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

Results EURL-*Salmonella* PT Typing 2023

serotyping & cluster analysis

Wilma Jacobs-Reitsma
EURL-*Salmonella*
With back-up from our
RIVM typing experts

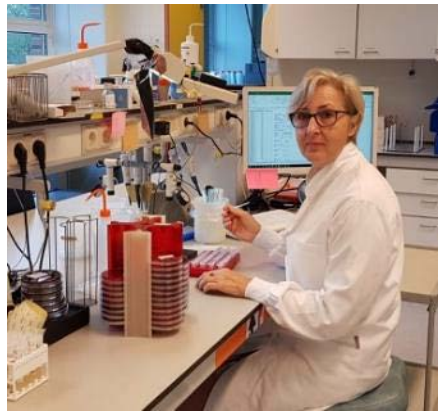
EURL Salmonella

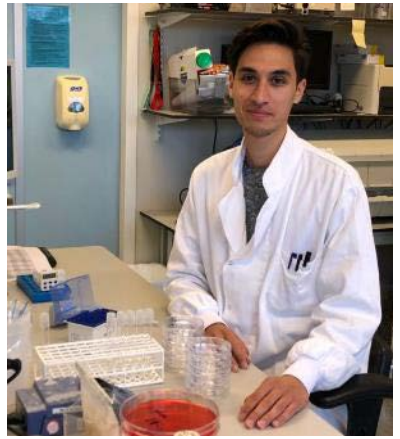
Co-funded by
the European Union

RIVM typing experts

- Serotyping:
  - Anjo Verbruggen
- Cluster Analysis:
  - Maaike van den Beld
  - Maren Lanzl
  - Angela van Hoek
  - Robin Diddens
  - The IDS-bioinformatics team
- Sample preparations:
  - Wendy van Overbeek

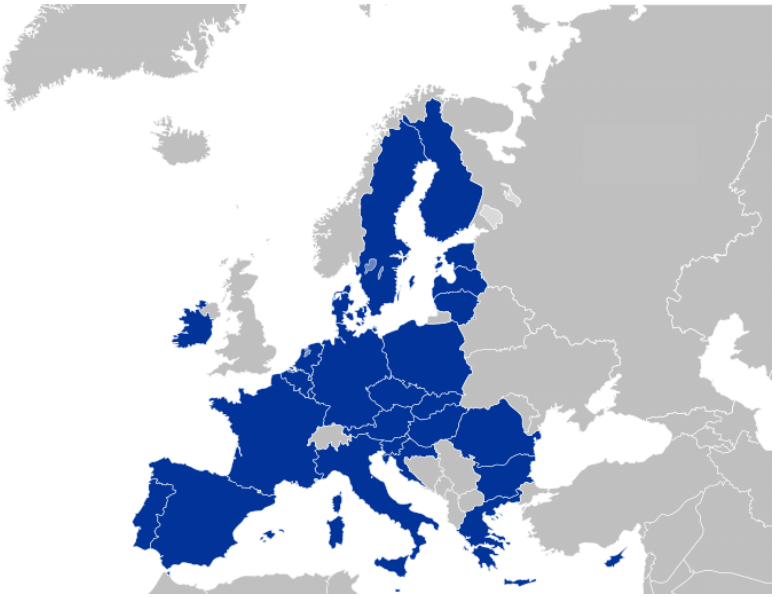
- EURL-*Salmonella* contact typing studies:
  - Wilma Jacobs





Participants PT Typing 2023

- National Reference Laboratories for *Salmonella* from all **27 EU** member states
- Moldova (EU candidate country)
- Iceland, Norway, Switzerland (EFTA countries)
- United Kingdom (third country)



- Serotyping
  - **32** participants
- Optional Cluster Analysis
  - **19** participants

Part on Serotyping (28th PT)

- 20 different serovars of *Salmonella enterica* subsp. *enterica*
- 1 additional strain of an uncommon type (optional)
- Serotyping and reporting in accordance with the White-Kauffmann-le Minor scheme (2007) and the published supplements
  - Supplement 2003–2007 (no. 47) Guibourdenche et al., 2010
  - Supplement 2008-2010 (no. 48) Issenhuth-Jeanjean et al., 2014
- Laboratories have to report only those results, on which the identification of serovar names is based.

Table 1. Examples of expected reporting (or not expected reporting)

O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar name*
9	g,m	-	Enteritidis
4,12	i	2	Typhimurium
4,5,12	i	-	4,5,12:i:-
4	i	-	4:i:-
6,7	-	1,5	6,7:-:1,5
42	g,t	-	42:g,t:-
4,12,27	l,v	e,n,z15	Brandenburg
1,4,[5],12	g,m,s	[1,2]	Hato

In grey: do not take over the notation "as it is" from the WKLM scheme but only report what was tested and detected, e.g. 4:g,m,s:- Hato.

Salmonella strains (1)

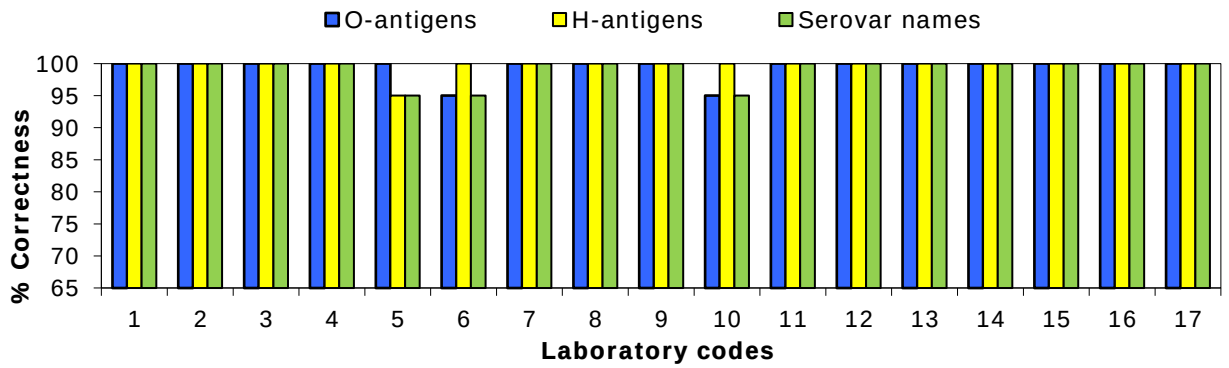
Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar
S12	1,4,[5],12	e,h	e,n,x	Chester
S11	1,4,[5],12	f,g	[1,2]	Derby
S16	**1,4,[5],12**	**i**	**1,2**	**Typhimurium**
S19	**1,4,[5],12**	**i**	**-**	**1,4,[5],12:i:-**
S5*	6,7	l,z_{13},[z_{28}]	e,n,z_{15}	Neukoelln
S4	**6,7,14**	**r**	**1,2**	**Virchow**
S9	**6,7,14**	**r**	**1,5**	**Infantis**
S7	6,7	r	1,7	Colindale
S2	6,7,14	z_{10}	e,n,z_{15}	Mbandaka
S13*	6,8	e,h	e,n,x	Fillmore
S1	**6,8**	**z_{10}**	**e,n,x**	**Hadar**
S3	6,8	z_{10}	e,n,z_{15}	Glostrup

