period 2010 – 2013, one in the 2014, 2015, 2018, and 2021 studies, and none in the 2016, 2017, 2019 and 2020 studies.

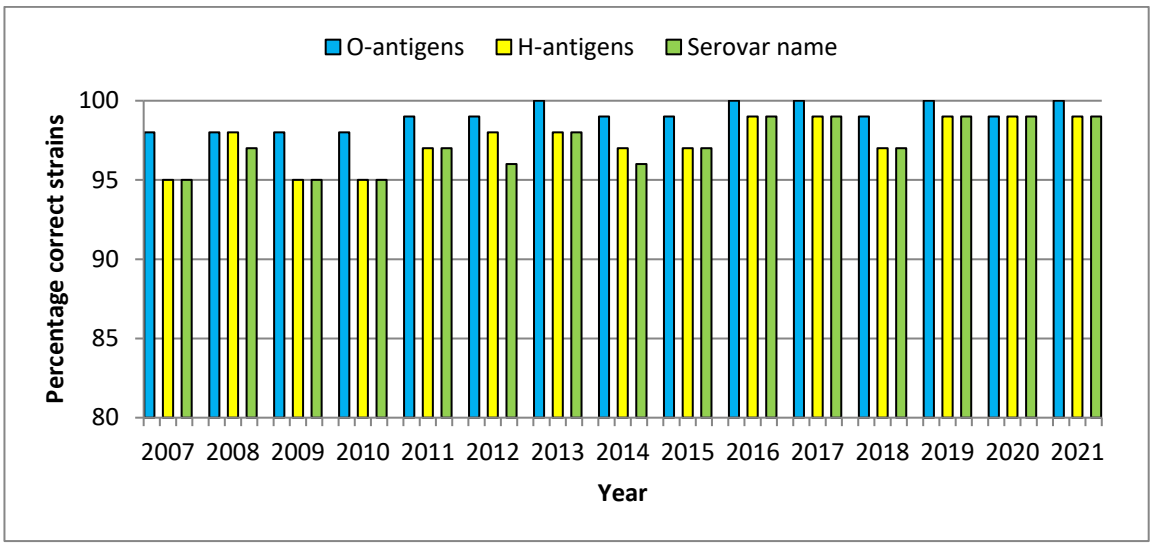


Figure 4.5 Serotyping results of the EU NRLs, based on the percentages of correctly typed strains

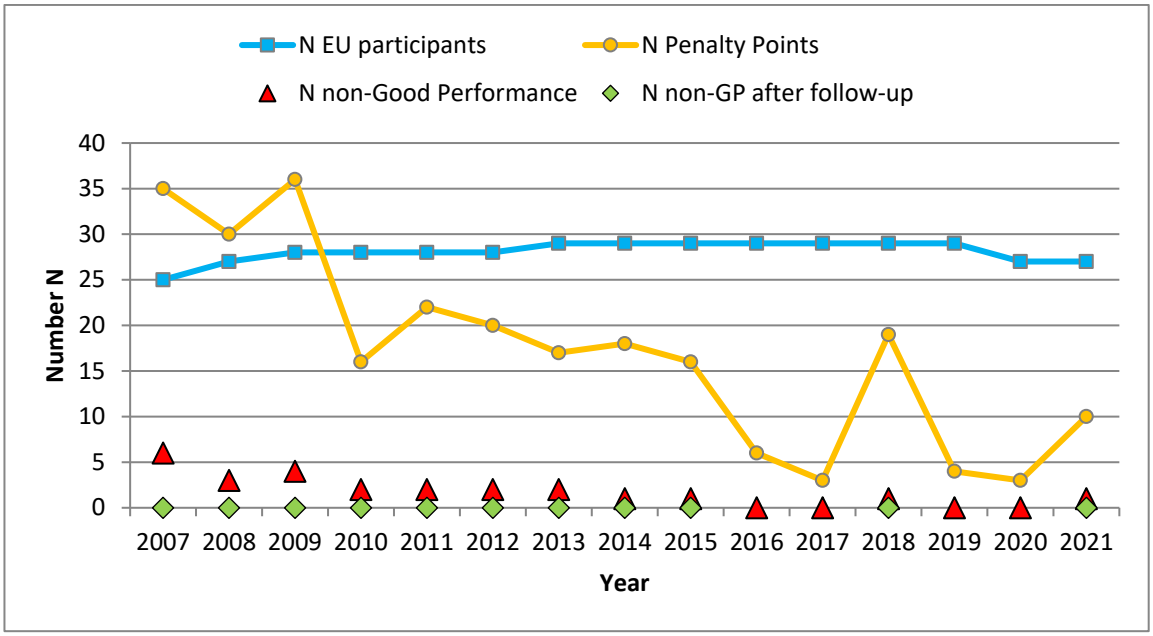


Figure 4.6 Serotyping results of the EU NRLs-Salmonella, based on the number (N) of Penalty Points and non-Good Performance (non-GP)

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 nineteen NRLs participated in the cluster analysis. All nineteen performed WGS analysis. Five participants additionally performed MLVA analysis.
All participants received both their individual laboratory evaluation report and the interim summary report on cluster analysis on 26 May 2022. Annex 5 gives an example of an individual laboratory evaluation report on cluster analysis results. The interim summary report is available on the EURL-*Salmonella* website:
https://www.eurlsalmonella.eu/documenten/interim-summary-report-eurl-salmonella-pt-cluster-analysis-2021

As a general question, the participants were asked if they serotyped the ten strains. Twelve participants indicated to have serotyped the 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 6, 7, 11, 23, and 35) submitted cluster analysis results based on MLVA data.
Annex 7 shows the allelic profiles submitted by the participants.

Participants were asked to report per strain if (yes or no) they found a clustering match with the reference outbreak strain (21SCA-REF) in the EURL-*Salmonella* PT Typing 2021:
Salmonella Enteritidis ST11, MLVA type 3-10-4-4-1.
The MLVA cluster definition for the PT Typing 2021 was set at no loci with a different number of repeats. Based on this cluster definition, MLVA-based results were expected to indicate strains 21SCA04,

21SCA06 (reference strain), 21SCA08 and 21SCA09 (technical duplicate of the reference strain) to be a clustering match with the REF outbreak strain.
All five participants reported the MLVA-based cluster analysis results fully as expected (Table 4.6). The technical duplicates 21SCA06/21SCA08 were expected to be reported as (part of) one cluster and all five participants did this (Table 4.6).

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

Lab code	21 SCA01	21 SCA02	21 SCA03	21 SCA04	21 SCA05	21 SCA06	21 SCA07	21 SCA08	21 SCA09	21 SCA10
Expected	**No**	**No**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**Yes**	**No**
6	No	No	No	Yes	No	Yes	No	Yes	Yes	No
7	No	No	No	Yes	No	Yes	No	Yes	Yes	No
11	No	No	No	Yes	No	Yes	No	Yes	Yes	No
23	No	No	No	Yes	No	Yes	No	Yes	Yes	No
35	No	No	No	Yes	No	Yes	No	Yes	Yes	No

4.3.3 Results cluster analysis based on WGS data

Nineteen participants (Table 4.8) submitted cluster analysis results based on WGS data. Three participants submitted both cgMLST-based and reference-based-SNP data, one of which in addition also submitted assembly-based-SNP data (23 submissions in total).

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 EURL-*Salmonella* performed DNA extraction, library preparation and sequencing in-house, except for participants 1, 6, 10 and 26 (library preparation and sequencing outsourced) and participant 11 (sequencing outsourced). Most participants used the Illumina MiSeq platform(10x), followed by the Illumina NextSeq (6x), and the Illumina NovaSeq (3x) or MiniSeq (1x). Including the EURL-*Salmonella*, fourteen participants used cgMLST for data analysis and ten participants used SNP-based analysis (6x reference-based and 4x assembly-based). Tools used for this analysis varied from in-house (chewBBaca-based) pipelines to commercial ones, most often Ridom SeqSphere^{+} (6x). The most commonly used tools for cluster analysis were Minimum Spanning Tree (MST, 12x), followed by Maximum Likelihood (ML, 7x), single linkage hierarchical clustering (3x) and Neighbor Joining (NJ, 2x).

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

Fourteen participants reported the md5 checksum for the *compressed* paired-end fastq files of the 21SCA-REF. These were correct for thirteen participants, indicating that the transfer of data from the secure ftp site went alright.

21SCA_REF_R1.fq.gz: 206064240c9fd1f9b89bbd74db4c31d6
21SCA_REF_R2.fq.gz: 4eb99f380cd3d3ca72642079528d4506

One participant reported md5 checksums that deviated from the expected ones, but these matched with the ones of the *uncompressed* paired-end fastq files:
21SCA_REF_R1.fq: e8e2aaff2830d4c2833d7ca203ecba01
21SCA_REF_R2.fq: 4078d467927d6ae2573984a2a8e02a0c

All participants' raw data (compressed fastq files) 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 (Figure 4.7). Annex 10 shows data per strain. Data for strain 21SCA09-Lab21 were excluded due to a 85,2% of good targets, which is below the quality threshold.

