

## **MSc in Bioinformatics for Health Sciences**

### **BCO. Elements of Biocomputing**

#### **Syllabus Information**

**Academic Course:** 2018/19

**Academic Center:** 804 - Official Postgraduate Programme in Biomedicine

**Study:** 8045 – Bioinformatics for Health Sciences - MSc

**Subject:** 30166 – BCO. Elements of Biocomputing

**Credits:** 5.0

**Course:** 1st

**Teaching languages:** English

**Teachers:** Abel David Gonzalez Perez, Nuria Lopez Bigas

**Teaching Period:** 1<sup>st</sup> term

#### ***Presentation***

Elements of Biocomputing is aimed at introducing the use of information technologies and computational approaches in biomedical research. The course is centered in the identification and understanding of the problems that can be tackled by computational approaches in a biomedical research project. Some of the basic approaches in computational biology will be introduced using real case studies as examples, such as research problems on health related topics. The intended training includes the acquisition of concepts and skills to find data and information over the public biomedical databases, the execution of basic analysis on high-throughput biomedical data, the functional annotation of mutations, and basic concepts in system biology (e.g. mapping of genes and proteins to interaction networks and network analysis).

The course is aimed at students interested in learning how to use basic bioinformatic software as users, but is not aimed at learning how to develop such tools. The course is adapted to both students with biological knowledge and those without (e.g. Engineers, Mathematicians, Physicist etc.).

#### ***Associated skills***

##### **General:**

##### **Instrumental:**

1. Proficient reading/writing/listening scientific English.
2. Knowledge of office software to do quality scientific presentations and reports.

##### **Interpersonal:**

1. Group work.
2. Ability to solve by yourself a given problem.

### Systemic:

1. Analysis and synthesis abilities.
2. Ability to search and manage information from different sources

### **Specific competences:**

1. to understand the kind of problems that can be tackled using computational approaches in biomedical research.
2. to understand the different kinds of data describing the molecules and processes in the life sciences that are assembled in different databanks and archives.
3. to appreciate the essential characteristics of a database: its coverage, organization and access routes to retrieve the information it contains.
4. to learn how to find biological information using the wide variety of online resources in biology.
5. to learn how to search for specific information about sequences and biological processes related to diseases.
6. to understand basic concepts of text mining (information retrieval, information extraction, terminologies and ontologies) and to use text mining tools.
7. to understand the concept of sequence alignment, and the basis of scoring schemes for string alignments, including substitution matrices and gap penalties.
8. to appreciate the differences between global alignments and local alignments, and the use of approximate methods for fast searching on sequence databases.
9. to learn how to use online resources for sequence similarity searches (BLAST).
10. to understand the significance of Z-scores and to learn how to interpret the P-values and E-values returned by database searches.
11. to learn how to assess the functional effect of variants in the genome.
12. to learn concepts of cancer genomics and how to find and interpret cancer genomics information.
13. to learn basic concepts of systems biology and how can be applied for the study of biomedical problems.
14. to learn basic concepts of network biology, to understand networks as a framework for data integration, and to appreciate the difference between static and dynamic analysis of networks.
15. to learn how to build, manipulate and visualize biological networks, and how to integrate additional information on the networks.
16. to become familiar with the strategies to retrieve biological networks from public resources.
17. to understand the kind of problems that high-throughput experimentation addresses, and the data analysis strategies required to deal with the data.
18. to become familiar with the methods of functional enrichment of gene lists.

## ***Contents***

### **Contents section 1: Finding information on biomedical archives: biological databases and literature databases**

- To understand the kind of problems that can be tackled using computational approaches in biomedical research.
- To understand the general kind of data describing molecules and processes related to life sciences, and how are they stored in databanks and archives.
- To learn how to retrieve information from these databases and archives, and how start analysis on the data retrieved. Introduction to high-throughput genomics technologies.

### **Contents section 2: Relationships between sequences: sequence alignment and similarity**

- To understand the basic concepts of sequence alignment and similarity.
- To learn how to use online resources for sequence similarity searches (BLAST) and multiple sequence alignment.
- To understand and learn how to annotate the functional effect of mutations.

### **Contents section 3: High-throughput data analysis and visualization**

- To learn how high-throughput data is generated in biological experiments.
- To learn how to perform basic analysis on this data to answer biological questions (Differential Expression Analysis and Enrichment Analysis).
- To learn how to visualize complex biological data in the form of Genome Coordinates and Interactive Heat-maps.

### **Contents section 4: Introduction to systems biology in biomedicine**

- To learn the basic concepts of systems biology and how they can be applied to address biomedical research problems.
- To learn basic concepts of network biology as a framework for data integration and analysis (static and dynamic).
- To learn how to represent biological data as networks, to visualize and analyze them, and to integrate additional information on the networks.

## ***Teaching methods***

The course is mainly a practical course, and each topic is organized around a brief lecture to present the basic concepts behind each topic, followed by hands-on exercises using publicly available resources. In addition, the students will conduct throughout the course an exercise around a real case problem in the biomedical area, where they will have to address different aspects of the problem using the bioinformatics approaches learned during the course.

## ***Evaluation***

The evaluation will be divided in two parts, with percentages of the total grade indicated in parenthesis:

1. Student's presentation and written article on the integration of the different practical exercises ("Integrative work") carried out during the course on a use case (45%).
2. Final exam (55%): practical exercise at the end of the course. This exam will be individual.

For 1, the class will be divided in small groups of people to carry out the integrative work and prepare the final presentation covering and integrating the exercises done during the course on the use case. In these group activities, the same grade will be given to each of the members of the group. During the practical exercises, the students will apply the different approaches learned during the course to analyze the real case study proposed. At the end, they will have to present the results obtained in each exercise in the context of this real use case (student's presentation on the Integrative exercise).