



GenEpi-BioTrain

WGS systems for One Health

Priyanka Nannapaneni, Scientific officer Bioinformatics, ECDC

04/06/2024

Intended Learning Objectives

Specific objectives of this session:

1. Understand the principles and applications of WGS technology in the context of One Health
2. Evaluate the potential of WGS for enhancing surveillance and early detection of foodborne and waterborne pathogens
3. Show the collaboration strategies between EFSA and ECDC to harmonize WGS methodologies and data sharing practices

Outline

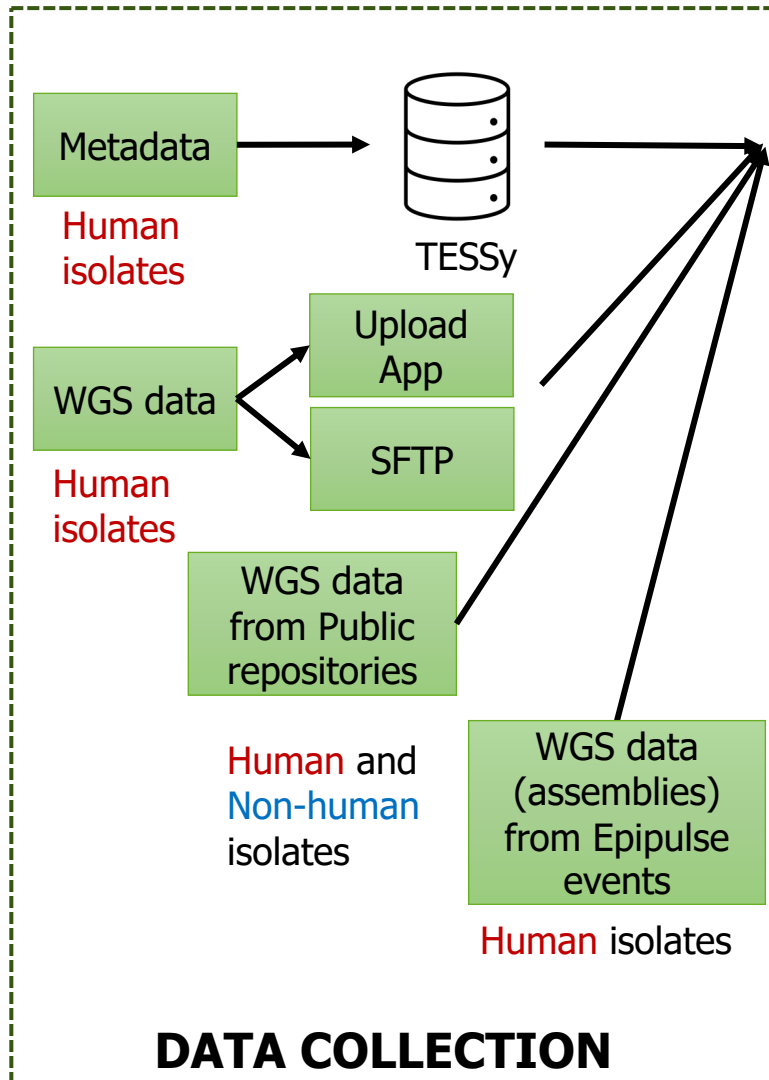
This session consists of the following elements

- EU/EEA molecular surveillance objectives
- Data flow at ECDC cloud system
- Species specific cluster criteria for automatic cluster detection
- Dissemination of cluster results
- Overview of clusters for Listeria and Salmonella
- Integration of food isolates from EFSA database
- Data Visibility rules for PH users
- Future developments

