

Bioinformatics

Code: 100780
ECTS Credits: 6

Degree	Type	Year	Semester
2500250 Biology	OB	3	2

Contact

Name: Daniel Yero Corona
Email: Daniel.Yero@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

Josep Antoni Pérez Pons
Rosa Maria Tarrío Fernandez
Raquel Egea Sánchez

Prerequisites

In order to take this subject, it is recommendable that the students have previously acquired enough solid knowledge on subjects like Genetics, Molecular Genetics, Microbiology, Structure and Functions of Biomolecules and Further Cell Biology.

We recommend a basic understanding of English, although it is not mandatory.

Objectives and Contextualisation

The treatment and computer analysis of molecular data has acquired a fundamental role in the modern Biology and the topics that will be taught in this subject are a basic introductory vision of bioinformatics. The main objectives are:

- To provide the basic bioinformatics knowledge that allows the use of search tools to interrogate the main public databases in Life Sciences and the different approaches for the computational analysis of nucleic acid and protein sequences.
- To give a perspective of the potential of this discipline in the field of research as well as in the professional field.

Skills

- Apply statistical and computer resources to the interpretation of data.
- Be able to analyse and synthesise
- Be able to organise and plan.
- Develop critical thinking and reasoning and communicate ideas effectively, both in the mother tongue and in other languages.

- Develop independent learning strategies.
- Obtain information, design experiments and interpret biological results.

Learning outcomes

1. Apply statistical and computer resources to the interpretation of data.
2. Be able to analyse and synthesise.
3. Be able to organise and plan.
4. Develop critical thinking and reasoning and communicate ideas effectively, both in the mother tongue and in other languages.
5. Develop independent learning strategies.
6. Use and interpret data sources and understand the fundamental principles of bioinformatic analysis in order to establish the corresponding relations between structure, function and evolution.

Content

Topic 1. Databases in Health and Life Sciences. Introduction to bioinformatics and molecular databases. Database Search Strategies. Integrated information retrieval and data submission to nucleotide databases. Sequence formats. NCBI databases. Tools for DNA sequence analysis, visualization and edition.

Topic 2. Sequence Alignments. Pairwise sequence alignment tools. Dot-Plot. Local and global alignments. Sequence alignment and dynamic programming. Substitution Matrices: identity and similarity. Scores, gaps and gap penalties. Multiple sequence alignment. Progressive method of multiple alignments. Applications of multiple sequence alignment.

Topic 3. Sequence Similarity Search. Heuristic algorithms. Bioinformatics tools for sequence similarity searching in sequence databases: BLAST and FASTA strategies. Types of BLAST searches and their applications. DNA or genome annotation by homology based search tools.

Topic 4. Introduction to programming for task automation in bioinformatics. The standard bioinformatic analysis. Scripts. Perl. Variables. Modules. Bioperl.

Topic 5. Gene and protein domain annotation. *Ab Initio* gene prediction. Protein domain annotations with InterProScan. Functional annotations using the Gene Ontology.

Topic 6. Molecular Phylogenetic Reconstruction. Biological evolution. Molecular phylogenetic. Methods of phylogenetic inference (UPGMA, Neighbor-Joining). Phylogenetic reconstruction: Examples.

Topic 7. Genetic Variation and Natural Selection. Types of genetic variation. Neutral theory of molecular evolution. Test of the neutral model of evolution (K_a/K_s). Examples of natural selection.

Topic 8. Comparative Genomics. Phylogenetic footprinting/shadowing. Identifying orthologs and paralogs. Synteny.

Topic 9. Proteins: Sequence Analysis. The sequence-structure-function relationship. Primary sequence database of protein. Sequence-based analysis of proteins and prediction of sub-cellular localization.

Topic 10. Proteins: Functional Analysis. Function prediction. Remote homology detection. Tools and resources for identifying protein families, domains and motifs. Secondary and integrated databases.

Topic 11. Proteins: Structural Analysis. The protein data bank (PDB). Searching for structural homologues. Prediction of structural features. Modeling of 3D Structures. Structural Classification of Proteins.

Methodology

The teaching methodology includes two types of differentiated activities: Lectures and practical sessions in the computer classroom. The learning will also have individual and/or collective mentoring of students that will serve as support to solve more specific problems or those requiring it for its complexity or difficulty.

Lectures

Lectures will address the basic ideas of the different topics and will also increase student motivation to participate actively. Lectures will motivate students to expand and confront autonomously the acquired knowledge as a personal work.

Computer practices

These practices are organized based on problems proposed by teachers that should be solved using different bioinformatic tools and analysis. These activities are of **obligatory attendance**.

Mentoring

Individual or small group tutorials to solve questions related to the subject. This type of activity will be carried out at the request of the students.

