COURSE DATA

Data Subject	
Code	44696
Name	Bioinformatic techniques
Cycle	Master's degree
ECTS Credits	2.5
Academic year	2022 - 2023

Study (s)

Degree	Center	Acad. year	Period
2224 - M.U. Investigación y Desarrollo en Biotecnología Biomedicina	Faculty of Biological Sciences	1	First term

Subject-matter

Degree	Subject-matter	Character
2224 - M.U. Investigación y Desarrollo en Biotecnología Biomedicina	1 - New technology	Obligatory

Coordination

Name	Department
BARRIO ESPARDUCER, ELADIO	194 - Genetics

SUMMARY

Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data. As an interdisciplinary field of science, bioinformatics combines computer science, statistics, mathematics, and engineering to analyze and interpret biological data: nucleotide and protein sequences, genome organization, structure and variability, as well as the prediction of biomolecule structure and function.

With the advent of the era of genomics, bioinformatics has extended its field of study to the massive analysis of biological data, including those derived from humans, and therefore, has become as a very important support for biomedical and biotechnology research.

The contents of this course have been designed as an advanced extension of the basic knowledge and training acquired by the students during their degrees. The extended knowledge and skills acquired in this course must be defined as bioinformatic techniques, since the aim is not to train bioinformaticians, but biomedicine and biotechnology researchers able to understand and use bioinformatic analysis to the advanced user level.

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

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

Other requirements

There are no enrollment restrictions with other subjects of the curriculum of the master or any prerequisites

OUTCOMES

2224 - M.U. Investigación y Desarrollo en Biotecnología Biomedicina

- 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.
- Students should possess and understand foundational knowledge that enables original thinking and research in the field.
- Be able to integrate new technologies in their professional and/or research work.
- Ser capaces de analizar de forma crítica tanto su trabajo como el de su compañeros.
- Capacidad de seleccionar y gestionar los recursos disponibles (instrumentales y humanos) para optimizar resultados en investigación.
- Ser capaces de realizar una toma rápida y eficaz de decisiones en situaciones complejas de su labor profesional o investigadora, mediante el desarrollo de nuevas e innovadoras metodologías de trabajo adaptadas al ámbito científico/investigador, tecnológico o profesional en el que se desarrolle su actividad.
- Ser capaces de acceder a la información necesaria en el ámbito específico de la materia (bases de datos, artículos científicos, etc.) y tener suficiente criterio para su interpretación y empleo.
- Aplicar el razonamiento crítico y la argumentación desde criterios racionales.
- Capacidad para preparar, redactar y exponer en público informes y proyectos de forma clara y coherente, defenderlos con rigor y tolerancia y responder satisfactoriamente a las críticas que pudieren derivarse de su exposición.

- Ser capaces de trabajar en equipo, sin discriminación entre hombres y mujeres, con eficiencia en su labor profesional o investigadora adquiriendo la capacidad de participar en proyectos de investigación y colaboraciones científicas o tecnológicas.
- Aprendizaje en la redacción de artículos científicos en los campos de la Biomedicina y la Biotecnología.
- Manejar adecuadamente las fuentes de información científica y poseer la habilidad de hacer una valoración crítica de las mismas, integrando la información para aportar conocimientos a grupos de investigación multidisciplinares.
- Utilizar adecuadamente las herramientas informáticas, métodos estadísticos y de simulación de datos, aplicando los programas informáticos y la estadística a los problemas biomédicos y biotecnológicos.
- Dominar el método científico, el planteamiento de protocolos experimentales y la interpretación de resultados en el ámbito biomédico y biotecnológico.
- Ser capaces de aplicar la experiencia investigadora adquirida tanto en la empresa privada como en organismos públicos.
- Saber diseñar estrategias experimentales multidisciplinares en el ámbito de las biociencias moleculares para la resolución de problemas biológicos complejos, especialmente los relacionados con salud humana.
- Adquirir destrezas en el manejo de las metodologías avanzadas empleadas en las biociencias moleculares y en el registro anotado de actividades.

LEARNING OUTCOMES

1. Understanding of the operation of the file system and how the information is managed by the Linux operating system

2. Knowing of the types of data generated by new generation sequencing techniques: sequence formats and files.

3. Knowing the types of genome assemblers and their algorithms.

4. Strategies and procedures of current genome annotation methods.

5. Understanding special statistical problems posed by high performance biological data.

6. Knowing how to apply basic statistical techniques and those adapted to the computational treatment of big experimental or clinical data.

7. Understanding the biological, mathematical and computational fundamentals of the bioinformatic methods used in the inference of evolutionary processes at the genome level.

