



POLITÉCNICA

INTERNATIONAL
CAMPUS OF
EXCELLENCE

COORDINATION PROCESS OF
LEARNING ACTIVITIES
PR/CL/001



E.T.S. de Ingenieros
Informáticos

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

203000025 - Genomics Data Analyses And Visualization

DEGREE PROGRAMME

20BC - Master Universitario En Biología Computacional

ACADEMIC YEAR & SEMESTER

2022/23 - Semester 1

Index

Learning guide

1. Description.....	1
2. Faculty.....	1
3. Prior knowledge recommended to take the subject.....	2
4. Skills and learning outcomes	2
5. Brief description of the subject and syllabus.....	4
6. Schedule.....	5
7. Activities and assessment criteria.....	8
8. Teaching resources.....	12
9. Other information.....	14

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 of the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	20BC - Master Universitario en Biología Computacional
Centre	20 - E.T.S. De Ingenieria Agronomica, Alimentaria Y De Biosistemas
Academic year	2022-23

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Joaquin Giner Lamia (Subject coordinator)	B30a CBGP	joaquin.giner@upm.es	Tu - 17:00 - 18:00
Jaime Huerta Cepas	B29	j.huerta@upm.es	Tu - 17:00 - 18:00
Miguel Angel Moreno Risueño	B10	miguelangel.moreno@upm.es	Sin horario.

* 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
Profesor Ayudante Doctor	email@upm.es	Departamento BBV
Jaime Huerta Cepas	j.huerta@upm.es	CBGP

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 área de la Biología Computacional, de formular conclusiones, hipótesis o líneas de trabajo a partir de la información disponible, y de formarse una opinión 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

La asignatura de Análisis y Visualización de Datos Genómicos aborda el manejo, análisis e interpretación de datos de secuenciación a escala genómica. En concreto, la asignatura ofrece una visión práctica y actualizada sobre los métodos y herramientas computacionales más utilizados en el análisis de genomas, transcriptomas, metagenomas y otros datos genómicos relacionados, cubriendo desde aspectos básicos como el ensamblaje y el control de calidad de secuencias, hasta el estudio comparado de genomas, la predicción funcional o la inferencia de redes de regulación. Además, la asignatura ofrece un visión general sobre el uso de dichos datos tanto en el contexto de estudios de ciencia básica (evolución, biología molecular, desarrollo) como en proyectos carácter más aplicado (medio ambiente, clínica, mejora genética).

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. Applications and current sequencing techniques.
2. Genome assembly and annotation. NGS data handling.
3. Variant Calling Population Genomics. SNPs and Genome Wide Association Studies.
4. Transcriptomics. RNA differential expression analysis.
5. Identification of Biomarkers from NGS data. Finding Molecular identity.
6. Single Cell transcriptomics. Identification of Cellular variants and transcriptional maps.
7. Functional and Regulatory Genomics.
8. Comparative Genomics
9. Metagenomics

6. Schedule

6.1. Subject schedule*

Week	Classroom activities	Laboratory activities	Distant / On-line	Assessment activities
1	Intro Genómica Duration: 01:00 Tecnologías y datos de Secuenciación Duration: 01:00 Tecnologías y datos de secuenciación Duration: 02:00			
2	Ensamblado de genomas Duration: 02:00 Ensamblado de genomas Duration: 02:00			
3	Genomica poblacional Duration: 02:00 Genomica poblacional Duration: 02:00			
4	Transcriptomica Duration: 02:00 Transcriptomica Duration: 02:00			
5	Biomarcadores Duration: 02:00 Biomarcadores Duration: 02:00			
6	Transcriptomica de célula única Duration: 02:00 Transcriptomica de célula única Duration: 02:00			

7	<p>Transcriptómica de célula única Duration: 02:00</p> <p>Genómica Funcional Duration: 02:00</p>			
8	<p>Genómica Funcional Duration: 02:00</p> <p>Genómica Comparada Duration: 02:00</p>			
9	<p>Genómica Comparada Duration: 02:00</p> <p>Genómica Comparada Duration: 02:00</p>			
10	<p>Genómica Comparada Duration: 02:00</p> <p>Genómica Comparada Duration: 02:00</p>			
11	<p>Genómica Comparada Duration: 02:00</p> <p>Genómica Comparada Duration: 02:00</p>			
12	<p>Metagenómica Duration: 02:00</p> <p>Metagenómica Duration: 02:00</p>			
13	<p>Metagenómica Duration: 02:00</p> <p>Metagenómica Duration: 02:00</p>			
14	<p>Metagenómica Duration: 02:00</p> <p>Metagenómica Duration: 02:00</p>			

15	<p>Metagenómica Duration: 02:00</p> <p>Metagenómica Duration: 02:00</p>			
16				
17				<p>Handling sequencing data</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p> <p>Genome Assembly and Population Genomics</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p> <p>Transcriptomics</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p> <p>Biomarkers and Single Cell</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p> <p>Comparative Genomics</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p> <p>Metagenomics</p> <p>Continuous assessment and final examination Not Presential Duration: 04:00</p>

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.

* The schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Handling sequencing data		No Presential	04:00	10%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Genome Assembly and Population Genomics		No Presential	04:00	13.4%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Transcriptomics		No Presential	04:00	10%	0 / 10	CT01 CT08 CE09 CT07 CE01 CE04 CG02 CE08 CG05 CG01
17	Biomarkers and Single Cell		No Presential	04:00	16.6%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04

							CE08
17	Comparative Genomics		No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Metagenomics		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
17	Handling sequencing data		No Presential	04:00	10%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Genome Assembly and Population Genomics		No Presential	04:00	13.4%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Transcriptomics		No Presential	04:00	10%	0 / 10	CT01 CT08 CE09 CT07 CE01 CE04 CG02 CE08

							CG05 CG01
17	Biomarkers and Single Cell		No Presential	04:00	16.6%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Comparative Genomics		No Presential	04:00	25%	0 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07 CE01 CE04 CE08
17	Metagenomics		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
Full exercise covering all topics in the course		Face-to-face	08:00	100%	5 / 10	CG05 CG01 CG02 CT01 CT08 CE09 CT07

La calificación en convocatoria extraordinaria se obtendrá por los conceptos detallados en la tabla de Evaluación Convocatoria Extraordinaria

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

<p>Maria S. Poptsova (2014) Genome Analysis: Current Procedures and Applications. ISBN: 978-1-908230-29-4. Caister Academic Press</p>	<p>Bibliography</p>	
<p>Peter N. Robinson, Rosario Michael Piro, Marten Jager (2017) Computational Exome and Genome Analysis. ISBN 9781498775984. Chapman and Hall/CRC</p>	<p>Bibliography</p>	
<p>Vijai Bhadauria (2016) Omics in Plant Disease Resistance. ISBN: 978-1-910190-35-7. Caister Academic Press</p>	<p>Bibliography</p>	
<p>Kenneth D. Birnbaum and Edo Kussell (2011) Measuring cell identity in noisy biological systems. Nucleic Acids Research, 2011, Vol. 39, No. 21 9093?9107</p>	<p>Bibliography</p>	
<p>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.</p>	<p>Bibliography</p>	

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.