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

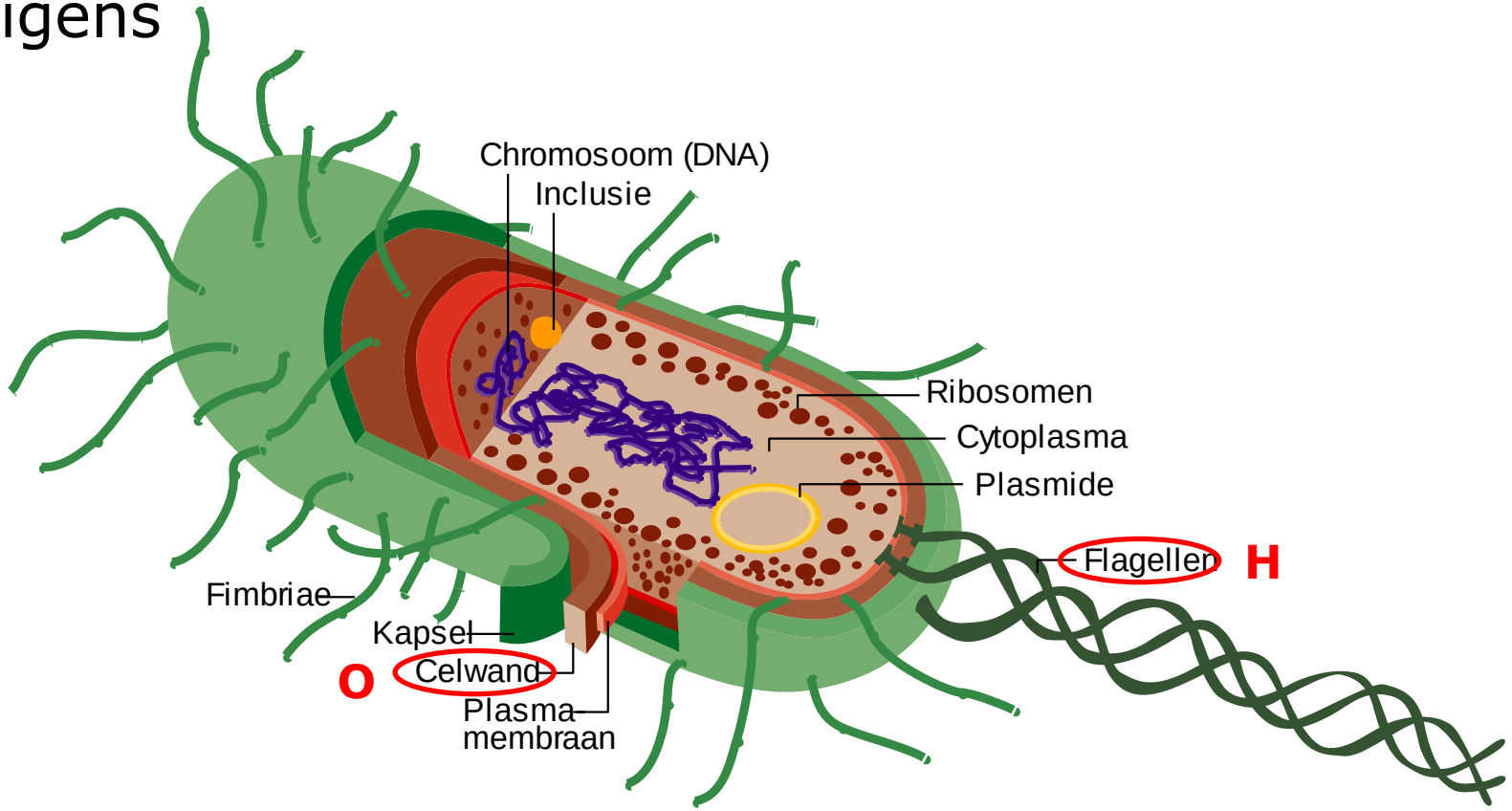
In-house validation of Salmonella serotyping with WGS

