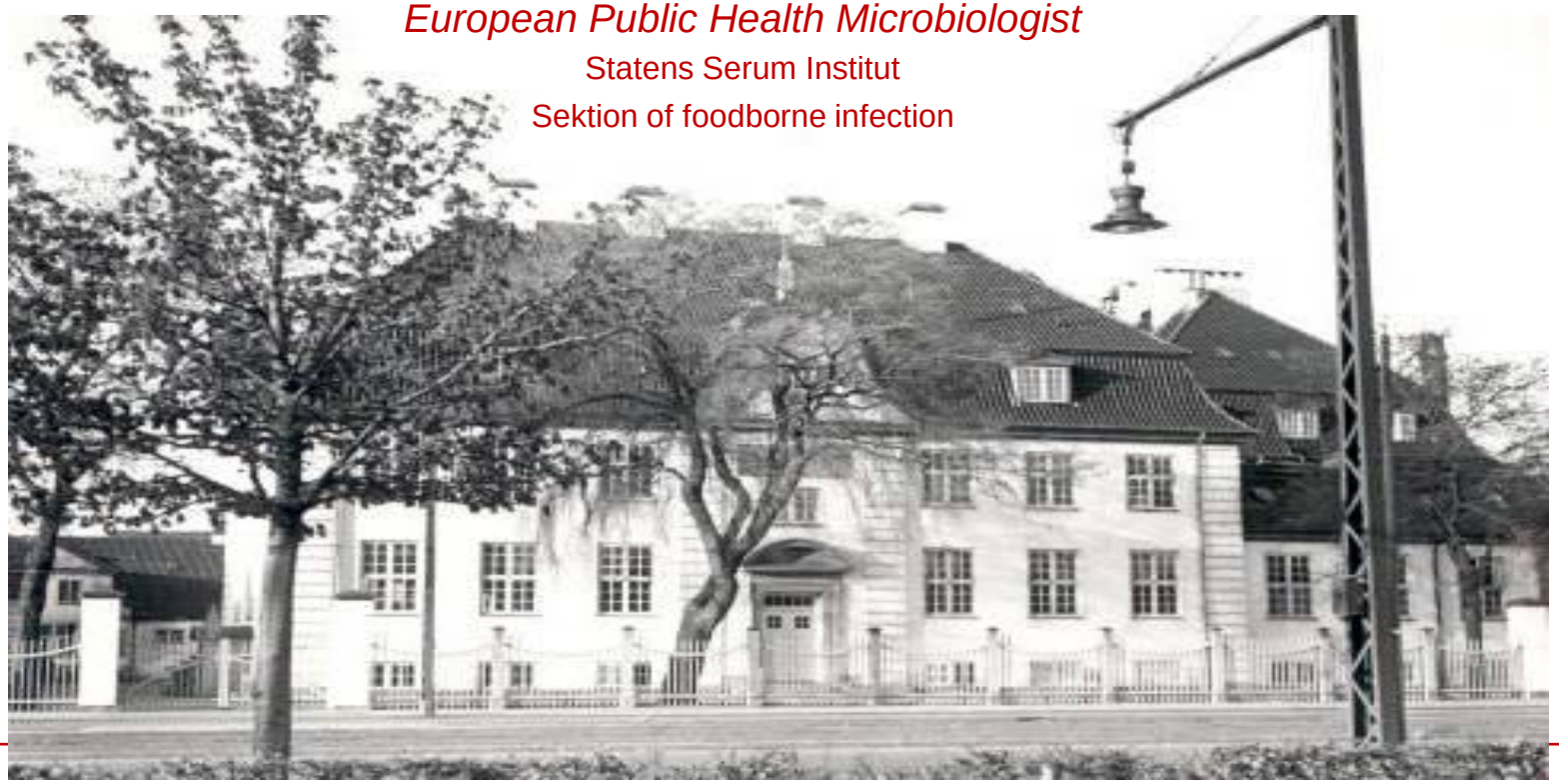
STATENS SERUM INSTITUT

EXTERNAL QUALITY ASSESSMENT OF MOLECULAR TYPING OF STEC 2018-2019

Susanne Schjørring, M.Sc, Ph.D
European Public Health Microbiologist
Statens Serum Institut
Sektion of foodborne infection



Molecular typing EQA´s

STATENS SERUM INSTITUT

Funded by European Centre for Disease Prevention and Control (ECDC)
Organised by Statens Serum Institut, Denmark, since September 2012

Objectives

Serotyping

Assess the **ability to assign correct O groups and H types** by using either serological (somatic 'O' and flagellar 'H' antigens) or molecular typing methods (PCR or WGS).

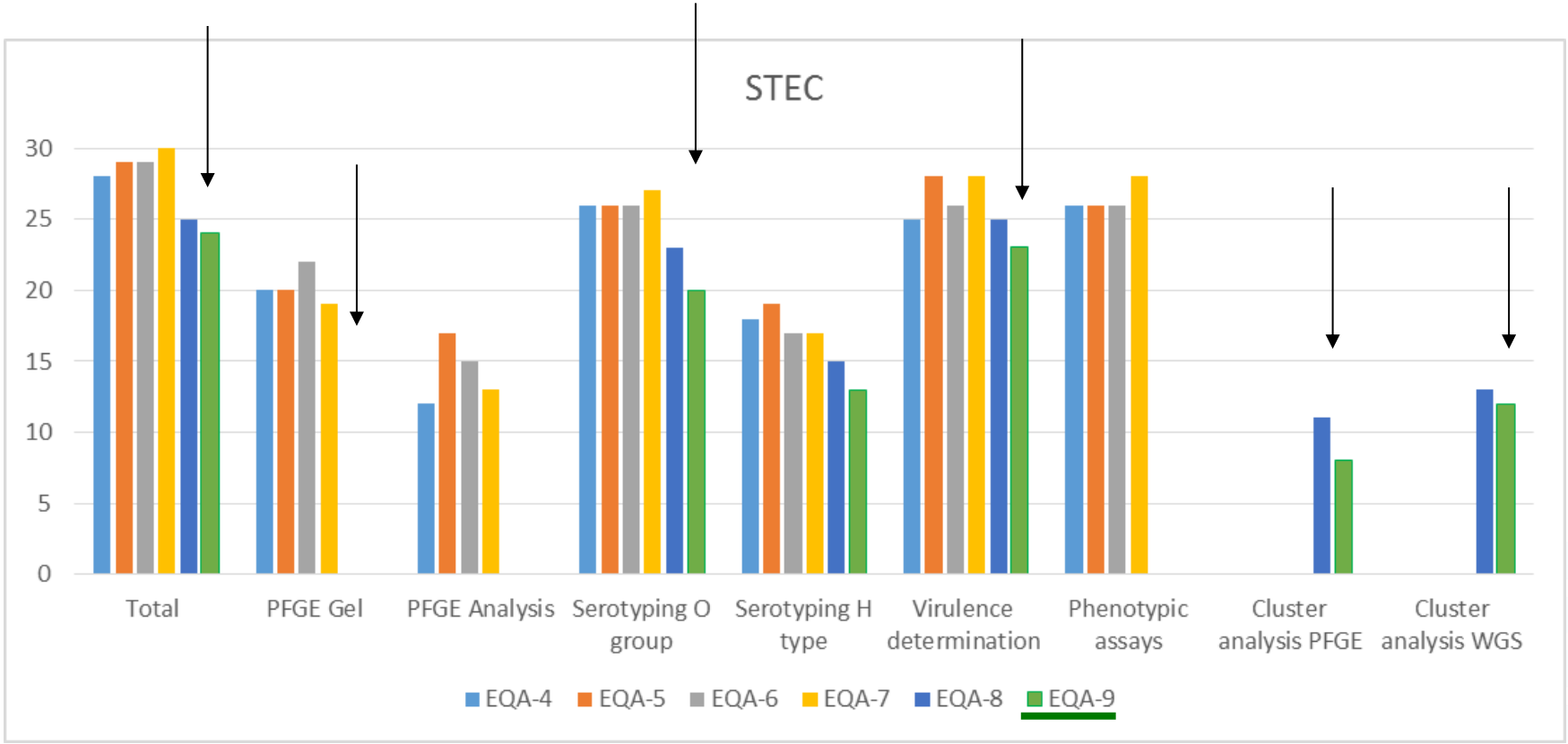
Virulence profile

Assess the **ability to assign the correct virulence profile**.
The presence/absence of *stx1*, *stx2*, *eae, aaiC* and *aggR* genes and subtyping of stx genes (*stx1a, stx1c* and *stx1d* and *stx2a* to *stx2g*).

