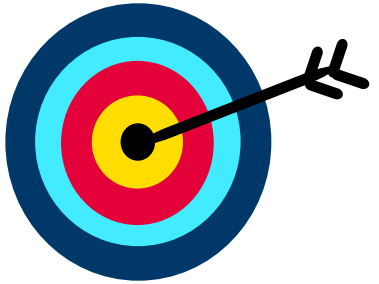
German Federal Institute for Risk Assessment

A *Salmonella* cgMLST validation study for accreditation in Germany



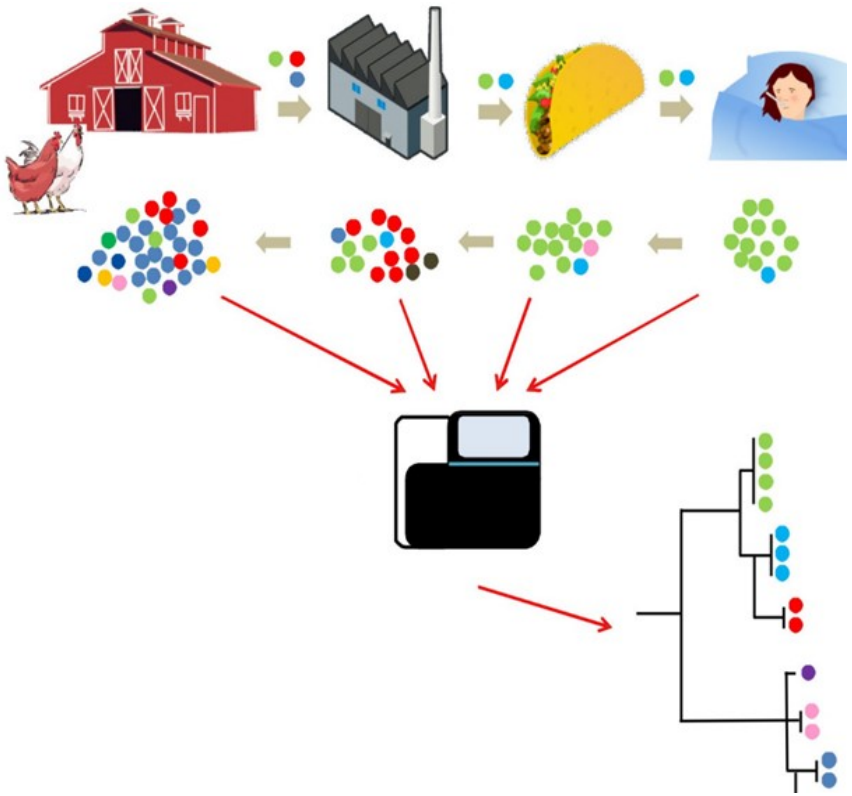
23.05.2023, EURL *Salmonella* Workshop 2023

Marina Lamparter
NRL for *Salmonella* Germany
Unit 42: Food Microbiology, Pathogen-Host Interactions
Dep. Biological Safety

National Reference Laboratory for *Salmonella* Germany

3000 - 4000 isolates per year

- Official veterinary & food inspection laboratories, animal health service
- Universities, research laboratories
- Private laboratories



Application of Whole Genome Sequencing for the Detection of Foodborne Disease Outbreaks, (BfR-Wissenschaft 1/2020), modified

Since 2018: Application of NGS-technology for Whole-Genome Sequencing (WGS)

>7000 sequenced *Salmonella spp* (May 2023)