Maaike van den Beld

EURL workshop | 28-5-2021

Salmonella serotyping

Detect specific antigens



Drawing: W. P. M. Hoekstra, from Microbiologie, p. 50

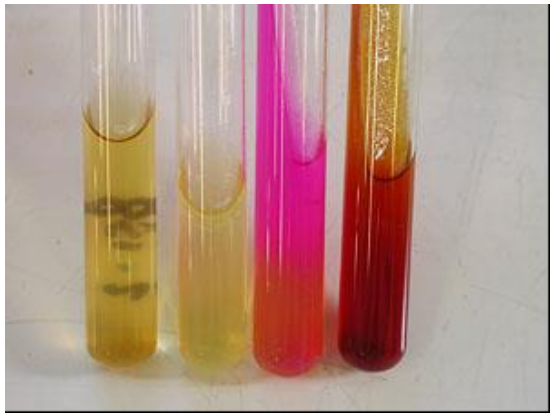
Salmonella serotyping at RIVM until 2020

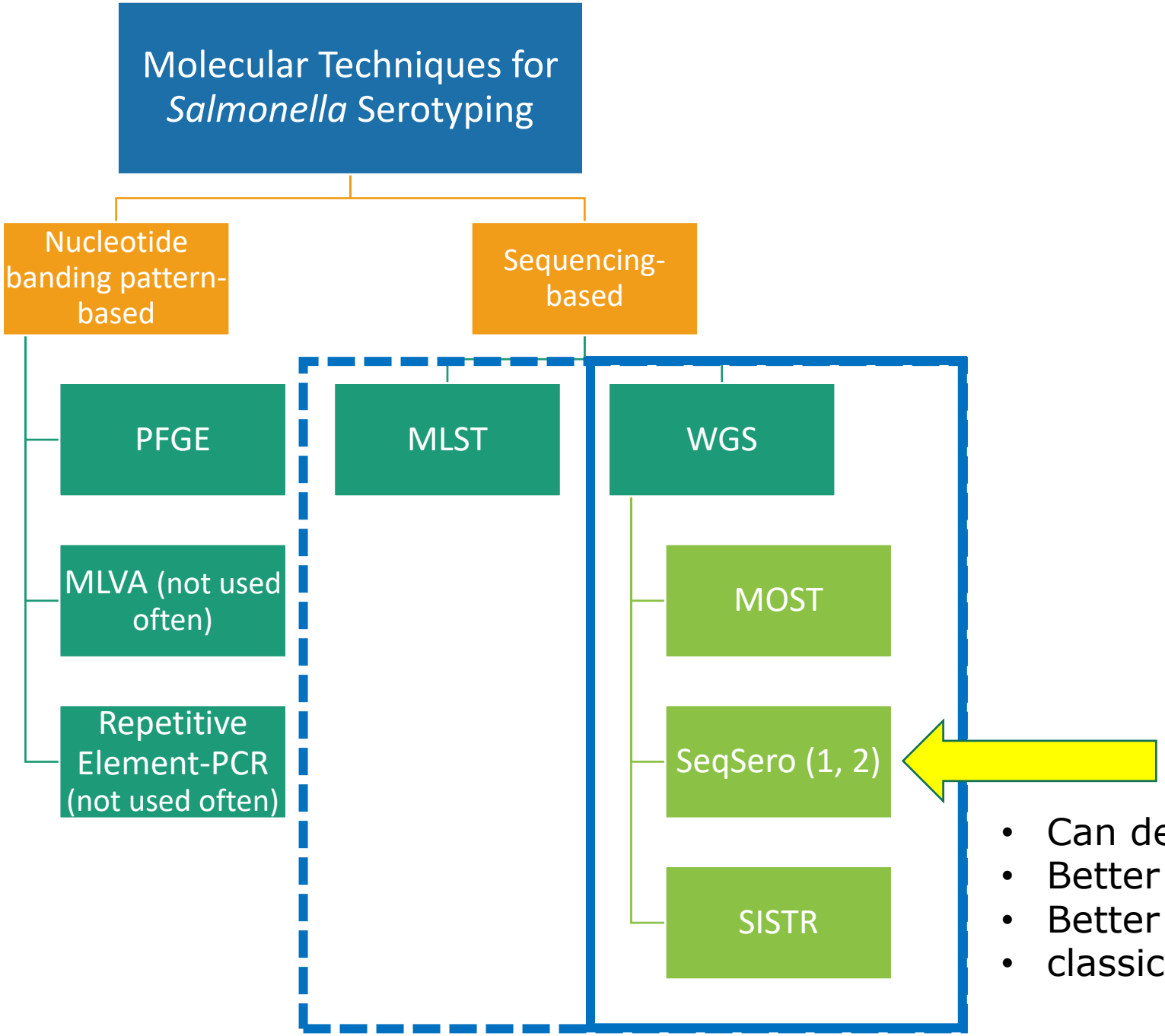
Screening with xMAP assay Luminex

Followed by classical slide agglutination with approx.
70 O-antisera and 100 H-antisera

Salmonella serotyping at RIVM until 2020

To distinguish different subspecies or confirm genus *Salmonella*
biochemical tests and MALDI-TOF

