Eiropas Molekulārās bioloģijas laboratorija un vai Latvijai jākļūst par tās dalībvalsti

Alvis Brāzma
European Bioinformatics Institute
European Molecular Biology Laboratory

EMBL-EBI

Runas plāns

- Kas ir EMBL?
- Kas ir EMBL-EBI un ko es tur daru?
- Kamdēļ Latvijai būtu jākļūst par EMBL dalībvalsti?

History of EMBL

"I believe that international activity is very important in building world peace." Sir John Kendrew, EMBL's 1st DG

Europe's centre of excellence in life science research, services and training

Founded in 1974 by 10 states as an intergovernmental organisation to promote the molecular life sciences in Europe and beyond

Founders Sir John Kendrew and James Watson

EMBL

EMBL is growing – the current Member States

Member states (27)

Austria 1974
Denmark 1974
France 1974
Germany 1974
Israel 1974
Italy 1974
Netherlands1974
Sweden 1974
Switzerland 1974
United Kingdom 1974
Finland 1984
Greece 1984
Norway 1985
Spain 1986
Belgium 1990
Portugal 1998
Ireland 2003
Iceland 2005
Croatia 2006
Luxembourg 2007
Czech Republic 2014
Malta 2016
Hungary 2017
Slovakia 2018
Montenegro 2018
Poland 2019
Lithuania 2019

Associate member states

Australia 2008
Argentina 2014
India applied

Prospect member states

Estonia 2019

EMBL

EMBL Sites – over 1700 people and more than 80 nationalities

Hinxton EMBL-EBI

Bioinformatics

Grenoble

Structural Biology

Barcelona

Tissue Biology and Disease Modelling

Hamburg

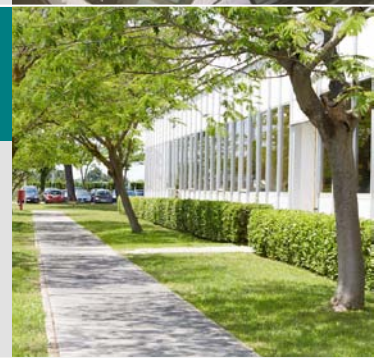
Structural Biology

Heidelberg

Life Sciences

Rome

Epigenetics and Neurobiology



EMBL

EMBL’s Core Principles

Scientific excellence

Internationality and diversity

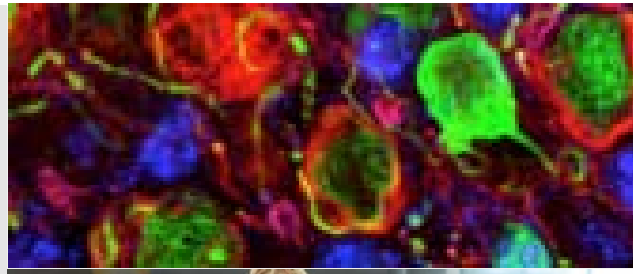
Collaboration

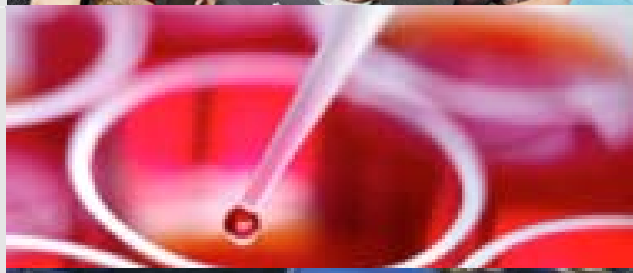
Scientific freedom

Cutting-edge infrastructure

Staff turnover

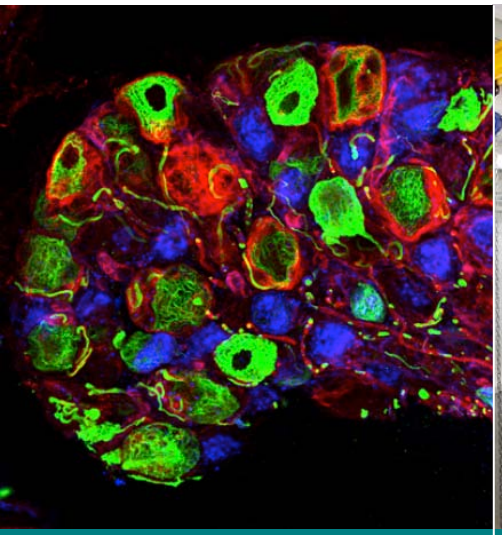
Young talent and early independence





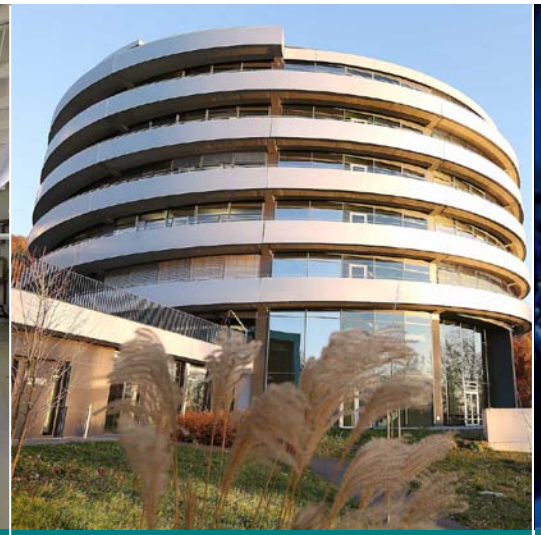
EMBL

EMBL’s missions



Fundamental research

Services



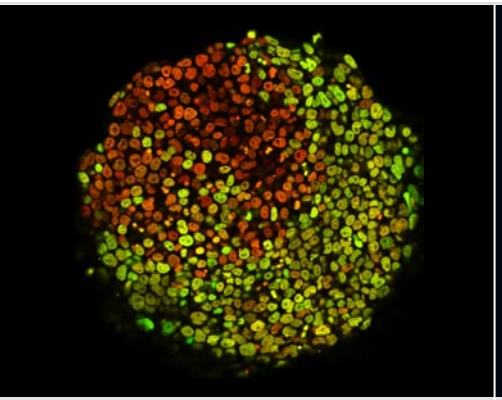
Advanced training

Technology development & transfer

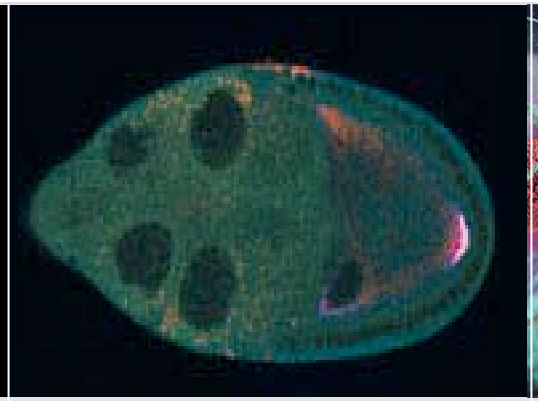
Integration of life science research

EMBL

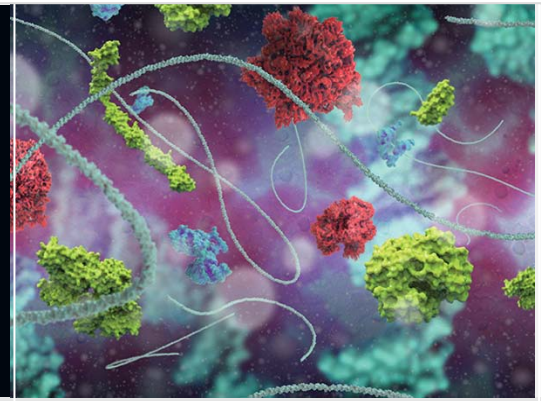
EMBL research units – over 80 independent research groups



