Implementing WGS data collection and analysis for communicable diseases at EU/EEA level

Erik Alm, Senior Expert Applied Molecular Epidemiology
Presented by: Saara Kotila, Scientific Officer Molecular Surveillance

Campylobacter activities at ECDC

- Annual case-based data collection and reporting at the EU/EEA-level
- Annual AMR isolate-based data collection and reporting at the EU/EEA-level
- Food- and waterborne diseases network (FWD-Net) meeting 2020, emphasis on *Campylobacter*
- Joint EURL-AR and FWD-Net workshop on *Salmonella* and *Campylobacter* AMR, also 2020
- WGS support for possible multi-country events through laboratory contractor
- WGS analysis tools for outbreak investigations

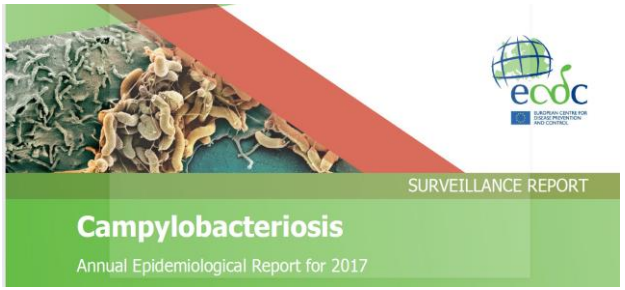
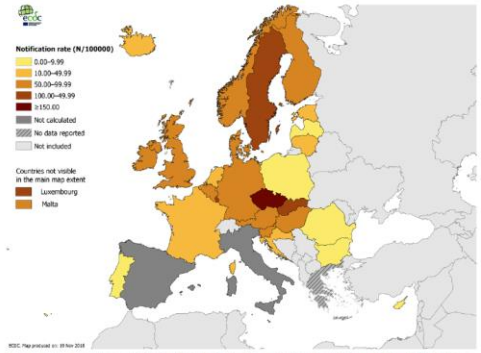


Table 1. Distribution of confirmed campylobacteriosis cases and rates per 100 000 population by year and country, EU/EEA, 2013–2017

Country	2013		2014		2015		2016		2017			
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Confirmed cases	Rate	ASR	Reported cases
Austria	5 731	67.8	6 514	76.6	6 258	72.9	7 003	81.4	7 204	82.1	84.0	7 204
Belgium	8148	-	8 086	-	9 066	80.7	10 055	88.9	8 649	76.2	75.4	8 649
Bulgaria	124	1.7	144	2.0	227	3.2	202	2.8	195	2.7	3.0	195
Croatia	0	0.0	1 647	38.8	1 393	33.0	1 524	36.4	1 686	40.6	43.3	1 694
Cyprus	56	6.5	40	4.7	29	3.4	21	2.5	20	2.3	2.3	20
Czech Republic	18 267	173.7	20 750	197.4	20 960	198.9	24 094	228.2	24 326	230.0	240.9	24 508
Denmark	3 772	67.3	3 773	67.0	4 327	76.5	4 712	82.6	4 255	74.0	74.8	4 255
Estonia	382	28.9	285	21.7	318	24.2	298	22.6	285	21.7	22.5	347
Finland	4 066	74.9	4 889	89.7	4 586	83.8	4 637	84.5	4 289	77.9	81.4	4 289
France	5 198	39.6	5 958	45.2	6 074	45.7	6 686	50.2	6 579	49.1	48.6	6 579
Germany	63 280	78.6	70 571	87.4	69 829	86.0	73 663	89.6	69 178	83.8	84.4	69 414
Greece												
Hungary	7 247	73.1	8 444	85.5	8 342	84.6	8 656	87.0	7 807	79.7	84.7	7 840
Iceland	101	31.4	142	43.6	119	36.2	128	38.5	119	35.2	36.0	119
Ireland	2 288	49.6	2 593	55.9	2 453	52.4	2 511	53.1	2 779	58.1	57.5	2 788
Italy	1 178	-	1 252	-	1 014	-	1 057	-	1 060	-	-	1 060
Latvia	9	0.4	37	1.8	74	3.7	90	4.6	59	3.0	3.0	61
Liechtenstein		-		-		-		-		-	-	
Lithuania	1 139	38.3	1 184	40.2	1 186	40.6	1 225	42.4	990	34.8	35.6	993
Luxembourg	675	125.7	873	158.8	254	45.1	516	89.9	613	103.8	104.1	613
Malta	246	58.2	288	67.1	248	56.4	212	47.1	231	50.2	51.7	231
Netherlands	3 702	42.4	4 159	47.5	3 778	43.0	3 383	38.3	2 890	32.5	31.0	2 890
Norway	3 291	65.2	3 386	66.3	2 318	44.9	2 317	44.5	3 884	73.9	74.4	3 884
Poland	552	1.5	650	1.7	653	1.7	773	2.0	874	2.3	2.4	874
Portugal		-		-	271	2.6	359	3.5	596	5.8	6.9	602
Romania	218	1.1	256	1.3	311	1.6	517	2.6	467	2.4	2.5	479
Slovakia	5 945	108.0	6 744	124.5	6 949	128.2	7 623	140.5	6 946	127.8	130.4	7 057
Slovenia	1 027	49.9	1 184	57.4	1 328	64.4	1 642	79.5	1 408	68.2	71.4	1 408
Spain	7 064	-	11 481	-	13 227	-	15 542	-	18 860	-	-	18 860
Sweden	8 114	84.9	8 288	85.9	9 180	94.2	11 021	111.9	10 608	106.1	107.7	10 608
United Kingdom	66 382	103.9	66 716	103.7	59 797	92.2	58 911	90.1	63 304	96.2	95.0	63 304
EU/EEA	219 102	69.0	249 346	96.3	234 571	92.7	249 362	66.0	250 101	64.9	65.5	250 026

