COllaborative Management Platform for detection and Analyses of (Re-) emerging and foodborne outbreaks in Europe

A global platform for the sequence-based rapid identification of pathogens

Prof. Frank M. Aarestrup, coordinator (Technical University of Denmark)
Prof. Marion Koopmans, deputy coordinator (Erasmus Medical Center, the Netherlands)

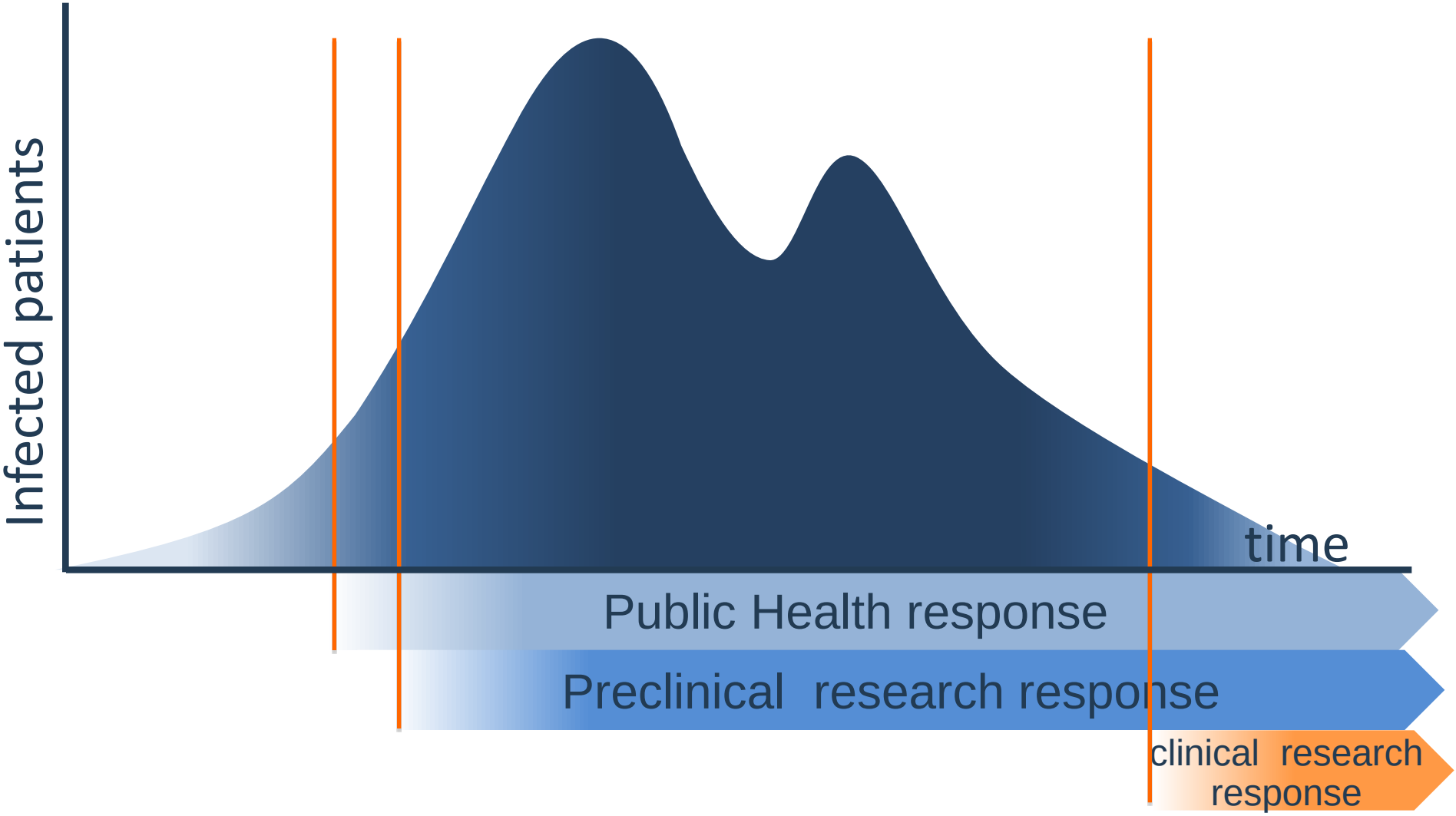
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.
research and innovation programme under grant agreement No 643

Infectious disease situation 2015

- Dynamics of common infectious diseases are changing
  - Demographic change, population density, anti vaccine, AMR, etc.
- New diseases emerge frequently
  - Deforestation, population growth, health system inequalities, travel, trade, climate change
- Effects are difficult to predict due to complexity of problems
  - Rapid flexible response
- Public health and clinical response depend on global capacity for disease surveillance
  - Rapid sharing, comparison and analysis of data from multiple sources and using multiple methodologies

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Clinical research response to ID outbreaks usually fragmented and too late



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

What the world needs

- Real-time data on occurences of all infectious agents
- (Automatic) detection of related clusters in time and space
- Possibility to observe trends in clones and species as well as virulence and resistance
- Ability to rapidly compare between all types of data

There can be no real-time surveillance without real-time data sharing

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

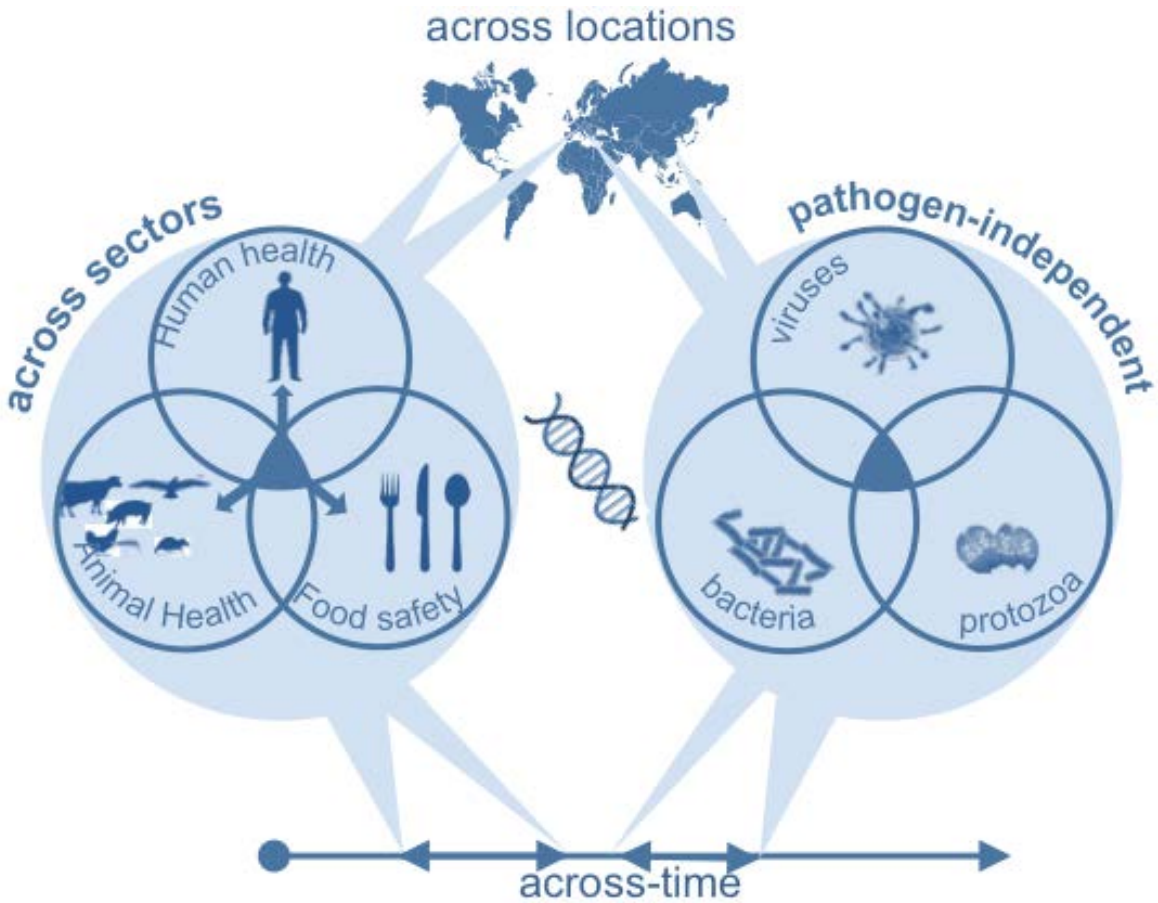
NGS advantages

- Laboratory diagnostics increasingly rely on (pathogen) genomic information
- RNA / DNA are common across pathogens, therefore, methods to analyse pathogen genomes are potentially universal
- Next generation sequencing capacity is developing fast, and costs are becoming competitive

