

High Throughput Biostatistical Analysis

ACOBIONM is a Biotechnology company specialized in the discovery of new **Biomarkers** and in the development of innovative **Diagnostics**. ACOBIONM's platform combines **molecular biology techniques** (NGS, high throughput sequencing, quantitative RT-PCR...) and specific tools & programs in **bioinformatics and biostatistics** to perform a large range of RNA/DNA analyses.

In case of RNA-sequencing (RNA-seq) analysis, there is a wide variety of applications. So, a single analysis pipeline doesn't "work" for all cases, but a customized approach is much more appropriate to identify the relevant Gene Markers associated with the biological question of interest.

Note that a crucial prerequisite for a successful RNA-seq study is that the generated data have the potential to answer to the biological questions of interest. This is achieved first by defining a good experimental design (choice of the library type, choice of the sequencing depth and the number of replicates appropriate to the biological system under study), and second by planning an adequate execution of the sequencing experiment itself, ensuring that data acquisition does not become contaminated with unnecessary biases.

Hence, Acobiom will design your RNA-seq biostatistical analysis according to your specifications, both for experimental design and data analysis.

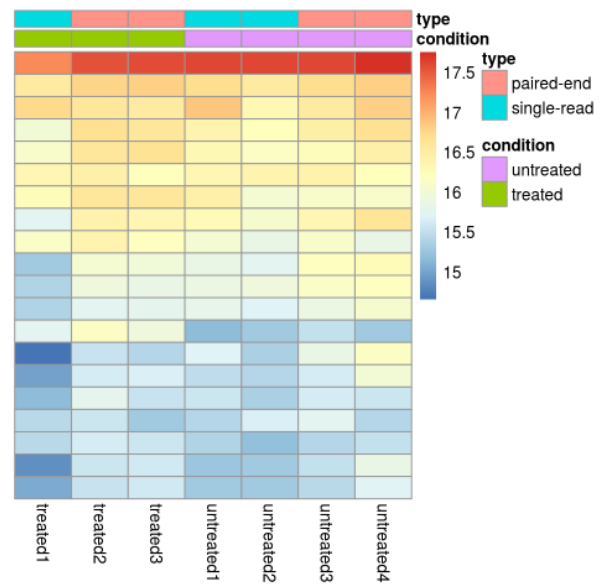
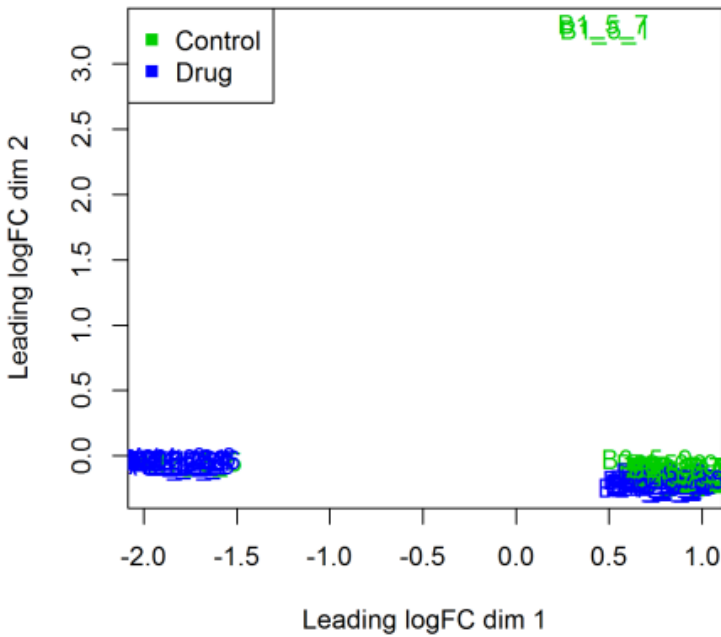
Below, a generic roadmap for RNA-seq data analysis is described according to different stages. The major analysis steps are listed above the lines for pre-analysis, core analysis and advanced analysis. At last, we present visualization methods to confirm the integrity of the data and the process.

- **Input/Output (I/O)**
 - Input: Data Matrix (RNA-Seq raw reads), gene-level annotation metadata, sample-level annotation metadata.
 - Output: List of differentially expressed genes corresponding to a transcriptome signature enabling samples stratification. *The Objective: answering an initial biologic question.*
- **Pre-processing: Reads alignment, gene-level expression quantitation, quality assessment/control, sample filtering.**
 - Quality assessment/control (QA/QC), Numerical and graphical summaries of gene-level read counts, gene filtering.
 - Normalization and expression quantitation
 - Exploratory Data Analysis (**Premium Analysis**)
- **Differential Expression (DE) analysis:** Identification of Gene Markers.

1. Pre-processing

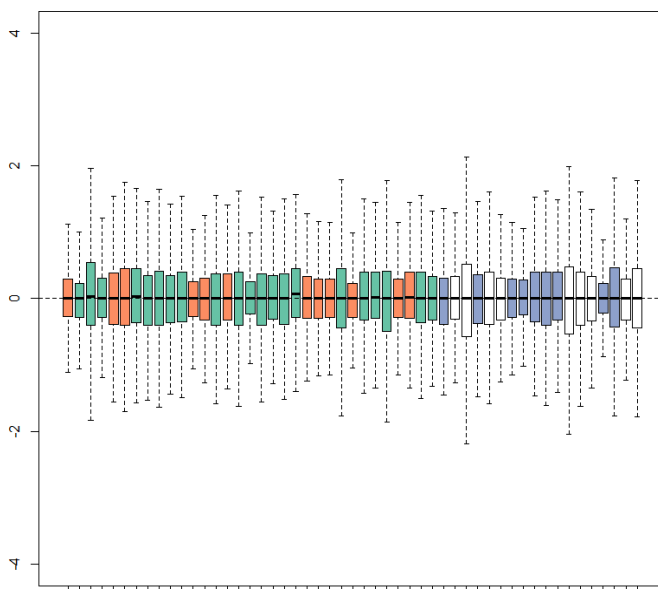
- QA/QC
 - Using MDS-Plot and other tools to visualize data.
 - Filtering out transcripts adaptively and optimizing the sequencing depth. (Premium Analysis)
- Using exploratory data analysis to assess the best normalization strategy. (Premium Analysis)