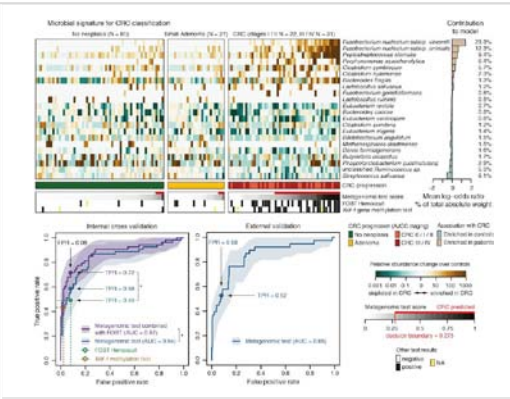
Cell biology and biophysics – Heidelberg



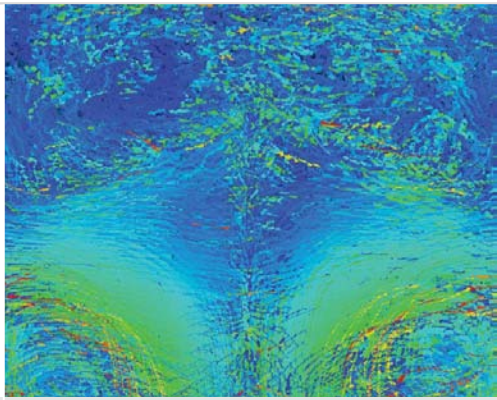
Developmental biology – Heidelberg



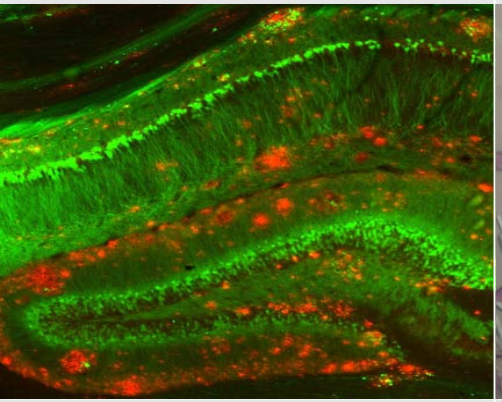
Genome biology – Heidelberg



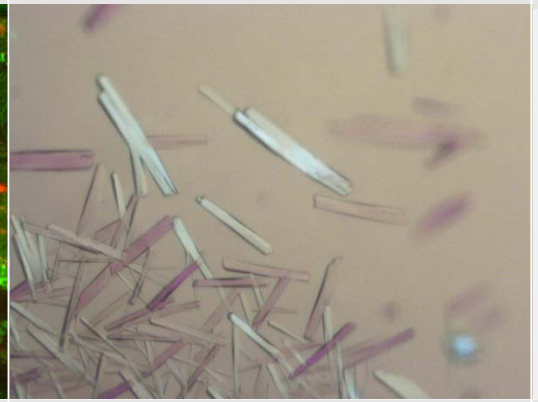
Structural and computational biology – Heidelberg



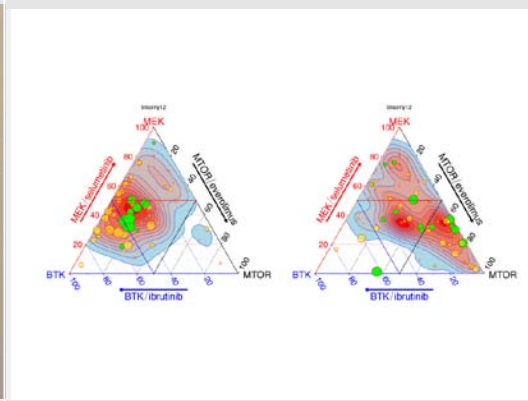
Directors' research – Heidelberg



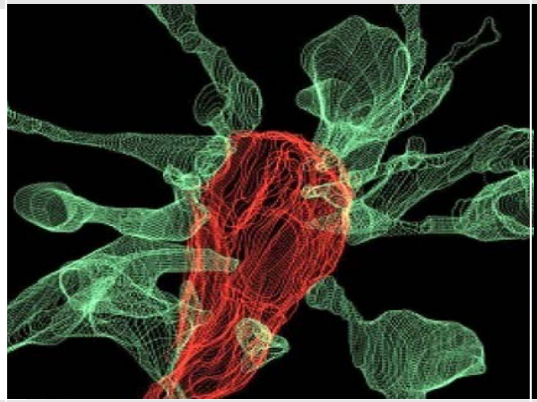
Structural biology – Hamburg



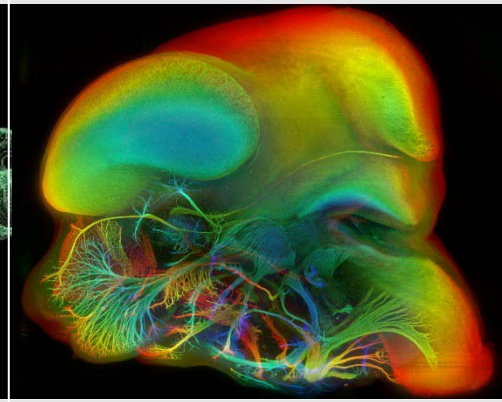
Structural biology – Grenoble



Bioinformatics – EMBL-EBI, Hinxton



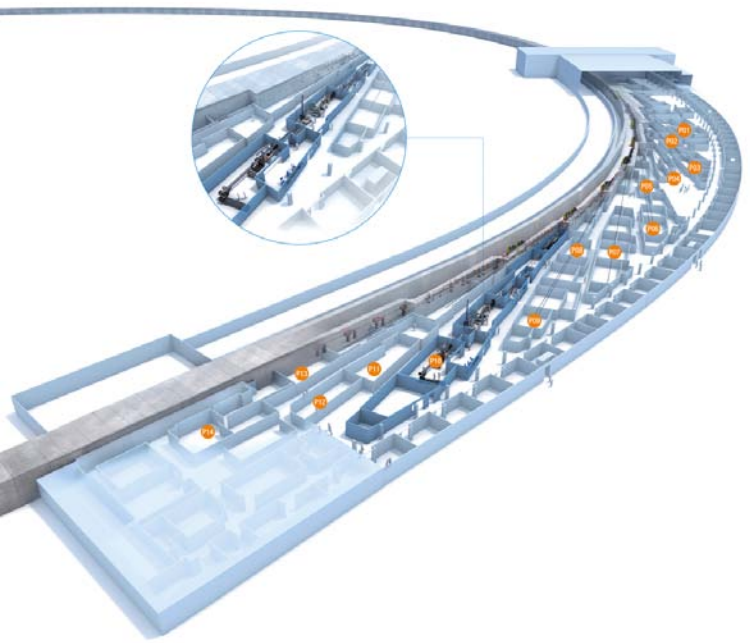
Epigenetics and neurobiology – Rome



Tissue biology and disease modelling – Barcelona

EMBL

Access to infrastructure and services



Structural biology services

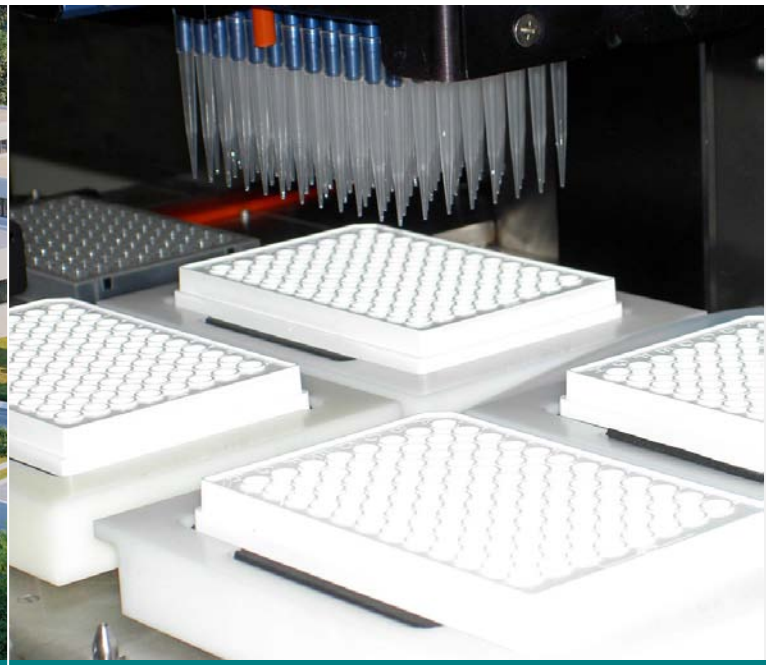
ESRF(Grenoble) and PETRA IV (Hamburg) upgrades