Activities

Title	Hours	ECTS	Learning outcomes
Type: Directed			
Computer practices	27	1.08	1, 4, 2, 6
Lectures	25	1	1, 4, 2, 6
Type: Supervised			
Mentoring	3	0.12	4, 2
Type: Autonomous			
Individual reading	10	0.4	1, 5, 4, 2, 6
Individual study	40	1.6	1, 5, 4, 2, 6
Literature search	4	0.16	1, 5, 4, 2, 6
Solving problems in collaborative groups	30	1.2	1, 5, 4, 2, 6

Evaluation

The evaluation schedule is organized into **four main assessment activities** and there will also be a **reassessment test** and an **optional activity** to get a **higher mark**. The success in meeting the course learning objectives will be evaluated as follows:

Main assessment activities

Midterm exams (60%).

- **Midterm exam 1.** Approximately **40%**.
- **Midterm exam 2.** Approximately **20%**.

Midterm exams are combined tests that can count on theoretical and/or practical questions, combining multiple-choice questions, written answer questions and/or problem solving.

The weights of each one of the exam will depend on the distribution of the sessions in the academic calendar. As a reference, the first evaluation will weigh approximately **40%** and the second one **20%**.

None of the assessment activities will account for more than 50% of the final mark.

Each of the two exams (1 and 2) require a **minimum grade of 4 points** (out of 10) in order to pass the course.

Continuous assessments (40%).

Throughout the course the teachers will pose problems or questions related to the subject taught (or with new content not necessarily introduced by the teachers) that the students will have to solve in the form of tests or sporadic deliveries. There will be two different types of assessments: solving-problem in groups (integrative problem) with periodical deliveries, and continuous assessment of contents through individual questionnaires at the end of each unit.

- **Problem-solving in groups (20%).**

This integrative problem will consist in the resolution of a problem that will include questions related to the different thematic blocks presented during the theoretical and practical classes.

This integrative problem will be solved autonomously in groups of 3-4 students. The teaching staff will supervise the work and solve doubts and general questions about the strategies to solve problem.

The students will periodically deliver partial results on this problem to the teachers following established guidelines on content, presentation format and deadlines.

This activity requires a **minimum grade of 4 points** (out of 10) in order to pass the course.

- **End-of-unit tests (20%).**

Tests of combined multiple-choice, numerical and/or short-answer questions to recapitulate worked contents in each unit. A Moodle platform will be used for questions with feedback. The non-execution of any of the assessment activities is a zero in that activity.

This activity requires a **minimum grade of 4 points** (out of 10) in order to pass the course.

Reassessment attempt

To be eligible for the retake process, the student should have been previously evaluated in a set of activities equaling **at least two thirds of the final score** of the course.

Reassessment of each midterm exam (1 and 2) can be done separately. The grade obtained at the reassessment exam will substitute the failed regular exam during the trimester and will be used to calculate the final grade according to the percentages reported in this teaching plan. **Getting a grade below 4 points in any of the reassessment exams results in failing the course.**

In case of presenting to the two reassessment exams, grade will be calculated taking all questions as a whole.

The qualification obtained for the End-of-unit tests can be retaken provided that the number of activities carried out is greater than or equal to 50% of the programmed ones. Note however that the grade obtained for the integrative problem **cannot be retaken**.

Getting a higher mark

In order to get a higher mark, those students who have passed the midterm exams 1 and 2 may opt for a final test. This test will include the **whole subject**. It is not possible to get a higher grade through a written work or other types of activities, or to get a higher grade for the continuous assessment activities.

Exam difficulty level will correspond to the objectives of the subject and, therefore, may be **higher than the midterm exams**.

The student who takes this exam **renounces the previous notes** and therefore, the grade of this test will be the one that will prevail in the final grade even if it is lower than the ones obtained in the midterm exams. The new grade obtained cannot be used to obtain "honors".

Calculation of Final Grade

Final grade = [(Midterm exam 1 x 0.40) + (Midterm exam 2 x 0.20) + (Integrative problem x 0.20) + (End-of-unit tests x 0.20)]

Passing the course

Overall, in order to successfully complete this course, the student must get a **minimum final grade of 5 points** (out of 10) and a minimum grade of 4 points (out of 10) in each of the four main activities.

Not evaluable

The student will be graded as "**Not evaluable**" if the weighting of all conducted evaluation activities is less than 67% of the final score.

Evaluation activities

Title	Weighting	Hours	ECTS	Learning outcomes
Continuous assessment (End-of-unit tests)	20	2	0.08	1, 5, 4, 2, 3, 6
Midterm Exam 1 (theoretical-practical contents)	40	3	0.12	1, 4, 2, 6
Midterm Exam 2 (theoretical-practical contents)	20	3	0.12	1, 4, 2, 6
Problem-solving in groups	20	3	0.12	1, 5, 4, 2, 3, 6

Bibliography

- 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.*)
- Cristianini, N. Y M. W. Hahn. 2007. Introduction to Computational Genomics. A case studies approach. Cambridge Univ. Press.
- Pevzer, P & R Shamir. 2011. Bioinformatics for biologists. Cambridge Univ. Press.
- Samuelsson, T. 2012. Genomics and Bioinformatics. Cambridge Univ. Press.
- Xiong, J. 2006. Essential bioinformatics. Cambridge Univ. Press.