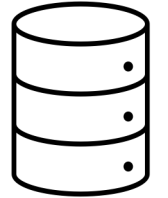
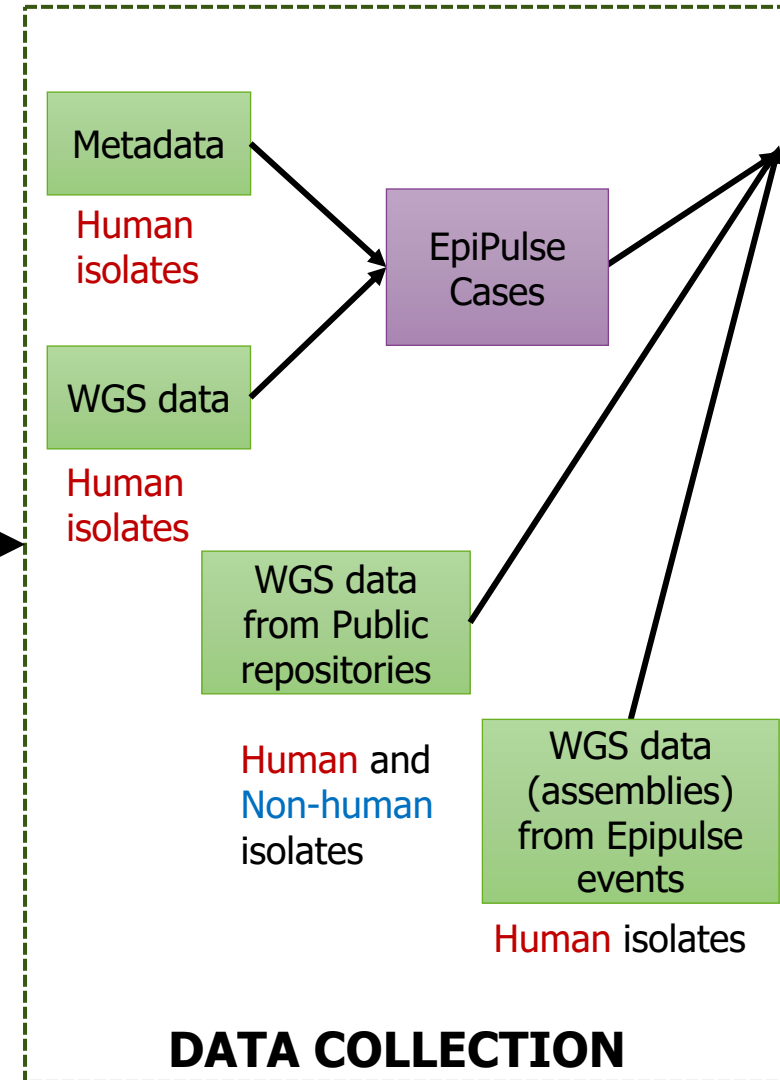
EU/EEA molecular surveillance objectives

- Detection of multi-country outbreaks and/or geographically dispersed clusters to trigger outbreak investigations and contribute to trace back and forward investigations so that appropriate control and preventive measures can be implemented
- Identification of persistent strains causing human infections in EU/EEA
- Identification of transmission chains, new risk factors for infection and severity of disease
- Monitoring of EU/EEA trends by selected indicators
- Support to national investigations

ECDC Data Flow – Upstream



From July 2024,
moving to EpiPulse
cases, for some
pathogens (including
Listeria)

