



A two-day dip into the EBI's data resources - Understanding your data 10-11 September 2007

EMBL-EBI 

Contact: Vicky Schneider, Scientific Training Officer e-mail: vicky@ebi.ac.uk T: +44 (0)1223 494448
Janet Copeland, Workshops and Exhibitions Organiser e-mail: copeland@ebi.ac.uk T: +44 (0)1223 492510

Course description

As we move towards understanding biology at the systems level, access to large data sets of many different types has become crucial. Technologies such as genome-sequencing, microarrays, proteomics and structural genomics have provided 'parts lists' for many living organisms.

The EMBL-EBI maintains the world's most comprehensive range of molecular databases. This course provides a broad introduction to EBI core resources and practical, trainer-guided demonstration of when, why and how to use them. Several resources will be covered spanning genomics to systems biology.

Specific resources presented:

EB-eye search engine

The EB-eye search system is a high-performance, full-featured text search engine that provides rapid access to the EBI's data resources. Users can search globally across all EBI databases or narrow their selection to individually in selected resources by using the advance search option.

EBI Literature resources

CiteXplore is the EBI's portal to the scientific literature, directing users to full text articles as well as relevant biological databases. It provides access to MEDLINE (NLM), Agricola (NAL), patent abstracts (EPO), Chinese Biological Abstracts (CBA) and CiteSeer (UPenn).

Ensembl (genomes): www.ensembl.org

A genomic sequence database providing a comprehensive and integrated source of annotation of mainly vertebrate genome sequences.

ArrayExpress (microarray data):

www.ebi.ac.uk/arrayexpress

EMBL-EBI's MIAME-compliant repository for gene expression data from microarray and transcriptomics experiments (MIAME - Minimum Information About a Microarray Experiment).

UniProt (protein sequences): www.uniprot.org

UniProt is a collaboratively maintained central repository of protein sequence and function. Users can access curated protein information through a central access point, the UniProt Knowledgebase (UniProtKB).

MSD (macromolecular structures): www.ebi.ac.uk/msd

The EBI's Macromolecular Structure Database (MSD) is Europe's representative in the worldwide Protein Data Bank (www.wwpdb.org). It contains structural data from crystallography, nuclear magnetic resonance spectroscopy and electron microscopy experiments, in addition

to derived information linking structure to biochemical function, protein sequence and taxonomy.

IntAct (protein–protein interactions): www.ebi.ac.uk/intact

IntAct is a molecular interaction database and analysis system. The project is a member of the International Molecular Exchange (IMEx) consortium for exchanging molecular interaction data and supports the development of the Human Proteome Organisation (HUPO) community standard data model.

PRIDE (proteomics): www.ebi.ac.uk/pride

The Proteomics Identifications Database is a public data repository for proteomic identifications and associated mass spectra, linked to supporting publications.

Reactome (pathways): www.reactome.org

Reactome is a curated knowledgebase of biological processes, containing reactions, pathways, proteins, DNA, small molecules and complexes. Reactome encompasses metabolism, signalling, cell cycle, apoptosis, gene expression, pathogen–host interactions and much more.

Understanding your data

In addition to presenting practical information on the range of EBI resources and how to use these tools, this course aims to illustrate their applicability to research by using a range of case studies. In this way, use of the resources is presented with reference to biological questions.

In order to tailor the course content to delegates' requirements and best meet your needs, it is vital that we have prior knowledge of your familiarity with the resources presented (or similar), the context under which you currently use these resources and your expectations of the training. The online registration form (accessed from www.ebi.ac.uk/training/hands-on) contains fields for this information and we ask you to be as detailed as possible.

Trainer profiles

David Croft is a bioinformatician and has been giving talks and training for the Reactome project since 2005. He is also involved in maintaining and updating the Reactome website. Prior to joining the EBI, he spent seven years in industry, designing bioinformatics and cheminformatics applications.

Phil Jones is the technical lead software engineer for the PRIDE project. He has a background in both biological sciences and software engineering, and holds a Post-Graduate Certificate in Education. He has held positions in software engineering and in science education. He has given courses and talks on various software development projects in numerous locations worldwide.

Sandra Orchard is a Senior Scientific Database Curator working across the Sequence Database, InterPro and Proteomics Services teams. She is responsible for the curation standards within the IntAct Molecular Interaction database as well as contributing to the annotation of the UniProtKB, InterPro and GOA databases and applies her experience to provide hand-on training in several resources including UniProtKB, IPI, InterPro and IntAct.

Gabriella Rustici is a research scientist in the Microarray Informatics group at the EBI and her research focuses on meta-analysis of microarray datasets. She is also involved in training activities and applies her experience to provide hands-on training in several resources including ArrayExpress and Expression Profiler.

Vicky Schneider is the EBI's Scientific Training Officer and responsible for developing and coordinating the in-house and external training activities of the EBI. She has an academic background in behavioural ecology and adaptive evolution. Prior to joining the EBI, Vicky held an Assistant Professor position at the University of Bern, Switzerland.

Giulietta Spudich is an outreach and training officer for Ensembl. She is also involved in providing technical support for the genome browser. She has a background in academic research (structural biology and biochemistry) and has taught lecture-style classes and day-long workshops for University of California, Berkeley (USA), University of Cambridge (UK) and gives world-wide workshops on Ensembl.