

EMBL-European Bioinformatics Institute
in a nutshell



Contents

The hub of bioinformatics in Europe	4
Who we are	6
Where we are	7
How we're funded	7
Services	8
Research	10
Training	12
Supporting industry	13
How to find us	14



The hub of bioinformatics in Europe

As we move towards understanding biology at the systems level, access to large data sets holding many different types of biological information has become crucial. However, the extensive use of next-generation sequencing and other high-throughput technologies also provides challenges in coping with the volume of data generated. It is vital that this information is collected, stored and curated in ways that allow its efficient retrieval and exploitation. The European Bioinformatics Institute (EBI), which is part of the European Molecular Biology Laboratory (EMBL), is one of the few places in the world that has the resources and expertise to undertake this important task.

EMBL-EBI grew out of the EMBL's pioneering work to provide public biological databases to the research community. Since being established in Hinxton in 1994, EMBL-EBI has been a leader in the bioinformatics revolution. We have diversified to provide data resources in all the major molecular domains, expanded to include a broad research base, developed unique ways of supporting our users both in industry and academia, and created a user-training programme to enable researchers to make the most of the EBI's data resources and tools.

Strength through collaboration

All our major data resources are the products of international collaborations. We work with other data providers to ensure that our data repositories, and those of our collaborators, are comprehensive and up to date. For example:

The European Nucleotide Archive (ENA) is produced as part of the International Nucleotide Sequence Database Collaboration involving GenBank in the USA and the DNA Databank of Japan.

The protein sequence and annotation data in the UniProt databases is maintained through the collaborative efforts of three centres, the EBI, the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR) at Georgetown University Medical Center.

The PDBe, our macromolecular structure database, represents the European arm of the worldwide Protein Data Bank (wwPDB), which also includes the Research Collaboratory for Structural Bioinformatics (RCSB) and the Biological Magnetic Resonance Data Bank (BMRB) in the United States and the Protein Data Bank of Japan (PDBj). These organisations act as deposition, data processing and distribution centres for macromolecular structure data.

The ArrayExpress Archive, a database of functional genomics information, imports data on a weekly basis from the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) in the USA.

Our mission

To provide freely available data and bioinformatics services to all facets of the scientific community in ways that promote scientific progress

To contribute to the advancement of biology through basic investigator-driven research in bioinformatics

To provide advanced bioinformatics training to scientists at all levels, from PhD students to independent investigators

To help disseminate cutting-edge technologies to industry

We also actively participate in international efforts to develop data standards, including the MIAME standard, which specifies the minimum information required to describe a microarray experiment, and the Human Proteome Organisation's Proteomics Standards Initiative (PSI), which develops data standards for molecular interactions, mass spectrometry, protein separations, protein modifications and protein informatics. Europe has always been at the forefront of bioinformatics research, but as we move towards the European Union's goal of a single European Research Area, there is a greater need than ever for bioinformatics experts and experimental biologists throughout Europe to work together towards common goals that will expedite biological research. To this aim, EMBL-EBI coordinates the preparatory phase of the ELIXIR project and several other EU-funded projects:

ELIXIR's objective (www.elixir-europe.org) is to secure funding commitments from government agencies, charities, industry and intergovernmental organisations throughout Europe, to strengthen and sustain a world-class infrastructure for the management and integration of information in the life sciences. ELIXIR's preparatory phase is funded from the European Commission's Framework 7 Capacities Programme for Research Infrastructures. ENFIN (www.enfin.org) brings together experimental and computational biologists to develop the next generation of informatics resources for systems biology. IMPACT (www.ebi.ac.uk/impact) is an EU-funded infrastructure project to improve protein annotation by developing and enhancing InterPro and its contributing databases.

The SLING Integrating Activity (www.sling-fp7.org) aims to serve life science information for the next generation by making available a comprehensive range of databases and services, enhancing them to keep pace with changing science, and providing user training.