Outbreaks in Germany

Foodborne outbreaks - dashboard

CAUSATIVE AGENTS

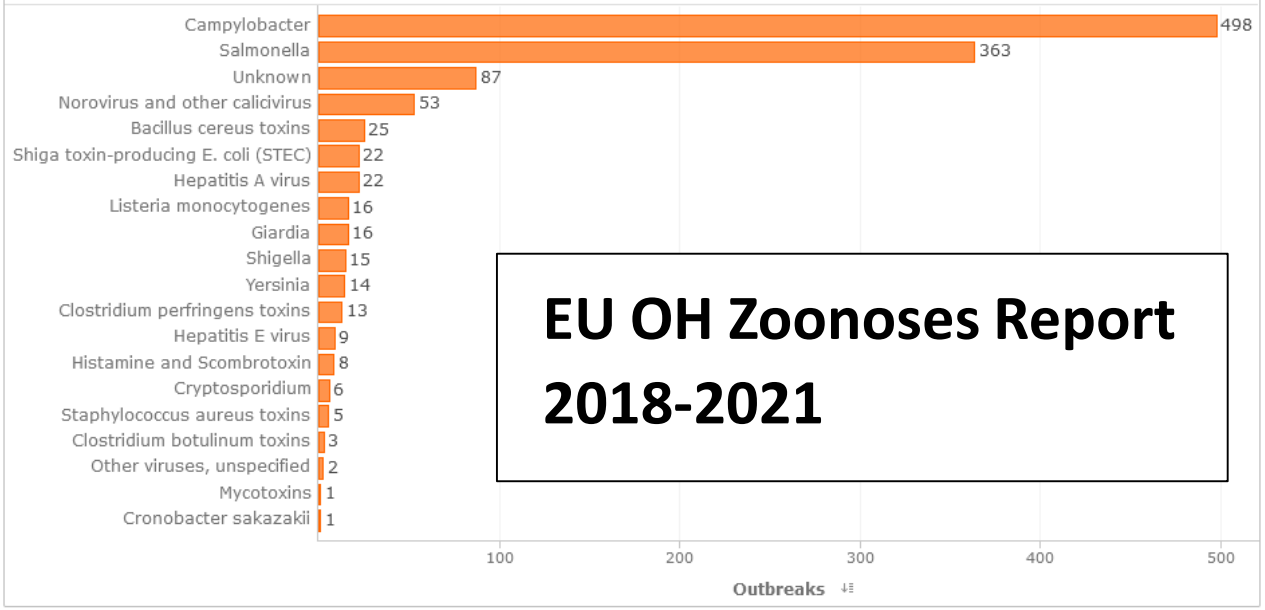
Caution in data interpretation

Reporting year	Strength of evidence	Type of outbreak	EU and non EU	Reporting country
2021, 2020, 2019, ...	(All)	(All)	(All)	Germany

Number of countries	Outbreaks	Cases	Hospitalisations	Deaths
1	1,179	6,768	1,079	22

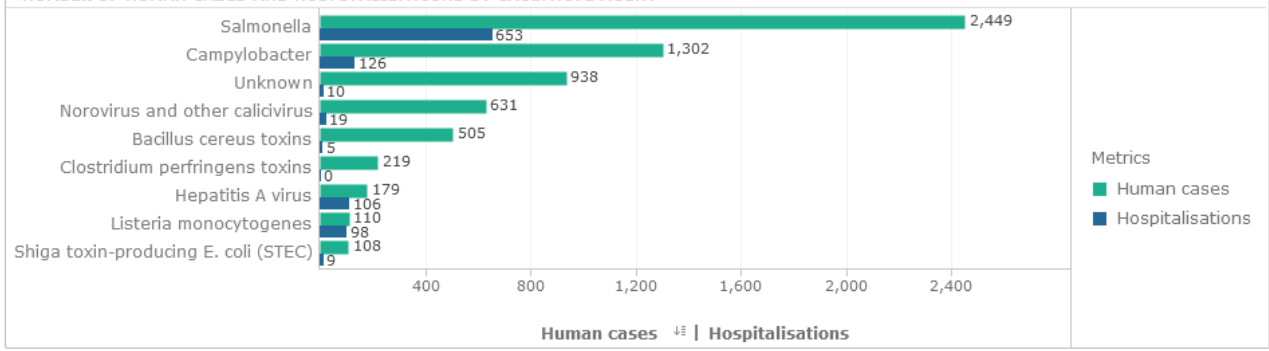
Each graph can be maximized with the ⤢ icon that appears in the right corner of the titlebar.