Sample Preparation and Characterisation Facility

Imaging Centre

Opening in 2021 at Heidelberg site

Developing integrative novel technologies with industry partners in framework collaborations



Core Facilities

From -omics to imaging

Supporting and rapidly evolving technologies

EMBL

Training

Internal

200 PhD students,

250 postdocs

EMBL International PhD Programme

EMBL Postdoctoral Programme

General Training and Development

External

~ 7000 guests per year

EMBL Courses and Conferences

EMBL Visitor Programme

Online training

European Learning Lab for the Life Sciences (ELLS)

EMBL

Internal Training

EMBL International PhD Programme

- > 200 students from over 50 countries
- ~ 50 students accepted each year
- Joint PhD degree with Cambridge and Heidelberg Universities

"The vast network of collaborations creates a hotbed for creative science"
Hernando Martínez, EMBL PhD student from Spain

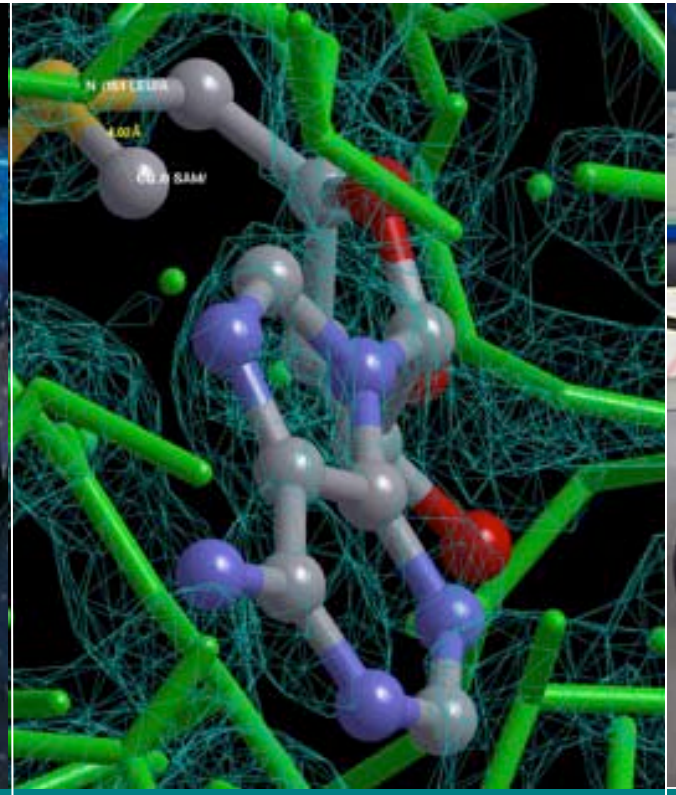
Postdoctoral programmes

> 250 postdocs

- Classical postdoctoral scheme
- EMBL Interdisciplinary Postdocs (EIPODs)
- EMBL-EBI Sanger Postdocs (ESPODs)
- EMBL-EBI BRC* Postdocs (EBPODs)

Technology transfer and development

Imaging technology



Software development

Synchrotron instrumentation

EMBL develops a broad spectrum of technology and instrumentation for life science research

EMBL makes its discoveries and inventions available to the scientific community and to society through EMBLEM

EMBL

Bioinformatics: EMBL services provide resources for big data

EMBL-EBI is the global leader in biological data management

Hosts archival data resources, knowledge-bases and driver of data standards

Data coordinators in large-scale international research consortia

Enabling data access to researchers and clinicians world-wide

Bioinformatics services: From people to microbes, from genomes to systems

EMBL

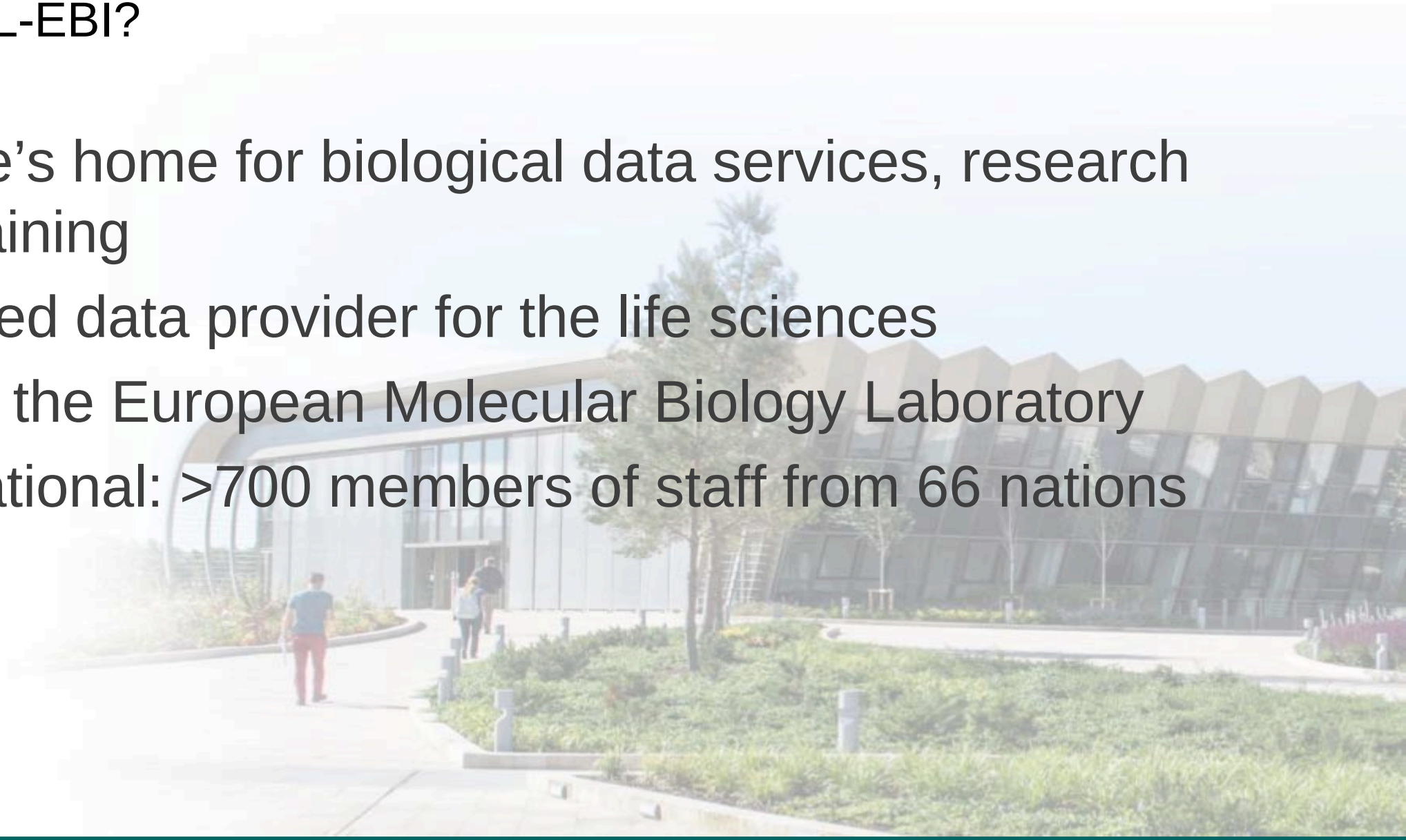
EMBL Cambridge – European Bioinformatics Institute

Bioinformatics – using computational methods and tools to tools to study biological systems

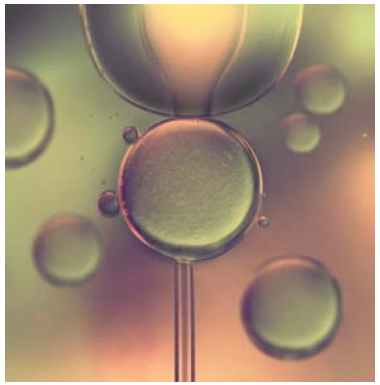
Services, Research, Training
Industry engagement
European Coordination

What is EMBL-EBI?

