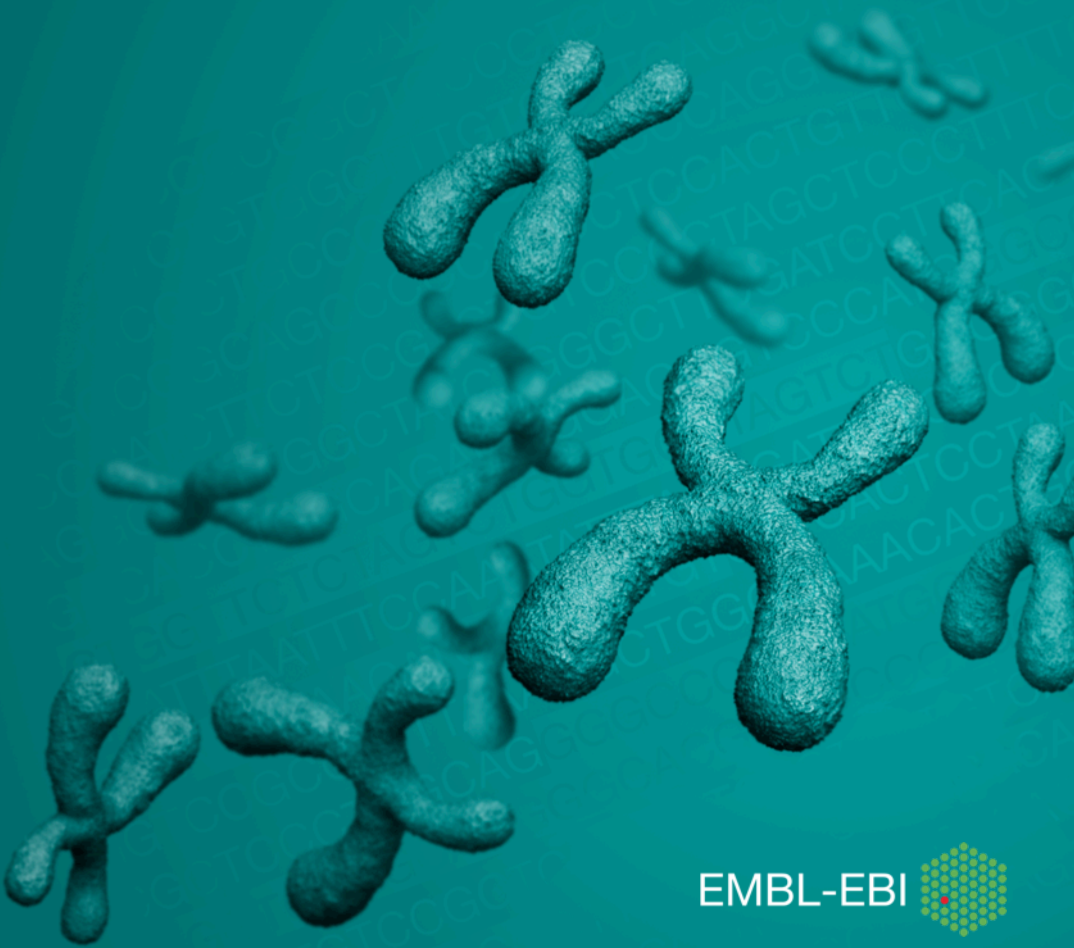
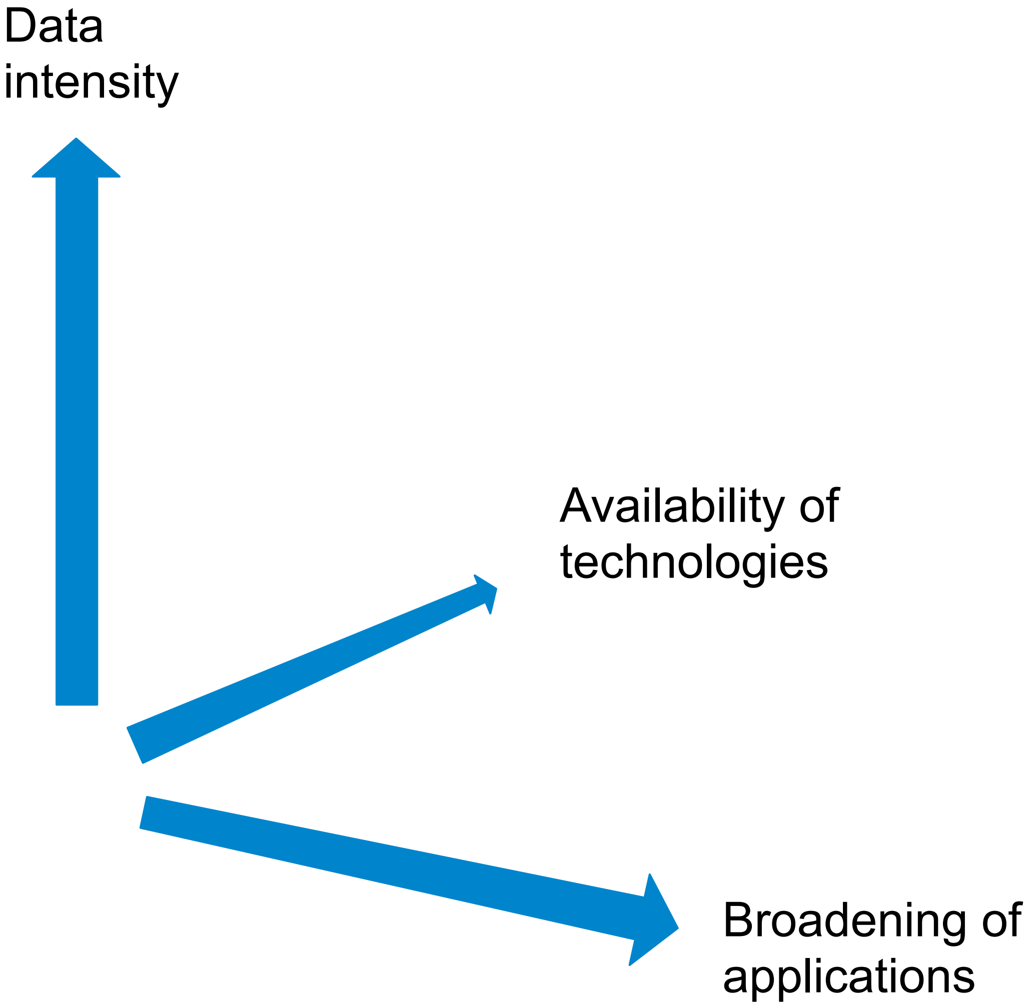
The challenge of NGS data

Guy Cochrane



EMBL-EBI



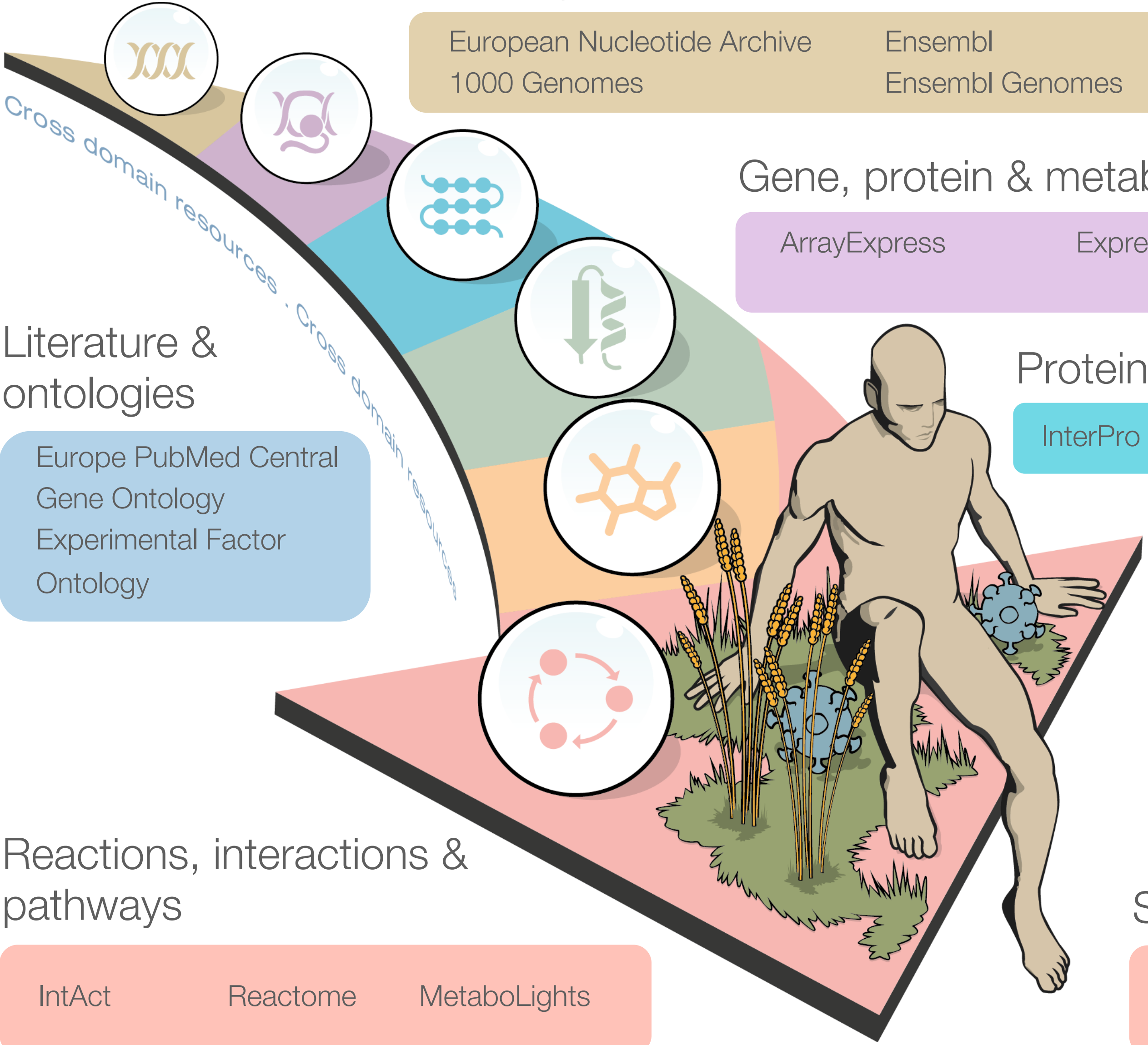
Data intensity

Availability of technologies

Broadening of applications

Context

EMBL European Bioinformatics Institute



Genes, genomes & variation

European Nucleotide Archive
1000 Genomes
Ensembl
Ensembl Genomes
European Genome-phenome Archive
Metagenomics portal

Gene, protein & metabolite expression

ArrayExpress
Expression Atlas
Metabolights
PRIDE

Literature & ontologies

Europe PubMed Central
Gene Ontology
Experimental Factor Ontology

Protein sequences, families & motifs

InterPro
Pfam
UniProt

Molecular structures

Protein Data Bank in Europe
Electron Microscopy Data Bank

Chemical biology

ChEMBL
ChEBI

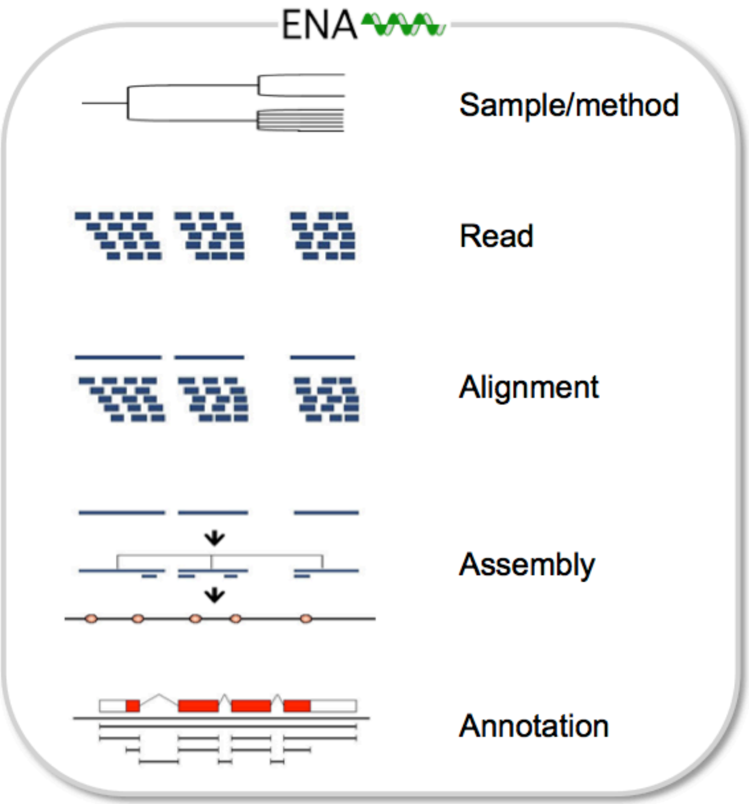
Reactions, interactions & pathways

IntAct
Reactome
MetaboLights

Systems

BioModels
Enzyme Portal
BioSamples

European Nucleotide Archive (ENA)

