MEDTEC SHOOL

Molecular and Computational Biology and Medical Genetics (MCBMG)

Year: 2nd
Period: 2nd semester
Credits: 9

Course faculty

Rosanna Asselta (Coordinator) – Medical Genetics
Linda Pattini – Bioinformatics and Computational Science
Letizia Straniero – Molecular Biology

Objectives

The course will merge molecular biology and medical genetics with computational science and bioinformatics to give an up-to-date vision of molecular genetics mechanisms involved in human diseases, with a particular focus on genetic disorders. Besides the theoretical part, large space will be given to practical sessions in which the students will learn how to use genomics data present on the web and how to analyze real DNA and RNA sequencing data.

The "Molecular biology" module will be focused on understanding structure, function, and turnover of macromolecules, the structure and regulation of genes and genomes, the mutability of the genome and the mechanisms of DNA repair. The impact of recombinant DNA technologies on basic and applied biomedical research will also be illustrated.

The "Medical genetics" module will illustrate the types and extent of genetic variation seen in the human genome, and how this variation affects disease susceptibility. The module will particularly emphasize how molecular and population genetics have impacted on our understanding of the mechanisms of human disease, providing us with novel diagnostic and therapeutic strategies.

The "Bioinformatics and Computational science" module will focus on fundamentals of computational methods used to analyze functional genomics data. It will illustrate techniques for comparison of biological sequences by considering specific algorithms, complexity issues and available resources. The module will also introduce computational strategies and challenges in data analysis to study genetic variation and transcription. Aspects of analytics for raw data processing will be illustrated highlighting their relatedness with technology. Fundamentals of data mining techniques applied to gene expression profiling will be presented along with tools for functional annotation. The module will also introduce the R programming environment and the implementation of specific pipelines to manipulate and process data through practical examples of solutions for specific biological problems.

Prerequisites

To be allowed to take the MCBMG exam students must have passed the exam of "Cell Biology, Embryology and Histology".

Contents

1. The mutability of the genome

Causes of mutation (Prof Straniero)

Spontaneous and induced causes of mutation. Agents that induce mutation. Common mutations due to DNA replication.

Learning objectives:

- Describe examples of the main categories of DNA damaging agents
- Describe mechanisms by which DNA can be damaged
- Discuss the relationship between DNA damage and DNA mutation
- Discuss the balance between mutation-inducing mechanisms and DNA repair and its role in evolution

Genetic variation (Prof. Asselta)

Polymorphisms and mutations: classification and functional consequences.

Learning objectives:

- Describe the different types of point mutations
- Discuss differences between germinal and somatic mutations, and between polymorphisms and mutations
- Describe the functional consequences of the different types of mutation at the RNA and protein levels

DNA repair mechanisms (Prof. Straniero)

Main mechanisms of DNA repair and related diseases.

Learning objectives:

- Discuss the cell responses to DNA damage
- Describe the main mechanisms of DNA repair and the specific damages they are apt to
- **Clinical drop**: Inherited diseases due to defects in the DNA repair systems

2. Mendelian genetics: from pedigrees to mutations

Mendelian pedigree patterns and their complications (Prof. Asselta)

How to build up a pedigree. Examples of autosomal dominant, autosomal recessive, X-linked, Y-linked and mitochondrial inheritance. Incomplete penetrance, expressivity,

male lethality, de-novo mutations, mosaicism, phenocopies, complementation, and mitochondrial inheritance.
Learning objectives:

- Being able to collect a genetic anamnesis and to draw a pedigree
- Describe the main patterns of inheritance
- Discuss the main complications of the classical Mendelian pattern of transmission

Genetic mapping of Mendelian traits (Prof. Asselta)
Polymorphisms as a tool for genetic mapping, recombinants and non-recombinants, two-point and multi-point mapping, the concept of LOD score.
Learning objectives:

- Understand how polymorphisms can be used to trace Mendelian traits
- Understand the meaning of LOD score in genetic analysis

Prototypic Mendelian diseases: cystic fibrosis and thalassemia (Prof. Asselta)
The quest for the gene causing cystic fibrosis. The organization of the globin loci. Sickle-cell anemia. Alpha and beta thalassemia.
Learning objectives:

- Describe a typical approach to identify disease genes
- Describe the gene arrangement of the globin loci
- Illustrate the molecular mechanisms underlying sickle cell anemia and thalassemia
- **Clinical drop**: neonatal screening and genetic diagnosis of cystic fibrosis