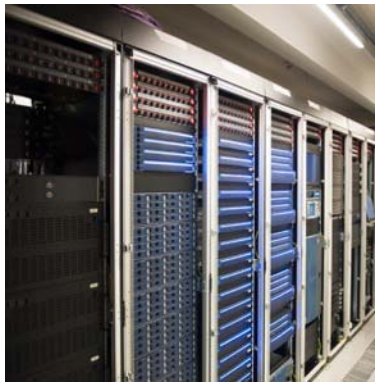
- Europe’s home for biological data services, research and training
- A trusted data provider for the life sciences
- Part of the European Molecular Biology Laboratory
- International: >700 members of staff from 66 nations



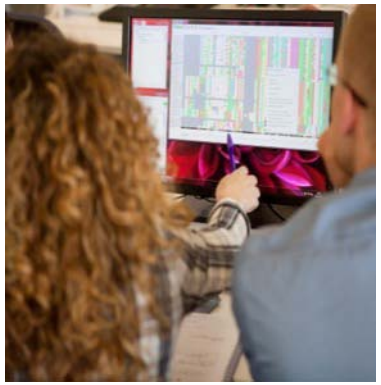
Our mission



Deliver excellent research



Deliver scientific services

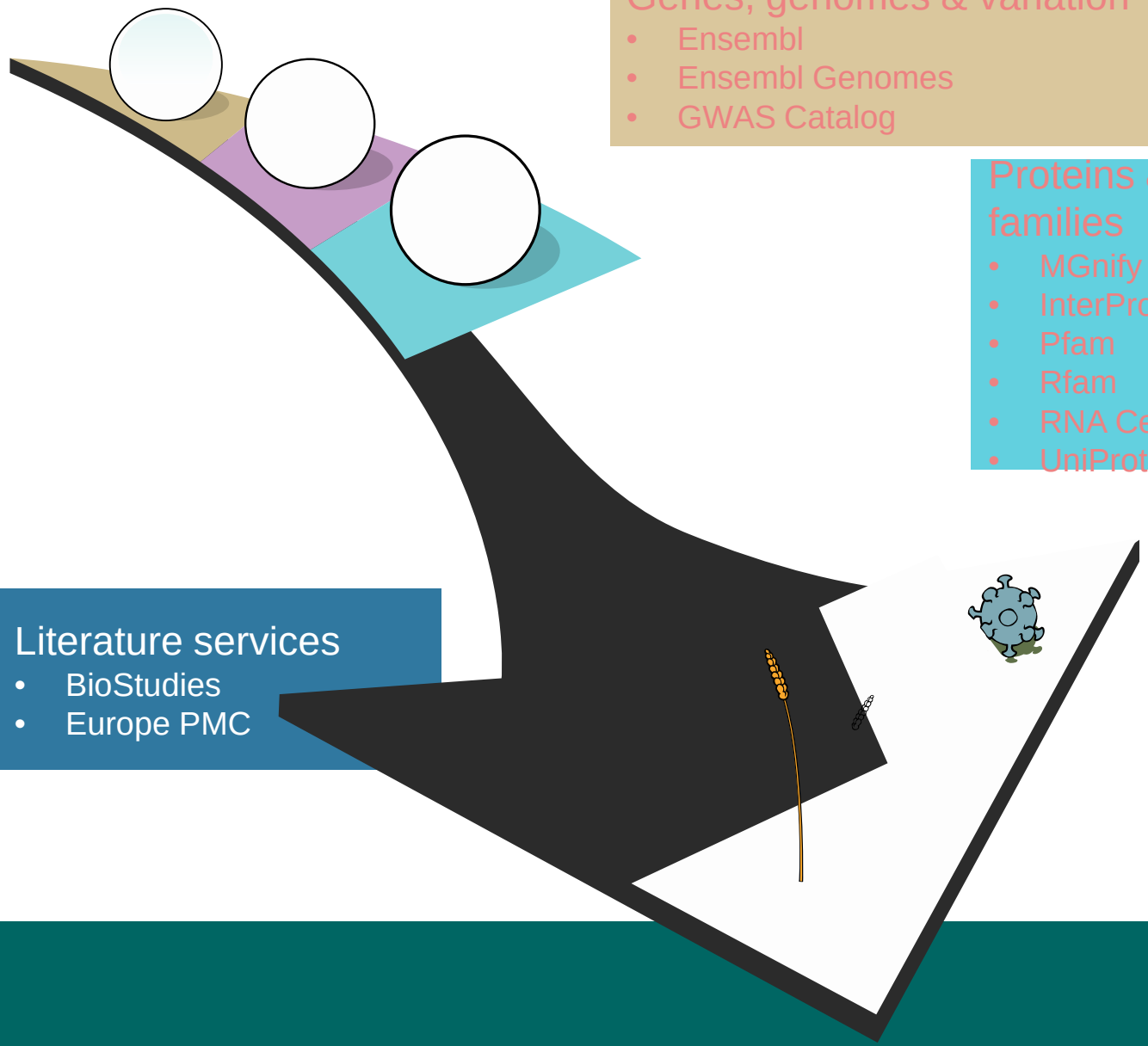


Train the next generation of scientists

Engage with industry

Coordinate bioinformatics in Europe

Services - data resources at EMBL-EBI



Genes, genomes & variation
- Ensembl
- Ensembl Genomes
- GWAS Catalog

Molecular atlas
- Array Express
- Expression Atlas
- PRIDE

Proteins & protein families
- MGnify
- InterPro
- Pfam
- Rfam
- RNA Central
- UniProt

Macromolecular & cellular structure
- Protein Data Bank in Europe (PDBe)
- PDBe-KB
- Electron Microscopy Data Bank
- EMPIAR

Literature services
- BioStudies
- Europe PMC

Chemistry services
- ChEBI
- ChEMBL
- MetaboLights
- SureChEMBL

Molecular systems
- BioModels
- IntAct
- OmicsDI
- Reactome

Molecular archives
- European Nucleotide Archive
- European Variation Archive
- European Genome-phenome Archive
- Experimental Factor Ontology
- BioSamples
- Mouse Resources