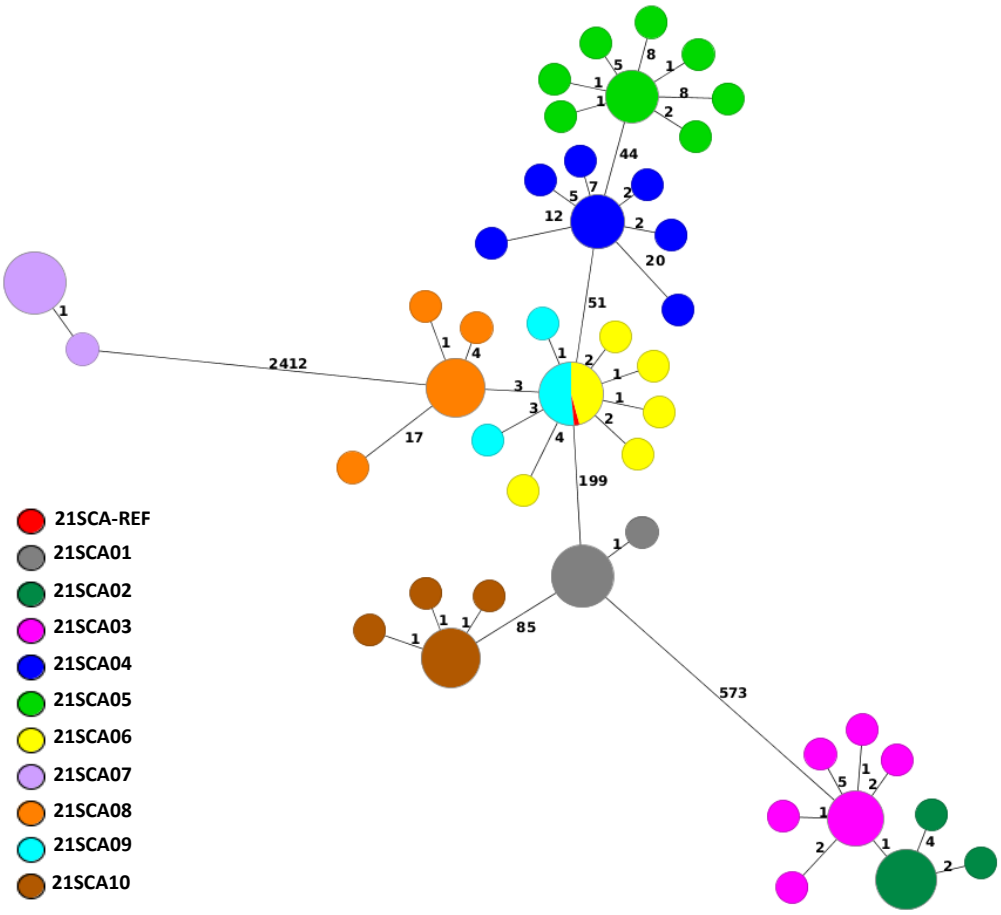


Figure 4.7 MST of the strains from the participants' raw data, processed with the in-house developed Juno-assembly pipeline (Ridom SeqSphere$^+$, cgMLST (3002), pairwise ignoring missing values)

An overview of the main QC parameters results on all in-house *de novo* assembled genomes (fasta files) is given in Table 4.7. Annex 11 shows detailed data per participant.

Table 4.7 Results QC parameters on the in-house de novo *assembled genomes, average per participant*

Laboratory code	Illumina Platform	Average # contigs	Average Largest contig	Average Total length	Average N50	Average Coverage
Lab01	NovaSeq	29	1291885	4740679	442147	333
Lab02	MiSeq	30	1251605	4741310	440259	64
Lab06	MiSeq	66	532471	4747424	216226	77
Lab07	MiSeq	120	273543	4738615	91042	89
Lab10	NovaSeq	28	1400018	4741572	456027	1142
Lab11	NextSeq	47	739956	4729025	258217	62
Lab12	MiSeq	336	1031772	5096253	351553	113
Lab14	MiniSeq	31	1153990	4739769	505746	46
Lab16	MiSeq	27	1346341	4741875	456211	78
Lab19	NextSeq	35	1055968	4739697	373170	80
Lab21	MiSeq	209	208161	4762167	62429	105
Lab22	NextSeq	31	1291895	4740739	435930	160
Lab23	MiSeq	37	852835	4742664	340196	138
Lab24	MiSeq	217	463159	4759701	195526	93
Lab26	NovaSeq	30	1276112	4740835	435752	263
Lab27	NextSeq	59	828347	4740385	304238	91
Lab30	MiSeq	41	1122732	4746658	380311	67
Lab34	NextSeq	122	249239	4737348	82829	71
Lab35	MiSeq	30	1423476	4746342	502131	67
EL	NextSeq	31	1214837	4740775	424874	151

Participants were asked to report per strain if (yes or no) they found a clustering match with the reference outbreak strain in the EURL-*Salmonella* PT Typing 2021: 21SCA-REF (*Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1).

The WGS cluster definition for the PT Typing 2021 was set at a maximum of seven allelic differences from the reference. Based on this (cgMLST-)cluster definition, WGS-based results were expected to indicate strains 21SCA06 (reference strain), 21SCA08, and 21SCA09 (technical duplicate of the reference strain) to be a clustering match with the provided 21SCA-REF outbreak strain (also see Figures 3.1 and 4.7).

Fourteen of the 23 submissions (three participants with multiple submissions) reported the WGS-based cluster analysis results fully as expected (Table 4.8).

Annex 12 shows, per submission, the participants' distance matrix data for their comparison of the 21SCA-REF with the ten tested strains. Based on these distance matrix data, all but one (Lab 12) of the thirteen cgMLST submissions were reported in accordance with the PT 2021 cluster definition of a maximum of seven allelic differences, even though this subsequently ended up in a deviation from the expected result (Table 4.6 and Annex 12).

No cluster definition was specified for SNP-based analysis. Therefore, the ten SNP submissions were based on the participants' internal criteria for a *Salmonella* cluster. The apparent variety in these internal criteria

may partly explain the differences in cluster analysis results reported for strain 21SCA08 (Annex 12).

The technical duplicates 21SCA06/21SCA09 were expected to be reported as (part of) one cluster and all 23 submissions did this (Table 4.8).

Table 4.8 Expected cluster analysis results and the cluster analysis results reported by the nineteen WGS participants

Lab code-method	21 SCA01	21 SCA02	21 SCA03	21 SCA04	21 SCA05	21 SCA06	21 SCA07	21 SCA08	21 SCA09	21 SCA10
Expected	**No**	**No**	**No**	**No**	**No**	**Yes**	**No**	**Yes**	**Yes**	**No**
1-SNPa	No	No	No	No	No	Yes	No	No	Yes	No
2-SNPa	No	No	No	No	No	Yes	No	Yes	Yes	No
6-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
6-SNPa	No	No	No	No	No	Yes	No	Yes	Yes	No
6-SNPr	No	No	No	No	No	Yes	No	Yes	Yes	No
7-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
10-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
10-SNPr	No	No	No	No	No	Yes	No	Yes	Yes	No
11-SNPr	No	No	No	No	No	Yes	No	Yes	Yes	No
12-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
14-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
16-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
19-cgMLST	No	No	No	No	No	Yes	No	No	Yes	No
19-SNPr	No	No	No	No	No	Yes	No	No	Yes	No
21-SNPr	No	No	No	No	No	Yes	No	No	Yes	No
22-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
23-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
24-cgMLST	No	No	No	No	No	Yes	No	Yes	Yes	No
26-cgMLST	No	No	No	No	No	Yes	No	No	Yes	No
27-SNPr	No	No	No	No	No	Yes	No	No	Yes	No
30-SNPa	No	No	No	No	No	Yes	No	No	Yes	No
34-cgMLST	No	No	No	No	No	Yes	No	No	Yes	No
35-cgMLST	No	No	No	No	No	Yes	No	No	Yes	No

SNPa: assembly-based SNP analysis, SNPr: reference-based SNP analysis.

Same laboratories, using different methods.

Deviation from the expected result.

5 Conclusions

5.1 Serotyping

- The overall results for the 35 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.
- One EU NRL-*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 eight 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 MLVA and WGS.
- A total of nineteen participants performed cluster analysis; five using MLVA analysis and nineteen using WGS analysis.
- All five participants reported the MLVA-based cluster analysis results fully as expected.
- Fourteen of the 23 submissions (three participants with multiple submissions) reported the WGS-based cluster analysis results fully as expected.
- The technical duplicates 21SCA06/21SCA09 were expected to be reported as (part of) one cluster all five MLVA and all 23 WGS submissions did this.

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 15-3-2023).

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 15-3-2023).

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 15-3-2023).

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

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

EN ISO 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.

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.

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 15-03-2023).

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 15-03-2023).

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 15-3-2023).

Mooijman, K.A., 2021. The 26th EURL-*Salmonella* workshop; 28 May 2021, Online. National Institute for Public Health and the Environment, Bilthoven, the Netherlands. RIVM Report no.: 2021-0130. https://www.rivm.nl/bibliotheek/rapporten/2021-0130.pdf (accessed 15-3-2023).