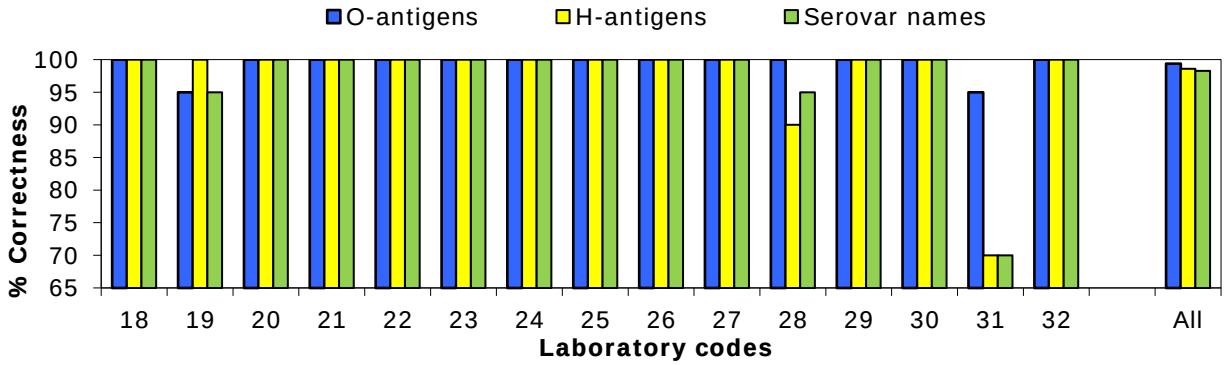
* Represented in an EURL-*Salmonella* PT Serotyping for the first time

Salmonella strains (2)

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar
S6	**1,9,12**	**g,m**	**-**	**Enteritidis**
S18	1,9,12[Vi]	g,p	-	Dublin
S17	9,12	l,v	1,7	Kapemba
S8	3,{10},{15}	l,v	1,6	London
S10	3,10	z_{35}	z_6	Cairina
S15	1,3,19	g,[s],t	-	Senftenberg
S14	13,23	d	l,w	Putten
S20	28	i	1,5	Cotham
S21	41	z	1,5	II 41:z:1,5

Percentages of correct serotyping results, per participant





Correct typing of strains

- 100% accurate serotype naming by 26/32 (81%) participants

- 11 serovars completely correct identified by all participants:
  - Hadar (S1), Virchow (S4), Enteritidis (S6), London (S8), Infantis (S9), Chester (S12), Fillmore (S13), Senftenberg (15), Typhimurium (S16), Kapemba (S17), and 1,4,[5],12:i:- (S19)

Overview

Lab: REF	S1 Hadar	S2 Mbandaka	S3 Glostrup	S4 Virchow	S5 Neukoelln	S6 Enteritidis	S7 Colindale	S8 London	S9 Infantis	S10 Cairina	S11 Derby	S12 Chester	S13 Fillmore	S14 Putten	S15 Senftenberg	S16 Typhimurium	S17 Kapemba	S18 Dublin	S19 1,4,[5],12:i:-	S20 Cotham	Lab: REF
1	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5:i:-	Cotham	1
2	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4:i:-	Cotham	2
3	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	3
4	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	Typhimurium, monophasic (4,5,12:i:-	Cotham	4
5	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Sinchew	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	Monophasic Typhimurium	Cotham	5
6	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Elisabethville	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	6
7	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	7
8	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	8
9	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4:i:-	Cotham	9
10	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	-:i:1,5	10
11	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	11
12	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:- (mST)	Cotham	12
13	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	13
14	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteriditis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	14
15	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	Monophasic variant (4,5,12:i:-)	Cotham	15
16	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	I 4:i:-	Cotham	16
17	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cothan	17
18	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	18
19	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	Typhimurium monophasic Variant	rough : i : 1,5	19
20	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	20
21	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	21
22	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	22
23	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	23
24	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5 : i : -	Cotham	24
25	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	1,4,5,12:i:-	Cotham	25
26	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublín	monophasic Typhimurium	Cotham	26
27	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12:i:-	Cotham	27
28	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kampemba	Rostock	4,5,12:i:-	Cotham	28
29	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	1,4,[5],12:i:-	Cotham	29
30	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	4,5,12: i: -	Cotham	30
31	Hadar	Djudu	Tshiongwe	Virchow	Braenderup	Enteritidis	Colindale	London	Infantis	Cairina	Essen	Chester	Fillmore	Handen	Senftenberg	Typhimurium	Kapemba	Dublin	Typhimurium monofaza	-	31
32	Hadar	Mbandaka	Glostrup	Virchow	Neukoelln	Enteritidis	Colindale	London	Infantis	Cairina	Derby	Chester	Fillmore	Putten	Senftenberg	Typhimurium	Kapemba	Dublin	Monophasic Salmonella typhimurium	Cotham	32
X	0	1	1	0	1	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	**X**

remark (e.g. spelling errror)
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

Details

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-2	**6,7,14**	**z10**	**e,n,z15**	**Mbandaka**	**REF**
S-2	6,7	z29	e,n,z15	Mbandaka	28
S-2	6,7	z10	e,n,x	Djudu	31
S-3	**6,8**	**z10**	**e,n,z15**	**Glostrup**	**REF**
S-3	6,8	e,h	e,n,z15	Tshiongwe	31
S-5	**6,7**	**l,z13,[z28]**	**e,n,z15**	**Neukoelln**	**REF**
S-5	6,7	e,h	e,n,z15	Braenderup	31
S-6	**1,9,12**	**g,m**	**-**	**Enteritidis**	**REF**
S-6	9,12	g,m	-	Enteriditis	14
S-6	9,12	g.m	-	Enteritidis	26
S-7	**6,7**	**r**	**1,7**	**Colindale**	**REF**
S-7	15	r	7	Elisabethville	6
S-10	**3,10**	**z35**	**z6**	**Cairina**	**REF**
S-10	3,10	l,v	z35	Sinchew	5
S-11	**1,4,[5],12**	**f,g**	**[1,2]**	**Derby**	**REF**
S-11	1,4,5,12	f.g	-	Derby	25
S-11	4,12	g,f	-	Derby	30
S-11	4,12	g,m	-	Essen	31
S-13	**6,8**	**e,h**	**e,n,x**	**Fillmore**	**REF**
S-13	6,8	e,h	e,nx	Fillmore	26
S-14	**13,23**	**d**	**l,w**	**Putten**	**REF**
S-14	13,23	d	1,2	Handen	31
S-16	**1,4,[5],12**	**i**	**1,2**	**Typhimurium**	**REF**
S-16	4,5,15	i	1,12	Typhimurium	28
S-17	**9,12**	**l,v**	**1,7**	**Kapemba**	**REF**
S-17	9,12	l,v	1.7	Kapemba	18
S-17	9,12	l,v	1,7	Kampemba	28
S-18	**1,9,12[Vi]**	**g,p**	**-**	**Dublin**	**REF**
S-18	9,12	g,p,u	-	Rostock	28
S-20	**28**	**i**	**1,5**	**Cotham**	**REF**
S-20	-	i	1,5	-:i:1,5	10
S-20	28	i	1,5	Cothan	17
S-20	-	i	1,5	rough : i : 1,5	19
S-20	-	-	-	-	31

