UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Metodologie Bioinformatiche

2627-1-F0803Q054

Aims

ELearning Objectives
Knowledge:
• To understand the main bioinformatic methodologies for the analysis of biological data, particularly omics data.
• To comprehend the theoretical foundations and algorithms underlying bioinformatic analysis tools.
Skills:
• To use bioinformatic tools for processing and interpreting biological data.
• To select and apply appropriate methodologies based on the specific biological question.

Expected Learning Outcomes
At the end of the course, the student will be able to:
• Knowledge and understanding: describe the fundamental principles of major bioinformatic methodologies, with reference to applications in genomics, transcriptomics, and omics data analysis.
• Applying knowledge and understanding: use software environments and tools for the bioinformatic analysis of real datasets, critically interpreting the results obtained.
• Making judgements: assess the appropriateness of different methodological approaches for specific biological or clinical questions, and propose alternative solutions where needed. This skill is developed through interactive activities in class, including guided discussion of case studies, interpretation exercises based on real data, and practical examples worked out collectively.
• Communication skills: clearly and coherently communicate methodological choices and analysis results, both in written form and through oral presentations.
• Learning skills: independently deepen their knowledge of bioinformatic tools and resources not covered in class, keeping up to date through scientific literature and technical documentation.

Contents

• Introduction to bioinformatics
• Data generation: from sequencing platforms to genome assembly and annotation, to multi-omics
• Biological data organization and management
• Sequence Analysis
• Transcriptomic data analysis
• Multi-omics data integration

Detailed program

1. Introduction to bioinformatics
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. Sequence analysis
   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. Analysis of the transcriptome
   a. From RNA sequencing to gene expression levels
   b. Single-cell RNA sequencing data pre-processing and analysis
   c. Unsupervised and Supervised Learning
6. Multi-omics data integration
   a. Enrichment analysis and integrated visualization of transcriptomic, proteomic, and metabolomic data
   b. Integration of omics data into metabolic network models

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.

Recommended background knowledge:

- Fundamentals of artificial intelligence, including basic machine learning concepts.
- Fundamentals of probability and statistics.
  These topics are covered in the parallel compulsory course Structures and Molecular Interactions and are not addressed as part of the present course. Although concurrent attendance of the two courses is not mandatory, students are strongly encouraged to take them in parallel, as the knowledge acquired in the other course provides the background required for several topics covered in this course.

Teaching form

35 hours of lectures, divided as follows:
– 30 hours in delivered mode (face-to-face), supported by in-class presentations;
– 5 hours in interactive mode, dedicated to the guided development and discussion of students' final project proposals.

10 hours of practical sessions held in person, focused on the use of biological databases and bioinformatic analysis platforms.

The language of instruction is either English or Italian, depending on students' preferences.

Textbook and teaching resource

Slides and video recordings of lectures and practical sessions will be available on the course e-learning page.
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.

Grade composition:
The final grade is composed of the project presentation (4/5) and the alignment exercise (1/5).
Grading scale:
18–19: Project covering a limited number of topics from the course syllabus; numerous gaps in knowledge and limited capacity for critical analysis.
20–23: Project covering some topics from the course syllabus; independent analysis limited to purely practical and execution-related issues; imprecise vocabulary and somewhat uncertain presentation skills.
24–27: Project covering a wide range of topics from the course syllabus; ability to independently develop arguments and critical analysis; ability to apply knowledge to specific contexts and link themes to concrete cases; correct vocabulary and competence in using discipline-specific terminology.
28–30/30L: Comprehensive project and thorough, exhaustive preparation on course topics; ability to independently address and critically analyze themes; capacity for reflection and self-reflection, as well as for linking themes to concrete cases and diverse contexts; excellent critical and independent thinking skills; full mastery of discipline-

specific terminology and a rigorous, articulate presentation style; strong argumentation and reflective skills; ability to draw connections to other disciplines.

Office hours

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

Sustainable Development Goals

GOOD HEALTH AND WELL-BEING