Molecular typing-based cluster analysis

Assess the **ability to detect a cluster of closely related isolates**.
Laboratories could perform the analyses using PFGE or derived data from WGS.

Participation since 2012 (ECDC funded)

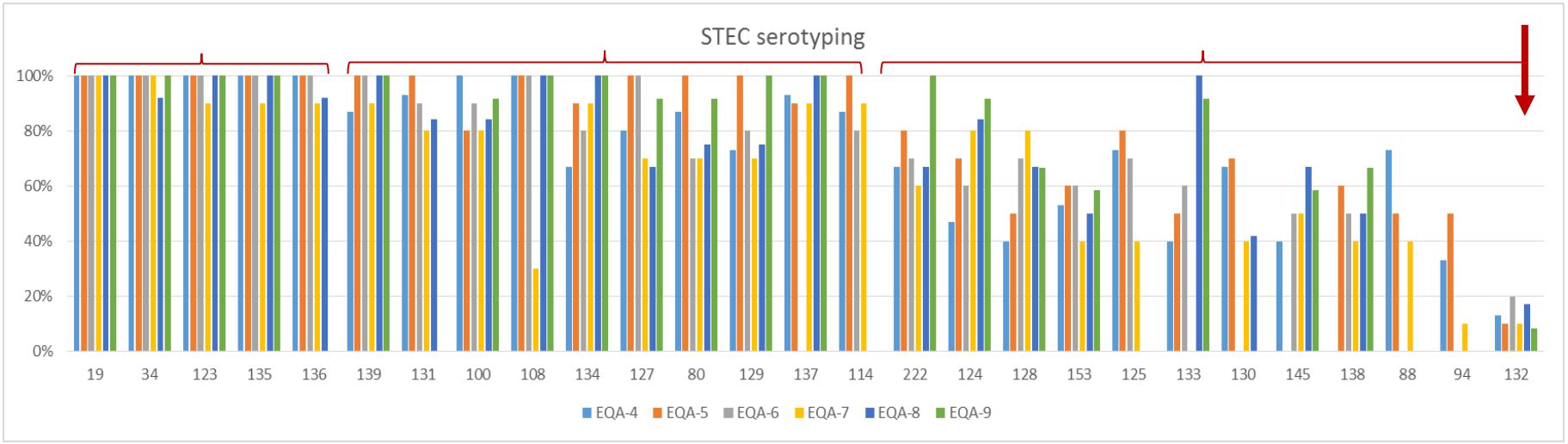


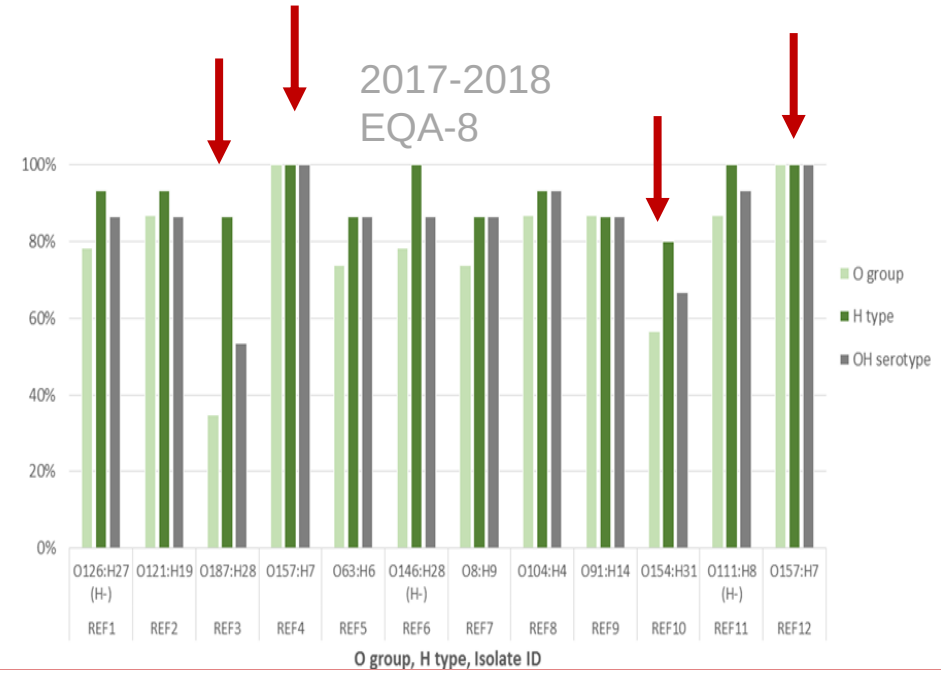
Participants 2018-2019 (n=24)

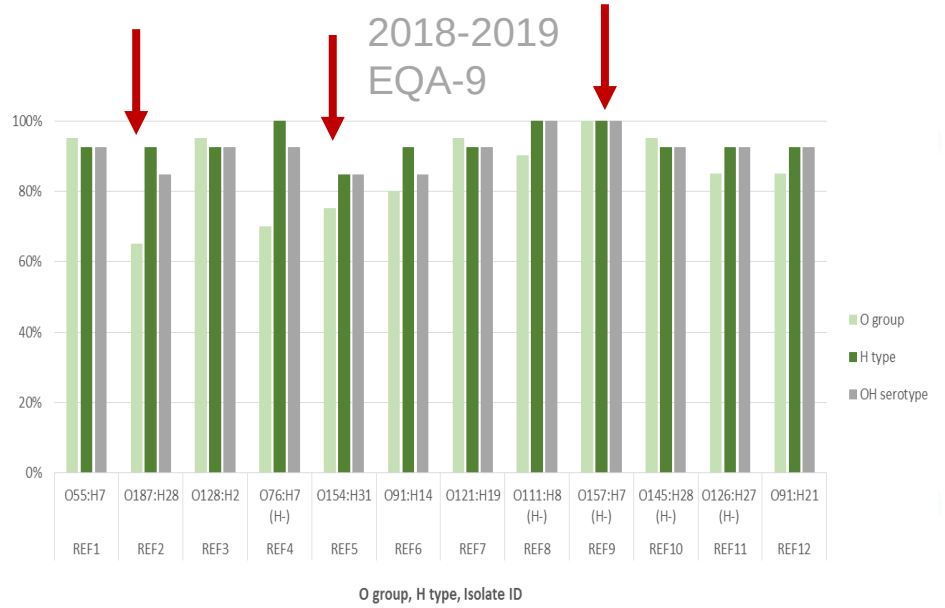
STATENS SERUM INSTITUT

Austria
Belgium
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Iceland
Ireland
Italy
Latvia
Lithuania
Luxembourg
Macedonia
Norway
Poland
Portugal
Romania
Slovenia
Sweden
The Netherlands
United Kingdom

