

# Bioinformatics We

EBI - NCBI

June 12-16, 2011

UCSC - Stanford

## Course Descriptions

[Download PDF](#)

### EBI Overview

#### **Description**

The European Bioinformatics Institute (EMBL-EBI) in Hinxton, UK, grew out of the work by the European Molecular Biology Laboratory (EMBL) to provide biological data to the research community. Today EBI is a centre for research and services in bioinformatics, housing a wide variety of databases of biological data, including genomic, proteomic, biochemical, structural, interaction, and expression data, amongst others. It provides a toolbox of bioinformatics software and search engines with which to collate the wealth of biological data. The EBI training workshop will provide a grounding in a selection of fundamental EBI genomic and proteomic database and appropriate analysis software.

### EBI - ArrayExpress

#### **Description**

The Microarray team at the EBI provides public resources for storing, managing and analysing microarray data, and is a public, online repository for microarray data that can be accessed via the Microarray Minimum Information for Experiment (MIAME) standard.

[Bioinformatics Week?](#)

[Help?](#)

[Publications](#)

[Workshops](#)

[Events](#)

[Bioinformatics](#)

[Courses](#)

Microarray Minimum Information for Microarray Experiment (MIAME) standard stores both normalized and well-annotated raw data from 40,000 hybridizations from more than 70 species. Data can be submitted on-line, and password protection to your data is provided for authors and reviewers. This course will examine how to use ArrayExpress, deposit data, and give yourself and other microarray community members access to your data at the level of the experiment or the gene. ExpressionProf, a software package written to conduct online microarray data analysis, will also be introduced.

### **Audience**

No prior experience is required.

### **Format**

The course consists of a 1-hour lecture OR a 2-hour hands-on workshop which covers the content of the lecture (see below). Lecture attendees may bring their wireless computer laptops and follow the lecturer, as conditions permit.

- **Lecture**

Registration for the lecture is unrestricted and open to all Stanford University staff, (including visiting faculty and staff) and graduate students. An open office hour period with EBI instructors for lecture participants will be available.

- **Hands-on Workshop**

The workshop duplicates the content of the lecture, but differs in that it focuses on practice of the lecture demonstration examples. Individual consultation with an instructor will be available to workshop participants immediately following the workshop.

**Workshop computer space is limited to a single session due to limited availability! Register early. Registration will be "first come - first served". A waiting list will be maintained if necessary.**

### **Registration**

Registration is required for either the ArrayExpress lecture or workshop.

### **Instructors**

[Helen Parkinson](#)

### **Times**

[View Schedule](#)

of macromolecular structure data. Since its inception in 1996, the MSD group partners around the world to improve the quality of PDB data, through a clean archive that addresses inconsistencies and inaccuracies in the legacy archive. The improvements in data quality in the legacy archive have been achieved largely through the creation of a new data archive, in the form of a relational database that stores all of the data in a new implementation of the MSD database, together with the parallel development of new tools and methodologies for data harvesting, validation and archival, has led to significant improvements in the quality of data that enters the archive. The MSD has developed new query systems in order to allow users to interact with the database and retrieve consistent data for their own use. In addition, the MSD has also developed new visualization methods to help analysis of structures and improved the deposit process. The latest release of AutoDep4.0 (<http://www.ebi.ac.uk/msd-srv/autodep4/>), an archival and web-based structure deposition tool. This workshop will introduce various services offered by the MSD that can help in solving research problems.

### **Audience**

No prior experience is required.

### **Format**

The course consists of a 1-hour lecture OR a 2-hour hands-on workshop which covers the content of the lecture (see below). Lecture attendees may bring their wireless computer laptops and follow the lecturer, as conditions permit.

- **Lecture**

Registration for the lecture is unrestricted and open to all Stanford University staff, (including visiting faculty and staff) and graduate students. An open period with EBI instructors for lecture participants will be available.

- **Hands-on Workshop**

The workshop duplicates the content of the lecture, but differs in that it focuses on practice of the lecture demonstration examples. Individual consultation with an instructor will be available to workshop participants immediately following the workshop.

**Workshop computer space is limited to a single session due to its limited availability! Register early. Registration will be "first come - first served". A waiting list will be maintained if necessary.**

### **Registration**

Registration is required for either the ArrayExpress lecture or workshop.

### **Instructors**

[Sameer Velankar](#)

### **Times**

[View Schedule](#)

## EBI - Reactome

### Description

Reactome is a collaborative project between the EBI and the Cold Spring Harbor Laboratory, USA, consisting of a curated knowledgebase of biological processes described by reactions, pathways, macromolecules, small molecules, complexes and catalytic activities. These processes include both well-defined metabolic pathways, such as glycolysis and the citric acid cycle, and a diverse range of other pathways, such as signalling pathways and cell cycle regulation. The Reactome web interface provides a textbook-like view of cellular processes, providing tools for keyword searching, integrating expression data and generating protein-protein interaction networks. This course describes the underlying concepts behind Reactome, and explores the features of the web interface using hands-on exercises. In addition, participants will learn through the process of setting up their own Reactome site, and shown how to use the tools available.

### Audience

No prior experience is required. Anyone with an interest in using and/or discussing bioinformatics services at Stanford is encouraged to attend.

### Format

The course consists of a 1 hour lecture followed by a 30 minute open consultation.

### Registration

None.