The European Molecular Biology Laboratory (EMBL) is a basic research institute funded by public research monies from 20 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, and associate member state Australia. Research at EMBL covers the spectrum of molecular biology. The Laboratory has five units: the main Laboratory in Heidelberg, and Outstations in Hinxton (EMBL-EBI), Grenoble, Hamburg and Monterotondo. The cornerstones of EMBL's mission are: to perform basic, investigator-driven research in molecular biology; to train scientists, students and visitors at all levels; to provide facilities and services for the scientific community; and to actively engage in technology transfer activities.





Janet Thornton, Director

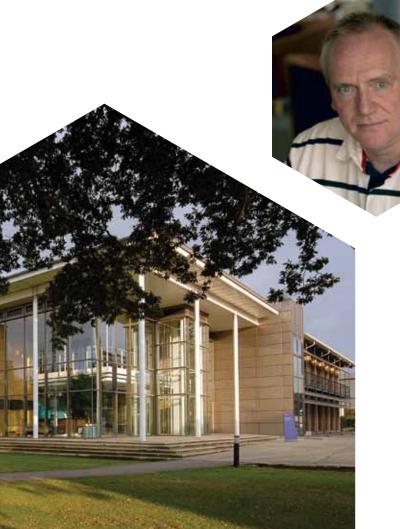
Janet has been Director of EMBL-EBI since
October 2001. Her research group focuses
on understanding biological processes from
a structural perspective using computational
approaches. After a physics degree, she moved
into biophysics at the National Institute for Medical
Research, Mill Hill, London. Before moving
to EMBL-EBI she held a joint appointment at
University College London and the Bernal Chair
in the Crystallography Department at Birkbeck
College. She is a Fellow of the Royal Society, a
Member of EMBO and a foreign associate of the
US National Academy of Sciences.

Who we are

EMBL-EBI has a dynamic and cosmopolitan atmosphere; our work is highly interdisciplinary and our staff members originate from all over the world. EMBL-EBI's combination of service provision and research offers unique advantages; our commitment to our users ensures that we enjoy access to state-of-the-art computational facilities and systems support in combination with the informal and intellectually stimulating atmosphere of a world-class academic research institute.

Graham Cameron, Associate Director

Graham has been Associate Director of EMBL-EBI since September 2001. He was responsible for EMBL-EBI's launch in 1993 and ran its information services until 1998, when he became EMBL-EBI's joint head (with Michael Ashburner). Previous to that he worked at EMBL-Heidelberg in the DNA Data Library, first in software and database design, then leading the project from 1986. His educational background is in psychology, and he has experience in data provision for many disciplines, including primate behaviour, health information, social behaviour and British economic behaviour.



Where we are

EMBL-EBI shares the Wellcome Trust Genome Campus in Hinxton, near Cambridge (UK), with the Wellcome Trust Sanger Institute and the Wellcome Trust Conference Centre. Set in 22 hectares of parkland on the banks of the river Cam, the campus is one of the world's foremost centres for genomics and bioinformatics. We have the advantage of being housed in modern buildings in a beautiful rural setting, whilst being only a few miles away from the academic centre of Cambridge.

The campus provides an exceptionally stimulating environment in which to do top-quality science. It is regularly visited by some of the greatest minds in biomedical research, with workshops, seminars and conferences held on campus open to EMBL-EBI staff. We have excellent computational facilities and a well-stocked library. The campus also enjoys a highly active social life. The sports and social club hosts events for all tastes, and there are inter- and intra-institute teams for a wide range of sports, including football, basketball and volleyball. We have a gym, a choice of places to eat, and free transport by shuttle bus to Cambridge, Saffron Walden and some of the surrounding villages.



