

Titulación	Tipo	Curso
4313473 Bioinformática / Bioinformatics	OB	0

## Contacto

Nombre: Javier Panadero Martinez

Correo electrónico: javier.panadero@uab.cat

## Equipo docente

Javier Panadero Martinez

## Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

## Prerrequisitos

Para el desarrollo general del curso, se recomienda tener un nivel B2 de inglés o similar.

Para el módulo de programación es muy recomendable tener nociones básicas de Linux como usuario (conocer las herramientas básicas de gestión y edición de carpetas y ficheros).

## Objetivos y contextualización

Los objetivos generales de este módulo son la aplicación de las herramientas y técnicas básicas para el desarrollo en esta área. Se trabajan las capacidades de resolver retos y adaptarse a las tecnologías y paradigmas de la bioinformática.

## Competencias

- Diseñar y aplicar la metodología científica en la resolución de problemas.
- Identificar las necesidades bioinformáticas de los centros de investigación y las empresas del sector de la biotecnología y la biomedicina.
- Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
- Trabajar individualmente y en equipo en un contexto internacional y multidisciplinario.

- Utilizar sistemas operativos, programas y herramientas de uso común en bioinformática, así como, manejar plataformas de cómputo de altas prestaciones, lenguajes de programación y análisis bioinformáticos.

## Resultados de aprendizaje

1. Diseñar y aplicar la metodología científica en la resolución de problemas.
2. Diseñar, analizar y evaluar las prestaciones de infraestructuras paralelas y grandes volúmenes de datos.
3. Gestionar plataformas paralelas y bases de datos bioinformáticas de acuerdo a las necesidades existentes.
4. Identificar las ventajas y limitaciones de la bioinformática y la importancia de la aplicación de nuevas tecnologías computacionales en investigaciones ómicas.
5. Implementar algoritmos y técnicas de cálculo estadístico, para la gestión de grandes volúmenes de datos.
6. Iniciarse en la creación, conexión, edición y consulta de bases de datos.
7. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
8. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
9. Trabajar individualmente y en equipo en un contexto internacional y multidisciplinario.
10. Utilizar los lenguajes de programación en Bioinformática R y Python.

## Contenido

### 1. Linux (comandos y shell scripting)

Comandos básicos, manejo a nivel de usuario, gestión de software y sistema de ficheros  
Herramientas de procesamiento de texto y manipulación de datos  
Redirecciones, tuberías y filtros  
Shell scripting en Bash

### 2. Lenguajes de programación

Introducción a la programación en Python para Bioinformática  
Variables, expresiones, tipos de datos, operadores, construcciones programáticas y contextos  
Funciones, módulos y subrutinas  
Programación recursiva  
Entrada/Salida  
Depuración de código  
Otros lenguajes de programación: R

### 3. Estructuras de datos y procesamiento de datos

Estructuras de datos básicas (cadenas, listas, tuplas, conjuntos y diccionarios)  
Estructuras de datos anidadas y objetos  
Árboles y grafos  
Modelado y representación de datos en bioinformática  
Formatos básicos en bioinformática (FASTQ, SAM, VCF)  
Expresiones regulares

### 4. Algoritmos en bioinformática

Introducción a la complejidad algorítmica  
 Algoritmos de divide y vencerás  
 Enumeración combinatoria y backtracking  
 Programación dinámica

#### 5. Herramientas y librerías en bioinformática

Herramientas para la visualización de datos  
 Introducción a Biopython  
 Introducción a NumPy y Pandas

### Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Solución de problemas en el aula	14	0,56	1, 7, 8, 5
Trabajo realizado en el aula	20	0,8	1, 2, 3, 7, 8, 5
Trabajo realizado en el laboratorio	12	0,48	1, 4, 3, 7, 5, 9
Tipo: Supervisadas			
Trabajo realizado en el laboratorio a partir de las lecturas recomendadas	15	0,6	1, 4, 7, 8, 9
Tipo: Autónomas			
Trabajo realizado de forma semanal sobre los entregables y materiales proporcionados	83	3,32	1, 2, 8, 5, 9