Dynamic mutations (Prof. Straniero)
The concept of microsatellite instability. Trinucleotide repeat expansion disorders and their classification. Anticipation. Fragile X syndrome, Huntington disease, myotonic dystrophy.
Learning objectives:

- Describe the general features of trinucleotide repeat expansion disorders
- Understand the concept of anticipation
- Illustrate the molecular characteristics of Fragile X syndrome, Huntington disease, and myotonic dystrophy

3. Cancer genetics

Sporadic, Familial and Hereditary cancers (Prof. Asselta)
Epidemiology of cancer & Cancer as genetic disease. Cancer as a hereditary disease. Genetic counselling of cancer.
Learning objectives:

- Describe the basic molecular mechanism leading to cancer

- Describe the genetic landscape of cancers
- Describe the most common inherited cancer syndromes and their molecular mechanisms
- Learn the basic steps that are taken during a genetic counselling for cancer predisposition

4. Recombinant DNA & DNA/RNA sequencing

Essentials of DNA recombinant technology (Prof. Straniero)
What is rDNA? Tools of rDNA technology, making a recombinant DNA molecule, DNA probes and hybridization, main applications of rDNA.
Learning objectives:

- Understand the principles of rDNA technology
- Understand the concept of genomic and cDNA library
- Be able to describe a cloning experiment
- Know what a molecular probe is and what is a hybridization experiment

Polymerase Chain Reaction (PCR) & DNA sequencing (Prof. Straniero)
The discovery of PCR, the principles of a PCR reaction, what can you do with PCR? Sanger sequencing, Next generation sequencing.
Learning objectives:

- Discuss why DNA is amplified during a PCR reaction
- Illustrate the main applications of PCR
- Describe the theory of Sanger sequencing
- Compare Sanger sequencing and next generation sequencing (NGS)
- Describe the concepts of clonal amplification and parallel sequencing
- List the main technologies for NGS
- Discuss the concept of sequencing depth

Sequence alignment (Prof. Pattini)
Importance of biological sequence comparison, definition of similarity and homology. Local or global pairwise alignment of two biomolecular sequences, sequence similarity search, exact and heuristic approaches
Learning objectives:

- Learn how sequence comparison is performed
- Use of available software tools to perform sequence alignment and result interpretation

Introduction to R environment for processing data, visualizing information and extracting knowledge (Prof. Pattini)
Learning objectives:

- Ability to manipulate and visualize data
- Ability to use specific libraries relevant to the topics of the course

Exploring genetic variation (Prof. Pattini)
Applications of Next-generation sequencing (NGS) in biomedicine. NGS of whole genomes and exomes as a powerful tool in biomedical research and clinical diagnostics. Genome informatics for DNA variant characterization.
Learning objectives:

- Compute and visualize genetic variation
- Learn basic computational methods for annotation of genetic variants

5. Regulation of gene expression

Introduction to the multilayer regulation of gene expression (Prof. Straniero)
How cells modulate gene expression. Levels of gene expression regulation: from chemical modification of DNA to post-transcriptional and post-translational regulation. The principles of DNA-protein interaction.
Learning objectives:

- Discuss the principles of transcriptional regulation in prokaryotes and eukaryotes
- Describe the main protein modules interacting with DNA

The transcriptional regulation (Prof. Straniero)
The transcriptional regulation of gene expression in eukaryotes. Enhancers, silencers and insulators. The epigenetic regulation of gene expression.
Learning objectives:

- Understand the concept of in-cis and in-trans regulation
- Discuss the interaction between modulation by transcription factor and chromatin status
- Explain how epigenetic information influences gene expression

Post-transcriptional regulation (Prof. Straniero)
Transcription attenuation. Alternative splicing, differential polyadenylation, RNA editing. Control of mRNA localization and stability. RNA interference.
Learning objectives:

- Explain the mechanism of attenuation in Bacteria
- Understand the importance of alternative splicing regulation
- Describe examples of post-transcriptional regulation
- Understand the role on nonsense-mediated mRNA decay
- Describe mechanisms regulating mRNA stability

Introduction to R and alignment with python for processing data, visualizing information and extracting knowledge (Prof. Pattini)

Learning objectives:

- Understanding how engineering methods for data access, manipulation and analysis can contribute to increasing biomedical knowledge
- Ability to use software tools prepare and manipulate data

Gene expression data analysis (Prof. Pattini)

High throughput technologies for gene expression profiling. Raw data processing. Statistics for differential analysis. Basics of data mining techniques for high level data analysis: unsupervised learning (clustering, dimensionality reduction). Functional annotation (sources and methods).

