

Bioinformàtica

Codi: 101951

Crèdits: 6

Titulació	Típus	Curs	Semestre
2500890 Genètica	OB	3	1

Professor/a de contacte

Nom: Sonia Casillas Viladerrams

Correu electrònic: sonia.casillas@uab.cat

Idiomes dels grups

Podeu accedir-hi des d'aquest [enllaç](#). Per consultar l'idioma us caldrà introduir el CODI de l'assignatura. Tingueu en compte que la informació és provisional fins a 30 de novembre de 2023.

Equip docent

Antoni Barbadilla Prados

Marta Coronado Zamora

Oscar Conchillo Solé

Raquel Egea Sanchez

Prerequisits

És convenient que l'alumnat repassi els conceptes bàsics assolits a les assignatures de segon:

- el mòdul de *Fonaments de programació* dins de l'assignatura Tècniques Instrumentals
- Genètica Molecular d'eucariotes
- Genètica Molecular de procariotes
- Genètica de poblacions

Així també, aquesta assignatura es complementa amb la assignatura Genòmica, proteòmica i interactòmica que es fa al mateix semestre, per la qual cosa es requereix un co-seguiment.

És molt recomanable coneixements d'anglès per a la lectura d'informació.

Serà imprescindible el coneixement a nivell bàsic d'algun llenguatge de programació (preferiblement Python) per tal de poder seguir les sessions, tan teòriques com pràctiques, i completar les activitats d'avaluació continuada.

Objectius

La bioinformàtica -el tractament, l'anàlisi informàtica i la interpretació i modelització de dades moleculars i genòmiques- ha adquirit un protagonisme fonamental a la genètica actual. La matèria impartida durant aquest curs constitueix una visió introductòria a la bioinformàtica. Els objectius principals són:

- Proporcionar a l'alumnat els coneixements bioinformàtics bàsics que els permeti tant l'ús d'eines per a realitzar cerques d'informació a les bases de dades com abordar l'anàlisi computacional de seqüències d'àcids nucleics, proteïnes i genomes
- Aplicar la programació de scripts per a l'automatització de tasques bioinformàtiques que s'han treballat prèviament a classe, amb la fi de desenvolupar el pensament computacional
- Donar una perspectiva del potencial d'aquesta disciplina tant en l'àmbit de la recerca com professional

Competències

- "Conèixer i aplicar les eines "òmiques" de la genòmica, la transcriptòmica i la proteòmica."
- Actuar amb responsabilitat ètica i amb respecte pels drets i deures fonamentals, la diversitat i els valors democràtics.
- Actuar en l'àmbit de coneixement propi avaluant les desigualtats per raó de sexe/gènere.
- Actuar en l'àmbit de coneixement propi valorant l'impacte social, econòmic i mediambiental.
- 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.
- Introduir canvis en els mètodes i els processos de l'àmbit de coneixement per donar respostes innovadores a les necessitats i demandes de la societat.
- 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. Actuar amb responsabilitat ètica i amb respecte pels drets i deures fonamentals, la diversitat i els valors democràtics.
2. Actuar en l'àmbit de coneixement propi avaluant les desigualtats per raó de sexe/gènere.
3. Actuar en l'àmbit de coneixement propi valorant l'impacte social, econòmic i mediambiental.
4. 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.
5. Desenvolupar l'aprenentatge autònom.
6. Explicar i aplicar els mètodes de l'anàlisi i l'anotació de genomes.
7. Introduir canvis en els mètodes i els processos de l'àmbit de coneixement per donar respostes innovadores a les necessitats i demandes de la societat.
8. Llistar i explicar el contingut de les bases de dades bioinformàtiques i fer recerques d'informació.
9. Raonar críticament.
10. Saber comunicar amb eficàcia, oralment i per escrit.
11. Tenir capacitat d'anàlisi i de síntesi.
12. Utilitzar i gestionar informació bibliogràfica o recursos informàtics o d'Internet en l'àmbit d'estudi, en les llengües pròpies i en anglès.

13. Utilitzar i interpretar els resultats de les aplicacions bioinformàtiques per a l'anàlisi molecular de seqüències.
14. Utilitzar les tècniques i les eines bioinformàtiques que permeten descriure i analitzar el genoma humà.
15. 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

Tema 1. Introducció a la bioinformàtica. La bioinformàtica a la recerca genòmica. Els tres vèrtexs de l'anàlisi bioinformàtica. Els reptes bioinformàtics d'integració de dades.

Tema 2. Bases de dades d'interès per a les biociències. Bases de dades bibliogràfiques. Construcció d'estratègies de cerca. Bases de dades moleculars. Motors de cerca.

Tema 3. Formats de seqüències. Extracció de seqüències de les bases de dades primàries. Formats de seqüències. Eines per la conversió entre formats.