### Instructors

### Times

[View Schedule](#)

## EBI - Web Services

EBI Web Services, such as SOAP (Simple Object Access Protocol), is an integration that provides information retrieval and data analysis tools for use with multiple databases without the need for manual searching through a web browser or for installing databases in-house. Moreover, information retrieval and analysis are linked, providing a combination and manipulation of search results. Using these technologies, you can build complex applications in their language of choice (such as Java, Perl, VB, Python, Ruby, R and PHP), with data access directly from and to your code. You can submit large and complex jobs straight from your data pipeline with all results returned directly to your software for processing at the next stage. The number of databases at the EBI that have Web Services access is now in double figures, making the adoption of this technology very advantageous for any laboratory involved in genomics research. You can retrieve data from multiple EBI databases using the WSDb and you can make use of sequence similarity tools such as Blast, Fasta and MBLAST to perform multiple sequence alignments using ClustalW, T-Coffee or MUSCLE. You can also explore the InterPro protein families and domains database using InterProScan and perform powerful queries over 40 different biological ontologies via the ontology lookup tool. In this lecture, the EBI Web Services will be demonstrated, providing an overview of what individual Web Services can offer.

**Audience:**

**Format:** The course consists of a 1-hour lecture/demonstration. The workshop is limited to 20 participants or less to maximize interactions between instructor and student. A waiting list will be maintained with one additional session added, if necessary during Bioinformatics Week.

**Registration:**

**Instructor**

**Times**

[View Schedule](#)

EBI - Ensembl / BioMart

macaque, mouse, rat, dog, cow, opossum, chicken, zebrafish, among others) imports the manually curated data from other model organisms to use as core (e.g. Drosophila from FlyBase, C. elegans from WormBase and yeast from SGC). Ensembl provides a unique gene analysis pipeline with each piece of genomic information predicted and supported by biological evidence. Additionally, ESTs, mRNAs, SNPs, proteins and other biological information are mapped to the genomes.

This course is designed to present different data types available in the browser, how to use "Views", how to compare and contrast various information from several sources, and how to get the most out of the Ensembl database with the unique data-mining tool, BioMart. BioMart is a simple, distributed data integration system with powerful genomic-based queries. BioMart can generate a number of different types of output, including sequence alignments and list data. This course will also introduce BioMart, focusing on how it can be used by non-programming, researchers.

### **Audience**

No prior experience is required. Basic familiarity with the steps and processes of BLAST sequence searching is helpful, but not necessary.

### **Format**

The course consists of a 2-hour lecture OR a 4-hour hands-on workshop which covers the content of the lecture (see below). Lecture attendees may bring their wireless computer laptops and follow the lecturer, as conditions permit.

- **Lecture**

Registration for the lecture is unrestricted and open to all Stanford University staff, (including visiting faculty and staff) and graduate students. An open period with EBI instructors for lecture participants will be available Friday, 10/15/04, immediately following the lecture.

- **Hands-on Workshop**

The workshop duplicates the content of the lecture, but differs in that it focuses on practice of the lecture demonstration examples. Individual consultation with an instructor will be available to workshop participants immediately following the workshop. **Workshop computer space is limited to a single session due to limited availability! Register early. Registration will be "first come - first served". A waiting list will be maintained if necessary.**

### **Registration**

Registration is required for either the Ensembl lecture or workshop.

### **Instructors**

[Xose Fernandez](#)

**Times**

[View Schedule](#)

**A Field Guide to NCBI GenBank and NCBI Molecular Biology Resources**

Registration is required for both the Field Guide lecture and workshop.

**Instructors**

[Peter Cooper](#), [Eric Sayers](#)

**Times**

[View Schedule](#)

Gene Ontology - EBI



own GO subsets. Individual consultation with the EBI instructor will be workshop participants immediately following the workshop. **Workshop is limited to a single session due to instructor availability! Registration will be "first come - first serve."** A waiting list will be necessary.

**Registration**

Registration is required for both the Gene Ontology lecture and workshop.

**Instructors**

[Emily Dimmer](#)

**Times**

[View Schedule](#)

**Proteomics Workshop - UniProt, InterPro, IntAct**

## Description

This workshop will introduce participants to three major EBI proteomic databases: InterPro, IntAct and PRIDE – and explore the wealth of information they contain to enhance your research. UniProt is a central repository of protein sequence and by combining the information in Swiss-Prot, TrEMBL and PIR, while InterPro aggregates free open source data and resources aimed at analysing proteomes and interactions. InterPro provides information on the function, annotation and classification of proteins by integrating major signature databases from Gene3D, Pfam, PIRSF, Prints, ProDom, Prosite, SuperFamily, Tigrfams, and PANTHER into a unified protein resource. InterPro also provides structural annotation by mapping individual proteins to PDB, MSD, CATH, SCOP and SwissModel, as well as functional annotation through GO mapping, literature citations, references, and links to several external databases (Pfam, CAZy, Merops, IUPUI, InterPro Cluster, Blocks, and Prosite doc). IntAct provides a comprehensive, annotated database of protein interactions derived from literature curation and through direct submissions. Interactions can be viewed individually, or as part of an interaction network, and can be used to help define the proteins involved in such networks.

## Audience

No prior experience is required. Basic familiarity with the steps and processes of sequence searching is helpful, but not necessary.

## Format

The course consists of a 2-hour hands-on workshop. Lecture attendees may bring their own registered wireless-access enabled computer laptops and follow the lecturer, as permitted.

## Hands-on Workshop

The workshop considerably duplicates the content of the lecture, but differs in the hands-on practice of the lecture demonstration examples. Both workshops are available, choose the one that best fits your schedule. Individual consultation with the EBI instructor is available to workshop participants immediately following the workshop. **Workshop space is limited to a single session due to instructor availability! Registration will be "first come - first serve."** A waiting list will be maintained.

## Registration

Registration is required for either the Ensembl lecture or workshop.

## Instructors

[Sandra Orchard](#)

[Jennifer McDowall](#)

## Times

Last Updated: May 26, 2005