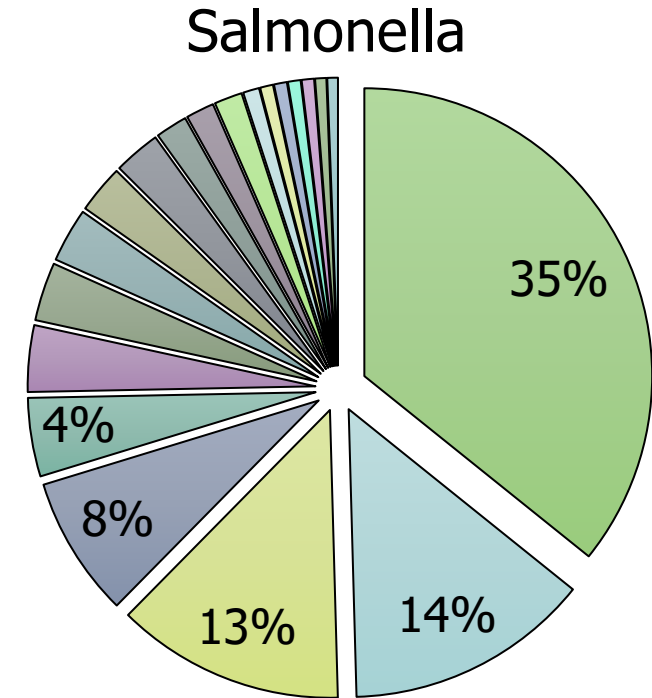


BioNumerics
interface

Azure Cloud

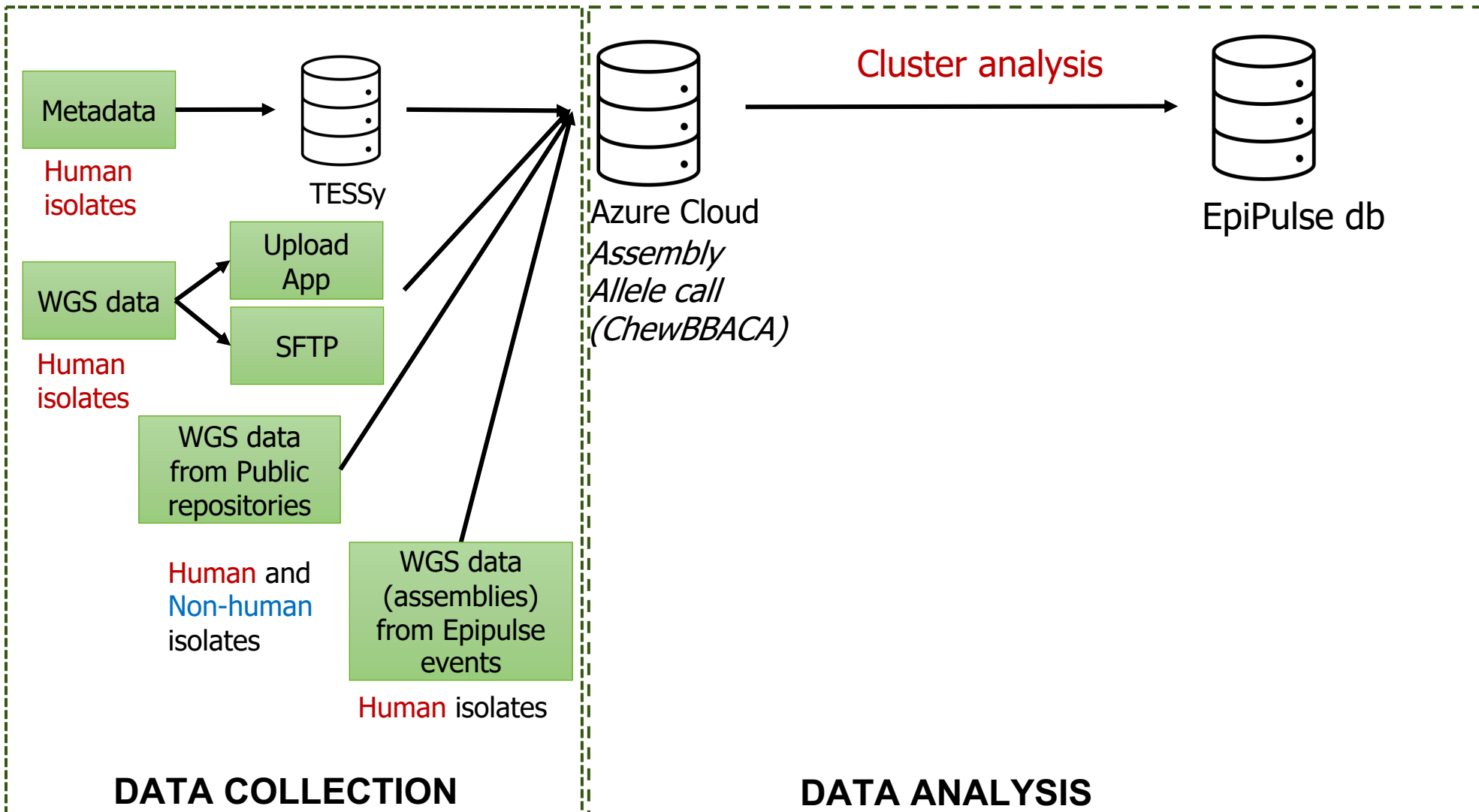
*Assembly and Allele
call (ChewBBACA)*

- In 2019, ECDC opened EU-EEA wide WGS-enhanced surveillance
- >6000 isolates in Listeria database, 32 countries
- >6000 isolates in Salmonella database, 30 countries
- Uneven distribution of data from member states