Assigning the O group







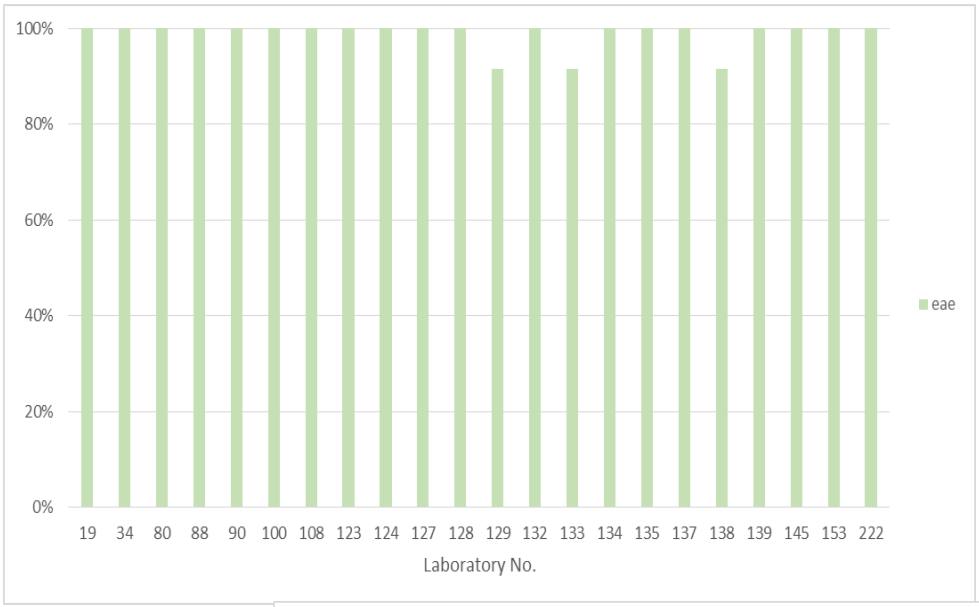
Average: 79%, 92% and 86%

Average:86%, 94% and 92%

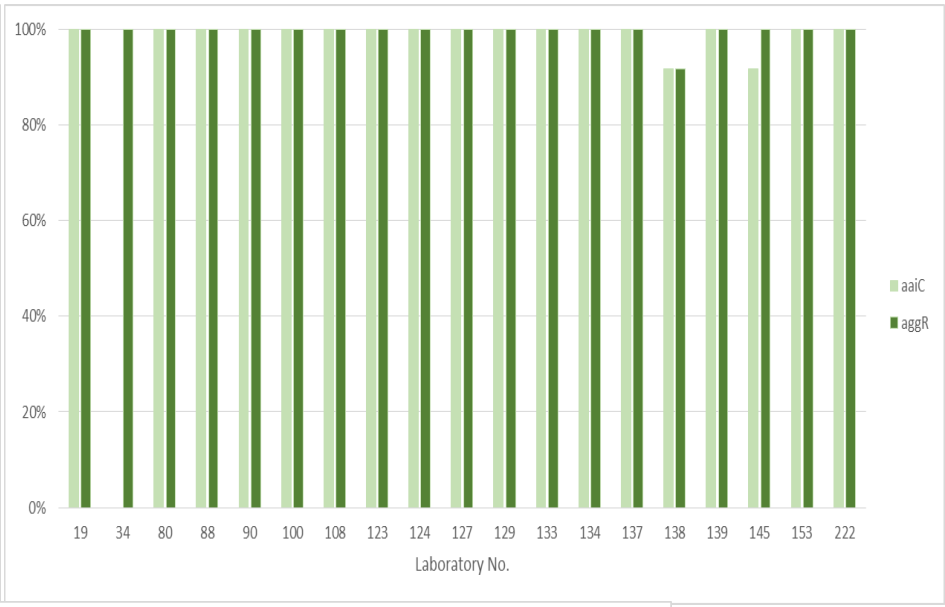
Virulence profile 2018-2019

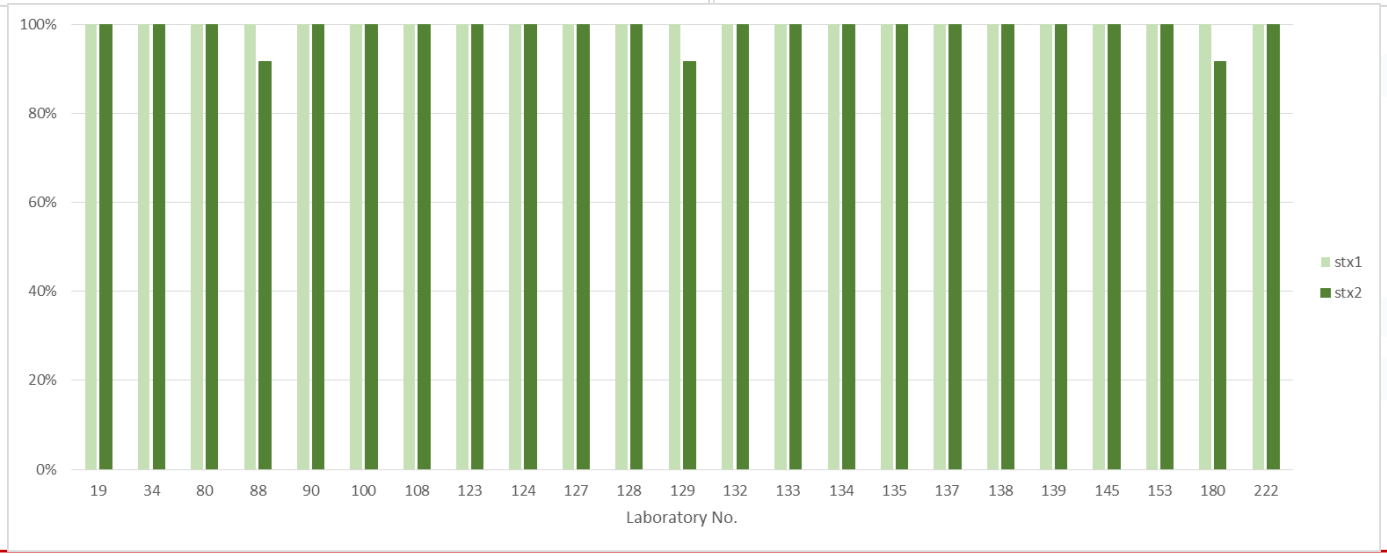
STATENS SERUM INSTITUT

eae 95% average



aaiC 99% aggR 99.5~100%



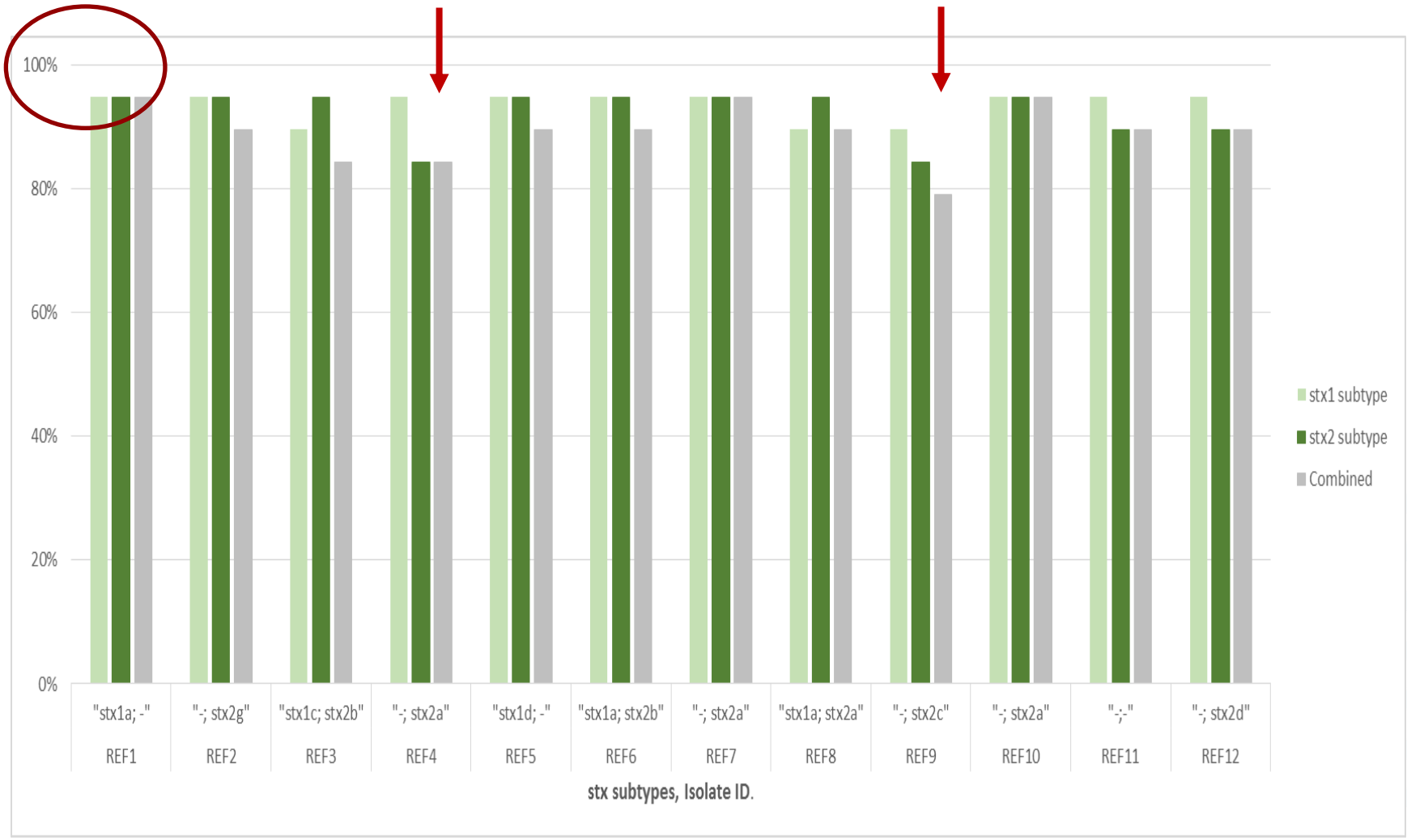


stx1 100%
stx2 99%

STEC stx subtyping

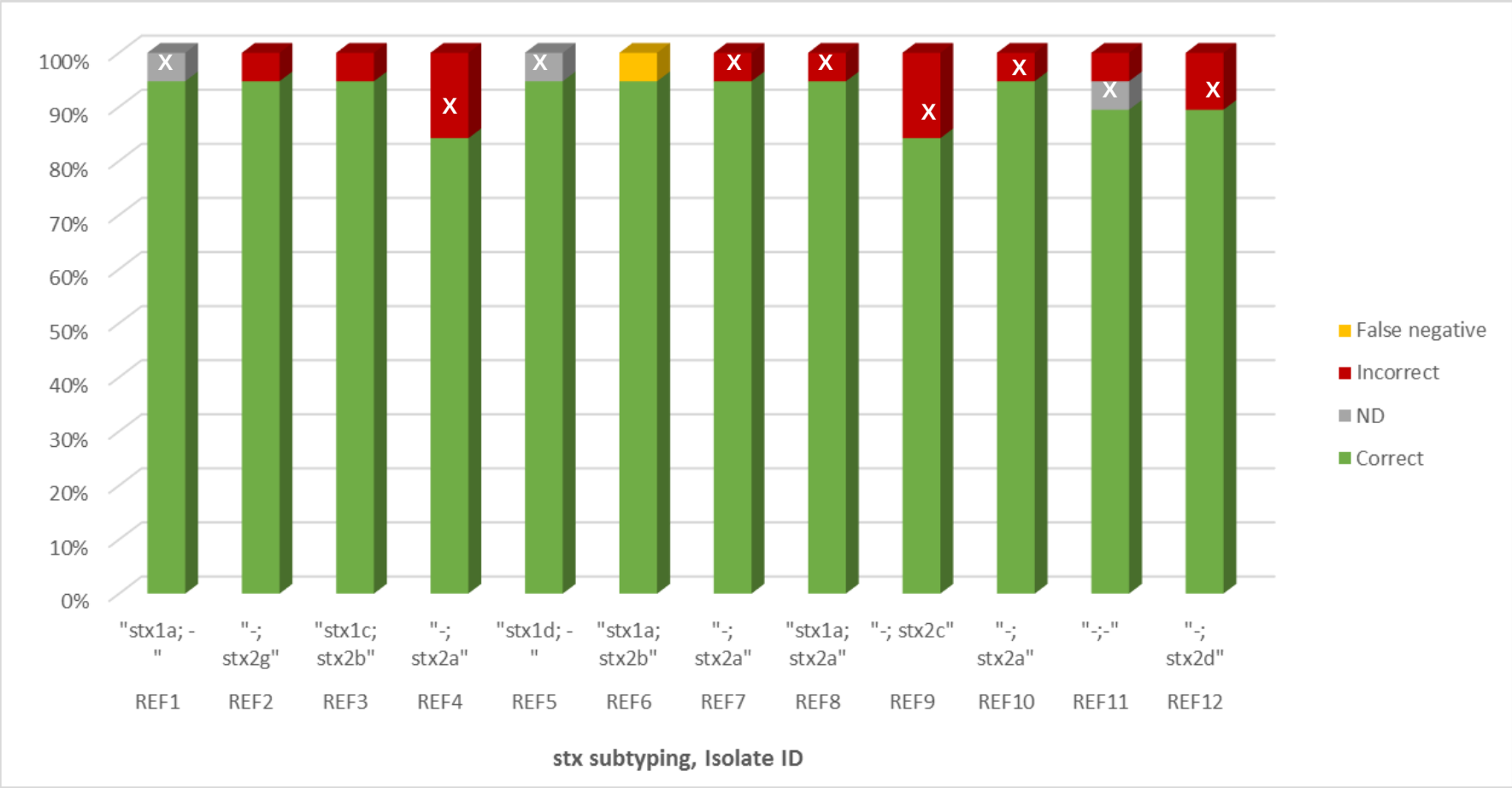
STATENS SERUM INSTITUT

2018-2019