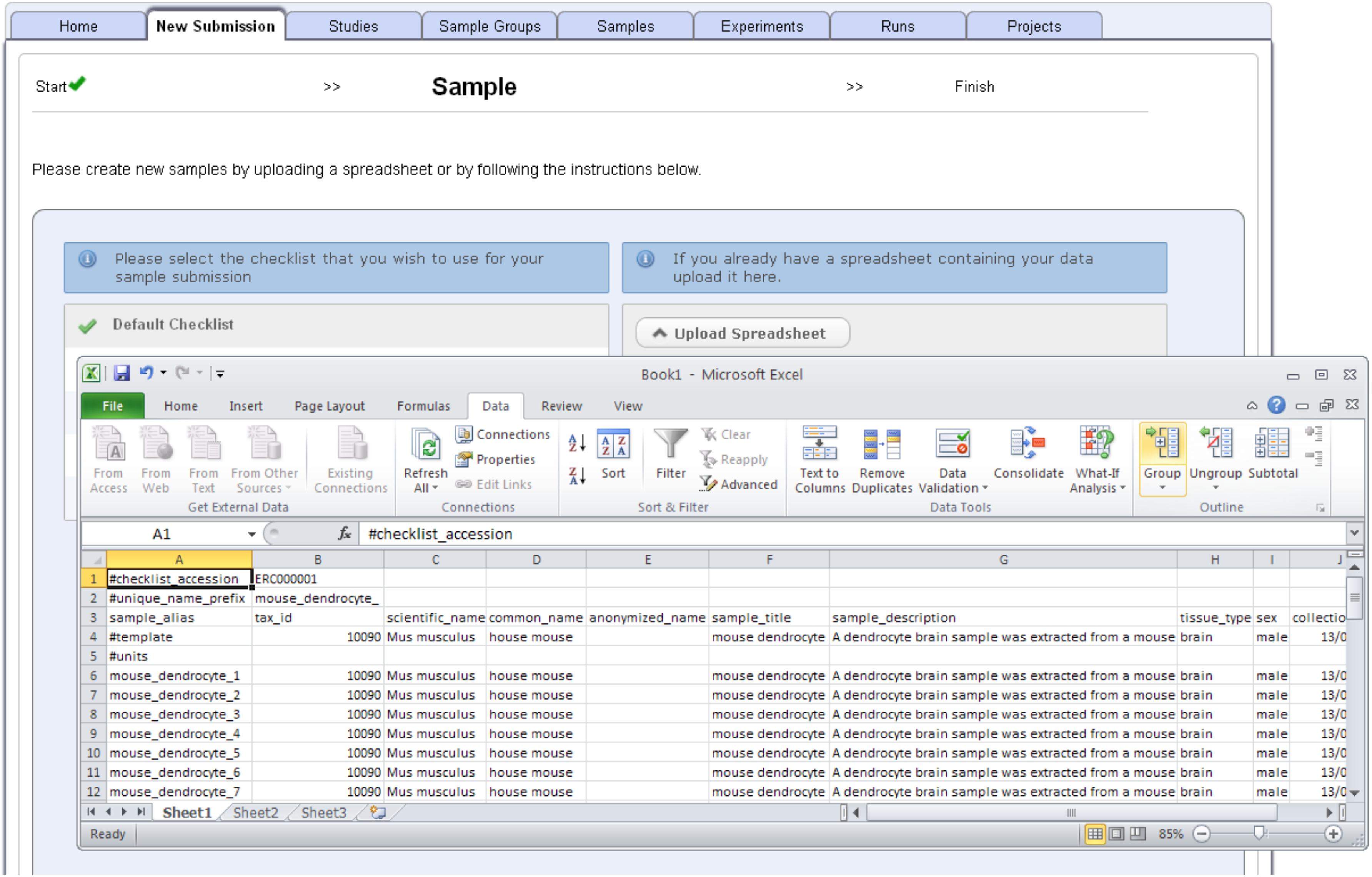


ENA

http://www.ebi.ac.uk/ena/

- A **broad platform** for the management, sharing, integration and dissemination of sequence data
- Established in the early 1980s, extended for **new technologies and applications**
- **Globally comprehensive scientific record** and European node of INSDC
- **Connectivity** with broader EMBL-EBI resources
- Sequence data **foundation**
- **Sustained** within EMBL-EBI under EMBL funding with additional support from EC, UK Research councils, Wellcome Trust, etc.
- **Substantial scale:** 1.3 petabase pairs across >1 million taxa, 2,000-5,000 active data providers, global consumer userbase
- Rich submission, discovery and retrieval **software, tools and services**

Submissions



Please create new samples by uploading a spreadsheet or by following the instructions below.

Please select the checklist that you wish to use for your sample submission

If you already have a spreadsheet containing your data upload it here.

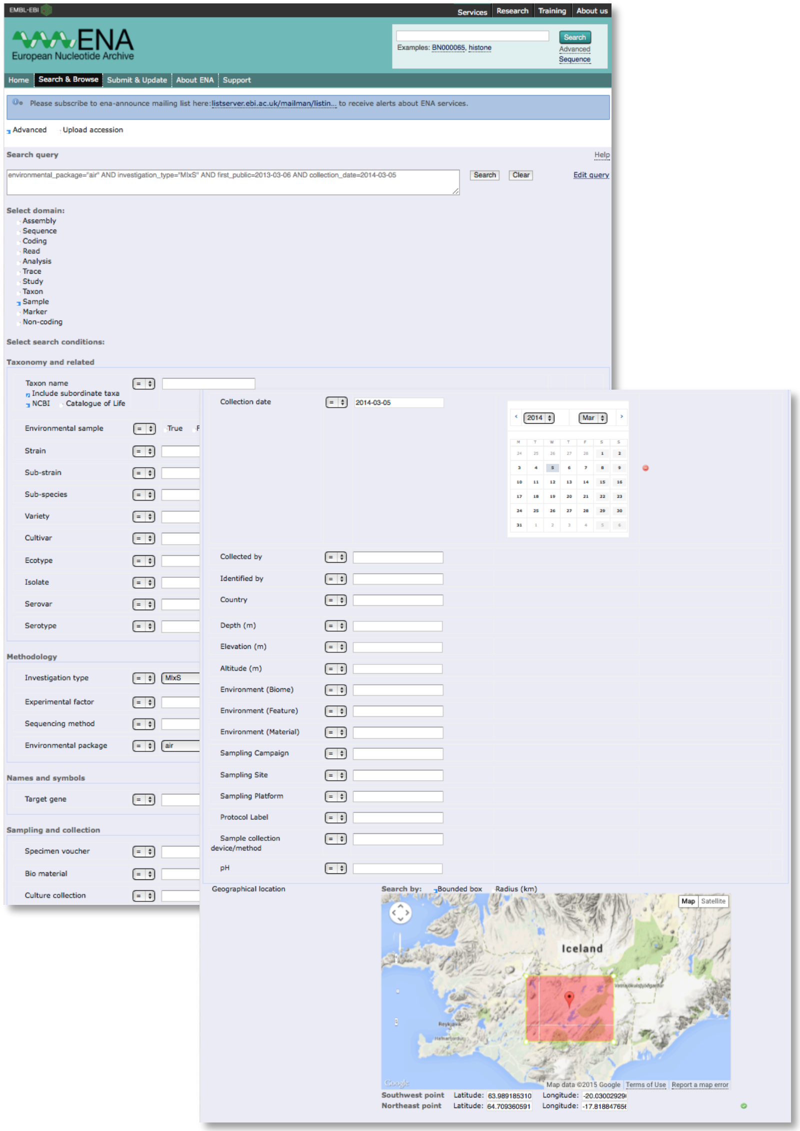
Default Checklist

Upload Spreadsheet

A1: #checklist_accession

	A	B	C	D	E	F	G	H	I	J
1	#checklist_accession	ERC000001								
2	#unique_name_prefix	mouse_dendrocyte_								
3	sample_alias	tax_id	scientific_name	common_name	anonymized_name	sample_title	sample_description	tissue_type	sex	collectio
4	#template	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
5	#units									
6	mouse_dendrocyte_1	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
7	mouse_dendrocyte_2	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
8	mouse_dendrocyte_3	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
9	mouse_dendrocyte_4	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
10	mouse_dendrocyte_5	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
11	mouse_dendrocyte_6	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0
12	mouse_dendrocyte_7	10090	Mus musculus	house mouse		mouse dendrocyte	A dendrocyte brain sample was extracted from a mouse	brain	male	13/0

Data discovery



Please subscribe to ena-announce mailing list here:listserver.ebi.ac.uk/mailman/listin... to receive alerts about ENA services.

Search query

```
environmental_package="air" AND investigation_type="MIxS" AND first_public=2013-03-06 AND collection_date=2014-03-05
```

Select domain:

- Assembly
- Sequence
- Coding
- Read
- Analysis
- Trace
- Study
- Taxon
- Sample
- Marker
- Non-coding

Select search conditions:

Taxonomy and related

Taxon name Include subordinate taxa NCBI Catalogue of Life	=	
Environmental sample	=	True F
Strain	=	
Sub-strain	=	
Sub-species	=	
Variety	=	
Cultivar	=	
Ecotype	=	
Isolate	=	
Serovar	=	
Serotype	=	

Methodology

Investigation type	=	MIxS
Experimental factor	=	
Sequencing method	=	
Environmental package	=	air

Names and symbols

Target gene	=	

Sampling and collection

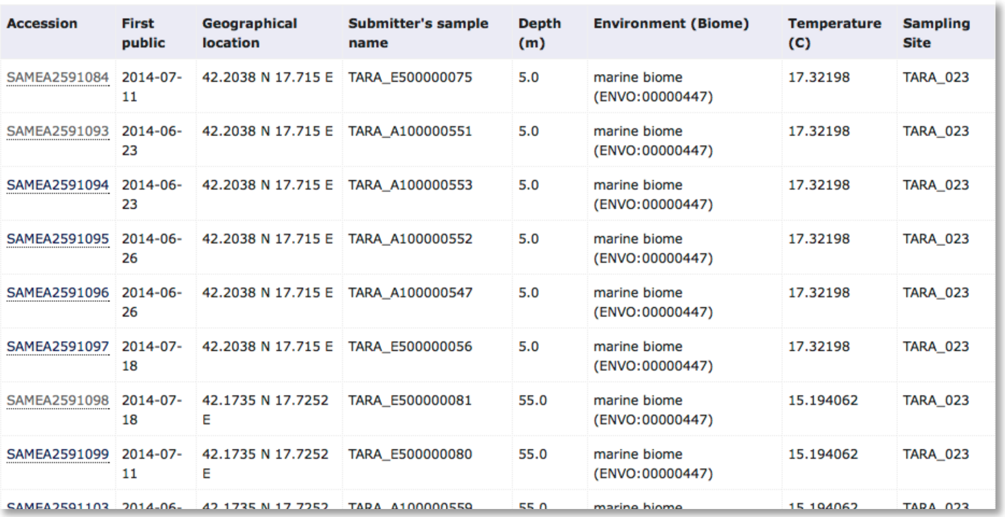
Specimen voucher	=	
Bio material	=	
Culture collection	=	
Collection date	=	2014-03-05
Collected by	=	
Identified by	=	
Country	=	
Depth (m)	=	
Elevation (m)	=	
Altitude (m)	=	
Environment (Biome)	=	
Environment (Feature)	=	
Environment (Material)	=	
Sampling Campaign	=	
Sampling Site	=	
Sampling Platform	=	
Protocol Label	=	
Sample collection device/method	=	
pH	=	

Geographical location

Search by: Bounded box Radius (km)

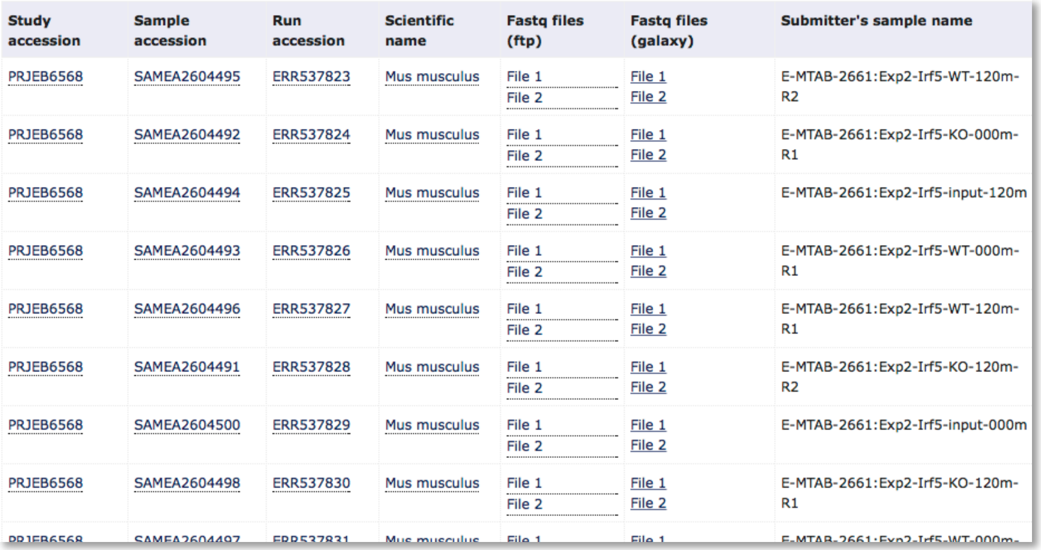
Southwest point	Latitude:	63.989185310	Longitude:	-20.03002929
Northeast point	Latitude:	64.709360591	Longitude:	-17.81884765

temperature>=10 AND temperature<=25 AND geo_box1(42, 17, 43, 18)



