**Bioinformatics University Specialist Programme**

In order for a student to be awarded the degree, 30 ECTSa credits must be completed: 22 taught ECTS and a 8 ECTS research/development project including a dissertation. 1 ECTS is equivalent to a student working 30 hours. In this context that means 10 tuition hours (Theory and/or lab practical work) and 20/15 homework hours.

Both Life Sciences oriented and Computer Science oriented students will be accepted, and specific tuition will be provided by dividing the first semester into two paths: (A) Life Sciences (for computer scientists) and (B) Computer Science (for life scientists). In the second semester, one of the optional modules must be chosen: an optional module represents a concrete application of Bioinformatics.

**1st semester (October – February)**

1A.- Introduction to genetics (Compulsory for the life sciences path, 3 ECTS). [Iratxe Zarraonaindia, 30hr]

* The Cell.
* DNA structure and Function.
* Mendelian and Quantitative Genetics.
* Linking genotype and phenotype.
* Genetics of bacteria and viruses.
* Population genetics.
* Genetics and evolution.

1B.- Programming and Data Bases (Compulsory for the computer science path, 3 ECTS).

* Bioinformatics programming in Perl: regular expressions and file parsing, Perl libraries for bioinformatics. [Mark Wilkinson, 10hr]
* Introduction to computational complexity and R language. [Igor Santos, 6hr]
* Object Oriented Programming with Python. [Javier Nieves,7hr ]
* Databases: managing and using relational databases, introduction to NoSQL databases. [Javier Nieves, 7hr]

2.- Advanced computing (Compulsory, both paths, 6 ECTS).

* Introduction to Data mining and Heuristic Optimization: principal techniques and bioinformatics applications. [Iñaki Inza, 20hr]
* UNIX: basic system administration (users, processes, file system), shell and shell scripting, implementation of UNIX pipelines and scheduled tasks. [Mikel Egaña/Jorge Langa, 10hr]
* High Performance Computing: Infrastructures, Algorithms. [A. Mendiburu/JM Alonso, 20hr]
* Web technologies: Web servers, Web APIs and mash-ups, Web services. Life Sciences Semantic Web. [Mark D. Wilkinson/Mikel Egaña, 10hr]

**2nd Semester (February - May)**

3.- Applied bioinformatics I. Variation Analysis (Compulsory, both paths, 5 ECTS).

* Research methods in bioinformatics. [Alberto Labarga, 10hr]
* Whole Genome Sequencing and de novo sequencing. [Iratxe Zarraonaindia/Fernando Rendo, 20hr]
* Introduction and data pre-processing.
* Data pre-processing.
* Re-sequencing (Biomedicine) .
* De novo sequencing (Environmental) .
* Mutation discovery.
* Molecular phylogeny .
* Reduced genome sequencing. [Iratxe Montes/Fernando Rendo, 20hr]
* Introduction and data pre-processing .
* Sequences capture. Exome analysis.
* RAD/GBS sequencing.
* De Novo transcriptome sequencing.
* Alignment based approaches and assembly based approaches.
* Gene structure: Identifying intron-exon boundaries (IEB). Determining complete transcripts. Mutation discovery.

4. Applied bioinformatics II. Expression analysis (Compulsory, both paths, 5 ECTS)

* Introduction to Big Data concerned with the analysis of DNA expression. [Marcos Araúzo, 10hr]
* Transcriptomics. RNA-Seq.
* Sequencing of protein-coding genes. Protein families. [Marcos Araúzo, 5hr]
* Sequencing RNA genes (miRNA, siRNA, rRNA, tRNA, snoRNA, snRNA). [Marcos Araúzo, 5hr]
* Gene expression microarrays: preprocess, differential expression and classification with R-Bioconductor. [I. Inza, 10hr]
* Funtional analysis of genes sets. Gene Ontologies (GO). [Marcos Araúzo, 5hr]
* Epigenomics. [Marcos Araúzo, 5hr]
* Methylomes: DNA methylation .
* Histone modification: ChIP-Seq .
* Proteomics. [Marcos Araúzo, 5hr]
* Global analysis of Big Data. PCA, hierarchical clustering. [Marcos Araúzo, 5hr]

5.- Bioinformatics applied to biomedicine (Optional for both paths, 3 ECTS).

* Human genome variation. [Iratxe Montes, 10hr]
* Variation analysis in humans (SNPs, chromosomal aberrations). [Marcos Araúzo, 5hr]
* Genome Wide Association Studies. (GWAS) in humans [José A. Lozano, 5hr]
* Epigenetic data analysis. [B. Calvo, 5hr]
* Chemoinformatics. Drug discovery. [Marcos Araúzo, 5hr]

6.- Bioinformatics applied to environment, populations, and non-model organisms (Optional for both paths, 3 ECTS).

* Metagenomics. [Iratxe Zarraonaindia, 20hr]
* Metabarcoding. [Aitor Albaina, 10hr]
* Population genomics in non-model organisms (SNPs, ...). [Iratxe Montes, 4hr]
* Comparative genomics. [Mikel Egaña/Jorge Langa, 4hr]
* Ecological informatics. [Peter B Pearman, 2 hr]

**Project (May - September)**

The project is (in the most part) related to the optional module undertaken by the student, and will probably include a placement in a partner institution (company or academia). The project is small, very focused and with a moderate technical difficulty, preferably touching upon content of various compulsory modules. The outcome will be a dissertation and an oral presentation summarizing it.

Two types of projects are available:

Research project: it focuses on answering a scientific question and the candidate will exploit bioinformatics as a means to obtain such answer. The project will be assessed in terms of how different technologies were combined and the merit of the scientific answer. The dissertation will be produced as a scientific paper.

Development project: it focuses on the the technical aspect of bioinformatics, and the candidate will develop a program or system to achieve an specific aim. The project will be assessed technically, thus the performance, overall design, implementation skills, etc. will be taken into account, rather than the scientific result. The result of the project will be a piece of software with enough quality for free distribution and that, potentially, it could be published in a bioinformatics journal. The dissertation will be produced as a technical report.