

Current Topics in Bioinformatics

Code: 105065
ECTS Credits: 3

Degree	Type	Year	Semester
2500890 Genetics	OT	4	0

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Marta Coronado Zamora
Email: Marta.Coronado@uab.cat

Use of Languages

Principal working language: spanish (spa)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Other comments on languages

45%

Teachers

Marta Coronado Zamora

External teachers

Nerea Carrón Rodas

Prerequisites

- It is recommended to have passed *Bioinformatics* (3rd year), *Genomics*, *Proteomics and Interactomics* (3rd year) and the module of *Databases and programming fundamentals* within *Instrumental Techniques* subject (2nd year) of the Degree of Genetics.
- It is essential to have some bases of some programming language (preferably Python) and be familiar with the Linux environment in order to follow the practical sessions and complete the activities of continuous evaluation.
- A level B1.2 of English or equivalent is recommended.

Objectives and Contextualisation

The purpose of this subject is to cover basic topics of bioinformatics in the form of practical lessons, workshops and lectures given by experts. It is not a cumulative subject but a transversal one, whose purpose is to provide students with the wide range of concepts and approaches that bioinformatics encompasses.

The main objective is to provide students with the knowledge and skills necessary to apply bioinformatics in different areas of genomic research and, by extension, other omics. The subject and the activities carried out during this course provide a global insight into the potential of bioinformatics field in both basic and applied research.

Competences

- Be able to analyse and synthesise.
- Be able to communicate effectively, orally and in writing.
- Describe and identify the structural and functional characteristics of nucleic acids and proteins including their different organisational levels.
- Describe the organisation, evolution, inter-individual variation and expression of the human genome.
- Develop self-directed learning.
- Know and apply the omic tools of genomics, transcriptomics and proteomics.
- Measure and interpret the genetic variation in and between populations from a clinical, conservational and evolutionary perspective, and from that of the genetic improvement of animals and plants.
- Perceive the strategic, industrial and economic importance of genetics and genomics to life sciences, health and society.
- Reason critically.

- Use and interpret data sources on the genomes and macromolecules of any species and understand the basics of bioinformatics analysis to establish the corresponding relations between structure, function and evolution.
- Use and manage bibliographic information or computer or Internet resources in the field of study, in ones own languages and in English.

Learning Outcomes

1. Be able to analyse and synthesise.
2. Be able to communicate effectively, orally and in writing.
3. Defend the relevance of progress in the generation and interpretation of data on a genomic scale for our understanding and technological manipulation of organisms.
4. Develop self-directed learning.
5. Explain and apply the methods for the analysis and annotation of genomes.
6. Explain how knowledge of human genetic variation is applied to personalised medicine, pharmacogenomics and nutrigenomics.
7. List and explain the content of bioinformatics databases and perform searches for information.
8. Reason critically.
9. Use and interpret the results of bioinformatics applications in the molecular analysis of sequences.
10. Use and manage bibliographic information or computer or Internet resources in the field of study, in ones own languages and in English.
11. Use bioinformatics techniques and tools to describe and analyse the human genome.
12. Use the techniques, tools and methodologies used to describe, analyse and interpret the enormous amounts of data produced by high performance technologies.

Content

The subject will be composed of theoretical-practical sessions, lectures and workshops given by recognized specialists in the different subjects and fields¹.

Theoretical-practical sessions (~12h)

They will take place in the computer room². Students will work both individually and as a group (3-4 students) promoting active learning that will allow them to develop the capacity for analysis and synthesis, critical reasoning and the capacity to solve problems.

We are about to start a journey into real Bioinformatics. This series of four hands-on training activities will show you the basic workflows in bioinformatics: from the data management and processing with Linux, visualization, and posterior functional analyses. The practicals are divided into two big parts: Part I: Basics on Bioinformatics workflows and Part II: Solving real cases in genomics.

These practicals pretend also to gain other skills, very valuable in research but rarely experienced during the Degree, such as collaborating, learning how to transform the data into effective graphs to communicate and doing reproducible research.

Title		Description and learning outcomes
Introduction		Subject presentation: organization, methodology, preparations, group creation, subject grading
P1. Basics on Bioinformatics workflows	Data management and processing	Learning Linux for Bioinformatics - Learning to manage raw data with basic commands of bash, a powerful Linux language.

	Data exploration and visualization	Data exploration and visualization - Learning how to represent biological data into knowledge graphs. We will use ggplot2, an R package.
P2. Solving real cases in genomics	Genome-Wide Association studies	Which are the variants behind three common human diseases? We will perform a genome wide association study with R to detect SNPs associated with complex human diseases.
	Transcriptomic analyses	Finding differential expressed genes in cancer. Perform a differential expression analyses using human cancer data.
	Mentoring*	*Two extra sessions of two hours will be implemented to be set according to the needs of the students and the difficulty of the cases.

Invited conferences (10h)

Attendance to at least four conferences (2h/conference) of invited experts in the field of bioinformatics that will be taught in English.