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

Additional strain S21

› Examined by 31/32 participants

Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar	Lab code
S-21	**41**	**z**	**1,5**	**II 41:z:1,5**	**REF**
S-21	6,7	z	1,5	6,7:z:1,5	28
S-21	6,7	z	1,5	II 6,7:z:1,5	15
S-21	6,7	z	1,5	Poitiers	32
S-21	18	z36	-	enterica II	31
S-21	41	-	1,5	41 : - :1,5 (II)	26
S-21	-	z	e,n,x,z15	-:z:e,n,x,z15	2
S-21	?	z	5	? : z : 5	24

S-21	41	z	1,5	41:z:1,5	8x

in variations:

S-21	41	z	1,5	*Salmonella enterica* subsp. *salamae* 41:z:1,5 (II)	16x

Evaluation of Performance

- **“Good Performance” (**Workshop Bilthoven, 2007)

- **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 5 serotypes to another strain.
- **1 penalty point:** Incorrect typing of all other *Salmonella* serotypes.

- For each NRL-*Salmonella* the total amount of penalty points is determined.
- “Good Performance” is: less than 4 penalty points.
- A follow-up is obligatory for EU-NRLs with 4 penalty points or more.

Evaluation of Performance

Lab code	Penalty points	Good performance
1	0	yes
2	0	yes
3	0	yes
4	0	yes
5	1	yes
6	1	yes
7	0	yes
8	0	yes
9	0	yes
10	0	yes
11	0	yes
12	0	yes
13	0	yes
14	0	yes
15	0	yes
16	0	yes

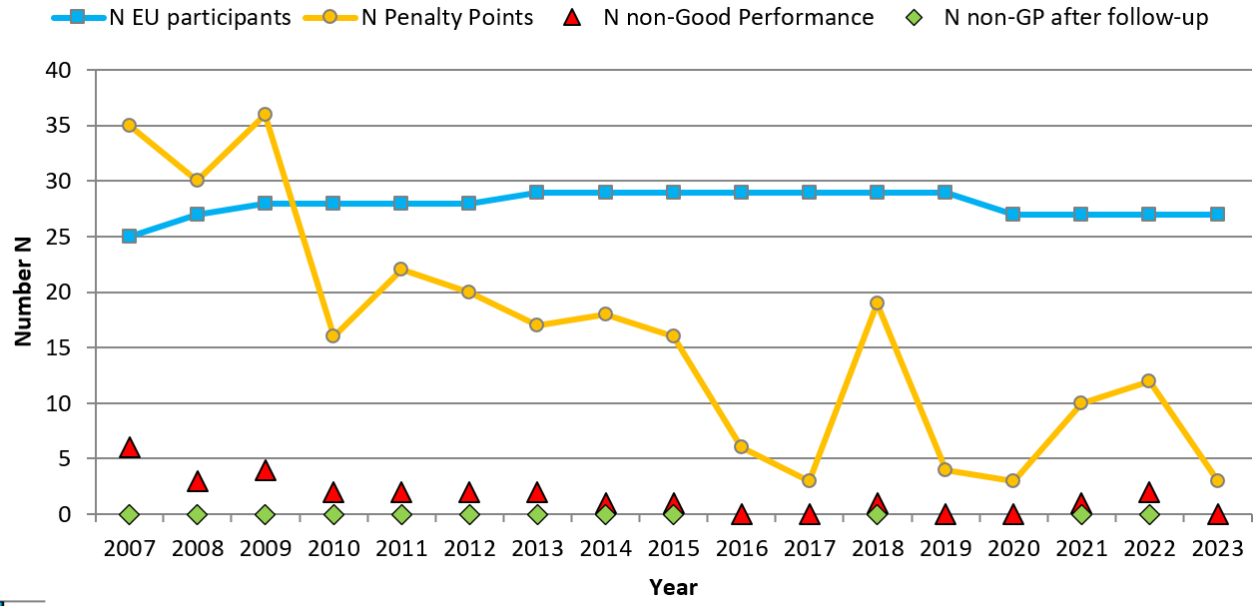
Lab code	Penalty points	Good performance
17	0	yes
18	0	yes
19	0	yes
20	0	yes
21	0	yes
22	0	yes
23	0	yes
24	0	yes
25	0	yes
26	0	yes
27	0	yes
28	1	yes
29	0	yes
30	0	yes
31	5	**NO**
32	0	yes