Accession	First public	Geographical location	Submitter's sample name	Depth (m)	Environment (Biome)	Temperature (C)	Sampling Site
SAMEA2591084	2014-07-11	42.2038 N 17.715 E	TARA_E500000075	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591093	2014-06-23	42.2038 N 17.715 E	TARA_A100000551	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591094	2014-06-23	42.2038 N 17.715 E	TARA_A100000553	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591095	2014-06-26	42.2038 N 17.715 E	TARA_A100000552	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591096	2014-06-26	42.2038 N 17.715 E	TARA_A100000547	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591097	2014-07-18	42.2038 N 17.715 E	TARA_E500000056	5.0	marine biome (ENVO:00000447)	17.32198	TARA_023
SAMEA2591098	2014-07-18	42.1735 N 17.7252 E	TARA_E500000081	55.0	marine biome (ENVO:00000447)	15.194062	TARA_023
SAMEA2591099	2014-07-11	42.1735 N 17.7252 E	TARA_E500000080	55.0	marine biome (ENVO:00000447)	15.194062	TARA_023
SAMEA2591103	2014-06-	42.1735 N 17.7252	TARA_A100000559	55.0	marine biome	15.194062	TARA_023

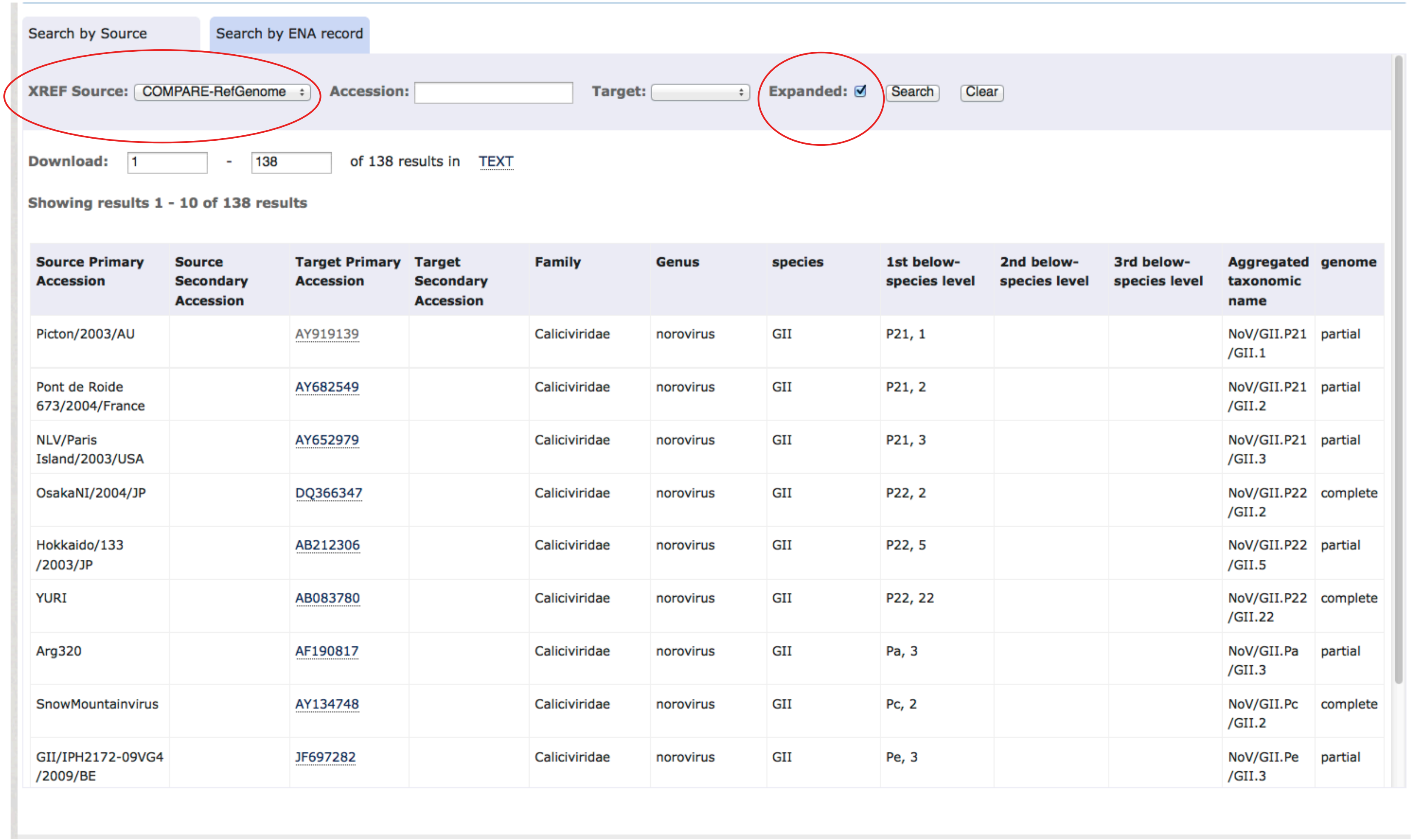
tax_tree(10090) AND library_source="GENOMIC" AND instrument_platform="ILLUMINA" AND library_strategy="ChIP-Seq"



Study accession	Sample accession	Run accession	Scientific name	Fastq files (ftp)	Fastq files (galaxy)	Submitter's sample name
PRJEB6568	SAMEA2604495	ERR537823	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-WT-120m-R2
PRJEB6568	SAMEA2604492	ERR537824	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-KO-000m-R1
PRJEB6568	SAMEA2604494	ERR537825	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-input-120m
PRJEB6568	SAMEA2604493	ERR537826	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-WT-000m-R1
PRJEB6568	SAMEA2604496	ERR537827	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-WT-120m-R1
PRJEB6568	SAMEA2604491	ERR537828	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-KO-120m-R2
PRJEB6568	SAMEA2604500	ERR537829	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-input-000m
PRJEB6568	SAMEA2604498	ERR537830	Mus musculus	File 1 File 2	File 1 File 2	E-MTAB-2661:Exp2-Irf5-KO-120m-R1
PRJEB6568	SAMEA2604497	ERR537831	Mus musculus	File 1	File 1	E-MTAB-2661:Exp2-Irf5-WT-000m-

Cross-references and tagging

GUI: http://www.ebi.ac.uk/ena/data/xref/search



XREF Source: COMPARE-RefGenome

Accession:

Target:

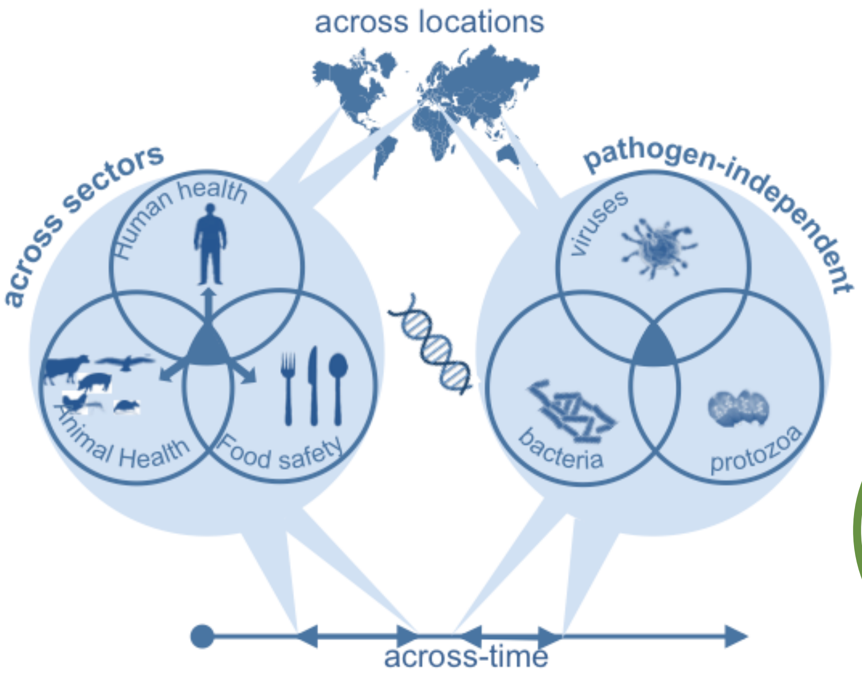
Expanded: ☑

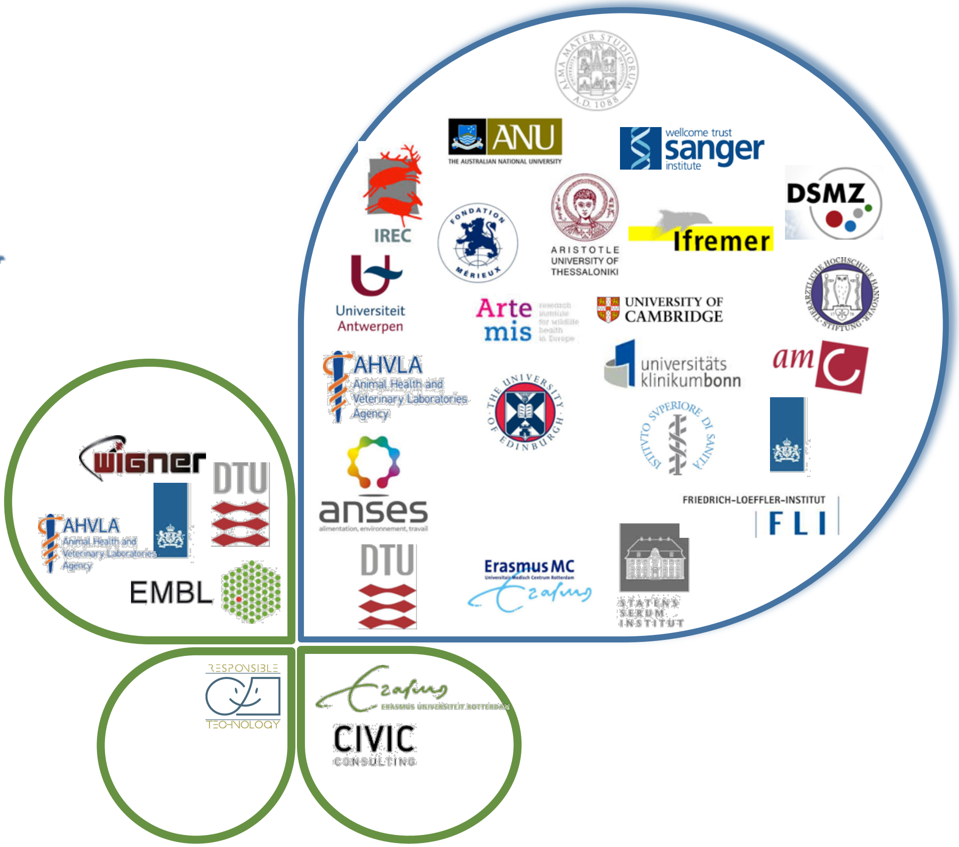
Download: 1 - 138 of 138 results in TEXT

Showing results 1 - 10 of 138 results

Source Primary Accession	Source Secondary Accession	Target Primary Accession	Target Secondary Accession	Family	Genus	species	1st below-species level	2nd below-species level	3rd below-species level	Aggregated taxonomic name	genome
Picton/2003/AU		AY919139		Caliciviridae	norovirus	GII	P21, 1			NoV/GII.P21/GII.1	partial
Pont de Roide 673/2004/France		AY682549		Caliciviridae	norovirus	GII	P21, 2			NoV/GII.P21/GII.2	partial
NLV/Paris Island/2003/USA		AY652979		Caliciviridae	norovirus	GII	P21, 3			NoV/GII.P21/GII.3	partial
OsakaNI/2004/JP		DQ366347		Caliciviridae	norovirus	GII	P22, 2			NoV/GII.P22/GII.2	complete
Hokkaido/133/2003/JP		AB212306		Caliciviridae	norovirus	GII	P22, 5			NoV/GII.P22/GII.5	partial
YURI		AB083780		Caliciviridae	norovirus	GII	P22, 22			NoV/GII.P22/GII.22	complete
Arg320		AF190817		Caliciviridae	norovirus	GII	Pa, 3			NoV/GII.Pa/GII.3	partial
SnowMountainvirus		AY134748		Caliciviridae	norovirus	GII	Pc, 2			NoV/GII.Pc/GII.2	complete
GII/IPH2172-09VG4/2009/BE		JF697282		Caliciviridae	norovirus	GII	Pe, 3			NoV/GII.Pe/GII.3	partial

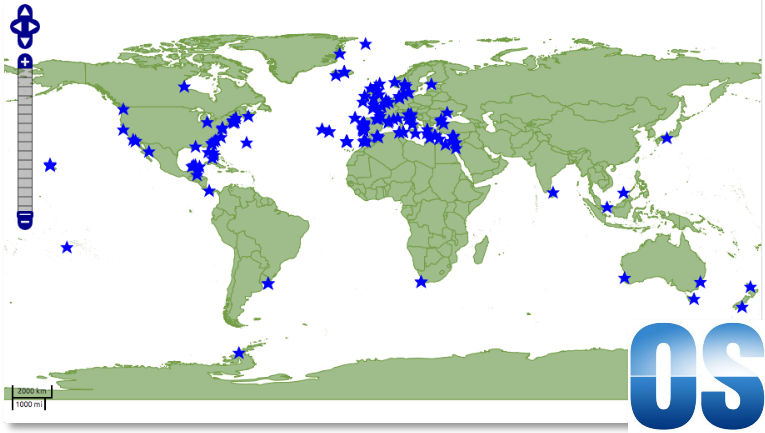
A global platform for the sequence-based rapid identification of pathogens



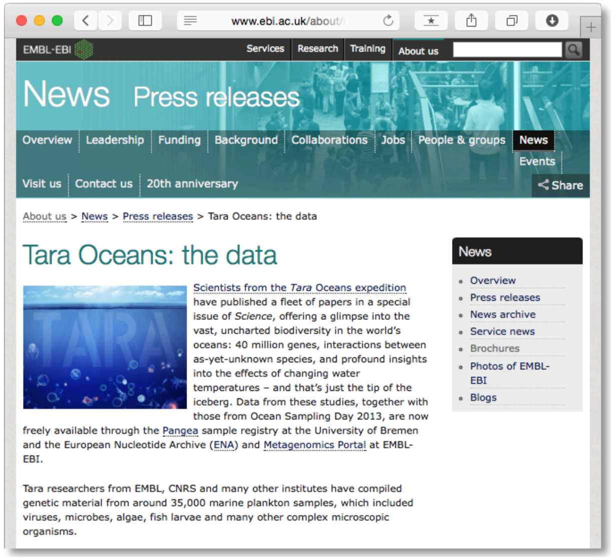

ALMA MATER STUDIORUM A.D. 1088
ANU THE AUSTRALIAN NATIONAL UNIVERSITY
wellcome trust sanger institute
DSMZ
IREC
FONDATION MERIEUX
ARISTOTLE UNIVERSITY OF THESSALONIKI
Ifremer
Universiteit Antwerpen
Artemis
UNIVERSITY OF CAMBRIDGE
AHVLA Animal Health and Veterinary Laboratories Agency
THE UNIVERSITY OF EDINBURGH
universitäts klinikumbonn
amc
ISTITUTO SUPERIORE DI SANITÀ
anses
DTU
Erasmus MC
STATENS SERUM INSTITUT
FRIEDRICH-LOEFFLER-INSTITUT
F L I
wigner
EMBL
RESPONSIBLE TECHNOLOGY
Erasmus Universiteit Rotterdam
CIVIC CONSULTING

COMPARE: the enabling system for rapid identification, containment and mitigation of emerging infectious diseases and foodborne outbreaks by generation and comparison of genomic information on samples and pathogens across sectors, time and locations, with additional contextual data.

