Campylobacter in the One Health WGS system.
- Comparing cgMLST Schemes.

Update from inter EURLsWG onNGS

Bo Segerman
EURL-*Campylobacter* workshop
Uppsala, October 22-23rd, 2024

SVA SWEDISH VETERINARY AGENCY
EURL Campylobacter
Co-funded by the European Union

Campylobacter cg/wg MLST schemes

- Requirements for a scheme to be implemented into the One Health WGS system:
  - To be publicly accessible
  - To be functionable with chewBBACA
  - Well populated with alleles
  - To be scientifically correct
- Additional considerations:
  - One scheme for both *C. coli* and *C. jejuni* or two separate?
  - Type of scheme, cg or wgMLST?

Summary of schemes *Campylobacter*

Scheme	General	Number of targets	Validated	Proposed cut-off
PubMLST ”Oxford scheme” Cody et al 2017 cgMLST scheme	Commonly used for both surveillance and outbreak detection Designed for both *C. jejuni* and *C. coli* Publically available from pubMLST Nomenclature in pubMLST and well populalated with alleles.	**Version 1:** 1,343 loci: a subset of the 1,643 loci in *C. jejuni* NCTC 11168, and present in >95% of 2,472 isolates of both *C. jejuni* and *C. coli* **Version 2:** 1,142 loci	Yes, tested on additional 1478 *C. jejuni* and *C. coli* isolates from Europe and North America. 99.5% of the isolates contained >95% loci	11 AD (LNS) 10 AD (PulseNet) 4 AD (SSI – national surveillance) EURL PT38 outbreak detection: 5, 10, 13 or 14.

Summary of schemes *Campylobacter*

Scheme	General	Number of targets	Validated	Proposed cut-off
cgMLST INNUENDO	Designed for population dynamics Designed *for C. jejuni* No continous deposit of new alleles Publically available from: https://zenodo.org/records/1322564	678 loci in *C. jejuni* Recommends only max 2% missing loci.	loci present in ≥ 99% of the strains tested. 6526 *C. jejuni* genomes tested	4 AD (L1)
wgMLST INNUENDO	Designed for outbreak detection Designed *for C. jejuni* Publically available from: https://zenodo.org/records/1322564 Avaliable from Chewie-NS	2795 loci in *C. jejuni*	6526 *C. jejuni* genomes tested	9 AD

Summary of schemes *Campylobacter*

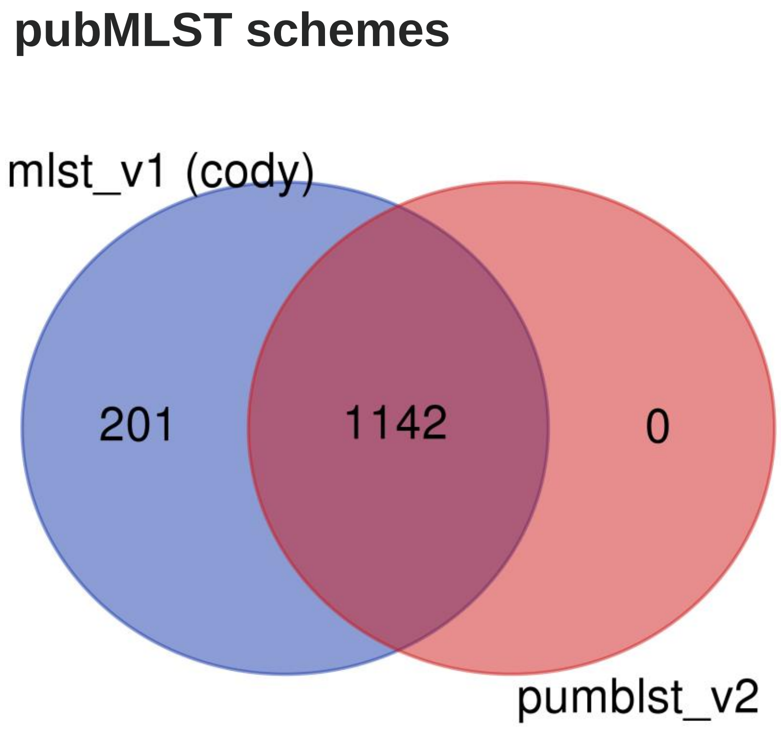
Scheme	General	Number of targets	Validated	Proposed cut-off
BioNumerics wgMLST scheme	Used by PulseNet in addition to the pubMLST scheme Not publically available (yet)	7-gene MLST, pubMLST + 5280 accessory loci from several camplobacter species = 6651 loci	?	10 AD
Ridom Seqsphere cgMLST scheme	Designed for both *C. jejuni and C. coli* Validated on a small set of strains Publically available from: Cgmlst.org	637 loci in *C. coli* and *C. jejuni* Designed from 12 complete genomes (2 *C. coli*)	Validated on 33 isolates from 3 outbreaks	13 AD
Ridom Seqsphere cgMLST scheme + accessory genes	Designed for both *C. jejuni and C. coli* Validated on a small set of strains Publically available from: Cgmlst.org	637 + 958 loci = 1595 loci Designed from 12 complete genomes (2 *C. coli*)	Validated on 33 isolates from 3 outbreaks	13 AD

- A very good correlatelation between wgMLST and cgMLST schemes
- In some cases, wgMLST has increased the resolution after a cluster has been detected

  - ➢ Suggest to use a cgMLST scheme for cluster detection of *Campylobacter* in the One Health WGS system

- All cgMLST schemes (PubMLST v1 and v2, Innuendo and Ridom) are publically accessible
- All cgMLST schemes are all functional with allele calling in chewBBCA
- Analysis of inclusivity of schemes for different genotypes of *C. jejuni* and *C. coli* is displayed in following slides