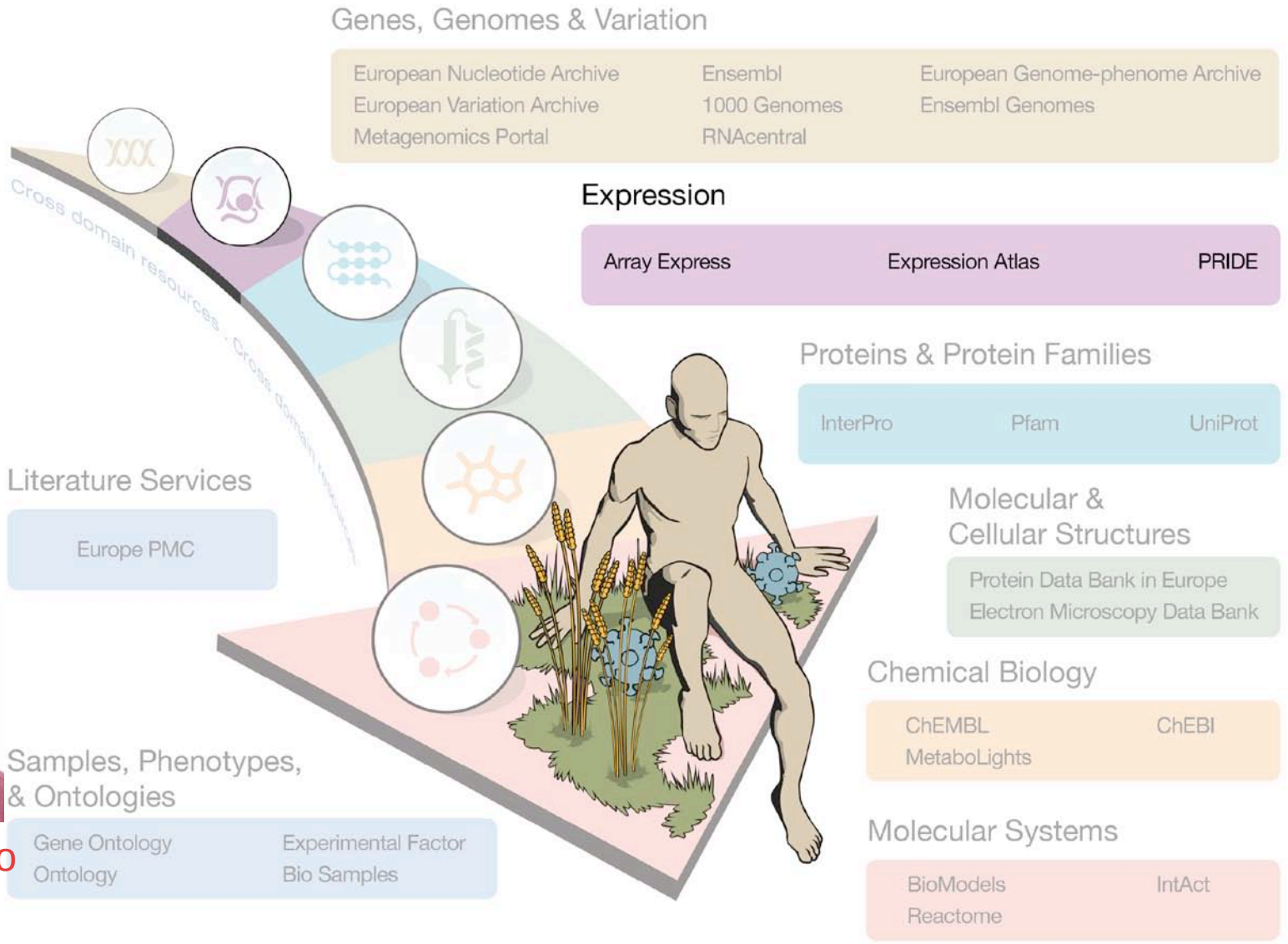
Molecular Atlas

Alvis Brazma

Ugis Sarkans

Irene Papaheudoru

Juan Antonio Vicziano



Nature 403, 699-700 (17 February 2000)

One-stop shop for microarray data

Is a universal, public DNA-microarray database a realistic goal?

Alvis Brazma, Alan Robinson, Graham Cameron and Michael Ashburner

Of the techniques that are being used to obtain the massive data sets of the molecules of life, the most visible is the DNA sequencing of the human genome. Following on from the publication of the human chromosome 22 sequence[1], a rough draft of the whole human genome should be available by the spring. But such advances can create the false impression that everything about life at the molecular level will soon be understood.

In reality, genome projects simply transfer digital information from DNA to computer file; this genetic 'parts-list' is a long way from providing an understanding of function. It took hundreds of years to advance from a fairly detailed understanding of human

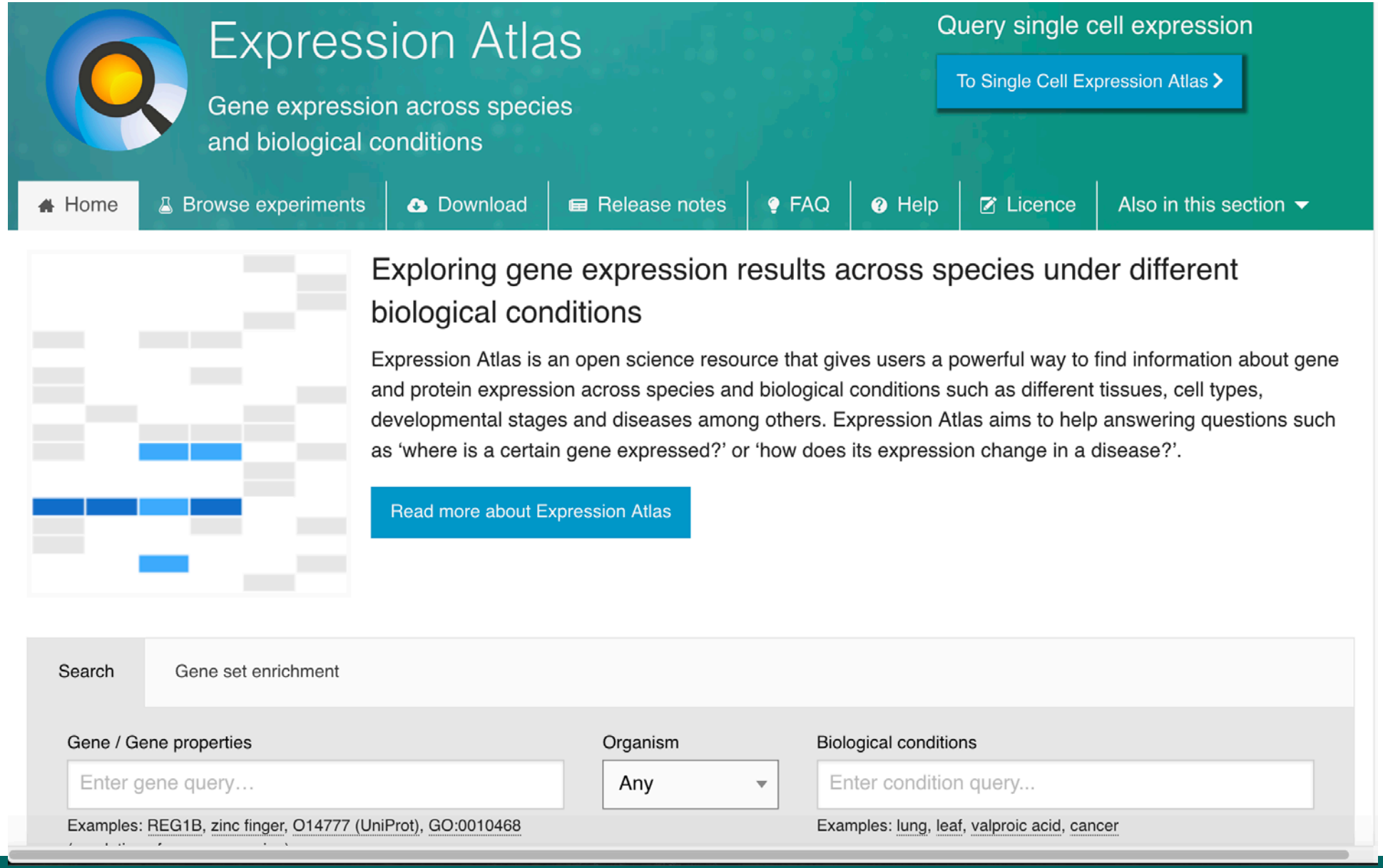
experiment looking at 40,000 genes from 10 different samples, under 20 different conditions, produces at least 8,000,000 pieces of information. Currently, these data are scattered among various independent Internet sites, or may not be publicly available at all, although conclusions drawn from the data will have been published. Details about how experiments were carried out are often incomplete. Yet the amount of information being produced in this way is set to explode as the cost of microarray technology falls.

The need for a public repository

It is time to create a public repository for microarray data, with standardized annotation (see Box 2, overleaf). But this is a complex and ambitious project, and is one of the biggest challenges that bioinformatics has yet faced. Major difficulties stem from the detail

One difficulty concerns the inherent fuzziness of gene-expression data. Essentially all current expression measurements are relative: we can tell which genes are expressed differently in an experiment only in comparison with another experiment, or in relation to another gene in the same experiment. Such methods tell us little about how many copies of a messenger RNA are present. Moreover, the transcription levels reported are an average over the whole cell population sampled.