MDS Plot: drug treatment colours



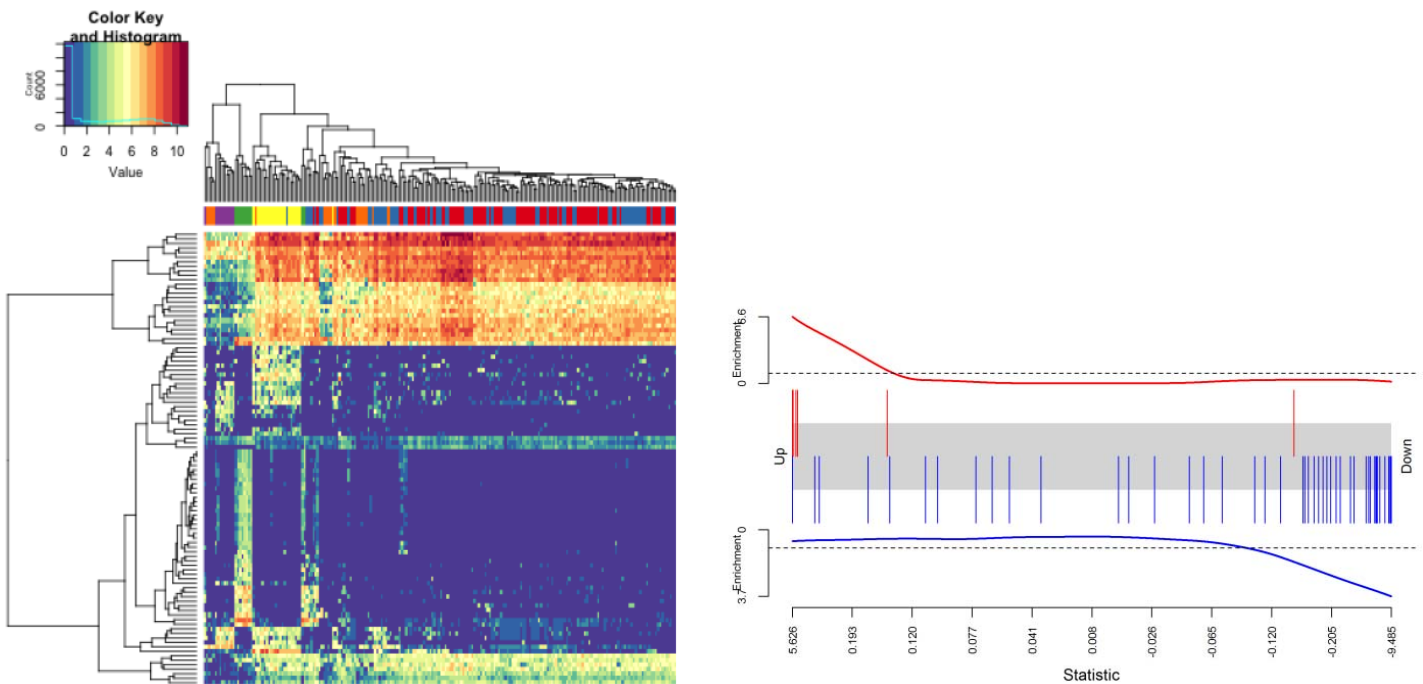
2. Normalization

- Normalization is crucial to ensure that differences observed in expression measures between samples are truly due to differential expression, but not to technical artifacts.
- ACOBIOM will use additional normalization's strategies in order to assess for differences in sequencing depths (i.e., total read counts) and underrated technical effects, e.g. library preparation, flow-cell effects or lane effects. **(Premium Analysis).**
- Acobiom's strategy for normalization is to test various methods (Full Quantile, Upper-quartile etc.) and compare the expression yielded to select the most reliable strategy.



3. Differential Expression (DE) analysis

- The Differential Expression analysis is done with a process adapted from two **R packages** “**edgeR**” and “**DESeq**”.
 Acobiom’s team combines the results of these two methods to select the most robust Gene Markers in order to answer your questions.
 Acobiom adds two more **R packages** (“**DEXSeq**” and “**DESeq2**”) to increase drastically the resolution of your results. (**Premium Analysis**)
- Results can be displayed in easy interpretable plot like heatmap or any other plot of your liking and pertinence for your presentation/publication usage.



- Once a list of primers has been determined, Acobiom will use another technology: **Real time PCR in order to create the future signature.**

Data resulting from performance of your project will be confidential: your scientific team will be the owner of the full data (raw and treated), that will be delivered on physical memory storage.

Acobiom’s team is engaged in a constant scientific monitoring process for using the latest version of software and performing on the most recent omics analysis tools (*R software, EdgeR, DESeq, DESeq2, DEXSeq, ddCt...*).

Contacts

- Pascal LAURENT, Project manager, Mob: +33(0)607.352.248 - laurent@acobiom.com

Note: Acobiom is labeled « *Crédit Impôt Recherche* » **Agreement** for the period: 2015-2019