Universitat Autònoma de Barcelona

2021/2022

Bioinformatics

Code: 100894
ECTS Credits: 3

Degree	Type	Year	Semester
2500252 Biochemistry	OB	3	2

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Margarida Julia Sape
Email: Margarita.Julia@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: No
Some groups entirely in Spanish: No

Teachers

Margarida Julia Sape

Prerequisites

There are no pre-requisites for this subject.

It is recommended that the student refreshes the core concepts from "Biologia Molecular" and "Química i Enginyeria de Proteïnes" (2nd year).

Objectives and Contextualisation

This course provides an introduction to bioinformatics.

Upon completion of the course, the student should be able:

-To understand the relevance of the availability of public and annotated databases in the deveopment of sequence-based predictive tools.

-To master the most important web-based and/or graphical user interface tools for the study of sequences.

-To obtain, align, visualise and compare sequences.

-To infer phylogenetic relationships among sequences.

-To be familiar with the different predictive tools for sequences, and to be able to choose the best suited to each different experimental question.

-To be able to design simple bioinformatics experiments to answer biochemical questions.

Competences

- Collaborate with other work colleagues.
- Design experiments and understand the limitations of experimental approaches.
- Interpret experimental results and identify consistent and inconsistent elements.
- Make an oral, written and visual presentation of ones work to a professional or non-professional audience in English and understand the language and proposals of other specialists.
- Manage bibliographies and interpret the information in the main biological databases, and also know how to use basic ICT tools.
- Manage information and the organisation and planning of work.
- Read specialised texts both in English and ones own language.
- Take responsibility for one's own learning after receiving general instructions.
- Think in an integrated manner and approach problems from different perspectives.
- Use ICT for communication, information searching, data processing and calculations.
- Use the basics of mathematics, physics and chemistry that are required to understand, develop and evaluate the chemical procedures of living matter.

Learning Outcomes

1. "interpret use information existing databases biological, patents;, market, etc."
2. Collaborate with other work colleagues.
3. Design experiments and understand the limitations of experimental approaches.
4. Interpret experimental results and identify consistent and inconsistent elements.
5. Make an oral, written and visual presentation of ones work to a professional or non-professional audience in English and understand the language and proposals of other specialists.
6. Manage information and the organisation and planning of work.
7. Model and quantitatively represent a biological system or process.
8. Read specialised texts both in English and ones own language.
9. Take responsibility for one's own learning after receiving general instructions.
10. Think in an integrated manner and approach problems from different perspectives.
11. Use ICT for communication, information searching, data processing and calculations.

Content

- Bioinformatics. Biomedical databases, public, centralised repositories, formats, controlled vocabularies and data standardisation for exchange and reanalysis.
- Alignment of two sequences. PAM and BLOSUM matrices. Alignment algorithms.
- "Basic Local Alignment Search Tool" (BLAST). BLAST search algorithm. Parameters and basic BLAST types. Evaluation of BLAST results.
- PSI-BLAST and other advanced types of BLAST searches. The "position-specific scoring matrix" (PSSM).
- Multiple sequence alignment (MSA). MSA strategies: exact, progressive, iterative, consistency-based or structure-based.
- MSA databases: Pfam and "Conserved domain database".
- Phylogenies. Phylogenetic tree types and components. Steps and methods to build and evaluate a phylogenetic tree.
- Domains. Protein modularity and development of search and/or prediction tools. Bioinformatic tools for predicting the physical properties, location and function of proteins.
- Principles of protein structural prediction. The Chou and Fasman algorithm. Homology-based, fold-recognition or "ab-initio" strategies. Structural visualisation tools. Protein databases (Unoprot, PDB), families, hierarchical categorisation.

Methodology

Classroom sessions. These will cover the teoretical fundamentals of the course, and will be evaluated in the exam.

Autonomous learning- MOODLE activities. During the 8 weeks of the course, the instructor will propose several activities, to be performed autonomously. Students will deliver the outcome of these activities through the MOODLE platform. The activities proposed will range from visualisation of tutorial videos, execution of procedures or questionnaires, depending on the subject matter. In general terms, activities will have a close relationship with the problems and theoretical aspects of the computer lab and classroom sessions. It might be necessary that students perform a specific activity on MOODLE previously to a particular computer lab session, in order to fully benefit from it.

Autonomous learning- Case study. The instructor will propose a case study to be solved in small groups (2-3 people) or individually. The exercise will be integrative and will not be necessarily the same for all the students. The case study will be autonomously solved by each workgroup. Workgroups will have several weeks to carry out the case study. Results will be delivered through a brief oral presentation (5-10 min), which will be also uploaded into the MOODLE platform.

Autonomous learning: Study.