NUMBER OF OUTBREAKS BY CAUSATIVE AGENT

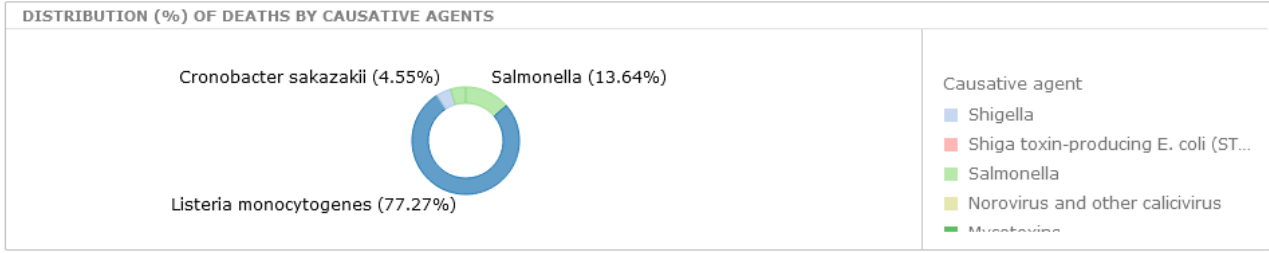


EU OH Zoonoses Report 2018-2021

NUMBER OF HUMAN CASES AND HOSPITALISATIONS BY CAUSATIVE AGENT



DISTRIBUTION (%) OF DEATHS BY CAUSATIVE AGENTS



European Food Safety Authority. (2022). User Guide – Dashboard on foodborne outbreaks. Zenodo. https://doi.org/10.5281/zenodo.7313523 Screenshot 07.02.2023

Outbreak Analysis at NRL for *Salmonella*

OuReMa

humane ReferenceSQ

Case	chewiesnake status	chewiesnake last update	chewiesnake report	Samples extern	Related BfR Samples	Related BfR Samples 10 alleles	minimum cgMLST dis BfR data
EURL_PT_2022	success	Fri Feb 3 04:42:43 CET 2023	cgMLST report	1	8	4	1
FWD-2022-00072_S.Ball	success	Fri Feb 3 04:43:35 CET 2023	cgMLST report	1	30	26	2
Mbandaka_2022_EFSA	success	Fri Feb 3 04:44:27 CET 2023	cgMLST report	2	4	0	18
SAL_Amsterdam_2021_NGS_1_Tahini	success	Fri Feb 3 04:45:26 CET 2023	cgMLST report	1	8	8	0
SAL_Chester_2021_NGS_1	success	Fri Feb 3 04:46:30 CET 2023	cgMLST report	1	10	5	3
SAL_Chester_2021_NGS_2	success	Fri Feb 3 04:47:21 CET 2023	cgMLST report	1	10	4	0
SAL_Choleraesuis_2022_NGS_1	success	Fri Feb 3 04:48:13 CET 2023	cgMLST report	1	40	10	2

All comparisonDB distances to all query samples below threshold

All distances above 500 have been omitted.

Copy | Print | Download | Column visibility

Show 10 entries

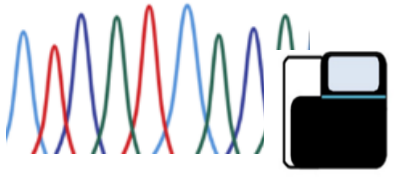
Sample_query	Sample_comparisonDB	Distance
All	All	All
SA21-00939	20-SA00210-0	0
SA21-00939	22-SA01789-0	0
SA21-00939	20-SA01732-0	2
SA21-00939	21-SA01490-0	2
SA21-00939	21-SA01548-0	2
SA21-00939	21-SA02768-0	3
SA21-00939	21-SA02771-0	3
SA21-00939	21-SA02769-0	4

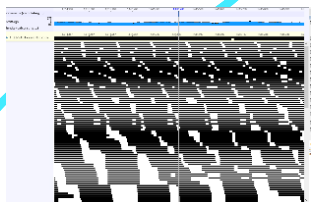
Screenshot OUREMA Pipeline, February 4th, 2023