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
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
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 2021

Number of penalty points: 1 **Evaluation: Good Performance**

	Reference Results				**Results NRL labcode:**			**1**
Strain	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar
S1	1,9,12	g,m	-	Enteritidis	9,12	g,m	-	Enteritidis
S2	11	e,h	1,2	Chingola	11	e,h	1,2	Chingola
S3	1,4,12,27	b	e,n,z15	Wagenia	4,12	b	e,n,z15	Wagenia
S4	4,[5],12	l,v	e,n,z15	Brandenburg	4,12	l,v	e,n,z15	Brandenburg
S5	3,{10}{15}{15,34}	y	1,5	Orion	3,10	y	1,5	Orion
S6	6,7,14	e,h	e,n,z15	Braenderup	6,7	e,h	e,n,z15	Braenderup
S7	{6,7,14}{54}	g,m,[p],s	[1,2,7]	Montevideo	6,7	g,m,s	-	Montevideo
S8[a)]	1,4,[5],12	i	-	4,5,12:i:-	4,5,12	i	-	4, 5,12:i:-
S9	1,4,12,[27]	b	l,w	Wien	4,12	b	l,w	Wien
S10	6,7,14	r	1,2	Virchow	6,7	r	1,2	Virchow
S11	1,9,12	l,z13	e,n,x	Napoli	9,12	l,z13	e,n,x	Napoli
S12	8,20	z38	-	Apeyeme	8	z38	-	Apeyeme
S13	6,7,14	r	1,5	Infantis	6,7	r	1,5	Infantis
S14	28	z10	e,n,x	Umbilo	28	z10	e,n,x	Umbilo
S15	6,8	z10	e,n,x	Hadar	6,8	z10	e,n,x	Hadar
S16	4,[5],12	a	1,7	Arechavaleta	4,5,12	a	1,2	Kisangani
S17	3,{10}{15}{15,34}	e,h	1,6	Anatum	3,10	e,h	1,6	Anatum
S18	6,8,20	r,[i]	1,5	Bovismorbificans	6,8	r	1,5	Bovismorbificans
S19	1,4,[5],12	i	1,2	Typhimurium	4,5,12	i	1,2	Typhimurium
S20	1,13,23	z29	-	Cubana	23	z29	-	Cubana
S21[b)]	50	r	1,5,(7)	50:r:1,5 (IIIb)	50	r	1,5,7	Salmonella enterica subspecies diarizonae serovar 50:r:1,5,7

a) Typhimurium, monophasic variant as determined by PCR.

b) *Salmonella enterica* subspecies *diarizonae*

Results

EURL-*Salmonella* PT Serotyping 2021

Remarks Lab 1: S8 is a *Salmonella* Typhimurium monophasic variant.
S5 was difficult to find the correct 2nd flagellar phase, due to agglutinate with H2 and H:5 antisera simultaneously.
Due to problems with the culture media of biochemicals, the results for S21 were sent by email on 26 January 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 26th EURL-*Salmonella* Workshop (28 May 2021, online), Strain S-21 was an additional strain to the study.
Testing of this strain was optional and results were not included in the evaluation (remarks in blue or grey only).
The evaluation of the serotyping results was performed as indicated in Table 1 of the Protocol as sent to the participants.
In addition to that, Good Performance was evaluated on the basis of penalty points as indicated below.

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

Good Performance is defined as < 4 penalty points.

Annex 2 Serotyping results per strain and per laboratory

Lab:	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
REF	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
1	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4, 5,12:i:-	Wien	Virchow
2	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	Typhimurium, monophasic (4,5,12:i:-)	Wien	Virchow
3	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
4	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	Monophasic Salmonella typhimurium	Wien	Virchow
5	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
6	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12 : i : -	Wien	Virchow
7	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
8	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	monophasic Typhimurium	Wien	Virchow
9	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,[5],12:i:-	Wien	Virchow
10	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
11	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,12:i:-	Wien	Virchow
12	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:e,n,x	Wien	Virchow
13	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12: i: -	Wien	Virchow
14	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
15	Enteritidis	Chingola	Wagenia	Brandenburg	/	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
16	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
17	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
18	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
19	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
20	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
21	Enteritidis	Chingola	Wagenia	Brandenbourg	Orion	Braenderup	Montevideo	4,5:i:-	Wien	Virchow
22	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
23	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4:i:-	Wien	Virchow
24	Enteriditis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
25	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
26	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	Typhimurium monophasic Variant	Wien	Virchow
27	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	Monophasic 4,5,12:i:-	Wien	Virchow
28	Enteritidis	Chingola	Abony	Brandenburg	Orion	Braenderup	Montevideo	1,4,5,12:i:-	Wien	Virchow
29	Enteritidis	Chingola	Wagenia	Kimuenza	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
30	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:-	Wien	Virchow
31	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	1,4,[5],12:i:-	Wien	Virchow
32	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5:i:-	Wien	Virchow
33	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4, 5, 12: i: -	Wien	Virchow
34	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:- (mST)	Wien	Virchow
35	Enteritidis	Chingola	Wagenia	Brandenburg	Orion	Braenderup	Montevideo	4,5,12:i:- (monophasic Typhimurium)	Wien	Virchow
X	0	0	1	1	0	0	0	1	0	0

S11 Napoli	S12 Apeyeme	S13 Infantis	S14 Umbilo	S15 Hadar	S16 Arechavaleta	S17 Anatum	S18 Bovismorbificans	S19 Typhimurium	S20 Cubana	Lab: REF
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Kisangani	Anatum	Bovismorbificans	Typhimurium	Cubana	1
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Archavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	2
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	3
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	4
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	5
Nordrhein	Albany	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	6
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	7
Lomalinda	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	8
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	9
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	10
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	11
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Archavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	12
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	13
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	14
Lomalinda	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	15
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	16
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	17
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	18
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	19
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	20
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	-:r:1,5	Typhimurium	Cubana	21
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	22
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	23
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	24
Napoli	8,20:HME:-	Infantis	Djibuti	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	13,23:HME:-	25
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	26
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	27
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	28
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubaba	29
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	30
Napoli	Apeyene	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	31
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	32
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	33
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	34
Napoli	Apeyeme	Infantis	Umbilo	Hadar	Arechavaleta	Anatum	Bovismorbificans	Typhimurium	Cubana	35
3	1	0	1	0	1	0	0	0	0	**X**

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

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

Results for Strain S21 are given in Annex 4.

Annex 3 Details per strain that caused problems in serotyping

Strain code	O-antigens	H-antigens		Serovar	Lab code
		(phase 1)	(phase 2)		
S-3	1,4,12,27	b	e,n,z15	Wagenia	REF
S-3	4,12,27	b	e,n,x	Abony	28
S-4	4,[5],12	l,v	e,n,z15	Brandenburg	REF
S-4	4	l,v	e,n,z15	Brandenbourg	21
S-4	4	l,v	e,nZ15	Brandenburg	23
S-4	4,12	l,v	e,n,x	Kimuenza	29
S-5	3,{10}{15}{15,34}	y	1,5	Orion	REF
S-5	3,10	y	/	/	15
S-6	6,7,14	e,h	e,n,z15	Braenderup	REF
S-6	6,7	e,h	e,nZ15	Braenderup	23
S-8	1,4,[5],12	i	-	4,5,12:i:-	REF
S-8	4,5,12	i	e,n,x	4,5,12:i:e,n,x	12
S-8	4,5,12	i	1,2	4,5,12:i:- (monophasic Typhimurium)	35
S-11	1,9,12	l,z13	e,n,x	Napoli	REF
S-11	9,46	l,z13,z28	e,n,z15	Nordrhein	6
S-11	9	a	e,n,x	Lomalinda	8
S-11	9	a	e,n,x	Lomalinda	15
S-12	8,20	z38	-	Apeyeme	REF
S-12	8,20	z4,z24	-	Albany	6
S-12	8,20	HME	-	8,20:HME:-	25
S-14	28	z10	e,n,x	Umbilo	REF
S-14	17	z10	e,n,x	Djibuti	25
S-15	6,8	z10	e,n,x	Hadar	REF
S-15	6,8	Z10	e,n,x	Hadar	11
S-16	4,[5],12	a	1,7	Arechavaleta	REF
S-16	4,5,12	a	1,2	Kisangani	1
S-16	4,5,12	a	1,7	Archavaleta	2
S-16	4,5,12	a	1,7	Archavaleta	12
S-18	6,8,20	r,[i]	1,5	Bovismorbificans	REF
S-18	-	r	1,5	-:r:1,5	21
S-18	6,8	r	s	Bovismorbificans	32
S-20	1,13,23	z29	-	Cubana	REF
S-20	13,23	HME	-	13,23:HME:-	25