Learning objectives:

- Learn the general pipeline for RNA-seq data analysis
- Understand advantages and potentialities of RNA-seq based transcriptomics
- Learn basic explorative analysis of gene expression profiling

Practical activity 1: Biological sequence comparison (Prof. Pattini)

Examples of sequence analysis and similarity search with Blast.

Objective: using tools for sequence comparison and interpreting results

Practical activity 2: Introduction to R environment (Prof. Pattini)

A brief introduction to R commands

Objective: basic data manipulation and library use

Practical activity 3: Analysis of Whole Exome Sequencing (WES) data (Prof. Pattini)

An example analysis pipeline on a clinical dataset

Objective: identifying and characterizing gene mutations

Practical activity 4: Analysis of transcriptomic data (Prof. Pattini)

Total RNAseq data from cells treated or not with different compounds acting on splicing will be analyzed.

Objective: identifying genes and splicing isoforms differentially expressed in treated vs untreated cells.

Noncoding RNAs (Prof. Straniero)

The non-coding RNA revolution. Overview on the main classes of non-coding RNAs. Post-transcriptional regulation by small RNAs in prokaryotic and eukaryotic organisms. Long non-coding RNAs: structural features and mechanism of action. Competing endogenous RNAs.

Learning objectives:

- Explain the importance of RNA-mediated gene expression regulation
- Describe the main classes of noncoding RNAs (small and long)
- Illustrate the mechanism of RNA interference
- Describe microRNA biogenesis and mechanism of action
- Illustrate the features of long noncoding RNAs
- Recognize the complexity of mechanisms of action of long noncoding RNAs

Meet the expert: Splicing and disease

Dr. Emanuele Buratti, ICGEB, Trieste, Italy.

6. Genes and genomes

Human genome organization (Prof. Straniero)

Nuclear and mitochondrial genomes. The C-value paradox. Classes of repetitive DNA elements. Gene families. Organization of repetitive elements in the human genome. Segmental duplications.

Learning objectives:

- Discuss the differences between the nuclear and mitochondrial genomes
- List the main classes of genomic sequences, their function and their origin
- Describe the concept of pseudogene and transposon

Practical Activity 5: Surfing the genome (Prof. Straniero)

These practical lessons will allow to acquire confidence with online tools for medical doctors and researchers in the field of molecular genetics. Students will become able to get information on genes, mutations, and associated diseases starting from an anonymous sequence of DNA.

Gene therapy (Prof. Straniero)

Overview on the main approaches used for gene therapy. Traditional approaches with their pros and cons and genome editing (CRISPR/Cas9)

Learning objectives:

- Be able to describe the concept of gene therapy
- Understand the differences between in-vivo and ex-vivo gene therapy

7. Genetics of complex traits

Population genetics (Prof. Asselta)

Allele frequencies in populations. Hardy-Weinberg equilibrium. Genetic drift, population bottlenecks and founder effects. Genetic selection. The concept of the heterozygote advantage.

Learning objectives:

- Understand the applications and limitation of Hardy-Weinberg law
- Distinguish heterozygous advantage from founder effect
- Being able to calculate heterozygote frequency from disease prevalence

Genetics of complex (multifactorial) diseases (Prof. Asselta)

The polygenic and multifactorial nature of common diseases. Estimating the contribution of genetic and environmental factors. Linkage disequilibrium. Principle of allelic association.

Learning objectives:

- Understand the difference between dichotomic and continuous traits
- Describe the concept of linkage disequilibrium.
- Understand the difference between linkage analysis and association analysis

NGS for the identification of the molecular basis of Mendelian diseases. From genome sequence to the causative variant (Prof. Asselta)

From candidate gene analysis to whole genome sequencing. Targeted resequencing vs whole genome sequencing. Exome sequencing: flowchart and examples on how to design a study. The big challenge to fish out the pathogenic variants. How to find a novel disease gene: from theory to practice.

Learning objectives:

- Compare different NGS-based approaches for the molecular diagnosis of inherited disease
- Illustrate what is an exome and what are the critical steps in a standard exome-sequencing experiment
- Understand the hypotheses underlying different study design choices

Genome-wide approaches to complex diseases (Prof. Asselta)

LD structure of human genome. Genome-wide association studies (with a focus on COVID-19). Missing heritability. Burden of rare variants.

