



## FITXA IDENTIFICATIVA

### DADES DE L'ASSIGNATURA

**Codi:** 33142  
**Nom:** Bioinformàtica  
**Cicle:** Grau  
**Crèdits ECTS:** 4,5  
**Curs acadèmic:** 2025-26

### TITULACIONS

Titulació	Centre	Curs	Període
1109 - Grau en Bioquímica i Ciències Biomèdiques	Facultat de Ciències Biològiques	4	Segon quadrimestre

### MATÈRIES

Titulació	Matèria	Caràcter
1109 - Grau en Bioquímica i Ciències Biomèdiques	Matèria d'assignatures optatives	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 àmplia. A l'hora de seleccionar i ordenar els continguts s'han seguit dos principis fonamentals. Primer, que qualsevol anàlisi bioinformàtica ha de ser reproduïble. 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 entenable 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 quasi 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

-

Accedir a les principals bases de dades biològiques i recuperar i usar la informació que contenen.

Aplicar correctament els mètodes d'inferència filogenètica i interpretar-ne els resultats.

Conèixer els mecanismes evolutius a escala molecular.

Conèixer els mètodes que permeten manejar grans quantitats de dades derivades de les tècniques òmiques.

Saber utilitzar els principals mètodes bioinformàtics.

Saber utilitzar les diferents fonts bibliogràfiques i bases de dades biològiques i usar les eines bioinformàtiques.

## 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**

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

#### **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. Grafs 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	24,00
Preparació d'activitats d'avaluació	0,00
Resolució de casos pràctics	0,00
<b>Total hores</b>	<b>67,50</b>

## METODOLOGIA DOCENT

Hi ha dos tipus de continguts teòrics: la teoria bàsica necessària per practicar la bioinformàtica (com funcionen els algoritmes i els programes utilitzats, etc) i la reflexió teòrica sobre la pràctica (quines eines són més adequades, per què cal garantir la reproduïbilitat, etc.). Per això, la relació entre pràctica i teoria serà bidireccional. 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 els procediments 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

### Referències bàsiques

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

Referència b2:

[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/)

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

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

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

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

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

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

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

### Referències complementàries

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

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

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

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

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

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

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

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