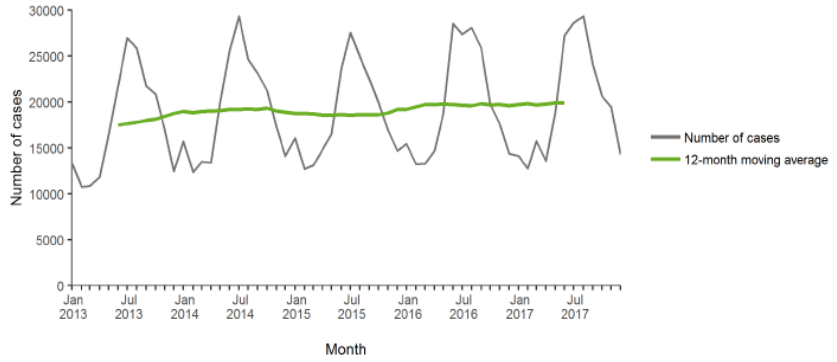
Source: country reports.
ASR: age-standardised rate
.: no data reported
-: no rate calculated.

Figure 1. Distribution of confirmed campylobacteriosis cases per 100 000 population by country, EU/EEA, 2017



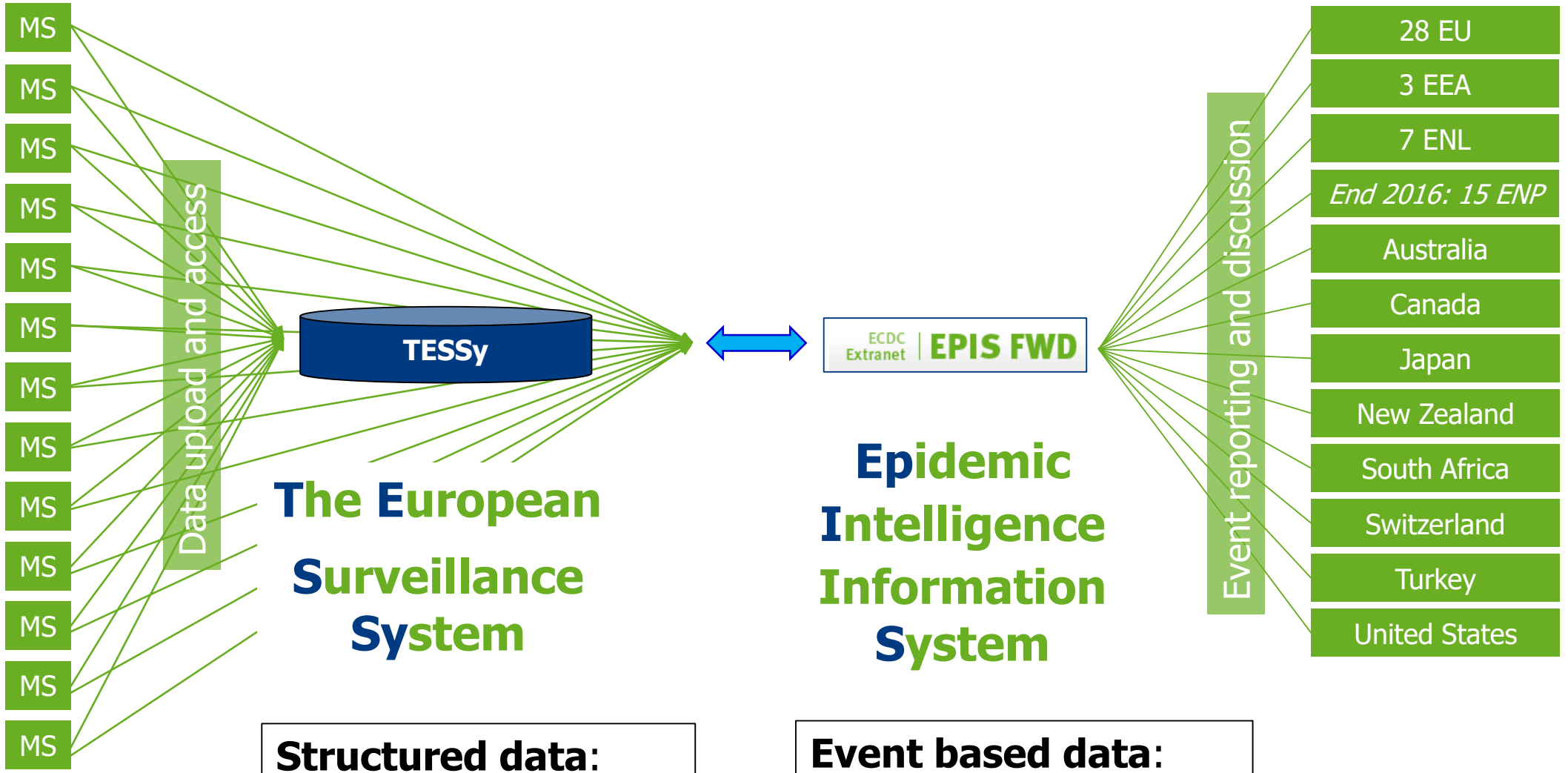
Source: country reports from Austria, Belgium, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden and the United Kingdom. No rates were calculated for Italy and Spain.

