

# Microarray Data Analysis

Barcelona | June 16, 17 and 18 | 2008

**IMIM**  
*hospital del mar*



Teachers from the Bioinformatics  
Department of the Centro de Investigación  
Príncipe Felipe in Valencia

Ignacio Medina  
David Montaner and  
Joaquín Dopazo



#### Registration:

Please send a mail to [mtora@imim.es](mailto:mtora@imim.es) with your cv and brief description of your interests. The course is restricted to a maximum of 20 participants.

#### Place:

UDIMAS. Facultat de Medicina, UAB  
C. Doctor Aiguader, 80 - BARCELONA  
Aula d'informàtica

#### Organization:

Serveis Científicotècnics,  
IMIM-Hospital del Mar

[www.imim.es](http://www.imim.es)

A three days theoretical and practical course on microarray data analysis oriented to experimentalists, end-users and PhD students who want to learn about the state-of-the-art of the data analysis methodologies and acquire the skills necessary for analysing their own data using one of the most advanced software packages, the GEPAS.

## Monday 16

10:00-13:00

### Presentation and Introduction

Structure of the course. Why microarrays? Pre- and post-genomics hypothesis testing: a note of caution. Design of experiments. Data preprocessing and normalization. Unsupervised analysis (clustering). Supervised analysis (gene selection, predictors). Functional profiling.

15:00-18:00

### Normalization

Getting rid of unwanted variability from sources other than the experimental conditions assayed. Methods for Affymetrix, two-color and one-color microarrays. Theory and practical exercises using GEPAS.

## Tuesday 17

10:00-13:00

### Gene selection

Methods for selecting genes differentially expressed among two or more experimental conditions, correlated to a continuous variable or correlated to survival time. How to deal with the multiple-testing problem. Theory and practical exercises using GEPAS.

15:00-17:00

### Predictors

Gene selection in the context of class prediction. How to deal with the selection bias problem. Different methods for class prediction. Estimating the error of classification. Interpretation of confusion matrices.

17:00-18:00

### Clustering

Different clustering methods: hierarchical clustering, SOM, SOTA and k-means. Pros and cons. Measures of cluster quality. Cluster visualization.

## Wednesday 18

10:00-12:00

### Gene Annotation Data Bases Overview

Most popular gene and protein annotation repositories be revised. These are some of the data that the Babelomics suit imports for the functional profiling of genomic experiments.

12:00-13:00 and 15:00-17:00

### Functional profiling of genomic experiments

Understanding the biological roles played by the genes in the experiments. Using different types of information for the functional profiling of microarray experiments: Gene Ontology, InterPro motifs, transcription factor binding sites, gene expression in other experiments, text-mining, etc. New trends in the analysis of microarray data: testing pathway-based or function-based hypothesis.