11 laboratories subtyped all 12 isolates correctly

Subtyping *stx2*



9 errors was reported by one laboratory

REF4 *stx2a* and /or *stx2c* ?

STATENS SERUM INSTITUT

Mapping:

”and” = half and half

”or” = EDD (*stx2a*) vs. END (*stx2c*)

J Clin Microbiol. 2012 Sep;50(9):2951-63. doi: 10.1128/JCM.00860-12. Epub 2012 Jul 3

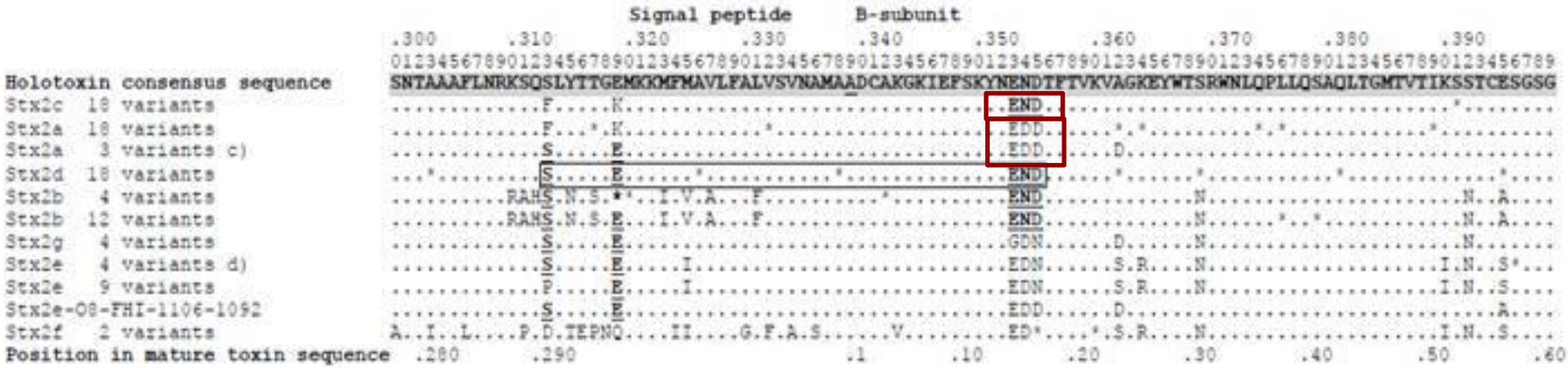
JCM Journals.ASM.org

Multicenter Evaluation of a Sequence-Based Protocol for Subtyping Shiga Toxins and Standardizing Stx Nomenclature