6

ECDC CLUSTER PROCESS FLOWCHART



Cluster detection criteria

Cluster detection

**Species specific cluster criteria for automatic cluster detection*

- Core cluster: Single-linkage cluster of more than two sequences within core allele difference (AD) threshold agreed per pathogen (only Human isolates define the core clusters)
- Extended cluster: Slightly broader cut-off than core cluster definition agreed per pathogen that usually includes the Food isolates
- Clusters include both single country clusters as well as multi-country clusters
- Cluster Status: A cluster status is OPEN as long as at least one of the isolates within the cluster has been reported to be collected within the past 12 months and it is **multi-country cluster**, otherwise CLOSED

Species specific cluster criteria for automatic cluster detection



Species specific cluster criteria for automatic cluster detection

- *L. monocytogenes*: core cluster within 4 AD, extended cluster 7 AD
- *Salmonella* spp.: core cluster within 5 AD, extended cluster 10 AD. For certain serotypes/subtypes, for example *S. Enteritidis*, a tighter core cluster could be needed for outbreak investigations. In addition, the single linkage clustering approach possibly needs to be replaced with a Neighbour joining clustering approach for endemic clones.
- STEC: core cluster within 5 AD, extended cluster 10 AD. For certain subtypes a tighter core cluster cut-off could be needed for outbreak investigations. *Campylobacter* spp.: core cluster within 5 AD, extended cluster 10 AD.
- *Shigella* spp./EIEC: core cluster within 5 AD, extended cluster 10 AD. For certain species/subtypes a tighter core cluster cut-off could be needed for outbreak investigations.

N.B Specific cluster cut-offs could be applied for multi country outbreak case definitions.

Dissemination of outputs and results in EpiPulse

Dissemination of outputs and results – EpiPulse Signals



- **Signals** in EpiPulse “Events Forum and News” are created for only multi-country clusters that are circulating in past 12 months (OPEN) within the ECDC weekly clustering
- Signals are only visible to involved countries
- Signals indicate clusters not assessed as posing a public health risk for the EU/EEA yet.
- Signals can be escalated to Events which allows all EpiPulse users to see it

Dissemination of outputs and results – EpiPulse Signals



← → ↺ 🏠 epipulse.ecdc.europa.eu/ebs/#/

ECDC Links | EFSA ONE HEALTH | Home Page - WGS | EpiPulse MTT | Manual installation... | Overview | Jupyter Notebook V... | Web Transfer Client | Issues · cov-lineage... | Cov-Lineages | plotPCA — deepTo... | GISAID - Initiative | SAVE WG.pptx | Docu...

ecdc EpiPulse Report Manage Explore Collaborate (0) 9

🏠

All
ARHAI
EI
LEGI
EVD
FWD
HEP
HIV
IRV
PREP
SoHO
SRV
STI
TB
VPD
Resp. duty
TALD
Daily CDTR
Weekly CDTR

Search by text in Please select Q Announcement Event Forum Long-Term Monitoring News Signal Threat Closed Discarded Open