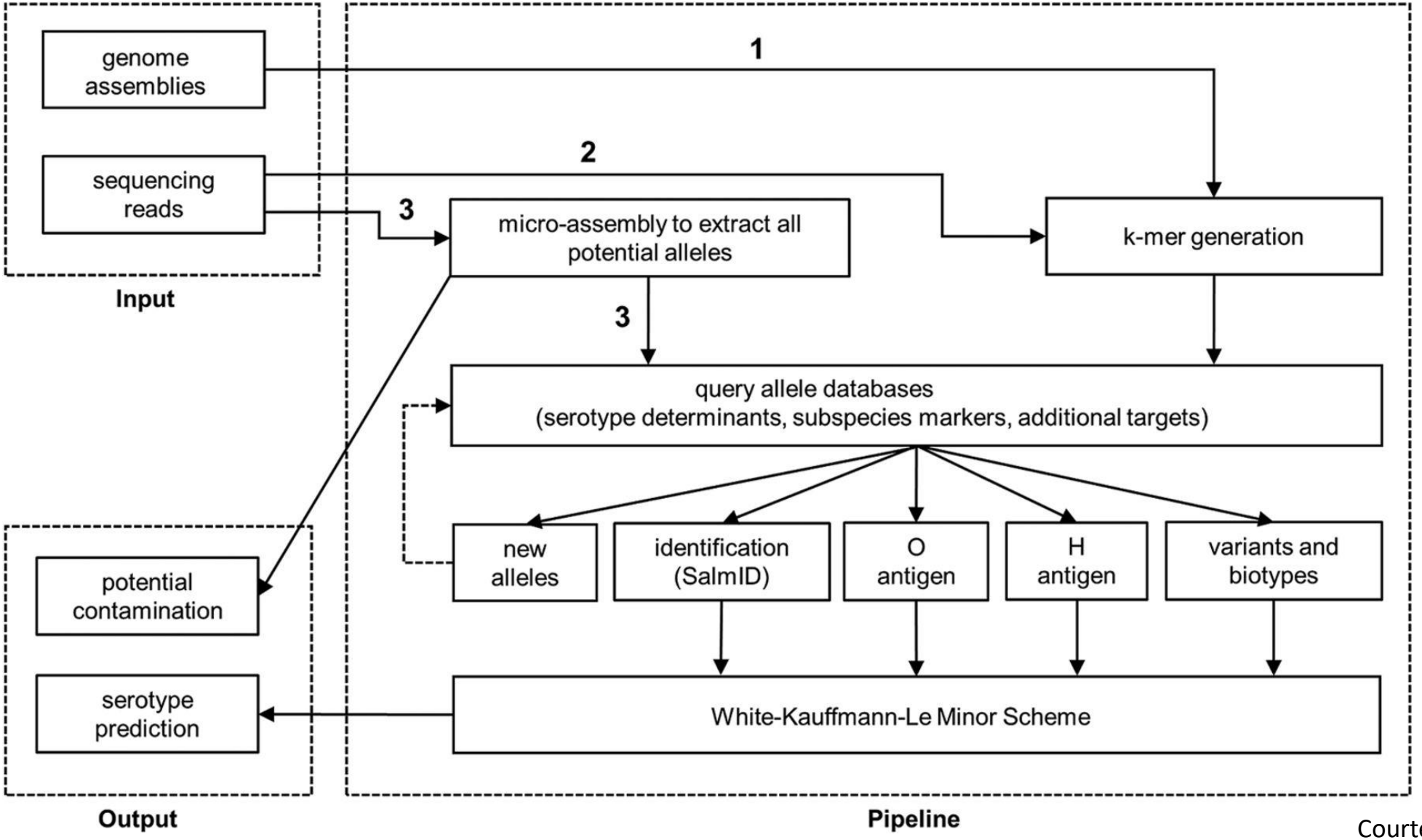




- Can detect mixed samples
- Better at detecting monophasic serotypes
- Better backwards compatibility with
- classical serotyping

Courtesy of Alejandra Hernandez Segura

SeqSero2 workflows



1) k-mer generation from genome assembly

2) k-mer generation from raw reads

3) Microassembly from raw reads

Figure from Zhang et al., 2019, Appl Environ Microbiol 14;85(23):e01746-19

Courtesy of Alejandra Hernandez Segura

Performance of the three SeqSero2 workflows

Prediction result	Raw read k-mer (no. [%])	Allele microassembly (no. [%])	Genome assembly (no. [%])
Expected serotype[a]	2,198 (96.4%)	2,250 (98.7%)	2,246 (98.5%)
Unexpected serotype[b]	73 (3.2%)	19 (0.8%)	23 (1.0%)
Partial or no serotype[c]	9 (0.4%)	11 (0.5%)	11 (0.5%)
Results of all tests	2,280	2,280	2,280

Figure from Zhang et al., 2019, Appl Environ Microbiol 14;85(23):e01746-19

Courtesy of Alejandra Hernandez Segura

In-house Salmonella pipeline

- Based on micro-assembly mode SeqSero 2
- Built on Snakemake 5.20.1
- Uses Miniconda and mamba to resolve software dependencies
- Locally installable
- Automated analysis
- Input: Illumina data, already filtered and trimmed with Trimmomatic (Juno assembly pipeline result)

Output, summarized and specific

