Universitat Autònoma de Barcelona

2022/2023

Bioinformatics

Code: 104415
ECTS Credits: 6

Degree	Type	Year	Semester
2503740 Computational Mathematics and Data Analytics	OT	4	1

Contact

Name: Natalia Isabel Vilor Tejedor
Email: nataliaisabel.vilor@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Teachers

Angel Gonzalez Wong

Prerequisites

None. Recommended to have taken the Bioinformatics subject.

Objectives and Contextualisation

The course aims to provide a view on the possibilities of big data analysis in bioinformatics. The course consists of two parts: 1) computational methodologies applied to drug discovery and 2) analysis of omics data. The course is part of the Mention in Statistics for Health Sciences.

Competences

- Calculate and reproduce certain mathematical routines and processes with ease.
- Formulate hypotheses and think up strategies to confirm or refute them.
- Make effective use of bibliographical resources and electronic resources to obtain information.
- Students must be capable of applying their knowledge to their work or vocation in a professional way and they should have building arguments and problem resolution skills within their area of study.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
- Using criteria of quality, critically evaluate the work carried out.
- Work cooperatively in a multidisciplinary context assuming and respecting the role of the different members of the team.

Learning Outcomes

1. Identify the utility of statistical knowledge in bioinformatics and in health sciences.
2. Make effective use of bibliographical resources and electronic resources to obtain information.
3. Recognise the statistical methods of inference most frequenty used in bioinformatics.
4. Students must be capable of applying their knowledge to their work or vocation in a professional way and they should have building arguments and problem resolution skills within their area of study.
5. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
6. Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
7. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
8. Use temporary-evolution data-summary graphs.
9. Using criteria of quality, critically evaluate the work carried out.
10. Work cooperatively in a multidisciplinary context, taking on and respecting the role of the distinct members in the team.

Content

PART 1. Big Data in Drug Design

1. Introduction to big data in drug design.
2. Structure of proteins and chemical space in small molecules (drug-like).
3. Protein-drug interactions.
4. Virtual screening.
5. Molecular dynamics.

PART 2. Big Data in Omics Analysis

1. Introduction to Bioconductor and bioinformatics tools for the analysis of omic data.
2. Genetic Association Studies and GWAS (Genome-wide association studies).
3. Multivariate Methods for Omics Data Analysis Integration and Big Data.

*Unless the requirements enforced by the health authorities demand a prioritization or reduction of these contents.

Methodology

The course is organized in sessions of 3 hours. Each session consists of a theoretical part (theory classroom) that will introduce the new concepts followed by a practical part (computer room) where the students will work on the implementation of concepts explained in the theoretical part. In each session the teacher will indicate the students some tasks to do autonomously, such as reading articles or sending reports. The material used by the teachers will be available on the Virtual Campus of the course.

**The proposed teaching methodology may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.*

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			

Practical sessions	21	0.84
Presentation of Research Project	3	0.12
Theory classes	21	0.84
Type: Supervised		
Tutoring	10	0.4
Type: Autonomous		
Preparation of Research Project	20	0.8
Study	70	2.8

Assessment

PART 1. Big Data in Drug Design (50%):

- practical exercises (10%)

- theoretical-practical test (20%)

- bioinformatics oral presentation of a project (20%)

BLOCK 2. Big Data in Data Analysis (50%):

- practical exercises (30%)

- theoretical-practical test (20%)

The minimum global qualification required to pass the subject will be 5 points. The minimum mark of each of the evaluated activities must be equal to or greater than 4 points. Students who have any of the parts suspended will be able to do the recovery exam where they can be re-examined from the suspended part.

Student's assessment may experience some modifications depending on the restrictions to face-to-face activities enforced by health authorities.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Presentation Research Project	20	0.5	0.02	6, 4, 5, 3, 10, 2, 8
Presentation of practicum reports	40	0.5	0.02	9, 1, 7, 6, 4, 5, 3, 10, 2, 8
Theoretical-practical exams	40	4	0.16	9, 1, 7, 6, 3

Bibliography

- Lesk A.M. *Introduction to Bioinformatics.* Oxford University Press 2005.
- Attwood, T.K., Parry-Smith, D.J*., Introducción a la Bioinformática.* Pearson Education, 2002.
- Foulkes A.S. *Applied Statistical Genetics with R. For Population-based Association Studies.*Springer Dordrecht Heidelberg London New York. ISBN 978-0-387-89553-6
- Gonzalez JR, Cáceres A. *Omic association studies with R and Bioconductor*. Chapman and Hall/CRC, ISBN 9781138340565, 2019.

- https://www.bioconductor.org/

Software

R