Flemming Scheutz,[a] Louise D. Teel,[b] Lothar Beutin,[c] Denis Piérard,[d] Glenn Buvens,[d] Helge Karch,[e] Alexander Mellmann,[e] Alfredo Caprioli,[f] Rosangela Tozzoli,[f] Stefano Morabito,[f] Nancy A. Strockbine,[g] Angela R. Melton-Celsa,[b] Maria Sanchez,[b] Søren Persson,[a] and Alison D. O'Brien[b]

WHO Collaborating Centre for Reference and Research on *Escherichia* and Klebsiella, Unit of Foodborne Bacteria and Typing, Department of Microbiology and Infection Control, Statens Serum Institut, Copenhagen, Denmark[a]; Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA[b]; National Reference Laboratory for Escherichia coli, Federal Institute for Risk Assessment (BfR), Berlin, Germany[c]; National Reference Center for Shiga Toxin/Verotoxin Producing *E. coli*, Department of Microbiology, Universitair Ziekenhuis Brussel, Brussels, Belgium[d]; Institute of Hygiene and National Consulting Laboratory for Hemolytic Uremic Syndrome, University of Münster, Münster, Germany[e]; Istituto Superiore di Sanità, Rome, Italy[f]; and National Escherichia and Shigella Reference Unit, Enteric Diseases Laboratory Branch, National Center for Emerging and Zoonotic Infectious Diseases (CDC), Atlanta, Georgia, USA[g]

When Shiga toxin-producing *Escherichia coli* (STEC) strains emerged as agents of human disease, two types of toxin were identified: Shiga toxin type 1 (Stx1) (almost identical to Shiga toxin produced by *Shigella dysenteriae* type 1) and the immunologically distinct type 2 (Stx2). Subsequently, numerous STEC strains have been characterized that express toxins with variations in amino acid sequence, some of which confer unique biological properties. These variants were grouped within the Stx1 or Stx2 type and often assigned names to indicate that they were not identical in sequence or phenotype to the main Stx1 or Stx2 type. A lack of specificity or consistency in toxin nomenclature has led to much confusion in the characterization of STEC strains. Because serious outcomes of infection have been attributed to certain Stx subtypes and less so with others, we sought to better define the toxin subtypes within the main Stx1 and Stx2 types. We compared the levels of relatedness of 285 valid sequence variants of Stx1 and Stx2 and identified common sequences characteristic of each of three Stx/Stx1 and seven Stx2 subtypes. A novel, simple PCR subtyping method was developed, independently tested on a battery of 48 prototypic STEC strains, and improved at six clinical and research centers to test the reproducibility, sensitivity, and specificity of the PCR. Using a consistent schema for nomenclature of the Stx toxins and *stx* genes by phylogenetic sequence-based relatedness of the holotoxin proteins, we developed a typing approach that should obviate the need to bioassay each newly described toxin and that predicts important biological characteristics.



Cluster analysis (2018-2019)

STATENS SERUM INSTITUT

Perform a cluster analyses by using PFGE- or WGS-derived data

- Report the isolates identified as being closely related (outbreak)
- Submit distance between one(cluster isolates) and the other test isolates
  - Band difference in PFGE (total bands /shared bands)
  - SNP distances or allele differences (wgMLST/cgMLST) (WGS)
  - If using WGS, the submission should include the fastq- files

Evaluation:

- The ability to detect a cluster of closely related isolates bases on a pre-defined categorization by the organizer / (PFGE / WGS)
  (mimicking an outbreak situation)
- The submitted raw reads were “evaluated“ by the SSI in-house quality control pipeline

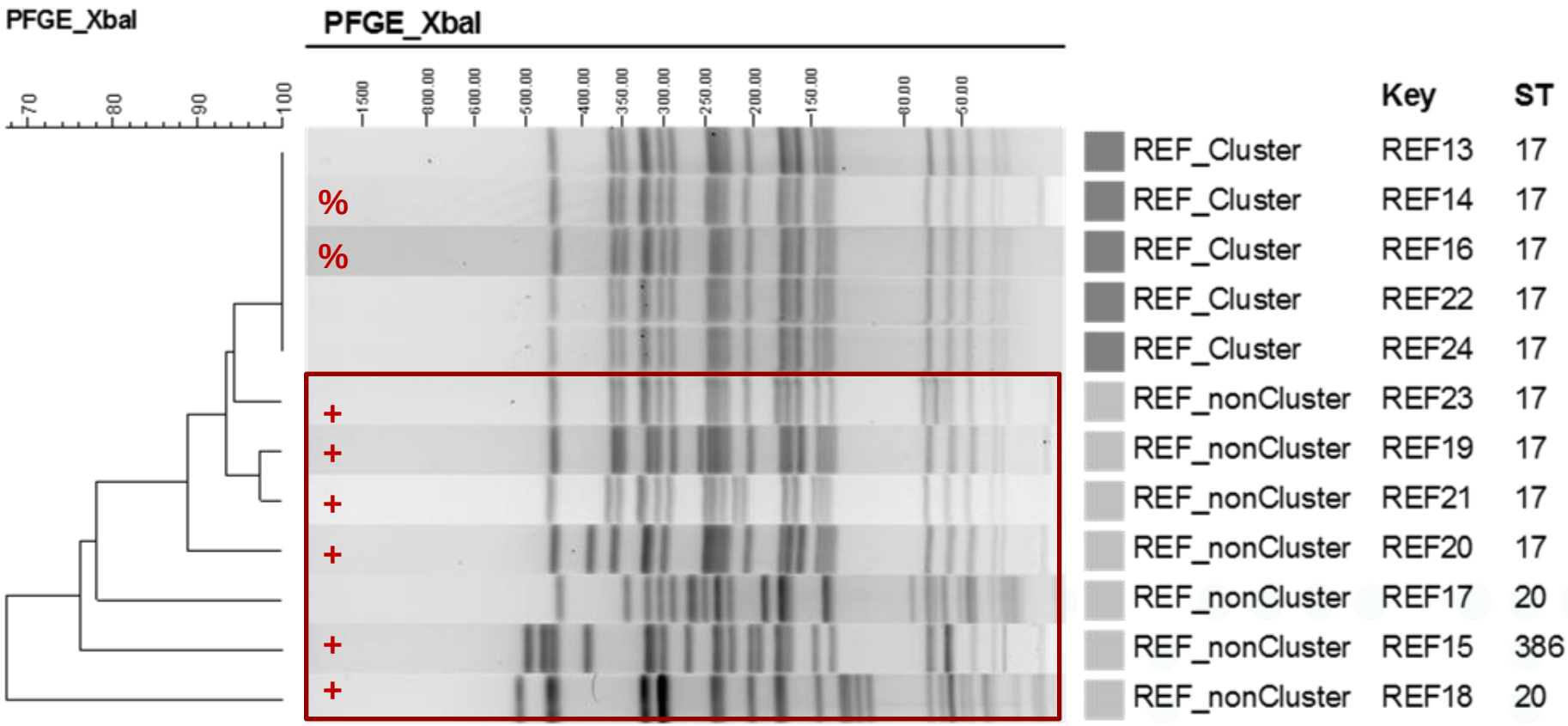
Cluster analysis (PFGE)

		Laboratory number							
Isolate ID	**ST**	**19**	**90**	**123**	**124**	**127**	**130**	**132**	**222**
REF13‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF14‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	No x	Yes
REF15	386	No	No	No	No	No	No	Yesx	No
REF16‡#	17	Yes	Yes	Yes	Yes	Yes	No x	No x	Yes
REF17	20	No	No	No	No	No	No	No	No
REF18	20	No	No	No	No	No	No	Yesx	No
REF19	17	No	No	No	No	No	Yes x	Yesx	No
REF20	17	No	No	No	No	No	Yes x	Yesx	No
REF21	17	No	No	No	No	No	Yes x	Yes	Yesx
REF22‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF23	17	No	No	No	No	No	Yes x	Yesx	No
REF24‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster-identified conclusion		Yes	Yes	Yes	Yes	Yes	No	No	No