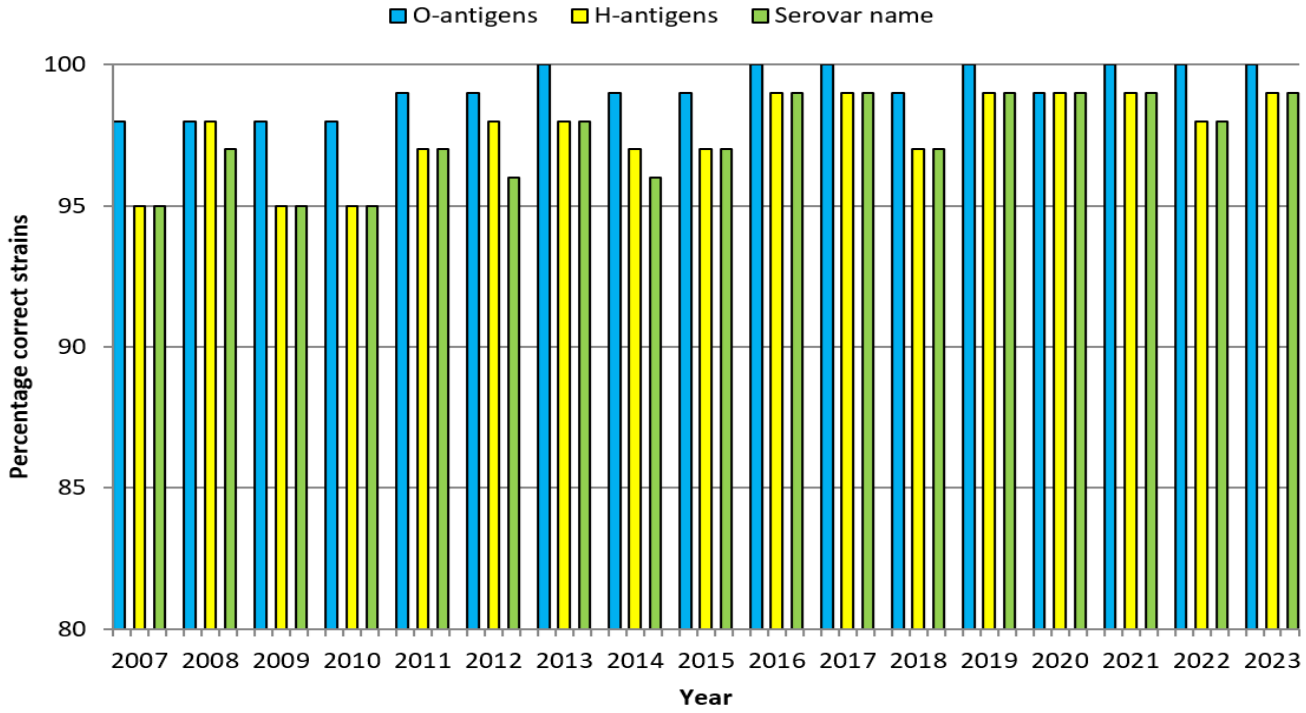
Follow-up study 1 participant

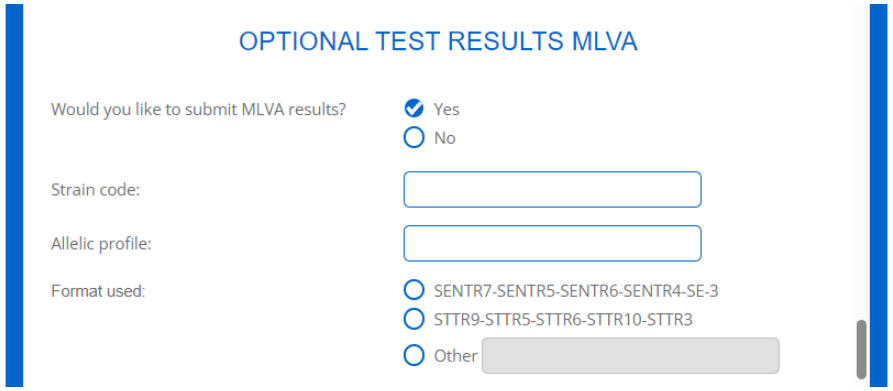
Strain code	O-antigens	H-antigens (phase 1)	H-antigens (phase 2)	Serovar
SF1	1,9,12	g,m	-	Enteritidis
SF2	6,8	z_{10}	e,n,z_{15}	Glostrup
SF3	1,4,[5],12	f,g,s	[1,2]	Agona
SF4	13,23	d	l,w	Putten
SF5	6,7,14	z_{10}	e,n,z_{15}	Mbandaka
SF6	1,4,[5],12	e,h	e,n,z_{15}	Sandiego
SF7	1,4,[5],12	f,g	[1,2]	Derby
SF8	6,7,14	d	l,w	Livingstone
SF9	6,7	l,z_{13},[z_{28}]	e,n,z_{15}	Neukoelln
SF10	1,4,[5],12	i	1,2	Typhimurium

- Evaluation still on-going

Results in time (EU NRLs only)







Optional MLVA results

- Optionally, MLVA on any applicable strain within serotyping part
- Serotyping result form for submission
- Results will be evaluated by a comparison among the MLVA participants

- **One** participant entered results for strains S6 (SE), S16 (STM), and S19 (mSTM)
- Comparison not possible

- **No MLVA options anymore** in future EURL-*Salmonella* PTs Typing
- Any need for NRLs to have the reference sets of MLVA strains available from us?

Conclusions PT serotyping **2023** (all participants)

- O-antigens: 99% of strains typed correctly
- H-antigens: 99% of strains typed correctly
- Serovar names: 98% of strains typed correctly

- 100% accurate serotyping by 26/32 (81%) participants
  - 3 participants with 1 penalty point
  - 1 participant with 5 penalty points
  - (2 participants with ‘rough’ results for strain S20)

- 31 participants “Good Performance”
- 1 participant currently Follow-up

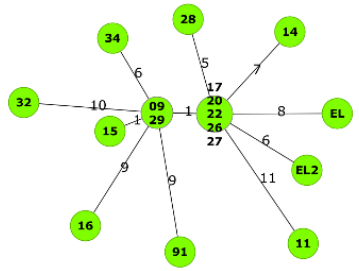
Part on Cluster Analysis (5th PT)

- Cluster analysis using WGS
  - The cluster analysis 2023 is mimicking an outbreak situation, with a *Salmonella* Bovismorbificans strain as the reference
  - Participants' own routine method(s) of choice

- PT Protocol
  - Downloading: fastq.gz files + md5checksums ('dry' part)
  - Uploading: distance matrix, raw reads (fastq.gz files) + (optional) md5checksums
  - Result form: wet-lab/dry-lab protocols, **cluster analysis**

‘Wet’ strains *Salmonella* Cluster Analysis PT 2023

- Pre-testing (WGS at t=0, t=1, t=2)
  - DNA extraction, library preparation, sequencing performed in-house; WGS platform: Illumina NextSeq. Raw data processing: in-house developed Juno-assembly pipeline which includes SPAdes 3.15.3. Cluster analysis: Ridom SeqSphere⁺, using the cgMLST Enterobase v2.0 scheme.
- Strains with stable and consistent cgMLST results selected for inclusion in PT
- Technical duplicates: 23SCA02/23SCA05
  - Shipment tubes prepared from the same blood agar plate culture



'Wet' strain code	Serovar	ST	Background	Origin
23SCA01	Typhimurium	19	2019 SCA03	Human
23SCA02	Bovismorbificans	142	techn duplo REF	Human
23SCA03	Bovismorbificans	142		Human
23SCA04	Bovismorbificans	142		Human
23SCA05	Bovismorbificans	142	REF	Human
23SCA06	Bovismorbificans	142		Human

