



FICHA IDENTIFICATIVA

DATOS DE LA ASIGNATURA

Código: 33142
Nombre: Bioinformática
Ciclo: Grado
Créditos ECTS: 4,5
Curso académico: 2025-26

TITULACIONES

Titulación	Centro	Curso	Periodo
1109 - Grado en Bioquímica y Ciencias Biomédicas	Facultat de Ciències Biològiques	4	Segundo cuatrimestre

MATERIAS

Titulación	Materia	Carácter
1109 - Grado en Bioquímica y Ciencias Biomédicas	Materia de asignaturas optativas	OPTATIVA

COORDINACIÓN

LUCAS LLEDO JOSE IGNACIO

RESUMEN

La bioinformática es un campo científico interdisciplinar que desarrolla métodos y herramientas informáticas para la interpretación de datos biológicos complejos y abundantes, como las secuencias genómicas o las medidas de expresión génica. La bioinformática tiene aplicaciones en biomedicina, biología de la conservación, biología de sistemas, etc.

La asignatura está necesariamente diseñada como una introducción a esta disciplina tan amplia. En la selección y ordenamiento de los contenidos se han seguido dos principios fundamentales. Primero, que cualquier análisis bioinformático debe ser reproducible. Esta exigencia implica la adopción de buenas prácticas básicas, también útiles en otros contextos. En segundo lugar concebimos la bioinformática como una actividad colectiva, destinada a la colaboración, a menudo en un ambiente multidisciplinar. Es, por tanto, necesario saber comunicar los resultados de forma comprensible y ser capaces de colaborar en su producción. Esta característica también condiciona los hábitos fundamentales de trabajo con el ordenador y las herramientas utilizadas.

La asignatura está orientada a la práctica, con casi el doble de tiempo en aula de informática que en aula de teoría.

CONOCIMIENTOS PREVIOS

RELACIÓN CON OTRAS ASIGNATURAS DE LA MISMA TITULACIÓN



No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

OTROS TIPOS DE REQUISITOS

No se precisan conocimientos previos de Linux ni de ningún lenguaje de programación. El curso parte de cero, sin asumir ninguna experiencia previa.

COMPETENCIAS / RESULTADOS DE APRENDIZAJE

1101 -

Acceder a las principales bases de datos biológicas y recuperar y emplear la información contenida en ellas.

Aplicar correctamente los métodos de inferencia filogenética e interpretar los resultados.

Conocer los mecanismos evolutivos a escala molecular.

Conocer los métodos que permiten manejar grandes cantidades de datos derivados de las técnicas ómicas.

Saber utilizar las diferentes fuentes bibliográficas y bases de datos biológicas y usar las herramientas bioinformáticas.

Saber utilizar los principales métodos bioinformáticos.

DESCRIPCIÓN DE CONTENIDOS

1. Archivos de texto y alineamientos

Inconvenientes de los entornos gráficos de usuaria. El texto plano y los formatos FASTA, FASTQ, SAM y VCF. La línea de comandos de Bash. Los lenguajes interpretados. Nociones de R o python. Automatización de procesos (scripts). Alineamiento de secuencias.

2. Principios de gestión de datos

Inconvenientes de las hojas de cálculo. Metadatos. Buenas prácticas en la organización de carpetas. Permisos y propiedades de los archivos en Linux. Los formatos XML y JSON como modelos de bases de datos. Nociones de bases de datos relacionales. Operaciones con tablas o marcos de datos. Test exacto de Fisher.

3. Bases de datos públicas

El archivo europeo de nucleótidos (ENA). Uniprot. Interpro. Otros repositorios permanentes. Consultas programáticas



mediante interfaz de programación de aplicaciones.

4. Visualización y exploración de datos

Herramientas avanzadas de representación gráfica (e.g., ggplot2). Análisis exploratorios de datos (PCA, biplots, etc.). Grafos.

5. Reproducibilidad

Pilares de la reproducibilidad en bioinformática. El control de versiones con git. Sistemas de control del ambiente de computación.

6. BLAST

El algoritmo BLAST y sus implementaciones como servicio en línea y como intérprete de comandos. Construcción de una base de datos de BLAST. Búsquedas básicas y avanzadas.

7. Reconstrucción filogenética

Cadenas de Markov y evolución molecular. Modelos de sustitución nucleotídica. Distancias genéticas. Neighbor-joining. Máxima verosimilitud. Métodos Bayesianos.

8. Ensamblaje de secuencias

Principales métodos y conceptos relacionados con el ensamblaje de genomas. Grafos de De Bruijn y k-meros. Control de calidad.

9. Anotación de genomas

Identificar y enmascarar repeticiones. Identificar genes por homología y ab initio. La ontología génica. Los formatos GFF/GTF y BED. Visualizadores genómicos.