Wellcome Trust, 4.3 M€ EBI Industry Programme, 0.6 M€ UK Research Councils, 2.2 M€ European Commission, 6.8 M€

How we're funded

EMBL-EBI attracted over €39 million in internal and external funding for 2009. The global importance of our work is reflected in the fact that we attract significant funds from external sources, including some beyond Europe. As part of EMBL, the largest part of our funding comes from the governments of EMBL's member states. Other major funders include the European Commission, Wellcome Trust, UK Research Councils, US National Institutes of Health and our industry partners.

The UK's Biotechnology and Biological Sciences Research Council (BBSRC) awarded a further €11.4m in August 2009 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 the EBI to efficiently manage an ever-expanding and diversifying range of services.

We are grateful for the support of our funders in allowing us to continue and expand our work.

Genomes

Services

Browse and compare vertebrate (and selected invertebrate) genomes in **Ensembl**, which is produced through a collaboration between EMBL-EBI and the Wellcome Trust Sanger Institute. Completely sequenced genomes, including those of prokaryotes and plants, are available through the five **Ensembl Genomes** interfaces. Find information on individual human genotypes in the **EGA** (European Genome-phenome Archive).

Nucleotide sequence

Find any DNA or RNA sequence in the public domain in the **ENA**, including completed genomes and sequences associated with patents.

EMBL-EBI maintains the world's most comprehensive range of molecular databases. We are the European node for globally coordinated efforts to collect and disseminate biological data. Many of our databases are household names to biologists – they include the ENA (DNA and RNA sequences), Ensembl (genomes), ArrayExpress (microarray data), UniProt (protein sequences), PDBe (macromolecular structures), IntAct (protein–protein interactions), Reactome (pathways) and CiteXplore (our portal to the scientific literature). The details of each database vary, but they all uphold the same principles of service provision [see box].

Our data resources are free

We are the custodians (not the owners) of biological data provided by the community, and progress in biological research depends on unrestricted open access to these data. Our website (www.ebi.ac.uk) provides direct access to all our data resources and tools, from quick links on the main page or from pull-down menus. This is complemented by our search engine, the EB-eye, which allows you to search all of our core databases at once. You can refine your search, confine the search to specific data resources, and move seamlessly from one data resource to another. We provide tools that allow you to submit your own data, perform complex queries across multiple databases, analyse data and view the results in different ways. You can also download data and software from ftp://ftp.ebi.ac.uk, and access a growing number of our resources programmatically using web services.

Supporting our users

We want to make sure that all our users, who are responsible for over 3.5 million requests on our website each day, have all the information they need to make the most of our resources. If you need some guidance on how to make the most of a service, we provide help pages at www.ebi.ac.uk/help as well as the 2Can user support portal at www.ebi.ac.uk/2can. You can also e-mail our helpdesk at support@ebi.ac.uk and we'll get back to you within two working days.

Gene expression

Access annotated microarray experiments in the ArrayExpress Archive, our MIAME-compliant gene expression database. The Gene Expression Atlas contains a subset of curated and re-annotated data from ArrayExpress, which can be used to search gene expression under different biological conditions.

Protein families, domains and motifs

Perform integrated searches for protein families, motifs and domains with InterPro, which combines the strengths of 11 member databases to yield a powerful resource for large-scale protein classification and annotation.

Protein-protein interactions

Access information on protein–protein interactions, with supporting evidence, in IntAct. Build graphical views of interaction networks and annotate them with functional information. IntAct is also MIMIx compliant, allowing researchers to submit their molecular interaction experiments in a format that complies with agreed community standards.

Principles of service provision

Accessibility – Our data and tools are freely available to the research community without restriction. However, in the case of potentially identifiable human genetic information, data access for appropriately authorised researchers is determined by research consent agreements.

Compatibility – EMBL-EBI has possibly done more than any other organisation in the world to promote the adoption of standards in bioinformatics; the development of these standards promotes data sharing.

Literature

Search the scientific literature and automatically add links to biological data resources using **CiteXplore**.

