# Welcome to the European Bioinformatics Institute

Part of the European Molecular Biology Laboratory

## Who we are and what we do

The European Bioinformatics Institute (EMBL-EBI) is an academic research institute that is part of EMBL, Europe's flagship laboratory for the life sciences. Our 520 staff hail from 37 countries, and we welcome a regular stream of visiting scientists throughout the year. We share a campus with the Wellcome Trust Sanger Institute near Cambridge in the UK.

EMBL-EBI provides freely available data from life science experiments to the scientific community, and we perform basic research in computational biology. Our extensive training programme helps researchers in academia and industry to make the most of the incredible amount of data being produced every day in life science experiments.

Watch a 2-minute video about the EBI

## What is bioinformatics?

Bioinformatics is the application of computer technology to the storage, management and analysis of data from life science experiments.

One of the biggest challenges in biology today is analysing the massive volumes of data created in "high-throughput" experiments, for example DNA sequencing. Bioinformatics makes it possible to extract meaningful information from a sea of data. It provides the means to pull together many different kinds of information so that we can begin piecing together the great puzzle of how biological systems work.

## Why is bioinformatics important for me?

Biological data is the bedrock of life science research. Here are a few examples of how it can be used in beneficial ways:

- Understanding plant genomes helps us identify which species will be most tolerant to drought, salt and pests while still providing optimum nutrition.
- If we can identify patterns of genes that are active in different tumours, we can diagnose and treat cancers earlier.
- Methicillin-resistant Staphylococcus aureus (MRSA) infection is a global problem. Small variations in DNA sequence can help track transmission - this technology can help identify the source of new outbreaks.
- Drug resistance is growing to the one medicine used to treat Schistosomiasis, a parasitic infection. Studying the Schistosome genome will help identify the targets of existing drugs.

- Short sections of DNA - barcodes - are used to identify an organism. The Barcode of Life Initiative is developing DNA barcoding as a global standard for identifying species, which will have applications in the protection of endangered species, sustaining natural resources through pest control and food labelling.

## Quick facts

- The storage capacity of computing hardware doubles every 18 months but new biological data are doubling every 9 months.
- Over a third (36.3%) of Europe's life science databases have no assured funding.
- Users of Europe's biological databases range from clinical specialists to school teachers, environmental researchers and computer scientists.
- In 2010, the website of EMBL-EBI was visited by approximately 3.4 million unique IP, or web addresses. That could represent many more people because, just like phone numbers, an IP address might represent an individual or an entire organisation.
- In September 2011, the data storage capacity of EMBL-EBI was 14 petabytes (PB).

## What is EMBL?

The European Molecular Biology Laboratory (EMBL) is at the forefront of innovation in life sciences research, technology development and transfer, and provides outstanding training and services to the scientific community in its member states. This publicly funded, non-profit institute is housed at five sites in Europe whose expertise covers the whole spectrum of molecular biology.

**EMBL research units**: Heidelberg, Germany (EMBL headquarters); EMBL-EBI in Hinxton, UK; Grenoble, France; Hamburg, Germany; and Monterotondo, Italy.

**EMBL member states**: Austria, Belgium, Croatia, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom. Associate member state Australia.

## How we're funded

As part of the European Molecular Biology Laboratory, the largest part of our funding comes from the governments of EMBL's member states. The global importance of our work is reflected in the fact that we also attract significant funds from external sources, including some beyond Europe.

Some of our other major funders include the European Commission, Wellcome Trust, UK Research Councils, US National Institutes of Health and our Industry Programme partners. The UK's Biotechnology and Biological Sciences Research Council (BBSRC) has awarded major grants in support of EMBL-EBI's planned role as the central hub of ELIXIR. These funds have been used to develop a robust compute infrastructure that will allow EMBL-EBI to efficiently manage an ever-expanding and diversifying range of services.

## Leadership and guidance

Professor Janet Thornton is the Director of EMBL-EBI. Janet also heads the EBI's Research Programme. After playing a key role in EMBL for 30 years, Associate Director and founding EMBL-EBI member Graham Cameron will retire in 2012. Rolf Apweiler and Ewan Birney will then take on stewardship services in April 2012.

EMBL-EBI has an established guidance structure in the form of a Bioinformatics Advisory Committee (BAC). The Committee gives advice to the institute with regard to scientific strategy, future directions and proposals on the realisation of its programme.

## Contact us

If you would like to speak with someone on our news team, please call one of the numbers below. An mail to contactpress@ebi.ac.uk will reach Mary Todd Bergman, Katrina Pavelin, Cath Brooksbank and others in the EMBL Office of Communications and Public Affairs.

Mary Bergman, EMBL-EBI Press Office: +44 (0)1223 494 665

Katrina Pavelin, EMBL-EBI Press Office: +44 (0)01223 494 452

Sonia Furtado Neves, EMBL Press Office: +49 6221 387452

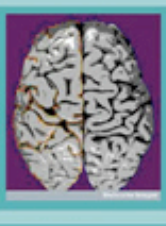
EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambs CB10 1SD, United Kingdom
Tel: +44 1223 494 444

Web: www.ebi.ac.uk/information
Twitter: @EMBLEBI
Facebook: /EMBLEBI

**What's a petabyte?**

# How many petabytes?

(1 000 000 000 000 000 bytes)

| | | |
|---|---|---|
| 1 PB | 20 million four-drawer filing cabinets filled with text |  |
| 2.5 PB | Storage capacity of the adult human hippocampus in binary data equivaent |  |
| 12 PB | Amount of extra storage space Apple has reportedly ordered to host video on its iTunes store |  |
| 14 PB | **EMBL-EBI's current data storage capacity** |  |
| 15 PB | Amount of data produced by the large hadron collider each year |  |
| 20 PB | Amount of data processed by Google each day | |

(all data comparisons here are approximate)

EMBL-EBI