10. Análisis de la expresión génica

La secuenciación de ARN. Demultiplexing. Normalización. Datos de recuento. Expresión diferencial.

11. Análisis bioinformáticos de proteínas

Predicción de estructuras secundarias y terciarias. Interacciones proteicas. La base de datos PDB. Visualización de estructuras proteicas.

VOLUMEN DE TRABAJO (HORAS)

**ACTIVIDADES PRESENCIALES**

Actividad	Horas
Tutorías	3,00
Teoría	16,00
Aula informática	26,00
Total horas	45,00

ACTIVIDADES NO PRESENCIALES

Actividad	Horas
Asistencia a otras actividades	0,00
Elaboración de trabajos individuales o en grupo	12,00
Estudio y trabajo autónomo	31,50
Preparación de clases	24,00
Preparación de actividades de evaluación	0,00
Resolución de casos prácticos	0,00
Total horas	67,50

METODOLOGÍA DOCENTE

Hay dos tipos de contenidos teóricos: la teoría básica necesaria para practicar la bioinformática (cómo funcionan los algoritmos y los programas utilizados) y la reflexión teórica sobre la práctica (qué herramientas son más adecuadas, por qué hace falta garantizar la reproducibilidad, etc.). Por eso, la relación entre práctica y teoría será bidireccional. Las clases de teoría, mediante lección magistral, informarán las sesiones prácticas. Y éstas a su vez motivarán la reflexión teórica. Las sesiones prácticas combinarán *live-coding* y trabajo autónomo supervisado.

Las cinco primeras unidades temáticas servirán para desarrollar las habilidades prácticas básicas: el uso pasivo de un lenguaje de programación, el control de versiones, la manipulación de tablas y la representación gráfica. La automatización de los procesos y la programación literaria serán facilitadas por la distribución de scripts previamente elaborados por el personal docente.

Una vez garantizada la familiaridad con las herramientas básicas, se abordarán retos bioinformáticos más complejos (búsquedas de BLAST, reconstrucción filogenética, etc). En esta segunda etapa, continuarán utilizándose las herramientas básicas (programación literaria, control de versiones, etc.), para reforzar los buenos hábitos bioinformáticos.

Las sesiones de tutoría servirán para ayudar a elaborar el cuaderno digital de prácticas y repasar o ampliar algún contenido, a demanda del alumnado.

EVALUACIÓN

Se utilizarán dos instrumentos de evaluación. En primer lugar, el **cuaderno digital de prácticas**, donde los y las estudiantes registraran los procedimientos y los resultados de todos los ejercicios de prácticas. El cuaderno supondrá el 50% de la nota final.

El segundo instrumento de evaluación será un **examen teórico-práctico**, que constituirá el otro 50% de la nota final. Para superar la asignatura, será necesario obtener al menos 4 puntos sobre 10 en el examen y 5 sobre 10 globalmente.

En segunda convocatoria, el procedimiento de evaluación será idéntico. En caso de haber obtenido una nota de 5 o más sobre 10 en el examen o en el cuaderno durante la primera convocatoria, esa nota se guardará para la segunda convocatoria, y sólo se tendrá que recuperar la parte suspendida.



BIBLIOGRAFÍA

Referencias básicas

Referencia b1: <https://software-carpentry.org/lessons/>

Referencia b2:

https://ucdavis-bioinformatics-training.github.io/2022_February_Introduction_to_R_for_Bioinformatics/

Referencia b3: <https://a-little-book-of-r-for-bioinformatics.readthedocs.io/en/latest/index.html>

Referencia b4: <https://doi.org/10.18637/jss.v059.i10>

Referencia b5: <https://www.ahl27.com/OtherTutorials/articles/BuildingTrees.html>

Referencia b6: <https://doi.org/10.1371/journal.pcbi.1000424>

Referencia b7: <https://www.ncbi.nlm.nih.gov/books/NBK279690/>

Referencia b8: Haddock, S.H.D & Dunn, C.W. 2011. Practical Computing for Biologists. Sinauer Associates, Inc. Sunderland (MA). 568 pp.

Referencia b9: Allesina, S. & Wilmes, M. 2019. Computing skills for biologists; a toolbox. Princeton University Press. Princeton (NJ). 441 pp.

Referencias complementarias

Referencia c1: <https://www.h3abionet.org/training/ibt>

Referencia c2: <https://doi.org/10.1093/bib/bbad375>

Referencia c3: <https://datasciencebox.org/>

Referencia c4: <https://datacarpentry.github.io/R-ecology-lesson-alternative/index.html>

Referencia c5: <https://www.ahl27.com/CompGenomicsBioc2022/>

Referencia c6: <https://www2.decipher.codes/Tutorials.html>

Referencia c7: <https://www.huber.embl.de/msmb/>

Referencia c8: <https://doi.org/10.1080/00031305.2017.1375989>