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	50	r	1,5,(7)	50:r:1,5,(7) (IIIb)	REF
S-21	50	r	1,5,7	Salmonella enterica subspecies diarizonae serovar 50:r:1,5,7	1
S-21	50	r	1,5,7	Salmonella enterica subspecies diarizonae serovar 50:r:1,5,7	2
S-21	50	r	1,5,7	50:r:1,5,7	3
S-21	50	r	1,5,7	III b	4
S-21	50	r	1,5	IIIb 50:r:1,5	5
S-21	OME	r	1,5,7	OME : r : 1,5,7 (IIIb)	6
S-21	50	r	5	S. IIIb (S. enterica subsp. diarizonae) 50:r:1,5,(7)	7
S-21	-	-	-	-	8
S-21	50	r	1,5,(7)	50:r:1,5,(7)	9
S-21	50 or 61	r	1,5,7	IIIb 50 or 61:r:1,5,7	10
S-21	50	r		S. enterica ssp. diarizonae IIIb	11
S-21	61	r	1,5,7	61:r:1,5,7	12
S-21	50	r	1,5	50:r:1,5 (IIIb)	13
S-21	50	r	1,5,7	50:r:1,5,7	14
S-21					15
S-21	50	r	1,5,7	50:r:1,2,7	16
S-21	50	r	1,5,7	IIIb 50:r:1,5,7	17
S-21	61	r	1,5,7	61:r:1,5,7 (IIIb)	18
S-21	50	r	1,5,7	IIIb 50:r:1,5,7	19
S-21	50	r	1,5,7	50:r:1,5,7 (IIIb)	20
S-21	50	r	1,5,7	IIIb, 50:r:1,5,7	21
S-21	50	r	1,5,7	IIIb 50:r:1,5,(7)	22
S-21	50	r	1,5,7	S.enterica subsp. diarizonae	23
S-21	50	r	1,5,7	50:r:1,5,7	24
S-21	OME	-	-	OME:-:-	25
S-21	50	r	1,5	IIIb 50 : r : 1,5,(7)	26
S-21	50	r	1,5,7	50:r:1,5,7 sg IIIb	27
S-21	50	r	1,5	Salmonella enterica subspecies diarizonae 50:r:1,5	28
S-21	50	l,v	z67	(VI)50:l,v:z67	29
S-21					30
S-21	50	r	1,5(7)	50:r:1,5,(7) - IIIb	31
S-21	61	r	5,7	61:r:5,7 (IIIb)	32
S-21					33
S-21	50	r	1,5,7	50:r:1,5,7 (IIIb)	34
S-21	50	r	1,5,7	IIIb 50:r:1,5,7	35

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

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

Evaluation

EURL-*Salmonella* PT Cluster Analysis 2021

Laboratory code: 6

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 2021).

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 2021 (www.eurlsalmonella.eu)

Did you serotype the strains: Yes
Methodology used: Classical serology

Strain:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**
Expected results:	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
Reported results:	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis

Strain:	**21SCA06[a)] (REF)**	**21SCA07**	**21SCA08**	**21SCA09[a)]**	**21SCA10**
Expected results:	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
Reported results:	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis

[a)]Technical duplicates

Submission of MLVA results: Yes

Strain:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**
Expected results:	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1
Reported results:	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1

Strain:	**21SCA06[a)] (REF)**	**21SCA07**	**21SCA08**	**21SCA09[a)]**	**21SCA10**
Expected results:	3-10-4-4-1	2-14-NA-7-NA	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2
Reported results:	3-10-4-4-1	2-14-NA-7-NA	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2

[a)]Technical duplicates

MLVA-based cluster identification in the PT Typing 2021 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 2021:
***Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1**
The cluster definition for MLVA is set at zero loci with a different number of repeats.

Strain:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
Expected results:	No	No	No	Yes	No	Yes	No	Yes	Yes	No
Reported results:	No	No	No	Yes	No	Yes	No	Yes	Yes	No

MLVA-based cluster identification as expected: Yes
Technical duplicates 21SCA06 and 21SCA09 reported within one cluster: Yes

Submission of WGS results: Yes
WGS platform used: Illumina MiSeq
Analysis used for WGS data: cgMLST-based *(Clustering analysis results by other methods were reported with the lab codes Lab76 and Lab86)*
Tool used for analysis: PyMLST v1
Method used or phylogenetic analysis: hierarchical clustering

	21SCA_REF_R1.fq.gz	21SCA_REF_R2.fq.gz
Expected md5 checksum:	206064240c9fd1f9b89bbd74db4c31d6	4eb99f380cd3d3ca72642079528d4506
Reported md5 checksum:	206064240c9fd1f9b89bbd74db4c31d6	4eb99f380cd3d3ca72642079528d4506

WGS-based cluster identification in the PT Typing 2021 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 2021:
21SCA_REF_R1.fq.gz and **21SCA_REF_R2.fq.gz** (*Salmonella* Enteritidis ST11, MLVA type 3-10-4-4-1)
The cluster definition for WGS is set at maximum 7 allele differences from the reference sequence.

Strain:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
Expected results:	No	No	No	No	No	Yes	No	Yes	Yes	No
Reported results:	No	No	No	No	No	Yes	No	Yes	Yes	No

WGS-based cluster identification as expected: Yes
Technical duplicates 21SCA06 and 21SCA09 reported within one cluster: Yes

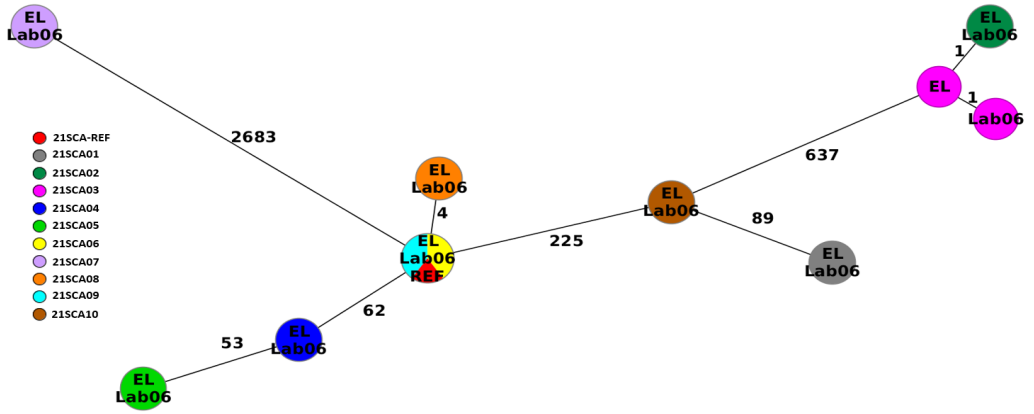


*Figure A6 Minimum Spanning Tree of the participants' results and the EURL-*Salmonella *(EL) results (Ridom SeqSphere+, cgMLST (3002), pairwise ignoring missing values)*

Annex 6 Serotyping results cluster analysis part

Lab code	Serotyping method(s) used	21SCA01	21SCA02	21SCA03	21SCA04	21SCA05	21SCA06	21SCA07	21SCA08	21SCA09	21SCA10
REF	Luminex/in-house Juno pipeline (SeqSero 2)	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
1	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:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-	Enteritidis 9:g,m:-
2	WGS, SISTR Pipeline v1.1.1 in Irida	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Berta|Pensacola|Sangalkam	Enteritidis	Enteritidis	Enteritidis
6	Classical serology	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis
7	Classical serology	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis
11	https://cge.cbs.dtu.dk/services/SeqSero/	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
16	SeqSero2 v1.2.1	Enteritidis	Gallinarum or Enteritidis	Gallinarum or Enteritidis	Enteritidis	Enteritidis	Enteritidis	Gallinarum or Enteritidis	Enteritidis	Enteritidis	Enteritidis
19/89	Classical serology/sistr 1.1.1	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. Enteritidis	S. enteritidis	S. Enteritidis (serology) S. Pensacola, S. Wangata or S. Noya (sistr)	S. Enteritidis	S. Enteritidis	S. Enteritidis
21	Sistr, SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis or Gallinarum	Enteritidis	Enteritidis	Enteritidis
22	WGS Illumina short reads, in-house pipeline based on SeqSero2	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
23	Luminex XMAP	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis
24	BioNumerics 8.0 Salmonella plugin	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis	S. enteritidis
26	http://www.denglab.info/SeqSero/	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis	Enteritidis

In blue: Deviation from the expected result.

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

Labcode	21SCA01	21SCA02	21SCA03	21SCA04	21SCA05
Expected	**2-9-9-4-2**	**2-11-9-3-1**	**2-11-9-3-1**	**3-10-4-4-1**	**3-10-5-4-1**
6	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1
7	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1
11	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1
23	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1
35	2-9-9-4-2	2-11-9-3-1	2-11-9-3-1	3-10-4-4-1	3-10-5-4-1

Labcode	21SCA06[a)]	21SCA07	21SCA08	21SCA09[a)]	21SCA10
Expected	**3-10-4-4-1**	**2-14-NA-7-NA**	**3-10-4-4-1**	**3-10-4-4-1**	**1-10-7-3-2**
6	3-10-4-4-1	2-14-NA-7-NA	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2
7	3-10-4-4-1	2-14 -NA-7-NA	3-10-4-4-1	3-10-4-4-1	NA-10-7-3-2
11	3-10-4-4-1	2-14-NA-7-NA	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2
23	3-10-4-4-1	2-14-0-7-0	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2
35	3-10-4-4-1	2-14-NA-7-NA	3-10-4-4-1	3-10-4-4-1	1-10-7-3-2

[a)] Technical duplicates.

Loci reported in the order: SENTR7-SENTR5-SENTR6-SENTR4-SE-3.

