

Temas Actuales de la Bioinformática

Código: 105065
Créditos ECTS: 3

| Titulación | Tipo | Curso | Semestre |
|------------------|------|-------|----------|
| 2500890 Genètica | OT | 4 | 1 |

Contacto

Nombre: Marta Coronado Zamora

Correo electrónico: marta.coronado@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: español (spa)

Algún grupo íntegramente en inglés: Sí

Algún grupo íntegramente en catalán: No

Algún grupo íntegramente en español: Sí

Otras observaciones sobre los idiomas

55%

Equipo docente

Marta Coronado Zamora

Prerequisitos

- Se recomienda haber superado la asignatura de *Bioinformática* (3r curso de Genética), *Genómica*, *Proteómica e Interactómica* (3r curso de Genética) y el módulo de *Bases de datos y fundamentos de programación* dentro de la asignatura *Técnicas instrumentales* (2º curso de Genética).
- Es imprescindible el conocimiento a nivel básico de algún lenguaje de programación (preferiblemente Python) y estar familiarizado con el entorno Linux para poder seguir las sesiones prácticas y completar las actividades de evaluación continuada.
- Se recomienda un nivel B1.2 de inglés o equivalente.

Objetivos y contextualización

El propósito de esta asignatura es cubrir temas básicos de la bioinformática en forma de talleres prácticos, *workshops* y conferencias impartidas por expertos. No es una asignatura acumulativa sino transversal, cuyo objeto es proporcionar a los estudiantes la amplia gama de conceptos y enfoques que abarca la bioinformática.

El objetivo principal es el de proporcionar a los estudiantes los conocimientos y habilidades necesarias para aplicar la bioinformática en diferentes áreas de la investigación genómica y por extensión otras ómicas. La materia impartida y las actividades realizadas durante este curso suministran una perspectiva global del potencial de la bioinformática tanto en el ámbito de la investigación básica como aplicada.

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.
- Medir e interpretar la variación genética dentro y entre poblaciones desde una perspectiva clínica, de mejora genética de animales y plantas, de conservación y evolutiva.
- 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 como se aplica el conocimiento de la variación genética humana a la medicina personalizada, la farmacogenómica o la nutrigenómica.
8. Explicar y aplicar los métodos del análisis y anotación de genomas.
9. Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.
10. Listar y explicar el contenido de las bases de datos bioinformáticas y realizar búsquedas de información.
11. Razonar críticamente.
12. Saber comunicar eficazmente, oralmente y por escrito.
13. Utilizar e interpretar los resultados de las aplicaciones bioinformáticas para el análisis molecular de secuencias.
14. Utilizar las técnicas y las herramientas bioinformáticas que permiten describir y analizar el genoma humano.
15. 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.
16. Utilizar y gestionar información bibliográfica o recursos informáticos o de Internet en el ámbito de estudio, en las lenguas propias y en inglés.

Contenido

La asignatura se compondrá de sesiones teórico-prácticas, conferencias y *workshops* impartidos por reconocidos especialistas en las diferentes materias y ámbitos.

Sesiones teórico-prácticas (~12h)

Se llevarán a cabo en el aula de informática. Los alumnos trabajarán tanto de manera individual como en grupo (3-4 alumnos) promoviendo un aprendizaje activo que permita desarrollar la capacidad de análisis y síntesis, el razonamiento crítico y la capacidad de resolución de problemas.

Estamos a punto de iniciar un viaje hacia la bioinformática real. La asignatura se divide en una serie de cuatro actividades de formación práctica que mostrarán los flujos de trabajo básicos en bioinformática: desde la gestión y el procesamiento de datos con Linux, la visualización y los análisis funcionales posteriores. Las prácticas se dividen en dos grandes partes. Parte I: Conceptos básicos sobre los flujos de trabajo de bioinformática y Parte II: Resolución de casos reales de genómica.

Estas prácticas pretenden también adquirir otras habilidades, muy valiosas en la investigación pero pocas veces experimentadas durante el Grado, como colaborar, aprender a transformar los datos en visualizaciones efectivas para comunicar y hacer una investigación reproducible.

| | Título | Descripción y resultados del aprendizaje |
|----------------------------------------|------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | Introduction | Presentación de la asignatura: organización, metodología, preparaciones previas, creación de grupos y evaluación de la asignatura |
| P1. Basics on Bioinformatics workflows | Data management and processing | Learning Linux for Bioinformatics - Aprenderemos a tratar datos crudos a través de comandos de <i>bash</i> , un lenguaje muy potente de Linux. |
| | Data exploration and visualization | Data exploration and visualization - Aprenderemos a representar datos biológicos en un mensaje claro, una visualización, y extraer información de las mismas. Utilizaremos <i>ggplot2</i> , un paquete de R. |
| P2. Solving real cases in genomics | Genome-Wide Association studies | Which are the variants behind three common human diseases? Realizaremos un estudio de asociación del genoma completo con R para detectar SNPs asociados con enfermedades humanas complejas. |
| | Transcriptomic analyses | Finding differentially expressed genes in cancer. Realizaremos un análisis de expresión diferencial utilizando datos de cáncer. |
| | Tutoria* | *Se implantaron dos sesiones extras de dos horas según las necesidades del estudiante y las dificultades de los casos. |

Conferencias invitadas (10h)

Asistencia a al menos cuatro conferencias (2h/conferencia) de expertos invitados en el ámbito de la bioinformática que serán impartidas en inglés.

Talleres (Workshops) (7h)

| Título | Ponente | Fecha aproximada ¹ |
|----------------------------------------------|--------------------------------------------------|-------------------------------|
| ¿Cómo podemos analizar el genoma del cáncer? | Claudia Arnedo Estudiante de doctorado al IRB | Noviembre 2021 |
| Cómo sobrevivir a un doctorado. | Edgar Garriga Estudiante de doctorado al CRG | Diciembre 2021 |
| TBA | TBA | TBA |

¹Las fechas definitivas se irán actualizando en el calendario y serán avisadas a través de las herramientas de comunicación del espacio del Moodle. Las conferencias serán virtuales y podrán ser grabadas.

Metodología

Actividades de aprendizaje presenciales y aprendizaje autónomo

Se implementará una experiencia de aprendizaje cooperativo, en concreto se seguirá la metodología Puzzle: se proporciona conjuntos de datos y procedimientos de manera distribuida a grupos reducidos. Cada grupo debe gestionar y resolver casos prácticos de forma autónoma.

Cada módulo trabajará en paralelo conceptos similares a través de casos prácticos reales, de forma que una vez finalizados los estudiantes intercambiarán información sobre la metodología escogida, su desarrollo y los resultados obtenidos, tratando de conseguir una cooperación efectiva entre los estudiantes. Los miembros de cada grupo conocerán a fondo la información que les corresponde. Se hará una exposición y/o redacción de un portafolio a través del cual el resto de grupos comprenderá las características y fundamentos de cada análisis. Las cuatro sesiones de prácticas estarán vinculadas entre sí, pues los resultados de una práctica o los métodos utilizados servirán para la siguiente práctica.

La participación activa, la gestión del trabajo, así como la discusión de los conocimientos adquiridos formarán una parte vital en el rol desempeñado de cada alumno.

Conferencias y talleres

Se impartirán un total de 5 conferencias por expertos en sus respectivos campos de investigación o laboral que ofrecerán una visión real sobre la bioinformática como pieza clave en la resolución de cuestiones de investigaciones biológicas básicas y aplicadas. Se enfatizará la importancia del tratamiento de datos en la era actual de los macrodatos.

Los tres talleres consistirán de tres sesiones que tratarán de tres aspectos de interés práctico, como sobrevivir el doctorado, como divulgar ciencia con un ejemplo real y cómo funciona el mundo científico.

*La metodología docente propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

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 | | | |
| Talleres | 7 | 0,28 | 5, 8, 11, 12 |
| Conferencias | 10 | 0,4 | 4, 7, 8, 10 |
| Sesiones teórico-prácticas | 12 | 0,48 | 4, 5, 7, 8, 10, 11, 12, 6, 16, 13, 14, 15 |
| Tipo: Supervisadas | | | |
| Portafolio | 20 | 0,8 | 4, 5, 7, 8, 10, 11, 12, 6, 16, 13, 14, 15 |
| Tipo: Autónomas | | | |
| Estudio/Resolución de problemas | 25 | 1 | 5, 10, 11, 13, 14, 15 |

Evaluación

La evaluación se realizará a través de la entrega de cuatro portafolios y la exposición de un tema de bioinformática a escoger por el alumno.

Portafolio (70%). En cada portafolio se expondrá los fundamentos básicos de los datos analizados, las herramientas utilizadas, el desarrollo de la metodología, así como una discusión sobre el resultado final de la entrega. Cada portafolio tendrá el mismo peso en la evaluación final.

Exposición (20%). Cada grupo realizará una exposición oral de 15 minutos.

Asistencia y participación (10%).

La asignatura se supera cuando la nota media de las actividades de evaluación es igual o superior a 5. El carácter continuado y transversal de esta evaluación hace que no se pueda evaluar la asignatura si la participación mínima del alumnado es inferior a un 80% de las sesiones propuestas.

Actividades de evaluación

| Título | Peso | Horas | ECTS | Resultados de aprendizaje |
|-----------------------------------|------|-------|------|-------------------------------------------------------|
| Asistencia y participación activa | 10% | 0 | 0 | 4, 5, 7, 8, 11, 12, 6, 16 |
| Portafolio | 70% | 0 | 0 | 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 6, 16, 13, 14, 15 |
| Presentación oral seminario | 20% | 1 | 0,04 | 1, 2, 3, 4, 7, 8, 9, 11, 12, 13 |

Bibliografía

Libros

- Pevzner, P. and R. Shamir. 2011. Bioinformatics for Biologists. Cambridge University Press
- Samuelsson, T. 2012. Genomics and Bioinformatics: An Introduction to Programming Tools for Life Scientists
- Lesk, A. 2014. Introduction to bioinformatics. Oxford University Press
- Claverie, J-M. 2007. Bioinformatics for dummies. Wiley, cop
- González, JR., Cáceres, A. 2019. Omic association studies with R and Bioconductor. CRC Press
- Hadley, W. 2009. ggplot2: elegant graphics for data analysis. Springer

Artículos

- [Himmelblau, E. \(2021\) A cartoon guide to bioinformatics by a novice coder. Nature](#)
- [Markowitz, F. \(2017\) All biology is computational biology. PLoS Biology 15\(3\):e2002050](#)
- [Webb, S. \(2018\) Deep learning for biology. Nature 554:555-557](#)
- [Mardis, E.R. \(2010\) The \\$1,000 genome, the \\$100,000 analysis? Genome Medicine 2:84](#)
- [Tyler-Smith C, Yang H, Landweber LF, Dunham I, Knoppers BM, Donnelly P, et al. \(2015\) Where Next for Genetics and Genomics? PLoS Biology 13\(7\):e1002216](#)

Enlaces

- [Web presentación Màster Bioinformàtica UAB](#)
- [Sueldos en bioinformática](#)
- [Bioinformatics: What? Why? Who?](#)

Software

- Sistema operativo: Linux
- Lenguajes de programación: bash, R
- Programas: RStudio y Jupyter Notebook
- Paquetes de R: ggplot2, shiny, rmarkdown, knitr, BiocManager, SNPassoc, SNPrelate, DESeq2, edgeR, snpStats, limma