‡: Closely related isolates (5 isolates) #: Technical triplicated isolates

(three also performed WGS)

Cluster analysis (PFGE)



cluster analysis (WGS)

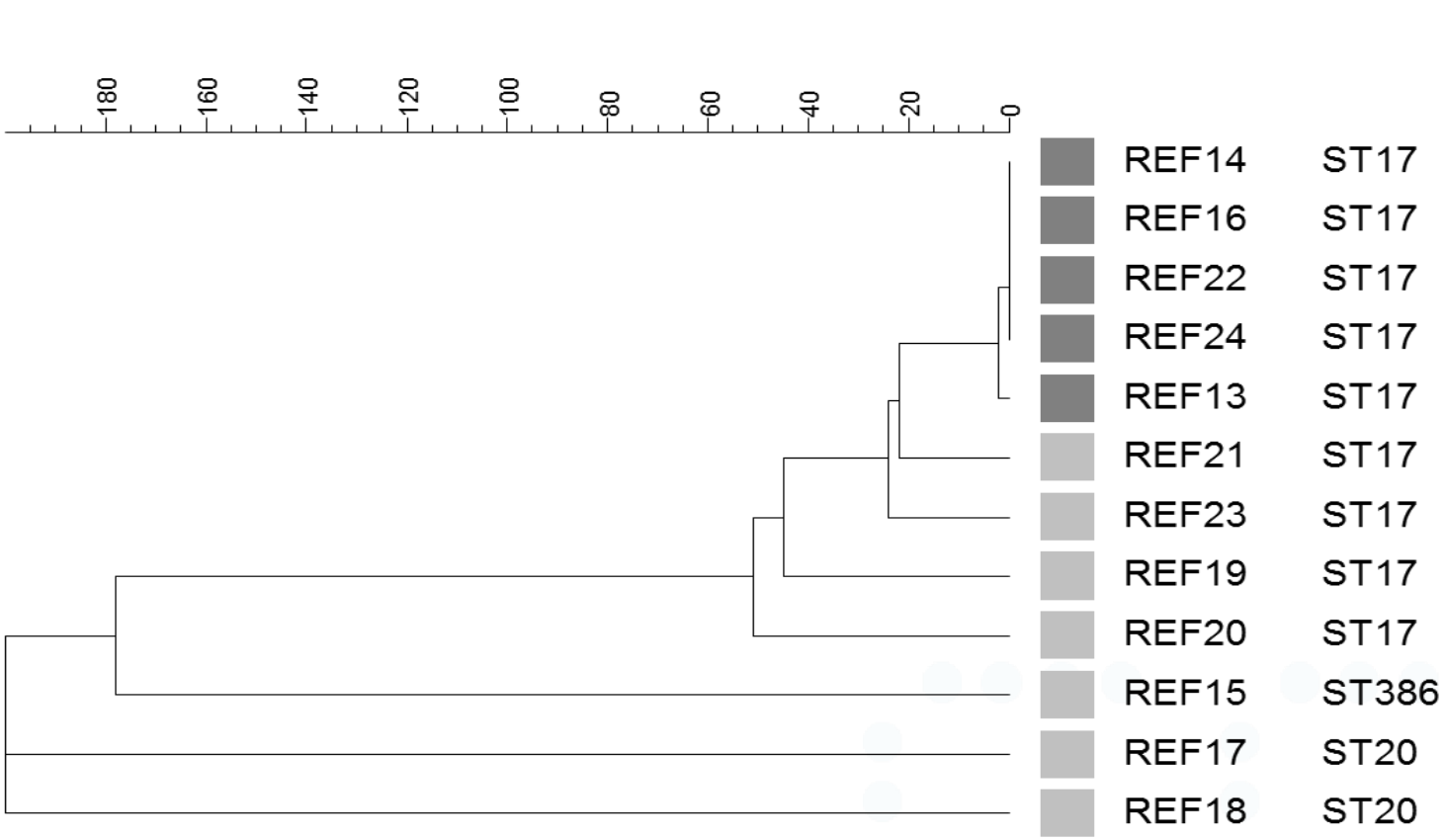
STATENS SERUM INSTITUT

O103:H2, *stx1a*

		Laboratory number											
Isolate ID	**ST**	**19**	**34**	**80**	**100**	**108**	**123**	**133**	**134**	**135**	**137**	**139**	**222**
REF13‡	17	Yes	Yes	Yes	Yes	No %	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF14‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF15	386	No	No	No	No	No	No	No	No	No	No	No	No
REF16‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No %
REF17	20	No	No	No	No	No	No	No	No	No	No	No	No
REF18	20	No	No	No	No	No	No	No	No	No	No	No	No
REF19	17	No	No	No	No	No	No	No	No	No	No	No	No
REF20	17	No	No	No	No	No	No	No	No	No	No	No	No
REF21	17	No	No	No	No	No	No	No	No	No	No	No	No
REF22‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF23	17	No	No	No	No	No	No	No	No	No	No	No	No
REF24‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main analysis		Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	SNP	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	SNP	Allele (cgMLST)	Allele (cgMLST)
Additional analysis													
Cluster-identified		Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No

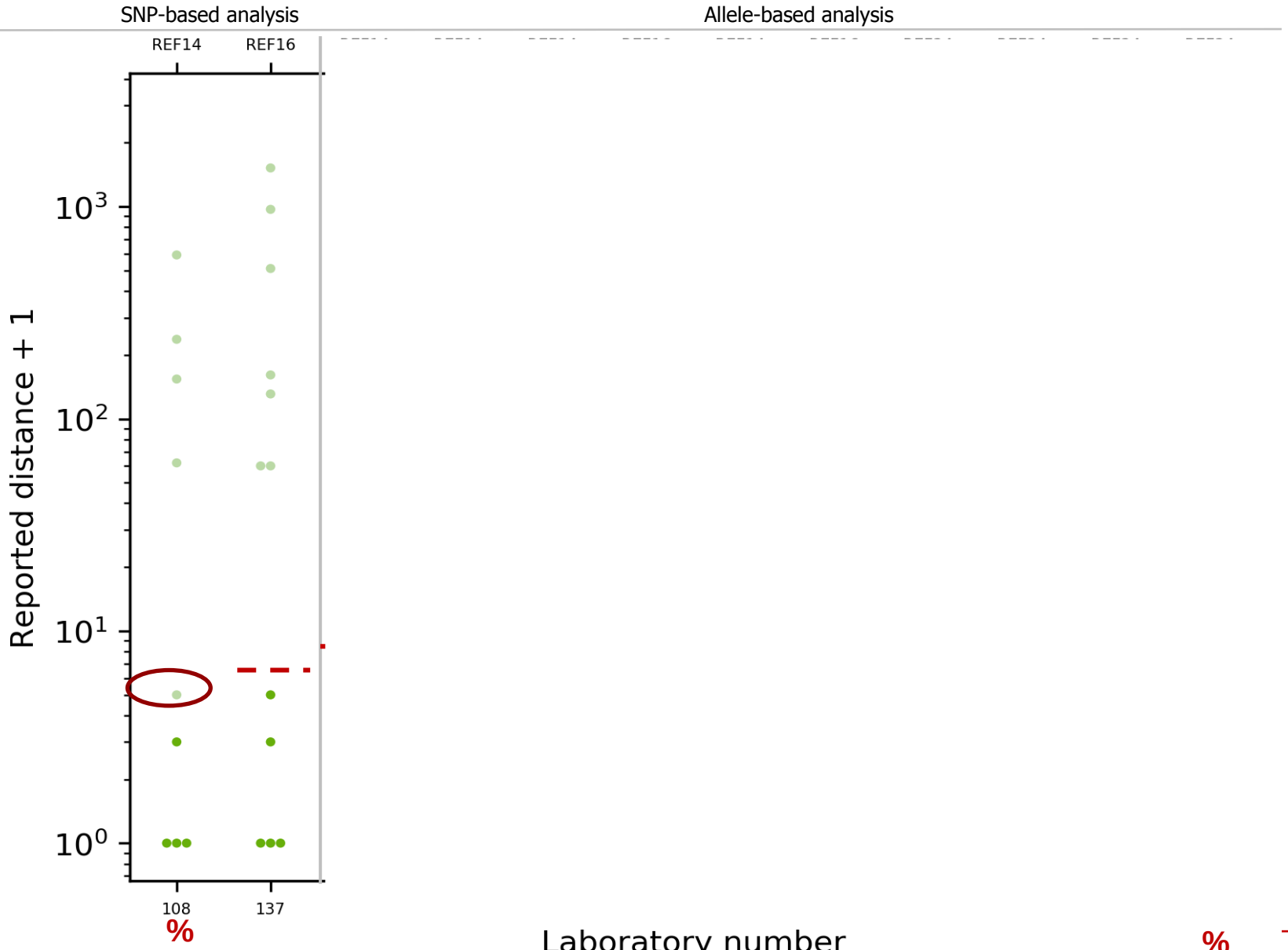
‡: Closely related isolates (5 isolates) #: Technical triplicated isolates

cgMLST (Enterobase Scheme)



