

## 1. DATOS BÁSICOS

<b>Asignatura</b>	METODOLOGÍA DE LA INVESTIGACIÓN I: MODELOS MATEMÁTICOS Y ESTADÍSTICOS EN BIOINFORMÁTICA Y BIOLOGÍA COMPUTACIONAL
<b>Titulación</b>	MÁSTER UNIVERSITARIO EN BIOINFORMÁTICA
<b>Escuela/ Facultad</b>	BIOMEDICINA Y SALUD
<b>Curso</b>	1º
<b>ECTS</b>	6 ECTS
<b>Carácter</b>	OBLIGATORIA
<b>Idioma/s</b>	ESPAÑOL
<b>Modalidad</b>	ONLINE
<b>Semestre</b>	PRIMERO
<b>Curso académico</b>	2022-2023
<b>Docente coordinador</b>	JAVIER PÉREZ FLORIDO

## 2. PRESENTACIÓN

En esta materia se estudiarán las principales herramientas computacionales para el análisis e investigación a partir de diferentes datos ómicos (genómica, transcriptómica, epigenómica) así como una aproximación integrativa y funcional de los mismos. Para ello, se introducirán conceptos de programación básicos de análisis de datos y estadística basados en el lenguaje R.

## 3. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

### Competencias básicas:

CB1. 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., incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

CB3. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

### Competencias transversales:

CT3. Competencia digital

CT5. Trabajo en equipo

CT6. Análisis crítico

### Competencias específicas:

CE4. Evaluar y testar las aplicabilidades de los diferentes modelos matemáticos, estadísticos o de inteligencia artificial en bioinformática.

**Resultados de Aprendizaje:**

- R1: Aplicar los fundamentos de la estadística y la probabilidad al ámbito de la bioinformática mediante el uso de los programas R y Rstudio.
- R2: Emplear los fundamentos del análisis inferencial a los estudios en Bioinformática mediante el uso de los programas R y Rstudio.
- R3: Analizar los fundamentos de las metodologías genómicas y sus principales aplicaciones
- R4: Evaluar las principales herramientas computacionales para el análisis de datos ómicos
- R5: Investigar y crear protocolos de extracción de datos de los experimentos de NGS
- R6: Crear redes de sistemas a partir de datos biológicos aplicando diferentes algoritmos matemáticos.

En la tabla inferior se muestra la relación entre las competencias que se desarrollan en la asignatura y los resultados de aprendizaje que se persiguen:

Competencias	Resultados de aprendizaje
CB1, CB3 CT3, CT5, CT6 CE4	Aplicar los fundamentos de la estadística y la probabilidad al ámbito de la bioinformática mediante el uso de los programas R y Rstudio.
	Emplear los fundamentos del análisis inferencial a los estudios en Bioinformática mediante el uso de los programas R y Rstudio
	Analizar los fundamentos de las metodologías genómicas y sus principales aplicaciones
	Evaluar las principales herramientas computacionales para el análisis de datos ómicos
	Investigar y crear protocolos de extracción de datos de los experimentos de NGS.
	Crear redes de sistemas a partir de datos biológicos

## 4. CONTENIDOS

**UNIDAD 1.** Introducción a R y entorno RStudio. Estructuras de control, tratamiento básico y lectura/Escritura de datos en R. Análisis estadístico en R

**UNIDAD 2.** Herramientas computacionales para el análisis de datos genómicos. Flujo de análisis de datos procedentes de secuenciación masiva. Identificación de variantes en enfermedades hereditarias y cáncer. Análisis de DNA-seq para datos de single-cell (scDNA-seq)

**UNIDAD 3.** Herramientas computacionales y procedimientos para el análisis de datos transcriptómicos. Introducción al análisis de expresión condatos desecuenciación ARN (RNA-Seq). Análisis de datos de ARN total y otras aplicaciones con datos de RNA-Seq. Análisis de RNA-Seq para datos de single-cell (scRNA-Seq).

**UNIDAD 4.** Herramientas computacionales para el análisis de datos epigenéticos. Introducción al análisis de datos epigenéticos. Análisis de datos de metilación. Análisis de datos regulación: CHIP-seq.

**UNIDAD 5.** Métodos estadísticos y de aprendizaje automático para extraer información de los datos de secuenciación de última generación NGS (ARNseq, Chip-Seq de modificaciones de histonas, y metilación de ADN).

**UNIDAD 6.** Redes en biología de sistemas: Teoría de grafos, Modelos de red, redes de expresión génica y redes de expresión metabólica.

## 5. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE

A continuación, se indican los tipos de metodologías de enseñanza-aprendizaje que se aplicarán:

- Clase magistral
- Aprendizaje cooperativo
- Aprendizaje basado en proyectos
- Aprendizaje basado en enseñanzas de taller
- Aprendizaje experiencial
- Entornos de simulación

## 6. ACTIVIDADES FORMATIVAS

A continuación, se identifican los tipos de actividades formativas que se realizarán y la dedicación en horas del estudiante a cada una de ellas:

**Modalidad online:**

<b>Clase magistral</b>	<b>8</b>
Seminario virtual	22
Elaboración de informes y escritos	12
Investigaciones y proyectos	10
Actividades en laboratorios virtuales	20
Estudios de contenidos y documentación complementaria	50
Tutoría virtual	18
Foro virtual	8
Pruebas presenciales de conocimiento	2
Clase magistral	8
<b>TOTAL</b>	<b>150 h</b>

## 7. EVALUACIÓN

A continuación, se relacionan los sistemas de evaluación, así como su peso sobre la calificación total de la asignatura:

**Modalidad online:**

<b>Sistema de evaluación</b>	<b>Peso</b>
Pruebas semipresenciales de conocimiento	60%
Informes y escritos	10%

Investigaciones y proyectos	10%
Cuaderno de prácticas de laboratorio	20%

En el Campus Virtual, cuando accedas a la asignatura, podrás consultar en detalle las actividades de evaluación que debes realizar, así como las fechas de entrega y los procedimientos de evaluación de cada una de ellas.