⊕ Advanced search criteria

Create item
Create output
Edit flags
🔔
Column visibility
Show 25 rows
📄
📄

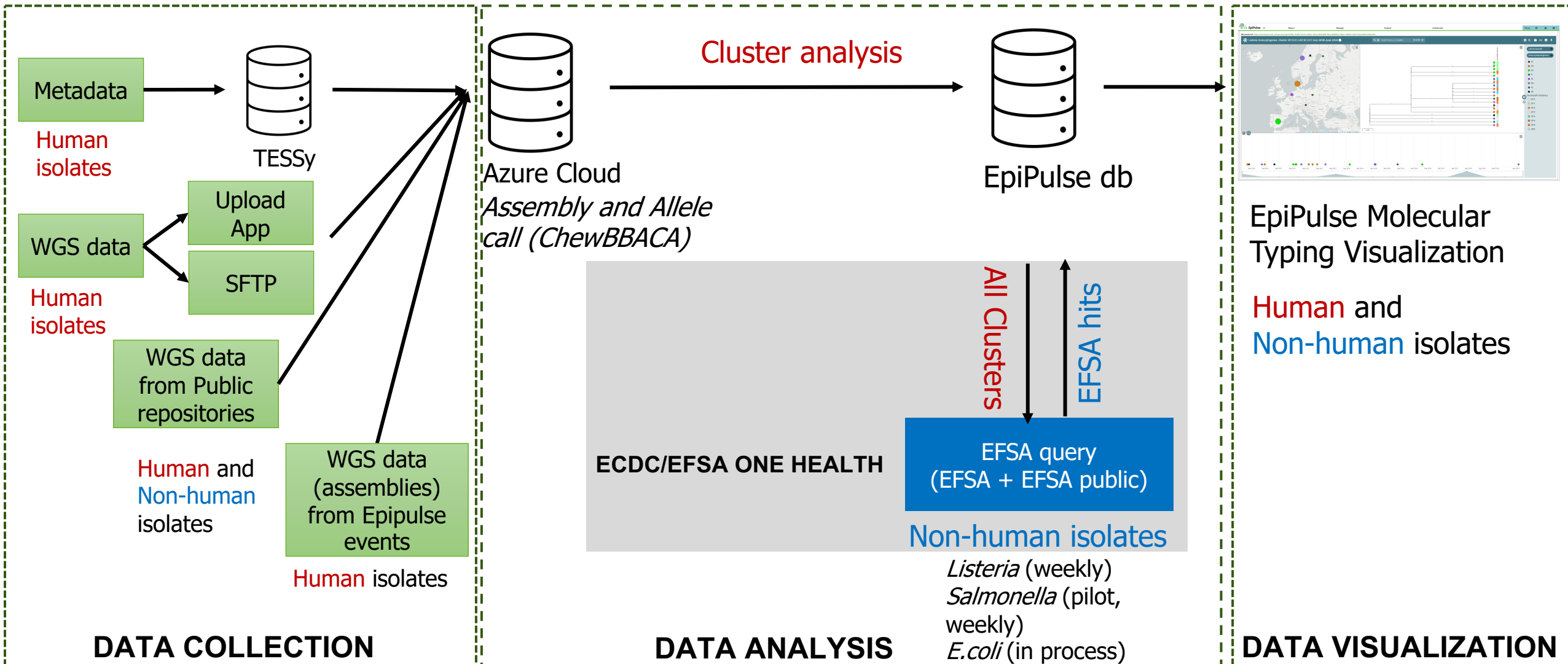
■	ID	Participating domain	Type	Title	Created by	Pathogens	Diseases	Modified time	Flags
	🔍	🔍 FWD	🔍 Signal	🔍	🔍	🔍	🔍	🔍	🔍
<input type="checkbox"/>	2024-FWD-00016	FWD	Signal	Molecular typing cluster 2024-03.LIST.01	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2023-FWD-00027	FWD	Signal	Molecular typing cluster 2017-10.LIST.88.CC412.Asci.0148.ApaI.0482	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2023-FWD-00095	FWD	Signal	Molecular typing cluster 2023-10.LIST.07.CC14	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2024-FWD-00005	FWD	Signal	Molecular typing cluster 2024-01.LIST.02.CC9	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2024-FWD-00029	FWD	Signal	Molecular typing cluster 2024-04.LIST.02	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2022-FWD-00011	FWD	Signal	Molecular typing cluster 2018-06.LIST.15.CC6	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:02	
<input type="checkbox"/>	2023-FWD-00020	FWD	Signal	Molecular typing cluster 2019-06.LIST.03.CC1	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2023-FWD-00038	FWD	Signal	Molecular typing cluster 2021-07.LIST.61.CC18	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2023-FWD-00102	FWD	Signal	Molecular typing cluster 2023-12.LIST.03	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2019-FWD-00109	FWD	Signal	Molecular typing cluster 2018-05.LIST.07.CC155	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2019-FWD-00130	FWD	Signal	Molecular typing cluster 2017-11.LIST.13.CC388	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2024-FWD-00004	FWD	Signal	Molecular typing cluster 2017-06.LIST.13.CC87	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	
<input type="checkbox"/>	2019-FWD-00131	FWD	Signal	Molecular typing cluster 2019-11.LIST.08.CC1	ECDC/Public Health	Listeria monocytogenes	Listeriosis	2024-05-07 16:01	