Consequently, gene-expression measurements from different technologies, or even from the same technology but from different laboratories, may not be quantitatively comparable. Two steps should allow data from different sources to be compared. First, relatively raw data should be stored to obviate any variation owing to, say, data-normalization methods. Second, standard sets of control


Expression Atlas
Gene expression across species and biological conditions
Query single cell expression
To Single Cell Expression Atlas
Home
Browse experiments
Download
Release notes
FAQ
Help
Licence
Also in this section
Exploring gene expression results across species under different biological conditions
Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how does its expression change in a disease?'.
Read more about Expression Atlas
Search
Gene set enrichment
Gene / Gene properties
Enter gene query…
Examples: REG1B, zinc finger, O14777 (UniProt), GO:0010468
Organism
Any
Biological conditions
Enter condition query...
Examples: lung, leaf, valproic acid, cancer

EMBL

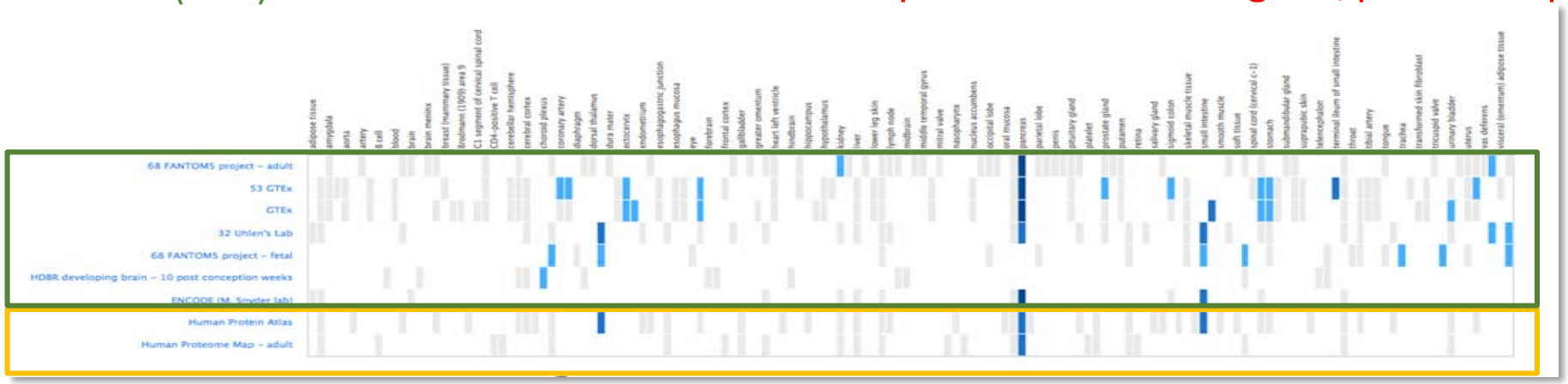
From gene expression in tissues to single cells

Baseline (bulk)

expression of REG1B gene/protein in pancreas

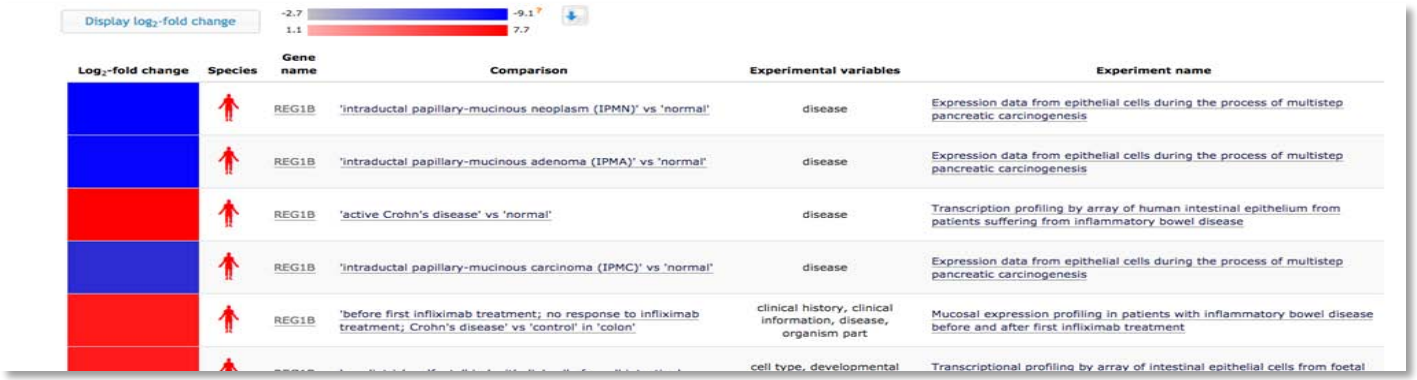
transcriptomics

proteomics



Differential (bulk)

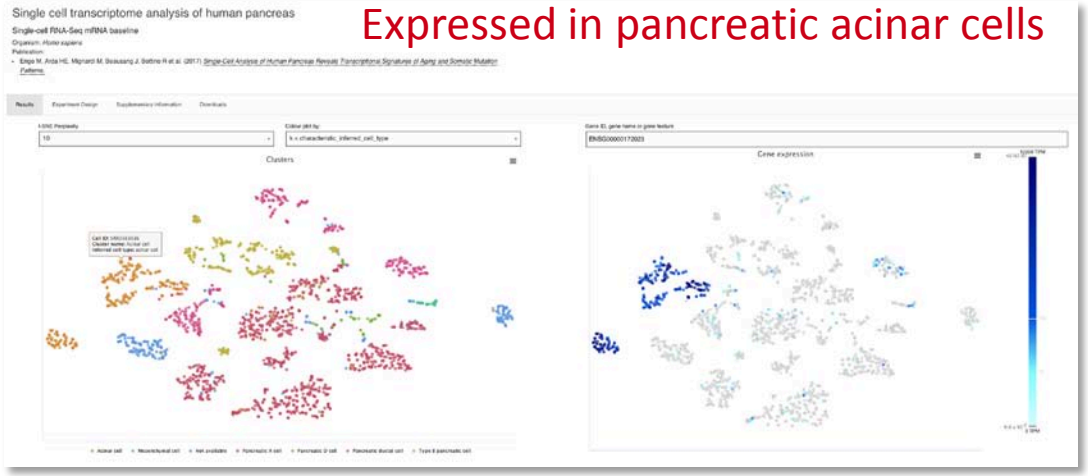
Down-regulated in pancreatic cancer



Log2-fold change	Species	Gene name	Comparison	Experimental variables	Experiment name
		REG1B	'intraductal papillary-mucinous neoplasm (IPMN)' vs 'normal'	disease	Expression data from epithelial cells during the process of multistep pancreatic carcinogenesis
		REG1B	'intraductal papillary-mucinous adenoma (IPMA)' vs 'normal'	disease	Expression data from epithelial cells during the process of multistep pancreatic carcinogenesis
		REG1B	'active Crohn's disease' vs 'normal'	disease	Transcription profiling by array of human intestinal epithelium from patients suffering from inflammatory bowel disease
		REG1B	'intraductal papillary-mucinous carcinoma (IPMC)' vs 'normal'	disease	Expression data from epithelial cells during the process of multistep pancreatic carcinogenesis
		REG1B	'before first infliximab treatment; no response to infliximab treatment; Crohn's disease' vs 'control' in 'colon'	clinical history, clinical information, disease, organism part	Mucosal expression profiling in patients with inflammatory bowel disease before and after first infliximab treatment
				cell type, developmental	Transcriptional profiling by array of intestinal epithelial cells from foetal

Single Cell

Expressed in pancreatic acinar cells



Research groups at EMBL-EBI

Alex Bateman | Ewan Birney | Pedro Beltrao | Alvis Brazma | Rob Finn | Paul Flicek | Andrew Leach | Moritz Gerstung

Nick Goldman | Zamin Iqbal | John Marioni | Evangelia Petsalaki | Oliver Stegle | Janet Thornton | Virginie Uhlmann | Daniel Zerbino

PERSPECTIVES

International network of cancer genome projects

The International Cancer Genome Consortium*

The International Cancer Genome Consortium (ICGC) was launched to coordinate large-scale cancer genome studies in tumours from 50 different cancer types and/or subtypes that are of clinical and societal importance across the globe. Systematic studies of more than 25,000 cancer genomes at the genomic, epigenomic and transcriptomic levels will reveal the repertoire of oncogenic mutations, uncover traces of the mutagenic influences, define clinically relevant subtypes for prognosis and therapeutic management, and enable the development of new cancer therapies.

