UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Bioinformatic Methodologies

2324-1-F0802Q054

Aims

The general objective of the course is to introduce students to the main problems and methodologies related to the computational analysis of biological sequences (DNA, RNA, proteins) to extract information of interest for biomedical research or biotechnology.
Knowledge and understanding
Students will acquire the basic knowledge and concepts related to computational methodologies and techniques for the collection, management and analysis of data in molecular biology, including data generated by single-cell Next Generation Sequencing technologies.
Applying knowledge and understanding
At the end of the course, students will master the main computational tools necessary to extract information of interest from the main biological databases.
Making judgements
Students will have to acquire full awareness of the significance of bioinformatics procedures for research, analysis and data processing; avoiding as much as possible a "black box" approach to using bioinformatics tools (in particular, those freely available online).
Communication skills
Use of an appropriate scientific vocabulary and ability in oral/written reports.

Learning skills
The teaching is not limited to a merely descriptive practical approach, but outlines research projects, appropriately sized to the skills to be acquired, which allow students not only to apply tools, but also to critically evaluate the results, understanding the differences that different types of tools and approaches can have on the data obtained.

Contents

• Introduction to bioinformatics
• Data generation: from sequencing platforms to genome assembly and annotation

• Biological data organization and management
• Comparison of sequences and reconstruction of phylogenetic trees
• Transcriptomic data analysis

Detailed program

1. Introduction to bioinformatics
   a. What is bioinformatics
   b. Review of computer science elements
   c. Review of statistics and probability elements
   d. Machine learning in a nuthsell
2. Data generation
   a. DNA sequencing platforms
   b. From chromatogram to reads: “base calling”
   c. Coverage, reads quality, data formats
   d. From reads to the sequence: genome assembly
   e. Genome annotation
3. Comparison of sequences
   a. Local and global alignment
   b. Exact and heuristic algorithms
4. Biological data organization and management
   a. Databases and DBMS: relational and flat file databases
   b. Biological databases
   i. Genomic databases (GenBank - ENA – DDBJ)
   ii. Proteomic dabases (UniprotKB, Swiss-Prot, TrEMBL – PDB)
   iii. Genome browsers: ENSEMBL, UCSC
5. Phylogenetic trees reconstruction methods
6. Analysis of the transcriptome
   a. From RNA sequencing to gene expression levels
   b. Single-cell RNA sequencing data pre-processing and analysis

Prerequisites

Concepts of cellular and molecular biology, and biochemistry provided in basic courses of bachelors in Biological or Biotechnological sciences will be given for granted.
Previous knowledge that is useful, which will be only briefly revised during the course:
• Elements of computer’s architecture and algorithms
• Elements of probability and statistics

Teaching form

Classroom lectures with slides.
Hands-on sessions on pc for the use of biological databases and bioinformatics platforms.

Textbook and teaching resource

All educational material will be available at the e-learning platform of the course.
Suggested textbook: Citterich, Ferré, Pavesi, Romualdi, Pesole. Fondamenti di Informatica. BIOLOGIA ZANICHELLI
Specialized research articles, surveys and book chapters will be recommended during the course.

Semester

First semester

Assessment method

Individual or group research project on a topic chosen by the student which involves the critical use of some tools seen in class
+
Oral exam in which the student presents the research project and demonstrates mastery of the concepts used in the project. During the oral exam, sequence alignment exercises (to be carried out on paper) will also be proposed.

There are no intermediate tests for attending students.

Office hours

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

Sustainable Development Goals