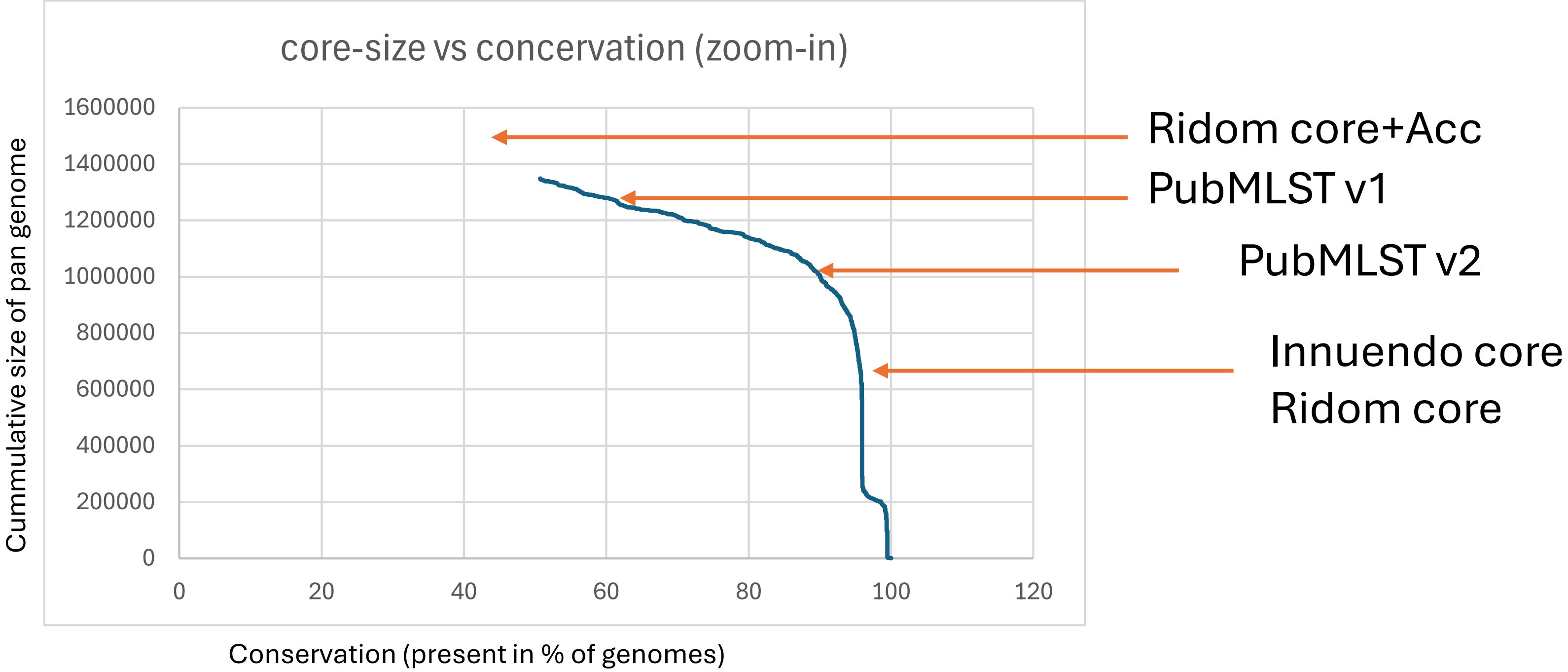
	loci	Alleles	Allels/loci
pubMLST_v1 (oxford,Cody)	1343	2003762	1492
pubMLST_v2	1142	1706531	1494
Innuendo cgMLST	678	169402	250
cgMLST.org (Ridom)	637	668154	1049



pubMLST schemes

Identical alleles exists for targets

Targets	Size (Mb)	overlapp (percent)	innuendo_core	PubMLST_v1	PubMLST_v2	Ridom_core	Ridom_coreAcc
678	0.63	innuendo_core	100,0	98,2	92,2	62,7	95,4
1343	1.26	PubMLST_v1	49,7	100,0	85,0	45,8	96,0
1142	1.06	PubMLST_v2	57,7	100,0	100,0	51,6	96,4
637	0.62	Ridom_core	66,6	96,4	92,3	100,0	100,0
1595	1.48	Ridom_coreAcc	40,9	80,8	69,0	39,9	100,0

