



Institut  
d'Estudis  
Catalans



Societat Catalana  
de **BIOLOGIA**

## Curs d'Introducció a la bioinformàtica per a viròlegs

10, 11 i 19 de juny de 2025

En línia

### Objectiu del curs

Aquest curs d'introducció a la bioinformàtica està enfocat a l'ús de R com a llenguatge de programació per l'anàlisi estadístic i visualització de dades. L'objectiu del curs és que els alumnes aprenguin els conceptes bàsics d'aquest llenguatge de codi obert àmpliament utilitzat en recerca científica. Es treballarà la introducció de dades a R, els diferents tipus de dades i el seu processament, així com la visualització de gràfics. A més, s'introduirà l'anàlisi de dades de transcriptòmica obtingudes per seqüenciació (RNA-sequencing) i l'anàlisi de variants en genomes de virus. A través de metodologies actuals, es cobrirà tot el procés, des de l'obtenció de dades crues fins a la interpretació dels resultats mitjançant eines com R i Galaxy. El curs està dirigit a investigadors i investigadores clínics i biomèdics que vulguin adquirir conceptes bàsics en l'ús aplicat de l'anàlisi i la interpretació de dades en R, amb especial èmfasi en tècniques utilitzades en virologia i transcriptòmica.

### Metodologia

El curs constarà de dues sessions teoricopràctiques en línia (4 hores) i una sessió totalment pràctica en format asíncron (2 hores) amb tutoria en línia (1 hora). A les sessions en línia es revisaran conceptes fonamentals d'estadística, l'anàlisi de dades amb R i l'anàlisi de dades de transcriptòmica basant-se en casos reals. A la pràctica asíncrona es desenvoluparan exemples que abracin tot el procés, des de l'obtenció de les dades fins a la seva anàlisi i interpretació dels resultats, que l'alumnat realitzarà amb l'ajut del personal docent.

### Suport / Material

El material necessari per al curs es penjarà amb antelació en una carpeta al núvol accessible per a l'alumnat.

### Organització

El curs està organitzat per la Secció de Virologia de la Societat Catalana de Biologia.

### Equip docent

Doctora Ana Maria Corraliza Márquez (Ajuntament de Barcelona)

Doctora Marta Coronado Zamora (Institut Botànic de Barcelona, CSIC-CMCNB)

## Programa

**Dimarts 10 de juny: 16-18 h (sessió en línia)**

**Sessió 1: Introducció al llenguatge de programació R i RStudio: paquet estadístic i visualització de dades**

- Funcionament bàsic de R i RStudio
- Tipus de dades i estructures en R
- Breu orientació estadística amb l'ús de R
- Visualització de dades amb ggplot2

**Dimecres 11 de juny: 16-18 h (sessió en línia)**

**Sessió 2: Transcriptòmica – RNA-sequencing**

- Base de dades GEO (estructura, recuperació de dades)
- Mètodes de normalització de dades d'expressió
- Normalitzacions
- Anàlisi d'expressió diferencial (amb el paquet DESeq2)
- Interpretació del *log-fold change*
- *P*-valor i problemes de multiplicitat (*p*-valor ajustat)
- Enriquiment funcional amb R (clusterProfiler)

**Sessió 3: Anàlisi bioinformàtic amb dades reals: *variant calling* en genomes virals i anàlisi de RNA-seq. Pràctica asincrònica.**

En aquesta sessió, explorarem dues pràctiques de recerca en virologia basades en dades reals.

**Identificació de mutacions en el virus SARS-CoV-2.** Utilitzarem dades de seqüenciació del SARS-CoV-2 per identificar i anotar mutacions en el genoma. Aprendre a descarregar dades massivament i de manera eficient mitjançant la plataforma Galaxy, així com a realitzar les anàlisis necessàries per identificar les variacions genètiques presents a les mostres seqüenciades.

**Anàlisi d'expressió diferencial amb dades de RNA-seq.** Analitzarem dades reals de RNA-seq en mostres control i mostres infectades amb un virus per detectar canvis en l'expressió gènica. Utilitzarem la plataforma Galaxy per obtenir i preprocessar les dades, incloent la neteja d'adaptadors, el control de qualitat dels *reads*, el mapeig al genoma i la quantificació de l'expressió gènica. A continuació, farem l'anàlisi d'expressió diferencial amb R per identificar gens diferencialment expressats i realitzar anàlisis d'enriquiment funcional.

Aquestes anàlisis combinaran l'ús de Galaxy i R, proporcionant un enfocament aplicat i versàtil que es pot extrapolar a altres problemes d'interès per als participants.

**Dijous 19 de juny: 16 a 17h (sessió en línia)**

Tutoria en línia 16-17 h (sessió en línia)

Espai per resoldre dubtes sobre el curs i la sessió pràctica realitzada.