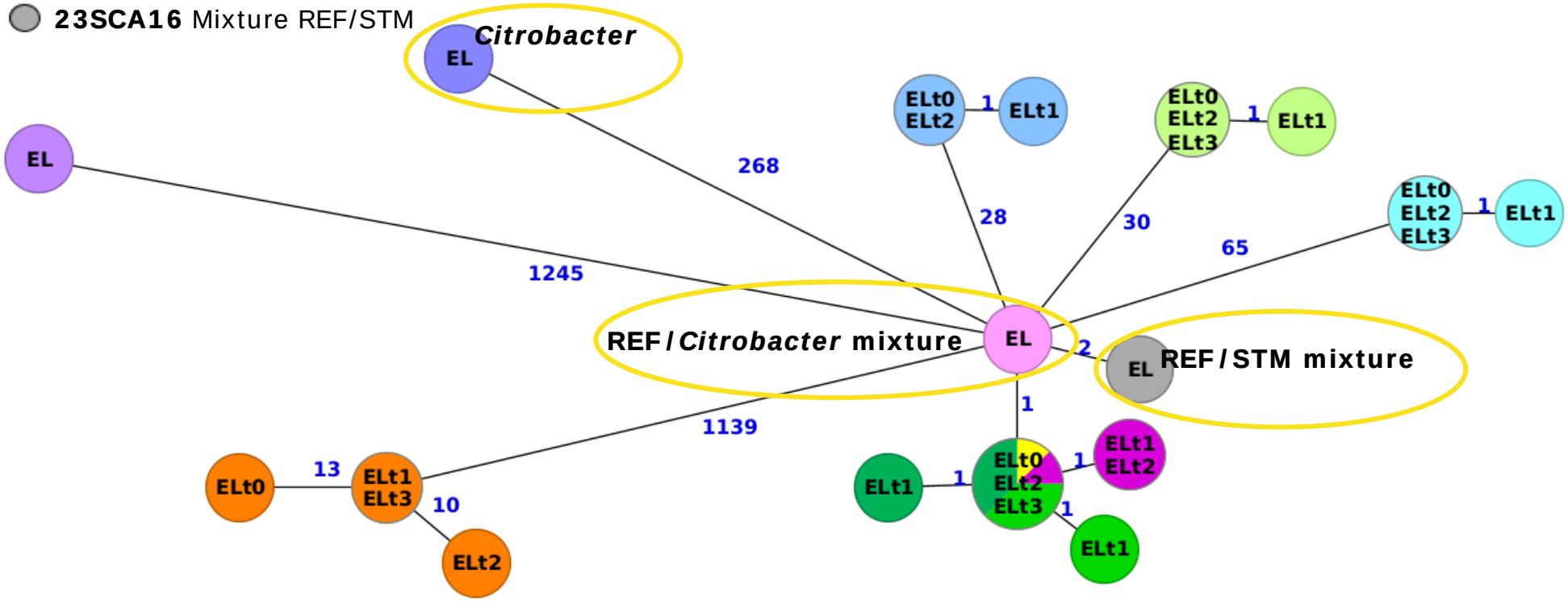
‘Dry’ strains *Salmonella* Cluster Analysis PT 2023

- Pre-testing (WGS at t=0, t=1, t=2)
  - DNA extraction, library preparation, sequencing performed in-house; WGS platform: Illumina NextSeq. Raw data processing: in-house developed Juno-assembly pipeline which includes SPAdes 3.15.3. Cluster analysis: Ridom SeqSphere⁺, using the cgMLST Enterobase v2.0 scheme.
- 6 strains/data selected for inclusion in the PT
- non-*Salmonella* strain 23SCA12 (*Citrobacter freundii*)
- Contaminated Reference strain 23SCA05
  - Wet-mixed on blood agar plates with 23SCA12 (Cf) and with 23SCA01 (STM)

'Dry' strain code	Serovar	ST	Background	Origin
23SCA11	Bovismorbificans	142		Human
23SCA12	non-*Salmonella*		*Citrobacter freundii*	ATCC 8090
23SCA13	Enteritidis	11	2021 pre-tested	Human
23SCA14	Mixed strains		23SCA12 + 23SCA05	
23SCA15	Bovismorbificans	142		Human
23SCA16	Mixed strains		23SCA01 + 23SCA05	

- **23SCA-REF** SBm
- **23SCA01** STM
- **23SCA02** duplicate REF
- **23SCA03** SBm
- **23SCA04** SBm
- **23SCA05** SBm REF
- **23SCA06** SBm
- **23SCA11** SBm
- **23SCA12** *Citrobacter*
- **23SCA13** SE
- **23SCA14** Mixture REF/Citro
- **23SCA15** SBm
- **23SCA16** Mixture REF/STM

‘Pre-view’



19 Participants in Cluster Analysis PT 2023

- 24 submissions
  - 1 participant: both cgMLST-based data and reference-based SNP data
  - 2 participants: 2x cgMLST-based data and 1x reference-based SNP data

	2019	2020	2021	2022	2023
PFGE	6	2	na	na	na
MLVA	8	6	5	5	na
WGS	14	21	19	20	19
WGS: cgMLST	9	15	13	17	18
WGS: SNP	5	8	10	10	6
Partipants overall*	18	21	19	20	19

*excluding multiple data submissions per participant

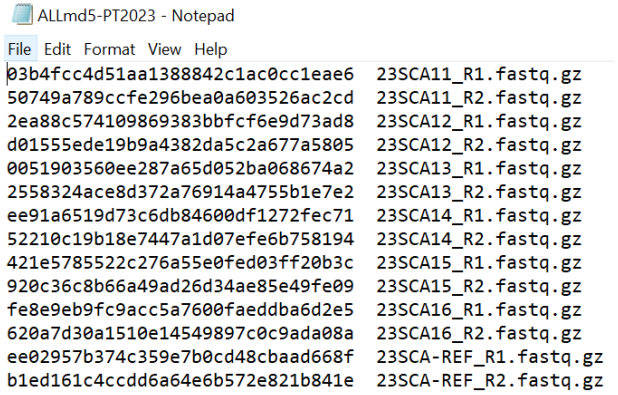
WGS protocols used by the 19 participants

Labcode	Wet lab[a)]	WGS platform	Data analysis[b)]	Tools for analysis	Method for cluster analysis
11	In-In-In	Illumina MiSeq	cgMLST-based	BioNumericsDTU-cgMLSTFinder 1.2	Complete linkage
21	In-In-In	Illumina MiSeq	cgMLST-based	chewBBACA + chewTree on Galaxy Aries, Enterobase as reference	Don´t know
14	In-In-In	Illumina MiSeq	cgMLST-based	BioNumerics	Minimum Spanning Tree (MST)
15-cgMLST1	In-In-In	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
15-cgMLST2	In-Out-Out	Illumina NovaSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
16	In-In-In	Illumina NextSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
25	In-In-In	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
26	In-In-In	Illumina MiSeq	cgMLST-based	chewBBACA	Minimum Spanning Tree (MST)
28-cgMLST	In-In-In	Illumina MiSeq	cgMLST-based	chewBBACA, scheme based on Enterobase, ReporTree for clustering	Minimum Spanning Tree (MST)
30	In-In-In	Illumina MiSeq	cgMLST-based	Ridom SeqSphere	Minimum Spanning Tree (MST)
27	In-In-In	Illumina MiniSeq	cgMLST-based	Ridom SeqSphere	MST to visualize, Distance Matrix for interpreting
9	In-In-In	Illumina MiSeq	cgMLST-based		MSTree V2 - Grapetree's modified version of MST
5	In-In-In	Illumina MiSeq	cgMLST-based	in-house galaxy	MSTreeV2
8	In-In-In	Illumina MiSeq	cgMLST-based	linux command line	Neighbor joining (NJ)
19	In-Out-Out	Illumina NovaSeq	cgMLST-based	cgMLSTFinder 1.2	Neighbor joining (NJ)
4	In-In-In	Illumina MiSeq	cgMLST-based	In-house pipeline with Chewbbaca and scheme from Enterobase	Single linkage
22-cgMLST1	In-In-In	Illumina NextSeq	cgMLST-based	inhouse chewieSnake pipeline (scheme: Enterobase)	single linkage hierarchical clustering
22-cgMLST2	In-In-In	Illumina NextSeq	cgMLST-based	inhouse chewieSnake pipeline (scheme: EFSA)	single linkage hierarchical clustering
23	In-In-In	Illumina MiSeq	SNPa-based	CSIPhylogeny 1.4	Maximum likelihood (ML)
15-SNPr	In-Out-Out	Illumina NovaSeq	SNPr-based	Snippy	Maximum likelihood (ML)
20	In-In-In	Illumina NextSeq	SNPr-based	TrimGalore, Snippy, snp-dist, Gubbins, RAxML	Maximum likelihood (ML)
28-SNPr	In-In-In	Illumina MiSeq	SNPr-based	Python-based pipeline using bowtie2 and bcftools	Minimum Spanning Tree (MST)
3	In-In-In	Illumina MiSeq	SNPr-based	EnterobaseGalaxy	Neighbor joining (NJ)
22-SNPr	In-In-In	Illumina NextSeq	SNPr-based	SnippySnake	single linkage hierarchical clustering