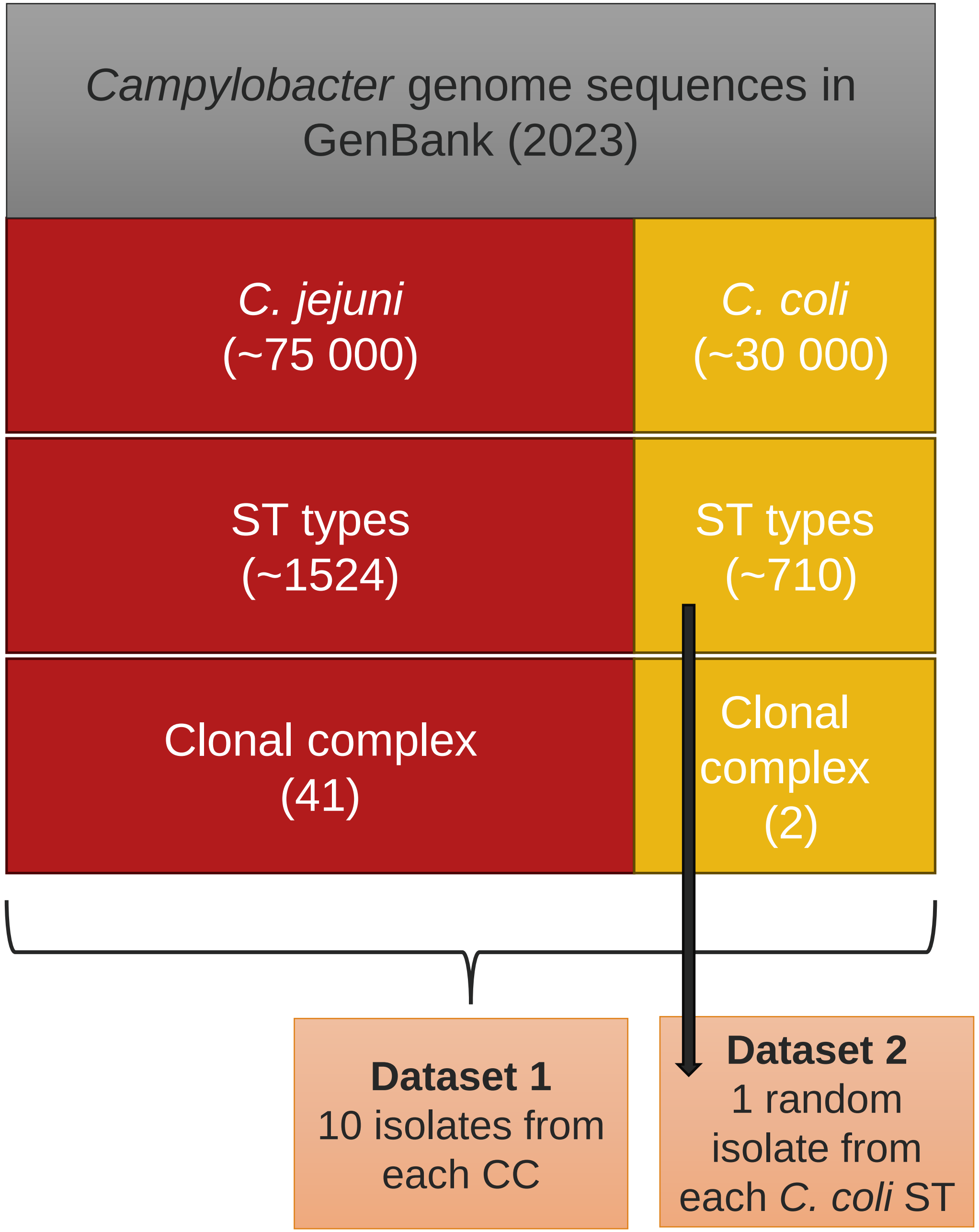


Ridom core+Acc

PubMLST v1

PubMLST v2

Innuendo core
Ridom core

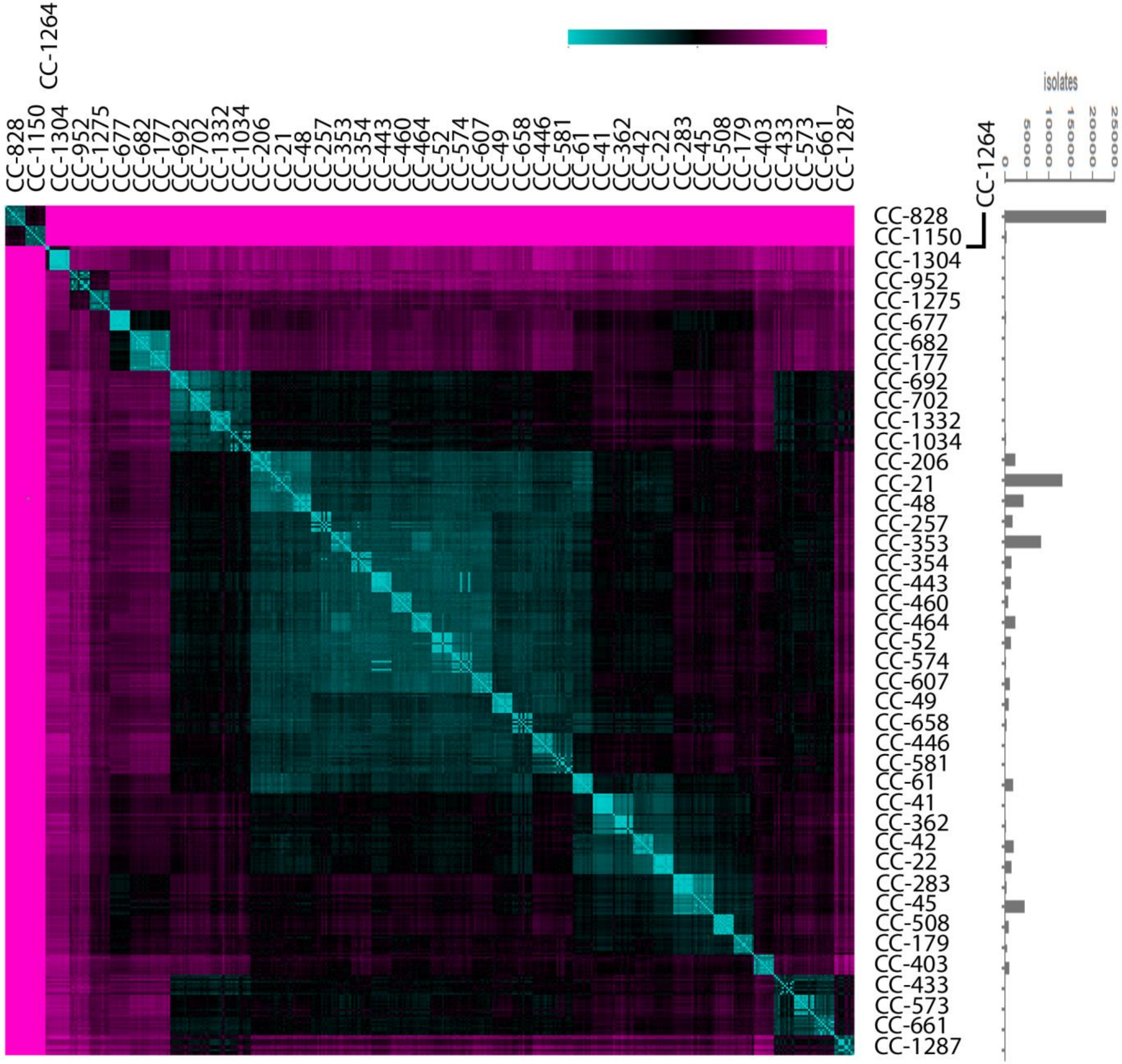


Campylobacter genome sequences in GenBank (2023)

C. jejuni (~75 000)	*C. coli* (~30 000)
ST types (~1524)	ST types (~710)
Clonal complex (41)	Clonal complex (2)

Dataset 1
10 isolates from each CC

Dataset 2
1 random isolate from each *C. coli* ST



Exact match alleles Dataset 1 (CC)