- ➢ Capturing NGS developments may provide a universal language that can be harnessed for early detection of outbreaks across disciplines and domains
- ➢ If the technology keeps developing, less equipped labs may leapfrog

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Our vision:
to build one system that serves all



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Epidemic preparedness research:
European Union-supported efforts

- Prediction
  - Emergence, surveillance, modelling
- Early recognition and containment
  - Surveillance, clinical awareness, infection control
- Data infrastructure
  - Data repositories, sharing
- clinical research
  - pathogen & disease characterisation
  - prevention & treatment
- funding
  - rapid responses

PREDEMICS, antigone, Impetus — 2009-2016, € 36 M

compare — 2015-2020, € 21 M

elixir — 2013-?, € >100 M

PREPARE (Platform foR European Preparedness Against (Re-)emerging Epidemics) — 2014-2019, € 24 M

GloPID-R — 2015-2020, € 3 M

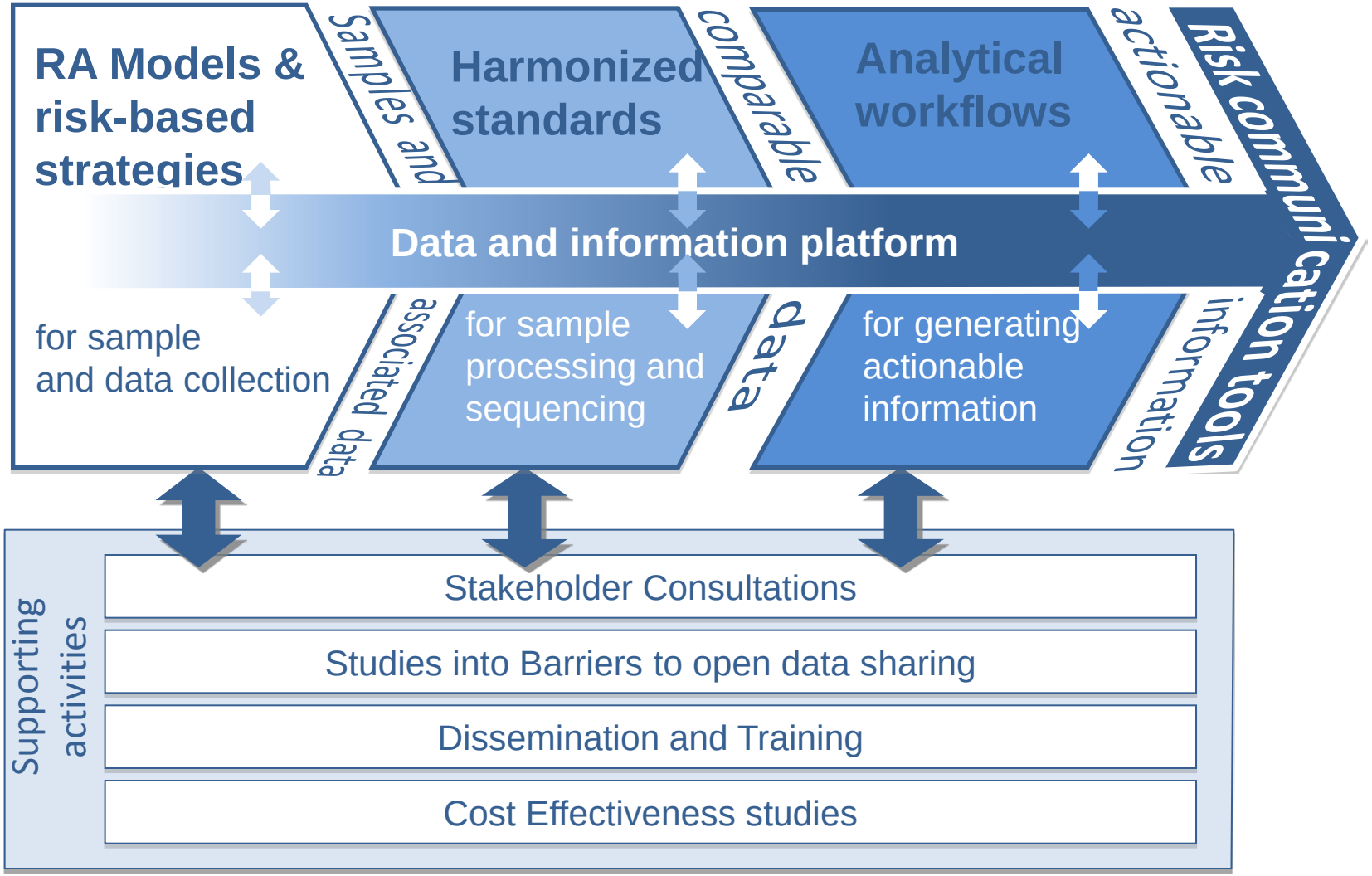
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

COMPARE principles

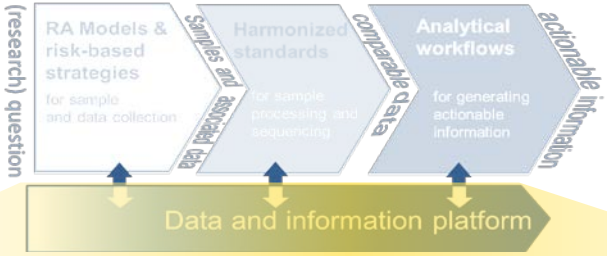
- COMPARE is sector-, domain- and pathogen**-independent**
- Analyzing **sequence-based pathogen data** in combination with **associated (clinical, epidemiological and other) data**
- Building on **established infrastructures (ENA, Elixir)**
- COMPARE is a **user-driven system**, designed with the information needs of its intended diverse group of future users and other stakeholders in mind
- COMPARE will make **optimal use of existing and future complementary systems, networks and databases,** ensuring compatibility where needed
- COMPARE is a **flexible, scalable and open-source based** information-sharing platform
- 1 December 2014 – 31 November 2019