Figure 2. Distribution of confirmed campylobacteriosis cases by month, EU/EEA, 2013–2017



Source: Country reports from Austria, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.

Indicator-based and event-based surveillance: TESSy and EPIS



Structured data:
- Isolates
- Notified cases

Event based data:
- Unusual increase of cases
- Outbreaks

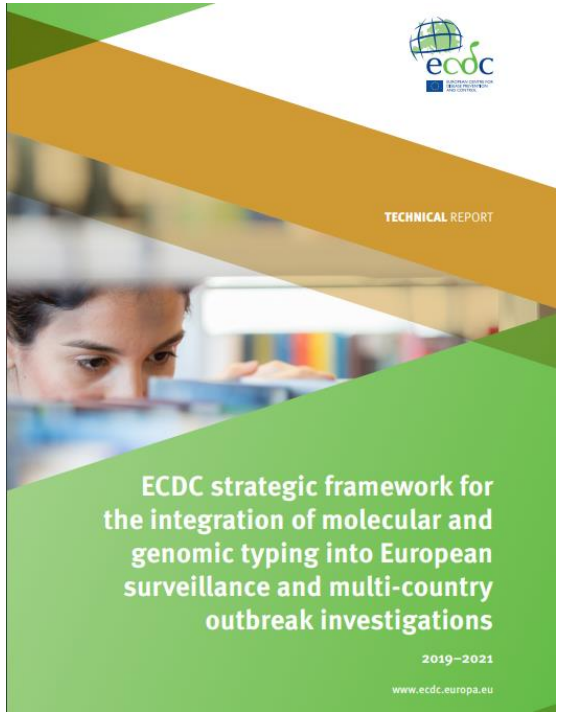
Whole genome sequencing (WGS) at ECDC

ECDC strategic framework for WGS by 2021:

- Cross-border outbreak investigations
- Control-oriented surveillance (cross-border outbreak detection)
- Strategy-oriented surveillance

Extensive list of diseases and pathogens to be implemented

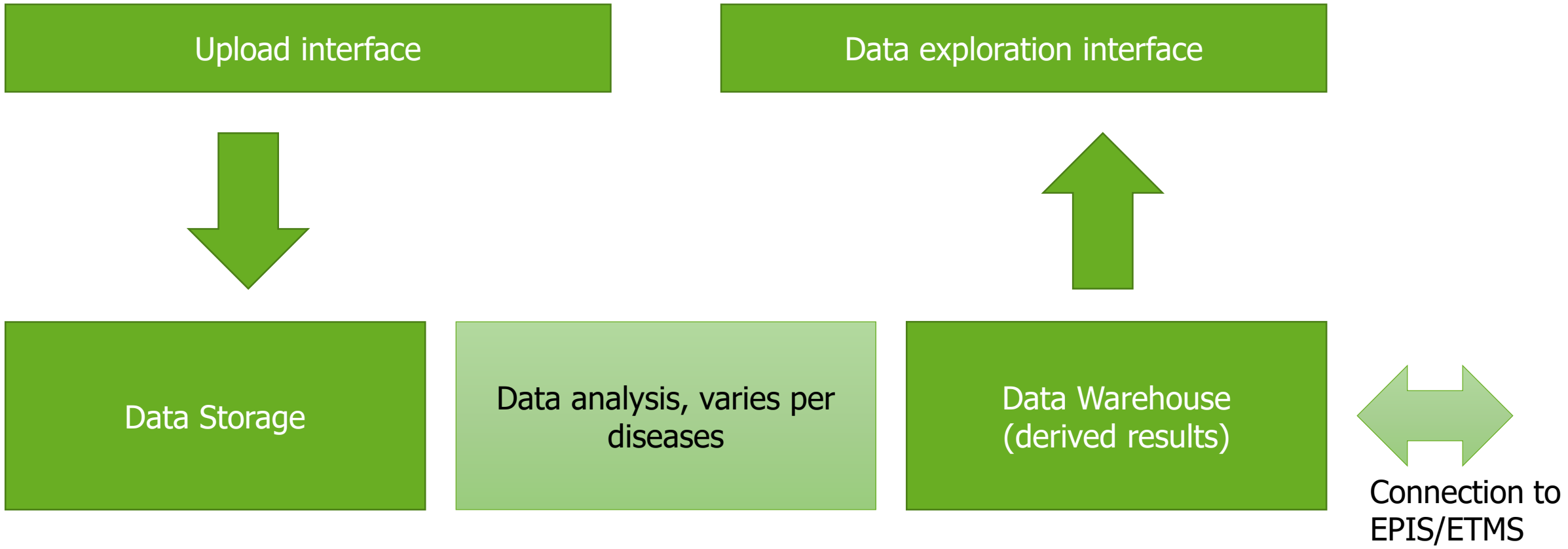
- Food- and waterborne diseases
- Vaccine-preventable diseases
- Multidrug-resistant tuberculosis
- Antimicrobial resistance
- Influenza
- Any emerging infectious disease threat at the EU level



