NCBI Mini Courses

National Center for Biotechnology Information ■ National Library of Medicine ■ National Institutes of Health ■ Department of Health and Human Services

NCBI provides several Mini Courses that provide practical introductions to bioinformatic procedures. NCBI Mini Courses are either problem-based such as "Identification of Disease Genes" or NCBI resource-based such as "BLAST Quick Start". The courses are two-and-one-half hours in length with the first hour-and-a-half devoted to an overview that is followed by a one hour hands-on session.

The courses are free and NCBI pays for the travel expenses of the instructors within the United States.

Researchers will learn how to apply NCBI resources to their research. The courses provide online bioinformatics modules for educational use.

Requirements for hosting the Mini Courses

- an audience of at least 15 people
- a multi-workstation computer classroom with internet access
- an instructor's workstation with internet access and projection capabilities

To host a Mini Course at your institute or for additional information:

contact Medha Bhagwat — bhagwat@ncbi.nlm.nih.gov.

www.ncbi.nlm.nih.gov/Class/minicourses

Mini Courses NCBI Resources one course at a time

Course Descriptions

1 Making Sense of DNA and Protein Sequences

In this course, we will find a gene within a eukaryotic DNA sequence. We will then predict the function of the derived protein by seeking sequence similarities to proteins of documented function using BLAST and other tools. Finally, we will find a 3D modeling template for this protein sequence using the Conserved Domain Search.

2 Unmasking Genes in the Human Genome

This Mini Course describes how to combine the output of multiple prediction programs to find genes, promoters and transcription-factor binding sites in human DNA sequences using a program called Greengene. Greengene integrates the output of several gene-finding tools and allows a coding sequence and accompanying protein translation to be assembled from the exons detected by these programs. Because the output of several programs is integrated, exon selection is more reliable.

3 Identification of Disease Genes

This Mini Course deals with the identification of a disease gene using NCBI's human genome assembly. The reference genome assembly, along with integrated maps, literature, and expression information comprises a powerful discovery system for exploring candidate human disease genes.

4 Correlating Disease Genes and Phenotypes

We will learn to determine what is known about a disease and the gene associated with it. We will then elucidate the biochemical and structural basis for the phenotype caused by the mutant protein.

5 BLAST Quickstart

A practical introduction to the BLAST family of sequence-similarity search programs. Exercises range from simple searches to creative uses of the BLAST programs.

6 Entrez Gene QuickStart

NCBI's Entrez Gene provides gene-based information such as chromosome location, sequence, expression, structure, function, and homology data. Each record represents a single gene from an organism. Entrez Gene includes organisms for which there is a RefSeq genome record.

In this course, we will learn how to obtain information about a human gene such as its mRNA and genomic sequence, gene structure (exon-intron locations), function and phenotypes associated with mutations. We will also learn how to determine whether the SNPs in the coding region of a gene are known to alter the function of the protein product .

Entrez Gene is the successor to LocusLink.

Structural Analysis Quickstart 7

This course covers how to visualize and annotate 3D protein structures using NCBI's Cn3D, identify conserved domain(s) present in the query protein, seach for other proteins containing similar domain(s), explore a 3D modeling template for the query protein and obtain distant sequence homologs that may not be identified by BLAST.

Map Viewer Quickstart 8

The Map Viewer can be used to visualize a number of genome assemblies at NCBI. In this course, we will use the human genome Map Viewer to locate a human gene, download its sequence along with its upstream sequence (to analyze promoter regions), obtain exon-intron coordinates, obtain a possible splice variant, and identify whether the variations in the gene are associated with a disease.

GenBank QuickStart 9

GenBank is a repository of nucleotide sequences from a large number of organisms. This course begins with a survey of different types of entries. Using a typical GenBank entry as a model, students will learn to understand the features annotated on it. The course will also cover how to submit sequences to GenBank and include an overview of the processing of the entries. Finally, students will learn how to efficiently search GenBank and download sequences.

Entrez Quickstart 10

Entrez is the integrated, text-based search and retrieval system used at NCBI for the major databases, including PubMed, Nucleotide and Protein Sequences, Expression, PubChem (biological activities of small molecules), Protein Structures, Complete Genomes, Taxonomy, and others. Entrez provides links to related records within the database and between other databases in Entrez.

This course will provide tips on effective searching in Entrez databases and accessing the records in various formats.

Microbial Genomes Quickstart 11

Microbial Genomes Quick Start: In this mini-course, you will learn how to access the microbial genome sequences and annotations, how to navigate through and download the gene and protein datasets, and will be introduced to the available genomic and comparative genomic analysis tools. The course will address practical discovery questions such as 'Are there identifiable genes in microbial genomes that may be horizontally transferred?' and 'What are the differences between closely-related pathogenic and non-pathogenic bacteria?'

NEW

12 The Prism of Bioinformatics and a Spectrum of Photoreceptors

The course addresses aspects of bioinformatics such as accessing, analyzing, and interpreting biological data using NCBI databases and tools. An analysis of the animal photoreceptor family is used to illustrate practical bioinformatics approaches to the study of sequence similarity, phylogenetic analysis, gene expression, homology, polymorphisms, 3-D structure and function.