UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Bioinformatic Methodologies

2122-1-F0802Q054

Aims

The main objectives of the course are to introduce students to the main problems related to the computational analysis of biological sequences (DNA, RNA, proteins) as well as to the synthesis and simulation of mathematical models in biology. Students will acquire the basic knowledge and concepts related to computational methods and techniques for collecting, managing and analyzing data in molecular biology, and will master the main computational tools necessary to extract information of interest for biomedical research from the main sequencing databases.

Contents

1) Introduction to bioinformatics: motivations, problems and methodologies

2) Organization and management of data: the main biological databases; access, query, data entry and data analysis

3) Sequence comparison: global, local and multiple alignment algorithms

4) Building of a phylogenetic tree

5) Structural genomics: mapping, sequencing, annotation, reconstruction

6) Trascription networks: graphs and computational models; transcriptome analysis

Detailed program

1) Basics of Informatics: computer acrhitecture, operating systems, algorithms, programming languages, databases

2) Sequence alignment. Global alignment: similarity score and homology; dot-matrix method; dynamic programming, traceback. Local alignment: substitution matrices (PAM and BLOSUM); Smith & Waterman algorithm; heuristic database searching (BLAST, FASTA). Multiple sequence alignment: progressive alignment (Clustal)

3) Biological databases: Genomic databases (EMBL – GenBank). Protein databases (SwissProt, PDB). Database query systems and bioinformatics platforms (NCBI, UCSC Genome Browser, Galaxy).

4) Building of phylogenetic trees. Distance-based methods: custering (UPGMUA and Neighbor Joining). Character-based methods (maximum parsimony, maximum likelihood)

5) Genome mapping, sequencing (Sanger, Next Generation sequencing, base calling), annotation (FASTA, FASTQ), reconstruction

6) Transcription networks. Transcriptome analysis: microarray, RNA-sequencing. Graph theory to describe transcription natworks: in- and out-degree distribution. Network motifs. Computational models of gene expression and regulation: the negative autoregulation case

Prerequisites

No prerequisites: all concepts from informatics or mathematics are provided in the course

Teaching form

Classroom lectures with slides.

Hands-on sessions on pc for the use of biological databases and bioinformatics platforms.

All lessons will be videorecorded and available on Moodle.

Textbook and teaching resource

Slides. All educational material will be available at the e-learning platform of the course.

Textbooks: Helmer Citterich M, Ferrè F, Pavesi G, Romualdi C, Pesole G. Fondamenti di bioinformatica. Zanichelli, 2018.

Semester

First semester

Assessment method

Students will discuss a scientific paper, previously given. Besides, oral questions on the programme

Office hours

Students are invited to contact the teacher by email to agree upon a date (possibly on WebEx)