Tema 4. Estadístiques de seqüències. Eines de visualització, edició i anàlisi de seqüències. Estadístiques de seqüències.

Tema 5. Alineament de seqüències. Comparació de seqüències. Alineament local i global per programació dinàmica. Matrius de puntuació: identitat i similitud. Penalitzacions per *gaps*. Dot-plots.

Tema 6. Cerques per similitud. Algoritmes heurístics per a la cerca per similitud contra bases de dades: BLAST. Aplicacions del BLAST. Cerca d'homòlegs remots per PSI-Blast. Eines per l'automatització de cerques per similitud.

Tema 7. Alineament múltiple. Alineament múltiple de seqüències. Alineaments progressius. Matrius de distància. Paquet CLUSTAL. Aplicacions de l'alineament múltiple. Eines per l'automatització d'alineaments múltiples. Matrius de posició (*pes*) i LOGOS.

Tema 8. Reconstrucció filogenètica molecular. La filogènia molecular. Mètodes d'inferència filogenètica. Eines per la reconstrucció filogenètica. Exemples de reconstrucció filogenètica.

Tema 9. Genòmica comparativa i funcional. Diferents nivells de comparació. Blocs sintènics. Hotspots de reorganització. Visualització comparativa mitjançant navegadors genòmics.

Tema 10. Assemblatge de genomes. Algoritmes de mapeig contra genomes de referència. Formats i eines de visualització d'assemblatges.

Tema 11. Variabilitat genètica. Descripció de la variabilitat genètica a partir d'arxius VCF. Càlcul de diferents mesures de variabilitat i proves de neutralitat. Bases de dades de SNPs i fenotips.

Metodologia

La metodologia docent inclou dos tipus d'activitats diferenciades: classes de teoria i classes pràctiques d'aula d'informàtica. L'aprenentatge també comptarà amb una tutorització individual i/o col·lectiva de l'alumnat que servirà de recolzament per a resoldre qüestions més concretes o que ho requereixin per la seva complexitat o dificultat.

Classes teòriques

Classes per transmetre els conceptes bàsics i la informació necessària per desenvolupar un aprenentatge autònom. Foment de la participació activa de l'alumnat.

Classes de Pràctiques d'aula d'informàtica

Aquestes pràctiques s'organitzaran a partir de problemes plantejats pel professorat que caldrà resoldre usant les diferents eines i anàlisis bioinformàtiques.

Seminaris i/o conferències

A les últimes sessions del curs hi haurà seminaris i/o conferències d'investigadors en el camp de la bioinformàtica

Tutories

Individuals o en grups petits per a la resolució de dubtes relacionats amb l'assignatura. Aquest tipus d'activitat es podrà realitzarà per petició de l'alumnat.

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			
Classes de teoria	25	1	4, 6, 8, 9, 11, 12, 13, 14, 15
Pràctiques aula informàtica	27	1,08	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Seminaris	3	0,12	4, 6, 8, 9, 11, 12, 13, 14, 15
Tipus: Supervisades			
Tutoria	4	0,16	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Tipus: Autònomes			
Elaboració de treballs	25	1	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Estudi	24	0,96	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Lectura de textos	2	0,08	4, 5, 9, 10, 11, 12
Resolució de problemes computacionals	40	1,6	4, 5, 6, 8, 9, 11, 12, 13, 14, 15

Avaluació

El sistema d'avaluació s'organitza en cinc activitats principals i hi haurà, a més, un examen de recuperació i una activitat opcional de millora de nota final. Els detalls de les activitats són:

Activitats d'avaluació principal

- Avaluacions parcials (1 i 2). Pes global 60%

Les avaluacions parcials són proves combinades que poden constar de preguntes de tipus test, de respostes escrites i de resolució de problemes. En aquestes proves s'avaluaran tots els continguts impartits a l'assignatura en forma de prova teòrico-pràctica.

Els pes de cadascuna de les avaluacions serà del 30%.

Per superar aquestes avaluacions cal assolir una nota mínima de 4,0 en cadascuna d'elles.

- Problemes computacionals (Rosalind). Pes global 25%

Al llarg de tot el curs el professorat plantejarà problemes computacionals o qüestions relatives a la matèria impartida que l'alumnat haurà de resoldre en forma d'avaluacions.

El caràcter continuat d'aquesta avaluació fa que no es pugui avaluar l'assignatura a no ser que hi hagi un participació mínima en un 50% de les proves proposades.

- ADAPTA'T-Bioinformàtica (treball integrador de tota l'assignatura). Pes global 15%

Aquesta activitat està vinculada al projecte d'innovació docent ADAPTA'T i dóna l'opció a participar en una experiència d'aprenentatge i servei en el segon semestre, reconeguda amb 2 crèdits de lliure elecció (<https://adaptat.omicsuab.org>).