Workshops (7h)

Title	Speaker	Tentative date ³
¿Cómo podemos analizar el genoma del cáncer?	Claudia Arnedo PhD student at IRB	November 2021
Cómo sobrevivir a un doctorado.	Edgar Garriga PhD student at CRG	December 2021
¡Esto no me lo habían dicho! Lo que no está escrito sobre la cultura científica.	Rosana Garrido Assistant professor UAB	January 2022

¹Unless the requirements enforced by the health authorities demand a prioritization or reduction of these contents.

²Unless the restrictions imposed by the health authorities require them to be performed in the normal theory classroom with a laptop.

³The final dates will be updated in the calendar and will be notified through the communication tools of the Moodle space. The lectures will be virtual and can be recorded.

Methodology

In-person learning activities and autonomous learning

A cooperative learning experience will be implemented, specifically the Puzzle methodology will be followed: data sets and procedures are provided in a distributed manner to small groups. Each group must manage and solve practical cases autonomously.

Each module will work in parallel on similar concepts through real practical cases, so that once finished the students will exchange information about the chosen methodology, its development and the results obtained, trying to achieve an effective cooperation among the students. The members of each group will know in depth the information they manage. Each group will make a lecture and/or writing of a portfolio through which the other groups understand the characteristics and foundations of each analysis. The four practical sessions will be linked, since the results or the methodology of one practical will be used for the next practical.

The active participation, the management of the work, as well as the discussion of the acquired knowledge will form a vital part in the role played by each student.

Conferences and workshops

A total of 5 lectures will be given by experts in their respective fields of research or work that will offer a real vision on how bioinformatics plays a key role in the resolution of basic and applied biological research questions. Emphasis will be placed on the importance of data processing in the current era of big data.

The three workshops will consist of three sessions that will deal with three aspects of practical interest, how to survive the doctorate, how to spread science with a real example and how the scientific world works.

*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			
Lectures	10	0.4	3, 6, 5, 7
Theoretical-practical sessions	12	0.48	3, 4, 6, 5, 7, 8, 2, 1, 10, 9, 11, 12
Workshops	7	0.28	4, 5, 8, 2
Type: Supervised			
Portfolio	20	0.8	3, 4, 6, 5, 7, 8, 2, 1, 10, 9, 11, 12
Type: Autonomous			
Study/Problem solving	25	1	4, 7, 8, 9, 11, 12

Assessment

The grading will be carried out through the delivery of four portfolios and a seminar about a bioinformatics topic chosen by the students.

Portfolio (70%). In each portfolio the basic fundamentals of the analyzed data will be exposed, the tools used, the development of the methodology, as well as a discussion on the final result of the delivery. Each portfolio will have the same weight in the final evaluation.

Seminar presentation (20%). Each group will make a 20-minute oral presentation.

Assistance and participation (10%).

The subject is passed when the average score of the assessment activities is equal to or greater than 5. The continuous and transversal nature of this evaluation means that the subject can not be evaluated if the minimum participation of the students is less than 80% of the students. proposed sessions.

*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
Portfolio	70%	0	0	3, 4, 6, 5, 7, 8, 2, 1, 10, 9, 11, 12
Seminar lecture	20%	1	0.04	3, 6, 5, 8, 2, 9
Soft skills	10%	0	0	3, 4, 6, 5, 8, 2, 1, 10

Bibliography

Books

- Pevzner, P. and R. Shamir. 2011. Bioinformatics for Biologists. Cambridge University Press
- Samuelsson, T. 2012. Genomics and Bioinformatics: An Introduction to Programming Tools for Life Scientists
- Lesk, A. 2014. Introduction to bioinformatics. Oxford University Press
- Claverie, J-M. 2007. Bioinformatics for dummies. Wiley, cop
- González, JR., Cáceres, A. 2019. Omic association studies with R and Bioconductor. CRC Press
- Hadley, W. 2009. ggplot2: elegant graphics for data analysis. Springer

Articles

- [Himmelblau, E. \(2021\) A cartoon guide to bioinformatics by a novice coder. Nature](#)
- [Markowetz, F. \(2017\) All biology is computational biology. PLoS Biology 15\(3\):e2002050](#)
- [Webb, S. \(2018\) Deep learning for biology. Nature 554:555-557](#)
- [Mardis, E.R. \(2010\) The \\$1,000 genome, the \\$100,000 analysis? Genome Medicine 2:84](#)
- [Tyler-Smith C, Yang H, Landweber LF, Dunham I, Knoppers BM, Donnelly P, et al. \(2015\) Where Next for Genetics and Genomics? PLoS Biology 13\(7\):e1002216](#)

Links

- [Web MSc in Bioinformatics UAB](#)
- [Bioinformatics salaries](#)
- [Bioinformatics: What? Why? Who?](#)

Software

- Operative system: Linux
- Programming languages: bash, R
- Software: RStudio y Jupyter Notebook
- R packages: ggplot2, shiny, rmarkdown, knitr, BiocManager, SNPAssoc, SNPRelate, DESeq2, edgeR, snpStats, limma