NGS test procedures

	Test Procedures	Tools
wet-lab	DNA isolation and quantification for NGS	
wet-lab	Library preparation	
wet-lab / dry-lab	WGS using short-read sequencing for typing and characterization of bacterial isolates (incl. QC & assembly)	**A**quamis
dry-lab	Genome comparison of bacterial isolates using WGS data (cgMLST) N E W	**B**akcharak **C**hewieSnake

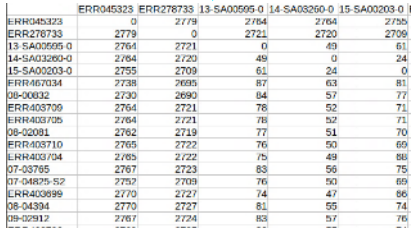
NGS documents by 4NSZ





Conda activate

ABC Pipeline



BfR

cgMLST at the NRL for *Salmonella*

Assembled sequences screened for loci (scheme), variants (alleles) are identified (database)

and translated to isolate-specific profiles, plus QC, profile comparison and dendrograms/matrices

ChewieSnake

In-house developed pipeline, implemented in automated outbreak analysis (OUREMA) in house, transparent

Scheme: Enterobase (modified) and EFSA

Deneke et al. 2021 Decentralized Investigation of Bacterial Outbreaks Based on Hashed cgMLST

RIDOM Seqsphere+

Commercial software

Scheme: Enterobase

Jünemann et al. 2013 Updating benchtop sequencing performance comparison

Test procedure - Main document

General Method

ChewieSnake

RIDOM Seqsphere+

Genome comparison of bacterial isolates using WGS data

- ✓ Application
- ✓ Principles
- ✓ Procedure (ChewieSnake and RIDOM Seqsphere+)
- ✓ Evaluation
- ✓ QC measures
- ✓ Maintenance, extension and functional controls
- ✓ Validation and verification

…

Attachment 01: Genome comparison analysis of *Listeria monocytogenes* using cgMLST

Attachment 02: Genome comparison analysis of *Salmonella enterica* using cgMLST

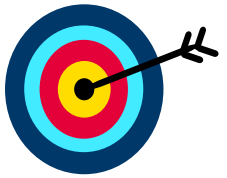
Test procedure - Attachment

Analysis of *Salmonella enterica*

Attachment 02: Genome comparison analysis of *Salmonella enterica* using cgMLST

- ✓ Purpose
- ✓ Procedure for *Salmonella enterica*
  - ✓ ChewieSnake: Enterobase or EFSA scheme
  - ✓ RIDOM Seqsphere+: Enterobase scheme
- ✓ Documentation
- ✓ QC measures
- ✓ Evaluation
- ✓ Validation and Verification

Validation – *Salmonella* analysis



Parameter

- Repeatability
- Reproducibility
- Robustness
- Trueness

Data Set

- Genus specific defined validation data (Aquamis) by 4NSZ
- EURL PT cluster 2020
- International Outbreak data

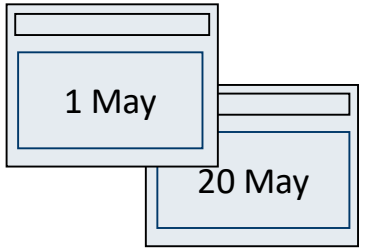
Material & Methods: Repeatability, Reproducibility & Robustness

Repeatability

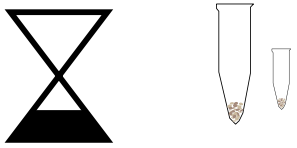
Test data

ID	Serovar	Matrix Category
SA-1	Typhimurium	Food
SA-2	Infantis	Feed
SA-3	Enteritidis	Environment
SA-4	Typhimurium monophasic	Animal