El treball consistirà en la resolució d'un problema que contindrà qüestions relatives als diferents blocs temàtics presentats durant les classes de teoria i de pràctiques.

Aquests treballs es realitzaran en grups de 3-5 alumnes i són de caràcter autònom. El professorat tutoritzarà els treballs i resoldrà dubtes i qüestions de plantejament.

L'alumnat lliurarà els treballs seguint les directrius de continguts, presentació i terminis fixats.

Aquesta activitat és de caràcter obligatori i la no participació comportarà la no avaluació de l'assignatura.

- Pràctiques aula d'informàtica. Poden restar fins a 1 punt de la nota final

Aquesta activitat és de caràcter obligatori i la falta no justificada o bé el no aprofitament de l'activitat poden restar fins a 1 punt de la nota final de l'assignatura.

El caràcter continuat d'aquesta avaluació fa que no es pugui avaluar l'assignatura a no ser que hi hagi un participació mínima en un 50% de les sessions proposades.

- Problemes opcionals o de creació pròpia de l'alumnat, i participació activa a classe/fòrum. Poden sumar fins a 1 punt sobre la nota final

Al llarg del curs el professorat proposarà problemes computacionals complexes que seran opcionals. La seva resolució podrà sumar fins a 1 punt sobre la nota final de l'assignatura.

Examen de recuperació

Per participar a la recuperació, l'alumnat ha d'haver estat prèviament avaluat en un conjunt d'activitats el pes de les quals equivalgui a un mínim de dues terceres parts de la qualificació total de l'assignatura.

Es podran recuperar les avaluacions 1 i/o 2 individualment i la nota, si és ≥ 4 , farà promig amb les aprovades en els parcials. En cas que en la/les proves recuperades no s'arribi a un mínim de 4,0, en no poder fer promig, no s'aprova l'assignatura.

Les avaluacions continuades, treball integrador i assistència a les pràctiques a l'aula d'informàtica, pel seu caràcter continuat, no són recuperables.

Millora de la qualificació final

L'alumnat que, havent superat les avaluacions 1 i 2, vulgui millorar la seva qualificació final, podrà optar a una prova final. Aquesta prova inclourà la totalitat de la matèria. No és possible millorar la nota mitjançant treballs o altres tipus d'activitats.

El grau de dificultat d'aquesta prova es correspondrà amb l'objectiu de la mateixa i, per tant, podrà ser superior al de les avaluacions parcials.

L'alumnat que es presenta a aquesta prova renuncia a les qualificacions prèvies i per tant, la nota d'aquesta prova de millora serà la que prevaldrà en la qualificació final encara que sigui inferior a l'obtinguda per parcials.

En cap cas l'alumnat que es presenti a aquesta prova de millora de nota optarà a la qualificació de Matrícula d'Honor.

Fórmula de ponderació de la nota final

Nota final = [(Avaluació 1 x 0,3) + (Avaluació 2 x 0,3) + (Avaluació continuada x 0,25) + (Treball integrador x 0,15)]

Aprovat

L'assignatura es considera aprovada si la nota final és ≥ 5.0

No avaluable

L'alumnat obtindrà la qualificació de "No Avaluable" quan les activitats d'avaluació realitzades tinguin una ponderació inferior al 67% en la qualificació final.

Avaluació única

L'avaluació única consisteix en una única prova de síntesi en la que s'avaluaran els continguts de tot el programa de teoria i pràctiques de l'assignatura en forma de prova teòrico-pràctica. La prova constarà de preguntes de tipus test, de respostes escrites i de resolució de problemes. La nota obtinguda en aquesta prova de síntesi suposarà el 60% de la nota final de l'assignatura. Per superar aquesta prova de síntesi cal assolir una nota mínima de 4,0.

L'avaluació dels Problemes computacionals (Rosalind) i ADAPTA'T-Bioinformàtica (treball integrador de tota l'assignatura) seguiran el mateix procés de l'avaluació continuada. La nota obtinguda suposarà el 25% i el 15% de la nota final de l'assignatura, respectivament. L'alumnat que s'aculli a l'avaluació única, podrà lliurar aquestes dues evidències juntes el mateix dia que el fixat per a la prova de síntesi.

La prova d'avaluació única es farà coincidint amb la mateixa data fixada en calendari per a la darrera prova d'avaluació parcial i s'aplicarà el mateix sistema de recuperació que per a l'avaluació convencional.

L'alumnat que s'aculli a l'avaluació única ha de fer les Pràctiques aula d'informàtica (PLAB) en sessions presencials i en les mateixes condicions que en l'avaluació convencional.

Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Avaluació continuada	25	0	0	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Avaluació parcial 1 (continguts teòric-pràctics)	30	0	0	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Avaluació parcial 2 (continguts teòric-pràctics)	30	0	0	4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15
Problemes opcionals o de creació	Poden sumar fins a 1 punt	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10,

pròpia de l'estudiant	sobre la nota final			11, 12, 13, 14, 15
Pràctiques aula d'informàtica	Poden restar fins a 1 punt de la nota final	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Treball integrador	15	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15

Bibliografia

Llibres

- Attwood, TK. i Parry-Smith, J. 1999. Introduction to Biocomputing. Longman, UK. (*Introducción a la Bioinformática. Attwood y Parry-Smith. 2002. Pearson Educación, S.A.*)
- Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC. (read online for free: <https://www.bioinformaticsalgorithms.org/>)
- Cristianini, N. Y M. W. Hahn. 2007. Introduction to Computational Genomics. A case studies approach. Cambridge Univ. Press. (read online: <https://eemblers.files.wordpress.com/2012/02/introduction-to-computational-genomics-a-case-studies-1>)
- Pevzer, P & R Shamir. 2011. Bioinformatics for biologists. Cambridge Univ. Press. (read online: <http://docshare04.docshare.tips/files/18157/181579469.pdf>)
- Samuelsson, T. 2012. Genomics and Bioinformatics. Cambridge Univ. Press.
- Scherer, S. 2008. A short guide to the Human Genome. Cold Spring Harbor Laboratory Press.
- Xiong, J. 2006. Essential bioinformatics. Cambridge Univ. Press.

MOOCS i recursos online:

- Biology meets Programming: Bioinformatics for Beginners (course by Compeau&Pevzner): <https://www.coursera.org/learn/bioinformatics> (fàcil! per a tothom!)
- Guions de pràctiques de l'assignatura: <https://stepik.org/course/1171/syllabus>
- Journey to the Frontier of Computational Biology (Specialization by Compeau&Pevzner): <https://www.coursera.org/specializations/computational-biology> (difícil! pels que volen un reptel!)
- Problemes computacionals a Rosalind: <http://rosalind.info/problems/locations/>
- Web associada al llibre "Compeau, P i Pevzner, P. 2018. Bioinformatics Algorithms. An active learning approach. 3rd Edition. Active Learning Publishers LLC": <https://www.bioinformaticsalgorithms.org/>

Programari

Programari bàsic

- Problemes computacionals a Rosalind: <http://rosalind.info/problems/locations>
- Llenguatge de programació Python (i BioPython)
- Anaconda
- Jupyter notebook

Bases de dades

- NCBI-Entrez Gene <http://www.ncbi.nlm.nih.gov/gene>
- NCBI-RefSeq <http://www.ncbi.nlm.nih.gov/refseq/>
- NCBI-Genome <https://www.ncbi.nlm.nih.gov/genome/>
- NCBI-Genbank <https://www.ncbi.nlm.nih.gov/genbank/>
- OMIM <https://www.omim.org/>
- GeneCards <https://www.genecards.org/>
- Geneid <https://genome.crg.es/geneid.html>
- Genscan <http://hollywood.mit.edu/GENSCAN.html>

- Pubmed <http://www.ncbi.nlm.nih.gov/pubmed>
- dbSNP <https://www.ncbi.nlm.nih.gov/snp/>
- ENSEMBL <http://www.ensembl.org/>
- PopHuman <https://pophuman.uab.cat/>
- PopHumanScan <https://pophumanscan.uab.cat/>
- EVA <https://www.ebi.ac.uk/eva/?Home>
- GWAS Catalogue <https://www.ebi.ac.uk/gwas/docs/about>
- ClinVar <https://www.ncbi.nlm.nih.gov/clinvar/intro/>
- DisGeNET <https://www.disgenet.org/help>
- Open Targets Platform <https://platform.opentargets.org/>
- RegulomeDB <https://regulomedb.org/>
- Human Genome Dating <https://human.genome.dating/>

Eines d'anàlisi

- EMBOSS <http://www.bioinformatics.nl/emboss-explorer/> , <http://www.ebi.ac.uk/Tools/emboss/>
- Galaxy <https://usegalaxy.org/> (i moltes eines integrades en aquesta plataforma)
- BLAST <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- Transeq https://www.ebi.ac.uk/Tools/st/emboss_transeq/
- PSA <https://www.ebi.ac.uk/Tools/psa/>
- MSA <https://www.ebi.ac.uk/Tools/msa/>
- Clustal Omega <https://www.ebi.ac.uk/Tools/msa/clustalo/>
- Clustal W <http://www.clustal.org/download/current/clustalw-2.1-win.msi>
- MEGA <http://www.megasoftware.net/>