Ocean Samping Day and Tara Oceans



OSD
Ocean Sampling Day



About us > News > Press releases > Tara Oceans: the data

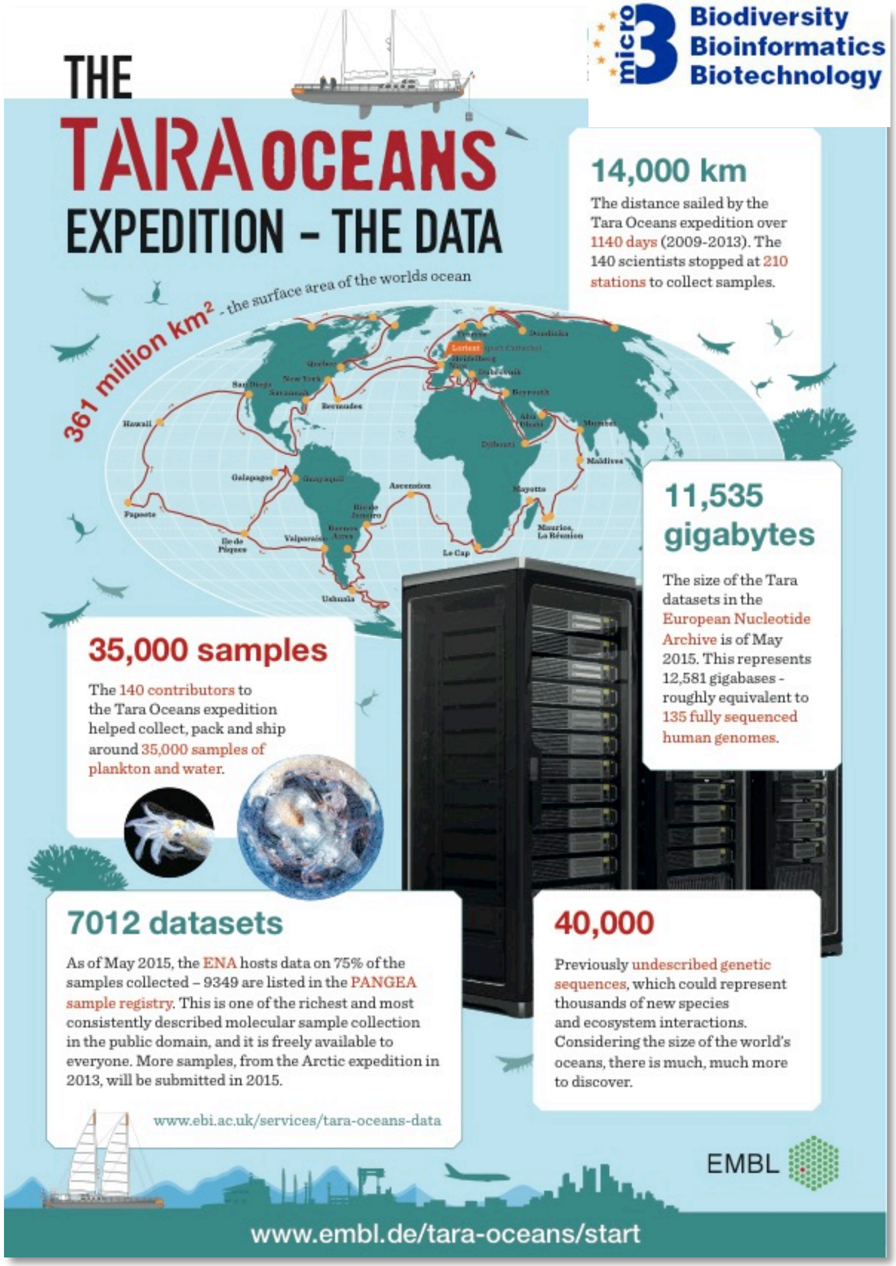
Tara Oceans: the data

Scientists from the *Tara* Oceans expedition have published a fleet of papers in a special issue of *Science*, offering a glimpse into the vast, uncharted biodiversity in the world's oceans: 40 million genes, interactions between as-yet-unknown species, and profound insights into the effects of changing water temperatures – and that's just the tip of the iceberg. Data from these studies, together with those from Ocean Sampling Day 2013, are now freely available through the Pangea sample registry at the University of Bremen and the European Nucleotide Archive (ENA) and Metagenomics Portal at EMBL-EBI.

Tara researchers from EMBL, CNRS and many other institutes have compiled genetic material from around 35,000 marine plankton samples, which included viruses, microbes, algae, fish larvae and many other complex microscopic organisms.

News

- Overview
- Press releases
- News archive
- Service news
- Brochures
- Photos of EMBL-EBI
- Blogs



THE TARA OCEANS EXPEDITION – THE DATA

14,000 km

The distance sailed by the Tara Oceans expedition over 1140 days (2009-2013). The 140 scientists stopped at 210 stations to collect samples.

11,535 gigabytes

The size of the Tara datasets in the European Nucleotide Archive is of May 2015. This represents 12,581 gigabases - roughly equivalent to 135 fully sequenced human genomes.

35,000 samples

The 140 contributors to the Tara Oceans expedition helped collect, pack and ship around 35,000 samples of plankton and water.

7012 datasets

As of May 2015, the ENA hosts data on 75% of the samples collected – 9349 are listed in the PANGEA sample registry. This is one of the richest and most consistently described molecular sample collection in the public domain, and it is freely available to everyone. More samples, from the Arctic expedition in 2013, will be submitted in 2015.

www.ebi.ac.uk/services/tara-oceans-data

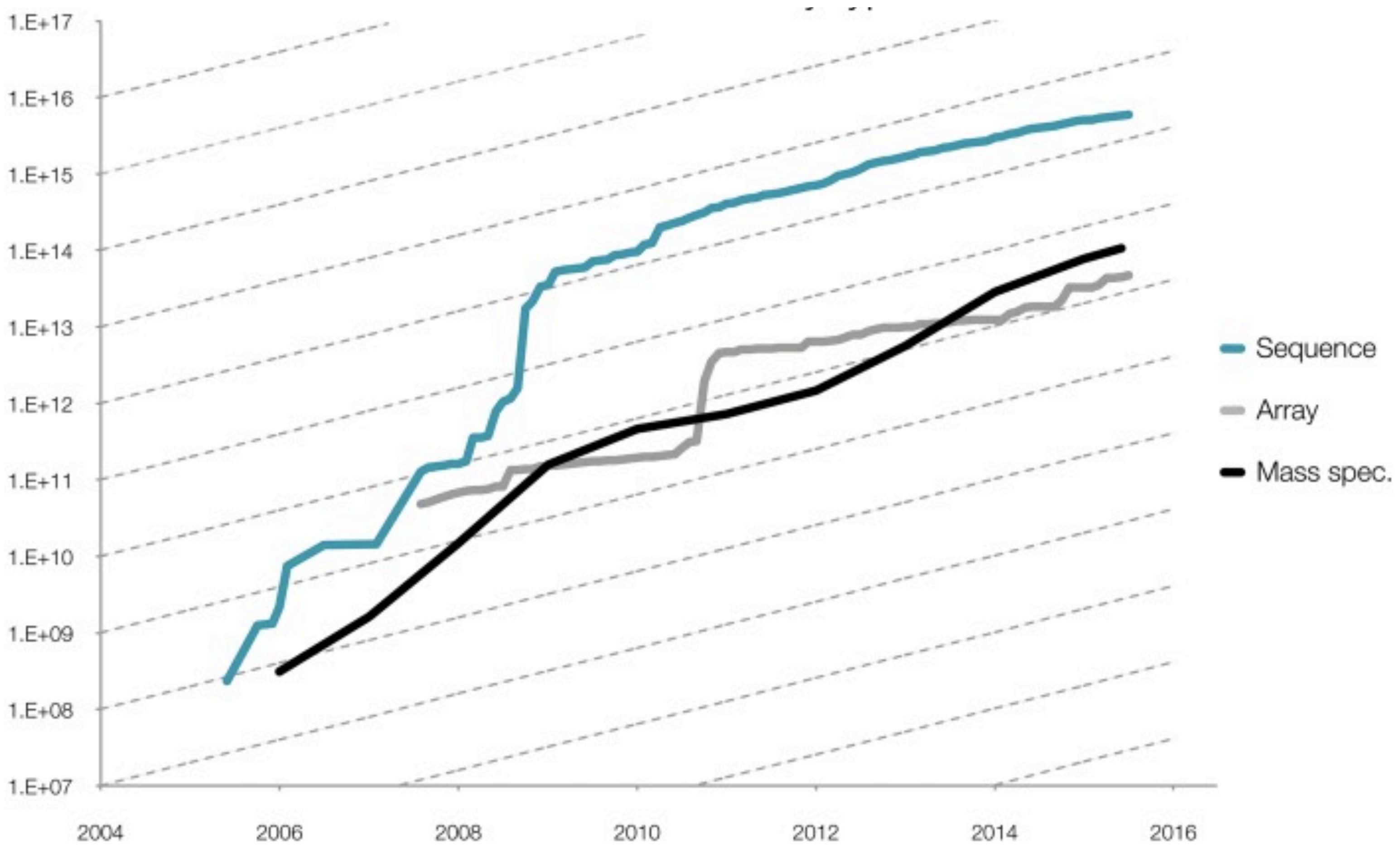
40,000

Previously undescribed genetic sequences, which could represent thousands of new species and ecosystem interactions. Considering the size of the world's oceans, there is much, much more to discover.

www.embl.de/tara-oceans/start

Technology

Data accumulation

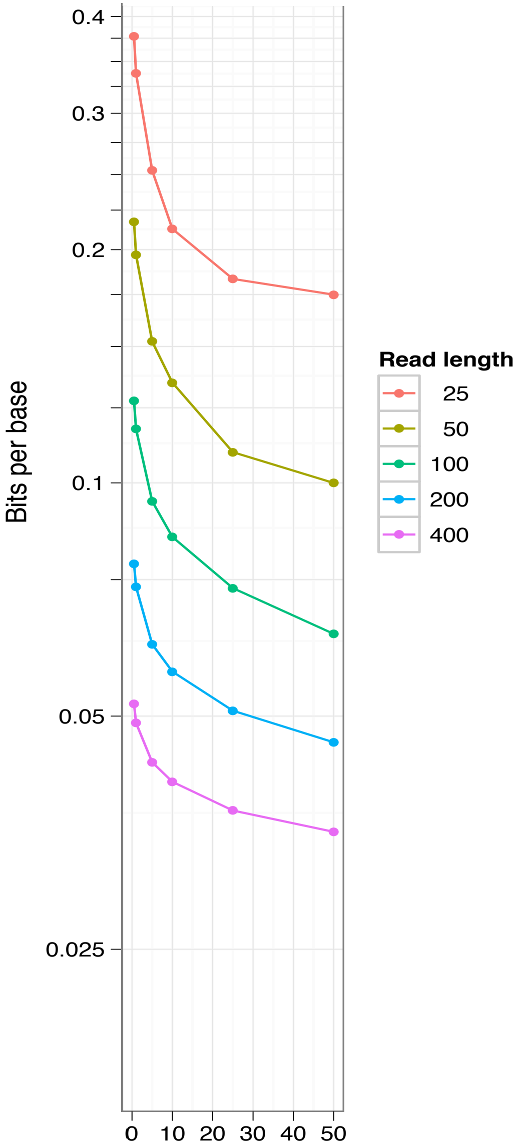


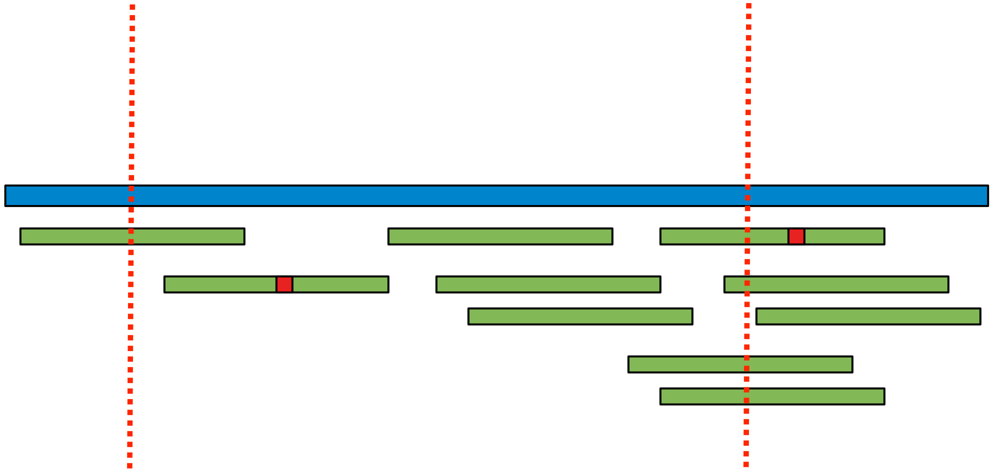
Cook CE, Bergman MT, Finn RD, Cochrane G, Birney E, Apweiler R. The European Bioinformatics Institute in 2016: Data growth and integration. Nucleic Acids Res. 2016 Jan;44(D1) D20-6. doi:10.1093/nar/gkv1352. PMID: 26673705; PMCID: PMC4702932.