Dissemination of results in EpiPulse Molecular Typing Tool



Integration of One Health system with EFSA

Integration of One Health system



Demo of ECDC-EFSA interaction

Create query Summary of query activity Incoming queries (EFSA->ECDC) Outgoing queries (ECDC->EFSA)

ECDC-EFSA cgMLST-based ad-hoc query tool v1.0

Matching parameters:

Subject:

Type of search:

Cluster/Event to match:

Match distance (max 50):

Perform matching in ECDC database: ☒

Perform matching in EFSA database: ☒

Before you start the matching, make sure that isolates are annotated with the correct cluster code/event code and that the isolates are visible in [EpiPulse Molecular typing](#). If there are still isolates you expect to show up in the query that do not, there may be some delay before they have been analysed with ChewBBACA which is required for the query. If the delay seems unusually long, contact the Bioinfo team for assistance. Note that only human isolates are used in the query (PB_SampleSource empty or equal to HUMAN)

Status:

Matching not yet started

Demo of ECDC-EFSA interaction

Query filtered on field: BN_Cluster=2023-08.LIST.60.

Number of query isolates: 17

Query performed: 2024-06-03 15:25:51

Number of EFSA hits: 8

Number of ECDC hits (not including query isolates): 0

[MicroReact link](#)

ECDC isolates used in the query

Column visibility

Copy

CSV

Search:

ECDC RecordId	Country	DateUsedForStatistics	DateReceivedReferencelabOrig	Sample source
	DK		2019	HUMAN
	DK		2019	HUMAN
	DK		2020	HUMAN
	DK		2021	HUMAN
	DE		2022	HUMAN
	DK		2022	HUMAN
	DK		2022	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2023	HUMAN
	DK		2024	HUMAN
	DK		2024	HUMAN
	IT	2023-03-06	2023-03-06	HUMAN

Showing 1 to 17 of 17 entries

Demo of ECDC-EFSA interaction

EFSA hits

Column visibility

Copy

CSV

Search:

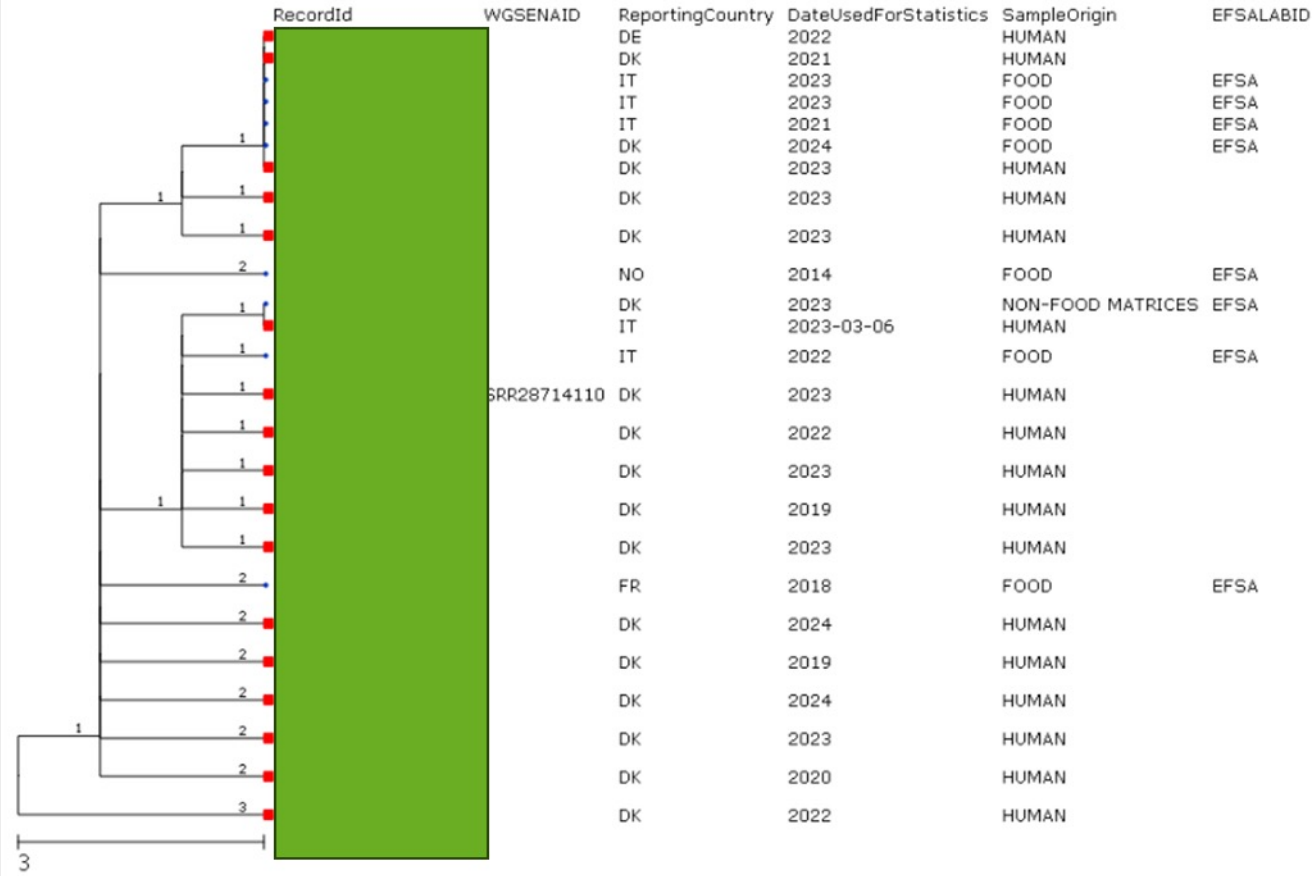
EFSA id	Country	Year	Sample matrix	Sample specifics	ENA accession	Closest distance to query	Closest query isolate(s)
	NO	2014	FOOD	Fish meat and products thereof		2	
	FR	2018	FOOD	Fish meat and products thereof		4	
	IT	2023	FOOD	Fish meat and products thereof		0	
	IT	2022	FOOD	Fish meat and products thereof		1	
	IT	2023	FOOD	Fish meat and products thereof		0	
	IT	2021	FOOD	Fish meat and products thereof		0	
	DK	2023	NON-FOOD MATRICES	Non-food matrices		0	
	DK	2024	FOOD	Fish meat and products thereof		0	