ECDC WGS system upgrade, design principles

- Re-use existing tools and resources for storage, calculation and visualisation
- Integrate data from several different calculation systems in the same Data Warehouse data model
- Encourage open data sharing while also fulfilling all data protection and access control requirements
- Focus on the user experience and service-oriented architecture (data submission and data exploration)

Data flow schematic



Building blocks

Built by ECDC:

- Data upload interface (API and GUI), TESSy
- Data warehouse for derived data, new
- Web service API for data consumption, new
- Cluster detection scripts, existing
- Web page to hold visualisation components, new

Re-used:

- Storage (currently ENA and network drive, cloud storage to be implemented in 2020)
- Calculation systems and operational databases (currently Bionumerics and BIGSdb, more will be added as needed)
- Interactive visualisation (currently MicroReact, we are also looking at NextStrain)
- Charts and tables libraries (charts.js, datatables.js)

Timeline

ECDC upgraded system expected to be operational Q1 2020

Currently focus on *Listeria*, in 2020 we expect to have support for continous surveillance for *Listeria, Salmonella*, STEC, *Neisseria meningitidis*, Influenza and MDR-TB

Outbreak support for *Campylobacter*

ECDC-EFSA joint database implementation will start Q1 2020

ECDC tool for analysing WGS data for clusters

ecdc Molecular Typing

Prototype version 0.17
Powered by MicroReact

LISTISO EBOV IMD ZIKV

Visualisation | Molecular clusters/datasets | My data | Settings

These table shows your uploaded data, you can choose to share or not share individual isolates with other users.

Save settings | Reload data from source (slow, you must refresh your browser to see the changes)

Column visibility | Copy | CSV | Show 100 entries | Search:

LISTISO submissions

Country	BatchId	RecordId	Created date	Modified date	QC (core loci detected)	QC (PASS/FAIL)	DateUsedForStatistics	Cluster code	Event	Share assembly
[redacted]			2019-09-19 14:52:02	2019-09-25 08:44:34	97.76	PASS	2019-06-18			☑
			2019-09-19 14:52:01	2019-09-25 08:44:31	98.56	PASS	2019-05-13			☑
			2019-09-19 14:52:01	2019-09-25 08:44:30	98.45	PASS	2019-05-26	2019-04.LIST.03.CC570		☑
			2019-09-18 18:12:43	2019-09-25 08:44:32	100.0	PASS	2019-08-08			☑
			2019-09-18 18:12:43	2019-09-25 08:44:28	100.0	PASS	2019-08-28	2017-10.LIST.82.CC1.AscI.0053.ApaI.0044		☑
			2019-09-18 17:02:27	2019-09-20 15:13:34	99.94	PASS			UI-490 Cluster B ST32	☑
			2019-09-18 17:02:27	2019-09-20 15:13:39	99.88	PASS			UI-490 Cluster B ST32	☑
			2019-09-18 17:02:27	2019-09-20 15:13:43	99.94	PASS			UI-490 Cluster B ST32	☑
			2019-09-18 12:19:37	2019-09-25 08:44:29	99.94	PASS	2018-09-20	2018-10.LIST.18.ST32		☑
			2019-09-18 12:19:37	2019-09-25 08:44:29	99.94	PASS	2019-07-18	2018-10.LIST.18.ST32	UI-490 Cluster B ST32	☑
			2019-09-18 12:19:36	2019-09-25 08:44:28	99.82	PASS	2019-06-15	2018-10.LIST.18.ST32	UI-490 Cluster B ST32	☑
			2019-09-06 10:53:35	2019-09-20 10:16:49	99.88	PASS		2019-09.LIST.01.CC388	UI594	☑
			2019-08-27 16:23:42	2019-08-28 14:30:00	98.51	PASS		2019-05.LIST.06		☑
			2019-08-27 16:23:42	2019-09-25 09:23:31	97.99	PASS	2019-04-17	2017-10.LIST.01.CC87		☑
			2019-08-27 16:23:42	2019-09-25 09:23:41	97.94	PASS	2019-04-19	2017-10.LIST.01.CC87		☑

ecdc Molecular Typing

Prototype version 0.17
Powered by MicroReact

LISTISO EBOV IMD ZIKV

Visualisation | Molecular clusters/datasets | My data | Settings

These tables shows molecular clusters detected by ECDC by applying automated algorithms, see respective surveillance protocols for details. Only multi-country clusters and clusters in your country are shown. Included are also generated datasets of particular interest.