Starting point

```
TGAGCTCTAAGTACC
329183050298757
```

Sequence compression

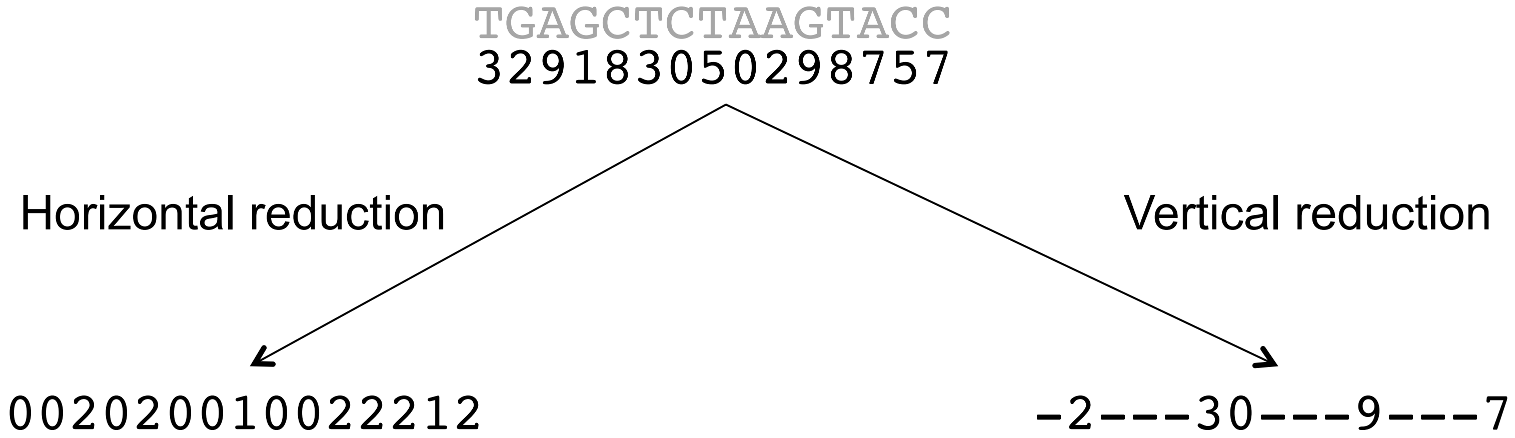




- Encoding of read starts and differences
- 3.5x–100x compression over existing formats
- Scales favourably with increasing read length and density

Fritz, M.H. Leinonen, R., et al. (2011) Efficient storage of high throughput DNA sequencing data using reference-based compression. *Genome Res.* 21 (5), 734-40.

Quality compression



TGAGCTCTAAGTACC

329183050298757

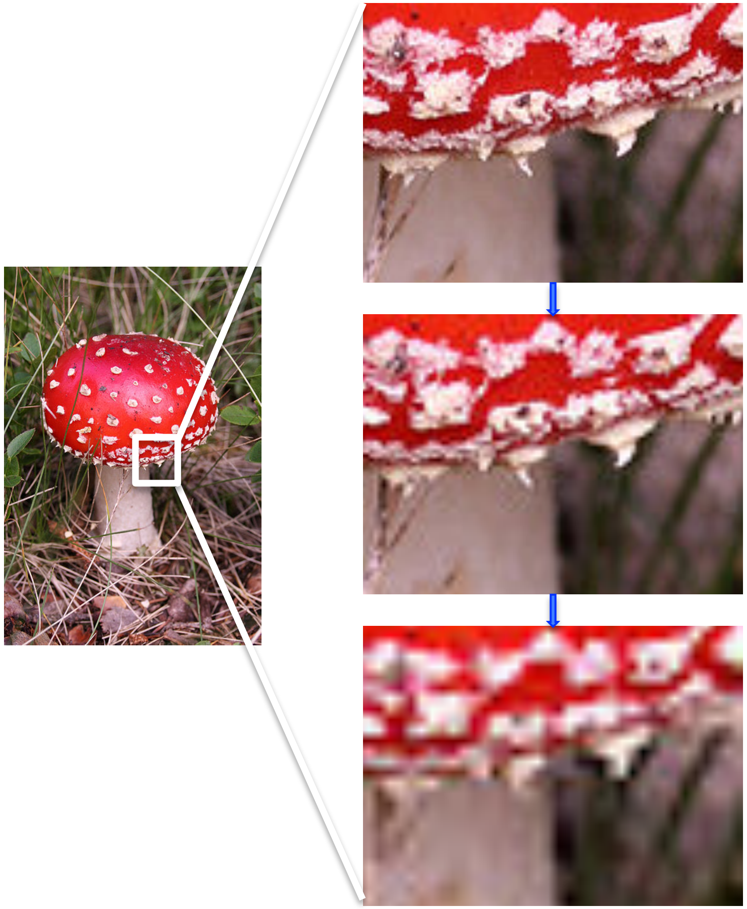
Horizontal reduction

Vertical reduction

```
002020010022212
```

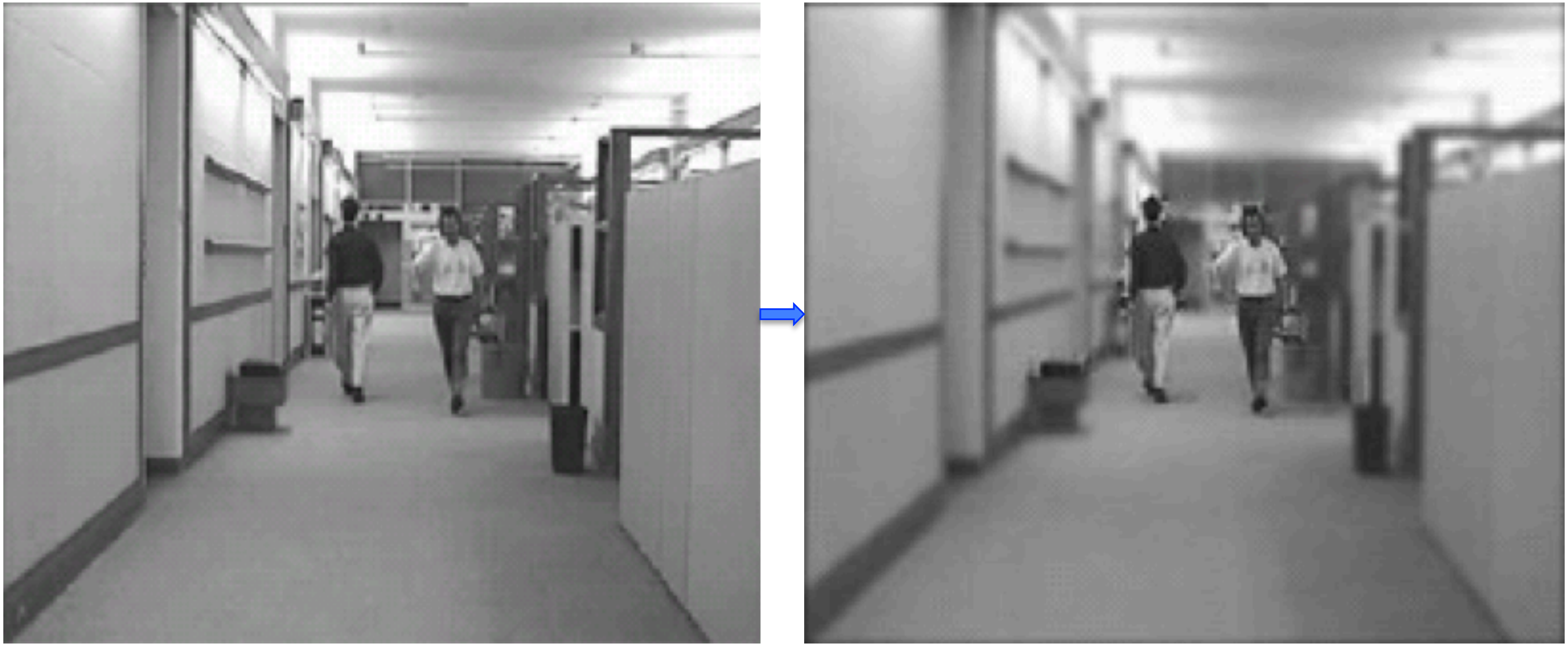
```
-2---30---9---7
```

Quality compression: simple, horizontal reduction



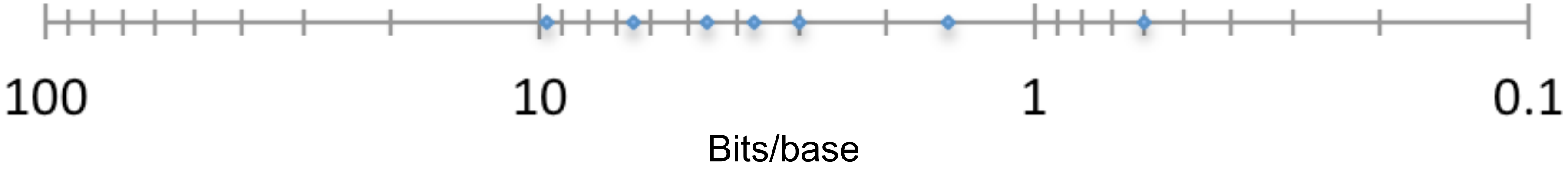
Photograph from MichaelMaggs, http://en.wikipedia.org/wiki/File:Amanita_muscaria_(fly_agaric).JPG

Models for data reduction

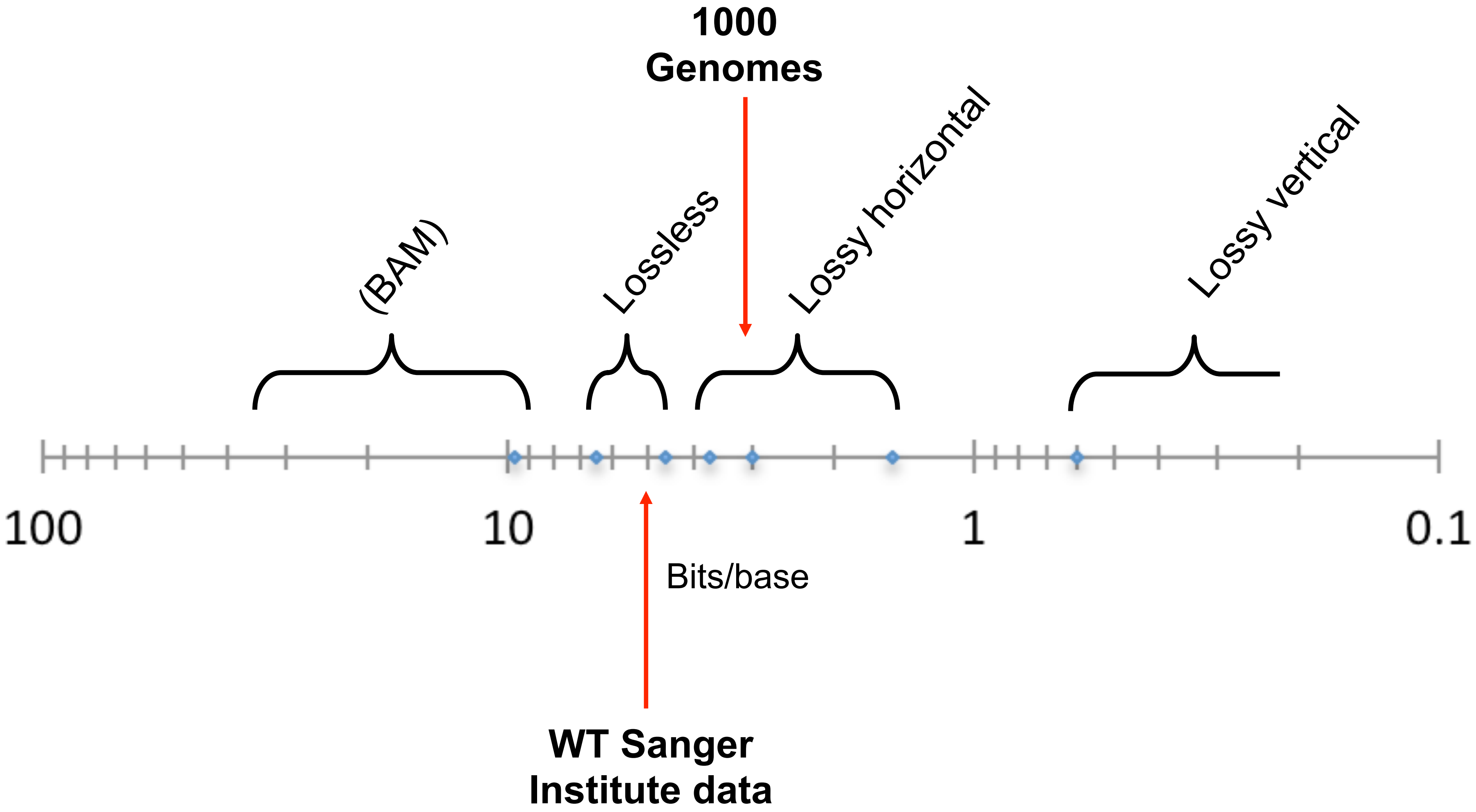


Jong-Seok Lee et al. (2009), http://mmspg.epfl.ch/files/content/sites/mmspl/files/shared/lee_icme.pdf