Showing 1 to 8 of 8 entries

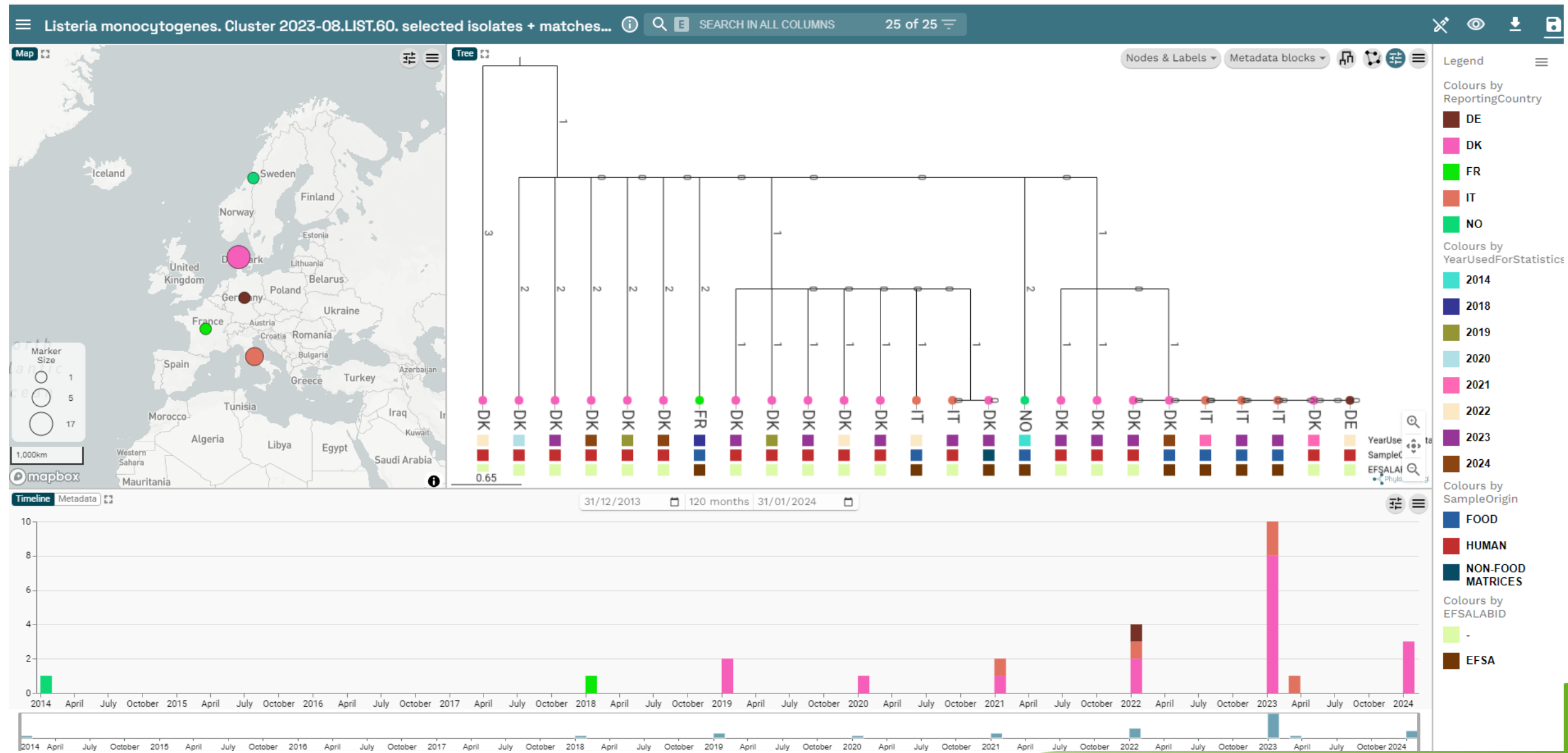
Demo of ECDC-EFSA interaction

Tree of all matching isolates (note that this includes all EFSA isolates previously matched even if no EFSA matching was performed). Query isolates are marked in red.

Listeria monocytogenes. Cluster 2023-08.LIST.60. selected isolates + matches within 7 (Generated 2024-06-03 15:27:28)



Demo of ECDC-EFSA interaction



Cluster statistics

Overview of TESSy multi-country clusters, Listeria

	Total
Total number of clusters	684
Number of single-country clusters detected	521
Number of multi-country clusters detected	163
OPEN clusters	26
Number of single-country clusters with non-human isolates (7 AD)	66
Number of multi-country clusters with non-human isolates (7 AD)	65



	Core cluster (within 4 cg-AD)
Median number of countries involved in clusters (range)	2 (2-10)
Median number of isolates (range)	4 (2-98)
Median duration in years (range)	2.8 (0-15.5)*

*calculated for clusters that have full start and end dates available

As of 15 February 2024

Overview of TESSy multi-country clusters, Salmonella

Total	
Total number of clusters	356
Number of single-country clusters detected	186
Number of multi-country clusters detected	170
OPEN clusters	16
Number of single-country clusters with non-human isolates (10 AD)	19
Number of multi-country clusters with non-human isolates (10 AD)	42



	Core cluster (within 5 cg-AD)
Median number of countries involved in clusters (range)	2 (2-25)
Median number of isolates (range)	7 (2-1373)
Median duration in years (range)	0.5 (0-22.9)*

*calculated for clusters that have full start and end dates available

As of 15 February 2024

Data visibility criteria

ECDC Data visibility criteria for PH users

- National identifiers (RecordId) can only be seen by ECDC and the submitting country
- Clusters and Signals can only be seen by involved countries (involved countries are those with human cases in the cluster), unless upgraded to an Event or Threat (it is however possible for countries to see other nearby clusters if they use the more exploratory features in EpiPulse)
- For EFSA data, country of origin can only be seen by ECDC and the same country, and further restrictions on data visibility can be applied by the EFSA providers
- For EFSA Public data, no visibility restrictions on country of origin

Outlook and Future developments

- Successful cross-sectoral collaboration at EU level
- Extremely valuable and aids in the timely assessment of outbreak investigations
- More pathogens in the pipeline for ECDC-EFSA interaction
 - SALM is in Pilot phase currently
 - *E.coli*, *Campy* are next in queue

Acknowledgements

The creation of this training material was commissioned by ECDC to Institut Pasteur with the direct involvement of Priyanka Nannapaneni