Key

Molecular components

Molecular behaviours

Other

Protein sequence

Use UniProt – the world's most comprehensive catalogue of information on proteins – to gain user-friendly access to richly annotated information on individual proteins, and to full proteome sets.

Proteomes

PRIDE is a centralised, standards compliant, public data repository for proteomics data, including information on post-translational modifications.

Protein structure

Find any of the Protein Data Bank's structures in the PDBe. View predicted quaternary structures, interactions with ligands and structure comparisons.

Chemical entities

Bridge the gap between the protein world and that of small molecules with ChEBI; explore structure–activity relationships of drugs and drug-like molecules with ChEMBL.

Pathways

Navigate a map of human biological pathways, ranging from metabolic processes to hormonal signalling, in **Reactome**.

Comprehensive data sets – Where several publicly available repositories exist, we have negotiated data-sharing agreements to ensure that our resources are comprehensive and up to date. We also negotiate with publishers to ensure that, wherever practicable, biological data are placed in a public repository as part of the publication process and cross-referenced in the relevant publication.

Portability – If practical our data sets are made available for download from the EMBL-EBI website. In many cases the entire

software system can be downloaded and installed locally.

Quality – Our databases are enhanced through annotation: features of the objects stored in them are extracted from other sources, defined and interpreted. Much of our annotation is performed by highly qualified biologists, and the automated annotation that we do is subjected to rigorous quality control. In many cases we also call on expertise outside the EBI for specialist annotation.

Research

EMBL-EBI provides a unique environment for bioinformatics research. We have a broad palette of research interests that complement our data resources, and these two strands of activity are mutually supportive. Although scientists worldwide have access to our data resources, being part of Europe's main centre for biological data provision offers clear advantages for bioinformatics research. EMBL-EBI researchers have the expertise of hundreds of bioinformaticians close at hand - even in the lunch queue or on the volleyball pitch. Our research teams aim to understand biology through the development of new approaches to interpreting biological data. The services teams, who develop and maintain our data resources, also perform research. Usually the focus of this research is to create new services and enhance our existing ones. Here we summarise the EBI's research activities performed in both the research and services groups.

Rolf Apweiler:

Protein data resources

Rolf's team is responsible for the protein and proteomics-related activities of the PANDA group, which include the production of protein sequence and protein family databases at EMBL-EBI (such as UniProt). The group's research activities involve developing methods to improve exploration of large biological datasets and improving the automatic annotation of UniProtKB/TrEMBL and integration with other resources.

Paul Bertone:

Differentiation and development

Paul's group investigates the cellular and molecular processes underlying mammalian stem cell differentiation and induced pluripotency. Using a combination of experimental and computational approaches, they are generating large-scale functional genomic and proteomic data to define the global regulatory influence of key transcription factors, signalling genes and non-coding RNAs in embryonic stem cells and tissue-specific cancer stem cells.



Ewan Birney:

Nucleotide data resources

Ewan's team is responsible for the nucleotide sequence resources of the PANDA group, which includes the ENA, Ensembl and Ensembl Genomes. The group's research activities are synergistic with the nucleotide service projects and focus on algorithmic methods for genome analysis, including new methods for generating genome assemblies and the broader scale evolution of regulatory regions and genetic epistasis models.

Alvis Brazma:

Functional genomics

Alvis's team is responsible for the development of the ArrayExpress Archive and Gene Expression Atlas resources for gene expression and related functional genomics data. In addition, the group's research focuses on the integration and meta-analysis of different data sets, including genomics and medical data. The goal of this analysis is to understand and model basic biological processes, including those that inform the development of diagnostics. Other research activities include high-throughput data analysis and development of algorithms for systems biology and biomedical informatics research.

Anton Enright:

RNA genomics

Anton's group has a particular interest in the function of small non-coding RNAs including microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs). The lab develops computational tools, systems and algorithms for functional genomics of small RNAs. The group collaborates extensively with experimental labs in order to build better data sets for helping us to understand these important regulatory molecules. The group also has interests in the visualisation of large data sets, graph-clustering and analysis of biological networks.