DESCRIPTION OF CONTENTS

1. Introduction to Bioinformatics

What is Bioinformatics? Multidisciplinary area. Bioinformatics applications. Basic knowledge on file formats and omics data analysis

2. Genome assembling methods

Types of data obtained by ultrasequencing (NGS), sequence formats and files. Sequence quality assessment and data pre-treatment. Genome assembly algorithms: 'de novo' and 'mapping'. Genome assembly quality assessment. Sequence data bases management: SRA. Re-sequencing data analysis: 'variant calling'

3. Genome annotation procedures

Genome annotation files and formats: GFF, EMBL, GenBank. Current annotation strategies: advance methods for gene prediction and annotation transfer. Gene identification algorithms. Genome annotation quality assessment

4. RNAseq gene expression analysis: data management.

Quality assessment of the RNAseq reads. Alignment of reads with a reference genome. Sequence structure and normalization. Gene transcription counting

5. Statistical concepts applied to bioinformatics

- Analysis design and data management
- Inferential Statistics. Confidence intervals and hypothesis testing
- Introduction to R and Bioconductor

6. Classical Statistics for omics data

- t test, ANOVA and non-parametric tests for omics data
- The problem of multiple comparisons. False Discovery Rate (FDR)
- Statistic analysis with R: performance and interpretation

7. Statistical modeling of omics data (I)

- Multivariable analysis concept
- Batch effect correction
- Clinical/Epidemiological data integration
- Lineal models

8. Statistical modeling of omics data (II)

- Negative binomial regression. RNAseq analysis.
- Statistic analysis with R: performance and interpretation

9. Predictive models with omics biomarkers

- Problem description
- Penalized regression methods
- Random Forest
- Penalized regression with R

10. Genome variation analysis

Phylogenomics: genome phylogenetic analysis. Population genomics: population structure analysis from genome variation. Selection detection at the genome level

WORKLOAD

ACTIVITY	Hours	% To be attended
Computer classroom practice	25,00	100
Attendance at events and external activities	1,00	0
Development of individual work	6,00	0
Study and independent work	8,00	0
Readings supplementary material	2,00	0
Preparation of evaluation activities	5,00	0
Preparing lectures	5,00	0
Preparation of practical classes and problem	3,00	0
Resolution of case studies	7,50	0
TOTAL	**62,50**	

TEACHING METHODOLOGY

Classes will be theoretical-practical, so they will be hold in a computer room.

For each topic, there will be:

1) An introduction of theoretical concepts that require prior preparation tasks by the students (by reading texts provided by the teacher).

2) Active learning through problem solving and practical case studies, through which the students will acquire skills on different aspects of the subject.

3) Self learning to solve individual problems and new theoretical and practical exercises.

4) We can consider attending a conference, seminar, panel discussion, on topics related to the subject of the course and organized at the Burjassot-Paterna campus by the Master CCA or departments, centers, research institutes from this campus.

EVALUATION

The evaluation will be based on a continuous assessment of each student, according to the different activities described in the Methodology section: attendance to all class activities, resolution, and presentation of all the works and activities, the degree of participation and involvement of the student in the process of learning. The specific aspects to evaluate are:

1. Objective exams on the contents of the course, either in the classroom or 'online' (based on a series of short questions or multiple-choice tests, regarding both theoretical and practical issues), and exercises and assignments along the course. These tests and exercises will be evaluated by each professor, of the different parts of the course. The total score will be a weighted mean of the scores of the different parts, according to the number of hours of each one. The score of this section will represent 85% of the final course scoring.

2. Assessment of participation in classroom activities, tutorials, and other group activities. Among other things, this section will assess the ability to ask questions, propose answers and lead the group discussion, as well as attendance. The score of this section will represent 15% of the final course scoring.

To pass the course, it is mandatory to get a final score of 5 out of 10.

REFERENCES

Basic

- Masoudi-Nejad A, Narimani Z, Hosseinkhan N. Next Generation Sequencing and Sequence Assembly: Methodologies and Algorithms. Springer Briefs in Systems Biology, 2013
- Soh J, Gordon PMK, Sensen CW. Genome Annotation. Chapman & Hall/CRC, 2012
- Krijnen, WP. Applied Statistics for Bioinformatics using R, 2009
- Lee, JK (Ed.) Statistical Bioinformatics A Guide for Life and Biomedical Science Researchers Wiley-Blackwell, 2010
- Toll, AP (Ed.) Structural Bioinformatics. Ceed Publishing, 2012
- Lemey P, Salemi M, Vandamme AM (Ed.) The phylogenetic handbook. A practical approach to phylogenetic analysis and hypothesis testing. 2nd ed. Cambridge: Cambridge University Press; 2009

- Yang, Z. Computational Molecular Evolution. Oxford: Oxford Univ. Press; 2006

Additional

- Ekblom, R., & Wolf, J. B. W. (2014). A field guide to whole-genome sequencing, assembly and annotation. Evolutionary Applications, 7(9), 10261042. http://doi.org/10.1111/eva.12178
- Baker, M. (2012). De novo genome assembly: what every biologist should know. Nature Methods, 9(4), 333337. http://doi.org/10.1038/nmeth.1935