Sample name	O antigen prediction	H1 antigen prediction(fliC)	H2 antigen prediction(fljB)	Predicted subspecies	Predicted antigenic profile	Predicted serotype	Potential inter-serotype contamination	Note
Sample 1	4	l,v	e,n,z15	I	4:l,v:e,n,z15	Brandenburg	no	
Sample 2	7	d	l,w	I	7:d:l,w	Livingstone	no	
Sample 3	9	d	z6	I	9:d:z6	Zega	no	
Sample 4	11	l,v	1,2	I	11:l,v:1,2	Stendal	no	
Sample 5	50	r	1,5,7	IIIb	50:r:1,5,7	IIIb 50:r:1,5,(7)	no	

O_contigs:

NODE_3_length_1313_cov_104.979773 O-11_wzy-new-from-CP019192; blast score: 1890.24 identity%: 98.24%; alignment from 1 to 1079 of antigen

H_contigs:

NODE_1_length_1745_cov_93.043165 fljB 1,2; blast score: 2760.01 identity%: 99.41%; alignment from 1 to 1520 of antigen

NODE_2_length_1351_cov_69.616170 fliC l,v; blast score: 2305.74 identity%: 83.03%; alignment from 256 to 1502 of antigen

Notes

- sdf gene not detected. The predicted serotypes share the same general formula: 9:g,m:-
- sdf gene detected
- The predicted serotypes share the same general formula: 4:a:1,5
- Detected a deletion that causes O5- variant of Typhimurium
- The SNP that causes d-Tartrate nonfermentating phenotype of Paratyphi B was not detected
- Detected the SNP for d-Tartrate nonfermenting phenotype of Paratyphi B
- This predicted serotype is not in the Kauffman-White scheme.

Notes signals

- The input genome cannot be identified as Salmonella. Check the input for taxonomic ID, contamination, or sequencing quality.

