



# UNIVERSITÀ DEGLI STUDI DI PALERMO

<b>DEPARTMENT</b>	
<b>ACADEMIC YEAR</b>	
<b>ANNO ACCADEMICO EROGAZIONE</b>	
<b>SUBJECT</b>	
<b>CODE</b>	
<b>SCIENTIFIC SECTOR(S)</b>	
<b>HEAD PROFESSOR(S)</b>	CHIODI MARCELLO      Professore Ordinario      Univ. di PALERMO
<b>OTHER PROFESSOR(S)</b>	
<b>CREDITS</b>	
<b>PROPAEDEUTICAL SUBJECTS</b>	
<b>MUTUALIZATION</b>	
<b>YEAR</b>	
<b>TERM (SEMESTER)</b>	
<b>ATTENDANCE</b>	
<b>EVALUATION</b>	
<b>TEACHER OFFICE HOURS</b>	<b>CHIODI MARCELLO</b> Tuesday    15:00    17:00    stanza del docente (edificio 13); eccezionalmente su teams Friday      12:00    13:00    stanza del docente (edificio 13); eccezionalmente su teams

DOCENTE: Prof. MARCELLO CHIODI

<b>PREREQUISITES</b>	Familiarity with computer programming and computational skills
<b>LEARNING OUTCOMES</b>	<p>The first part of the course focuses on the statistical analysis of gene expression. The statistical methodology will be contextualized within the fields of molecular biology, genetics, and transcriptomics. The student will be involved in the analysis of gene expression data using the R software.</p> <p>The second part of the course focuses on technical aspects related to parallel programming and the use of the computer terminal. These aspects are of fundamental importance for managing and analysing large amounts of data (regardless of the nature of the data of interest). These topics will be accompanied by bioinformatic applications such as genome alignment, prediction of molecular interactions, and estimation of cellular fractions in tissues. The student will become familiar with the main statistical methodologies for gene expression analysis. Furthermore, students can independently connect to a computer server and analyze large amounts of data through remote programming and parallel computing. The student will be able to create an executable R program.</p>
<b>ASSESSMENT METHODS</b>	<p>The exam consists of an oral test about the course topics. The student has to present a statistical analysis of genetics data. Before the exam, the student will receive data analysis instructions. The student must show the data analysis results during the oral test.</p> <p>Evaluation criteria: 30 cum laude: excellent knowledge of the topics, excellent use of technical language, excellent ability to autonomously apply the acquired knowledge. 30: excellent knowledge of the topics, excellent use technical language, good ability to autonomously apply the acquired knowledge. 26-29: Good knowledge of the topics, good use of the technical language, moderate ability to autonomously apply the acquired knowledge. 24-25: Moderate knowledge of the main topics, limited ability to autonomously apply the acquired knowledge. 21-23: Knowledge of the course's main topics, but without a thorough understanding. Limited use of the technical language, and poor ability to autonomously apply the acquired knowledge. 18-20: Minimal knowledge of the course's main topics and technical language, very limited ability to autonomously apply the acquired knowledge. Insufficient: The student does not possess an acceptable knowledge of the topics covered during the course and has no application skills.</p>
<b>EDUCATIONAL OBJECTIVES</b>	The course aims to introduce students to the main statistical methods for gene expression analysis and provide basic computer skills for analysing and managing large amounts of data.
<b>TEACHING METHODS</b>	Frontal lessons and computer exercises using R software and the computer terminal
<b>SUGGESTED BIBLIOGRAPHY</b>	materiale fornito in aula

## SYLLABUS

Hrs	Frontal teaching
2	Introduction to computational biology, genetics, and transcriptomics. Overview of molecular biology; central dogma of biology and protein synthesis.
5	Large scale inference in computational biology Introduction to gene expression data
6	Differential expression analysis: Significant analysis of microarray, moderated t-test, Mann-Whitney test, Kruskal-Wallis one-way ANOVA. Identification of genetic signatures. Gene set enrichment analysis. Gene Ontology, Reactome pathway analysis, GSEA software
6	Gene expression harmonization. Batch effect Analysis, PCA and UMAP clustering and unsupervised cluster analysis The bootstrap approach in hypothesis testing. Bootstrap t-test for differential expression analysis. Analisi di sopravvivenza per valutare il potere prognostico delle firme geniche sulle casistiche di tumori
5	Introduction to bioinformatics applications. Introduction to FASTA and FASTQ files. Introduction to sequence alignment problem. Genome alignment using STAR. Introduction to the mechanisms of gene regulation. Prediction of bonds between messenger RNA molecules and microRNAs In silico estimation of cell fractions in tumor tissues. Use of CIBERSORTX
Hrs	Workshops
8	Advanced computer tools for the analysis of large datasets. Introduction to the terminal. Introduction to remote programming.
6	Running software through the terminal. Creating executable R programs
4	Parallel computing in R. Monitoring processes and RAM usage through the terminal.