

**COURSE DATA****Data Subject**

Code	43459
Name	Bioinformatics
Cycle	Master's degree
ECTS Credits	3.0
Academic year	2024 - 2025

Study (s)

Degree	Center	Acad. Period	year
2210 - Master's Degree in Research in Molecular, Cellular and Genetics Bio	Faculty of Biological Sciences	1	First term

Subject-matter

Degree	Subject-matter	Character
2210 - Master's Degree in Research in Molecular, Cellular and Genetics Bio	4 - Bioinformatics	Obligatory

Coordination

Name	Department
PALERO PASTOR, FERRAN	194 - Genetics

SUMMARY

Bioinformatics is fundamentally practical and for this reason the theoretical knowledge will be taught simultaneously with the practical knowledge in the computer science classroom.

Originally, bioinformatics was defined as an interdisciplinary subject that covered the fields of biology, computer science, mathematics and statistics and whose objective was to analyze biological sequences, the contents and structures of genomes, and the prediction and function of proteins. With the advent of the "-omics" techniques, bioinformatics has extended its field of study to the analysis of large datasets, or Big Data, including those derived from complete genomes, and therefore is currently of great relevance in biomedical research.



PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

Students should have basic knowledge of molecular biology, statistics, and sequence analysis.

2210 - Master's Degree in Research in Molecular, Cellular and Genetics Bio

- Students should communicate conclusions and underlying knowledge clearly and unambiguously to both specialized and non-specialized audiences.
- Be able to access the information required (databases, scientific articles, etc.) and to interpret and use it sensibly.
- Ser capaz de buscar en la red información sobre secuencias de ácidos nucleicos y proteínas y adquirir la capacidad de manejar software relacionado con el análisis de secuencias.
?
- Ser capaz de procesar y extraer información a partir de los datos proporcionados por un servicio de secuenciación y convertir dicha información a un formato que permita su análisis con diferentes programas de análisis de secuencias.
- Ser capaz de recolectar información acerca de un organismo o elemento biológico, a partir de la almacenada en los servidores públicos, organizarla y sintetizarla.

The student, in order to pass this subject, must:

1. Be able to handle a sequence assembly program, from the initial steps of preparing sequence reads, to final assembly and mutation analysis.
2. Be able to work at a basic level with sequences derived from next-generation sequencing techniques using assembly software and be able to distinguish among sequence assembly algorithms.
3. Be able to annotate genomes of prokaryotes and eukaryotes, recognizing the diversity of patterns that can be found in different organisms.



DESCRIPTION OF CONTENTS

1. Introduction to Bioinformatics with GALAXY (Week 4: 4h)

1. Introduction to computer science and use of GALAXY.
2. Sequencing techniques: quality of sequences and file formats.

2. Databases and sequence search (Week 5: 4h)

3. Introduction to using Genbank and ENA
4. Sequence Search Using BLAST

3. Alignments, genetic distances and phylogenies (Week 6: 2h)

5. Alignment of sequences, importance and most frequent methods. Calculation of genetic distances. Molecular phylogenies.

4. Genome Assembly (Weeks 6 & 7: 4h)

7. Most frequent concepts and methods/algorithms. Assembly of sequences derived from NGS technologies
8. Evaluation of the assembled genome. Descriptive parameters of the assembly and genomic databases.

5. Annotation of genomes and genomic viewers (Weeks 7 and 8: 4h)

9. Annotation of bacterial and eukaryotic genomes
10. Genomic Viewers

6. Sequence mapping, variant analysis and population genomics (Weeks 8 and 10: 4h)

11. Sequence mapping: main methodologies and concepts. Use of minimap2.
12. Steps involved in variant calling. Types of data formats encountered during variant calling. Use command line tools to perform variant calling.

7. Transcriptome analysis (Week 12: 4h)

13. Introduction to R. Statistical packages for bioinformatics: BIOCONDUCTOR.
14. Analysis of gene expression data. Principal Component Analysis. Cluster analysis.

**WORKLOAD**

ACTIVITY	Hours	% To be attended
Computer classroom practice	26,00	100
Other activities	4,00	100
Development of group work	15,00	0
Study and independent work	15,00	0
Preparation of practical classes and problem	15,00	0
TOTAL	75,00	

TEACHING METHODOLOGY

The development of the subject is structured in face-to-face and independent work:

Face-to-face work:

- A) Classroom class sessions with a theoretical introduction to the master class system and a practical part based on the resolution of exercises and problems.
- B) Tutorials
- C) Review

Independent work:

- A) Practical exercises and bioinformatic problems solved outside class hours.
- B) Study of the contents and prior preparation of the classes.

EVALUATION

The assessment of student learning in the first call will be made by assessing the following sections:

- 1) An exam in the computer lab that may comprise both knowledge questions on theory and exercises that can be solved using the programs studied in the course. This test will earn up to 7 points and will be made after the end of classes.
- 2) Continuous Assessment based on class work and in solving exercises and bioinformatics problems. This section will be worth up to 3 points.



In the second call, the student will have two options: 1) an exam where you can collect up to 10 points and 2) if the student has obtained a note for the continuous evaluation on the first call, these points will be added to those of the test, which will be worth up to 7 points.

REFERENCES

Basic

- Bioinformatics and Functional Genomics, por Jonathan Pevsner (2015) publicado por Wiley_Blackwell. Tercera edición. Una introducción a la bioinformática y la genómica fácil de seguir y de entender los conceptos. Incluye muchos ejercicios prácticos y direcciones web., y está disponible en línea (https://trobes.uv.es/permalink/34CVA_UV/1b8uv2g/alma991009853672206258).

Additional

- Bioinformatics with Python Cookbook : Use Modern Python Libraries and Applications to Solve Real-World Computational Biology Problems (2022), también disponible en línea (https://trobes.uv.es/permalink/34CVA_UV/1bttdu2/alma991009923653906258).
- R Bioinformatics Cookbook : Utilize R Packages for Bioinformatics, Genomics, Data Science, and Machine Learning (2023), disponible en línea (https://trobes.uv.es/permalink/34CVA_UV/1bttdu2/alma991010329590506258)