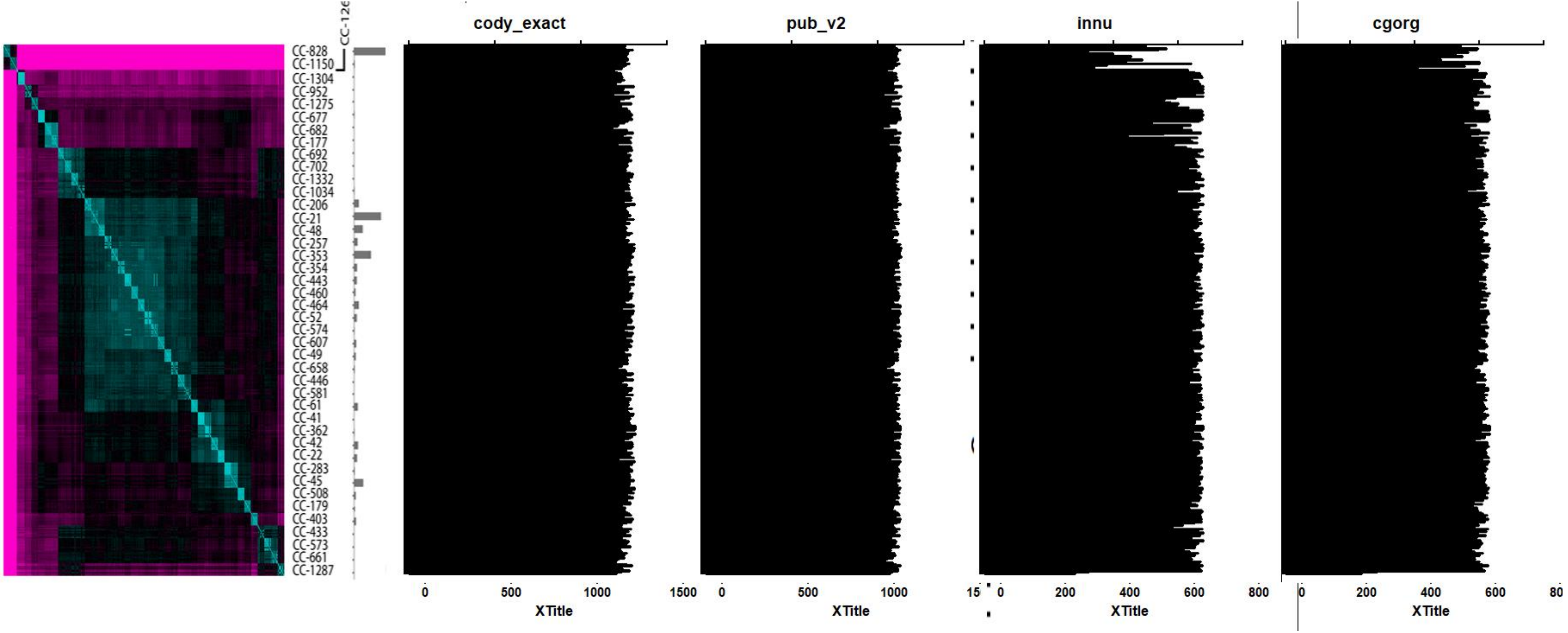
C.jejuni+C.coli

pubMLST_v1
1343 loci

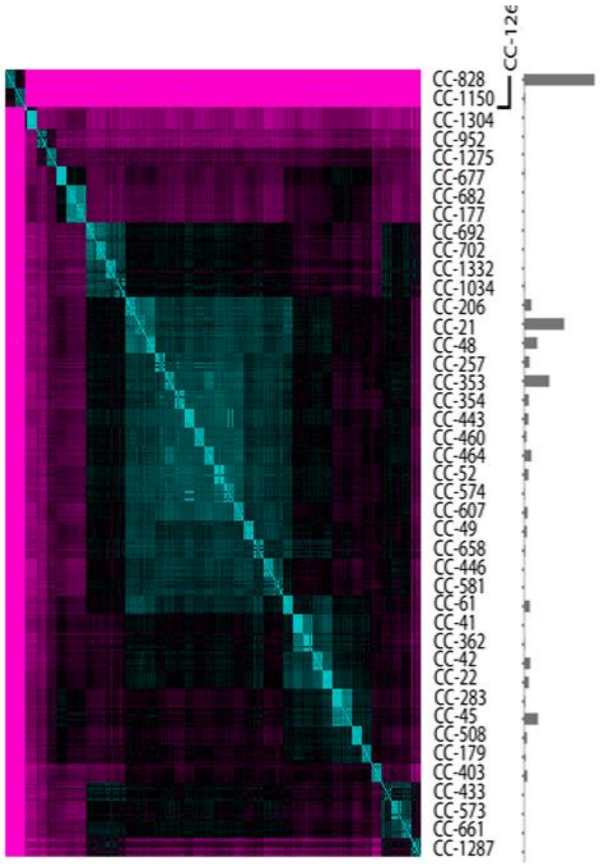
pubMLST_v2
1142 loci

Innuendo cgMLST
678 loci

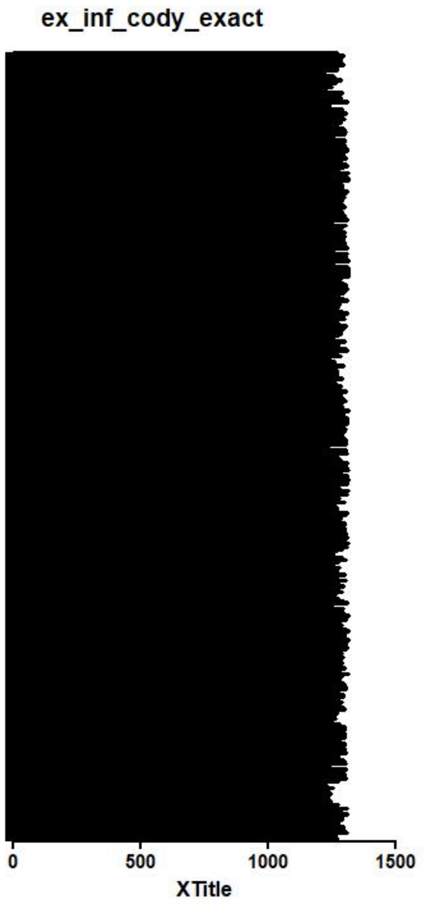
Ridom cgMLST
637 loci



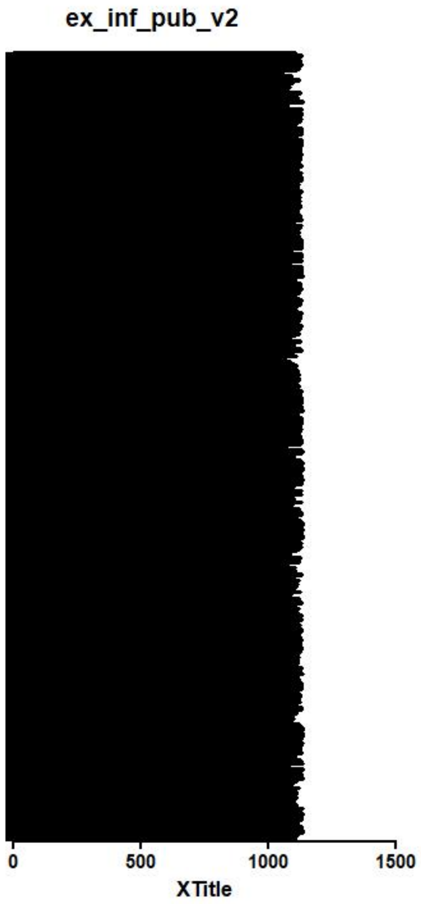
Exact match alleles
+ inferred alleles
Dataset 1 (CC)