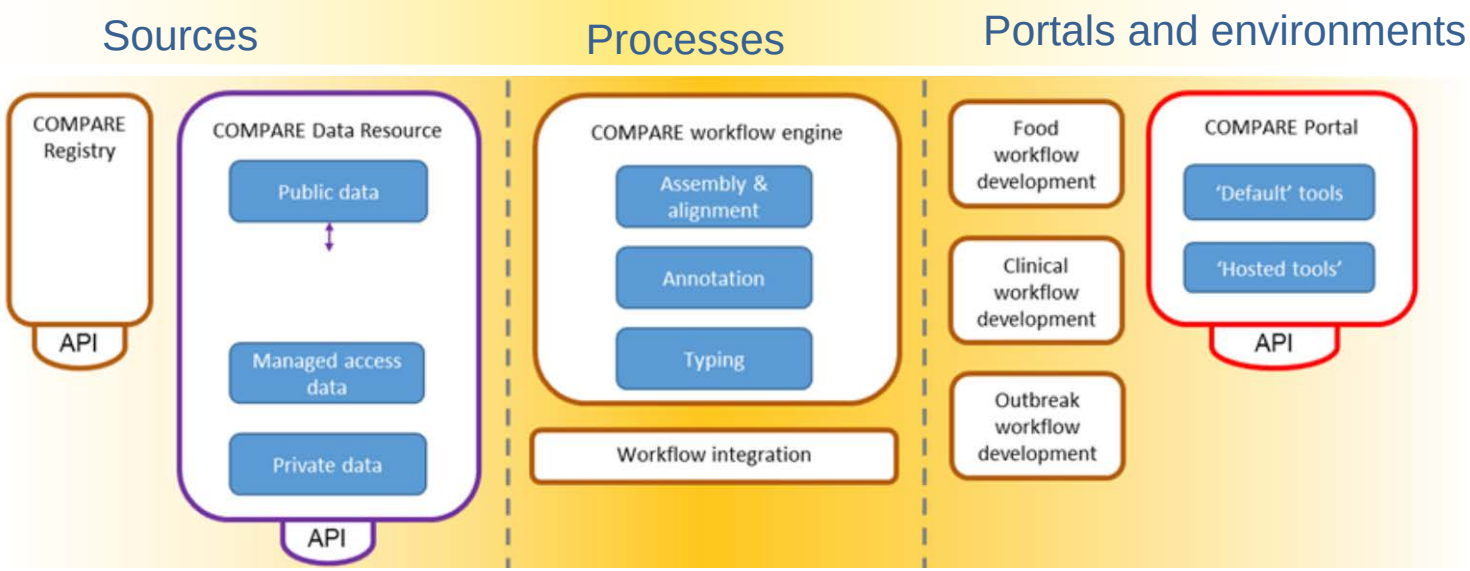
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Project structure



