

Presentation

We have learned in class how microarray technology provides a powerful tool for biomedical research by allowing to quantify, simultaneously, the expression levels of thousands of genes.

The goal of this practical exercise is two-fold:

1. Starting from a problem statement and an appropriate dataset you must analyze them by following the standard approach discussed during the lectures.
2. After doing the analysis you must write a report following the traditional structure of a scientific work (see *Guidelines of the Report*).

The data for the analysis

You have been asked to collect one dataset and this is what you should use for your assessment. Ideally the analysis can be nicer if you have at least two groups, but this can or cannot be your situation so don't worry particularly about this.

Pipeline for the Analysis

As we learned in lecture classes a microarray -and most omics- data analysis follows a pipeline of ordered steps. We saw that the standard *pipeline analysis* consists of

1. Data capture from images or text files
2. Quality assessment of raw data
3. Normalization
4. Quality assessment of normalized data
5. Filtering data (optional)
6. Identification of differentially expressed genes
7. Annotation
8. Multiple comparisons
9. Functional interpretation

Thus, in order to perform a whole analysis, you have to write an R code that collects all the chunks used for. In order for the code to be reproducible I don't expect a script but a Rmarkdown document and its renderization into an html document using the knitr package.

Delivery format

In order to evaluate your work, you must deliver three files:

- 1) The original Rmarkdown document with the code for the analysis and the explanations.
- 2) The R code that you extract from the Rcmd file with the purl command

- 3) An .html file with your report.
- 4) Any additional file that you consider appropriate.

The submission must be done two-fold

- Updating your github repository: The exercise should go to folder “Exercise_3”
- Uploading the delivery to the campus using this task. The report should contain the link to clone the repository. When you upload your results include them in a compressed file named with your surname and name (e.g. SanchezPlaAlex-reportMDA.zip).

Guidelines for the Report

Traditional structure of a scientific work should contain well organized the following sections

1. Abstract: should have no more than five lines.
2. Objectives: what is the goal of the study and what are the specific objectives
3. Material and methods:
 - Data type, type of experiment, experimental design, type of [microarrays] used,...
 - Methods that have been used in the analysis
 - General procedure of analysis (steps, “workflow” or “pipeline” that you have followed)
 - Procedures applied in each step (you do not need to provide the details of each method). Instead provide a qualitative description indicating why you have carried out this step, what is the input provided and what output has been produced.
4. Results: what has been obtained as a result of the analysis
5. Discussion: what limitations do you think that study has
6. Conclusions: of your work, NOT from the biological problem analyzed.

The preceding description is a basic proposal. You can adapt it at your convenience.

The report should be written in English.