Universitat Autònoma de Barcelona

2023/2024

Bioinformática

Código: 101951
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2500890 Genètica	OB	3	1

Contacto

Nombre: Sonia Casillas Viladerrams

Correo electrónico: sonia.casillas@uab.cat

Idiomas de los grupos

Puede consutarlo a través de este enlace. Para consultar el idioma necesitará introducir el CÓDIGO de la asignatura. Tenga en cuenta que la información es provisional hasta el 30 de noviembre del 2023.

Equipo docente

Antoni Barbadilla Prados

Marta Coronado Zamora

Oscar Conchillo Solé

Raquel Egea Sanchez

Prerrequisitos

Es conveniente que el alumnado repase los conceptos básicos adquiridos en las asignaturas de segundo:

- El módulo de Fundamentos de programación dentro de la asignatura Técnicas Instrumentales
- Genética Molecular de eucariotas
- Genética Molecular de procariotas
- Genética de poblaciones

Así también, esta asignatura se complementa con la asignatura Genómica, proteómica y interactómica que se hace en el mismo semestre, por lo que se requiere un co-seguimiento.

Es muy recomendable conocimientos de inglés para la lectura de información.

Será imprescindible el conocimiento a nivel básico de algún lenguaje de programación (preferiblemente Python) para poder seguir las sesiones, tanto teóricas como prácticas, y completar las actividades de evaluación continuada.

Objetivos y contextualización

La bioinformática -el tratamiento, el análisis informático y la interpretación y modelización de datos moleculares y genómicos- ha adquirido un protagonismo fundamental en la genética actual. La materia impartida durante este curso constituye una visión introductoria a la bioinformática. Los objetivos principales son:

- Proporcionar al alumnado los conocimientos bioinformáticos básicos que les permita tanto el uso de herramientas para realizar búsquedas de información en las bases de datos como abordar el análisis computacional de secuencias de ácidos nucleicos, proteínas y genomas
- Aplicar la programación de scripts para la automatización de tareas bioinformáticas que se han trabajado previamente en clase, con el fin de desarrollar el pensamiento computacional
- Dar una perspectiva del potencial de esta disciplina tanto en el ámbito de la investigación como profesional

Competencias

- Actuar con responsabilidad ética y con respeto por los derechos y deberes fundamentales, la diversidad y los valores democráticos.
- Actuar en el ámbito de conocimiento propio evaluando las desigualdades por razón de sexo/género.
- Actuar en el ámbito de conocimiento propio valorando el impacto social, económico y medioambiental.
- Capacidad de análisis y síntesis.
- Conocer y aplicar las herramientas ómicas de genómica, transcriptómica y proteómica.
- Desarrollar el aprendizaje autónomo.
- Describir e identificar las características estructurales y funcionales de los ácidos nucleicos y proteínas incluyendo sus diferentes niveles de organización.
- Describir la organización, evolución, variación interindividual y expresión del genoma humano.
- Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.
- Percibir la importancia estratégica, industrial y económica, de la genética y genómica en las ciencias de la vida, la salud y la sociedad.
- Razonar críticamente.
- Saber comunicar eficazmente, oralmente y por escrito.
- Utilizar e interpretar las fuentes de datos de genomas y macromoléculas de cualquier especie y comprender los fundamentos del análisis bioinformático para establecer las relaciones correspondientes entre estructura, función y evolución.
- Utilizar y gestionar información bibliográfica o recursos informáticos o de Internet en el ámbito de estudio, en las lenguas propias y en inglés.

Resultados de aprendizaje

1. Actuar con responsabilidad ética y con respeto por los derechos y deberes fundamentales, la diversidad y los valores democráticos.
2. Actuar en el ámbito de conocimiento propio evaluando las desigualdades por razón de sexo/género.
3. Actuar en el ámbito de conocimiento propio valorando el impacto social, económico y medioambiental.
4. Argumentar la trascendencia de los avances en la generación e interpretación de datos a escala genómica para la comprensión y la manipulación tecnológica de los organismos.
5. Desarrollar el aprendizaje autónomo.
6. Desarrollar la capacidad de análisis y síntesis.
7. Explicar y aplicar los métodos del análisis y anotación de genomas.
8. Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.
9. Listar y explicar el contenido de las bases de datos bioinformáticas y realizar búsquedas de información.
10. Razonar críticamente.
11. Saber comunicar eficazmente, oralmente y por escrito.

12. Utilizar e interpretar los resultados de las aplicaciones bioinformáticas para el análisis molecular de secuencias.
13. Utilizar las técnicas y las herramientas bioinformáticas que permiten describir y analizar el genoma humano.
14. Utilizar las técnicas, las herramientas y las metodologías que permiten describir, analizar e interpretar la enormes cantidades de datos producidos por la tecnologías de gran rendimiento.
15. Utilizar y gestionar información bibliográfica o recursos informáticos o de Internet en el ámbito de estudio, en las lenguas propias y en inglés.