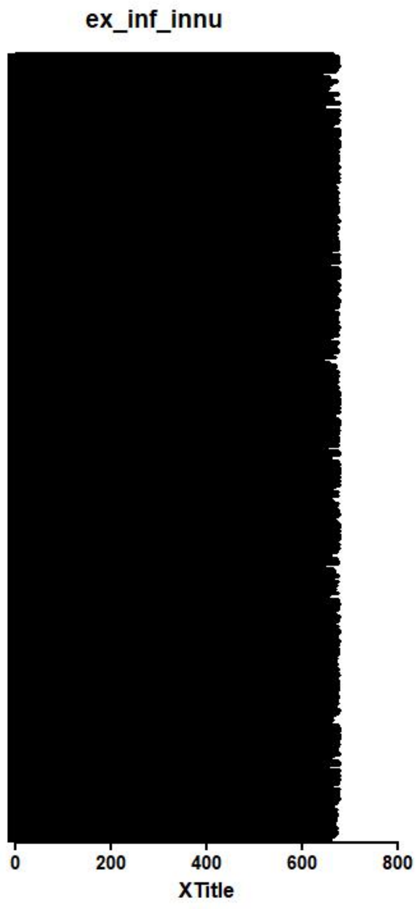
pubMLST_v1
1343 loci



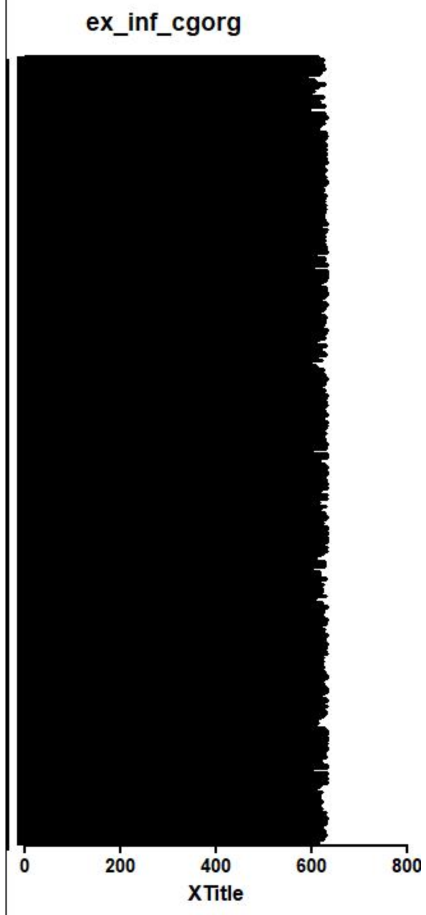
pubMLST_v2
1142 loci



Innuendo cgMLST
678 loci



Ridom cgMLST
637 loci



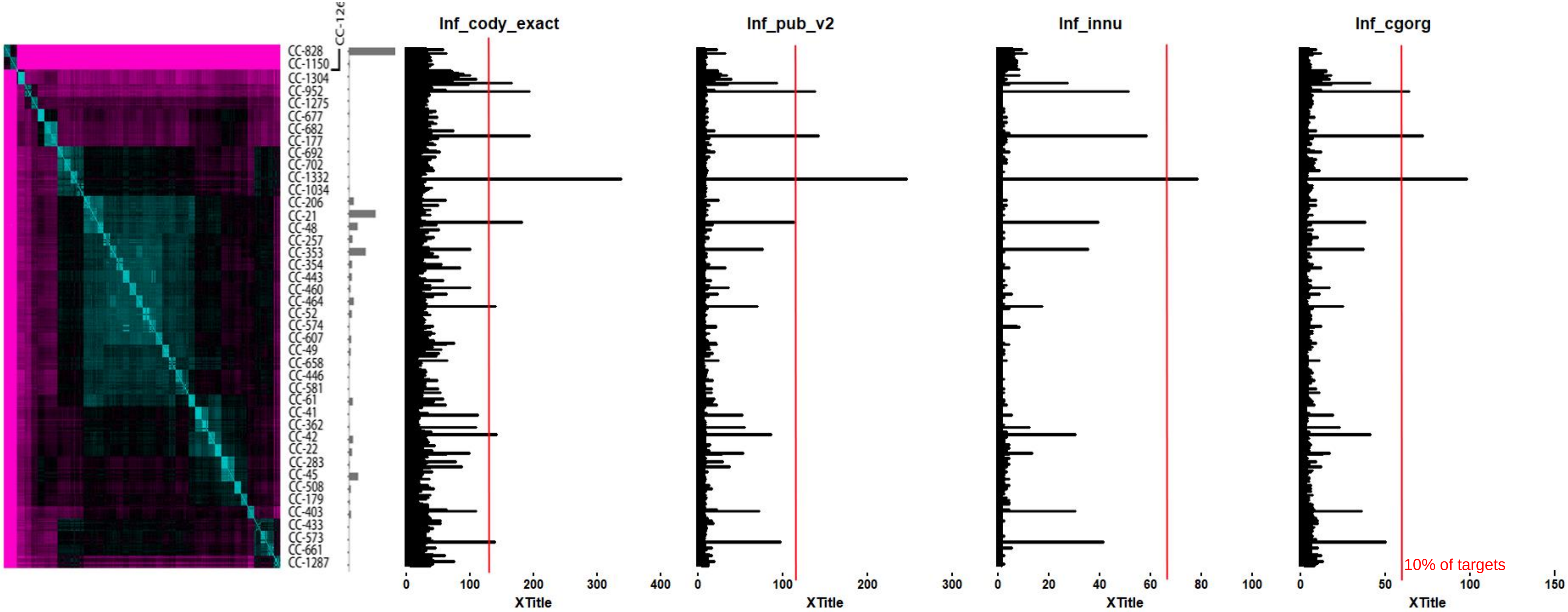
Locus Not Found
Dataset 1 (CC)

pubMLST_v1
1343 loci

pubMLST_v2
1142 loci

Innuendo cgMLST
678 loci

Ridom cgMLST
637 loci



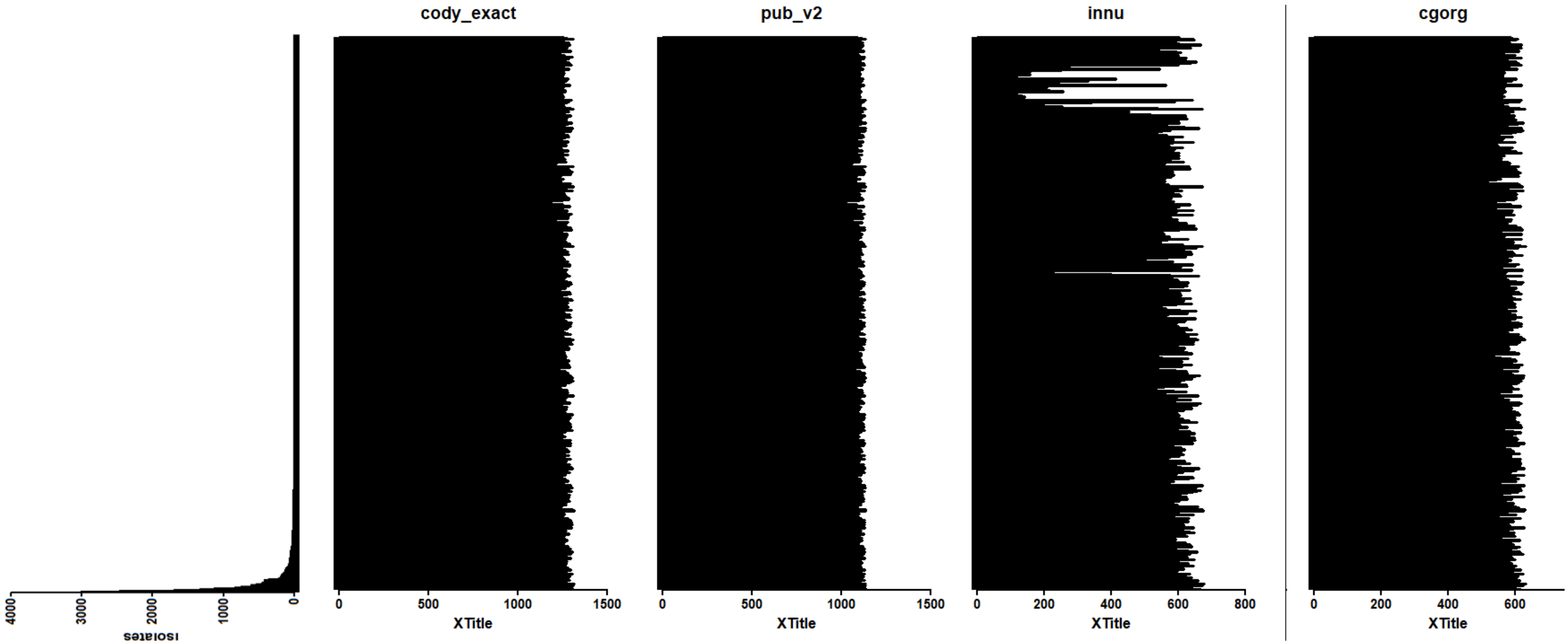
Exact match alleles
Dataset 2 (*C. coli* ST)

