ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

203000025 - Genomics Data Analyses And Visualization

DEGREE PROGRAMME

20BC - Master Universitario En Biologia Computacional

ACADEMIC YEAR & SEMESTER

2025/26 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	203000025 - Genomics Data Analyses And Visualization
No of credits	6 ECTS
Type	Compulsory
Academic year ot the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	20BC - Master Universitario en Biologia Computacional
Centre	20 - Etsi Agronómica, Aliment. Y Biosistemas
Academic year	2025-26

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Carlos Perez Cantalapiedra (Subject coordinator)	U. Bioquímica	carlos.cantalapiedra@upm.es	Tu - 17:15 - 19:00

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

2.3. External faculty

Name and surname	Email	Institution
Jaime Huerta Cepas	j.huerta@csic.es	CBGP - UPM-INIA-CSIC

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- Basic use of Linux Terminal and shell scripting

- Basic programming in Python

- Medium level in Biology and Evolution

- Basic level in R

4. Skills and learning outcomes *

4.1. Skills to be learned

CE01 - Comprender las bases moleculares y las técnicas experimentales estándares más comunes en las investigaciones ómicas (genómica, transcriptómica, proteómica, metabolómica, interactómica, etc.).

CE04 - Utilizar diferentes bases de datos (incluidos los bigdata), conocer sus estructuras y ontologías, aplicar la estadística a su análisis, siendo capaz de utilizar herramientas de representación y visualización.

CE08 - Capacidad de integrar tecnologías y sistemas propios de la Inteligencia Artificial, con carácter generalista, y en contextos más amplios y multidisciplinares.

CE09 - Capacidad de interpretar los modelos de clasificación supervisada y no supervisada obtenidos al aplicar las técnicas de Aprendizaje Automático para un conjunto de datos.

CG01 - Poseer los conocimientos que constituyen la base científica y tecnológica de la Biología computacional, lo

que permitirá el desarrollo de ideas originales en este campo, en un contexto de investigación o desarrollo.

CG02 - Familiarizarse con el trabajo y los métodos de la Biología Computacional en condiciones reales, adquiriendo la capacidad de diseñar aplicaciones/experimentos de forma independiente y describir, cuantificar, analizar y evaluar críticamente los resultados obtenidos.

CG05 - Que los estudiantes sean capaces de integrar conocimientos en el area de la Biología Computacional, de formular conclusiones, hipótesis o líneas de trabajo a partir de la información disponible, y de formarse una opinion fundamentada sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos.

CT01 - Capacidad para aplicar de forma profesional a su trabajo los conocimientos adquiridos considerando sus impactos en un contexto global y social.

CT07 - Ser capaz de manejar las tecnologías de la información y comunicación en un contexto profesional.

CT08 - Tener capacidad de análisis y síntesis para interpretar datos relevantes y abordar los problemas desde diferentes perspectivas.

4.2. Learning outcomes

RA11 - Adquisición de conocimientos para el ensamblaje y anotación de genomas y variantes genómicas y estudios de expresión génica diferencial

RA10 - Adquirir conocimientos para el manejo de tecnologías Next Generation Sequencing (NGS) y sus aplicaciones a diferentes campos experimentales y tecnológicos

RA12 - Aprendizaje de métodos de análisis genómico y su visualización.

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

The course on Genomic Data Analysis and Visualization deals with the management, analysis and interpretation of sequencing data at genomic scale. Specifically, the course offers a practical and updated view on the most commonly used computational methods and tools in the analysis of genomes, transcripts, metagenomes and other related genomic data, covering from basic aspects such as assembly and quality control of sequences, to the comparative study of genomes, functional prediction or inference of regulatory networks. In addition, the course offers an overview of the use of such data both in the context of basic science studies (evolution, molecular biology, development) and in more applied projects (environment, clinical, genetic improvement).