Paul Flicek:

Vertebrate genomics

Paul's team is a combined service and research group that brings together the tools and resources of the Ensembl genome analysis framework for comparative genomics, variation and functional genomics. The team also develops the large-scale bioinformatics infrastructure for human variation projects such as the 1000 Genomes Project and European Genome-phenome Archive (EGA) as a first step towards annotating all human variation. The team's research focuses on primary and integrative data analysis for computational epigenetics and the evolution of genome regulation using comparative functional genomic methods.

Nick Goldman:

Evolutionary analysis of sequence data

Nick's group is developing improved methods for analysing DNA and amino acid sequences to study evolution. The team uses information from evolutionary studies to increase the accuracy of other data analyses in bioinformatics, including the use of next-generation sequencing technology. They also focus on devising better mathematical models to describe molecular sequences and sequence evolution. By seeing which biological features add most to the accuracy of the models, they hope to find out which proposed evolutionary forces are responsible for the patterns of sequence variation within and among species.

Nicolas Le Novère:

Computational systems biology

Nicolas's group aims to gain a better understanding of signal transduction in neurons by building computational models. A strong focus is the molecular and cellular basis of neuroadaptation in neurons of the basal ganglia. Modelling requires various types of software, including design environments and simulators. Nicolas has been instrumental in the development of Systems Biology Markup Language (SBML), which is designed to facilitate the exchange of biological models between different types of software. His group also hosts the BioModels database, a repository for annotated biological models.

Nick Luscombe:

Genome-scale analysis of regulatory systems

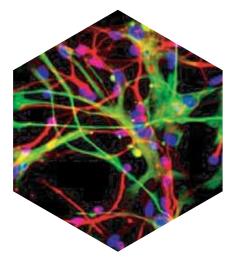
Nick's group uses computational techniques to analyse transcriptional regulatory systems on a genomic scale. Much of their work focuses on the regulatory system in yeast. By integrating diverse information – from its genome sequence to the results of functional genomics experiments – they can study the regulatory system of an entire organism. They have also expanded their interests to understanding regulation in enterobacteria and humans.

John Marioni:

Evolutionary genomics

John's group is motivated by a desire to understand how changes in genetic architecture (e.g. the relationship between sequence and gene expression) provide insights into the evolution of primates and other mammals. Recently, technological developments have led to enormous increases in the amount of data that can help us identify and interpret these changes. With this in mind, John's group develops statistical methods that will exploit data from these new technologies to the fullest extent. By doing this, and through close collaborations with wet lab researchers, the group hopes to use functional genomics data to answer pertinent questions in evolutionary biology.





Dietrich Rebholz-Schuhmann:

Automatic extraction of facts from the scientific literature

Dietrich's group performs research on text-mining methods to automatically extract information from the scientific literature and semantically align the results with the content of EMBL-EBI's databases. This work contributes to the UK PubMed Central and the CALBC (Collaborative Annotation of a Large-scale Biomedical Corpus) projects. The best-performing technical solutions are made freely available as research tools, such as Whatizit, a modular text-mining infrastructure. The group also works towards the automatic generation of ontological resources and fact databases. Collaborations with ChEBI, ArrayExpress and other groups supports the development of these projects.

Julio Saez-Rodriguez:

Modelling signal transduction

Julio's group is interested in how the dynamics of signal transduction ultimately influence cell fate decisions. The group builds predictive mathematical models using high-throughput experimental data collected after applying many different perturbations to the pathways of interest to get at the underlying network structure. Specifically, the group's research aims to combine statistical methods with models describing the mechanisms of signal transduction either as logical or as physico-chemical systems. They then use these models to better understand how signalling is altered in human disease and predict effective therapeutic targets.