[a)] Wet lab preparations: DNA extractions, Library preparation, Sequencing. In: In-house, Out: Outsourced. [b)] SNPa: assembly-based, SNPr: reference-based

Data quality criteria used by the participants

- Comparable to the PT Typing 2019, 2020, 2021, 2022
- Complete overview will be included in the Final Report PT Typing 2023
- **MD5**: full evaluation will be included in Interim summary report and Final report
  - MD5 can be used as a checksum to verify data integrity against unintentional corruption, e.g. as a result of a faulty file transfer. The MD5 hashes (*message digests*) are typically represented as a sequence of 32 hexadecimal digits.
  - Failures (including juno-pipeline) for Lab 23 strains 23SCA01, 23SCA02, 23SCA04.


ALLmd5-PT2023 - Notepad

File Edit Format View Help

```
03b4fcc4d51aa1388842c1ac0cc1eae6  23SCA11_R1.fastq.gz
50749a789ccfe296bea0a603526ac2cd  23SCA11_R2.fastq.gz
2ea88c574109869383bbfcf6e9d73ad8  23SCA12_R1.fastq.gz
d01555ede19b9a4382da5c2a677a5805  23SCA12_R2.fastq.gz
0051903560ee287a65d052ba068674a2  23SCA13_R1.fastq.gz
2558324ace8d372a76914a4755b1e7e2  23SCA13_R2.fastq.gz
ee91a6519d73c6db84600df1272fec71  23SCA14_R1.fastq.gz
52210c19b18e7447a1d07efe6b758194  23SCA14_R2.fastq.gz
421e5785522c276a55e0fed03ff20b3c  23SCA15_R1.fastq.gz
920c36c8b66a49ad26d34ae85e49fe09  23SCA15_R2.fastq.gz
fe8e9eb9fc9acc5a7600faeddba6d2e5  23SCA16_R1.fastq.gz
620a7d30a1510e14549897c0c9ada08a  23SCA16_R2.fastq.gz
ee02957b374c359e7b0cd48cbaad668f  23SCA-REF_R1.fastq.gz
b1ed161c4ccdd6a64e6b572e821b841e  23SCA-REF_R2.fastq.gz
```

MD5 checksum ‘dry’ strains PT 2020: 11/21
PT 2021: 14/19
PT 2022: 16/20
PT 2023: 16/19

MD5 checksum ‘wet’ strains PT 2023: 13/19

Evaluation cluster analysis (1)

- Participants were asked to report per strain:
  - whether the data passed their Quality Control (QC) criteria or not
  - whether a clustering match with the reference strain (23SCA-REF) was found or not

- Comparing the participants' results to the expected results
  - pre-defined by the EURL-*Salmonella* (Protocol PT Typing 2023)

For this particular PT 2023 situation, the cgMLST-based cluster definition is set at maximum 5 allelic differences from the reference sequence.

- 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 23SCA02/23SCA05 to be (part of) one cluster.
  - Also, strains were expected to be excluded from the cluster analysis (distance matrix) if the data did not pass the participant's QC.
- Deviations from the expected results are indicated in blue

Evaluation WGS-based cluster analysis (2)

Labcode-method	Strain code											
	23 SCA01	23 SCA02	23 SCA03	23 SCA04	23 SCA05	23 SCA06	23 SCA11	23 SCA12	23 SCA13	23 SCA14	23 SCA15	23 SCA16
	STM	Ref2 Bm	Bm	Bm	Ref Bm	Bm	Bm	Citro	SE	Ref+Citro	Bm	Ref+STM
Expected	**No**	**Yes**	**No**	**Yes**	**Yes**	**No**	**No**	**n.a.**	**No**	**n.a.**	**Yes**	**n.a.**
3-SNPr	No	Yes	Yes	Yes	Yes	Yes	Yes	n.a.	n.a.	n.a.	Yes	No
4-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
5-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
8-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
9-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
11-cgMLST	No	Yes	No	Yes	Yes	No	No	No	No	n.a.	Yes	n.a.
14-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
15-cgMLST1	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
15-cgMLST2	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
15-SNPr	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
16-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
19-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
20-SNPr	n.a.*	Yes	No	Yes	Yes	No	No	n.a.	n.a.*	n.a.	No	n.a.
21-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
22-cgMLST1	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
22-cgMLST2	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
22-SNPr	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	n.a.
23-SNPa	No	Yes	No	Yes	Yes	No	No	No	No	n.a.	Yes	n.a.
25-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
26-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
27-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes
28-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	Yes	No	Yes
28-SNPr	No	Yes	No	Yes	Yes	No	No	n.a.	No	Yes	No	Yes
30-cgMLST	No	Yes	No	Yes	Yes	No	No	n.a.	No	n.a.	Yes	Yes

n.a.: Not applicable (QC not passed) n.a.*: Other serovar than Bm

- Technical duplicates 23SCA02 and 23SCA05 within one cluster in 24/24 submissions
- 11/24 submissions reported completely as expected

Excluding non-*Salmonella* data 23SCA12

Lab code	Strain 23SCA12 passed QC	Reason(s) not passing QC	Strain 23SCA12 excluded from cluster analysis
Expected	**No**	**Not *Salmonella* but *Citrobacter freundii***	**Yes**
3-SNPr	No	Per sequence GC content	No
4-cgMLST	No	This is Citrobacter freundii, not passing the contamintion/species check	Yes
5-cgMLST	No	Contamination: Citrobacter freundii	Yes
8-cgMLST	No	contamination with Citrobacter freundii	Yes
9-cgMLST	No	Citrobacter genome	Yes
11-cgMLST	Yes		Yes
14-cgMLST	No	Failed Core Percent quality parameter. Kmer Finder confirmed isolate was not Salmonella.	Yes
15-cgMLST1	No	Non Salmonella spp. : Citrobacter	Yes
15-cgMLST2	No	Not a Salmonella, detected Citrobacter freundii	Yes
15-SNPr	No	Not a Salmonella, detected Citrobacter freundii	Yes
16-cgMLST	No	contamination too high, contaminated with Citrobacter	Yes
19-cgMLST	No	Citrobacter sp.	Yes
20-SNPr	No	Contaminated: KmerID largely Citrobacter	Yes
21-cgMLST	No	Identified as Citrobacter by kmerfinder	Yes
22-cgMLST1	No	Read Fraction Majority Genus: no Salmonella (AQUAMIS), fraction of loci found in cgMLST below 95% (86%) - ChewieSnake	Yes
22-cgMLST2	No	Read Fraction Majority Genus: no Salmonella (AQUAMIS), fraction of loci found in cgMLST below 95% (86%) - ChewieSnake	Yes
22-SNPr	No	Read Fraction Majority Genus: no Salmonella (AQUAMIS), fraction of loci found in cgMLST below 95% (86%) - ChewieSnake	Yes
23-SNPa	Yes*		No
25-cgMLST	No	Citrobacter freundii	Yes
26-cgMLST	No	The 89% of the reads are classified as Citrobacter sp.	Yes
27-cgMLST	No	ID species = 100% Citrobacter freundii	Yes
28-cgMLST	No	Not Salmonella enterica	Yes
28-SNPr	No	Not Salmonella enteritidis	Yes
30-cgMLST	No	Different species: Citrobacter freundii	Yes