The genomes of all cancers accumulate somatic mutations[1]. These include nucleotide substitutions, small insertions and deletions, chromosomal rearrangements and copy number changes that can affect protein-coding or regulatory components of genes. In addition, cancer genomes usually acquire somatic epigenetic 'marks' compared to non-neoplastic tissues from the same organ, notably changes in the methylation status of cytosines at CpG dinucleotides.

A subset of the somatic mutations in cancer cells confers oncogenic properties such as growth advantage, tissue invasion and metastasis, angiogenesis, and evasion of apoptosis[2]. These are termed 'driver' mutations. The identification of driver mutations will provide insights into cancer biology and highlight new drug targets and diagnostic tests. Knowledge of cancer mutations has already led to the development of specific therapies, such as trastuzumab for *HER2* (also known as *NEU* or *ERBB2*) positive breast cancers[3] and imatinib, which

incomplete studies; (3) lack of standardization across studies could diminish the opportunities to merge and compare data sets; (4) the spectrum of many cancers is known to vary across the world; and (5) an international consortium will accelerate the dissemination of data sets and analytical methods into the user community.

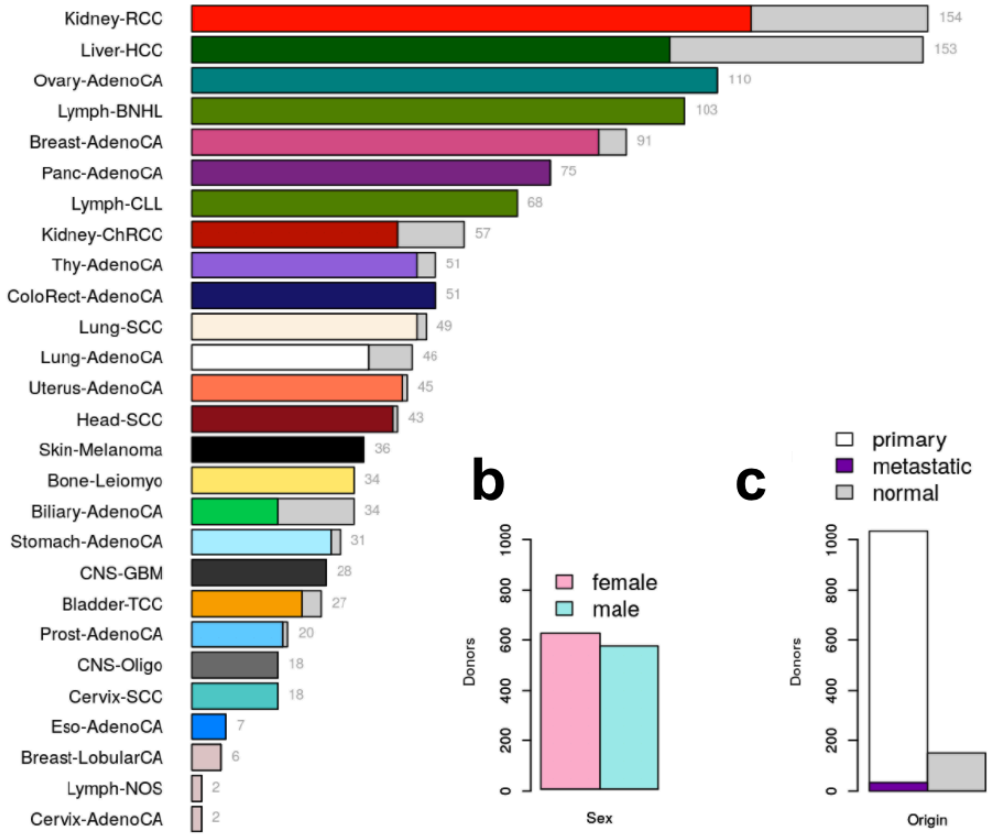
Working groups were created to develop strategies and policies that would form the basis for participation in the ICGC. The goals of the consortium (Box 1) were released in April 2008 (http://www.icgc.org/files/ICGC_April_29_2008.pdf). Since then, working groups and initial member projects have further refined the policies and plans for international collaboration.

Bioethical framework

ICGC members agreed to a core set of bioethical elements for consent as a precondition of membership (Box 2). The Ethics and Policy

CAGEKID - genomic approaches to identify biomarkers of most common form of **clear-cell renal carcinoma**

Pan-Cancer Analysis of Whole Genomes: PCAWG-3 integrating genomes with RNA



- Donors: 1188
- Samples
  - Cancer : 1209
  - Normals: 150
- 27 cancer types

- Use of GTEx data

PCAWG
PanCancer Analysis
OF WHOLE GENOMES

EMBL

Genomic basis for RNA alterations revealed by whole-genome analyses of 27 cancer types

PCAWG Transcriptome Core Group; Claudia Calabrese[1*], Natalie R. Davidson[2,3,4,5,6*], Nuno A. Fonseca[1*], Yao He[7*], André Kahles[2,3,5,6*], Kjong-Van Lehmann[2,3,5,6*], Fenglin Liu[7*], Yuichi Shiraishi[8*], Cameron M. Soulette[9*], Lara Urban[1*]; Deniz Demircioğlu[10,11], Liliana Greger [1], Siliang Li[12,13], Dongbing Liu[12,13], Marc D. Perry[14,15], Linda Xiang[14], Fan Zhang[7], Junjun Zhang[14], Peter Bailey[16], Serap Erkek[17], Katherine A. Hoadley[18], Yong Hou[12,13], Helena Kilpinen[19], Jan O. Korbel[17], Maximillian G. Marin[9], Julia Markowski[20], Tannistha Nandi[11], Qiang Pan-Hammarström[12,21], Chandra Sekhar Pedamallu[22], Reiner Siebert[23], Stefan G. Stark[2,3,5,6], Hong Su[12,13], Patrick Tan[11,24], Sebastian M. Waszak[17], Christina Yung[14], Shida Zhu[12,13], PCAWG Transcriptome Working Group, Philip Awadalla[14,25], Chad J. Creighton[26], Matthew Meyerson[22,27,28], B.F. Francis Ouellette[29], Kui Wu[12,13], Huangming Yang[12], ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Network, Alvis Brazma[1,33#], Angela N. Brooks[9,22,27,33#], Jonathan Göke[11,30#], Gunnar Rätsch[2,3,4,5,6,33#], Roland F. Schwarz[1,20,31,32#], Oliver Stegle[1,17#], Zemin Zhang[7#]

*Co-first author

Senior authors

[33]Correspondence: Alvis Brazma (brazma@ebi.ac.uk), Angela N. Brooks (anbrooks@ucsc.edu), Gunnar Rätsch (Gunnar.Ratsch@ratschlab.org)

[1]European Molecular Biology Laboratory, Hinxton, CB10 1SD, UK , [2]ETH Zurich, Zurich, 8092, Switzerland, [3]Memorial Sloan Kettering Cancer Center, New York, 10065, USA, [4]Weill Cornell Medical College, New York, 10065, USA, [5]SIB Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland, [6]University Hospital Zurich, Zurich, 8091, Switzerland, [7]Peking University, Beijing, 100871, China, [8]The University of Tokyo, Minato-ku, 108-8639, Japan, [9]University of California, Santa Cruz, Santa Cruz,