Reproducibility



Robustness



BfR

Material & Methods: Repeatability & Reproducibility

Basic Sequencing data - GS

Repeatability - WP

1) Same condition sample preparation and sequencing

Reproducibility - VP

1) Library prepared by different persons, sequenced in same run

2) Same library sequenced in different run

3) Different library prep and sequencing

Example Repeatability

ID	Library Prep	Sequencing
SA-2-GS	Technician A 20.05.21	NextSeq 2x149 cycles 21.05.2021 Technician A
SA-2-WP1		
SA-2-WP2		
SA-3-GS		
SA-3-WP1		
SA-3-WP2		

Example Reproducibility (same library, different runs)

ID	Library Prep	Sequencing date
SA-2-GS	Technician A 20.05.2021	21.05.2021
SA-2-VP2		11.06.2021
SA-4-GS		21.05.2021
SA-4-VP2		11.06.2021

BfR

Material & Methods: Robustness

Robustness - RB

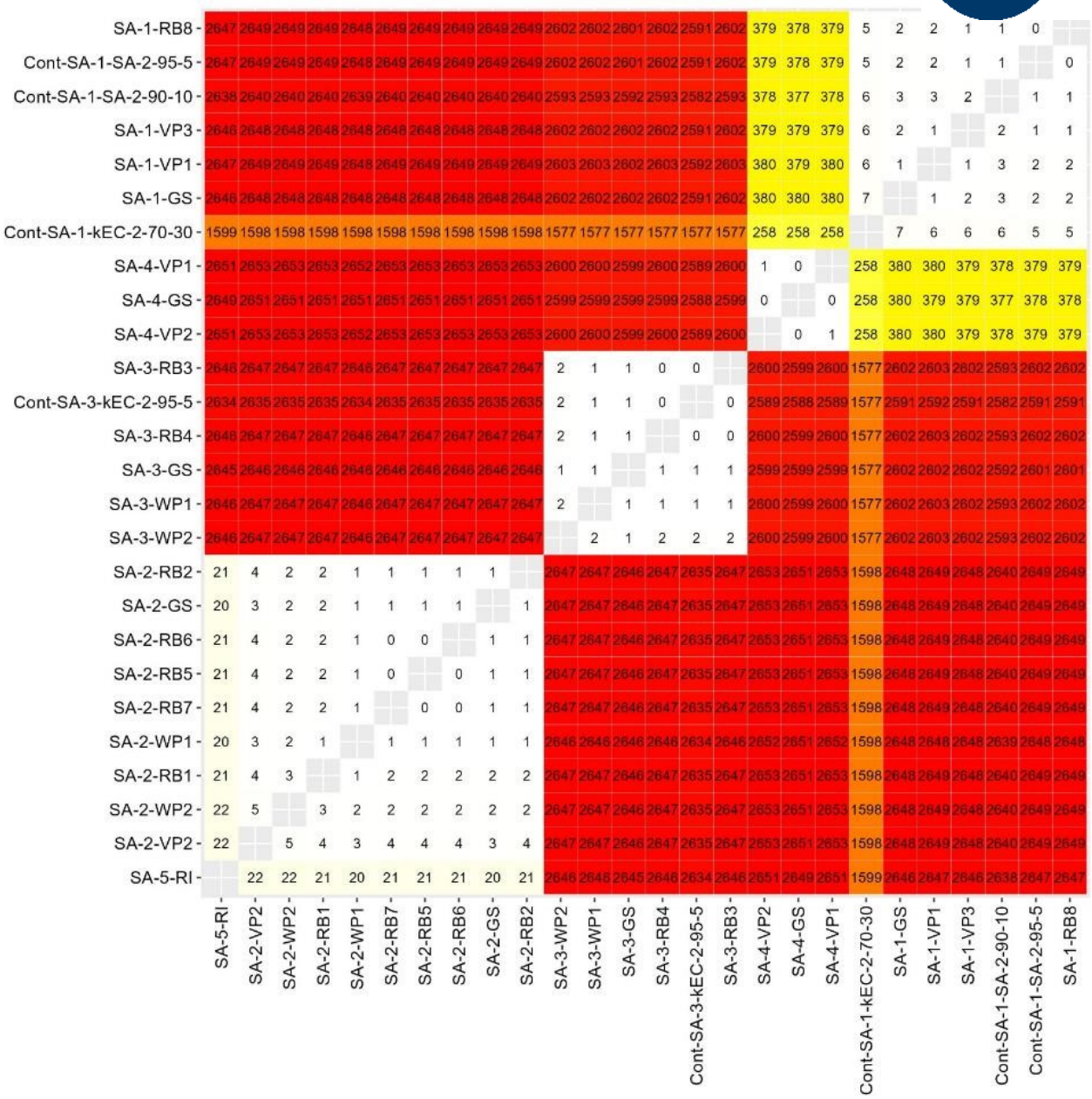
1) Full and half library volume
2) Low DNA concentration/higher amplification number
3) Different sequencer
4) Read length
5) Different indices kits
6) Expired kits
7) WGS and amplicon library mix
8) Detection of contaminated samples (Selectivity)