CRAM performance

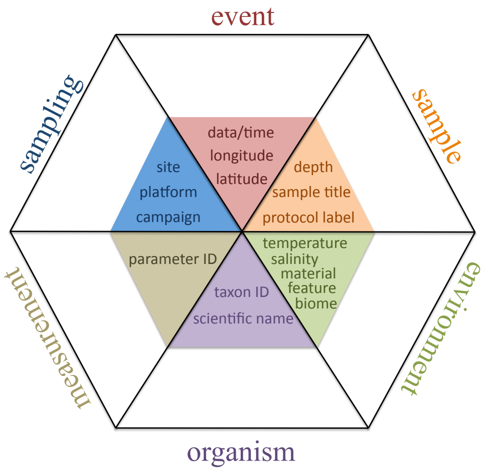


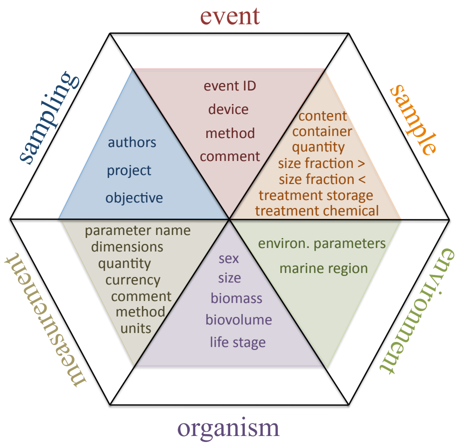
CRAM performance

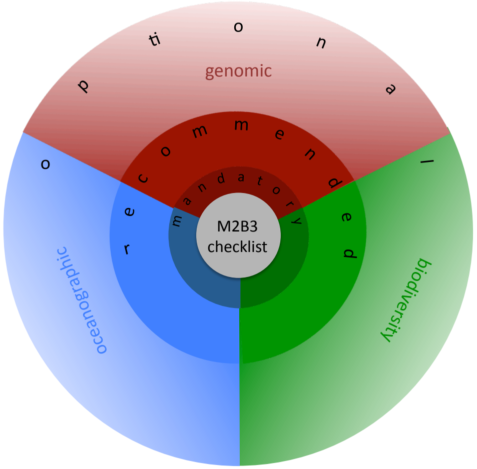


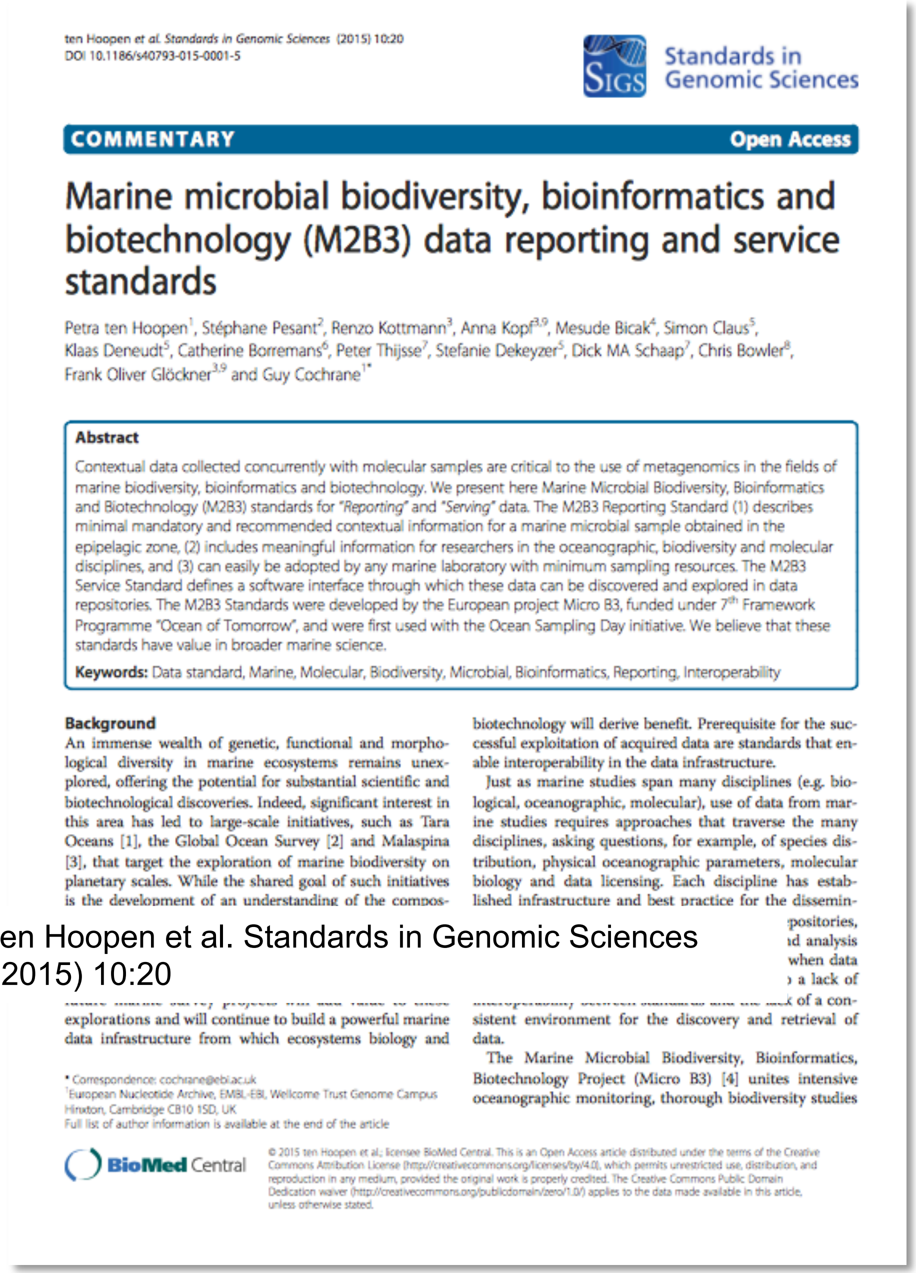
Human aspects

Standards









ten Hoopen *et al. Standards in Genomic Sciences* (2015) 10:20
DOI 10.1186/s40793-015-0001-5

Standards in Genomic Sciences

COMMENTARY Open Access

Marine microbial biodiversity, bioinformatics and biotechnology (M2B3) data reporting and service standards

Petra ten Hoopen[1], Stéphane Pesant[2], Renzo Kottmann[3], Anna Kopf[3,9], Mesude Bicak[4], Simon Claus[5], Klaas Deneudt[5], Catherine Borremans[6], Peter Thijsse[7], Stefanie Dekeyzer[5], Dick MA Schaap[7], Chris Bowler[8], Frank Oliver Glöckner[3,9] and Guy Cochrane[1*]

Abstract

Contextual data collected concurrently with molecular samples are critical to the use of metagenomics in the fields of marine biodiversity, bioinformatics and biotechnology. We present here Marine Microbial Biodiversity, Bioinformatics and Biotechnology (M2B3) standards for *"Reporting"* and *"Serving"* data. The M2B3 Reporting Standard (1) describes minimal mandatory and recommended contextual information for a marine microbial sample obtained in the epipelagic zone, (2) includes meaningful information for researchers in the oceanographic, biodiversity and molecular disciplines, and (3) can easily be adopted by any marine laboratory with minimum sampling resources. The M2B3 Service Standard defines a software interface through which these data can be discovered and explored in data repositories. The M2B3 Standards were developed by the European project Micro B3, funded under 7th Framework Programme "Ocean of Tomorrow", and were first used with the Ocean Sampling Day initiative. We believe that these standards have value in broader marine science.

Keywords: Data standard, Marine, Molecular, Biodiversity, Microbial, Bioinformatics, Reporting, Interoperability

Background

An immense wealth of genetic, functional and morphological diversity in marine ecosystems remains unexplored, offering the potential for substantial scientific and biotechnological discoveries. Indeed, significant interest in this area has led to large-scale initiatives, such as Tara Oceans [1], the Global Ocean Survey [2] and Malaspina [3], that target the exploration of marine biodiversity on planetary scales. While the shared goal of such initiatives is the development of an understanding of the compos-

explorations and will continue to build a powerful marine data infrastructure from which ecosystems biology and biotechnology will derive benefit. Prerequisite for the successful exploitation of acquired data are standards that enable interoperability in the data infrastructure.

Just as marine studies span many disciplines (e.g. biological, oceanographic, molecular), use of data from marine studies requires approaches that traverse the many disciplines, asking questions, for example, of species distribution, physical oceanographic parameters, molecular biology and data licensing. Each discipline has established infrastructure and best practice for the dissemin-

positories,
nd analysis
when data
a lack of
of a consistent environment for the discovery and retrieval of data.

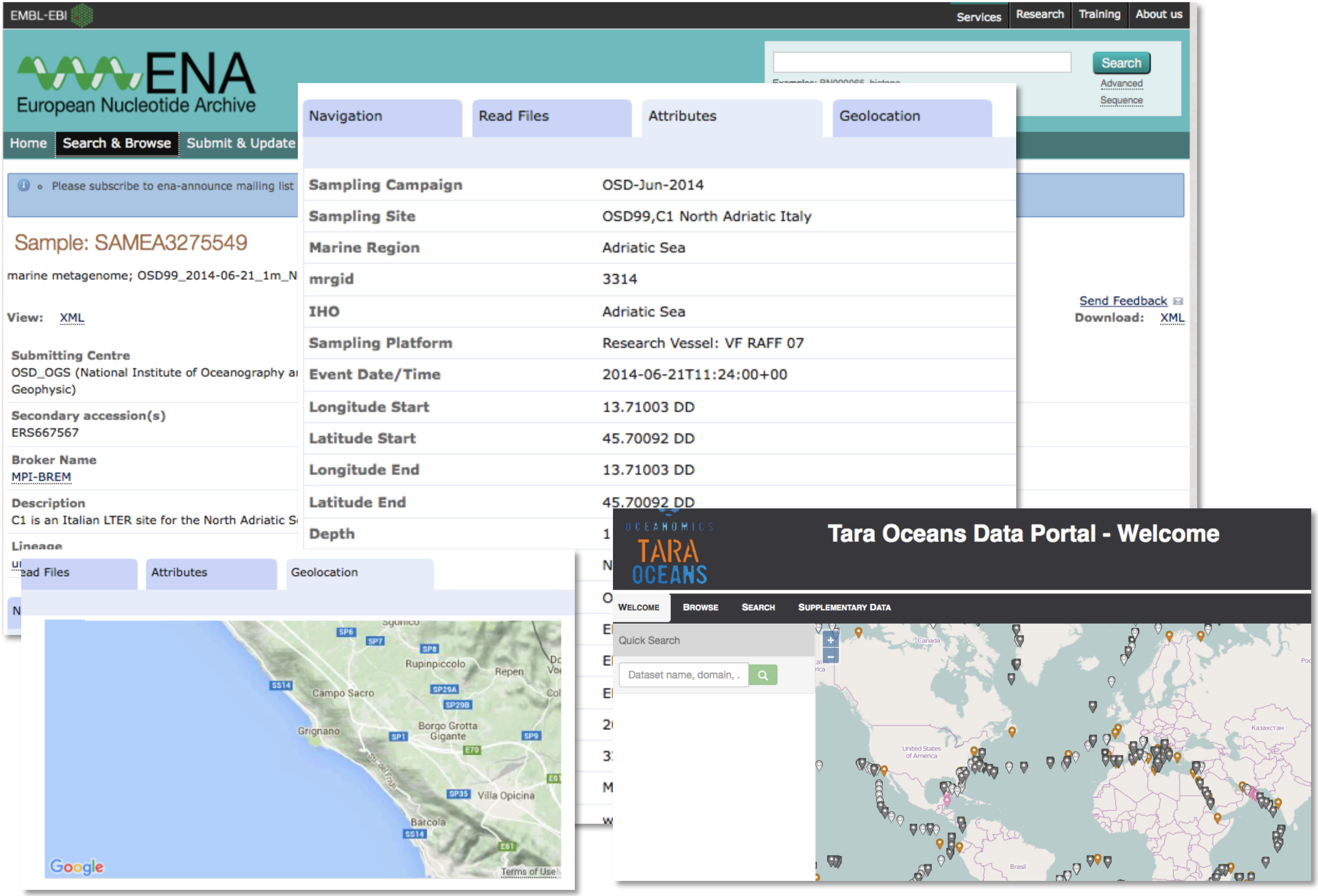
The Marine Microbial Biodiversity, Bioinformatics, Biotechnology Project (Micro B3) [4] unites intensive oceanographic monitoring, thorough biodiversity studies

[*] Correspondence: cochrane@ebi.ac.uk
[1] European Nucleotide Archive, EMBL-EBI, Wellcome Trust Genome Campus Hinxton, Cambridge CB10 1SD, UK
Full list of author information is available at the end of the article

BioMed Central

ten Hoopen et al. Standards in Genomic Sciences (2015) 10:20

micro B3 Biodiversity Bioinformatics Biotechnology



Please subscribe to ena-announce mailing list

Sample: SAMEA3275549

marine metagenome; OSD99_2014-06-21_1m_N

View: XML

Submitting Centre
OSD_OGS (National Institute of Oceanography a
Geophysic)