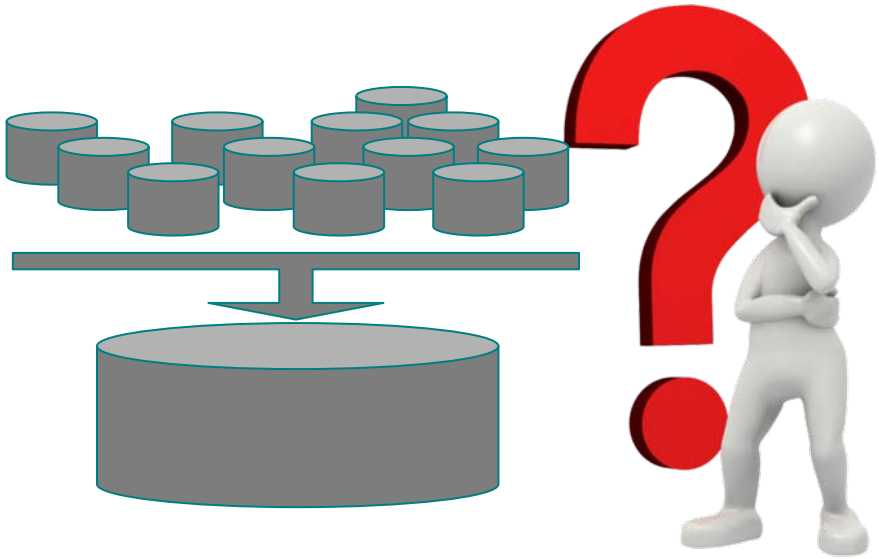
Accepted in for publication in Nature

EMBL

Projekts: “sRNAflow - rīks mazo RNS sekvenēšanas datu analīzei bioloģiskajos šķidrumos”
Starpposma uzdevums:
Saģenerēt metagenomu, kas apvieno sugu genomus:

- Cilvēka
- 44000 baktēriju
- 400 sēņu
- 200 protistu.

Izveidotā metagenoma izmērs: 192GB.



Indeksācijai nepieciešamā aparatūra:
1.5 TB operativa atmiņa, ~8000 CPU/stundas, 600 GB glabātuve.

EMBL-EBI

Pavels Zajakins (visited EBI for two months and more is planed)

LV BMC

Latvijas Biomedicīnas pētījumu un studiju centrs
biomedicīnas pētījumi un izglītība no gēniem līdz cilvēkam

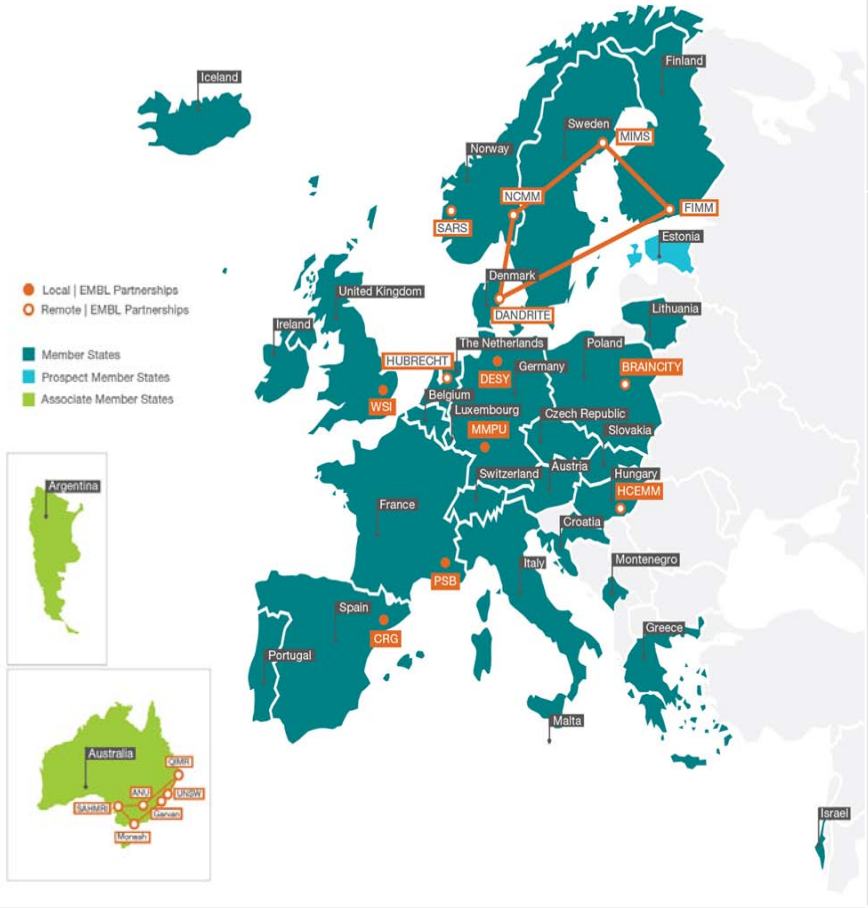
Exporting the EMBL model in Member States: EMBL Partnerships

Close cooperative affiliation with national institutes in EMBL member states

Exploit complementarity or synergy & transfer know-how

No net transfer of EMBL resources possible

EMBL Partnerships



Establish network of international centers of scientific excellence and advanced training modelled on EMBL

Tailor-made to serve national interest

Exports EMBL model and implements it nationally

EMBL

Should Latvia join EMBL?

Are there any advantages for a small country in joining an international science organisation?

Why a small country should join an international science organisation?

- The role of international science organisations:
  - To serve science
  - To serve the members and member states
- Through participating scientists
  - Build closer links to each other internationally (facilitating joint projects, fundraising and other activities)
  - Benefit from joint infrastructure (including training)
- How does this help a small country and its scientists?
  - Building the visibility and prestige of the science and the country
  - Making contacts to scientists and policy makers of other countries

Why a small country should join an international science organisation?

- Cons
  - It costs money that could otherwise be spent on science at home
- However, it is an investment
  - It may help to bring in funds by increasing success rate of applications to international science funding agencies
  - Perhaps this can raise science profile at home and through this more national funding for science?
  - I shouldn’t necessarily cost very much

There were nine Member States that reported R & D expenditure that was below 1.00 % of their GDP in 2016, (…), with the lowest R & D intensities recorded in Cyprus (0.50 %), Romania (0.48 %) and Latvia (0.44 %).

(Source http://ec.europa.eu/eurostat/)

Can we do better, can EMBL membership help?

- Participation in EMBL will improve the opportunities of obtaining international science funding
- Participation in international science organisation can advance visibility and standing of science in Latvia
- Will this improve the government’s willingness to fund science in Latvia? (A naïve thought?)

Prospect Membership

The aim: Attract European countries to consider acceding to EMBL

- 3 year-transitional scheme towards full membership status
- Broad access to the EMBL services and facilities under agreed terms
- More opportunities for PhD students, postdoctoral researchers and visitors
- Observer status in the EMBL Council with the right to speak
- Awareness-raising campaign on EMBL and its opportunities in the country
- Collaborations in H2020 grants, including Teaming and Twinning, and any other collaborations of interest to Estonian community
- Accompanying proposal for membership fee reduction during the first 5 years of full membership

EMBL

Prospect Membership

- No financial contribution
- Reduced membership contribution during the first 5 years of Full Membership
- Very successful - 5 countries (SK, HU, PL, LT, EE) joined in recent years

	Prospect Member (Observer)			**Full Member** (Voting rights)					
Year	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9
Annual Membership contribution	0%	0%	0%	30%	40%	50%	80%	80%	100%
Special contribution towards capital expenditure (entry fee)	/	/	/	20%	20%	20%	20%	20%	/

EMBL

EMBL Member States

Member states (27)

Austria 1974
Denmark 1974
France 1974
Germany 1974
Israel 1974
Italy 1974
Netherlands1974
Sweden 1974
Switzerland 1974
United Kingdom 1974
Finland 1984
Greece 1984
Norway 1985
Spain 1986
Belgium 1990
Portugal 1998
Ireland 2003
Iceland 2005
Croatia 2006
Luxembourg 2007
Czech Republic 2014
Malta 2016
Hungary 2017
Slovakia 2018
Montenegro 2018
Poland 2019
Lithuania 2019

Associate member states

Australia 2008
Argentina 2014
India applied

Prospect member states

Estonia 2019

EMBL

Thank you!

EMBL