In blue: Deviation from the expected result.

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

Labcode	Wet lab [a)]	WGS platform	Data analysis	Tool for analysis	Method for cluster analysis
1	In-Out-Out	Illumina NovaSeq	SNP-based - assembly-based	CSIPhylogeny 1.4	Maximum likelihood (ML)
2	In-In-In	Illumina MiSeq	SNP-based - assembly-based	In house pipeline[b)]	Maximum likelihood (ML)
6-cgMLST	In-Out-Out	Illumina MiSeq	cgMLST-based	PyMLST v1	hierarchical clustering
6-SNPa	In-Out-Out	Illumina MiSeq	SNP-based - assembly-based	Roary and Prank	Maximum likelihood (ML)
6-SNPr	In-Out-Out	Illumina MiSeq	SNP-based - reference-based	BWA, bcftools, RAxML	Maximum likelihood (ML)
7	In-In-In	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
10-cgMLST	In-Out-Out	Illumina NovaSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
10-SNPr	In-Out-Out	Illumina NovaSeq	SNP-based - reference-based	in-house pipeline iVARCall2	Maximum likelihood (ML)
11	In-In-Out	Illumina NextSeq	SNP-based - reference-based	MINTyper 1.0	Minimum Spanning Tree (MST)
12	In-In-In	Illumina MiSeq	cgMLST-based	Linux cgmlst finder	Minimum Spanning Tree (MST)
14	In-In-In	Illumina MiniSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
16	In-In-In	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
19-cgMLST	In-In-In	Illumina NextSeq	cgMLST-based	inhouse automated ChewieSnake Pipeline	single linkage hierarchical clustering
19-SNPr	In-In-In	Illumina NextSeq	SNP-based - reference-based	inhouse automated SnippySnake Pipeline	single linkage hierarchical clustering
21	In-In-In	Illumina MiSeq	SNP-based - reference-based	In-house pipeline	Minimum Spanning Tree (MST)
22	In-In-In	Illumina NextSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
23	In-In-In	Illumina MiSeq	cgMLST-based	ChewBBaCa	Minimum Spanning Tree (MST)
24	In-In-In	Illumina MiSeq	cgMLST-based	BioNumerics	Minimum Spanning Tree (MST)
26	In-Out-Out	Illumina NovaSeq	cgMLST-based	chewbbaca	Neighbor joining (NJ)
27	In-In-In	Illumina NextSeq	SNP-based - reference-based	Snippy, Gubbins	Maximum likelihood (ML)
30	In-In-In	Illumina MiSeq	SNP-based - assembly-based	CGE CSIPhylogeny 1.4	ML included in CSYPhylogeny

Labcode	Wet lab [a)]	WGS platform	Data analysis	Tool for analysis	Method for cluster analysis
34	In-In-In	Illumina NextSeq	cgMLST-based	Enterobase chewBBACA	MSTreeV2
35	In-In-In	Illumina MiSeq	cgMLST-based	in-house Galaxy pipeline	Neighbor joining (NJ)
EURL-Salm	In-In-In	Illumina NextSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)

[a)] Wet lab preparations: DNA extraction, Library preparation, sequencing. IN: In-house, Out: Outsourced.

[b)] based on parSNP, Gubbins, creating a ML tree in IQTree, creating a SNP distance matrix with snp-dists.

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

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
1	%GC	Spades and Quast	according to the bacterial species
1	coverage	Spades and Quast	30x
1	Genome size	Spades and Quast	Total length not exceding 20% genome size
1	N. contigs	Spades and Quast	N.contigs<500
1	N50	Spades and Quast	N50>15000
1	Number of reads	Command line	e.g. similar number of reads in all samples
1	Preprocessing and percentage after filtering	Trimmomatic	LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36
1	Raw reads quality	FastQC	Quality score of bases
2	%GC	Data from multiQC	Not an exact threshold, but will give you an idea if you have sequenced the right species. For Salmonella ~50-52%
2	Genome coverage	Data exported from fastQC and calculated manually in Excel	Usually aim for about 30X coverage
2	N50	Quast, based on assembly	No absolute threshold for this, but at least 15.000 bp (would probably be a bit sceptical to a dataset with lower than 50.000 bp, but will probably depend on species sequenced)
2	Number of contigs	Quast	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.

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
2	Quality of raw reads	FastQC/MultiQC	Not a real threshold on this, also depending on read length etc but will be evaluated
2	total length of assembly	Quast	If this differs too much from what to expect. We do not have an exact threshold for this, but lean towards suggested from EU-RL AMR +/- 20%
6/76	%GC	Quast	>51 and <53
6/76	assembly length	Quast	>4600000
6/76/86	Contamination	kmerfinder	
6/76	coverage	Quast	>30
6/76	N50	Quast	>150000
6/76	number of contigs	Quast	<100
6/76/86	serotype	SeqSero	
86	Depth	Qualimap	>35X
86	Horizontal coverage	Qualimap	>90%
86	Number of reads	Qualimap	>900000
7	allele calling result - percentage of good targets	Ridom SeqSphere	98%
7	assembly lenght	Ridom SeqSphere	~5MBases for Salmonella
7	coverage	Ridom SeqSphere	minimum 20-30x
7	No. of. contigs	Ridom SeqSphere	200 bases (contigs shorter than 200 have to be ignored)
10/80	Breath coverage	Python	min coverage : 80%
10/80	Contamination	Confidr	around 10% (appreciation)
10/80	Coverage	BBtool	min 30X, max 100X
10/80	De novo assembly	Spades	
10/80	Gap Closing	Gap Closer	

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
10/80	Genome Assembly Evaluation	Quast	
10/80	N50	Quast	Appreciation (no threshold)
10/80	Number of contigs	Quast	Appreciation (no threshold)
10/80	Scaffolding	MeDuSa	Delete scaffolds <200b
10/80	Trimming	Trimmomatic	Min lenght : 50pb, Phred score < 20
11	Average read length	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Should be similar to the expected read length from the sequencing platform.
11	Contamination of genomic sequences by KmerFinder	KmerFinder tool; https://cge.cbs.dtu.dk/services/KmerFinder	Pure bacterial culture
11	Depth of coverage	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Coverage >30x
11	FASTQC- quality parameters	FastQC v0.11.9	All parameters pass
11	N50	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	>30 000 bp
11	Number of reads	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	The number of reads refers to the sequence yield, how much was sequenced. (No criteria established).
11	Size of assembled genome	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	Deviation <0,5 million bp from the expected genome size.
11	Total number of contigs (after assembly)	SPAdes Assembly website; http://cab.spbu.ru/software/spades/	<500 contigs
12	coverage mean	qualimap	>30
12	insert size median	qualimap	340-400
12	N50	quast	>15000
12	number of contigs	quast	<300

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
12	total lenght	quast	4,8-5,3*10exp6
14	Allele calling	cgMLST CT7 Enterobase & cgMLST Statistics	cgMLST alleles found and called > 95%
14	Avg. coverage (assembled)	Assembly Statistics in SeqSphere	50x but if it's less, the % of good targets should be >95%
14	Contamination check	Mash Screen - SeqSphere	identity >=0.95
14	Genome size	Assembly Statistics in SeqSphere	length of contigs assembled < ref genome + 10%
16	Assembled genome size	QUAST	+/-10% of the median genome size for species in NCBI Genome database
16	Contamination	Kraken 2	Majority of taxonomically classified reads should be assigned to the target species.
16	Mean coverage	bwa, samtools, QualiMap	>= 30x (if other criteria are fulfilled, <30x could be accepted)
16	N50	QUAST	>=10 Kbp
19/89	confirmation of serotyping	SISTR	confirmed serotype
19/89	coverage depth	shovill - quast	>30
19/89	duplicated orthologs	shovill - quast	almost no duplicated orthologs
19/89	fraction of reads uniquely assigned to Salmonella enterica	KRAKEN	> 0.90
19/89	number of contigs	shovill - quast	> 200
19/89	predicted species	mash	Salmonella species
19/89	Q30 base fraction	fastp	> 0.80
19/89	single copy orthologs (genome completeness)	shovill - quast	nearly all single copy orthologs
19/89	total length	shovill - quast	4.5-5.5 Mb for Salmonella
21	Assembly size	Python SeqIO package	Warning if not between 4,25 and 5,75 Mbp

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
21	Contamination	Kraken2	Warning if < 80 % Salmonella
21	Coverage	Estimation using linux commands paste, cut and wc	> 10x
21	Raw data quality	Python SeqIO package	Average phred quality > 20
21	Read length	Trimmomatic	36 bases
22	average coverage	Bbtools	>30
22	completeness	CheckM	>96
22	contamination	CheckM	<4
22	GC%	QUAST	51.6 - 52.3
22	N50	QUAST	>20000
22	number of contigs >0bp	QUAST	<300
22	phred score	FastQC	>30
22	total length of filtered scaffolds	QUAST	4.54 - 5.21 Mb
23	contamination	mash	best hit has to match with expected species
23	Cumulative Assembly Len	in house python script	4Mbase < assembly Len < 6Mbases
23	N50	in house python script	N50 > 250000
23	num of Contigs	in house python script	num of contigs < 200
24	Core	BioNumerics 8.0	95
24	Coverage	BioNumerics 8.0	>30
24	Genome size (contamination)	BioNumerics 8.0	4.5 to 5.4
24	N50	BioNumerics 8.0	> 15,000
24	Number of Contigs	BioNumerics 8.0	< less or equal to 400
26	basic statistics fastqc (for the reads)	fastqc	pass