REF13, 14, 16, 22, 24: Cluster of closely related isolates (5 isolates)
(0-2 allele differences (Enterobase scheme / 0-4 SNP distances (NASP pipeline)
REF14, 16 and 22 is technical triplicated isolates

SNP Distance / allele difference



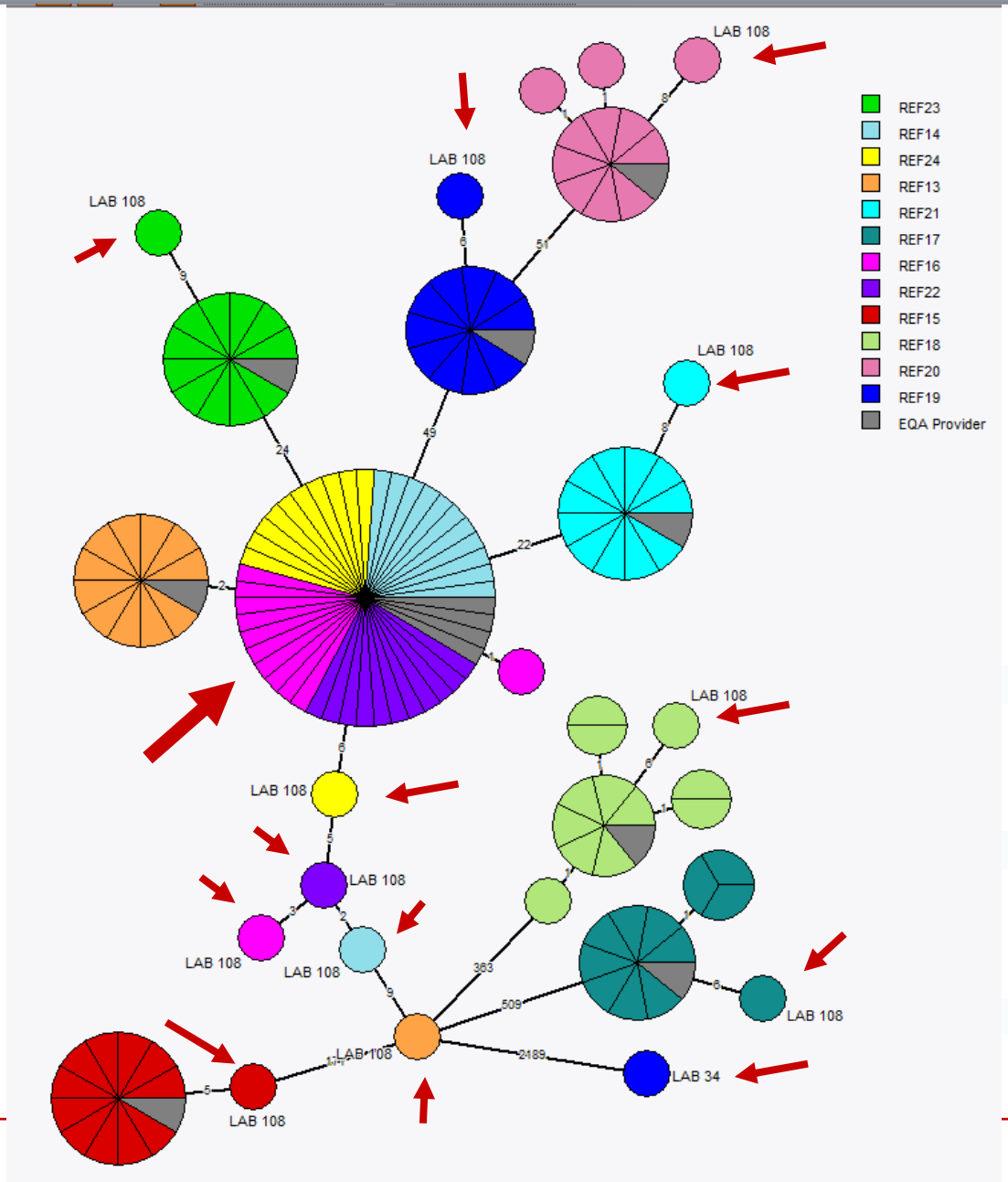
Quality Control pipeline

STATENS SERUM INSTITUT

- In general the QC of the data were ok

		Laboratory number											
Parameters	**Ranges***	**19**	**34**	**80**	**100**	**108**	**123**	**133**	**134**	**135**	**137**	**139**	**222**
Detected species	{Ec} or {Sf}	*Ec*	Ec / Sf	Ec	Ec	Ec	Ec / Sf	Ec	Ec	Ec / Sf	Ec	Ec	Ec
Species 1 (%)		83.1-92.8	56.2-98.2	76.8-94.9	87.7-96.1	90.8-96.0	83.0-95.6	74.7-93.9	86.8-96.2	82.0-95.4	83.0-92.4	85.4-96.1	89.6-96.9
Species 2 (%)		0.9-2.7	0.1-13.0	0.8-3.4	0.8-2.3	0.9-1.6	0.7-7.7	0.9-3.2	0.8-2.8	0.7-8.8	1.0-2.5	0.9-2.8	0.8-1.5
Unclassified reads (%)	{<100}	5.1-12.1	1.5-41.0	1.9-18.9	1.8-7.9	1.9-5.6	2.2-13.3	4.0-20.7	1.8-8.8	2.1-14.0	5.9-12.9	2.0-9.9	0.8-6.6
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.0-5.4	5.1-5.5	4.7-5.4	5.1-5.5	4.5-5.0	5.1-5.5	0.9-5.5	5.1-5.5	5.1-5.4	4.9-5.2	5.1-5.5	5.1-5.4
Length [0–25] x min. coverage (kbp)	{<250}	2.7-18.7	0.0-15.5	0.0-673.4	0.0-99.9	0.0-20.0	0.0-1.2	0.0-3843.4	0.0-2.1	9.9-62.4	0.0-251.4	0.0-0.0	0.0-2.8
Number of contigs at 25 x min. coverage	{>0}	341-534	158-290	124-211	149-225	2417-5182	210-316	185-553	163-247	153-247	829-1008	153-255	343-2122
Number of contigs [0–25] x min. coverage	{<1000}	5-32	0-14	0-74	0-14	0-62	0-3	0-3130	0-3	6-38	0-236	0-0	0-12
Average coverage	{>50}	88-168	40-88	37-130	46-145	25,5-95,5	85-136	21-135	36-96	77-108	36-204	45-78	52-266,5
Number of reads (x1000)		3332-6194	973-2240	741-3094	1087-3397	490-1752	1775-2982	774-3857	1267-3397	2795-3949	2047-10851	1611-2712	963,5-4602
Average read length		134-141	204-243	252-288	229-243	272-294	232-259	153-238	151-151	142-148	98.5-99.5	151-151	290-326
Average insert size		211-246	265-359	286-418	284-372	NA	255-313	152-269	298-380	342-418	302-355	298-328	NA
N50 (kbp)		30-48	54-166	86-135	86-140	1-3	73-116	2-127	88-140	58-135	18-26	82-140	4-40