*Comment Lab 23: Strain 23SCA12 was included in the cluster analysis as it passed in the quality control.
However this strains is not Salmonella enterica but was identified by kmerFinder as Citrobacter freundii.

Excluding contaminated data 23SCA14

Lab code	Strain 23SCA14 passed QC	Reason(s) not passing QC	Strain 23SCA14 excluded from cluster analysis
Expected	**No**	**Wet mixture of 23SCA05 REF *S.* Bovismorbificans and 23SCA12 *Citrobacter freundii***	**Yes**
3-SNPr	No	Per sequence GC content	No
4-cgMLST	No	not passing contamination/species check - contaminated, also very large assembly (total length)	Yes
5-cgMLST	No	number of contigs, total length, contamination	Yes
8-cgMLST	No	contamination with Citrobacter freundii	Yes
9-cgMLST	No	Citrobacter contamination	Yes
11-cgMLST	No	Total length too big, low GC%, too many contigs, low N50	Yes
14-cgMLST	No	Failed on Genome Length. Kmer Finder confirmed contamination.	Yes
15-cgMLST1	No	Total length of assembly incorrect >5 Mb + number of contigs>150 + contamination : Citrobacter	Yes
15-cgMLST2	No	Contamination by 13% of Citrobacter freundii	Yes
15-SNPr	No	Contamination by 13% of Citrobacter freundii	Yes
16-cgMLST	No	Contamination too high, contaminated with Citrobacter, Total length assembly too high	Yes
19-cgMLST	No	assembly length (9722959), number of contigs >1000: 314	No
20-SNPr	No	Contaminated: Assembly > 5.8Mbp and large portion is Citrobacter	Yes
21-cgMLST	No	Genome size 9,7MBp and Identified as mix of Citrobacter and Salmonella by kmerfinder	Yes
22-cgMLST1	No	total length >9000000kb, Inter species contamination (3750 SNVs), read and contig fraction majority genus below cut-off	Yes
22-cgMLST2	No	total length >9000000kb, Inter species contamination (3750 SNVs), read and contig fraction majority genus below cut-off	Yes
22-SNPr	No	total length >9000000kb, Inter species contamination (3750 SNVs), read and contig fraction majority genus below cut-off	Yes
23-SNPa	No	contaminated; Citrobacter freundii and Salmonella enterica	Yes
25-cgMLST	No	contamination > 10% - Citrobacter freundii	Yes
26-cgMLST	No	Assembly exceeds 5.5Mb and is contaminated with Citrobacter sp. (12% of the reads)	Yes
27-cgMLST	No	Mixed culture was detected with Citrobacter freudii >2%	Yes
28-cgMLST	Yes		No
28-SNPr	Yes		No
30-cgMLST	No	Double the assembly length, failed cgMLST	Yes

Excluding contaminated data 23SCA16

Lab code	Strain 23SCA16 passed QC	Reason(s) not passing QC	Strain 23SCA16 excluded from cluster analysis
Expected	**No**	**Wet mixture of 23SCA05 REF *S.* Bovismorbificans and 23SCA01 *S.* Typhimurium**	**Yes**
3-SNPr	Yes		No
4-cgMLST	Yes		No
5-cgMLST	Yes*		No
8-cgMLST	No	Contamination with Klebsiella pneumoniae	Yes
9-cgMLST	No	14% contamination	Yes
11-cgMLST	No	too large total length, low GC%, a lot of contigs, low N50	Yes
14-cgMLST	No	Failed on Genome Length. Kmer Finder confirmed contamination.	Yes
15-cgMLST1	No	Number of contig >150 + suspected contamination with other serotypes of Salmonella	Yes
15-cgMLST2	No	Contaminated with other Salmonella, Out of EFSA-recommended thresholds: 5,5Mb (max 5,3Mb); 190 contigs (max 150)	Yes
15-SNPr	No	Contaminated with other Salmonella, Out of EFSA-recommended thresholds: 5,5Mb (max 5,3Mb); 190 contigs (max 150)	Yes
16-cgMLST	No	Indication of multiple serotypes	No
19-cgMLST	Yes*		No
20-SNPr	No	Contaminated: multiple serotypes (Bovismorbificans and Bsilla)	Yes
21-cgMLST	No	Genome size 5,4Mbp	Yes
22-cgMLST1	No	total length above threshold, Intra species contamination (340 SNVs)	Yes
22-cgMLST2	No	total length above threshold, Intra species contamination (340 SNVs)	Yes
22-SNPr	No	total length above threshold, Intra species contamination (340 SNVs)	Yes
23-SNPa	No	N. contigs>500; Total length higher than expected	Yes
25-cgMLST	Yes		No
26-cgMLST	Yes		No
27-cgMLST	No	number of contigs = 443 (warning) but all the other criteria are OK	No
28-cgMLST	Yes		No
28-SNPr	Yes		No
30-cgMLST	Yes		No

*Comment Lab 5: extra care should be taken with the conclusion of Strain 23SCA16 since the threshold of 10% missing loci was almost passed.

*Comment Lab 19: 23SCA16 did not completely pass the quality check (contigs >1000: 154, N50: 99482), but was so close, that it was processed for further analyses.