Computer lab sessions. These will be focused into the practical aspects of the course, and it is expected that students will become proficient in data search, procedural aspects and in the analysis of the information obtained by the predictors and databases covered throughout the lessons. The approach will be dynamic, and problems will be solved during the development of each session. Aspects covered in computer lab sessions will be evaluated in the exam.

Tutorial sessions. Individual or small group sessions, focused in doubts related to the course. Upon demand from students. Date/hourto be agreed with the instructor. Particularly recommended when students will be working in the case study.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Classroom sessions	10	0.4	3, 8, 7, 1, 10
Computer lab sessions	16	0.64	11, 2, 3, 6, 4, 7, 1, 10, 5, 9
Type: Supervised			
Tutorial sessions	5	0.2	3, 4, 8, 7, 1, 10
Type: Autonomous			
Case study	15	0.6	11, 2, 3, 6, 8, 7, 1, 10, 9
MOODLE activities	9	0.36	11, 6, 8, 7, 1, 9
Study	10	0.4	11, 3, 6, 4, 7, 1, 10, 9

Assessment

There will be four different evaluation activities:

1) Exam. The exam will evaluate both theoretical and applied concepts covered during the sessions and can include: short questions, questionnaires or data analysis tasks or results proposed by the instructor. The exam

will contribute a 50% of the final mark, and will be solved individually by each student. In order to pass the exam, the mark must be equal or higher than 5/10, and it will be possible to retake it at the programmed date at the end of the semester.

2) Delivery of tasks through MOODLE. All activities will weigh in total a 20% of the final mark. It will not be possible to retake any of them if they are delivered after the deadline or the student fails in any of them.

3) Case study. This part will contribute a 30% of the final mark, which in turn will be split into:

- 25% for the oral presentation and the power point, evaluated by the instructor. It can not be retaken if failed or delivered after the deadline.

- 5% for the peer assessment by fellow students. It can not be retaken if failed or delivered after the deadline.

Each student must accomplish both the following criteria, in order to pass the course:

a) Pass the exam with a mark that is equal or higher than 5/10 and,

b) Obtain a mark that is equal to or higher than 5/10, after doing the weighted average of all four evaluation activities (exam, MOODLE, case study evaluated by instructor and by peers).

Retaking the exam and improving your mark

Those students who fail the exam can retake it on the planneddate at the end of the semester. The text will have a similar format than the first exam. Those who wish to improve their marks can sit again for it, but specifically withdraw their former mark. In that case, the difficulty of the exam could be higher than for those students who failed. Students wishing to sit again for the exam must inform the instructor at least 48 h before the date, in order to plan the logistics (number of rooms needed, etc). Both exams will be revised on a date and place announced by the instructor, between 1 and 7 working days from the publication of the marks. Those students that are not able to attend the exam(s) due to a justified cause and provide the corresponding documentation to the Degree Coordinator, will have the right to sit for an exam on another date. The test might combine the resolution of problems with an oral exam. The Degree Coordinator will see for the practical implementation of that with the instructor. Any other aspect that is not specifically covered in this guide, will follow the general regulations for evaluation of the Bioscience Faculty. In order to retake the exam, students must have previously been evaluated in a set of activities that amount for two thirds of the total weight of the subject. Therefore, when evaluated activities are less than 67% of the total weight, the student will receive the "No Avaluable" mark.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Case study	25%	2	0.08	11, 2, 3, 6, 4, 8, 7, 1, 10, 5
Exam	50%	4	0.16	11, 3, 6, 4, 8, 7, 1, 10, 9
MOODLE activities	20%	2	0.08	11, 2, 3, 6, 4, 7, 1, 10, 9
Peer assessment (case study)	5%	2	0.08	2, 4

Bibliography

- Pevsner, Jonathan. 2015. Bioinformatics and functional genomics, 3rd edition. Wiley-Blackwel. ISBN: 978-1-118-58178-0.
- Lesk, Arthur. 2014. Introduction to Bioinformatics 4th edition. Oxford University Press. ISBN: 9780199651566.
- Pazos, Florencio; Chagoyen, Mónica. 2015. Practical protein bioinformatics. Springer international publishing. ISBN: 978-3-319-12726-2

- Web resources suggested by the instructor

Software

Programs:

Jalview: https://www.jalview.org/

MEGA X: https://www.megasoftware.net/

Notepad++: https://notepad-plus-plus.org/downloads/ç

Icn3d: https://www.ncbi.nlm.nih.gov/Structure/icn3d/icn3d-3.2.0.zip

PyMol: https://pymol.org/2/

We pages and Webservers:

https://www.ncbi.nlm.nih.gov/

https://services.healthtech.dtu.dk/

https://www.expasy.org/

https://bio.tools/

https://www.ebi.ac.uk/services