Example Robustness (expired kits)		
ID	**Library Prep**	**Sequencing**
SA-2-RB6	Technician A 01.06.2021 CD Indices	iSeq, 2 x 151 cycles, 04.06.2021, Technician A
SA-2-RB7		iSeq, 2 x 151 cycles, 08.06.2021, Technician A

Results: Repeatability, Reproducibility, Robustness

ChewieSnake Enterobase

- Repeatability: 0-2 AD
- Reproducibility: 0-3 AD
- Robustness: 0-2 AD
- 0-3 AD between approaches



Screenshot Distance Matrix Validation Results, ChewieSnake Enterobase

Material & Methods: Trueness

Re-analysis of NRL for *Salmonella* samples from:

A) EURL proficiency test cluster analysis 2020

- ChewieSnake Enterobase
  - v 2.0.0
  - v 3.1.1
- Ridom Seqsphere+ Enterobase

B) International Outbreak cluster

- ChewieSnake Enterobase
  - v 3.1.1
- Ridom Seqsphere+ Enterobase

Trueness

Reference-Results from EURL *Salmonella*

Results: Trueness

A) EURL PT cluster analysis 2020

S. Typhimurium (monophasic)

- Cluster (threshold ≤ 6 AD) correctly identified
- Max. distance to reference results: (AD <10) was 1

Tabelle 50: *ChewieSnake* cgMLST Distanzmatrix (Enterobase-basiertes Schema, 3000 Loci), im EURL PT 020 eingereichte Ergebnisse, ChewieSnake Version 2.0.0

	SCA10	SCA07	SCA01	SCA05	SCA08	SCA04	20SCA	SCA06	SCA03	SCA02	SCA09
SCA10	0	383	350	362	363	363	363	363	362	362	362
SCA07	383	0	315	328	329	330	329	329	327	328	322
SCA01	350	315	0	53	53	54	53	53	50	51	39
SCA05	362	328	53	0	0	1	0	0	34	34	58
SCA08	363	329	53	0	0	2	1	1	35	35	59
SCA04	363	330	54	1	2	0	2	2	36	35	60
20SCA	363	329	53	0	1	2	0	1	35	35	59
SCA06	363	329	53	0	1	2	1	0	35	35	59
SCA03	362	327	50	34	35	36	35	35	0	5	57
SCA02	362	328	51	34	35	35	35	35	5	0	58
SCA09	362	322	39	58	59	60	59	59	57	58	0

Tabelle 53: cgMLST Distanzmatrix des EURL (siehe auch Anhang 1)

