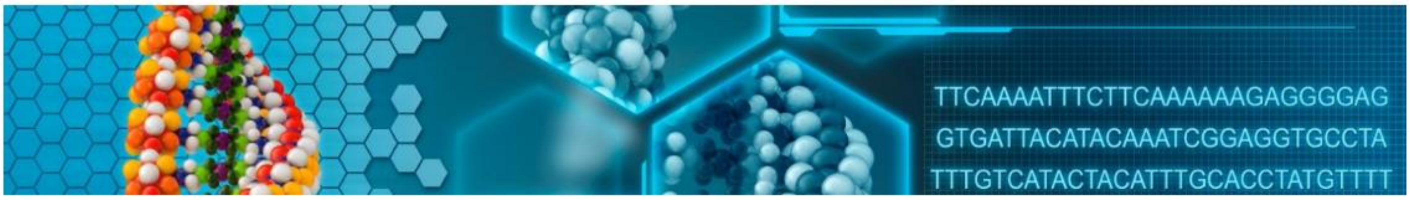

UVIC Msc in Omics Data Analysis

Location: Barcelona (location)
Language: English
Period: Classes: October to March; Master Project: April to September
Timetable: 3pm to 7pm Monday through Thursday

PROGRAM (60 ECTS)

The program of the master is divided into 3 compulsory blocks: Omics (35 ECTS), Applications (10 ECTS) and Final Master Project (15 ECTS).

OMICS (35 ECTS)

BIONFORMATICS (4 ECTS)

Course: Programming and database management for Bioinformatics
Course: Statistical and data mining methods for omics data analysis

GENOMICS (7 ECTS)

Course: Genome Bioinformatics
Course: Analysis of Complex Disease Association Studies

TRANSCRIPTOMICS (7 ECTS)

Course: Transcriptomics: Analysis of Microarray gene expression data
Course: Next Generation Sequencing analysis

EPIGENOMICS (5 ECTS)

Course: Epigenomics

PROTEOMICS (5 ECTS)

Course: Proteomics
Course: Metabolomics

INTERACTOMICS (7 ECTS)

Course: Interactomics, Systems Biology
Course: Integrative Genomics

APPLICATIONS (10 ECTS)

Course: Metagenomics
Course: Scientific communication
Conferences: Nutrigenomics, neurogenomics, pharmacogenomics, etc.

FINAL MASTER PROJECT (15 ECTS)

Programming and database management for bioinformatics

The aim of this course is to provide some important programming and database management skills that are essential for bioinformatics: using Linux operating system, managing databases, programming in Python, using and writing statistical functions in R.

Program:

- Linux
- Programming Languages for Bioinformatics: Python
- Database Management. MySQL
- R and Bioconductor

Lecturers and trainers:

M.Luz Calle. Biosciences Dept (UVic-UCC)) http://mon.uvic.cat/bms/members/malu/

Arnau Cordomí Laboratory of Computational Medicine, School of Medicine, UAB http://lmc.uab.cat/?page_id=6

Natalia Vilor-Tejedor Bioinformatic group in genetic epidemioloy. ISGlobal Barcelona Institute for Global Health

Statistical and data-mining methods for omics data analysis

The aim of this course is to introduce the most important statistical and data mining methods for bioinformatics and omics data analysis. The course combines lectures with hands-on sessions using R for illustration of the different methodologies.

Program:

1. Exploratory data analysis – Descriptive statistics
2. Important distributions
3. Principles of statistical inference
4. Important statistical tests
5. Multiple testing
6. Resampling methods for inference
    a. Bootstrap estimates and confidence intervals
    b. Permutation tests
7. Regression models
    a. Linear regression
    b. One and Two-way analysis of variance
    c. Logistic regression
8. Resampling methods for model selection and validation
    a. Bootstrap validation
    b. Cross-validation
9. Models for survival analysis
    a. Nonparametric estimators of the survival and cumulative hazard functions
    b. Semiparametric Cox's proportional hazards model
10. Unsupervised methods: Cluster analysis and PCA
    a. Linkage cluster analysis
    b. K-means cluster analysis
    c. Dimension reduction: Principal Component Analysis
11. Supervised data-mining methods for classification
    a. Classification. Measures of classification accuracy: classification error, sensitivity and specificity, ROC curve, AUC
    b. Classification and Regression trees (CART)
    c. Random Forest (RF)
    d. Support Vector Machine
    e. Neural Networks