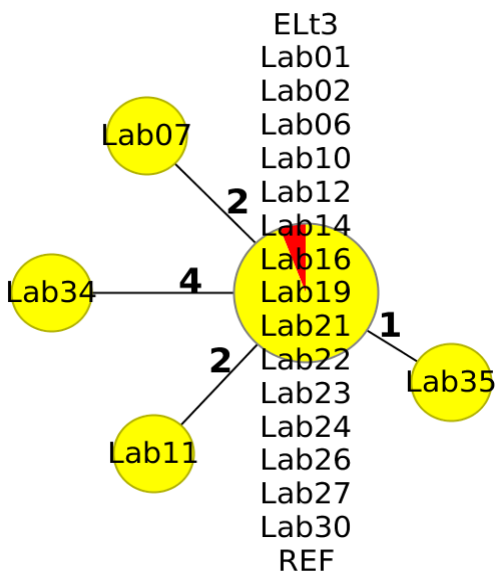
Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
26	coverage	spades	>60
26	GC %	quast	52
26	N50	quast	>50000
26	number of contigs (>1000)	quast	<50
26	total length	quast	4.6-5 Mb
27	contamination	KmerID	>75%
27	Coverage	Qualimap	>30
27	N50	Quast	>25000
27	Number of contigs	Quast	<600
27	Serotyping	SeqSero2	serovar identification
30	Coverage	Manual calculation	Minimum 20X
30	GC%	Quast	Approx. 52%
30	Genome size	Quast	4,5-5 Gbp
34	Coverage	custom	>30
34	Number of bases	seqkit	4-5.8 Mbp
34	Number of contigs	seqkit	<600
34	phred quality	fastp,fastqc	>=20
35	Contamination check	Kraken2	>5% contaminating species = fail
35	fastQC	fastQC	
35	Median coverage	bowtie2 map 2.4.1 - samtools 1.9	>20
35	N50	Quast	
35	Total lenght	Quast	around 5Mpb
35	Total number of contigs	Quast	<500
EL	% good targets cgMLST	Ridom SeqSphere	>95%
EL	Contamination	CheckM	<4%

Lab code	Criterion	Tool (if applicable)	Threshold (if applicable)
EL	Coverage	Formula: (total reads * length of read)/length of genome sequenced	>30
EL	N50	QUAST	
EL	Number of contigs	QUAST	<300
EL	Total length assembly	Ridom SeqSphere / QUAST	4.5 - 5.2 Mbases

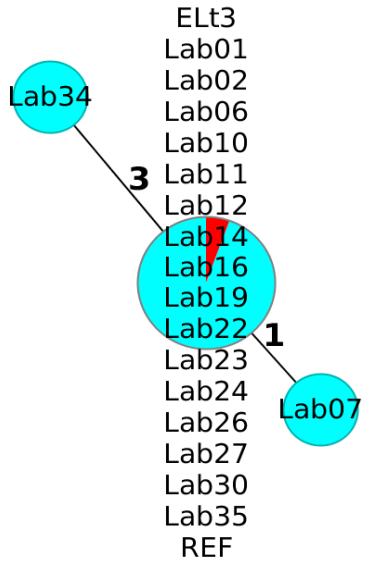
Annex 10 WGS results cluster analysis part, Minimum Spanning Tree per strain

MST for each strain, using all participants' raw data, processed with the in-house developed Juno-assembly pipeline (Ridom SeqSphere+, cgMLST (3002), pairwise ignoring missing values). Data for strain 21SCA09-Lab21 were excluded due to a 85.2% of good targets, which is below our quality threshold.

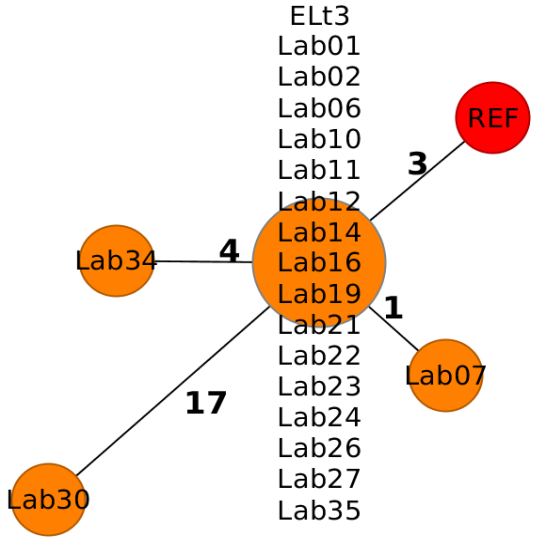
21SCA06 plus REF

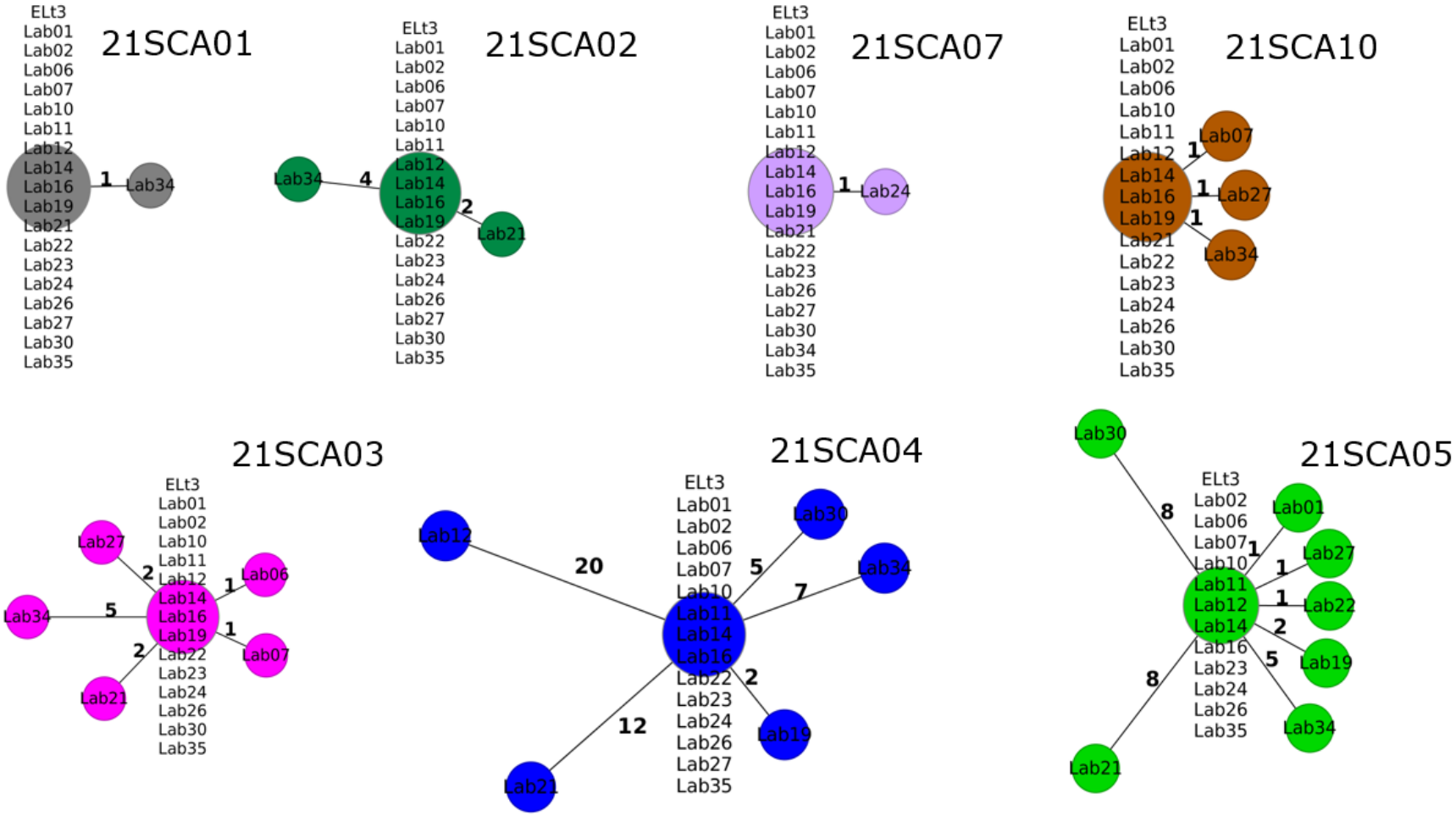


21SCA09 plus REF



21SCA08 plus REF




21SCA01
ELt3
Lab01
Lab02
Lab06
Lab07
Lab10
Lab11
Lab12
Lab14
Lab16
Lab19
Lab21
Lab22
Lab23
Lab24
Lab26
Lab27
Lab30
Lab35
1
Lab34
21SCA02
4
2
21SCA07
21SCA10
Lab07
Lab27
21SCA03
Lab06
5
21SCA04
20
12
7
21SCA05
8