Cluster ID	Event ID	Cluster status	Number of isolates	Countries	Updated (7d)	First date	Last date	Last updated	Extended cluster (7AD)	Neighborhood (50 AD)
Search Cluster ID	Search Event ID	Search Cluster stat	Search Number of	Search Countries	Search Updated (7	Search First date	Search Last date	Search Last updat	Search Extended c	Search Neighborhc
2017-10.LIST.82.CC1.AscI.0053.ApaI.0044		OPEN	5			2012-07-15	2019-08-28	2019-09-25	2017-10.LIST.82.CC1.AscI.0053.ApaI.0044	2017-10.LIST.82.CC1.AscI.0053.ApaI.0044
2018-10.LIST.18.ST32	UI-490 Cluster B ST32	OPEN	8			2018-01-27	2019-07-18	2019-09-25	2018-10.LIST.18.ST32	2018-10.LIST.18.ST32
2019-03.LIST.02.CC1		OPEN	8			2018-02-25	2019-06-12	2019-08-06	2019-03.LIST.02.CC1	2019-03.LIST.02.CC1
2017-10.LIST.01.CC87		OPEN	37			2011-12-13	2019-06-05	2019-09-25	2017-10.LIST.01.CC87	2017-10.LIST.01.CC87
2019-06.LIST.10.CC8	UI563	OPEN	5			2019-05-15	2019-05-15	2019-07-09	2019-06.LIST.10.CC8	2019-06.LIST.10.CC8
2017-10.LIST.69.CC1		OPEN	3			2010-08-27	2019-02-19	2019-04-30	2017-10.LIST.69.CC1	2017-10.LIST.69.CC1
2016-11.LIST.01	UI452	OPEN	25			2014-07-23	2019-02-18	2019-05-29	2016-11.LIST.01	2016-11.LIST.01
2017-06.LIST.70.CC1.AscI.0053.ApaI.0040		OPEN	48			2015-08-26	2019-02-18	2019-07-03	2017-06.LIST.70.CC1.AscI.0053.ApaI.0040	2017-06.LIST.70.CC1.AscI.0053.ApaI.0040
2019-07.LIST.14.CC87		OPEN	2			2018-08-15	2018-11-30	2019-08-07	2019-07.LIST.14.CC87	2019-07.LIST.14.CC87
2019-07.LIST.12.CC2		OPEN	2			2014-09-26	2018-11-26	2019-08-07	2019-07.LIST.12.CC2	2019-07.LIST.12.CC2
2019-07.LIST.13.CC6		OPEN	2			2013-10-15	2018-11-21	2019-08-07	2019-07.LIST.13.CC6	2019-07.LIST.13.CC6
2018-05.LIST.07.CC155		OPEN	7			2015-08-30	2018-10-18	2019-08-07	2018-05.LIST.07.CC155	2018-05.LIST.07.CC155
2015-04.LIST.04.CC6.AscI.0002.ApaI.0072	UI583_test	OPEN	26			2010-01-27	2015-03-26	2019-09-10	2015-04.LIST.04.CC6.AscI.0002.ApaI.0072	2015-04.LIST.04.CC6.AscI.0002.ApaI.0072
2019-06.LIST.11.CC6	UI552	OPEN	3					2019-07-09	2019-06.LIST.11.CC6	2019-06.LIST.11.CC6
2019-08.LIST.01.CC9		CLOSED	2			2018-01-17	2019-06-04	2019-08-06	2019-08.LIST.01.CC9	2019-08.LIST.01.CC9
2019-04.LIST.03.CC570		CLOSED	3			2019-05-26	2019-05-26	2019-09-25	2019-04.LIST.03.CC570	2019-04.LIST.03.CC570
2017-11.LIST.11.CC87		CLOSED	7			2015-05-30	2019-05-01	2019-07-10	2017-11.LIST.11.CC87	2017-11.LIST.11.CC87
2019-03.LIST.07.CC6		CLOSED	3			2017-03-24	2019-02-20	2019-07-08	2019-03.LIST.07.CC6	2019-03.LIST.07.CC6
2019-07.LIST.01.CC2		CLOSED	2			2016-05-24	2019-02-10	2019-07-10	2019-07.LIST.01.CC2	2019-07.LIST.01.CC2

Molecular Typing

Prototype version 0.17
Powered by MicroReact

LISTISO | EBOV | IMD | ZIKV

Visualisation | Molecular clusters/datasets | My data | Settings

Pathogen: Listeria

Clustering method: Single linkage

Include data from SRA and ENA
Include data from all countries
Sampled within X days:
Uploaded within X days:
Search for specific cluster/dataset:
Search for specific UI:
Search for specific RecordIds:

After filtering, also include all matches within X genetic differences:

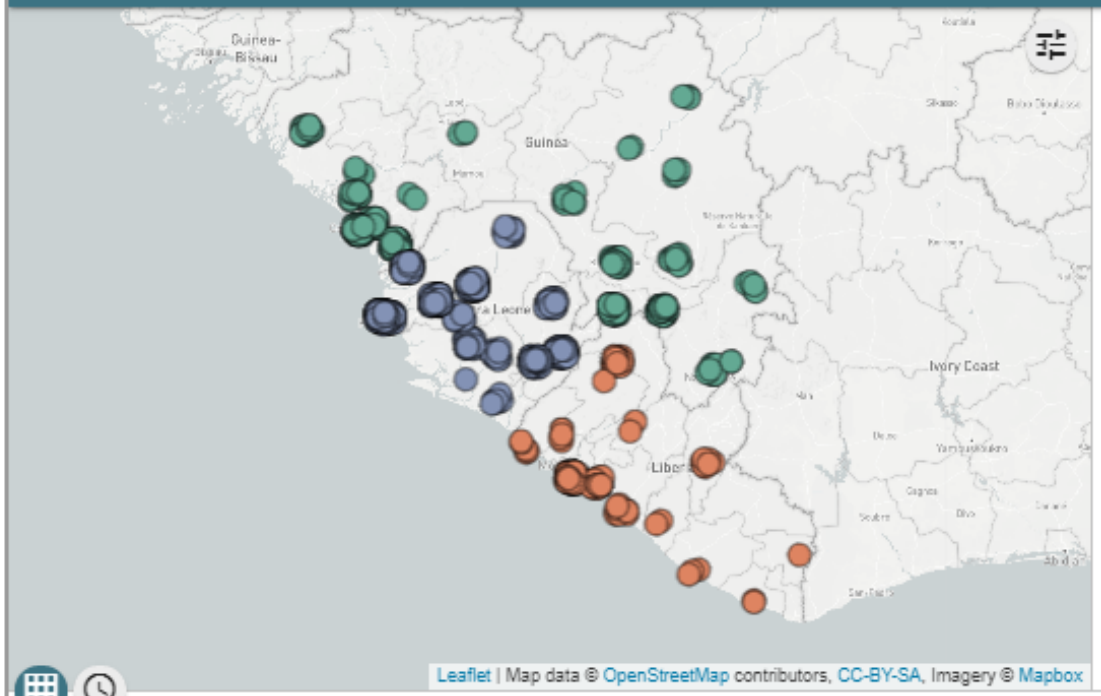
Update

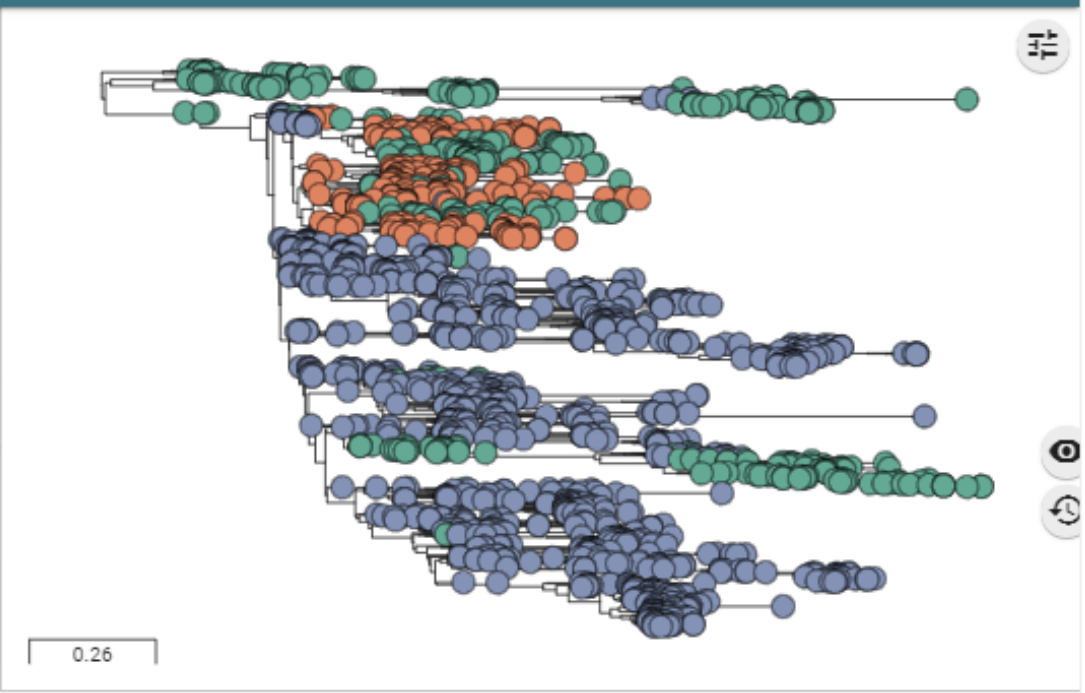
URL for sharing:
URL for recreating:

MicroReact | Tables and figures | Distance matrix | Isolate table and download

West African Ebola epidemic 2013-2016

SEARCH IN ALL COLUMNS 1610 of 1610





Columns | Comfortable

ID	number	country	location	locImputation	nonJitLat	nonJitLong	
EBOV|20140008|KR653251|SL...	1	SLE	Kenema		7.858614095	-11.18956765	2014-08-22
EBOV|20140024|KR653252|SL...	2	SLE	PortLoko		8.685188747	-12.53675216	2014-08-20
EBOV|20140038|KR653267|SL...	3	SLE	Tonkolili		8.465489914	-12.21193614	2014-08-23
EBOV|20140091|KR653239|SL...	4	SLE	WesternRural		8.44632216	-13.16852491	2014-08-22