WP9 Information sharing platform

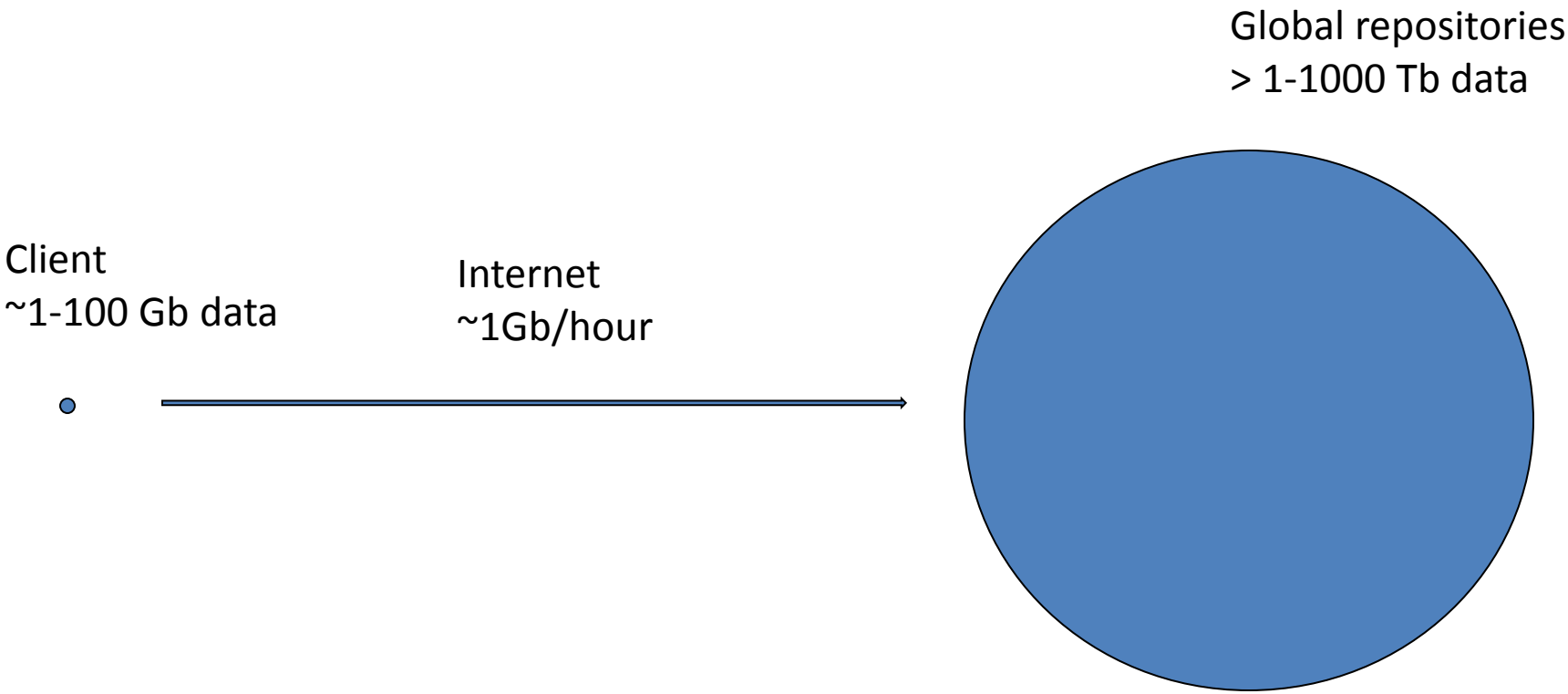




Building on the EU ESFRI Elixir, EMBL and DTU infrastructures

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Data comparison problem

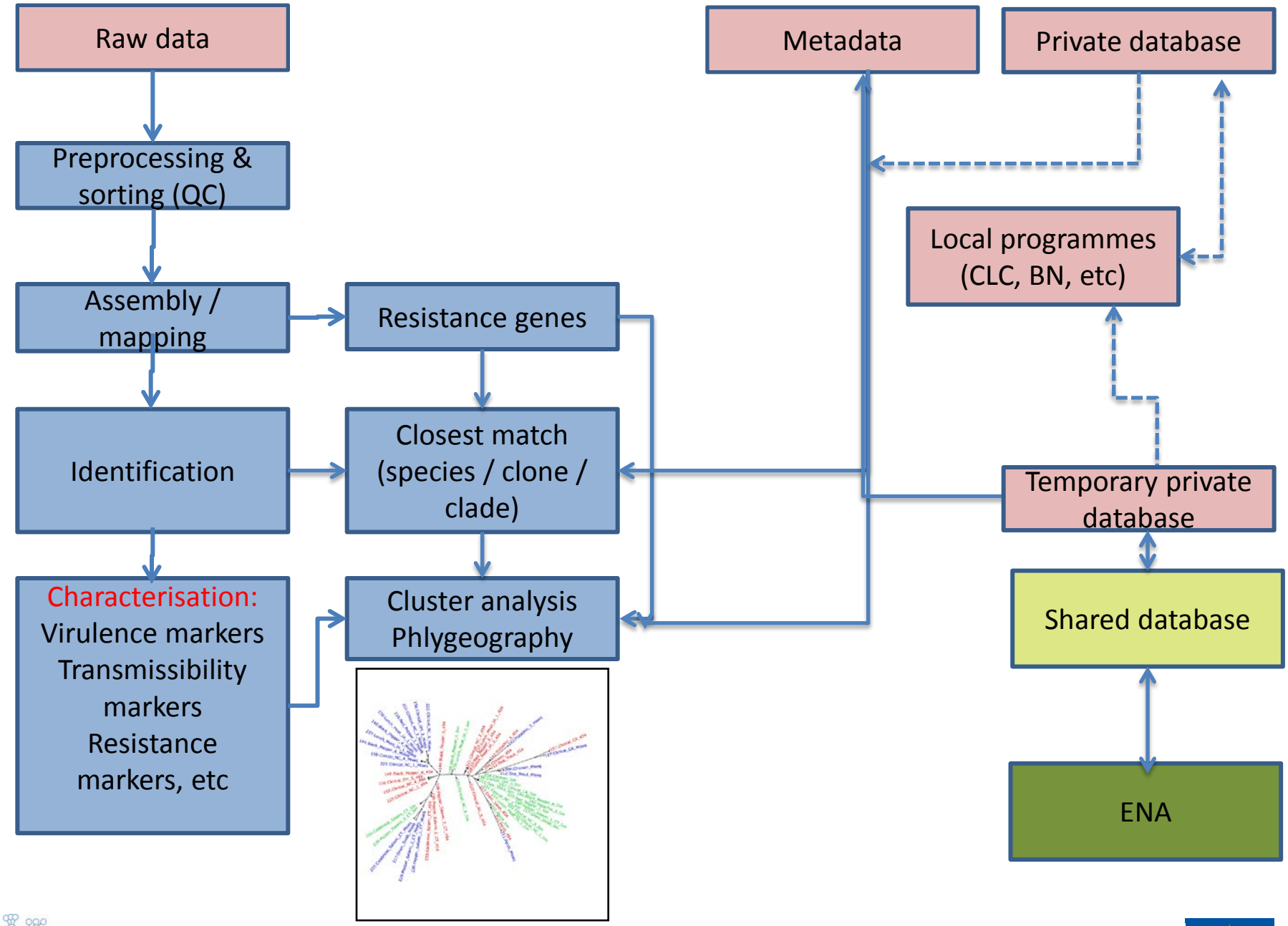


This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Workpackage 9: Data infrastructure design

- Sharing of highly structured and well-described data
  - COMPARE standards (e.g. checklists relating to isolates)
  - Data reporting tools that support the structuring and validation of data
- Support for a spectrum of data types
  - Raw NGS to assemblies
  - Derived data: typing information, AMR profiles, etc.
- Data availability
  - Early pre-publication private access, where required, for defined user groups according to explicit data-sharing agreements
  - Rapid flow of data to full public availability and global presentation
- Data discovery and retrieval systems, taking full advantage of data and metadata structures
  - Cloud-based autonomous analytical workflows (assembly, typing, phylogenetics, etc.)
  - Unifying data portal

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Current status - IT

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

The **CO**llaborative **M**anagement **P**latform for detection and **A**nalyses of (**R**e-) emerging and foodborne outbreaks in **E**urope is a collaboration of 29 institutions with experience in outbreak detection and response in areas of human health, animal health and food safety.

COMPARE Reference Genomes

This COMPARE Reference Genomes page offers a curated selection of published reference sequences covering viral, bacterial and protozoan genomes. These sequences can be searched and retrieved via the following URLs as tagged records in the European Nucleotide Archive (ENA). The complete COMPARE Reference Genomes dataset can be retrieved via the following URL:

http://www.ebi.ac.uk/ena/data/xref/search?source=COMPARE-RefGenome

ENA sequence or sample accessions for a single sample/isolate in the dataset can be returned using the following URL:

http://www.ebi.ac.uk/ena/data/xref/search?source=*<source>*&source_accession=*<source_accession>*

where source_accession is the isolate/sample name as shown in the table below, for example:

http://www.ebi.ac.uk/ena/data/xref/search?source=COMPARE-RefGenome&source_accession=Beijing/55028/2007/CHN

The ENA record, shown in the 'Target primary accession' column of the result from the above URL, can be retrieved with the following URL:

http://www.ebi.ac.uk/ena/data/view/<Target_primary_accession>

where Target_primary_accession has been inserted from the response to the previous URL (e.g. http://www.ebi.ac.uk/ena/data/view/GQ856465).

More extensive functions are described for REST services relating to the COMPARE Reference Genomes. Users should note that records in the dataset are served from ENA and are denoted as belonging to the dataset through ENA cross-reference annotations.

The below table serves as a quick overview of the COMPARE Reference Genomes set with direct hyperlinks to the INSDC records. Text in brackets next to a taxon name represents serovar or genotype information.

Sample/Isolate ID	Aggregated taxonomic name or Taxon name	Genome	INSDC record(s)
Norwalk/1968/US	NoV/GI.P1/GI.1	complete	M87661
Southampton/1991/UK	NoV/GI.P2/GI.2	complete	L07418
NLV/VA98115/1998	NoV/GI.P3/GI.3	partial	AB287450
Chiba407/198/JP	NoV/GI.P4/GI.4	complete	AB042808
Musgrove/1989/UK	NoV/GI.P5/GI.5	partial	AJ277614
BS5/1997/DE	NoV/GI.P6/GI.6	complete	AF093797

Reference genomes

- Curated selection of published reference sequences covering viruses, bacteria and protozoa
- Summary page and entry point at http://www.ebi.ac.uk/ena/about/compare-reference-genomes
- Searchable through webservices
- Launched 18.6.15, extended 25.8.15