Annex 11 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: NovaSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	23	30	28	31	30	23	53	23	23	22
Largest contig	1508903	1082370	1082372	1326024	1325852	1549186	436389	1549198	1549186	1509370
Total length	4696494	4911965	4911812	4749135	4714992	4701429	4620448	4701715	4701935	4696867
GC (%)	52,14	52,2	52,2	52,08	52,14	52,13	52,08	52,13	52,13	52,14
N50	491595	406594	479293	489733	401324	491608	189262	489949	489949	492165
Input read pairs	8750052	10862818	9484420	9538320	11024518	11575668	11806234	11693080	11368226	9195588
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	279,5	331,7	289,6	301,3	350,7	369,3	383,3	373,0	362,7	293,7

	Laboratory code: 02			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	25	44	31	30	29	23	54	22	20	24
Largest contig	1509211	679094	1082372	1326024	1325852	1549641	436489	1549198	1549641	1508530
Total length	4697307	4912749	4912391	4748991	4714616	4702565	4620723	4702857	4702977	4697919
GC (%)	52,14	52,2	52,2	52,07	52,14	52,13	52,09	52,13	52,13	52,14
N50	491595	349138	514862	491392	401183	491608	189262	491608	491608	490334
Input read pairs	1412082	1065676	906036	842140	785450	1011404	843364	1114648	1111244	1005310
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	90,2	65,1	55,3	53,2	50,0	64,5	54,8	71,1	70,9	64,2

	Laboratory code: 06			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	55	35	47	72	38	79	107	59	83	82
Largest contig	1231166	794676	463821	604137	477202	340955	310585	448724	282270	371177
Total length	4712320	4911980	4908812	4753617	4716623	4698233	4634810	4712382	4714986	4710472
GC (%)	52,11	52,2	52,21	52,08	52,13	52,16	52,08	52,12	52,13	52,13
N50	438799	433357	221071	140918	326440	129424	88195	172165	114911	96984
Input read pairs	1239532	949908	1040198	1274364	1254250	1253964	1328944	1177974	1224044	1392606
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	78,9	58,0	63,6	80,4	79,8	80,1	86,0	75,0	77,9	88,7

	Laboratory code: 07			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	126	87	109	124	120	161	162	121	130	58
Largest contig	316852	406160	228400	244923	244947	198730	164473	233180	167897	529865
Total length	4693797	4914639	4913751	4748477	4714842	4695850	4610601	4698396	4697973	4697827
GC (%)	52,16	52,19	52,2	52,09	52,17	52,16	52,13	52,16	52,15	52,14
N50	75675	128430	97432	83100	84821	58465	60973	80534	68643	172347
Input read pairs	1196488	1699690	1776736	1453104	1434392	1245746	1103468	1080850	1167512	1869612
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	76,5	103,8	108,5	91,8	91,3	79,6	71,8	69,0	74,6	119,4

	Laboratory code: 10			Platform used: NovaSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	26	25	26	31	31	23	50	22	22	24
Largest contig	1509311	1623998	1624000	1326024	1325852	1549186	436489	1549198	1549186	1506934
Total length	4697466	4913305	4913369	4750230	4716523	4702335	4621489	4702731	4702835	4695438
GC (%)	52,14	52,2	52,2	52,08	52,14	52,13	52,09	52,13	52,13	52,14
N50	490088	514862	514862	489733	401283	489949	189262	489950	489950	490334
Input read pairs	33172952	30788616	31808594	32254170	37211112	36011942	40722300	36513482	38846002	42966020
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	1059,3	940,0	971,1	1018,5	1183,4	1148,7	1321,7	1164,6	1239,0	1372,6

	Laboratory code: 11			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	33	63	51	47	44	51	71	34	35	39
Largest contig	950810	478486	752107	910805	650329	411592	435916	927450	1278502	603565
Total length	4688088	4893448	4894437	4740834	4704162	4692072	4604164	4692155	4694215	4686670
GC (%)	52,13	52,19	52,19	52,06	52,12	52,12	52,08	52,12	52,12	52,13
N50	286684	195685	228116	304337	227854	184068	139265	305434	406239	304491
Input read pairs	6246242	3729004	3134372	3935108	3160026	1623430	4933018	4363034	4008504	3980114
Read Length	75	75	75	75	75	75	75	75	75	75
Coverage	99,9	57,2	48,0	62,3	50,4	25,9	80,4	69,7	64,0	63,7

	Laboratory code: 12			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	92	138	25	88	28	24	50	2828	56	32
Largest contig	767670	305084	1624000	323757	1325852	1549741	668735	694610	1549741	1508530
Total length	4996759	5014420	4915968	4747177	4717881	4706997	4626608	7795125	4738026	4703569
GC (%)	52,13	52,13	52,19	52,08	52,13	52,12	52,08	51,34	52,12	52,12
N50	433478	167744	514862	105208	433809	489949	194101	196100	489949	490334
Input read pairs	2096462	1795292	1768638	1524504	1778286	1849904	2229152	1894094	2103072	1697532
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	125,9	107,4	107,9	96,3	113,1	117,9	144,5	72,9	133,2	108,3

	Laboratory code: 14			Platform used: MiniSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	23	28	30	33	36	25	53	26	28	24
Largest contig	1509124	1049328	1082372	1325850	1320934	1548892	436257	781216	1548345	937586
Total length	4696555	4910184	4910582	4749530	4714094	4700439	4617635	4701266	4700394	4697010
GC (%)	52,14	52,2	52,19	52,08	52,14	52,13	52,08	52,13	52,13	52,14
N50	491595	680848	679541	489733	373261	489839	188932	694233	479220	490255
Input read pairs	1749252	1536152	1741588	1692544	1252596	1342210	1273680	1537362	1103326	1437394
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	55,9	46,9	53,2	53,5	39,9	42,8	41,4	49,1	35,2	45,9

	Laboratory code: 16			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	29	23	26	28	31	19	47	21	18	26
Largest contig	1509311	1625663	1082372	1326038	1325852	1549641	437623	1549198	1549186	1508530
Total length	4696820	4913357	4912603	4751650	4715438	4702926	4621615	4702147	4704018	4698175
GC (%)	52,14	52,2	52,2	52,08	52,14	52,13	52,09	52,13	52,13	52,14
N50	490085	514862	514862	490160	401043	491608	189262	489949	489949	490334
Input read pairs	1422610	1539100	1443536	974136	1969410	995444	1580342	1168020	1832400	1912444
Read Length	250	250	250	250	250	250	250	250	250	250
Coverage	75,7	78,3	73,5	51,3	104,4	52,9	85,5	62,1	97,4	101,8

	Laboratory code: 19			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	30	33	40	33	41	28	58	28	26	30
Largest contig	937677	752319	752351	1325850	537061	1498454	436257	1548827	1548761	1222118
Total length	4695348	4910876	4909414	4751280	4712725	4701028	4620093	4700314	4700672	4695216
GC (%)	52,14	52,19	52,2	52,07	52,13	52,13	52,09	52,13	52,13	52,14
N50	421579	406496	371272	421989	239019	421481	153434	410690	479182	406556
Input read pairs	3220016	2673576	2527662	2754858	1818086	2313454	2176976	2255252	2690098	2786864
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	102,9	81,7	77,2	87,0	57,9	73,8	70,7	72,0	85,8	89,0

	Laboratory code: 21			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	153	179	188	167	217	134	202	128	653	72
Largest contig	227645	178000	204557	197937	107441	276626	146806	269282	66564	406754
Total length	4858731	4907045	4957609	4741265	4721943	4707702	4631802	4855826	4542334	4697412
GC (%)	52,15	52,21	52,2	52,09	52,16	52,13	52,09	52,12	51,97	52,14
N50	62131	55573	48112	51299	44236	68435	51719	85790	15296	141701
Input read pairs	2813526	1847410	1540804	1028836	1077206	1838548	1903014	4237700	698348	3150532
Read Length	250	250	250	250	250	250	250	250	250	250
Coverage	144,8	94,1	77,7	54,2	57,0	97,6	102,7	218,2	38,4	167,7

	Laboratory code: 22			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	25	32	30	32	33	27	53	26	24	24
Largest contig	1508903	1082370	1082372	1326024	1325852	1549186	436489	1549198	1549186	1509370
Total length	4697258	4911305	4912048	4748379	4714633	4702392	4621097	4701917	4701739	4696622
GC (%)	52,14	52,2	52,2	52,07	52,14	52,13	52,08	52,13	52,13	52,14
N50	491595	445656	445655	491371	400934	491608	189064	491609	489949	421859
Input read pairs	5074994	5920996	5014052	4934286	4848242	5757656	5018048	4858722	4712666	4421210
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	162,1	180,8	153,1	155,9	154,3	183,7	162,9	155,0	150,3	141,2

