



# Microarray Course

15<sup>th</sup> and 19<sup>th</sup> February 2021

## Organization

The course consists of lectures in the morning (09.00 - 11.00) and exercises in the afternoon (12.00 - 16.30) Central European Time, UTC+1

## Lectures

### 1: Introduction

- Microarrays
- R vs Excel: Why use R
- Reproducibility
- First steps in R
- Statistic measures

### 2: Normalization

- Bias vs variance
- MA-plots
- asinh, quantile, RMA normalization

### 3: Screening

- Comparing conditions
- Testing
- p-values
- Multiple testing
- QQ-plots

### 4: Clustering and Classification

- Principle
- Hierarchical, k-means, PAM
- Silhouette score
- Training and test set
- Overfitting
- Cross-validation
- Regularized regression (LASSO)

### 5: Local R usage

- Setting up R and RStudio
- Coding best practices
- Introduction to RStudio
- Reports

## Exercises

### 1: Introduction

- 01 Jupyter notebooks
- 02 Introduction to R
- 03 Tidyverse
- 03 Data visualization

### 2: Normalization

- 04 Loading data
- 04 Inspecting microarrays
- 04 Preprocessing
- 05 Quality control (plots!)

### 3: Differential gene expression

- 06 Filtering
- 06 Testing
- 06 Plots
- 06 Multiple testing
- 06 Annotation

### 4: Clustering and classification

- 07 Hierarchical, dendrograms
- 08 K-means, PAM
- 09 Training and test set
- 09 LASSO

### 5: Local R usage

- Set up R on your PC
- Usage, analysis, plotting
- Automated reports