	SCA10	SCA07	SCA01	SCA05	SCA08	SCA04	20SCA	SCA06	SCA03	SCA02	SCA09
SCA10	0	379	341	356	356	356	356	356	353	354	354
SCA07	379	0	311	327	327	328	327	327	323	325	320
SCA01	341	311	0	50	50	51	50	50	46	48	35
SCA05	356	327	50	0	0	1	0	0	33	34	5
SCA08	356	327	50	0	0	1	0	0	33	34	57
SCA04	356	328	51	1	1	0	1	1	34	34	58
20SCA	356	327	50	0	0	1	0	0	33	34	57
SCA06	356	327	50	0	0	1	0	0	33	34	57
SCA03	353	323	46	33	33	34	33	33	0	4	54
SCA02	354	325	48	34	34	34	34	34	4	0	56
SCA09	354	320	35	57	57	58	57	57	54	56	0

Results: Trueness

B) International Outbreak data from EURL *Salmonella*: NRL for *Salmonella* data (fastq) linked to 3 outbreaks (7 serovars)

- Comparable but NOT identical allelic distances!
- 0-3 AD (common cluster threshold <10 AD) observed between methods/workflows

Tabelle 55: Ergebnisse (Distanzmatrices) der Analysen des EURL, des BfR mit Ridom SeqSphere+der BfR-ChewieSnake Pipeline (Enterobase basiertes Schema, 3000 Loci, chewieSnake version 3.1.1) zum internationalen Ausbruchscluster UI-644 (*S.* Enteritidis). Keine Teilprobenangabe.

EURL	18-SA04214	20-SA00763	20-SA02116	21-SA00528	SRR11741596	SRR8333779
18-SA04214	0	3	1	4	1	0
20-SA00763	3	0	2	5	2	3
20-SA02116	1	2	0	3	0	1
21-SA00528	4	5	3	0	3	4
SRR11741596	1	2	0	3	0	1
SRR8333779	0	3	1	4	1	0

BfR Ridom	18-SA04214	20-SA00763	20-SA02116	21-SA00528	SRR11741596	SRR8333779
18-SA04214	0	3	2	5	1	0
20-SA00763	3	0	3	6	2	3
20-SA02116	2	3	0	4	1	2
21-SA00528	5	6	4	0	4	5
SRR11741596	1	2	1	4	0	1
SRR8333779	0	3	2	5	1	0

BfR chewie	18-SA04214	20-SA00763	20-SA02116	21-SA00528	SRR11741596	SRR8333779
18-SA04214	0	4	2	5	1	0
20-SA00763	4	0	4	7	3	4
20-SA02116	2	4	0	4	1	2
21-SA00528	5	7	4	0	4	5
SRR11741596	1	3	1	4	0	1
SRR8333779	0	4	2	5	1	0

Summary

- All regular samples showed more than 95% valid cgMLST targets
- Testing of repeatability, reproducibility and robustness with 4NSZ pipeline dataset applying ChewieSnake Enterobase- and EFSA scheme, and RIDOM with Enterobase scheme
  - 0-3 AD within testing
  - 0- 3 AD difference between applied methods
- Trueness-testing with ChewieSnake (v. 2.0.0 and v. 3.1.1) and RIDOM Seqsphere+ both with Enterobase scheme
  - EURL Reference data
  - Successful detection of clusters within trueness testing with PT 2020 (threshold 6 AD)
  - 3 International outbreaks (7 serovars): comparable but not identical results (0-3 AD)
  - **METHOD FIT FOR PURPOSE!**

BfR

Outlook

- Communication – communication – communication
  - Context of metadata
- Harmonization of workflows: Valuable help for confirmation
  - (e.g. GenoSalmSurv in Germany, EFSA WGS in Europe)

- PT 2021
- Skesa assembly
- SNP-analysis

Thank you for your Attention

Jennie Fischer

Angelina Groger

Istvan Szabo

& Team

NRL for *Salmonella*

Kirsten Mojimann

Wilma Jacobs-Reitsma

Angela van Hoek

Robin Diddens

EURL *Salmonella*

Sabine Haby

Jana Sachsenröder

Britta Kraushaar

Mibi QM

Maria Borowiak

Carlus Deneke

Katharina Thomas

and Team

Sequencing core unit - 4NSZ

Stefanie Lüth

Sylvia Kleta

NRL *Listeria*

Sandra Simon

Michael Pietsch

Robert Koch Institute