- Co-existence of multiple serotypes detected, indicating potential inter-serotype contamination. See 'Extracted_antigen_alleles.fasta' for detected serotype determinant alleles.

- O antigen was not detected. This result may be due to a rough strain that has deleted the rfb region.

Standards validation ISO 16140-6:2019

Method comparison study:

<u>Inclusivity study</u>

genus/species: 150 target isolates

serovar: if more than 11 > minimum 250 isolates (5 per serovar)

<u>Exclusivity study</u>

genus/species: 100 non-target isolates (minimum 75 from same family, minimum 2 *S.* bongori)

serovar: 100 non-target isolates (minimum 25 from same family, minimum 75 non-target serovars)

Standards validation ISO 15189:2012

ISO 15189 ACCREDITATION

MEDICAL LABORATORY

Accuracy: Degree of agreement with golden standard

- Golden standard: traditional serotyping

Measurement trueness: Degree of closeness of the measurement to its identity

- For example: retesting proficiency tests

Analytical specificity: Degree of interfering substances that can result in false-positives

- For example: unintentional cross reactions with non-Salmonella

Measurement precision: Degree in dispersal in multiple measurements in one sample

- Repeat same analysis multiple times

ISO 15189:2012: measurement trueness

100%

O8 serotypes

O8 serotypes sometimes differ by presence of O6 only

O6 is variably expressed

Genetically they are indistinguishable (Mikoleit 2012, Zhang 2019)

Table S3. Summary of optimized interpretation of antigenic formula implemented in SeqSero2

Antigen formula	SeqSero1 prediction	SeqSero2 prediction
8:k:1,5	Haardt or Blockley	Blockley
8:r:1,5	Hindmarsh or Bovismorbificans	Bovismorbificans
8:z10:e,n,x	Hadar or Istanbul	Hadar
8:l,v:1,2	Pakistan or Litchfield	Litchfield
8:d:1,5	Yovokome or Manhattan	Manhattan
8:d:1,2	Virginia or Muenchen	Muenchen
8:e,h:1,2	Newport or Bardo	Newport

Figure from Zhang et al., 2019, Appl Environ Microbiol 14;85(23):e01746-19

Inclusivity study/accuracy

Retrospective study (n=332)

1. 10 most found serotypes from different sources
2. Multiple serotypes from all subspecies
3. Complemented to include all O- and H-types

Prospective study (n=171)

All samples from Oct 2020 ran in duplicate: traditional and WGS

Total (n=503)

181 serotypes, 2 species, 6 subspecies (125 serotypes from subsp. I)

Accuracy/inclusivity (n=503)

With exclusion of non-expression of genetic factors:

(sub)species: 503 of 503 in accordance = 100%

Serovar: 500 of 503 in accordance = 99.4%

Inclusivity study

(sub)species: inclusivity deviation is 0

Serovar: inclusivity deviation is 3

Exclusivity study/analytical specificity (n=100)

Listeria monocytogenes (n=15)

Escherichia coli (n=20)

Shigella flexneri (n=8)

Shigella sonnei (n=12)

Klebsiella pneumonia (n=15)

Enterobacter cloacae complex (n=15)

Citrobacter freundii (n=15)

analytical specificity/excusivity (n=100)

None of non-Salmonella rendered a serovar

“The input genome cannot be identified as Salmonella. Check the input for taxonomic ID, contamination, or sequencing quality”

Analytical specificity: 100%

Exclusivity deviation: 0

ISO 15189:2012: measurement precision

Comparison of md5sums

Algorithm that uses hashes, identical files have identical hashes.

All md5sums are identical

100%

Conclusions

Fulfilled ISO16140-6 criteria (accept genes vs antigens)

Fulfilled ISO15189 criteria

Implemented as standard serotyping method at RIVM as of 1 January 2021

Ended Luminex screening, traditional agglutination still available

Anjo Verbruggen
Kim van der Zwaluw
Thijs Bosch

Eelco Franz

Kirsten Mooijman
Wilma Jacobs-Reitsma
Angela van Hoek

thank you

Alejandra Hernandez Segura

Sjoerd Kuiling
James Groot
Elske van Logchem
Sietze Brandes
Linda van de Nes