pubMLST_V1
”Oxford scheme”
”Cody scheme”
1343 loci

pubMLST_V2
1142 loci

Innuendo cgMLST
678 loci

cgMLST.org
637 loci



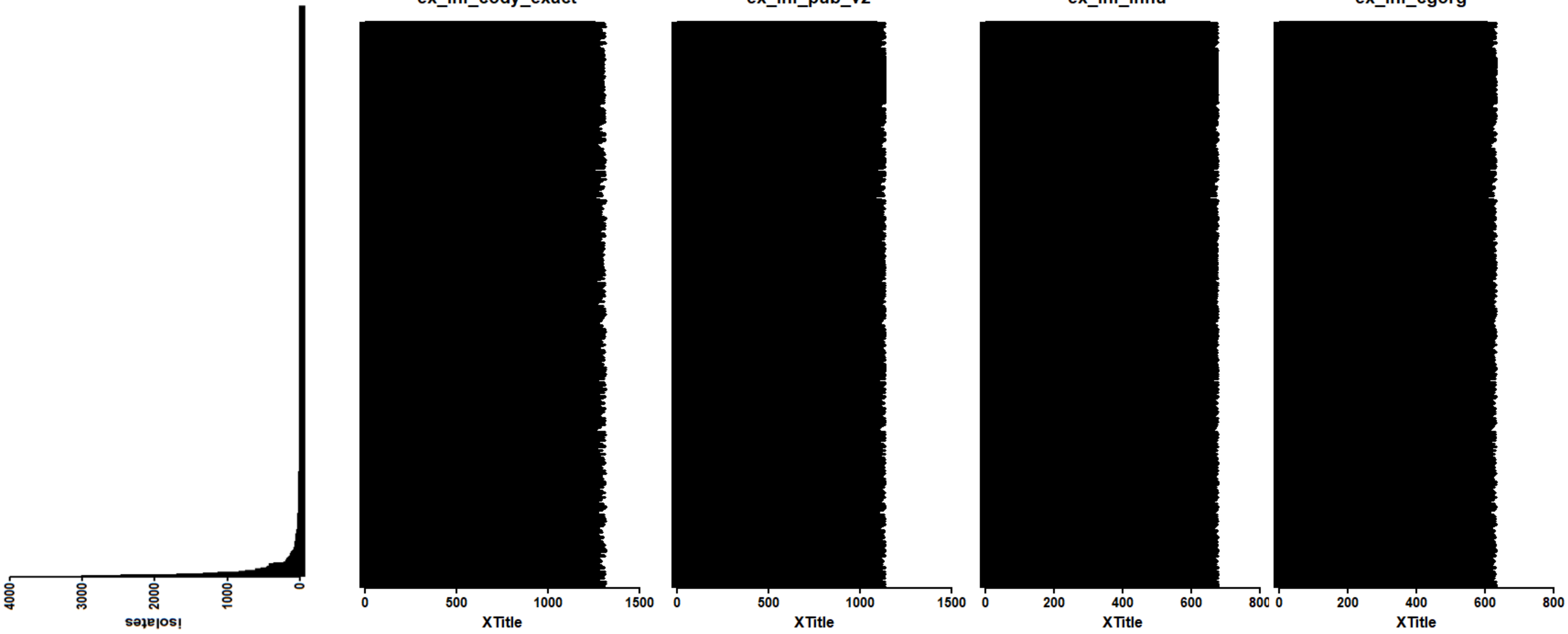
Exact match alleles + inferred alleles
Dataset 2 (*C. coli* ST)

pubMLST_v1
1343 loci

pubMLST_v2
1142 loci

Innuendo cgMLST
678 loci

Ridom cgMLST
637 loci



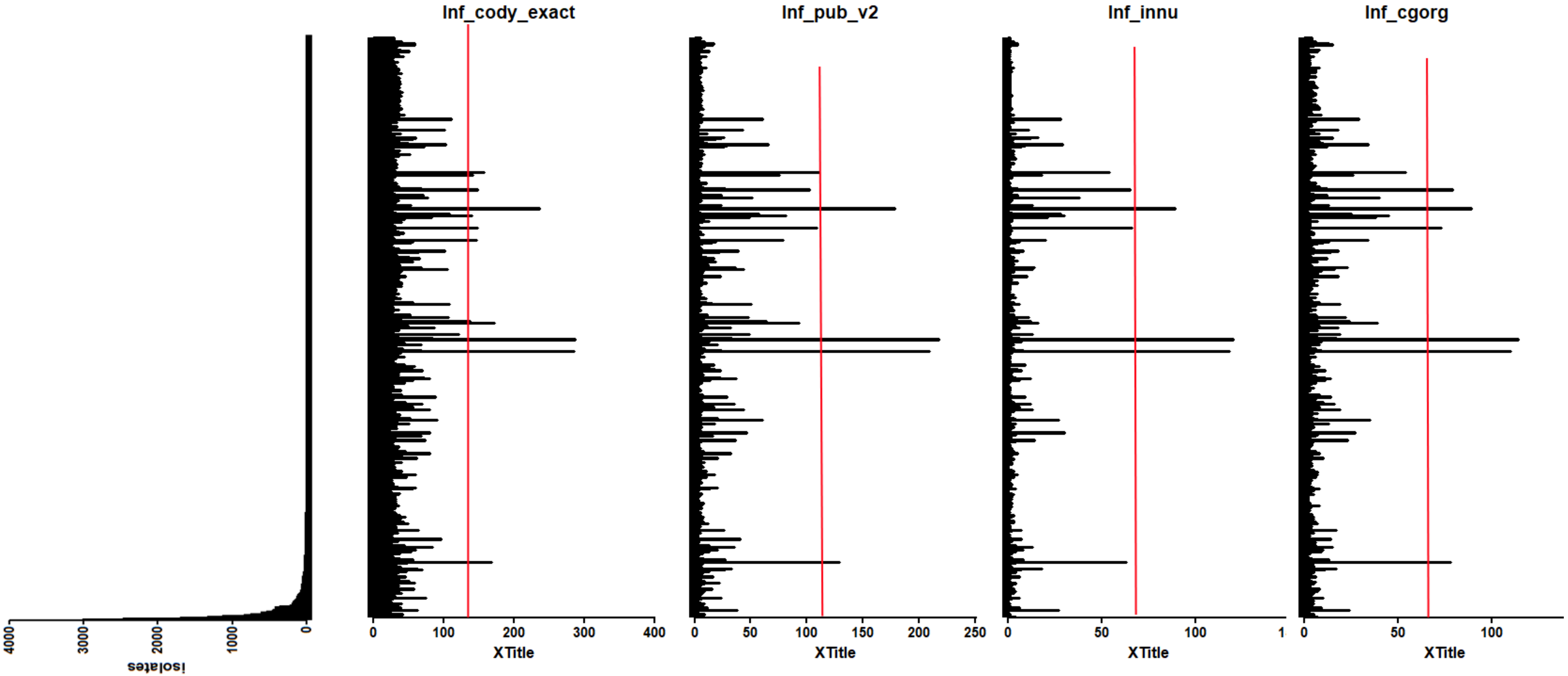
Locus Not Found
Dataset 2 (*C. coli* ST)