Contenido

Tema 1. Introducción a la bioinformática. La bioinformática en la investigación genómica. Los tres vértices del análisis bioinformática. Los retos bioinformáticos de integración de datos.

Tema 2. Bases de datos de interés para las biociencias. Bases de datos bibliográficas. Construcción de estrategias de búsqueda. Bases de datos moleculares. Motores de búsqueda.

Tema 3. Formatos de secuencias. Extracción de secuencias de las bases de datos primarias. Formatos de secuencias. Herramientas para la conversión entre formatos.

Tema 4. Estadísticas de secuencias. Herramientas de visualización, edición y análisis de secuencias. Estadísticas de secuencias.

Tema 5. Alineamiento de secuencias. Comparación de secuencias. Alineamiento local y global para programación dinámica. Matrices de puntuación: identidad y similitud. Penalizaciones por gaps. Dot-plots.

Tema 6. Búsquedas por similitud. Algoritmos heurísticos para la búsqueda por similitud contra bases de datos: BLAST. Aplicaciones del BLAST. Búsqueda de homólogos remotos para PSI-Blast. Herramientas para la automatización de búsquedas por similitud.

Tema 7. Alineamiento múltiple. Alineamiento múltiple de secuencias. Alineamientos progresivos. Matrices de distancia. Paquete CLUSTAL. Aplicaciones del alineamiento múltiple. Herramientas para la automatización de alineamientos múltiples. Matrices de posición (peso) y LOGOS.

Tema 8. Reconstrucción filogenética molecular. La filogenia molecular. Métodos de inferencia filogenética. Herramientas para la reconstrucción filogenética. Ejemplos de reconstrucción filogenética.

Tema 9. Genómica comparativa y funcional. Diferentes niveles de comparación. Bloques sinténicos. Hotspots de reordenación. Vista comparativa mediante navegadores genómicos.

Tema 10. Ensamblaje de genomas. Algoritmos de mapeo contra genomas de referencia. Formatos y herramientas de visualización de ensamblajes.

Tema 11. Variabilidad genética. Descripción de la variabilidad genética a partir de archivos VCF. Cálculo de diferentes medidas de variabilidad y pruebas de neutralidad. Bases de datos de SNPs y fenotipos.

Metodología

La metodología docente incluye dos tipos de actividades diferenciadas: clases de teoría y clases prácticas de aula de informática. El aprendizaje también contará con una tutorización individual y/o colectiva del alumnado que servirá de apoyo para resolver cuestiones más concretas o que lo requieran por su complejidad o dificultad.

<u>Clases teóricas</u>

Clases para transmitir los conceptos básicos y la información necesaria para desarrollar un aprendizaje autónomo. Fomento de la participación activa del alumnado.

Clases de prácticas de aula de informática

Estas prácticas se organizarán a partir de problemas planteados por el profesorado que habrá que resolver usando las diferentes herramientas y análisis bioinformáticos.

Seminarios y/o conferencias

En las últimas sesiones del curso habrá seminarios y/o conferencias de investigadores en el campo de la bioinformática.

Tutorías

Individuales o en grupos pequeños para la resolución de dudas relacionadas con la asignatura. Este tipo de actividad se podrá realizar por petición del alumnado.

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.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases de teoría	25	1	4, 7, 9, 10, 6, 15, 12, 13, 14
Prácticas aula informática	27	1,08	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Seminarios	3	0,12	4, 7, 9, 10, 6, 15, 12, 13, 14
Tipo: Supervisadas			
Tutoría	4	0,16	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Tipo: Autónomas			
Elaboración de trabajos	25	1	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Estudio	24	0,96	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Lectura de textos	2	0,08	4, 5, 10, 11, 6, 15
Resolución de problemas computacionales	40	1,6	4, 5, 7, 9, 10, 6, 15, 12, 13, 14

Evaluación

El sistema de evaluación se organiza en cinco actividades principales y habrá, además, un examen de recuperación y una actividad opcional de mejora de nota final. Los detalles de las actividades son:

Actividades de evaluación principal

- Evaluaciones parciales (1 y 2). Peso global 60%

Las evaluaciones parciales son pruebas combinadas que pueden constar de preguntas de tipo test, de respuestas escritas y de resolución de problemas. En estas pruebas se evaluarán todos los contenidos impartidos en la asignatura en forma de prueba teórico-práctica.

Los peso de cada una de las evaluaciones será del 30%.

Para superar estas evaluaciones hay que alcanzar una nota mínima de 4,0 en cada una de ellas.

- Problemas computacionales (Rosalind). Peso global 25%

A lo largo de todo el curso el profesorado planteará problemas computacionales o cuestiones relativas a la materia impartida que el alumnado deberá resolver en forma de evaluaciones.