Molecular Typing

Prototype version 0.17
Powered by MicroReact

LISTISO | EBOV | IMD | ZIKV

Visualisation | Molecular clusters/datasets | My data | Settings

Pathogen: Listeria

Clustering method: Single linkage

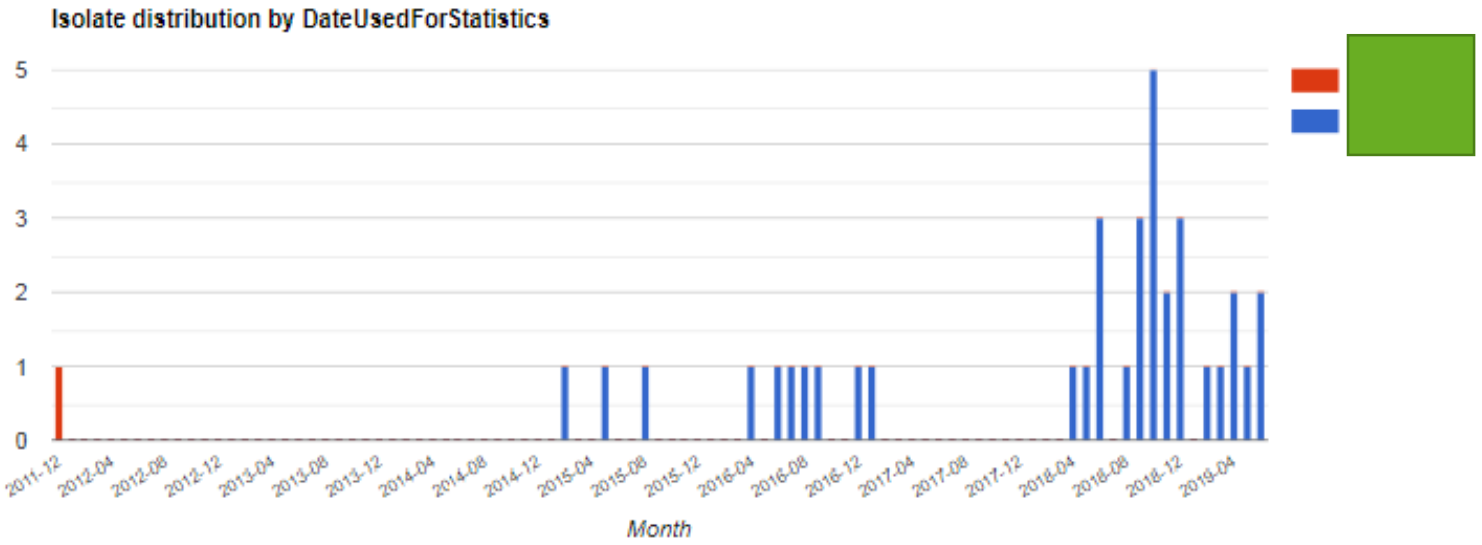
Include data from SRA and ENA
Include data from all countries
Sampled within X days:
Uploaded within X days:
Search for specific cluster/dataset: 2017-10.LIST.01.CC87
Search for specific UI:
Search for specific RecordIds:

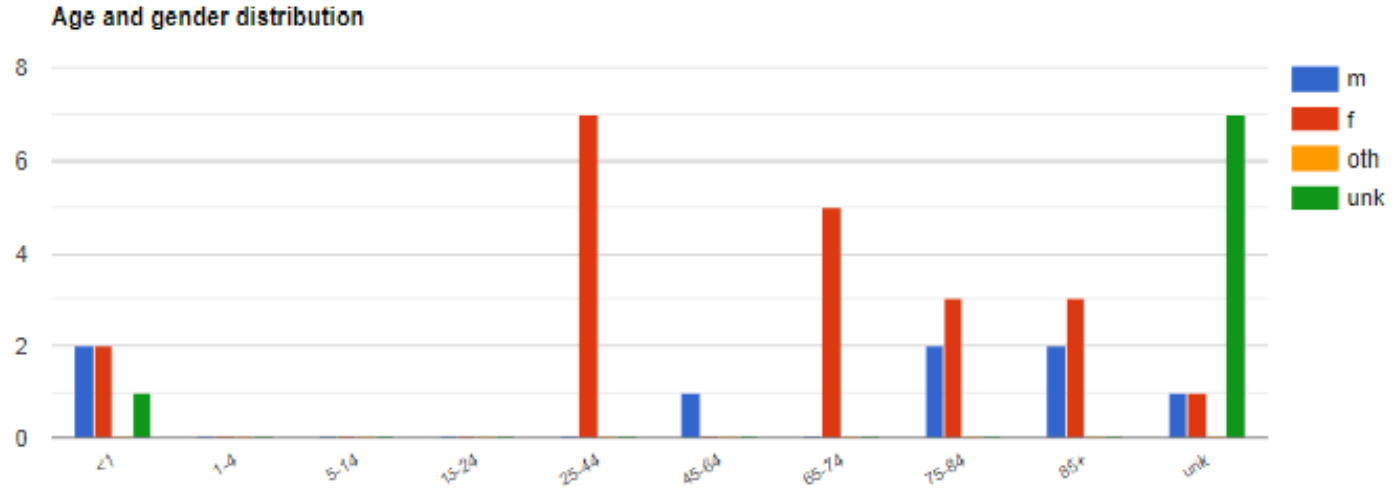
After filtering, also include all matches within X genetic differences: 7

Update

URL for sharing: http://microreact.ecdcnet.e
URL for recreating: http://zbiodev1.idmdevdmz

