

**FITXA IDENTIFICATIVA****DADES DE L'ASSIGNATURA**

**Codi:** 33190  
**Nom:** Bioinformàtica  
**Cicle:** Grau  
**Crèdits ECTS:** 4,5  
**Curs acadèmic:** 2025-26

**TITULACIONS**

Titulació	Centre	Curs	Període
1111 - Grau en Biotecnologia	Facultat de Ciències Biològiques	4	Segon quadrimestre

**MATÈRIES**

Titulació	Matèria	Caràcter
1111 - Grau en Biotecnologia	Optativitat	OPTATIVA

**COORDINACIÓ**

LUCAS LLEDO JOSE IGNACIO

**RESUM**

La bioinformàtica és un camp científic interdisciplinari que desenvolupa mètodes i eines informàtiques per la interpretació de dades biològiques complexes i abundants, com són les seqüències genòmiques o les mesures d'expressió gènica. La bioinformàtica té aplicacions en: biomedicina, biologia de la conservació, biologia de sistemes, etc.

L'assignatura està necessàriament dissenyada com una introducció a esta disciplina tan ampla. A l'hora de seleccionar i ordenar els continguts s'han seguit dos principis fonamentals. Primer, que qualsevol anàlisi bioinformàtica ha de ser reproducible. Aquesta exigència implica l'adopció de bones pràctiques bàsiques, que també resulten útils en altres contextos. En segon lloc, concebim la bioinformàtica com una activitat col·lectiva, destinada a la col·laboració, sovint en un ambient multidisciplinari. Per tant, cal saber comunicar els resultats de forma intel·ligible i ser capaços de col·laborar en la seua producció. Aquesta característica també condiona els hàbits fonamentals de treball amb l'ordinador i les eines utilitzades.

L'assignatura està orientada a la pràctica, amb el doble de temps en aula d'informàtica que en aula de teoria.

**CONEIXEMENTS PREVIS****RELACIÓ AMB ALTRES ASSIGNATURES DE LA MATEIXA TITULACIÓ**



No s'ha especificat restriccions de matrícula amb altres assignatures del pla d'estudis.

## ALTRES TIPUS DE REQUISITS

No calen coneixements previs de Linux ni de cap llenguatge de programació. El curs parteix de zero, sense assumir cap experiència prèvia.

## COMPETÈNCIES / RESULTATS D' APRENENTATGE

-

Actuar con autonomía en el aprendizaje, tomando decisiones fundamentadas en diferentes contextos, emitiendo juicios en base a la experimentación y el análisis y transfiriendo el conocimiento a nuevas situaciones

Analitzar a nivell molecular el resultat de la manipulació d'un organisme.

Capacitat d'interpretar dades rellevants.

Capacitat per transmetre idees, problemes i solucions dins la biotecnologia.

Colaborar eficazmente en equipos de trabajo, asumiendo responsabilidades y funciones de liderazgo y contribuyendo a la mejora y desarrollo colectivo

Contribuir en el diseño, desarrollo y ejecución de soluciones que den respuesta a demandas sociales, teniendo en cuenta como referente los Objetivos de Desarrollo Sostenible

Demostrar razonamiento crítico y autocrítico en el ámbito de la titulación, considerando aspectos tales como la ética profesional, los valores morales y las implicaciones sociales de las diferentes actividades realizadas

Posseir i comprendre els coneixements en biotecnologia.

Proposar solucions creatives i innovadores a situacions o problemes complexos, propis de l'àmbit de coneixement, per a donar resposta a les diverses necessitats professionals i socials

Saber aplicar aquests coneixements al món professional.

Saber comunicarse de manera efectiva, tanto de forma oral como escrita, adaptándose a las características de la situación y de la audiencia

Saber utilitzar la llengua anglesa en la redacció d'informes i per interpretar informació a partir de protocols, manuals i bases de dades.

Ser capaç d'abordar l'anàlisi de l'estructura de macromolècules a fi de modificar-la amb finalitats biotecnològiques.

## DESCRIPCIÓ DE CONTINGUTS



## 1. Arxius de text i alineaments

Inconvenients dels entorns gràfics d'usuària. El text pla i els formats FASTA, FASTQ, SAM i VCF. La línia de comandaments de Bash. Els llenguatges interpretats. Nocions d'R o python. Automatització de processos (scripts). Alineament de seqüències.

## 2. Principis de gestió de dades

Inconvenients dels fulls de càlcul. Metadades. Bones pràctiques en l'organització de carpetes. Permisos i propietats dels arxius en Linux. Els formats XML i JSON com a models de bases de dades. Nocions de bases de dades relacionals. Operacions amb taules o marcs de dades. Test exacte de Fisher.

## 3. Bases de dades públiques

L'arxiu europeu de nucleòtids (ENA). Uniprot. Interpro. Altres repositoris permanents. Consultes programàtiques mitjançant interfície de programació d'aplicacions.

## 4. Visualització i exploració de dades

Predicció d'estructura secundària i terciària. Interaccions proteiques. La base de dades PDB. Visualització d'estructures proteiques.