5.2. Syllabus

1. Introduction to Genomics Analysis
   1.1. Biological sequences in computers
   1.2. Current DNA sequencing techniques
2. Preprocessing of sequencing data
3. Read mapping methods
4. Variant Calling for Population Genomics
5. Genome assembly and annotation
6. Transcriptomics. RNA differential expression analysis.
7. Single Cell transcriptomics. Identification of Cellular variants and transcriptional maps.
8. Functional and Regulatory Genomics.
9. Comparative Genomics
10. Metagenomics

6. Schedule

6.1. Subject schedule*

Week	Type 1 activities	Type 2 activities	Distant / On-line	Assessment activities
1	**Introduction to Genomic Analysis** Duration: 01:00 Lecture **Biological sequences in computers** Duration: 01:00 Lecture **Current DNA sequencing technologies** Duration: 02:00 Lecture			
2	**Preprocessing of sequencing data** Duration: 02:00 Lecture **Preprocessing of sequencing data** Duration: 02:00 Problem-solving class			
3	**Read mapping methods** Duration: 02:00 Lecture **Read mapping methods** Duration: 02:00 Problem-solving class			
4	**Variant Calling for Population Genomics** Duration: 02:00 Lecture **Variant Calling for Population Genomics** Duration: 02:00 Problem-solving class			**Read preprocessing, mapping and variant calling** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00
5	**Genomic assembly and annotation** Duration: 02:00 Lecture **Genome assembly** Duration: 02:00 Problem-solving class			
6	**Genome annotation** Duration: 02:00 Problem-solving class **Transcriptomics. RNA differential expression analysis.** Duration: 02:00 Lecture			**Genome assembly and annotation** Individual work Progressive assessment and Global Examination Not Presential Duration: 03:00

7	**RNAseq mapping, counting and normalization** Duration: 02:00 Problem-solving class **RNA differential expression analysis.** Duration: 02:00 Problem-solving class			**RNAseq analysis** Individual work Progressive assessment and Global Examination Not Presential Duration: 06:00
8	**Single-cell transcriptomics** Duration: 02:00 Lecture **Single-cell transcriptomics** Duration: 02:00 Problem-solving class			
9	**Functional and regulatory genomics** Duration: 02:00 Lecture **Functional and regulatory genomics** Duration: 02:00 Problem-solving class			
10	**Comparative Genomics** Duration: 02:00 Lecture **Comparative Genomics** Duration: 02:00 Problem-solving class			
11	**Comparative Genomics** Duration: 02:00 Lecture **Comparative Genomics** Duration: 02:00 Problem-solving class			
12	**Comparative Genomics** Duration: 02:00 Problem-solving class **Comparative Genomics** Duration: 02:00 Lecture			
13	**Metagenomics** Duration: 02:00 Lecture **Metagenomics** Duration: 02:00 Problem-solving class			
14	**Metagenomics** Duration: 02:00 Lecture **Metagenomics** Duration: 02:00 Problem-solving class			

15	**Metagenomics** Duration: 02:00 Lecture **Metagenomics** Duration: 02:00 Problem-solving class			
16	**Metagenomics** Duration: 02:00 Lecture **Metagenomics** Duration: 02:00 Problem-solving class			
17				**Individual Project: evaluation of sequencing data** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00 **Individual Project: variant calling** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00 **Individual Project: genome assembly** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00 **Individual Project: transcriptomics** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00 **Individual Project: comparative genomics** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00 **Individual Project: metagenomics** Individual work Progressive assessment and Global Examination Not Presential Duration: 04:00

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
4	Read preprocessing, mapping and variant calling	Individual work	No Presential	04:00	10%	5 / 10	CT07 CE01 CG01 CG02
6	Genome assembly and annotation	Individual work	No Presential	03:00	10%	5 / 10	CG01 CG02 CT07 CE01
7	RNAseq analysis	Individual work	No Presential	06:00	10%	5 / 10	
17	Individual Project: evaluation of sequencing data	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: variant calling	Individual work	No Presential	04:00	5%	0 / 10	CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08 CG05
17	Individual Project: genome assembly	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04

							CE08
17	Individual Project: transcriptomics	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: comparative genomics	Individual work	No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: metagenomics	Individual work	No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08