Minimum spanning tree cgMLST



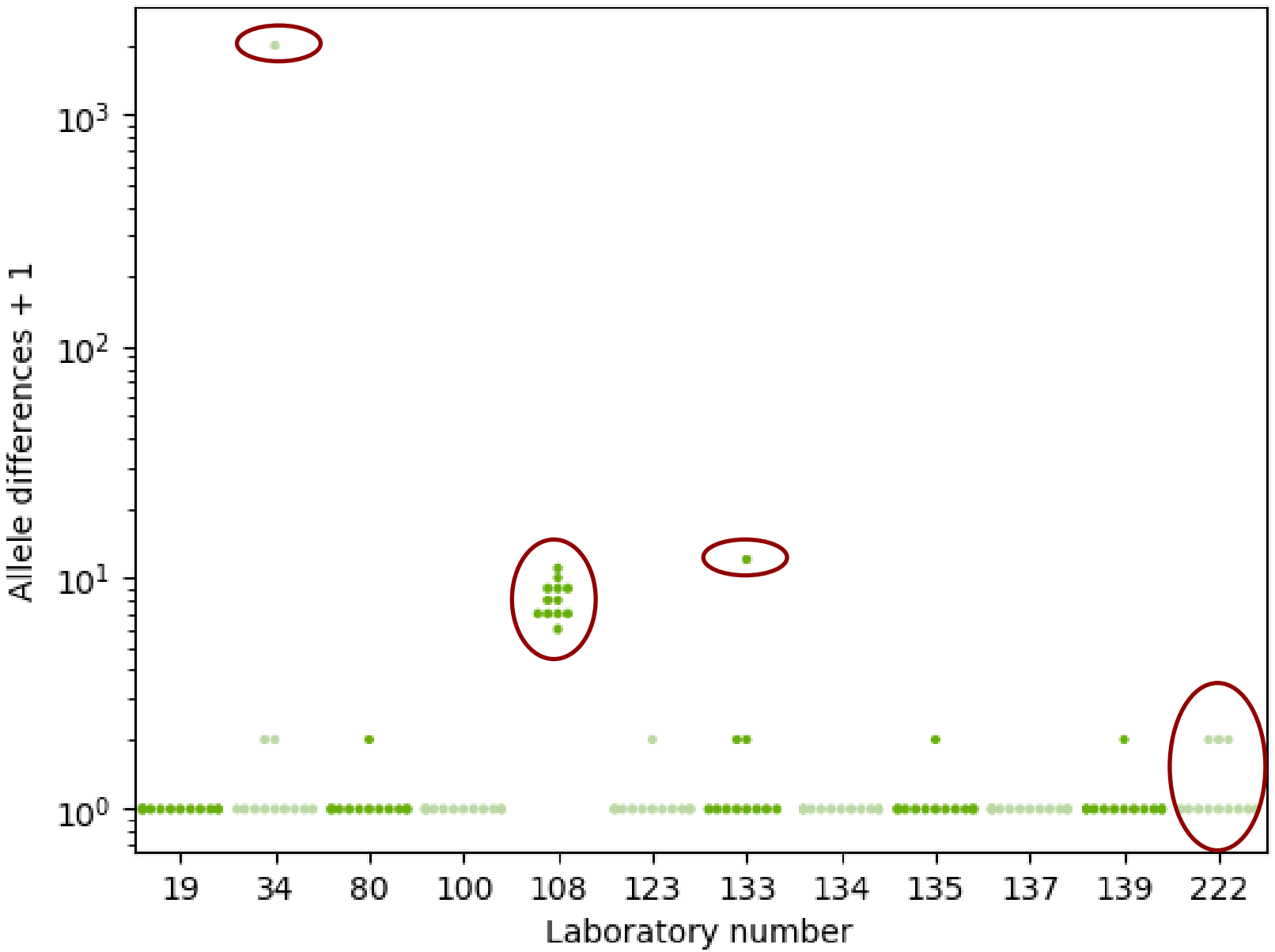
One sequence from Lab 133 discarded due to low quality

REF19 from Lab 34
Not an EQA isolate

Lab 108 show high number of differences
(Ion Torrent data)

Lab 222 REF16 ”identical” with the EQA provider (20 missing alleles)
(Ion Torrent data/lower number of loci in the scheme)

Allele difference from EQA provider “REF” (STEC)



Cluster analysis (WGS)

STATENS SERUM INSTITUT

- Allele based increased from 73% (2017-2018) to 83% (2018-2019)
- SNP based decreased from 27% to 17% (n=2)

- In general the performance were high ~ 83% ….. (92%)
- The laboratories that identified the correct cluster, reported similar results as the EQA provider (allelic difference /SNP distances within the cluster)

- Both SNP and allele based methods seem to be useful for interlaboratory comparability
  - cgMLST results were at a comparable level (very high degree of homogeneity) within each species, as long it is analyzed with the same scheme
  - The two SNP results could be used for interlaboratory comparability and communication
  - The reported results only gives a hint that using only 2360 loci (instead of 2513 as in Enterobase) and perhaps a different allele calling gives this impact on the conclusions.

Published Reports (2012 ->

STATENS SERUM INSTITUT

Listeria

STEC

Salmonella

https://www.ecdc.europa.eu/en/publications-data

Acknowledgements

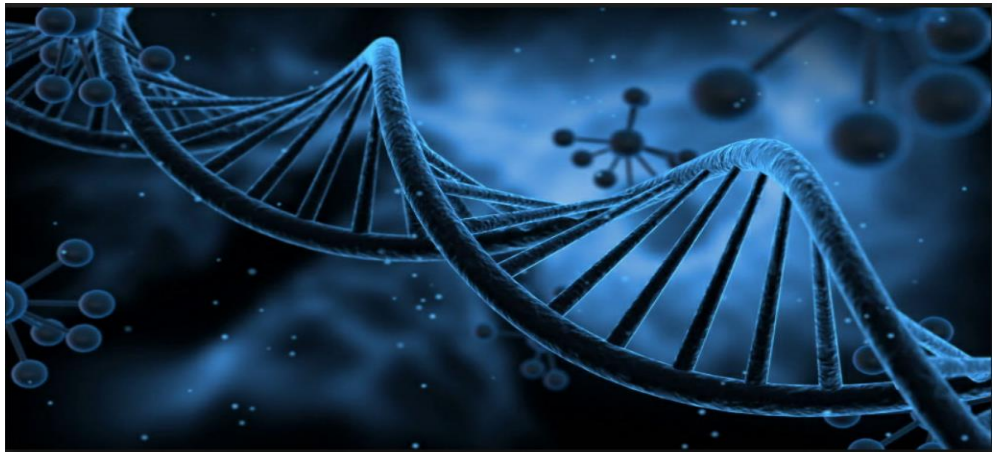
All the Public Health Reference Laboratories participating in the STEC EQA

Taina Niskanen

EQA-team STEC:
Flemming Scheutz, Gitte Sørensen and Kristoffer Kiil
Section of Foodborne Infections

STATENS SERUM INSTITUT

THANK YOU FOR YOUR ATTENTION



SUSANNE SCHJØRRING (SSC@SSI.DK)

Cluster analysis (PFGE)

		Laboratory number							
Isolate ID	**ST**	**19**	**90**	**123**	**124**	**127**	**130**	**132**	**222**
REF13‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF14‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
REF15	386	No	No	No	No	No	No	Yes	No
REF16‡#	17	Yes	Yes	Yes	Yes	Yes	No	No	Yes
REF17	20	No	No	No	No	No	No	No	No
REF18	20	No	No	No	No	No	No	Yes	No
REF19	17	No	No	No	No	No	Yes	Yes	No
REF20	17	No	No	No	No	No	Yes	Yes	No
REF21	17	No	No	No	No	No	Yes	Yes	Yes
REF22‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF23	17	No	No	No	No	No	Yes	Yes	No
REF24‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster-identified conclusion		Yes	Yes	Yes	Yes	Yes	No	No	No

‡: Closely related isolates (5 isolates) #: Technical triplicated isolates