Secondary accession(s)
ERS667567

Broker Name
MPI-BREM

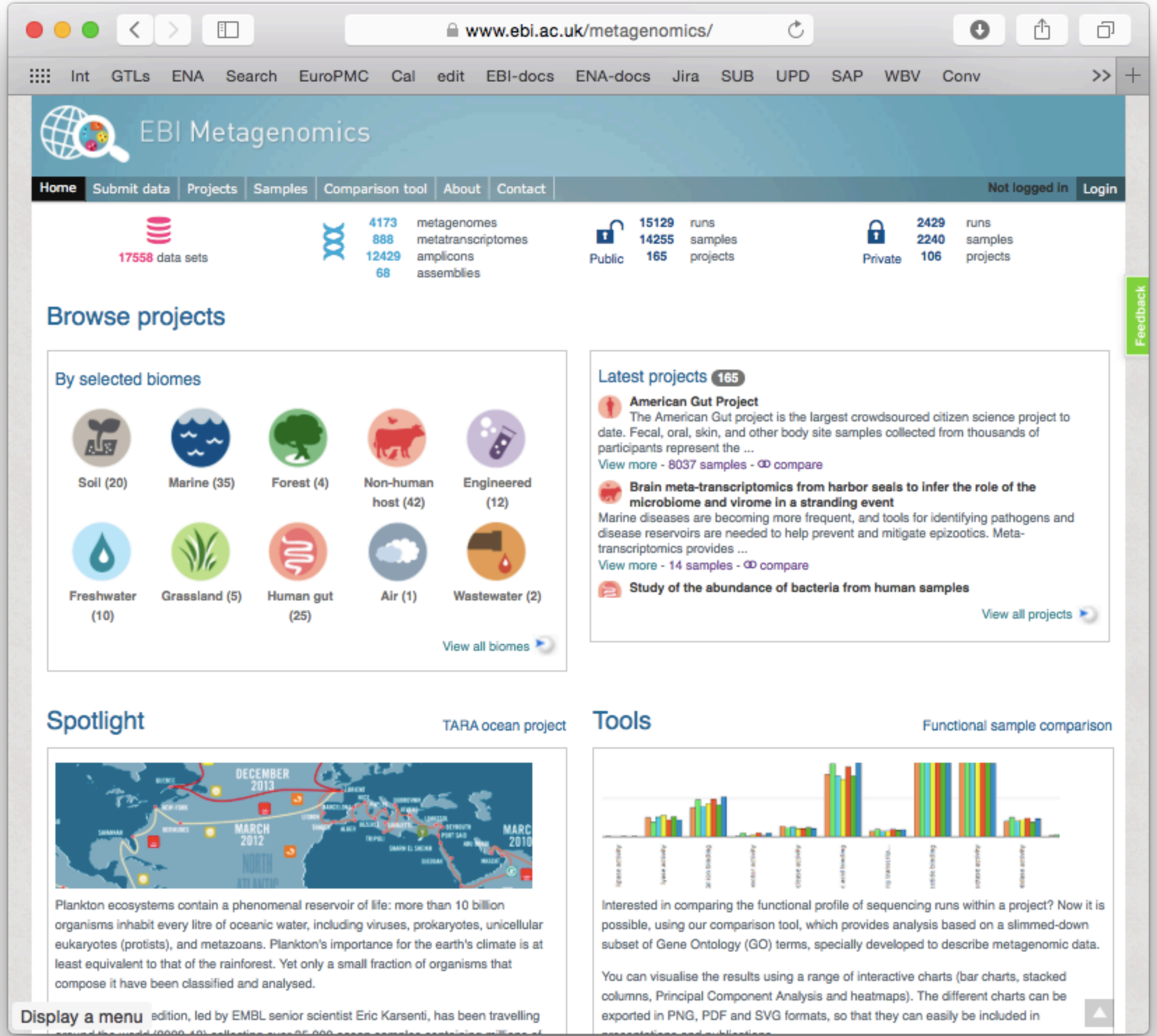
Description
C1 is an Italian LTER site for the North Adriatic S

Lineage

Sampling Campaign	OSD-Jun-2014
Sampling Site	OSD99,C1 North Adriatic Italy
Marine Region	Adriatic Sea
mrgid	3314
IHO	Adriatic Sea
Sampling Platform	Research Vessel: VF RAFF 07
Event Date/Time	2014-06-21T11:24:00+00
Longitude Start	13.71003 DD
Latitude Start	45.70092 DD
Longitude End	13.71003 DD
Latitude End	45.70092 DD
Depth	1

Tara Oceans Data Portal - Welcome

Metagenomics



EBI Metagenomics

17558 data sets

4173 metagenomes
888 metatranscriptomes
12429 amplicons
68 assemblies

Public
15129 runs
14255 samples
165 projects

Private
2429 runs
2240 samples
106 projects

Browse projects

By selected biomes

Soil (20)
Marine (35)
Forest (4)
Non-human host (42)
Engineered (12)
Freshwater (10)
Grassland (5)
Human gut (25)
Air (1)
Wastewater (2)

View all biomes

Latest projects 165

American Gut Project
The American Gut project is the largest crowdsourced citizen science project to date. Fecal, oral, skin, and other body site samples collected from thousands of participants represent the ...
View more - 8037 samples - compare

Brain meta-transcriptomics from harbor seals to infer the role of the microbiome and virome in a stranding event
Marine diseases are becoming more frequent, and tools for identifying pathogens and disease reservoirs are needed to help prevent and mitigate epizootics. Meta-transcriptomics provides ...
View more - 14 samples - compare

Study of the abundance of bacteria from human samples

View all projects

Spotlight

TARA ocean project

Plankton ecosystems contain a phenomenal reservoir of life: more than 10 billion organisms inhabit every litre of oceanic water, including viruses, prokaryotes, unicellular eukaryotes (protists), and metazoans. Plankton's importance for the earth's climate is at least equivalent to that of the rainforest. Yet only a small fraction of organisms that compose it have been classified and analysed.

edition, led by EMBL senior scientist Eric Karsenti, has been travelling

Tools

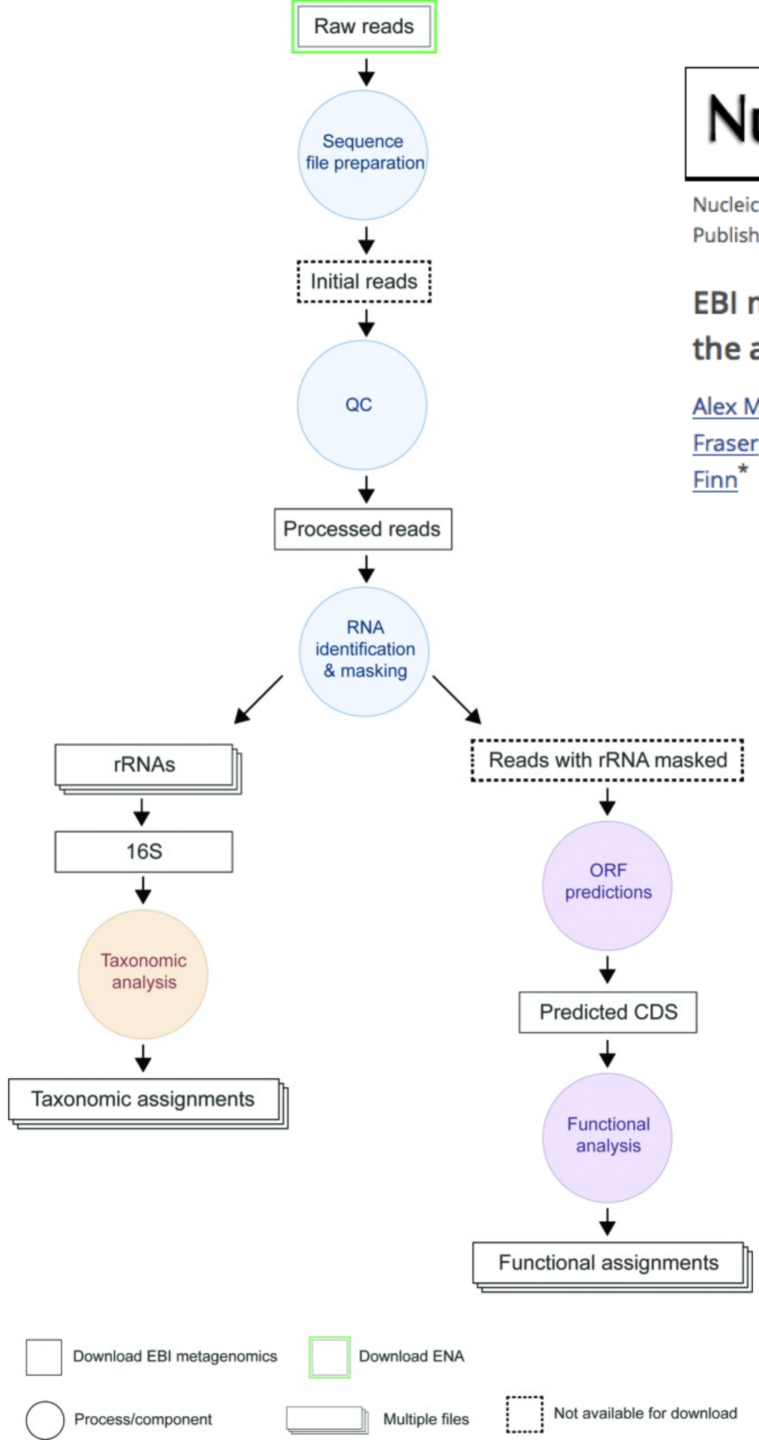
Functional sample comparison

Interested in comparing the functional profile of sequencing runs within a project? Now it is possible, using our comparison tool, which provides analysis based on a slimmed-down subset of Gene Ontology (GO) terms, specially developed to describe metagenomic data.

You can visualise the results using a range of interactive charts (bar charts, stacked columns, Principal Component Analysis and heatmaps). The different charts can be exported in PNG, PDF and SVG formats, so that they can easily be included in

https://www.ebi.ac.uk/metagenomics/

EMBL-EBI

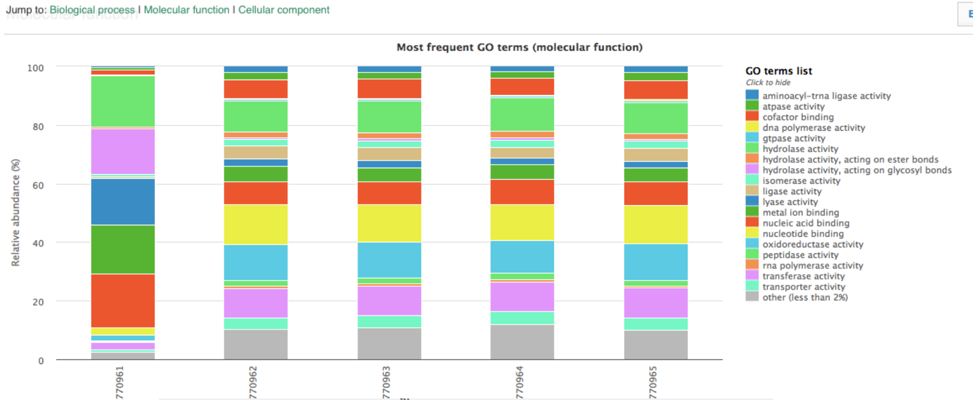


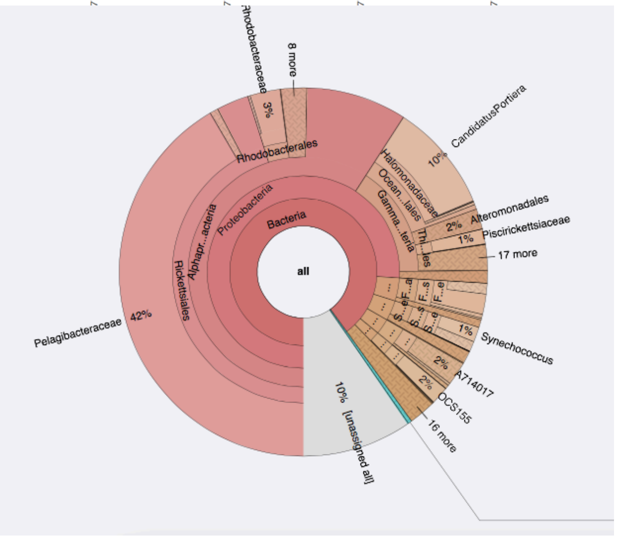
Nucleic Acids Research

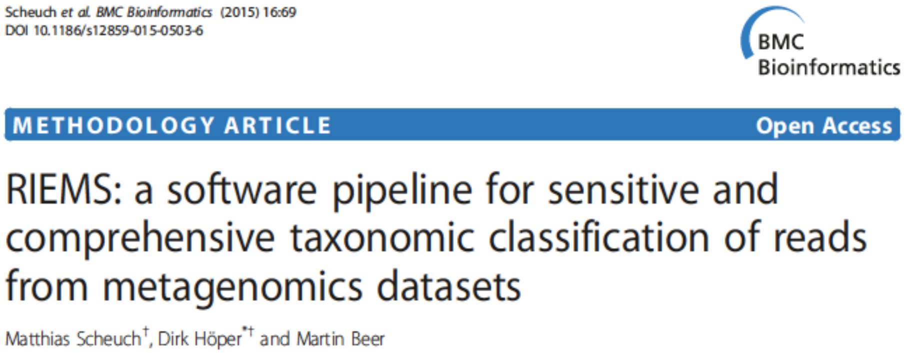
Nucleic Acids Res. 2016 January 4; 44(Database issue): D595–D603. PMCID: PMC4702853
Published online 2015 November 17. doi: 10.1093/nar/gkv1195