La metodología combinará el trabajo en el aula, la solución de problemas en el aula de forma supervisada, el trabajo en el aula de forma no supervisada y el trabajo a realizar de forma individual a partir de las lecturas recomendadas y entregas a realizar. Se utilizará una plataforma virtual para la entrega de los informes de los trabajos.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

### Evaluación

#### Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Evaluación del trabajo realizado por el alumnado durante el	10%	1	0,04	1, 4, 8, 5, 9

módulo

Exámen final	30%	1	0,04	1, 7, 8
Exámenes individuales teóricos y prácticos	50%	2	0,08	1, 2, 3, 7, 8, 10
Trabajo en el laboratorio. Individual y en grupo	10%	2	0,08	1, 2, 4, 6, 3, 8, 5, 9

La metodología combinará el trabajo en clase, la solución de problemas en clase, el trabajo autónomo en el laboratorio y fuera del mismo. Se utilizarán las plataformas virtuales para el seguimiento del curso. Ninguna de las actividades de evaluación individuales representará más del 50% de la nota final.

#### Examen de recuperación

Para poder participar en el proceso de recuperación, el alumnado deberá previamente haber participado en como mínimo el equivalente a dos tercios de la nota final del módulo en actividades de evaluación. El profesorado informará de los procedimientos y plazos para el proceso de recuperación. Nótese que las actividades realizadas dentro de clase de forma continua no pueden recuperarse.

#### No evaluable

El alumnado será calificado como "No evaluable" cuando el peso de la evaluación en la que ha participado sea inferior al equivalente al 67% de la nota final del módulo.

Esta asignatura/módulo no contempla el sistema de evaluación única.

## Bibliografía

- Rajkumar Buyya, "High Performance Cluster Computing: Programming and Applications", PH, 1999.
- Bell, Charles; Kindahl, Mats; Thalmann, Lars. "MySQL High Availability". O'Reilly, 2010.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A. & Wheeler, D. L. (2002).
- GenBank. *Nucl. Acids Res.*, 30(1):17-20. URL <http://nar.oupjournals.org/cgi/content/abstract/30/1/17>.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.*, 28(1):235-242.
- Bessant, C., Shadford, I., Oakley, D. "Building Bioinformatics Solutions with Perl, R and MySQL", Oxford University Press, 2009
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. & Schneider, M. .The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, 31(1):365-370.
- Christiansen, P., Wall, L., Orwant, J., "Programming Perl". 4th Edition, O'Reilly, 2012
- Mäkinen et al., Genome-Scale Algorithm Design: Biological Sequence Analysis in the Era of High-Throughput Sequencing. Cambridge Univ. Press, 2015.
- Matloff, N., "The Art of R Programming". No Starch Press Inc., 2011
- Lutz, M., "Learning Python", O'Reilly, 5th edition, 2013
- Siever, E., Figgins, S., "Linux in a nutshell" O'Reilly 2009.
- Sobell, M., "A Practical Guide to Linux. Commands, editors and shell programming". Prentice Hall, 2009.
- Tindall, James., Begining Perl for Bioinformatics. O'Reilly 2012.

#### Páginas web de referencia

- <http://mscbioinformatics.uab.cat>
- <https://cv.uab.cat>

Recursos de bioinformática e informática en la biblioteca electrónica de la UAB

- <http://www.uab.cat/biblioteques/trobador>
- <http://pagines.uab.cat/bctdigital/>

## Software

Linux (Ubuntu, Bash, linux-tools, etc)  
Python 2.7/3.x  
Jupyter Notebook / PyCharm  
R/RStudio  
Matplotlib/Seaborn  
Numpy/Pandas

## Lista de idiomas

Nombre	Grupo	Idioma	Semestre	Turno
(PLABm) Prácticas de laboratorio (máster)	1	Inglés	primer cuatrimestre	mañana-mixto
(TEm) Teoría (máster)	1	Inglés	primer cuatrimestre	mañana-mixto