COURSE DATA

Data Subject	
Code	42590
Name	In silico studies in biomedicine
Cycle	Master's degree
ECTS Credits	6.0
Academic year	2023 - 2024

Study (s)

Degree	Center	Acad. year	Period
2116 - M.U. en Bioinformática 12-V.1	School of Engineering	2	First term

Subject-matter

Degree	Subject-matter	Character
2116 - M.U. en Bioinformática 12-V.1	8 - In silico studies in biomedicine	Obligatory

Coordination

Name	Department
ARNAU LLOMBART, VICENTE	240 - Computer Science

SUMMARY

In modern data analysis processes bioinformatics is essential to acquire the necessary knowledge processing techniques for microarray data, both expression and SNP, array-CGH and ChIP-on-Chip, next-generation sequencing applied to: resequencing , RNA-seq, ChIP-seq, structural variation, genome assembly, etc..
Perform association studies, search for biomarkers, predictors of response or class, discover classes based on omics data, linking omics data among themselves and with phenotypes, and be able to give a functional interpretation of the relationships found.
Managing public genomic databases overlooking the in silico studies.

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

None

OUTCOMES

2116 - M.U. en Bioinformática 12-V.1

- Students should apply acquired knowledge to solve problems in unfamiliar contexts within their field of study, including multidisciplinary scenarios.
- Students should be able to integrate knowledge and address the complexity of making informed judgments based on incomplete or limited information, including reflections on the social and ethical responsibilities associated with the application of their knowledge and judgments.
- Students should communicate conclusions and underlying knowledge clearly and unambiguously to both specialized and non-specialized audiences.
- Students should demonstrate self-directed learning skills for continued academic growth.
- Be able to access the information required (databases, scientific articles, etc.) and to interpret and use it sensibly.
- Students should possess and understand foundational knowledge that enables original thinking and research in the field.
- Be able to access to information tools in other areas of knowledge and use them properly.
- To be able to assess the need to complete the scientific, historical, language, informatics, literature, ethics, social and human background in general, attending conferences, courses or doing complementary activities, self-assessing the contribution of these activities towards a comprehensive development.
- Desarrollar la iniciativa personal y ser capaces de realizar una toma rápida y eficaz de decisiones en su labor profesional y/o investigadora.
- Trabajar en equipo con eficiencia en su labor profesional y/o investigadora y con personas de diferente procedencia.
- Conocer las técnicas para el procesamiento de datos de microarrays, tanto de expresión como de SNPs, array-CGH y ChIP-on-Chip; secuenciación de nueva generación aplicada a resecuenciación, RNA-seq, Chip-seq, variación estructural, ensamblado de genomas y otras.
- Adquirir los conocimientos para realizar estudios in silico de asociación, búsqueda de biomarcadores, predictores de respuesta o clase, descubrir clases basadas en datos ómicos, relacionar datos ómicos entre sí y con fenotipos; y ser capaces de dar una interpretación funcional de las relaciones encontradas.

LEARNING OUTCOMES

Acquire the necessary knowledge processing techniques for microarray data, both expression and SNPs, array-CGH and ChIP-on-Chip, next-generation sequencing applied to: resequencing, RNA-seq, ChIP-seq, structural variation , genome assembly, etc..
Perform association studies, search for biomarkers, predictors of response or class, discover classes based on omics data, linking omics data among themselves and with phenotypes, and be able to give a functional interpretation of the relationships found.
Managing public genomic databases overlooking the in silico studies.

DESCRIPTION OF CONTENTS

1. Bases de dades (GEO, ArrayExpress, SRA): maneig i extracció de dades

Introduction to the main experimental genomic data repositories.

2. Genomic information systems: Biomart

Introduction to the tools that allow access to Biomart, a network of genomic databases spread across multiple repositories located in different parts of the world.

3. Primary processing data: expression microarray normalization, batch effect.

Quality control and population structure of genotyping chip. Quality control, mapping and variant calling in next generation sequencing.
Introduction to basic tools of genomic data preprocessing: R / Bioconductor, Babelomics, FastQC ...

4. Analysis of expression: differential expression tests expression microarrays and RNA-seq. Biomarkers. Molecular signatures.

Introduction to methodology and basic tools for differential expression analysis and gene selection: R / Bioconductor / limma.

5. Binders and gene expression-based predictors. Search classes (clustering and biclustering)

Introduction to supervised classification methods: SVM, KNN, LDA ...
Introduction to unsupervised classification methods: UPGMA, SOTA, K-means

6. Analysis of structural variation: with chips SNPs, tiling arrays and next generation sequencing. Amplifications / deletions, LOH, translocations and inversions.

Introduction to SNPs data preprocessing. Estimated number of copies of each genomic region. Introduction to methods of segmentation.

7. Analysis of association of SNPs and variants in case-control studies and families.

Introduction to different methodologies of association and linkage analysis.
Introduction to PLINK analysis tool.

8. Mutational analysis (resequencing). Search for disease genes.

Introduction to bioinformatics tools and databases available to predict or infer the pathogenicity of genomic variants.

9. Functional analysis: functional enrichment methods, gene-set, network analysis and prioritization methods.

Introduction to different methodologies to combine experimental data with the information available in databases.

10. Data integration and simple modeling of biological

Combined analysis of different genomic measures: expression, copy number variants ...

WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	21,00	100
Laboratory practices	9,00	100
Attendance at events and external activities	10,00	0
Development of group work	10,00	0
Development of individual work	20,00	0
Study and independent work	20,00	0
Readings supplementary material	20,00	0
Preparation of evaluation activities	15,00	0
Preparing lectures	25,00	0
Preparation of practical classes and problem	20,00	0
Resolution of case studies	10,00	0
TOTAL	**180,00**	

TEACHING METHODOLOGY

MD1 - Task training of the teaching-learning environment interaction in the classroom through expository sessions. Previous assignments include preparation (information search, reading texts supplied by teachers), teaching sessions themselves and the later work of deepening.
MD2 - Learning through problem solving and case studies, through which it is acquiring skills on different aspects of materials and subjects.
MD3 - Activities labs. Include preparation, implementation of the monitoring practices and teacher support, online freelance work and reporting practices.
MD4 - Cross-disciplinary skills. Include attendance at courses, conferences or round tables organized by the CEC of the Master and / or conduct of a bibliographic work on issues that contribute to the integral. It produces a report of activities.

EVALUATION

In the two calls:

SE1 Continuous assessment: minimum 5 and maximum 15.
SE2 Activities: minimum 10 and maximum 50.
SE3 Laboratory: minimum 25 and maximum 50.

REFERENCES

Basic

- Bioinformatics and Computational Biology Solutions Using R and Bioconductor
  Robert Gentleman 2005
- http://books.google.es/books?id=e8-CfJ5RIgAC&dq=bioconductor&hl=en&sa=X&ei=QC7oUZzSBcHbPaGqgNgL&redir_esc=y