MicroReact | Tables and figures | Distance matrix | Isolate table and download



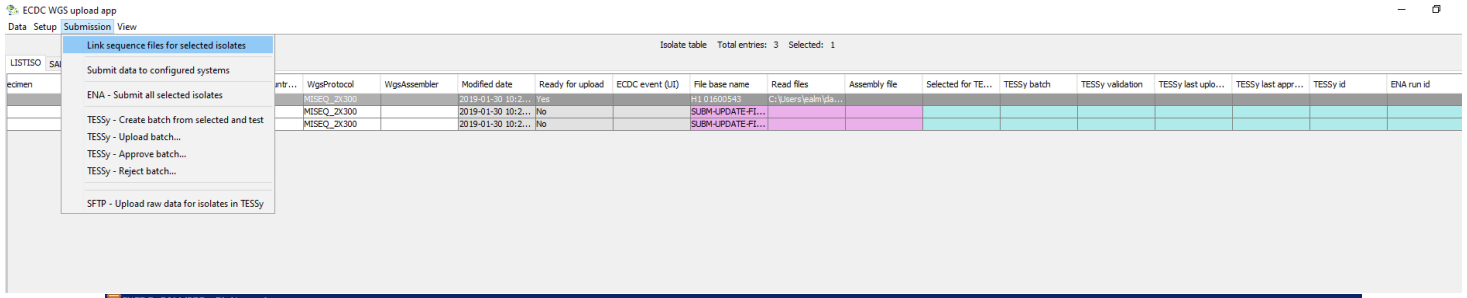


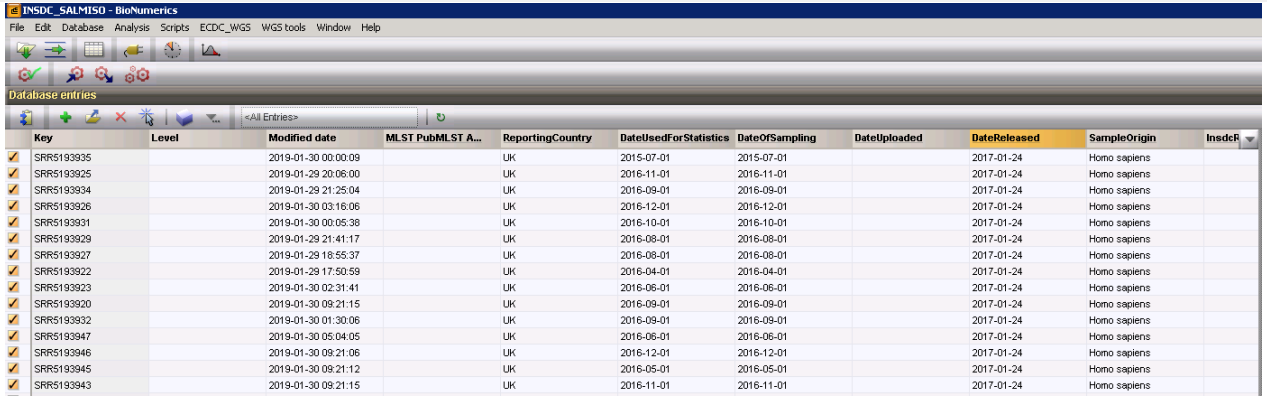
Thank you!!

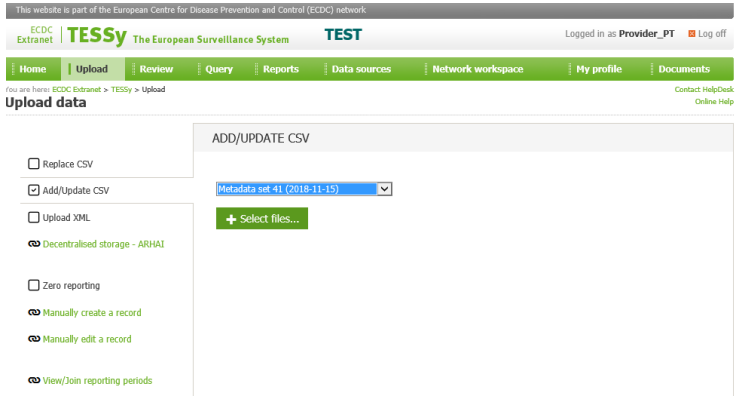
Software options for WGS data upload

- ECDC WGS upload application
- ECDC Bionumerics client plugin
- Direct TESSy submission

All three options include mandatory submission to TESSy isolate-based subjects



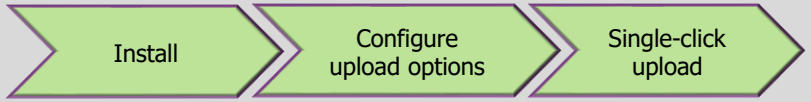




ECDC WGS upload application

How to use

- Download and install the application
- Read the manual and configure the data upload for each pathogen, personal support is available through the ECDC FWD mailbox
- Configure data import/mapping from Excel/csv, your LIMS system, or enter data directly
- Start uploading in a few clicks according to the reporting protocol
- The epidemiological data are submitted to TESSy only, the WGS reads/assembly is submitted to the selected system(s)



Features

- Can be configured to import data from databases or local files (MySQL, SQL Server, SQLite, Excel, csv)
- Configure only once, single click upload
- Can upload assemblies to TESSy and SFTP, raw reads to SFTP and ENA (configurable)
- Data sharing through SFTP and ENA