

**** Title ****

Bioinformatics and Systems Biology (BSB)

**** CFU ****

9 fundamental

**** Learning outcomes ****

The course provides students with knowledge and understanding of the main computational methodologies for the analysis of biological data and for the investigation of complex biological phenomena. Students will learn the basics of molecular biology and biochemistry and will become aware about methods for the collection and organization of biological data. Then, they will learn bioinformatics methods for the processing and analysis of sequencing data as well as systems biology methods for the analysis of biological networks such as cell pathways, gene regulatory networks and protein interaction networks. Finally, they will make an experience with a molecular dynamics simulator in which concepts presented in theoretical lessons are exemplified in the context of drug design.

**** Syllabus ****

Based on recommendations from:

“An overview of bioinformatics courses delivered at the academic level in Italy: Reflections and recommendations from BITS”, PLOS Computational Biology, 2023
(<https://doi.org/10.1371/journal.pcbi.1010846>)

1) Biological context (12h):

1. Introduction to Biology: from genomics and molecular biology to biochemistry.
2. Introduction to bioinformatics and basic concepts of molecular biology and biochemistry. Nucleic acid sequencing platforms: the 3 generations. Isolation, purification and identification proteins.
3. Data generation and Analysis: quality of reads, FASTQ and FASTA formats, genome annotation and genomics, proteomics and interactomics data, hints at NGS, Microarray, Mass Spectrometry technologies and protein structure determination techniques.
4. Data organization and management: Genomic, proteomic and tissue databases: entry file format, sequence entry, sequence search.
5. Lipidomics, metabolomics, and integrated imaging-omics techniques: quantitative analysis, correlation between imaging-derived data and quantitative targeted/untargeted mass spectrometry techniques

2) Analysis of biological data (24 hours):

1. From Reads to Sequence: Assembly Algorithms in the 3 generations of sequencing platforms
2. Comparison of sequences:
 - a. Alignment (local or global): definitions; optimal alignment, alignment scores, Substitution Matrices
 - b. Exact alignment algorithms: dynamic programming
 - c. Heuristic algorithms (BLAST, FASTA), similarity searches in databases
 - d. Multiple alignments

3. Phylogenies and pattern inference
4. Transcriptome analysis:
 - a. Annotation of alternative genes and transcriptomes
 - b. Analysis of RNA-seq data

3) Systems biology: the role of models (24 hours):

1. Modelling and analysis of (bio)chemical reaction networks
 - a. ODE and stochastic modeling approaches, simulation based on numerical integration, Gillespie algorithm and its variants
 - b. Graphical representation of reaction networks and biochemical pathways (e.g. in terms of Petri nets) and related analysis approaches
 - c. Flux balance analysis of metabolic pathways
2. Modelling and analysis of gene regulation networks
 - a. Boolean network models of gene regulation: simulation methods and attractor analysis, variants of Boolean networks
 - b. Inference of Boolean network models
3. Protein-Protein Interaction (PPI) networks
 - a. Proteomics and protein interactions, protein interaction databases, protein interaction prediction methods
 - b. Identification of modules in PPI networks
 - c. Network-based analysis of PPI networks and applications to drug target discovery
 - d. Hints on protein structure and function prediction

4) Applications to drug discovery and drug design (12 hours):

1. Basis of computer-aided drug design
2. Molecular Dynamics simulations in drug design
3. Laboratory on Ligand-protein graphical representation (UCSF Chimera) and MD simulations (Amber22)

**** Course organization & Assessment****

Lectures: 75%

Laboratory/practice: 25%

Assessment: written + oral exam

**** Prerequisites ****

Algorithms, Data analytics, AI basics