Learning objectives:

- Describe the modalities for conducting a genome-wide association analysis
- List possible genetic approaches to fill the gap of the missing heritability

What we have learned from omics approaches (Prof. Asselta)

Genomic consortia, Databases. From phenotypes to mutations, to mutations without a phenotype: redefining the classical concept of Mendelian diseases. The problem of incidental findings.

Learning objectives:

- Illustrate the key experiments that lead to the discovery of the LD structure of human genome
- List the principal consortia and publicly available databases with genomic data

Practical activity 6: Analysis of common germline variants (Prof. Asselta)
We will perform allelic association analyses at the genome-wide level. We will produce Manhattan and QQplots, and we will learn how to go deeper in association analyses (genotype analysis, epistasis, meta-analysis).

8. An additional level of complexity: Epigenetics

Epigenetics (Prof. Asselta)
Introduction to epigenetics. DNA methylation during development & Genomic imprinting. Chromatin structure & spatial organization of chromosomes. Regulatory RNAs: X-inactivation and sex determination. Environment & Epigenome
Learning objectives:

- Describe apparently unexplained phenomena now enclosed in the term Epigenetics
- Describe the Waddington epigenetic landscape
- Learn the difference between epigenetic and genetic modification
- Describe different functions of DNA methylation
- Describe the developmental process from the DNA methylation point of view
- Describe how genomic imprinting works
- Learn the most common genetic diseases associated with genomic imprinting problems
- Understand the different mechanisms leading to UPD
- Learn the most common genetic diseases associated with aberrant pattern of histone modifications
- Describe the organization of the nucleus, in terms of pores, lamina, and internal structures/granules
- Describe chromosome territories and TADs
- Learn the most common genetic diseases associated with a dys-organization of the nucleus structure
- Learn the basic modes of sex determination
- Describe the molecular mechanisms leading to X inactivation
- Describe how environmental factors can influence the epigenome

RNA-based pathogenic mechanisms (Prof. Asselta)
MicroRNA biogenesis, mechanism of action, and nomenclature. MicroRNA-mediated pathogenic mechanisms: implications in Mendelian and complex diseases. The ceRNA hypothesis
Learning objectives:

- Describe the biogenesis of microRNA
- Describe the 4 main mechanisms leading to miRNA-based genetic disorders

- Learn how a ceRNA network works, and how they can be related to genetic disorders

---

Teaching Methods

Lectures will mainly in presence, with the exceptions of few "meet-the-expert" seminars. Practical classes are also scheduled.

- Questions are always welcome during and after the lectures.
- We strongly encourage you to take notes during lessons, to share materials, and, when possible, to study together, helping each other.
- Slides of the lectures will be uploaded in the LMS platform, where additional material (video, papers, animations) will be posted by lecturers.

Assessment

Students' evaluation will be assessed through a written examination with multiple choice questions. The exam will include 33 questions, proportionally distributed among the different modules. Each question will score 1 point; no penalties will be applied for wrong answers. The threshold score for passing the exam will be 18 points.

Textbook

- Strachan & Read. Human Molecular Genetics 5th edition, Garland, 2018
- Alberts B. et al. Essential Cell Biology 5th edition, Norton, 2018
- Watson J et al. Molecular biology of the gene. 7th edition, Pearson, 2013
- Thompson & Thompson. Genetics in medicine. 8th edition, Elsevier, 2015
- Alberts B et al. Molecular biology of the cell. 6th edition, Garland Science, 2014

Code of conduct

- Students who falsify attendance to classes or any teaching activities will not be allowed to register for any of the dates published for the first exam session of the semester. They will receive an official warning letter from the Dean of the degree program and the Rector of the University. This breach of honor will be officially communicated to all the members of the Teaching Committee. If such behavior is repeated after receiving the first warning, the student will be required to repeat the academic year and the breach will be added to his/her academic career record. Additionally, the student will not be allowed to graduate with laude or with honors.

- Students who do not reach 75% of attendance will be required to pass an allocated part of the exam prior to being admitted to the final exam of the course. This part of the exam must be taken on one of the dates of the session preceding the date they intend to take the final exam. The student will receive a pass or fail evaluation, however, the result will not affect the final grade. Students who do not reach 50% of attendance in all the courses of the semester will have to repeat the academic year unless otherwise

decided in the light of specific and well documented reasons, submitted to a commission of the Teaching Committee.