Lecturers and trainers:

M.Luz Calle. Biosciences Dept (UVic-UCC) http://mon.uvic.cat/bms/members/malu/

Jordi Solé. Dept. of Information and Digital Technologies,
(UVIC) http://mon.uvic.cat/eps/research/research-groups/research-group-on-data-and-signal-processing/

Genome Bioinformatics

The aim of this course is to introduce the most important methods and tools for sequence analysis and sequence alignment in the context of comparative genomics and functional genomics.

Program:

- Concepts of genomics. Functional elements of the genome
- Genome Databases
- Introduction to proteins
- Introduction to protein databases
- Pairwise Sequence Alignments - Matrices - Dot Plots
- Pairwise Sequence Alignments - DP - Global - Local
- Similarity Search in Sequence Databases - BLAST
- Multiple Sequence Alignments -CLUSTAL
- Phylogenetics

Lecturers and trainers:

Ángel González. Laboratory of Computational Medicine, School of Medicine, UAB
http://lmc.uab.cat/?page_id=6

Mireia Olivella Biosciences Dept (UVic-UCC) http://mon.uvic.cat/bms/members

Josep M. Serrat Biosciences Dept (UVic-UCC) http://mon.uvic.cat/bms/members

Analysis of complex disease association studies

The aim of this course is to introduce the most important methodologies for the analysis of the genetic component of complex diseases. It is a practical course that combines lectures with practical sessions using R for illustration of the different methodologies.

Program:

- Variation in the Human Genome
- Population Genetics and Linkage Disequilibrium
- The International HapMap Project
- SNP prioritization and Tag SNP selection
- Genotyping platforms and Next Generation sequencing
- Association studies: Candidate Gene Studies, Candidate Region Studies, GWA Studies
- Data Quality Control: Population Stratification, Hardy-Weinberg Equilibrium
- Single-locus Tests of Association Studies: Chi-square test and logistic regression
- Haplotype analysis in Association Studies
- Confounding and Population Stratification
- Genome - Wide Associations Studies
- Genotype imputation methods
- Copy Number Variant Association Studies
- Analysis of Gene-environment and gene-gene interactions

Lecturers and trainers:

Mariona Bustamante Bioinformatics and Genomics Program. Center for Genomic Regulation (CRG) http://www.crg.eu/en/programmes-groups/genomics-and-disease

Juan Ramón González Bioinformatic Research Group in Genetic Epidemiology (BRGE), ISGlobal Barcelona Institute for Global Health http://www.creal.cat/brge.htm

Transcriptomics: Analysis of Microarray gene expression data

The main objective of this course is to introduce the most important methods of processing (preprocessing) and analyzing microarray expression data analysis. It aims to find out the main problems that can be studied with microarrays and how to design, process and analyze the corresponding experiments. Appropriate software to carry out each stage of the process will be introduced.

Program:

- Concepts of gene regulation
- Gene expression measurement
- Gene expression databases
- Experiments with DNA microarrays. Design and execution
- Data preprocessing: Exploration, normalization, filtering
- Detection of differentially expressed genes and related statistical problems (power, multiple comparisons, etc.)
- Classification and prediction from microarray data
- Functional analysis and biological interpretation

Lecturers and trainers:

Lara Nonell Microarray Analysis Service (IMIM) http://www.imim.cat/ofertadeserveis/en_sam.html

Eulàlia Puigdecanet Microarray Analysis Service (IMIM)
http://www.imim.cat/ofertadeserveis/en_sam.html

Josep M. Serrat Biociences Dept, EPS (UVic-UCC) http://mon.uvic.cat/bms/members

Next Generation Sequencing analysis

This is a hands-on training course with the aim of introducing the most important methodologies for NGS data analysis. The course starts with a brief introduction to NGS technologies and covers data analysis of RNA-Seq, ChIP-Seq and whole-exome sequencing experiments for variants detection.