pubMLST_v1
1343 loci

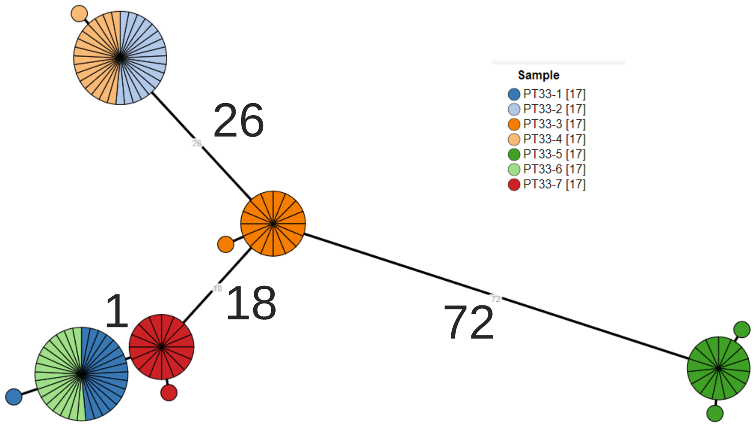
pubMLST_v2
1142 loci

Innuendo cgMLST
678 loci

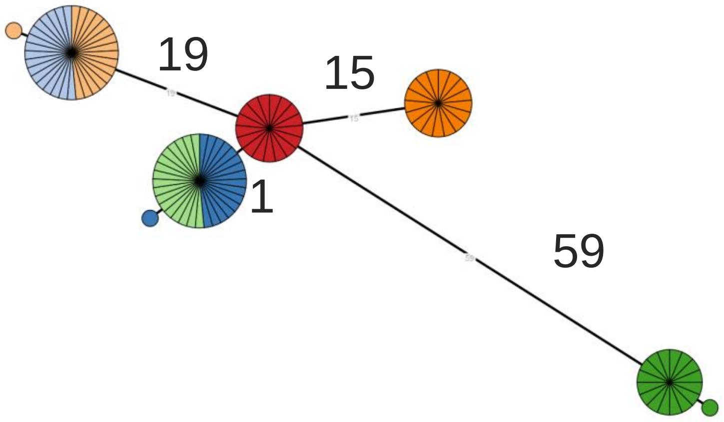
Ridom cgMLST
637 loci



PubMLST_v1



PubMLST_v2



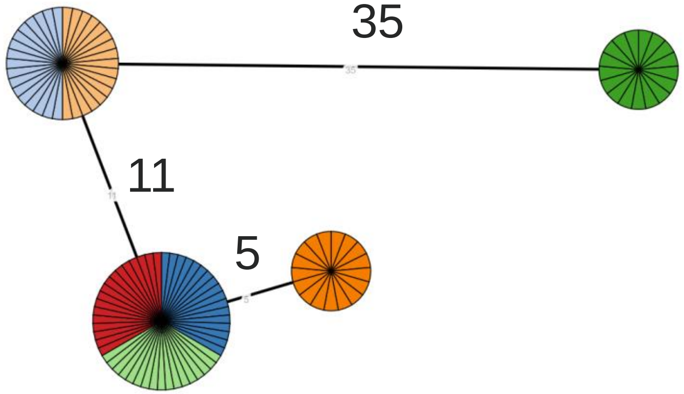
Sample

● PT33-1 [17] *
● PT33-2 [17]
● PT33-3 [17]
● PT33-4 [17]
● PT33-5 [17]
● PT33-6 [17] *
● PT33-7 [17]

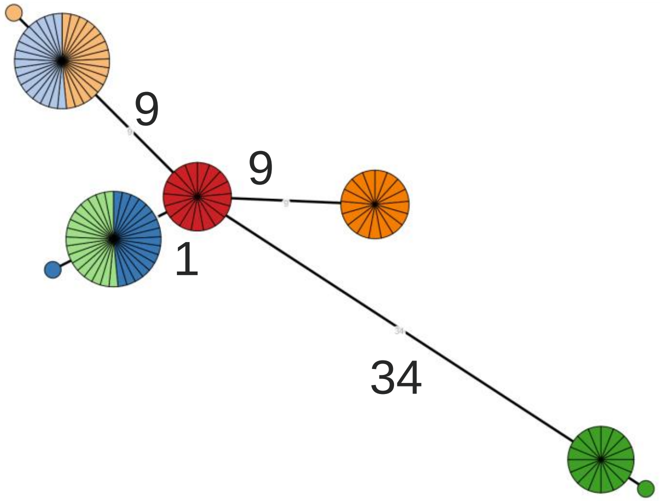
*= same DNA

Sample	Strain	Matrix	Location	Sampling
PT33-1, PT33-6	20C120	Chicken caeca	Sweden, farm A	October, 2020
PT33-2	20C028	Chicken caeca	Sweden, farm B	July, 2020
PT33-3	20C126	Chicken caeca	Sweden, farm C	July, 2020
PT33-4	20C060	Chicken caeca	Sweden, farm A	July, 2020
PT33-5	Val_Cj015	Milk filter	Sweden, farm D	2011
PT33-7	20C102	Chicken caeca	Sweden, farm E	October, 2020