7.1.2. Global examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
4	Read preprocessing, mapping and variant calling	Individual work	No Presential	04:00	10%	5 / 10	CT07 CE01 CG01 CG02
6	Genome assembly and annotation	Individual work	No Presential	03:00	10%	5 / 10	CG01 CG02 CT07 CE01
7	RNAseq analysis	Individual work	No Presential	06:00	10%	5 / 10	
17	Individual Project: evaluation of sequencing data	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07

							CE01 CE04 CE08
17	Individual Project: variant calling	Individual work	No Presential	04:00	5%	0 / 10	CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08 CG05
17	Individual Project: genome assembly	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: transcriptomics	Individual work	No Presential	04:00	5%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: comparative genomics	Individual work	No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Individual Project: metagenomics	Individual work	No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Individual Project and oral exam	Individual work	Face-to-face	26:00	100%	5 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08

7.2. Assessment criteria

Progressive assessment

The assessment activities will consist of carrying out exercises and solving problems raised during the classes of each thematic block, as well as in class participation and attendance. This evaluation will have a weight of 30% of the final grade of each block. Note however that it is mandatory to submit the individual works and achieve at least 5/10 score in those works. The remaining 70% will consist of carrying out a piece of work as an individual project using techniques and concepts covered throughout the course.

Global examination

The global examination will consist on the submission of exercises raised during the classes of each thematic block, with a weight of 30% of the final grade. It will also include a individual project, covering techniques and concepts addressed throughout the course, with a weight of 70%.

Extraordinary exam call

The extraordinary exam will require the submission of the individual project, covering techniques and concepts addressed throughout the course. Note that, in this extraordinary exam call, an oral exam will be also carried out, including questions about the different sections of the individual project submitted by the student. The individual project and the oral exam will be scored as a whole, corresponding to 100% of the final grade. At least 5/10 score will be required to pass the subject.

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Sequence - Evolution - Function: Computational Approaches in Comparative Genomics. Koonin EV, Galperin MY. Boston: Kluwer Academic; 2003. Free online materia: https://www.ncbi.nlm.nih.gov/books/NBK20260/	Bibliography	provides an understanding of the principles and approaches of functional genomics and of the potential and limitations of computational and experimental approaches to genome analysis. Free version at https://www.ncbi.nlm.nih.gov/pubmed/21089240
Practical Computing for Biologists	Bibliography	Steven H. D. Haddock, Casey W. Dunn. The book shows how to use many freely available computing tools to work more powerfully and effectively. http://practicalcomputing.org/
Modern Statistics for Modern Biology	Bibliography	by Susan Holmes and Wolfgang Huber. From raw data to beautiful illuminating output, you learn to write your own scripts in the R language and to use advanced statistics packages. Online free version at https://www.huber.embl.de/msmb

Beginner?s guide to comparative bacterial genome analysis using next-generation sequence data	Web resource	Microbial Informatics and Experimentation Beginner?s guide to comparative bacterial genome analysis using next-generation sequence data
Maria S. Poptsova (2014) Genome Analysis: Current Procedures and Applications. ISBN: 978-1-908230-29-4. Caister Academic Press	Bibliography	
Peter N. Robinson, Rosario Michael Piro, Marten Jager (2017) Computational Exome and Genome Analysis. ISBN 9781498775984. Chapman and Hall/CRC	Bibliography	
Vijai Bhadauria (2016) Omics in Plant Disease Resistance. ISBN: 978-1-910190-35-7. Caister Academic Press	Bibliography	
Kenneth D. Birnbaum and Edo Kussell (2011) Measuring cell identity in noisy biological systems. Nucleic Acids Research, 2011, Vol. 39, No. 21 9093?9107	Bibliography	
Idan Efroni, Pui-Leng Ip, Tal Nawy, Alison Mello and Kenneth D Birnbaum (2015) Quantification of cell identity from single-cell gene expression profiles. Genome Biology (2015) 16:9.	Bibliography	

9. Other information

9.1. Other information about the subject

The computational analysis of genomic data has become an essential part of environmental and health studies. Comparative genomics, for instance, is a powerful tool to investigate what makes a species, population, or even ecosystem unique; which stands as a fundamental pillar for developing sustainability programs around OSD13-15.