Program:

- Next Generation Sequencing Technologies
- Bioconductor for high-throughput sequence analysis
- Short read formats
- Alignment of reads to a reference genome
- Alignment formats
- Summarization
- Exom sequencing
- DNA-seq experiments for variant calling: SNP and rare variant detection
- RNA-seq experiments for differential gene expression
- Gene set enrichment for RNA-seq differential expression results
- ChIP-Seq analysis of DNA regions of interest
- Annotating ChIP peaks

Lecturers and trainers:

Eulàlia Puigdecanet Microarray Analysis Service (IMIM)
http://www.imim.cat/ofertadeserveis/en_sam.html

Juan Ramón González Bioinformatic Research Group in Epidemiology (BRGE), Center for Research in Environmental Epidemiology http://www.creal.cat/brge.htm

Epigenomics

This course provides an overview of the epigenetic mechanisms and their link to gene regulation. The aim of the course is to introduce the most important methods for epigenomics data analysis.

Program:

- Epigenetic mechanisms of gene regulation
- DNA methylation
- Histone modifications
- Epigenetics databases
- Analysis of epigenetic profiles
- NGS epigenetic data analysis

Lecturers and trainers:

Antonio Gómez CRG (Centre for Genomic Regulation) http://www.crg.eu/es

Sebastián Morán Cancer Epigenetics (Idibell) http://www.idibell.cat/modul/cancer-epigenetics/en

Proteomics

The aim of this course is to introduce the necessary concepts and the most important computational methods for proteomics data analysis. The course combines lectures with hands-on sessions using R for illustration of the different methodologies.

Program:

Part 1. Introduction:
- Course introduction.
- What are our goals (and what are not)?
- Mass spectrometry applied to proteomics: A brief Overview.

Part 2. Quantitative Proteomics:
- Gel-based techniques (DIGE).
- Gel-free techniques:
  a. Label free.
  b. Labeling-based techniques:
    i. Isobaric labeling with tags.
    ii. Non-isobaric labeling.

Part 3. Data acquisition and protein ID:
- How to obtain data: DDA (data-dependent acquisition) vs. DIA (data-independent acquisition).
- Data output formats:
  - Proprietary formats.
  - Standard (free) formats.
- Protein ID:
  - Search in databases.
  - Peptide validation (Statistical approaches: Percolator and FDR).
  - Alternatives to database search (De Novo sequencing).

Part 4. Data analysis with R:
- Why R?
- Statistical Analysis of Data.
- R packages for Proteomics

Lecturers and trainers:

Joan Josep Bech-Serra Signalling and Cell Cycle Laboratory, IRB - Institute for Research in Biomedicine, Parc Científic de Barcelona, http://www.irbbarcelona.org/en/profile/joan-josep-bech

Metabolomics

The aim of this course is to introduce the most important concepts and computational methods for metabolomics data analysis. The course combines lectures with hands-on sessions using R for illustration of the different methodologies.

Program:

- Experimental design
- Sample analysis & metabolite identification by mass spectrometry (MS) and nuclear magnetic resonance (NMR)
- Data processing for NMR, LC/MS and GC/MS
- Statistical and chemometric analysis
- Pathway analysis

Lecturers and trainers:

Òscar Yanes Center for Omics Sciences URV http://www.yaneslab.com/lab-members/

Maria Vinaixa Synbiochem. Manchester Synthetic Biology Research Center http://synbiochem.co.uk/people/

Miguel Ángel Rodríguez Center for Omics Sciences URV http://www.yaneslab.com/lab-members/

Interactomics: Systems Biology

This course, with a strong practical focus, is intended to provide students with the capacity to use data from databases of molecular interactions to build and analyze biological networks. The aim of the course is the analysis of network topology, the modeling of motif dynamics, and to establish relationship between topology and biological function.