In progress

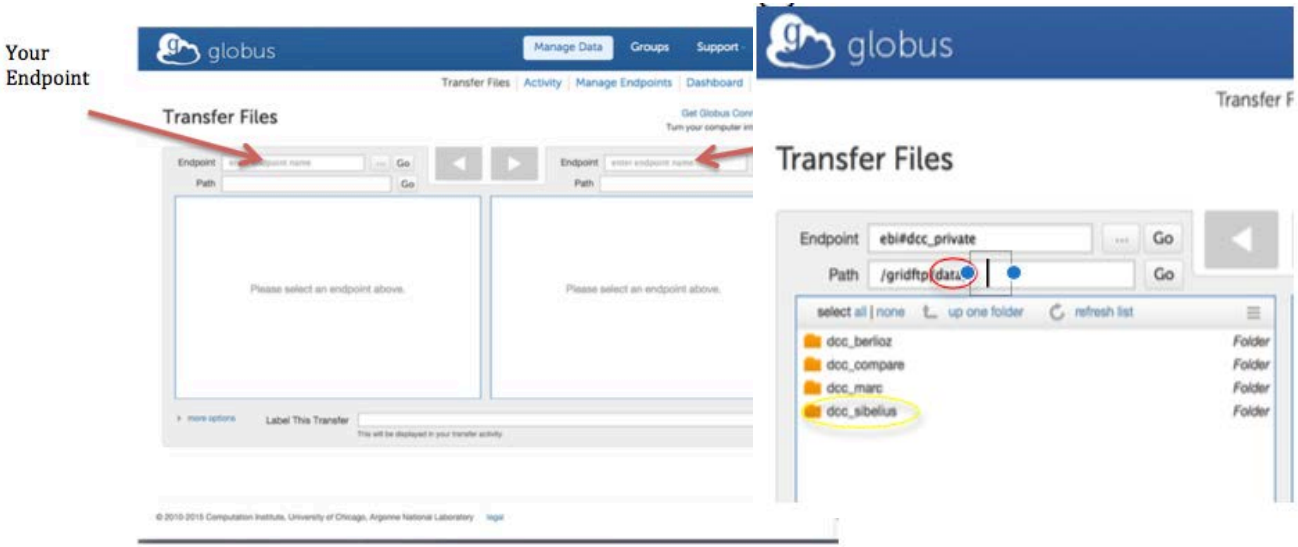
Protocols for sampling and handling

Ongoing benchmarking and ring trials

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

COMPARE data hubs

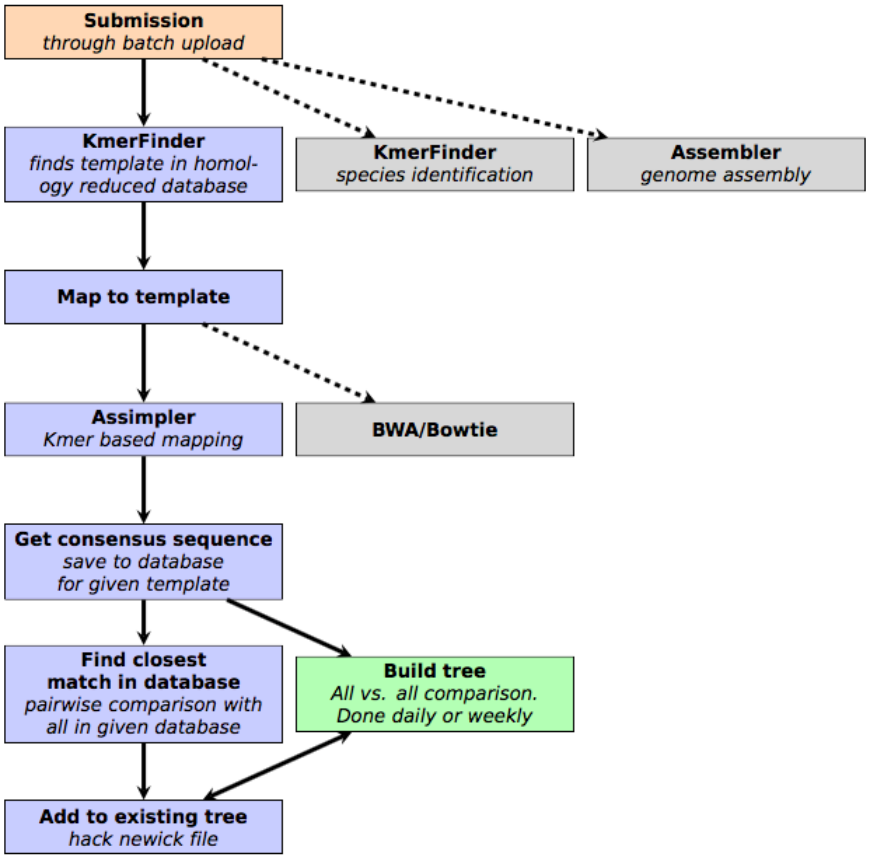
- On-request service to share pre-publication data
  - Set up: provider, consumer accounts, scope, etc.
  - Providers report data through existing COMPARE interfaces
  - Consumers retrieve metadata (web or spreadsheet manifest) and data (Globus FTP)
- Several existing hubs, including
  - **dcc_sibelius**, for the Influenza H5N8 pilot - pre-publication read data and metadata
  - **dcc_compare**, for replication of overall COMPARE public content, e.g. for external clouds or other infrastructure; currently all bacteria, viruses and some parasites, in time refinements expected



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

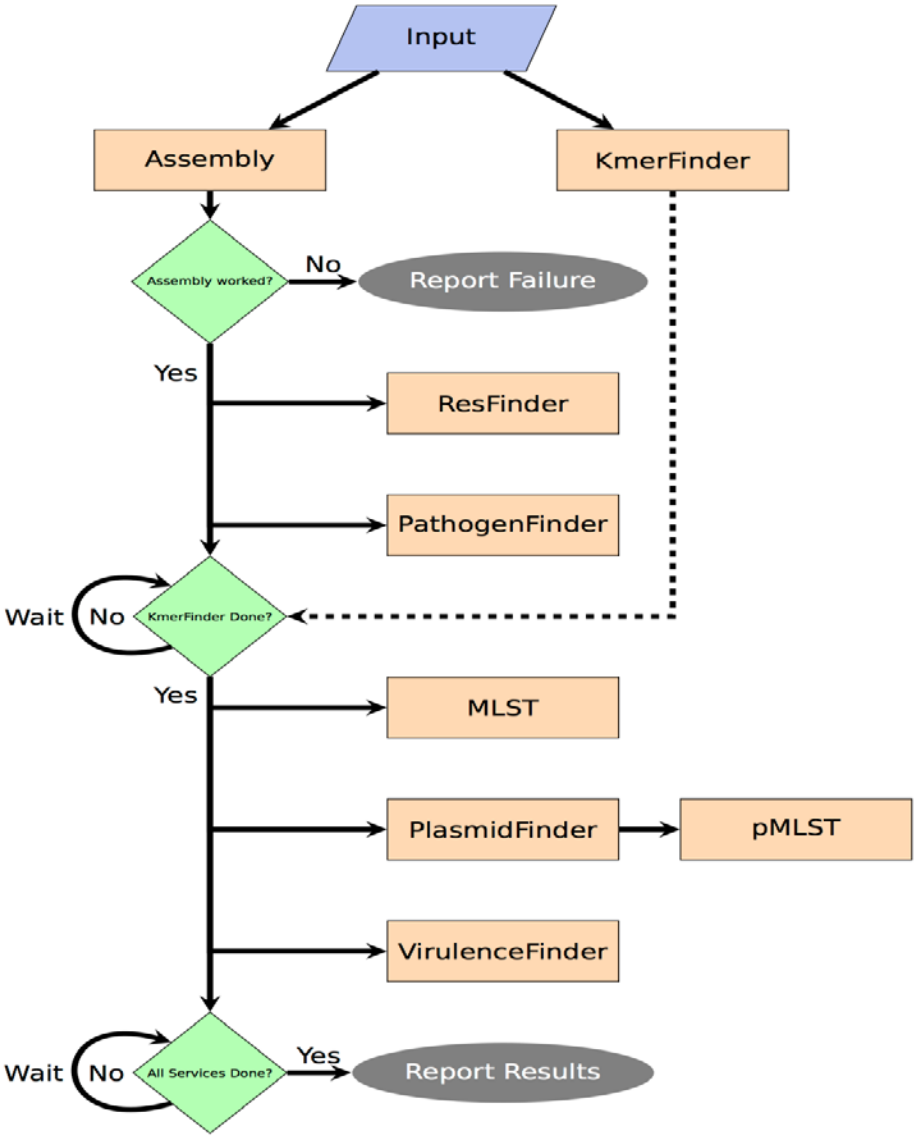
Cloud compute

- Embassy environment
  - Available to COMPARE
    - OpenStack framework (in place)
    - BioLinux standard machine (in place)
  - In development
    - CGE Docker
    - Discussions with CLIMB, NECTAR, IFB-Cloud (Galaxy)
- Development of analysis workflows
  - Evergreen
  - Assembly (Velvet + Prokka)
  - Metagenomics
  - CGE
  - iPython environment for engagement of workflow environments