Janet Thornton:

Proteins – structure, function and evolution

Janet's research group aims to understand more about how biology works at the molecular level. They analyse how enzymes perform catalysis, how these molecules recognise their cognate ligands, and how proteins and organisms have evolved to create life. They develop and use novel computational methods, gathering data either from the literature or by mining the data resources, to answer specific questions. Much of their research is collaborative, involving either experimentalists or other computational biologists.

Training

Training future bioinformaticians

Bioinformatics is a rapidly expanding discipline that touches on all areas of biology – pure and applied. EMBL-EBI aims to enable bioscientists from all disciplines, regardless of whether they have any formal training in computational biology, to harness the power of bioinfomatics.

Graduate training: We have a lively graduate community that is part of EMBL's International PhD Programme. Students can graduate from the university of their choice. Many of EMBL-EBI's PhD students join the University of Cambridge; they participate fully in Cambridge graduate life and leave EMBL-EBI with a PhD from the University of Cambridge. We encourage potential candidates to visit www.ebi.ac.uk/training/studentships/for more information.

Postdocs: Postdoctoral fellowships are available in our research groups. We are classed as an 'academic analogue' in the UK, and can apply to the Wellcome Trust and the UK research councils for funding for postdoctoral fellowships. We encourage potential candidates to contact the group leader of their choice to discuss the possibility of doing postdoctoral research at EMBL-EBI. Our postdocs can choose to take part in the EMBL Interdisciplinary Postdocs (EIPOD) Initiative. EIPODs promote research projects bringing together scientific fields that are usually separate or involve transferring techniques to a novel context. To find out more about the programme, please visit www.embl.de/training/postdocs/eipod/. We are also launching the EBI-Sanger Collaborative Postdocs (ESPOD) programme to encourage close collaboration between scientists at EMBL-EBI and Wellcome Trust Sanger Institute with the aim to further our understanding of molecular biology and disease. The first fellowships will begin in late 2010.

Visitors: There are opportunities for researchers at all stages of their career to spend time at the EBI. See www.ebi.ac.uk/training/Visitors_Programme/ for more information.

Open days: We hold open days so that early-stage researchers from institutions across Europe can gain direct experience of career opportunities at EMBL-EBI. Information about these and other EMBL-EBI events is available at www.ebi.ac.uk/training.

User training

Our user-training programme aims to equip biomedical researchers with the skills to make the most of their data. We have a three-pronged approach to user training:

Hands-on training at EMBL-EBI

This programme, held in our purpose-built IT training suite, is especially designed to give experimental biologists expert tuition that will help them to get to grips with their data using the EBI's data resources. Our courses, which are typically 3–5 days long, feature: a combination of hands-on and lecture-based tuition; direct contact with expert trainers; high-quality training materials; and a well-equipped training room. See the full programme at www.ebi.ac.uk/training/handson. We also run a range of other workshops on site; see www.ebi.ac.uk/training/onsite.html.

The Bioinformatics Roadshow

As part of the EU-funded SLING Integrating Activity, the EBI coordinates the Bioinformatics Roadshow, in collaboration with the Swiss Institute of Bioinformatics (SIB), the European Patent Office and the BRENDA project. The roadshow provides a mobile training programme covering Europe's most widely used biological data resources. Roadshows are typically held over two days (this is decided by the host) and are modular in nature. The roadshow 'modules' can be mixed and matched to create bespoke courses suitable for the host institute. Learn more about how to host a roadshow at www.ebi.ac.uk/training/roadshow.

eLearning

We also recognise the time and financial restrictions faced by researchers and are developing an eLearning portal to support and complement our face-to-face training programmes. To find out more, contact: elearning@ebi.ac.uk.