Program:

- Biological interactions: classes and biological relevance.
- Methods for the determination of molecular interactions.
- Molecular interaction databases.
- Parsing interaction databases.
- Building graph objects.
- Topological analysis of graphs.
- Genetic circuits dynamics
- Noise in genetic circuits
- Multicellular genetic circuits
- Integrative omic analyses
- Interactome properties and analyses
- Systems-level view of cancer

Lecturers and trainers:

Lara Nonell Microarray Analysis Service (IMIM) http://www.imim.cat/ofertadeserveis/en_sam.html

Integrative Genomics

The aim of this course is to introduce some relevant methodologies for Integrative Genomics in order to gain some insight into the underlying biology. It is a practical course that combines lectures with hands-on sessions, mainly in R, for illustration of the different methodologies. Several omics data sets will be analyzed, integrated and visualized.

Program:

- Concepts and tools overview
- Intregration based on correlations:
    - mRNA & miRNA data integration
    - genomic and transcriptomic data integration data
- An n-dimensional approach:
- Functional integrative genomics
- Visualization

Lectures and trainers:

Lara Nonell Microarray Analysis Service (IMIM) http://www.imim.cat/ofertadeserveis/en_sam.html

Introduction to Metagenomics

This course will introduce the field of metagenomics as well as the areas where it can be applied. The aim is to understand what questions can be explored with this technique and get an idea of the existing tools and approaches for data analyses. The course will also include hands-on sessions. The main focus will be microbial metagenomics, a field that has grown dramatically during the last five years.

Program:

- Introduction to metagenomics and its relation to/from Next generation Sequencing. The rDNA gene. ITS and 18s.
- Metagenetics: The rDNA gene. Amplicon sequencing design in Illumina technology. The tree of life.
- Data Analysis:
    - Quality Control. Chimera removal.
    - Sequence Clustering. Available approaches.
    - Operative Taxonomic Unit, what is it?
    - Alpha Diversity. Richness and diversity indices.
    - Rarefaction and normalization. Advantages and Caveats.
    - OTU artifacts and the importance of filtering. OTU table sparsity
- Hand-On Session(2h):
    - Introduction to Mothur. Installation and conventions.
    - Knowing your data.
    - Sequence analysis and OTU generation. From Sequences to OTU counts
    - OTU table data analysis. Descriptive analysis.
    - Alpha Diversity Analysis. Richness and Diversity tests
- Data Analysis(continuation, 2h)
    - Taxonomic classification: Databases & Algorithms and their advantages and limitations
    - Beta Diversity: Ecological distance and indices and their relation to data transformation.
    - Ordination Analysis: Polar Ordination, Principal Component analysis and Principal Coordinate analysis.
    - Differential Abundance.
    - OTU correlation with metadata.
- Hand-On Session(2h): 2
    - Taxonomic classification. From OTUs to bacterial genus.
    - Ordination analysis. Showing Patterns in bacterial communities.
    - Differential Abundance. Detecting disease associated bacteria.
- Metagenomics: Shotgun Sequencing. Experimental design
- Human Microbiome. Ongoing Projects and available web resources
- Concepts in Data analysis:
    - Sequence Quality Control
    - Metagenomics for taxonomical classification.
    - Metagenomics assembly
    - Gene prediction & Annotation
    - Gene Richness clinical utility
    - Co-Abundance groups and Metagenomic Species.
- Hands-on session (Continuation, 2h):
    - Functional Inference using 16s marker with PiCrust.

Lecturers and trainers:
Marc Noguera-Julian IrsiCaixa Research Institute; AIDS Chair (UVic-UCC)

Scientific communication

The goal of this course is to provide the knowledge to successfully

- write a scientific review paper
- make a presentation
- design a scientific poster
- review a scientific document

Lecturers and trainers:

Xavier Rovira, Biosciences Dept (UVic-UCC)