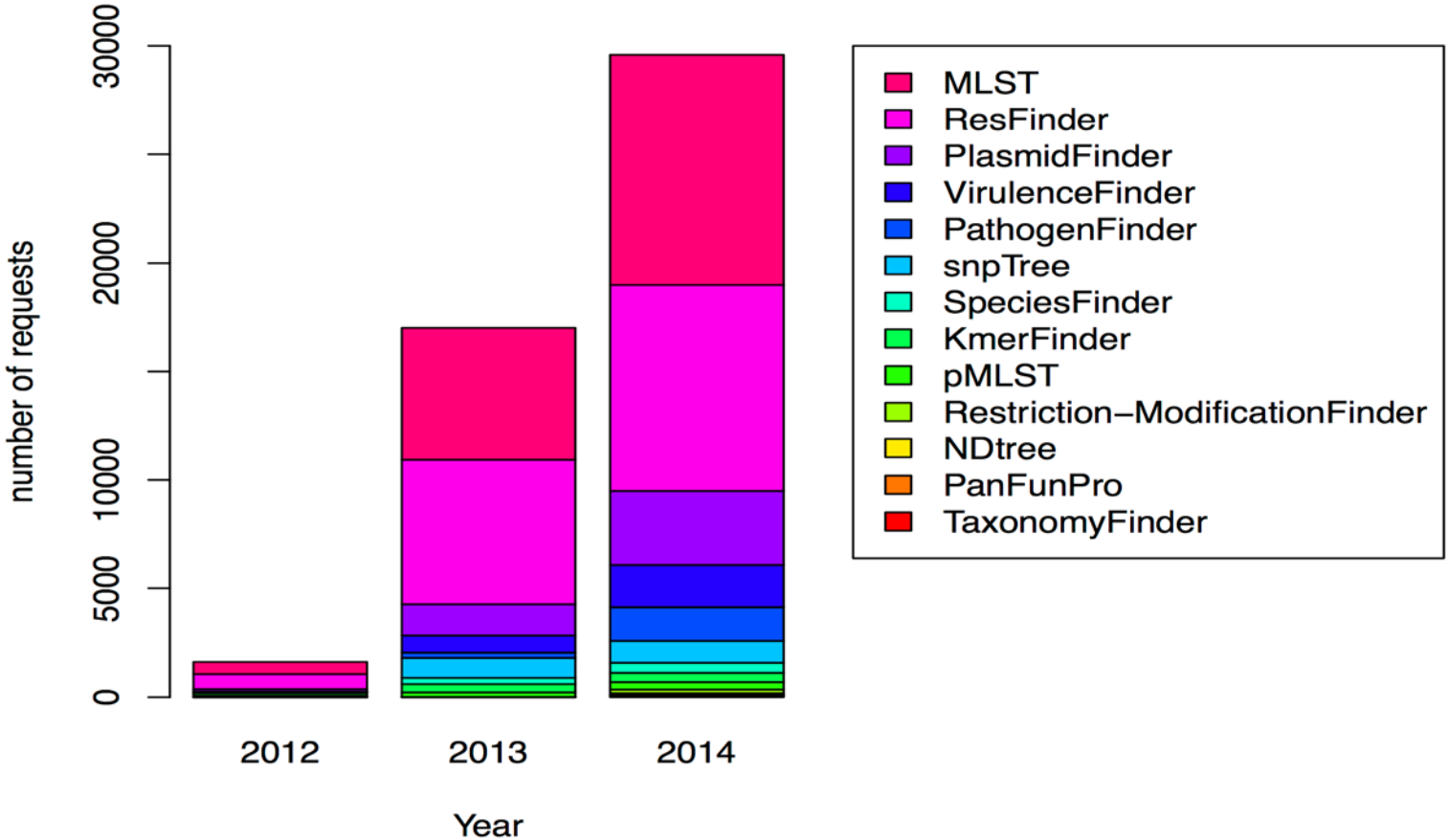
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Bacterial Analytic Pipeline



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

User Statistics



Until now: >175,000 submissions

From + 8,000 IP-adresses

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

CGE Server

https://cge.cbs.dtu.dk/services/cge/

Center for Genomic Epidemiology

Welcome **fmaa**

Home | Services | Instructions | Example | Article abstract

Bacterial Analysis Pipeline - Batch Upload

The CGE Bacterial Analysis Pipeline executes a workflow of services with predefined parameters and stores the submitted data and result in the database at the user's disposal.
View the version history of this server.

STEP 1: Download Metadata Template Template

STEP 2: Fill out template

STEP 3: Upload Metadata File

STEP 4: Select Files

STEP 5: Submit

Supported browsers

	Windows XP	Windows Vista	7	Windows 8	OS X Yosemite	ubuntu
≥ V 8.0					✓	
V 10.0			✓	✓		
edge			✓	✓		
≥ V 36.0				✓	✓	✓
≥ V 41.0	✓	✓	✓	✓	✓	✓

Progress Overview

Name	Size	Progress	Status

Remove all

Support | Scientific problems | Technical problems

Center for Genomic Epidemiology, DTU, Kemitorvet, Building 204, 2800 Kgs. Lyngby, Denmark
Funded by: The Danish Council for Strategic Research
Last modified August 25, 2015 13:32:18 GMT

DA 11:39 16-11-2015

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

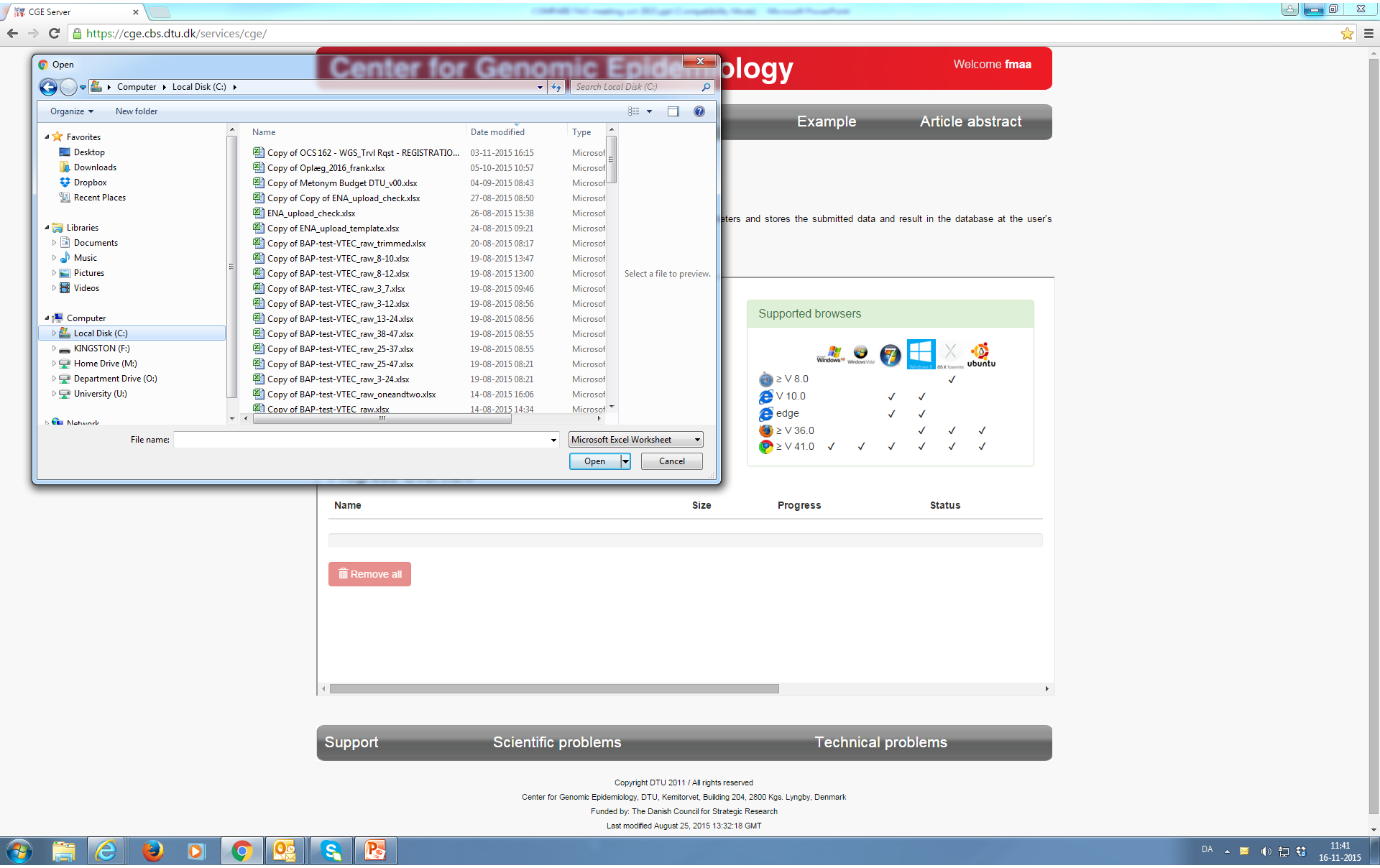
ENA_upload_template.xlsx - Microsoft Excel