Innuendo cgMLST



Ridom Seqsphere cgMLST



Assemblies provided by PT-users

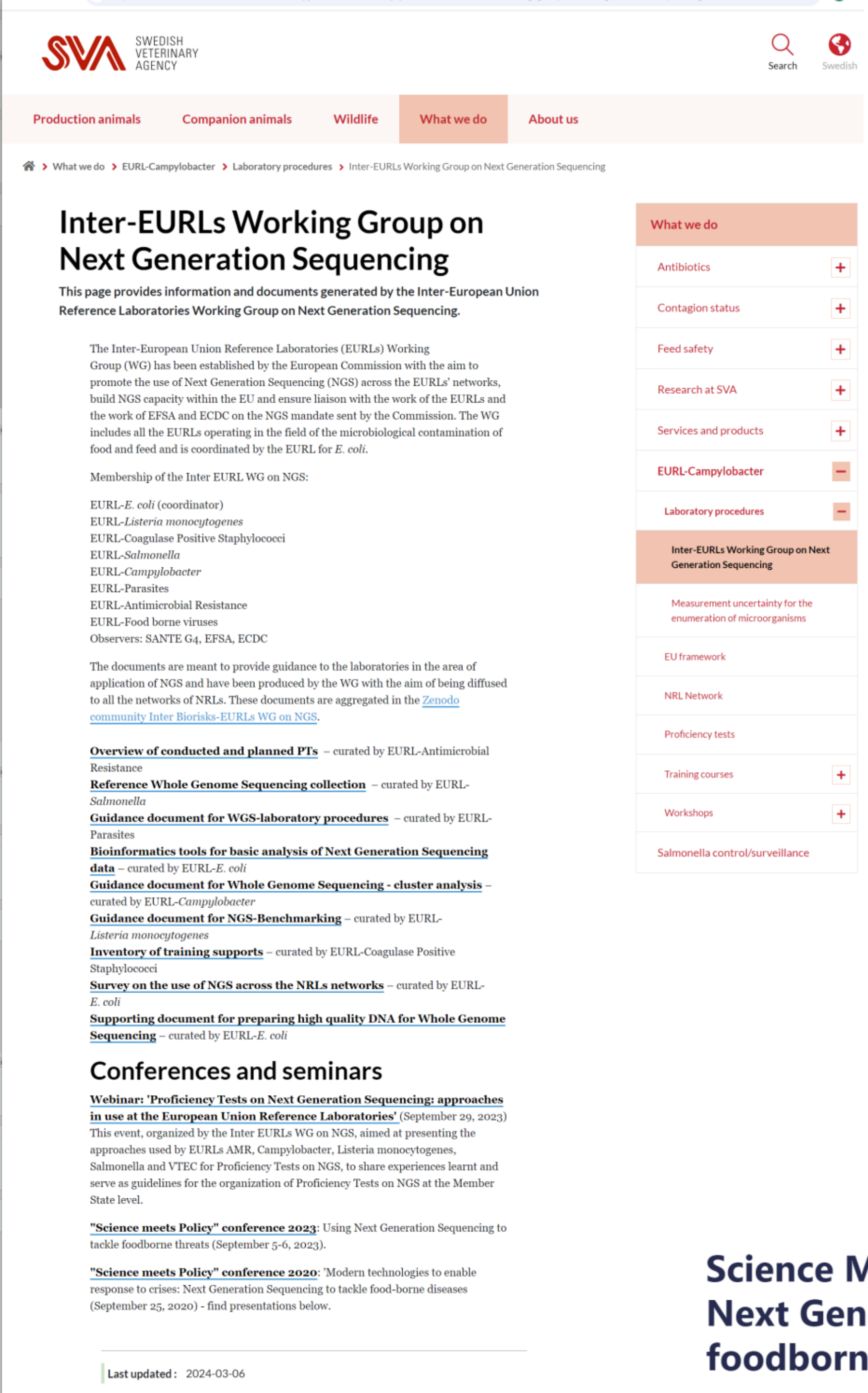
Conclusion

- The schemes have a high overlap in targets and differ mainly in size
- The PubMLST schemes and Ridom scheme are well populated with alleles
- All schemes showed an overall good functionality for both *C. jejuni* and *C. coli* – one scheme can be used for both species
- pubMLST 'Oxford' scheme v2 has a size covering the whole core part of the *C. jejuni* and *C. coli* pan-genome

Co-funded by the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor HaDEA can be held responsible for them.

Guidance documents



Inter-EURLs Working Group on Next Generation Sequencing

This page provides information and documents generated by the Inter-European Union Reference Laboratories Working Group on Next Generation Sequencing.

The Inter-European Union Reference Laboratories (EURLs) Working Group (WG) has been established by the European Commission with the aim to promote the use of Next Generation Sequencing (NGS) across the EURLs' networks, build NGS capacity within the EU and ensure liaison with the work of the EURLs and the work of EFSA and ECDC on the NGS mandate sent by the Commission. The WG includes all the EURLs operating in the field of the microbiological contamination of food and feed and is coordinated by the EURL for *E. coli*.

Membership of the Inter EURL WG on NGS:

EURL-*E. coli* (coordinator)
EURL-*Listeria monocytogenes*
EURL-Coagulase Positive Staphylococci
EURL-*Salmonella*
EURL-*Campylobacter*
EURL-Parasites
EURL-Antimicrobial Resistance
EURL-Food borne viruses
Observers: SANTE G4, EFSA, ECDC

The documents are meant to provide guidance to the laboratories in the area of application of NGS and have been produced by the WG with the aim of being diffused to all the networks of NRLs. These documents are aggregated in the Zenodo community Inter Biorisks-EURLs WG on NGS.

Overview of conducted and planned PTs – curated by EURL-Antimicrobial Resistance
Reference Whole Genome Sequencing collection – curated by EURL-*Salmonella*
Guidance document for WGS-laboratory procedures – curated by EURL-Parasites
Bioinformatics tools for basic analysis of Next Generation Sequencing data – curated by EURL-*E. coli*
Guidance document for Whole Genome Sequencing - cluster analysis – curated by EURL-*Campylobacter*
Guidance document for NGS-Benchmarking – curated by EURL-*Listeria monocytogenes*
Inventory of training supports – curated by EURL-Coagulase Positive Staphylococci
Survey on the use of NGS across the NRLs networks – curated by EURL-*E. coli*
Supporting document for preparing high quality DNA for Whole Genome Sequencing – curated by EURL-*E. coli*

Conferences and seminars

Webinar: 'Proficiency Tests on Next Generation Sequencing: approaches in use at the European Union Reference Laboratories' (September 29, 2023)
This event, organized by the Inter EURLs WG on NGS, aimed at presenting the approaches used by EURLs AMR, Campylobacter, Listeria monocytogenes, Salmonella and VTEC for Proficiency Tests on NGS, to share experiences learnt and serve as guidelines for the organization of Proficiency Tests on NGS at the Member State level.

"Science meets Policy" conference 2023: Using Next Generation Sequencing to tackle foodborne threats (September 5-6, 2023).

"Science meets Policy" conference 2020: 'Modern technologies to enable response to crises: Next Generation Sequencing to tackle food-borne diseases (September 25, 2020) - find presentations below.

Last updated : 2024-03-06

Overview of conducted and planned PTs – curated by EURL-Antimicrobial Resistance
Reference Whole Genome Sequencing collection – curated by EURL-*Salmonella* * Update coming
Guidance document for WGS-laboratory procedures – curated by EURL-Parasites
Bioinformatics tools for basic analysis of Next Generation Sequencing data – curated by EURL-*E. coli*
Guidance document for Whole Genome Sequencing - cluster analysis – curated by EURL-*Campylobacter* * Update coming
Guidance document for NGS-Benchmarking – curated by EURL-*Listeria monocytogenes*
Inventory of training supports – curated by EURL-Coagulase Positive Staphylococci * Update coming
Survey on the use of NGS across the NRLs networks – curated by EURL-*E. coli*
Supporting document for preparing high quality DNA for Whole Genome Sequencing – curated by EURL-*E. coli*

Future document:
Quality Control of NGS data

Joint NGS training courses

New edition of joint basic training course 2025

joint exercises

Conferences/seminars/webinars

Science Meets Policy conference: Using Next Generation Sequencing to tackle foodborne threats

Webinar: 'Proficiency Tests on Next Generation Sequencing: approaches in use at the European Union Reference Laboratories' (September 29, 2023)

SWEDISH
VETERINARY
AGENCY