Supporting industry

A comprehensive programme for big business

Advances in bioinformatics are having a major impact on industry. We have been supporting the needs of industry for over 12 years through our Industry Programme, which: kick-starts research of importance to industry; provides expert training; develops bioinformatics standards; helps its partners with technical development; and provides regular networking opportunities. The programme is driven by its partners and membership is by invitation. Our industry partners include large multinational companies in the biotech, pharmaceutical, agricultural, nutrition, personal-care and medical devices industries. The programme is funded by annual member subscriptions.



Support for small companies

In addition to supporting large companies, EMBL-EBI also seeks to serve small-to-medium enterprises (SMEs; www. ebi.ac.uk/industry/SME). The SME Support Forum aims to provide beneficial support at minimal cost to participants. The Support Forum meets on an annual basis, offering discussion and networking opportunities. Forum members also have access to: expert tuition; opportunities for technical development through collaboration; and consultancy on an ad hoc basis. Members of the SME Support Forum range from drug discovery and biotech start-ups to bioinformatics service providers.

For more information on how your company can benefit from involvement in the Industry Programme, visit the Industry Support pages at www.ebi.ac.uk/industry/ and download the programme brochure at www.ebi.ac.uk/Information/Brochures/. To learn more about joining the programme, please contact Dominic Clark, the Industry Programme Manager: clark@ebi.ac.uk.



Industry Partners

AstraZeneca

Bayer Schering Pharma

Boehringer Ingelheim

Eli Lilly & Company

Galderma

GlaxoSmithKline

F. Hoffmann-La Roche

Johnson & Johnson Pharmaceutical Research & Development

Merck Serono

Nestlé Research Centre

Orion Pharma

Philips Research

Pfizer Ltd

Sanofi-Aventis

Syngenta Ltd

Unilever

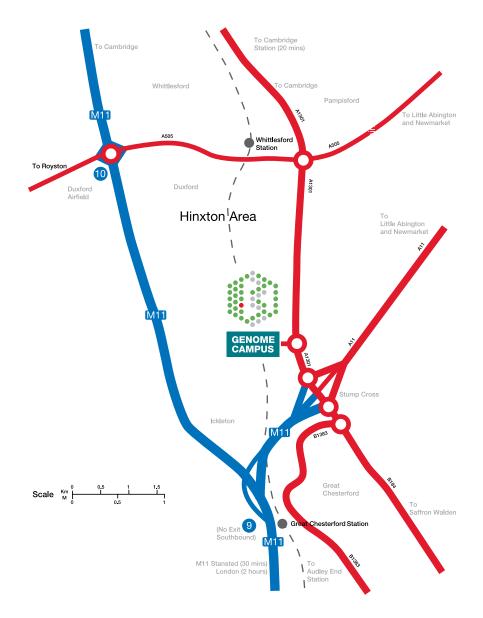
Joint pre-competitive initiatives

In addition to pre-competitive projects run through the Industry Programme, we have strong collaborations with industry to help stimulate pre-competitive research. These projects aim to improve the translation of research discoveries into advances in medicine, health and agriculture for the benefit of society. For example, EMBL-EBI is a member of the Pistoia Alliance (www.pistoiaalliance.org). This is a not-for-profit organisation established by life science companies with the aim to streamline non-competitive elements of the pharmaceutical drug discovery process. The Innovative Medicines Initiative (IMI; http://imi.europa.eu) is a joint initiative to foster pre-competitive collaboration within the pharmaceutical industry and is co-sponsored by the European Commission and EFPIA (European Federation of Pharmaceutical Industries and Associations). EMBL-EBI is a partner on two projects that are funded through the first call of the IMI. These are eTox (www.etoxproject. eu/) and EMTRAIN (www.emtrain.eu/).

How to find us



EMBL-European Bioinformatics Institute Wellcome Trust Genome Campus Hinxton, Cambridge CB10 1SD, UK





EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambridge
CB10 1SD, UK

Tel: +44 (0) 1223 494444 Fax: +44 (0) 1223 494468 e-mail: support@ebi.ac.uk web: www.ebi.ac.uk