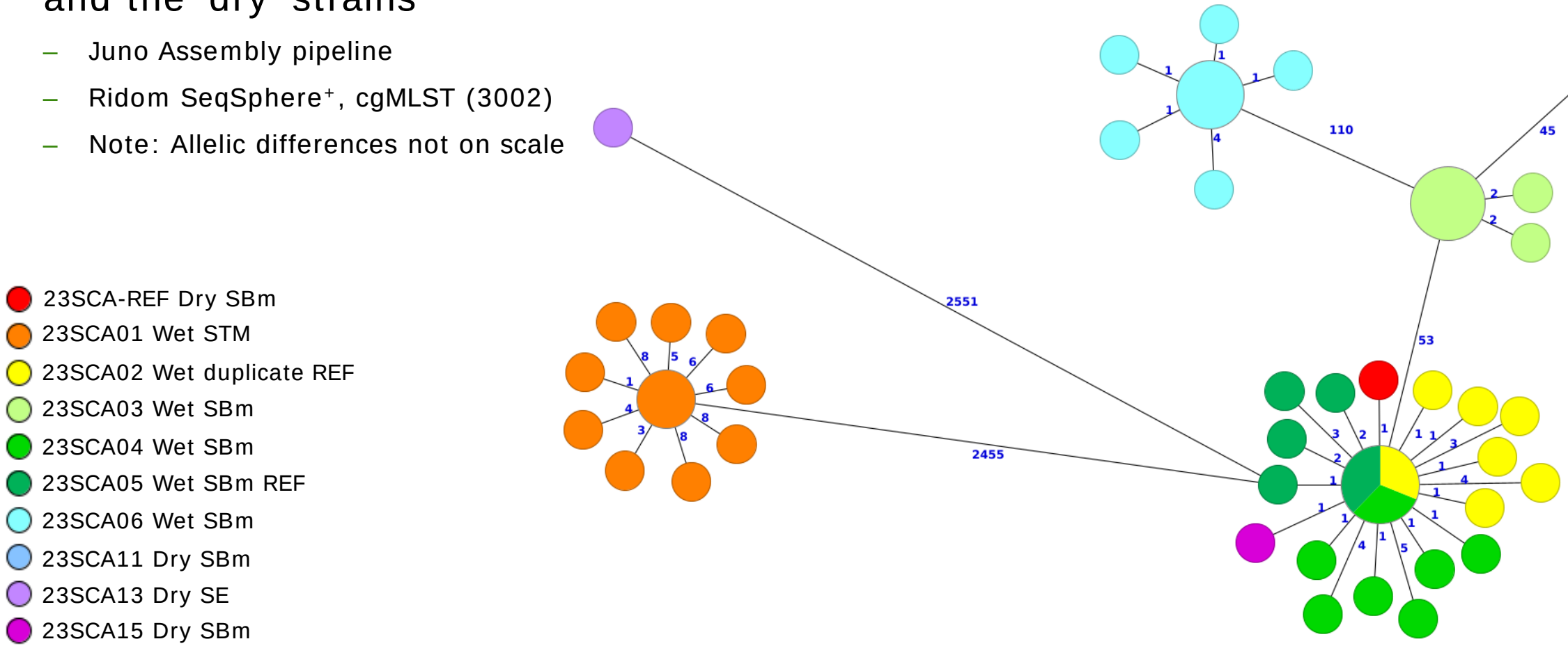
Per submission: *Distance matrix* data

Labcode-method	Strain code												
	23 SCA-REF	23 SCA01	23 SCA02	23 SCA03	23 SCA04	23 SCA05	23 SCA06	23 SCA11	23 SCA12	23 SCA13	23 SCA14	23 SCA15	23 SCA16
	Ref Bm	STM	Ref2 Bm	Bm	Bm	Ref Bm	Bm	Bm	Citro	SE	Ref+Citro	Bm	Ref+STM
EL-cgMLST	0	2595	**1**	56	**1**	**1**	128	60		2635		**2**	
4-cgMLST	0	2383	**0**	35	**0**	**0**	101	46		2459		**0**	**0**
5-cgMLST	0	2472	**0**	57	**0**	**0**	135	58		2504		**3**	**4**
8-cgMLST	0	2576	**5**	54	**1**	**0**	129	57		2624		**2**	
9-cgMLST	0		**1**	65	**2**	**1**	151	69				**4**	
11-cgMLST	0		**1**	59	**3**	**1**	135	61				**3**	
14-cgMLST	0	200	**0**	54	**0**	**0**	134	60		200		**1**	
15-cgMLST1	0		**0**	55	**0**	**0**	128	58				**1**	
15-cgMLST2	0	2594	**1**	55	**1**	**0**	127	58		2630		**1**	
16-cgMLST	0	2595	**1**	56	**1**	**1**	128	60		2635		**2**	4
19-cgMLST	0	2425	**0**	52	**0**	**0**	124	57		2442	14	**0**	**1**
21-cgMLST	0	2559	**0**	53	**0**	**0**	127	57		2562		**3**	
22-cgMLST1	0	2569	**1**	53	**1**	**1**	128	57		2605		**2**	
22-cgMLST2	0	2875	**1**	61	**2**	**1**	147	66		2922		**3**	
25-cgMLST	0	2599	**0**	55	**0**	**1**	125	58		2641		**1**	**0**
26-cgMLST	0	2822	**0**	19	**0**	**0**	96	32		2862		**0**	**0**
27-cgMLST	0	2595	**0**	55	**0**	**0**	127	58		2622		**1**	**0**
28-cgMLST	0	2548	**3**	58	**2**	**3**	132	62		2592	**1**	6	**4**
30-cgMLST	0	2595	**1**	56	**1**	**1**	128	59		2635		**2**	**1**
23-SNPa	0	11170	**3**	46	**2**	**2**	101	52	12277	11403		**6**	
3-SNPr	0	8	**0**	**2**	**0**	**0**	**7**	**4**	2971	447	0	**0**	0
15-SNPr	0	39206	**1**	269	**4**	**0**	671	645		39141		**9**	
20-SNPr	0		**0**	111	**2**	**0**	241	103				6	
22-SNPr	0		**0**	141	**2**	**1**	550	245				**7**	
28-SNPr	0	4487	**0**	55	**1**	**0**	137	46		21385	**0**	6	**0**

Numbers in **bold** were reported to be clustering with the REF strain

Evaluation WGS-based cluster analysis (3)

- MST of the strains from the participants' processed raw data and the 'dry' strains
  - Juno Assembly pipeline
  - Ridom SeqSphere⁺, cgMLST (3002)
  - Note: Allelic differences not on scale



Conclusions PT Cluster Analysis 2023

- **19** participants, with **24** submissions
- Pre-testing for strain selection
- 6 ‘wet’ strains and 6 ‘dry’ strains
- Pre-defined “Cluster definitions”, (REF strain, expected results)
- Strains 23SCA12, 23SCA14, 23SCA16 (failing QC) had to be excluded from cluster analysis (distance matrix)
- No performance criteria were set for this PT
  - Technical duplicate strains 23SCA02/23SCA05 reported to be (part of) one cluster: 24/24
  - Strains 23SCA12, 23SCA14, 23SCA16 excluded from cluster analysis: 11/24
  - Participants’ results compared to expected results cluster analysis: 13/24
- **Interim summary report expected to be emailed next week**

- **6th PT on (optional) Cluster Analysis, WGS only, in 2024**
  - Defined cluster analysis on 10-12 *Salmonella* strains
    - Simulation of an outbreak-related request from the EURL-*Salmonella* (EFSA/ECDC) to the NRL-network
    - A WGS reference sequence will be provided (plus md5sum values)
    - Proposal: 5-6 strains ‘wet’ analysis plus 5-6 strains ‘dry’ analysis

Thank you for your attention and your participation !
Hartelijk dank !