C30: The insert size must be given as an integer (i.e. no decimals)

Attribute Description

Field Name	Mandatory	Description
sample_name	No	The name of the isolate the user uses to identify the sample
group_name	Not Yet Used	The name of the group. Adding a group name will incorporate the sample in the analysis pipeline execution of that group
file_names	Yes	Name of all files associated to this sample. Multiple filenames should be seperated by a space
sequencing_platform	Yes	Choose between: LS454, Illumina, Ion Torrent, ABI SOLiD and unknown
sequencing_type	Yes	Choose between: single, paired, mate-paired, and unknown
pre_assembled	Yes	Has the uploaded sample data been assembled? yes / no
sample_type	Yes	Choose between:*isolate* or*metagenomic*
organism	Yes	Write 'unknown' if the organism name is unknown
strain	No	*Strain type ID*
subtype	No	Subspecific genetic linage, i.e MLST, serovar and biotype
country	Yes	
region	No	
city	(No)	
zip_code	(No)	Please provide as much information as possible. Low resolution locations reduce the usability. The minimal recommended option is to provide either city, zip_code or longitude and latitude coordinates.
longitude	(No)	
latitude		
location_note	No	Additional relevant details about the location
isolation_source	Yes	The host from which the sample/isolate has been taken. This should be a proper scientific name or one of the aliases found in the host sheet of this document
source_note	No	Additional relevant details about the isolation source. i.e. blood, laboratory experiment or urine
pathogenic	Yes	Is the organism decreed pathogenic? yes / no / unknown
pathogenicity_note	No	Additional relevant details about the organism's pathogenicity
collection_date	Yes	The date of the sample collection. Use one of the following format: YYYY-MM-DD or YYYY-MM or YYYY
collected_by	No	Name of the institute or person who took the sample
usage_restrictions	Yes	Choose either 'private' or 'public'. Note that private data will be deleted after some time to free disk space on the server.
release_date	Not Yet Used	Write the date from when the data and results should be public available. Format: YYYY-MM-DD
email_address	Not Yet Used	Email adress of the uploader
notes	No	Any additional information can be added
insert_size	Yes	The insert size must be given as an integer (i.e. no decimals)

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Sample Overview

https://cge.cbs.dtu.dk/tools_new/client/platform/sample/

Center for Genomic Epidemiology

Welcome **fmaa**

Home | Services | Batch Upload | MapViewer

Sample Overview

Download all data in an Excel spreadsheet?
Download the resistance data in a more detailed Excel spreadsheet format?

	Name	Date	Country	City	Origin	Action
-	VTEC47	2012-01-01	Denmark		human	Analyse Edit Remove

Service	Date	Status	Action
KmerFinder-2.1	2015-08-27	Success	Remove
Assembler-1.0	2015-08-27	Success	Remove
ResFinder-2.1	2015-08-27	Success	Remove
VirulenceFinder-1.2	2015-08-27	Success	Remove
ContigAnalyzer-1.0	2015-08-27	Success	Remove
PlasmidFinder-1.2	2015-08-27	Success	Remove
MLST-1.6	2015-08-27	Success	Remove
pMLST-1.4	2015-08-27	Success	Remove

+	VTEC46	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC45	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC44	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC43	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC42	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC41	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC40	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC39	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC38	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC37	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC36	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC35	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC34	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC33	2012-01-01	Denmark		human	Analyse Edit Remove
+	VTEC29	2012-01-01	Denmark		human	Analyse Edit Remove