El carácter continuado de esta evaluación hace que no se pueda evaluar la asignatura a no ser que haya un participación mínima en un 50% de las pruebas propuestas.

- ADÁPTATE-Bioinformática (trabajo integrador). Peso global 15%

Esta actividad está vinculada al proyecto de innovación docente ADÁPTATE y da la opción a participar en una experiencia de aprendizaje y servicio en el segundo semestre, reconocida con 2 créditos de libre elección (https://adaptat.omicsuab.org).

El trabajo consistirá en la resolución de un problema que contendrá cuestiones relativas a los diferentes bloques temáticos presentados durante las clases de teoría y de prácticas.

Estos trabajos se realizarán en grupos de 3-5 alumnos y son de carácter autónomo. El profesorado tutorizará los trabajos y resolverá dudas y cuestiones de planteamiento.

El alumnado entregará los trabajos siguiendo las directrices de contenidos, presentación y plazos fijados.

Esta actividad es de carácter obligatorio y la no participación comportará la no evaluación de la asignatura.

- Prácticas aula de informática. Pueden restar hasta 1 punto de la nota final

Esta actividad es de carácter obligatorio y la falta no justificada o bien el no aprovechamiento de la actividad pueden restar hasta 1 punto de la nota final de la asignatura.

El carácter continuado de esta evaluación hace que no se pueda evaluar la asignatura a no ser que haya una participación mínima en un 50% de las sesiones propuestas.

- Problemas opcionales o de creación propia del alumnado, y participación activa en clase/fórum. Pueden sumar hasta 1 punto sobre la nota final

A lo largo del curso el profesorado propondrá problemas computacionales complejos que serán opcionales. Su resolución podrá sumar hasta 1 punto sobre la nota final de la asignatura.

<u>Examen de recuperación</u>

Para participar en la recuperación, el alumnado debe haber estado previamente evaluado en un conjunto de actividades el peso de las cuales equivalga a un mínimo de dos terceras partes de la calificación total de la asignatura.

Se podrán recuperar las evaluaciones 1 y/o 2 individualmente y la nota, si es >=4, hará promedio con las aprobadas en los parciales. En caso de que en la/las pruebas recuperadas no se llegue a un mínimo de 4,0, al no poder hacer el promedio, no se aprueba la asignatura.

La evaluación continua, trabajo integrador y asistencia a las prácticas en el aula de informática, por su carácter continuado, no son recuperables.

<u>Mejora de la calificación final</u>

El alumnado que, habiendo superado las evaluaciones 1 y 2, quiera mejorar su calificación final, podrá optar a una prueba final. Esta prueba incluirá la totalidad de la materia. No es posible mejorar la nota mediante trabajos u otros tipos de actividades.

El grado de dificultad de esta prueba se corresponderá con el objetivo de la misma y, por tanto, podrá ser superior al de las evaluaciones parciales.

El alumnado que se presenta en esta prueba renuncia a las calificaciones previas y por tanto, la nota de esta prueba de mejora será la que prevalecerá en la calificación final aunque sea inferior a la obtenida por parciales.

En ningún caso el alumnado que se presente a esta prueba de mejora de nota optará a la calificación de Matrícula de Honor.

Fórmula de ponderación de la nota final

Nota final = [(Evaluación 1 x 0,3) + (Evaluación 2 x 0,3) + (Evaluación continua x 0,25) + (Trabajo integrador x 0,15)]

Aprobado
La asignatura se considera aprobada si la nota final es > = 5.0

No evaluable

El alumnado obtendrá la calificación de "No Avaluable" cuando las actividades de evaluación realizadas tengan una ponderación inferior al 67% en la calificación final.

Evaluación única

La evaluación única consiste en una única prueba de síntesis en la que se evaluarán los contenidos de todo el programa de teoría y prácticas de la asignatura en forma de prueba teórico-práctica. La prueba constará de preguntas tipo test, de respuestas escritas y de resolución de problemas. La nota obtenida en esta prueba de síntesis supondrá el 60%de la nota final de la asignatura. Para superar esta prueba de síntesis hay que alcanzar una nota mínima de 4,0.

La evaluación de los Problemas computacionales (Rosalind) y ADÁPTATE-Bioinformática (trabajo integrador) seguirán el mismo proceso de la evaluación continua. La nota obtenida supondrá el 25% y el 15% de la nota final de la asignatura, respectivamente. El alumnado que se acoja a la evaluación única, podrá entregar estas dos evidencias juntas el mismo día que el fijado para la prueba de síntesis.

La prueba de evaluación única se realizará coincidiendo con la misma fecha fijada en calendario para la última prueba de evaluación parcial y se aplicará el mismo sistema de recuperación que para la evaluación convencional.

El alumnado que se acoja a la evaluación única debe realizar las Prácticas aula de informática (PLAB) en sesiones presenciales y en las mismas condiciones que en la evaluación convencional.

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Evaluación continuada	25	0	0	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Evaluación parcial 1 (contenidos teórico-prácticos)	30	0	0	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14

Evaluación parcial 2 (contenidos teórico-prácticos)	30	0	0	4, 5, 7, 9, 10, 11, 6, 15, 12, 13, 14
Problemas opcionales o de creación propia del estudiante	Pueden sumar hasta 1 punto sobre la nota final	0	0	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 6, 15, 12, 13, 14
Prácticas aula informática	Pueden restar hasta 1 punto de la nota final	0	0	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 6, 15, 12, 13, 14
Trabajo integrador	15	0	0	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 6, 15, 12, 13, 14

Bibliografía

Libros

- Attwood, TK. i Parry-Smith, J. 1999. Introduction to Biocomputing. Longman, UK. (*Introducción a la Bioinformática. Attwood y Parry-Smith. 2002. Pearson Educación, S.A.*)
- Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC. (read online for free: https://www.bioinformaticsalgorithms.org/)
- Cristianini, N. Y M. W. Hahn. 2007. Introduction to Computational Genomics. A case studies approach. Cambridge Univ. Press. (read online: https://eembdersler.files.wordpress.com/2012/02/introduction-to-computational-genomics-a-case-studies-a)
- Pevzer, P & R Shamir. 2011. Bioinformatics for biologists. Cambridge Univ. Press. (read online: http://docshare04.docshare.tips/files/18157/181579469.pdf)
- Samuelsson, T. 2012. Genomics and Bioinformatics. Cambridge Univ. Press.
- Scherer, S. 2008. A short guide to the Human Genome. Cold Spring Harbor Laboratory Press.
- Xiong, J. 2006. Essential bioinformatics. Cambridge Univ. Press.

MOOCS y recursos online:

- Biology meets Programming: Bioinformatics for Beginners (course by Compeau&Pevzner): https://www.coursera.org/learn/bioinformatics (fácil! para todos!)
- Guions de pràctiques de l'assignatura: https://stepik.org/course/1171/syllabus
- Journey to the Frontier of ComputationalBiology (Specialization by Compeau&Pevzner): https://www.coursera.org/specializations/computational-biology (difícil! para los que quieren un reto!)
- Problemes computacionals a Rosalind: http://rosalind.info/problems/locations/
- Web associada al llibre "Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC": https://www.bioinformaticsalgorithms.org/

Software

Programario básico

- Problemes computacionals a Rosalind: http://rosalind.info/problems/locations
- Llenguatge de programació Python (i BioPython)
- Anaconda
- Jupyter notebook

Bases de datos

- NCBI-Entrez Gene http://www.ncbi.nlm.nih.gov/gene
- NCBI-RefSeq http://www.ncbi.nlm.nih.gov/refseq/
- NCBI-Genome https://www.ncbi.nlm.nih.gov/genome/

- NCBI-Genbank https://www.ncbi.nlm.nih.gov/genbank/
- OMIM https://www.omim.org/
- GeneCards https://www.genecards.org/
- Geneid https://genome.crg.es/geneid.html
- Genscan http://hollywood.mit.edu/GENSCAN.html
- Pubmed http://www.ncbi.nlm.nih.gov/pubmed
- dbSNP https://www.ncbi.nlm.nih.gov/snp/
- ENSEMBL http://www.ensembl.org/
- PopHuman https://pophuman.uab.cat/
- PopHumanScan https://pophumanscan.uab.cat/
- EVA https://www.ebi.ac.uk/eva/?Home
- GWAS Catalogue https://www.ebi.ac.uk/gwas/docs/about
- ClinVar https://www.ncbi.nlm.nih.gov/clinvar/intro/
- DisGeNET https://www.disgenet.org/help
- Open Targets Platform https://platform.opentargets.org/
- RegulomeDB https://regulomedb.org/
- Human Genome Dating https://human.genome.dating/

Herramientas de análisis

- EMBOSS http://www.bioinformatics.nl/emboss-explorer/ , http://www.ebi.ac.uk/Tools/emboss/
- Galaxy https://usegalaxy.org/ (i moltes eines integrades en aquesta plataforma)
- BLAST http://blast.ncbi.nlm.nih.gov/Blast.cgi
- Transeq https://www.ebi.ac.uk/Tools/st/emboss_transeq/
- PSA https://www.ebi.ac.uk/Tools/psa/
- MSA https://www.ebi.ac.uk/Tools/msa/
- Clustal Omega https://www.ebi.ac.uk/Tools/msa/clustalo/
- Clustal W http://www.clustal.org/download/current/clustalw-2.1-win.msi
- MEGA http://www.megasoftware.net/