### 7.1. Convocatoria ordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de actividades.

#### Actividades obligatorias:

o En esta asignatura existen actividades obligatorias. La no realización de alguna de ellas implica no superar la asignatura en convocatoria ordinaria, y tener que presentarse a la convocatoria extraordinaria. La nota máxima que aparecerá en las actas de convocatoria ordinaria será un 4,0 sobre 10,0.

o Las actividades obligatorias son:

- **Actividad 1.** Herramientas en R para análisis exploratorio
- **Actividad 2.** Análisis de datos genómicos
- **Actividad 3.** Análisis de expresión diferencial funcional
- **Actividad 4.** Análisis de activación diferencial

### 7.2. Convocatoria extraordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de actividades.

Se deben entregar las actividades no superadas en convocatoria ordinaria, tras haber recibido las correcciones correspondientes a las mismas por parte del docente, o bien aquellas que no fueron entregadas.

## 8. CRONOGRAMA

En este apartado se indica el cronograma con fechas de entrega de actividades evaluables de la asignatura:

Actividades evaluables	Fecha
<b>Actividad 1.</b> Herramientas en R para análisis exploratorio	06/02/2023

<b>Actividad 2.</b> Análisis de datos genómicos	09/02/2023
<b>Actividad 3.</b> Análisis de expresión diferencial funcional	10/04 /2023
<b>Actividad 4.</b> Análisis de activación diferencial	10/04 /2023
<b>Test final de repaso (Examen de Módulo)</b>	10%

Este cronograma podrá sufrir modificaciones por razones logísticas de las actividades. Cualquier modificación será notificada al estudiante en tiempo y forma.

Seminarios	Fecha
Uso de herramientas de R y RStudio para investigación Biomédica	27/02/2023
Herramientas computacionales para el análisis de datos genómico	01/03/2023
Herramientas computacionales para el análisis de datos transcriptómicos	06/03/2023
Herramientas computacionales para el análisis de datos epigenéticos	08/03/2023
Análisis funcional de datos ómicos	13/03/2023
Análisis mecanístico de enfermedades con Hipathia	15/03/2023

## 9. BIBLIOGRAFÍA

A continuación, se indica bibliografía recomendada:

Alkan, C. et al. (2011). Genome structural variation discovery and genotyping. *Nature reviews Genetics*, 12:363-376

Chen, S., Zhou, Y., Chen, Y., Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics*, 34(17):i884–i890

Conesa A., Madrigal P., Tarazona S., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1):13

Gentleman, R. (2008). *R Programming for Bioinformatics*, CRC Press

Danecek, P. et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2) giab008

Dharshini S.A.P., Taguchi Y.H., Gromiha M. (2019) Identifying suitable tools for variant detection and differential gene expression using RNA-seq data. *Genomics*, 112(3):2166-2172

Evrny, G.D. et al. (2021). Applications of Single-cell DNA sequencing. *Annual Review of Genomics and Human Genetics*, 22:171-197

Goodwin, S. et al. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews*, 17: 333-351

Haque A., Engel J., Teichmann S.A., Lönnberg T. (2017) A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications

Liu Y., Ruan H., Li S., et al. (2020). The genetic and pharmacogenomic landscape of snoRNAs in human cancer. *Molecular Cancer*, 19(1), 108

Loftus, S.C. (2021). Basic Statistics with R, Academic Press

López-Domingo, F. J., Florido, J. P., Rueda, A., Dopazo, J. & Santoyo-Lopez, J. (2014). ngsCAT: a tool to assess the efficiency of targeted enrichment sequencing. *Bioinformatics*, 30(12):1767-1768

Nakato, R., & Sakata, T. (2021). Methods for ChIP-seq analysis: A practical workflow and advanced applications. *Methods (San Diego, Calif.)*, 187, 44–53.

Pabinger, S. et al. (2013). A survey of tools for variant analysis of next generation genome sequencing data. *Briefings in Bioinformatics*, 15(2):256-278

Singer B. D. (2019). A Practical Guide to the Measurement and Analysis of DNA Methylation. *American journal of respiratory cell and molecular biology*, 61(4), 417–428.

Wickham, H., Grolemund, G. (2017). R for Data Science. O'Reilly

## 10. UNIDAD DE ATENCIÓN A LA DIVERSIDAD

Estudiantes con necesidades específicas de apoyo educativo:

Las adaptaciones o ajustes curriculares para estudiantes con necesidades específicas de apoyo educativo, a fin de garantizar la equidad de oportunidades, serán pautadas por la Unidad de Atención a la Diversidad (UAD).

Será requisito imprescindible la emisión de un informe de adaptaciones/ajustes curriculares por parte de dicha Unidad, por lo que los estudiantes con necesidades específicas de apoyo educativo deberán contactar a través de: [unidad.diversidad@universidadeuropea.es](mailto:unidad.diversidad@universidadeuropea.es) al comienzo de cada semestre.

## 11. ENCUESTAS DE SATISFACCIÓN

¡Tú opinión importa!

La Universidad Europea te anima a participar en las encuestas de satisfacción para detectar puntos fuertes y áreas de mejora sobre el profesorado, la titulación y el proceso de enseñanza-aprendizaje.

Las encuestas estarán disponibles en el espacio de encuestas de tu campus virtual o a través de tu correo electrónico.

Tu valoración es necesaria para mejorar la calidad de la titulación.

Muchas gracias por tu participación.