	Laboratory code: 23			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	32	33	33	39	36	30	78	33	27	24
Largest contig	919735	903416	808657	759061	1062648	812012	406539	695336	1271382	889568
Total length	4697424	4911234	4912606	4749981	4716963	4701978	4629890	4704579	4704247	4697738
GC (%)	52,14	52,2	52,2	52,08	52,14	52,13	52,1	52,13	52,13	52,14
N50	306666	326927	386262	268453	401283	406274	142230	401322	328928	433615
Input read pairs	2337312	2308328	2185440	1867470	2408638	1928052	2061334	2426916	2247360	2078224
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	149,3	141,0	133,5	117,9	153,2	123,0	133,6	154,8	143,3	132,7

	Laboratory code: 24			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	25	31	322	424	351	39	329	243	369	33
Largest contig	1239819	779012	151584	98375	87556	717776	89259	153102	80228	1234877
Total length	4697258	4911845	4936808	4780733	4741126	4699041	4643325	4721812	4771010	4694055
GC (%)	52,14	52,2	52,2	52,09	52,15	52,14	52,09	52,13	52,13	52,15
N50	491595	406694	29354	22279	27284	401322	26810	33724	25860	490333
Input read pairs	1520670	2135466	1623192	1490508	1383174	989114	1348968	1499840	1358952	1484170
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	97,1	130,4	98,6	93,5	87,5	63,1	87,2	95,3	85,5	94,9

	Laboratory code: 26			Platform used: NovaSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	26	28	29	31	30	24	54	26	24	26
Largest contig	1508903	1082370	1082372	1326024	1325852	1549186	436489	1549198	1549186	1351536
Total length	4695690	4912076	4912820	4749192	4714535	4702483	4621261	4702434	4701797	4696061
GC (%)	52,14	52,2	52,2	52,07	52,14	52,13	52,09	52,13	52,13	52,14
N50	491595	479293	479393	491393	401283	489949	189262	421596	421593	492165
Input read pairs	8454578	7205650	8066862	8308166	7612582	7886382	8778214	8625788	9132618	8966492
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	270,1	220,0	246,3	262,4	242,2	251,6	284,9	275,1	291,4	286,4

	Laboratory code: 27			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	24	40	142	58	136	34	55	30	29	37
Largest contig	1509036	545920	227998	889700	198494	1386330	435925	880988	1549000	660080
Total length	4696591	4910092	4902261	4770997	4703513	4701330	4620491	4701492	4701073	4696009
GC (%)	52,14	52,2	52,21	52,06	52,18	52,13	52,09	52,13	52,13	52,14
N50	421560	443033	75154	367537	67319	376073	183928	421588	421590	264597
Input read pairs	3659638	3325384	1213746	3599198	2494684	3397018	2768630	2850932	2938118	2610918
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	116,9	101,6	37,1	113,2	79,6	108,4	89,9	91,0	93,7	83,4

	Laboratory code: 30			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	23	29	32	103	43	25	53	50	24	24
Largest contig	1509211	994304	779195	1435991	995247	1549186	436489	469980	1549186	1508530
Total length	4698222	4911794	4911541	4803317	4715446	4703476	4619559	4702638	4702923	4697667
GC (%)	52,14	52,2	52,2	51,97	52,14	52,13	52,08	52,13	52,13	52,14
N50	491595	406694	406258	401034	228173	489949	189262	209817	489950	490375
Input read pairs	1693392	1614998	1444202	602994	541528	1439644	1047936	500450	809232	1004732
Read Length	300	300	300	300	300	300	300	300	300	300
Coverage	108,1	98,6	88,2	37,7	34,5	91,8	68,1	31,9	51,6	64,2

	Laboratory code: 34			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	93	138	165	207	98	128	57	138	107	85
Largest contig	310735	242307	180668	134957	321994	177048	436489	153302	209119	325769
Total length	4695168	4904371	4902624	4745604	4711770	4701474	4618243	4699834	4700826	4693570
GC (%)	52,14	52,19	52,19	52,07	52,13	52,12	52,08	52,13	52,13	52,14
N50	116919	72807	50753	35897	99751	63551	153412	59415	74193	101594
Input read pairs	2016822	2186248	1587408	1418768	2210850	1779706	4764142	1734802	1995992	2645486
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	64,4	66,9	48,6	44,8	70,4	56,8	154,7	55,4	63,7	84,5

	Laboratory code: 35			Platform used: MiSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	28	24	45	32	24	23	50	21	26	27
Largest contig	1509311	1623998	1624000	1326024	1325852	1549467	668635	1549753	1549186	1508530
Total length	4700791	4919170	4928040	4758702	4716414	4702934	4619883	4706063	4707598	4703824
GC (%)	52,14	52,19	52,24	52,07	52,14	52,13	52,08	52,13	52,13	52,13
N50	490086	514862	994204	489733	433415	438896	189031	490376	490376	490334
Input read pairs	1767870	1434062	1152878	1158186	1373168	763204	1062288	1301766	1208832	1433172
Read Length	250	250	250	250	250	250	250	250	250	250
Coverage	94,0	72,9	58,5	60,8	72,8	40,6	57,5	69,2	64,2	76,2

	Laboratory code: EL			Platform used: NextSeq						
Parameters:	**21SCA01**	**21SCA02**	**21SCA03**	**21SCA04**	**21SCA05**	**21SCA06**	**21SCA07**	**21SCA08**	**21SCA09**	**21SCA10**
# contigs	25	30	32	32	31	27	55	25	25	28
Largest contig	1508903	779012	779195	1326024	1325852	1549186	436489	1548849	1549012	1345851
Total length	4696151	4911728	4911404	4748963	4715577	4702284	4621938	4701797	4700638	4697272
GC (%)	52,14	52,2	52,2	52,07	52,14	52,13	52,08	52,13	52,13	52,14
N50	478804	406595	388225	491393	401183	421587	189064	489949	491609	490334
Input read pairs	4719760	6919986	5022402	6287956	4752080	5011936	3981888	3482630	4269944	3320196
Read Length	150	150	150	150	150	150	150	150	150	150
Coverage	150,8	211,3	153,4	198,6	151,2	159,9	129,2	111,1	136,3	106,0

Annex 12 WGS results cluster analysis part, Per submission, the participants' distance matrix data for their comparison of the 21SCA-REF with the ten tested strains

Labcode-method	21 SCA-REF	21 SCA01	21 SCA02	21 SCA03	21 SCA04	21 SCA05	21 SCA06	21 SCA07	21 SCA08	21 SCA09	21 SCA10	21SCA08 (Table 4.8)
1-SNPa	0	427	1444	1444	136	152	5	29352	11	4	444	No
2-SNPa	0	427	1461	1461	133	149	1	40860	10	1	450	Yes
6-SNPa	0	288	974	977	97	106	0	29800	3	0	302	Yes
30-SNPa	0	420	1435	1436	142	166	0	29291	44	2	441	No
6-SNPr	0	453	1563	1561	140	157	0	40941	9	0	481	Yes
10-SNPr	0	455	1571	1567	136	155	2	45908	10	1	476	Yes
11-SNPr	0	383	1361	1360	129	147	0	29299	9	0	414	Yes
19-SNPr	0	431	1464	1461	133	146	0	44867	9	0	449	No
21-SNPr	0	376	1265	1262	105	126	0	37250	8	0	390	No
27-SNPr	0	401	1378	1378	126	141	1	39910	11	1	424	No
6-cgMLST	0	233	682	680	69	83	0	2745	6	0	234	Yes
7-cgMLST	0	232	681	681	67	82	0	2749	4	0	232	Yes
10-cgMLST	0	232	680	680	67	81	0	2746	4	0	232	Yes
12-cgMLST	0	243	707	708	76	89	2	2767	9	3	248	Yes
14-cgMLST	0	232	680	680	67	81	0	2745	4	0	232	Yes
16-cgMLST	0	232	681	680	67	81	0	2750	4	0	232	Yes
19-cgMLST	0	233	691	692	70	85	1	2716	9	1	238	No
22-cgMLST	0	232	680	681	67	82	0	2750	4	0	232	Yes
23-cgMLST	0	228	691	691	69	85	0	2703	6	0	232	Yes
24-cgMLST	0	200	200	200	69	86	0	200	5	0	200	Yes

Labcode-method	21 SCA-REF	21 SCA01	21 SCA02	21 SCA03	21 SCA04	21 SCA05	21 SCA06	21 SCA07	21 SCA08	21 SCA09	21 SCA10		21SCA08 (Table 4.8)
26-cgMLST	0	314	928	926	109	127	4	3467	8	3	318		No
34-cgMLST	0	236	697	700	72	86	1	2737	10	2	241		No
35-cgMLST	0,0000	0,0825	0,2425	0,2415	0,0257	0,0307	0,0000	0,9289	0,0039	0,0000	0,0825		No
ELt3-cgMLST	0	232	681	681	67	81	0	2750	4	0	232		Yes

SNPa: assembly-based SNP analysis, SNPr: reference-based SNP analysis.

Deviation from the expected result Deviation based on internal result

The last column shows per submission the reported 'yes/no' if a clustering match was found for strain 21SCA08 (Table 4.8). All reported 'yes/no' results were in line with the reported distance matrix results, except for participant 12. The reported nine AD would have been expected as a 'no' instead of a 'yes', because this did not meet the given PT2021 cluster definition of a maximum of seven AD.

Published by:

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

May 2023

Committed to
health and sustainability