DA 14:23 03-09-2015

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

user-result-summary (3).xls - Microsoft Excel

F25 | Escherichia coli

organism	strain	subtype	country	region	city	zip_code	longitude	latitude	location_r	isolation_	source_nc	pathogeni	pathogeni	collection	collected_	notes	r_#contigs	r_#bases	r_n50	r_species	r_mlst	r_mlst_ge	r_pathoge	r_resistan	r_virulenc	r_plasmid	r_plasmid	r_pmlsts	r_spa	r_serotype
Escherich	C570-12	genes=vtx	Denmark				None	None		human		yes		2012			244	5302767	110998	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C541-12	genes=eae	Denmark				None	None		human		yes		2012			248	5283092	104387	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C641-12A	genes=eae	Denmark				None	None		human		yes		2012			268	5013998	73908	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,eae,e	IncFII(pSE	IncX1	IncF[F79*:	NA	NA
Escherich	C641-12B	genes=eae	Denmark				None	None		human		yes		2012			270	5017975	65597	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,eae,e	IncFII(pSE	IncX1	IncF[F79*:	NA	NA
Escherich	C696-12	genes=eae	Denmark				None	None		human		yes		2012			397	5437311	78570	Escherichi	ecoli[ST-2	ecoli[ADK	NA		astA,celb,i	IncFIB(AP(	ColRNAI,C	IncF[F-:A-:	NA	NA
Escherich	C697-12A	genes=eae	Denmark				None	None		human		yes		2012			218	5214489	103647	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncFIB(AP0019		IncF[F23:A	NA	NA
Escherich	C697-12B	genes=eae	Denmark				None	None		human		yes		2012			233	5210364	96034	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncFIB(AP0019		IncF[F23:A	NA	NA
Escherich	C748-12	genes=vtx	Denmark				None	None		human		yes		2012			279	5363379	83509	Escherichi	ecoli[ST-2	ecoli[ADK	NA		celb,ehxA,	IncFIA,Inc	IncB/O/K/	IncF[F25*:	NA	NA
Escherich	C749-12	genes=vtx	Denmark				None	None		human		yes		2012			147	4905062	116559	Escherichi	ecoli[Unkr	ecoli[ADK	NA		cdtB,espP,	IncFII(pSE11),IncFIB(		IncF[F74*:	NA	NA
Escherich	C750-12	genes=vtx	Denmark				None	None		human		yes		2012			396	5255403	35791	Escherichi	ecoli[Unkr	ecoli[ADK	NA		ehxA,espI,	IncFIB(AP(	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C751-12	genes=vtx	Denmark				None	None		human		yes		2012			323	4986040	39882	Escherichi	ecoli[ST-5	ecoli[ADK	NA	sul2	iha,iroN,is	IncFII(pRS	Col(BS512	IncF[F6*:A	NA	NA
Escherich	C659-12	genes=vtx	Denmark				None	None		human		yes		2012			441	5000054	26095	unknown	ecoli[ST-5	ecoli[ADK	NA	aadA1,dfr	astA,iha,p	IncFII(pRS	Col(BS512	IncF[F25*:	NA	NA
Escherich	C757-12	genes=eae	Denmark				None	None		human		yes		2012			429	5170629	29561	unknown	ecoli[ST-1	ecoli[ADK	NA		cif,eae,efa	IncI1,IncFIB(AP0019		IncF[F-:A-:	NA	NA
Escherich	C760-12	genes=vtx	Denmark				None	None		human		yes		2012			253	5005040	43519	Escherichi	ecoli[ST-5	ecoli[ADK	NA	blaTEM-1(	astA,celb,i	IncFII(pRS	Col(BS512	IncF[F25*:	NA	NA
Escherich	C767-12	genes=;ML	Denmark				None	None		human		yes		2012			313	5288716	48840	Escherichi	ecoli[ST-8	ecoli[ADK	NA		celb,gad,il	IncFII(29),	Col(MG82	IncF[F16*:	NA	NA
Escherich	C770-12	genes=vtx	Denmark				None	None		human		yes		2012			544	5501587	22843	unknown	ecoli[ST-7	ecoli[ADK	NA		astA,iha,ir	IncFII(29)	Col(MP18	IncF[F35*:	NA	NA
Escherich	C813-12	genes=eae	Denmark				None	None		human		yes		2012			647	5331283	21906	Escherichi	ecoli[Unkr	ecoli[ADK	NA	aph(3')-Ia	cif,eae,efa	IncFIB(AP(	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C812-12	genes=eae	Denmark				None	None		human		yes		2012			225	5247710	124133	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C816-12	genes=eae	Denmark				None	None		human		yes		2012			202	5157068	121282	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,cif,ea	IncFIB(AP(	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C818-12	genes=eae	Denmark				None	None		human		yes		2012			230	5256065	96061	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C819-12	genes=eae	Denmark				None	None		human		yes		2012			212	5259174	144767	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C820-12	genes=eae	Denmark				None	None		human		yes		2012			277	5436245	79411	Escherichi	ecoli[ST-2	ecoli[ADK	NA		astA,cba,c	IncI1,IncFIB(AP0019		IncI1[Unki	NA	NA
Escherich	C821-12	genes=eae	Denmark				None	None		human		yes		2012			260	5235519	115457	Escherichi	ecoli[ST-6	ecoli[ADK	NA		astA,eae,e	IncI1,IncF	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C849-12	genes=eae	Denmark				None	None		human		yes		2012			220	5245766	126677	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Morganell	C848-12	genes=;ML	Denmark				None	None		human		yes		2012			52	4055781	176604	NA	ecoli[Unkr	ecoli[ADK	NA					NA	NA	NA
Escherich	C850-12	genes=eae	Denmark				None	None		human		yes		2012			254	5065655	81376	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,cif,ea	IncFIB(AP001918)		IncF[F-:A-:	NA	NA
Escherich	C852-12	genes=eae	Denmark				None	None		human		yes		2012			243	5261562	104254	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C857-12	genes=eae	Denmark				None	None		human		yes		2012			250	5375983	90792	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,cif,ea	IncFIB(AP(	IncY,IncB/	IncF[F-:A-:	NA	NA
Escherich	C874-12	genes=vtx	Denmark				None	None		human		yes		2012			287	5474038	97468	Escherichi	ecoli[ST-4	ecoli[ADK	NA		ehxA,gad,i	IncFIB(AP(	Col156,Co	IncF[F25*:	NA	NA
Escherich	C884-12	genes=eae	Denmark				None	None		human		yes		2012			200	5155599	121342	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,cif,ea	IncFIB(AP(	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C885-12A	genes=eae	Denmark				None	None		human		yes		2012			115	4755036	103805	Escherichi	ecoli[ST-4	ecoli[ADK	NA		cif,eae,espA,espF,esp		Col156	NA	NA	NA
Escherich	C885-12B	genes=eae	Denmark				None	None		human		yes		2012			240	5253067	72400	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,eae,e	IncFIA(HI1	Col156	IncF[F5:A6	NA	NA
Escherich	C886-12	genes=eae	Denmark				None	None		human		yes		2012			258	5264368	96518	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,cif,ea	IncI1,IncF	IncB/O/K/	IncF[F-:A-:	NA	NA
Escherich	C887-12	genes=;ML	Denmark				None	None		human		yes		2012			77	4488733	153149	NA	ecoli[Unkr	ecoli[ADK	NA		gad,prfB			NA	NA	NA
Escherich	C891-12	genes=eae	Denmark				None	None		human		yes		2012			232	5212876	116338	Escherichi	ecoli[ST-2	ecoli[ADK	NA		astA,eae,e	IncFII,IncF	ColRNAI	IncF[F23:A	NA	NA
Escherich	C892-12	genes=vtx	Denmark				None	None		human		yes		2012			145	4939191	102286	Escherichi	ecoli[ST-2	ecoli[ADK	NA		celb,gad,k	IncFII(pCoo)		IncF[F12*:	NA	NA
Escherich	C893-12	genes=vtx	Denmark				None	None		human		yes		2012			208	5306779	82670	Escherichi	ecoli[Unkr	ecoli[ADK	NA		cba,cma,e	IncFIB(AP(	ColRNAI	IncF[F25*:	NA	NA
Escherich	C894-12	genes=eae	Denmark				None	None		human		yes		2012			206	5277292	94653	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncFIB(AP0019		IncF[F23:A	NA	NA
Escherich	C895-12	genes=vtx	Denmark				None	None		human		yes		2012			185	5204668	85195	Escherichi	ecoli[ST-6	ecoli[ADK	NA		gad,iha,lp	IncFII(pRS	Col156	IncF[F1:A-	NA	NA
Escherich	C896-12A	genes=eae	Denmark				None	None		human		yes		2012			247	5264467	68277	Escherichi	ecoli[ST-3	ecoli[ADK	NA		astA,eae,e	IncFIA(HI1	Col156	IncF[F5:A6	NA	NA
Escherich	C896-12B	genes=eae	Denmark				None	None		human		yes		2012			124	4758033	104256	Escherichi	ecoli[ST-4	ecoli[ADK	NA		cif,eae,espA,espF,esp		Col156	NA	NA	NA
Escherich	C904-12A	genes=vtx	Denmark				None	None		human		yes		2012			593	5523711	43914	Escherichi	ecoli[ST-6	ecoli[ADK	NA		celb,ehxA,	IncFIB(AP(	Col156	IncF[F25*:	NA	NA
Escherich	C905-12	genes=eae	Denmark				None	None		human		yes		2012			304	5058172	86262	Escherichi	ecoli[ST-1	ecoli[ADK	NA		astA,eae,e	IncFII,IncFIB(AP0019		IncF[F23:A	NA	NA

user-result-summary (3)

Ready

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Current developments - IT

- Ability to create own shared site
- Combining map, phylogeny and analysis (microreact)
- Evergreen trees
- Bring your own tools (already ability for own DB)

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Conclusions

- WGS/NGS is rapidly entering diagnostic and public health arena, with near real-time data generation
- Sequence platforms rapidly developing, cheaper, simpler
- Bottleneck at level of bioinformatics, particularly for intergroup comparison, national, international
- COMPARE aims to develop infrastructure and ICT to meet the coming demand
- In the coming years, we will be seeking partners for pilot projects

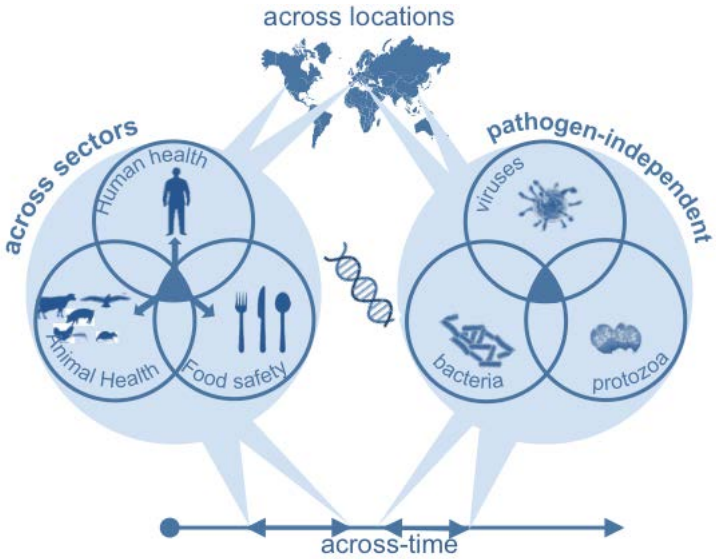
This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Pilot projects in 2016

- Cross cutting*: Escherichia coli,* ESBL, norovirus, metagenomics
- Clinical WP: AMR
- Public health WP: 6 food-borne pathogens
- Emerging infections WP: Influenza and MERS CoV
- Ad hoc

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Our vision: one system serves all



Guiding principles:

- Cross sector, cross domain, open source (not commercial)
- Interaction with the rest of the world (all inclusive)
- Data for action (actionable outputs)
- Central repository (ENA, DDJ, NCBI) (bring the tools to the data)

There can be no real-time disease detection & surveillance without real-time data sharing

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.