Eines avançades de representació gràfica (e.g. ggplot2). Anàlisis exploratòries de dades (PCA, biplots, etc.). Grafes.

## 5. Reproduïbilitat

Pilars de la reproduïbilitat en bioinformàtica. El control de versions amb git. Sistemes de control de l'ambient de computació.

## 6. BLAST

L'algorisme BLAST i les seues implementacions com un servei en línia i com interfície de línia d'ordres. Construcció d'una base de dades de BLAST. Cerques bàsiques i avançades.

## 7. Reconstrucció filogenètica

Cadenes de Markov i evolució molecular. Models de substitució. Distàncies genètiques. Neighbor-joining. Màxima versemblança. Mètodes Bayesians.

## 8. Assemblatge de seqüències

Principals mètodes i conceptes relacionats amb l'assemblatge de genomes. Grafes de De Bruijn i  $k$ -mers. Controls de qualitat.



## 9. Anotació de genomes

Identificar i emmascarar repeticions. Identificar gens per homologia i ab initio. L'ontologia gènica. Els formats GFF/GTF i BED. Visualitzadors genòmics.

## 10. Anàlisi de l'expressió gènica

La seqüenciació d'ARN. Demultiplexing. Normalització. Dades de recompte. Expressió diferencial.

## 11. Anàlisis bioinformàtiques de proteïnes

Predicció d'estructura secundària i terciària. Interaccions proteiques. La base de dades PDB. Visualització d'estructures proteiques.

### VOLUM DE TREBALL (HORES)

#### ACTIVITATS PRESENCIALS

Activitat	Hores
Tutories	3,00
Teoria	16,00
Aula informàtica	26,00
<b>Total hores</b>	<b>45,00</b>

#### ACTIVITATS NO PRESENCIALS

Activitat	Hores
Assistència a altres activitats	0,00
Elaboració de treballs individuals o en grup	12,00
Estudi i treball autònom	31,50
Preparació de classes	0,00
Preparació d'activitats d'avaluació	24,00
Resolució de casos pràctics	0,00
<b>Total hores</b>	<b>67,50</b>

### METODOLOGIA DOCENT

Les classes de teoria, mitjançant lliçó magistral, informaran les sessions pràctiques. I aquestes alhora motivaran la reflexió teòrica. Les sessions pràctiques combinaran *live coding* i treball autònom supervisat.

Les cinc primeres unitats temàtiques serviran per desenvolupar les habilitats pràctiques bàsiques: l'ús passiu d'un llenguatge de programació, el control de versions, la manipulació de taules i la representació gràfica. L'automatització dels processos i la programació literària estaran facilitades per la distribució d'scripts prèviament elaborats pel personal docent.

Una vegada garantida la familiaritat amb les eines bàsiques, s'abordaran repetes bioinformàtics més complexos (cerques de BLAST, reconstrucció filogenètica, etc). En aquesta segona etapa, continuaran utilitzant-se les eines bàsiques (la programació literària, el control de versions, etc), per tal de reforçar els



bons hàbits bioinformàtics.

Les sessions de tutoria serviran per ajudar a elaborar el quadern digital de pràctiques, i repassar o ampliar algun contingut, a demanda de l'alumnat.

## AVALUACIÓ

S'utilitzaran dos instruments d'avaluació. En primer lloc, el **quadern digital de pràctiques**, on els i les estudiants registraran el procediment i els resultats de tots els exercicis de pràctiques. El quadern suposarà el 50% de la nota final.

El segon instrument d'avaluació serà un **examen teòrico-pràctic**, que constituirà l'altre 50% de la nota final. Per superar l'assignatura, cal obtenir almenys 4 punts sobre 10 a l'examen i 5 sobre 10 globalment.

En segona convocatòria el procediment d'avaluació serà idèntic. En cas d'haver arribat a una nota de 5 o més en l'examen o en el quadern durant la primera convocatòria, eixa nota es guardarà per a la segona convocatòria, i només caldrà recuperar la part suspesa.

## BIBLIOGRAFIA

- <https://software-carpentry.org/lessons/>
- [https://ucdavis-bioinformatics-training.github.io/2022\\_February\\_Introduction\\_to\\_R\\_for\\_Bioinformatics](https://ucdavis-bioinformatics-training.github.io/2022_February_Introduction_to_R_for_Bioinformatics)
- <https://a-little-book-of-r-for-bioinformatics.readthedocs.io/en/latest/index.html>
- <https://doi.org/10.18637/jss.v059.i10>
- <https://www.ahl27.com/OtherTutorials/articles/BuildingTrees.html>
- <https://doi.org/10.1371/journal.pcbi.1000424>
- <https://www.ncbi.nlm.nih.gov/books/NBK279690/>
- <https://www.h3abionet.org/training/ibt>
- <https://doi.org/10.1093/bib/bbad375>
- <https://datasciencebox.org/>
- <https://datacarpentry.github.io/R-ecology-lesson-alternative/index.html>
- <https://www.ahl27.com/CompGenomicsBioc2022/>
- <https://www2.decipher.codes/Tutorials.html>
- <https://www.huber.embl.de/msmb/>
- <https://doi.org/10.1080/00031305.2017.1375989>