EBI metagenomics in 2016 - an expanding and evolving resource for the analysis and archiving of metagenomic data

Alex Mitchell, Francois Bucchini, Guy Cochrane, Hubert Denise, Petra ten Hoopen, Matthew Fraser, Sebastien Pesseat, Simon Potter, Maxim Scheremetjew, Peter Sterk, and Robert D. Finn*







Scheuch *et al. BMC Bioinformatics* (2015) 16:69
DOI 10.1186/s12859-015-0503-6

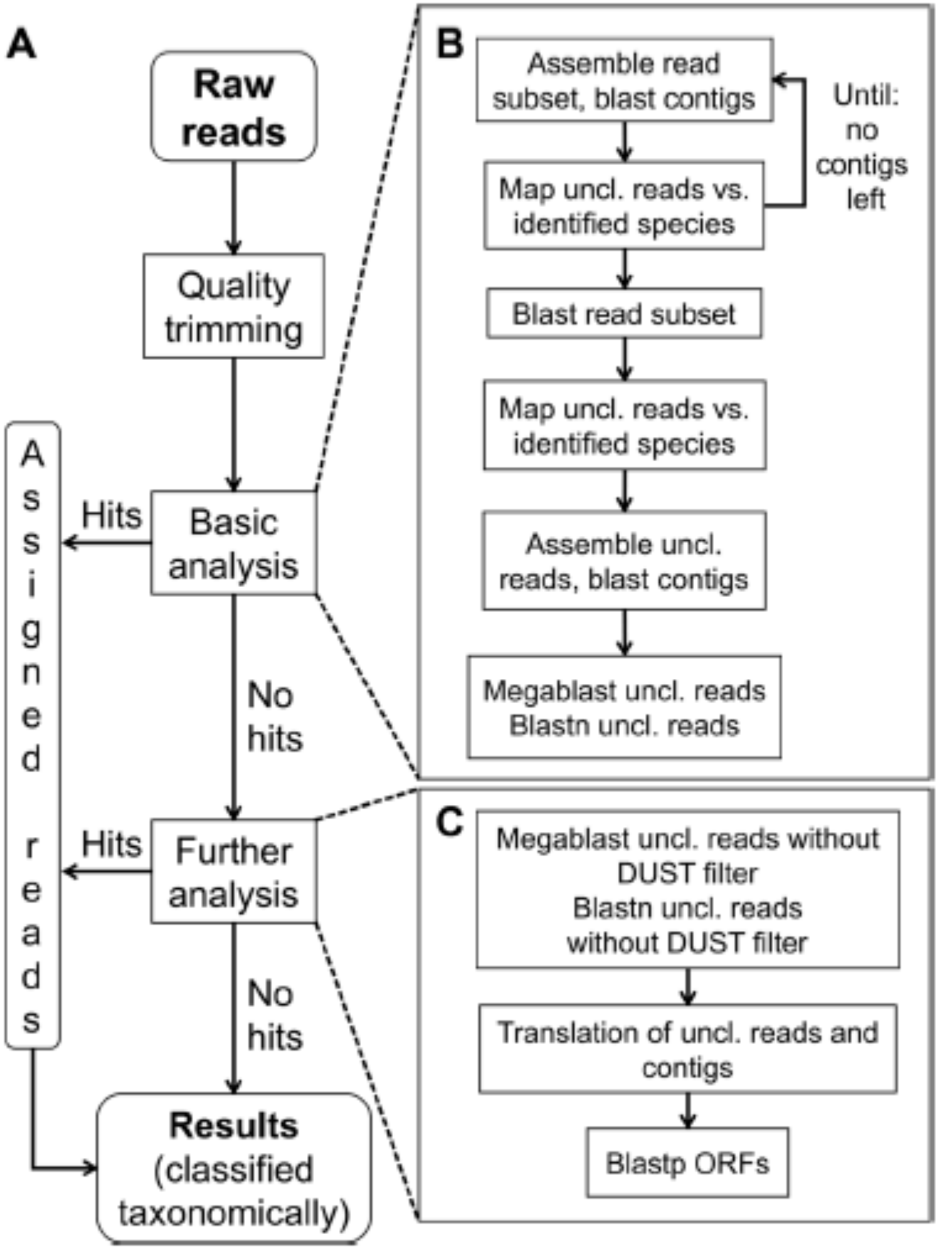
BMC Bioinformatics

METHODOLOGY ARTICLE Open Access

RIEMS: a software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets

Matthias Scheuch[†], Dirk Höper[*†] and Martin Beer

Scheuch *et al*. (2015) BMC Bioinformatics 16:69



A

Raw reads → Quality trimming → Basic analysis

Basic analysis: Hits → Assigned reads

Basic analysis: No hits → Further analysis

Further analysis: Hits → Assigned reads

Further analysis: No hits → Results (classified taxonomically)

Assigned reads → Results (classified taxonomically)

B

Assemble read subset, blast contigs

Map uncl. reads vs. identified species

Until: no contigs left

Blast read subset

Map uncl. reads vs. identified species

Assemble uncl. reads, blast contigs

Megablast uncl. reads
Blastn uncl. reads

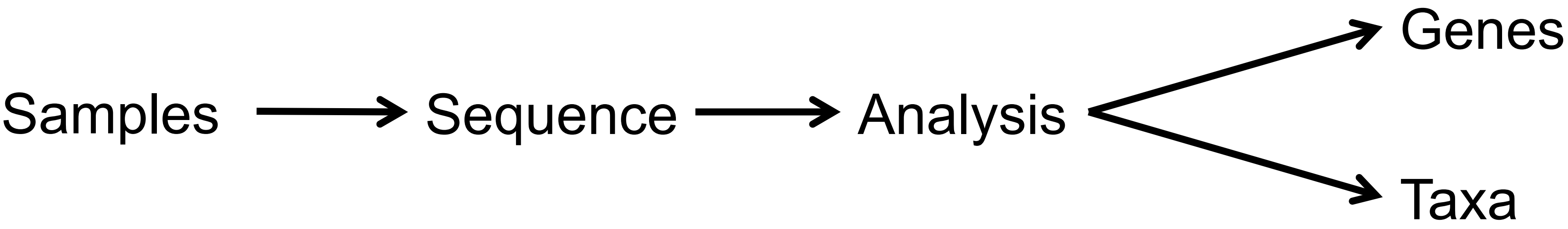
C

Megablast uncl. reads without DUST filter
Blastn uncl. reads without DUST filter

Translation of uncl. reads and contigs

Blastp ORFs

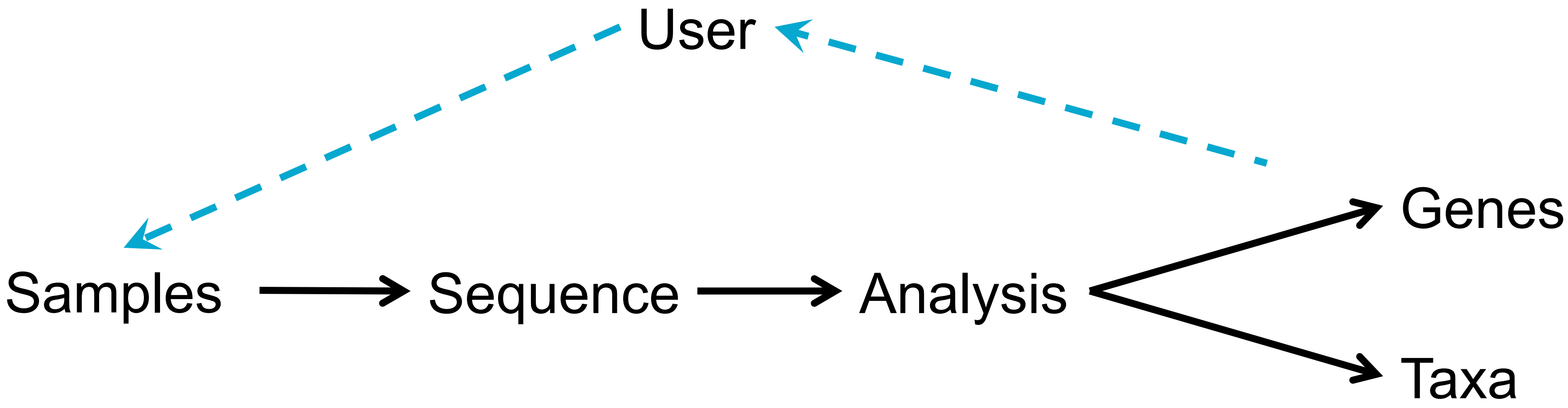
Identification data



Identification data accumulation

- Identification against context is informative
- (Molecular observations may be tentative or high confidence)
- Coincidental observations (time, place, virulence, host phenotypes)
- ‘Capture’ identifications to ‘connect’ coincidental observations

Identification data



User

Samples

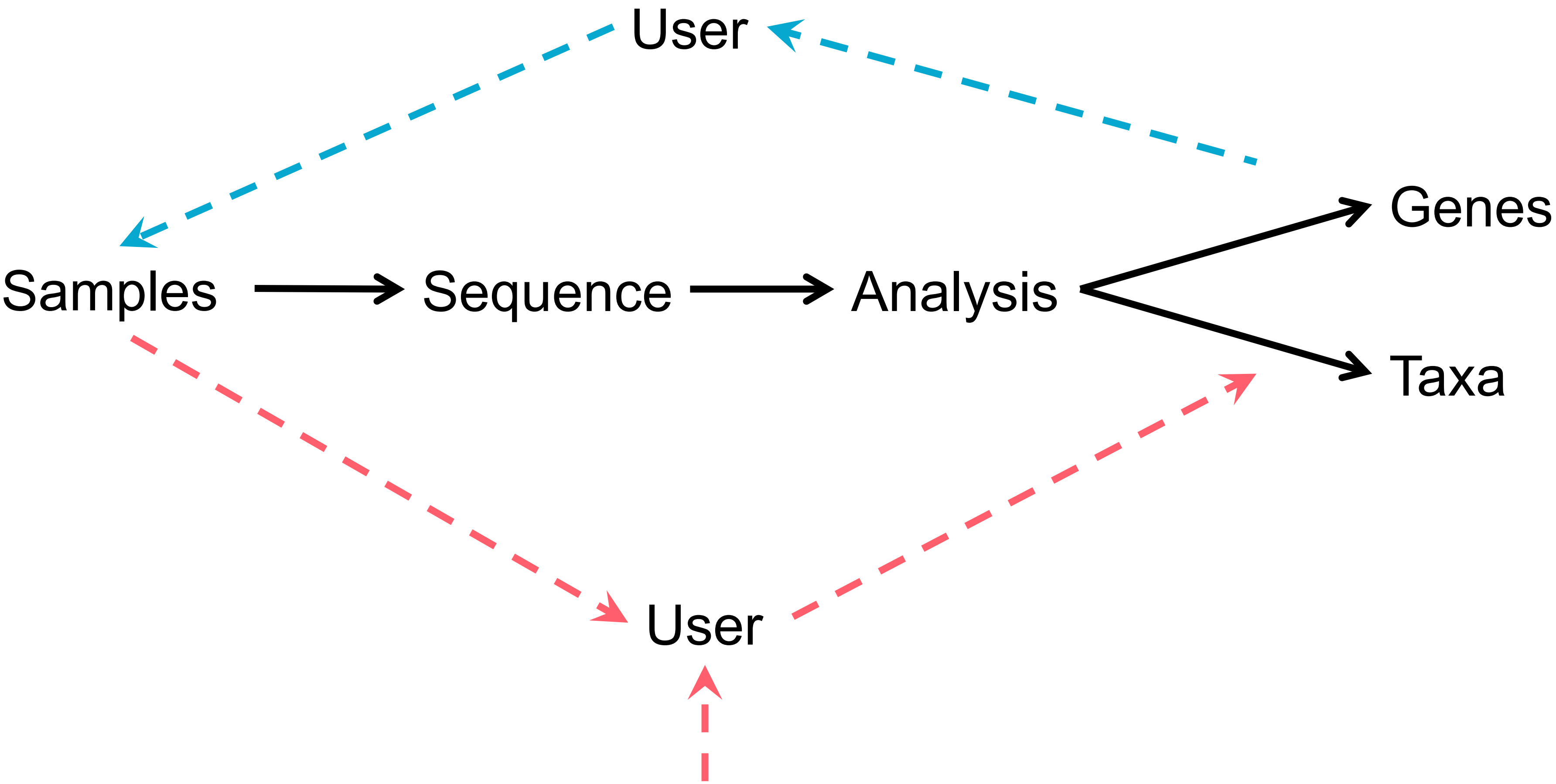
Sequence

Analysis

Genes

Taxa

Identification data



Acknowledgements

Standards & support

Ana Cerdeño-Tárraga, Ana Luisa Toribio, **Petra ten Hoopen**, Marc Rosello, Richard Gibson, Jeena Rajan, **Clara Amid**

Compression technology

Vadim Zalunin, Ewan Birney, ***James Bonfield***, Rasko Leinonen

ENA technical services

Blaise Alako, Simon Kay, Nima Pakseresht, Xin Liu, Suran Jayathilaka, Nicole Silvester, Daniel Vaughan, Neil Goodgame, Iain Cleland, Dmitriy Smirnov, Kethi Reddy, Vadim Zalunin, **Rasko Leinonen**

ENA - http://www.ebi.ac.uk/ena/

BBSRC
bioscience for the future

SEVENTH FRAMEWORK PROGRAMME

EMBL

wellcome trust

HORIZON 2020
THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION