

**Temes Actuals de la Bioinformàtica**

Codi: 105065

Crèdits: 3

Titulació	Típus	Curs	Semestre
2500890 Genètica	OT	4	0

La metodologia docent i l'avaluació proposades a la guia poden experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

**Professor/a de contacte**

Nom: Marta Coronado Zamora

Correu electrònic: Marta.Coronado@uab.cat

**Utilització d'idiomes a l'assignatura**

Llengua vehicular majoritària: espanyol (spa)

Grup íntegre en anglès: No

Grup íntegre en català: Sí

Grup íntegre en espanyol: No

**Equip docent**

Marta Coronado Zamora

**Equip docent extern a la UAB**

Nerea Carrón Rodas

**Prerequisits**

- Es recomana haver superat l'assignatura de Bioinformàtica (3r curs de Genètica), Genòmica, Proteòmica i Interactòmica (3r curs de Genètica) i el mòdul de *Bases de dades* i fonaments de programació dins de l'assignatura Tècniques instrumentals (2n curs de Genètica).
- És imprescindible el coneixement a nivell bàsic d'algun llenguatge de programació (preferiblement Python) i estar familiaritzat amb l'entorn Linux per poder seguir les sessions pràctiques i completar les activitats d'avaluació continuada.
- Es recomana un nivell B1.2 d'anglès o equivalent.

**Objectius**

El propòsit d'aquesta assignatura és cobrir temes bàsics de la bioinformàtica en forma de tallers pràctics, *workshops* i conferències impartides per experts. No és una assignatura acumulativa sinó transversal, on l'objectiu és proporcionar als estudiants l'àmplia gamma de conceptes i enfocaments que abarca la bioinformàtica.

L'objectiu principal és el de proporcionar als estudiants els coneixements i habilitats necessàries per aplicar la bioinformàtica en diferents àrees de la investigació genòmica i per extensió altres òmiques. La matèria impartida i les activitats realitzades durant aquest curs subministren una perspectiva global del potencial de la bioinformàtica tant en l'àmbit de la recerca bàsica com l'aplicada.

## Competències

- "Conèixer i aplicar les eines ""òmiques"" de la genòmica, la transcriptòmica i la proteòmica."
- Descriure i identificar les característiques estructurals i funcionals dels àcids nucleics i les proteïnes incloent-hi els seus diferents nivells d'organització.
- Descriure l'organització, l'evolució, la variació interindividual i l'expressió del genoma humà.
- Desenvolupar l'aprenentatge autònom.
- Mesurar i interpretar la variació genètica dins i entre poblacions des d'una perspectiva clínica, de millora genètica d'animals i plantes, de conservació i evolutiva.
- Percebre la importància estratègica, industrial i econòmica de la genètica i de la genòmica en les ciències de la vida, la salut i la societat.
- Raonar críticament.
- Saber comunicar amb eficàcia, oralment i per escrit.
- Tenir capacitat d'anàlisi i de síntesi.
- Utilitzar i gestionar informació bibliogràfica o recursos informàtics o d'Internet en l'àmbit d'estudi, en les llengües pròpies i en anglès.
- Utilitzar i interpretar les fonts de dades de genomes i macromolècules de qualsevol espècie i comprendre els fonaments de l'anàlisi bioinformàtica per establir les relacions corresponents entre estructura, funció i evolució.

## Resultats d'aprenentatge

1. Argumentar la transcendència dels avenços en la generació i interpretació de dades a escala genòmica per a la comprensió i la manipulació tecnològica dels organismes.
2. Desenvolupar l'aprenentatge autònom.
3. Explicar com s'aplica el coneixement de la variació genètica humana a la medicina personalitzada, la farmacogenòmica o la nutrigenòmica.
4. Explicar i aplicar els mètodes de l'anàlisi i l'anotació de genomes.
5. Llistar i explicar el contingut de les bases de dades bioinformàtiques i fer recerques d'informació.
6. Raonar críticament.
7. Saber comunicar amb eficàcia, oralment i per escrit.
8. Tenir capacitat d'anàlisi i de síntesi.
9. Utilitzar i gestionar informació bibliogràfica o recursos informàtics o d'Internet en l'àmbit d'estudi, en les llengües pròpies i en anglès.
10. Utilitzar i interpretar els resultats de les aplicacions bioinformàtiques per a l'anàlisi molecular de seqüències.
11. Utilitzar les tècniques i les eines bioinformàtiques que permeten descriure i analitzar el genoma humà.
12. Utilitzar les tècniques, les eines i les metodologies que permeten descriure, analitzar i interpretar les enormes quantitats de dades produïdes per les tecnologies de gran rendiment.

## Continguts

L'assignatura es compon de sessions teòriques-pràctiques, conferències i workshops impartits per reconeguts especialistes en les diferents matèries i àmbits<sup>1</sup>.

Sessions teoricopràctiques (~12h)

Es duran a terme a l'aula d'informàtica<sup>2</sup>. Els alumnes treballaran tant de manera individual com en grup (3-4 alumnes) promovent un aprenentatge actiu que permeti desenvolupar la capacitat d'anàlisi i síntesi, el raonament crític i la capacitat de resolució de problemes.

Estem a punt d'iniciar un viatge cap a la bioinformàtica real. L'assignatura es divideix en una sèrie de quatre activitats de formació pràctica que mostraran els fluxos de treball bàsics en bioinformàtica: des de la gestió i el processament de dades amb Linux, la visualització i les anàlisis funcionals posteriors. Les pràctiques es divideixen en dues grans parts. Part I: Conceptes bàsics sobre els fluxos de treball de bioinformàtica i Part II: Resolució de casos reals de genòmica.

Aquestes pràctiques pretenen també adquirir altres habilitats, molt valuoses en la investigació però poques vegades experimentades durant el Grau, com ara col·laborar, aprendre a transformar les dades en visualitzacions efectives per comunicar i fer una investigació reproducible.

	Títol	Descripció i resultats de l'aprenentatge
	Introduction	Presentació de l'assignatura: organització, metodologia, preparacions prèvies, creació de grups i avaluació de l'assignatura
P1. Basics on Bioinformatics workflows	Data management and processing	Learning Linux for Bioinformatics - Aprendrem a tractar dades crues a través de comandes de <i>bash</i> , un llenguatge molt potent de Linux.
	Data exploration and visualization	Data exploration and visualization - Aprendrem a representar dades biològiques en un missatge clar, una visualització, i extreure'n informació de les mateixes. Utilitzarem <i>ggplot2</i> , un paquet de R.

P2. Solving real cases in genomics	Genome-Wide Association studies	Which are the variants behind three common human diseases? Realitzarem un estudi d'associació del genoma complet amb R per detectar SNPs associats amb malalties humanes complexes.
	Transcriptomic analyses	Finding differentially expressed genes in cancer. Realitzarem un anàlisi d'expressió diferencial utilitzant dades de càncer.
	Tutoria*	*S'implantaren dues sessions extres de dues hores segons les necessitats de l'estudiant i les dificultats dels casos.

### Conferències convidades (10h)

Assistència a almenys quatre conferències (2h/conferència) d'experts convidats en l'àmbit de la bioinformàtica que seran impartides en anglès.

### Tallers (**workshops**) (7h)

Títol	Ponent	Data aproximada <sup>3</sup>
¿Cómo podemos analizar el genoma del cáncer?	Claudia Arnedo Estudiant de doctorat al IRB	Novembre 2021
Cómo sobrevivir a un doctorado.	Edgar Garriga Estudiant de doctorat al CRG	Desembre 2021
¡Esto no me lo habían dicho! Lo que no está escrito sobre la cultura científica.	Rosana Garrido Professora associada UAB	Gener 2022

<sup>1</sup>Llevat que les restriccions imposades per les autoritats sanitàries obliguin a una prioritació o reducció d'aquests continguts.

<sup>2</sup>Llevat que les restriccions imposades per les autoritats sanitàries obliguin a realitzar-les en l'aula de teoria normal amb un ordinador portàtil.

<sup>3</sup>Les dates definitives s'aniran actualitzant al calendari i seran avisades a través de les eines de comunicació de l'espai del Moodle. Les conferències seran virtuals i podran ser gravades.

## Metodologia

Activitats d'aprenentatge presencials i aprenentatge autònom

S'implementarà una experiència d'aprenentatge cooperatiu, en concret es seguirà la metodologia *Puzzle*: es proporciona conjunts de dades i procediments de manera distribuïda a grups reduïts. Cada grup ha de gestionar i resoldre casos pràctics de forma autònoma.

Cada mòdul treballarà en paral·lel conceptes similars a través de casos pràctics reals, de manera que un cop finalitzats els estudiants intercanviaran informació sobre la metodologia escollida, el seu desenvolupament i els resultats obtinguts, tractant d'aconseguir una cooperació efectiva entre els estudiants. Els membres de cada grup coneixeran a fons la informació que els correspon. S'efectuarà una exposició i/o redacció d'un portafolis a través del qual la resta de grups compregui les característiques i fonaments de cada anàlisi. Les quatre sessions de pràctiques estaran vinculades entre si, ja que els resultats o la metodologia d'una pràctica serviran per a la següent.

La participació activa, la gestió del treball, així com la discussió dels coneixements adquirits formaran una part vital en el paper exercit de cada alumne.

#### Conferències i tallers

S'impartiran un total de 5 conferències per experts en els seus camps d'investigació o laboral que oferiran una visió real de la bioinformàtica com una peça clau en la resolució de qüestions d'investigacions biològiques bàsiques i aplicades. Es farà èmfasi en la importància del tractament de dades en l'era actual del *big data*.

Els tres tallers consistiran de tres sessions que tractaran de tres aspectes d'interès pràctic, com sobreviure el doctorat, com divulgar ciència amb un exemple real i com funciona el món científic.

\*La metodologia docent proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

### Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Tallers	7	0,28	2, 4, 6, 7
Conferències	10	0,4	1, 3, 4, 5
Sessions teòrica-pràctiques	12	0,48	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Tipus: Supervisades			
Portafoli	20	0,8	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Tipus: Autònomes			
Estudi/resolució problemes	25	1	2, 5, 6, 10, 11, 12

### Avaluació

L'avaluació es realitzarà a través del lliurament de quatre portafolis i l'exposició d'un tema de bioinformàtica a escollir per l'alumne.

Portafolis (70%). A cada portafolis s'exposarà els fonaments bàsics de les dades analitzades, les eines utilitzades, el desenvolupament de la metodologia, així com una discussió sobre el resultat final del lliurament. Cada portafolis tindrà el mateix pes en l'avaluació final.

Exposició (20%). Cada grup realitzarà una exposició oral de 20 minuts.

Assistència i participació (10%).

L'assignatura se supera quan la nota mitjana de les activitats d'avaluació és igual o superior a 5. El caràcter continuat i transversal d'aquesta avaluació fa que no es pugui avaluar l'assignatura si la participació mínima de l'alumnat és inferior a un 80% de les sessions proposades.

\*L'avaluació proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

## Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Assistència i participació activa	10%	0	0	1, 2, 3, 4, 6, 7, 8, 9
Exposició seminari	20%	1	0,04	1, 3, 4, 6, 7, 10
Portafolis	70%	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

## Bibliografia

### Llibres

- 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

### Articles

- [Himmelblau, E. \(2021\) A cartoon guide to bioinformatics by a novice coder. Nature](#)
- [Markowetz, 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](#)

### Enllaços

- [Web presentació Màster Bioinformàtica UAB](#)
- [Sous en bioinformàtica](#)
- [Bioinformatics: What? Why? Who?](#)

## Programari

- Sistema operatiu: Linux
- Llenguatges de programació: bash, R
- Programes: RStudio y Jupyter Notebook
- Paquets de R: ggplot2, shiny, rmarkdown, knitr, BiocManager, SNPassoc, SNPRelate, DESeq2, edgeR, snpStats, limma