Omics Data Biostatistical Analysis

Acobiom is a Biotechnology Company specialized in discovering **new Biomarkers** and in developing innovative **Diagnostics**. Acobiom's platform combines **molecular biology techniques** (NGS sequencing, quantitative RT-PCR...) and specific tools/programs in **bioinformatics and biostatistics** to perform a large range of RNA/DNA analyses.

qPCR statistical analysis pipeline

- **Input/Output (I/O)**

  Input: Data Matrix (Cq raw reads), sample-level annotation metadata.
  Output: List of significant markers combined in a robust signature leading to a sample stratification/selection.

- **Pre-processing:** Missing values, quality assessment/control, sample filtering.
- **Quality assessment / control (QA/QC):** Numerical and graphical summaries of gene-level cq, gene filtering.
- **Normalization and expression quantitation**.
- **Fitting into an adjusted model** after selection using the most adequate solution to your problem.
- **Optimization on the signature model** (using FDA-approved method).

1. Pre-processing and Normalization

- **QA/QC**
  - Each Cq values are filtered in order to yield sufficient robust information across all predictive variables (conditions, pathological status, quantitative measure, etc.).
  - Housekeeping genes are evaluated though an adaptated M-values method (Vandesompele *et al*, 2002).
  - Each run is evaluated through its replicates by controlling multiple parameters of distribution.
- **Normalization**

  Classic step of 2ΔΔCq normalization (Livak et al, 2001).
  $\Delta Cq_{GOI} = [Mean\ Cq_{GOI}] - [Mean\ Cq_{HKG}]$
- **Expression quantitation**

  All analyses are run on the ΔCq values only (after the first normalization of the 2ΔΔCq method).

2. Model fitting

Elastic-Net regression

- ACOBIOM uses an application based on the « elastic-net » regression (Zou, Hui; Hastie, Trevor 2005).

ACOBIOM | 1682, rue de la Valsière - CS77394 - Cap Delta – Biopole Euromédecine II | 34184 Montpellier Cedex 4 | T: 04 67 41 97 48 | info@acobiom.com | www.acobiom.com

- The algorithm uses cyclical coordinate descent, which successively optimizes the objective function over each parameter with others fixed, and cycles repeatedly until convergence.
- The algorithm associates a range of genomic variables randomly comprised between *1* and *n* genes (any threshold of your liking can be used) until convergence in order to predict the output variable (survival, class selection, quantitative markers etc.).

Acobiom's team compute hundreds of thousands models (the optimized and sufficient number of computed models will be selected on line with your data) and all pertinent models related to the objective will be sorted.

For this selection, Acobiom associates for each gene a score based on its weight across all the models in order to discard the random effect association of variables and select a more "reality-closed" model.

ACOBIOM then computes a "key" or a "new omics variable" to combine the selected genomic variables (the signature) in an « **omics-score** » or « index » associated to each sample.

3. Optimization

Omics score

Once the signature and the « omics-score » have been computed, the Company can optimize the threshold of segregation between two (or more) populations: for example « good responders » and « bad responders » to a treatment and define a buffer-zone between to the two populations to help the decision in a safe way.

In the same manner as described above, Acobiom uses several simulations to estimate the most optimized parameters.

All this steps use cross validation methods based on hold-out method (proportion will be evaluated depending on your needs and your dataset). This will result in obtaining the most optimized model (specificity and sensitivity wise).

The omics score can be used in addition to several variables, e.g. in personalized medicine the clinical variables, as the pain, localization of tumor, etc.

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

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

Contacts

- David PIQUEMAL, CSO, Mob: +33(0)610.430.140 - piquemal@acobiom.com
- Didier RITTER, CEO, Mob: +33(0)662.636.396 - ritter@acobiom.com
- Pascal LAURENT, Project manager, Mob: +33(0)607.352.248 - laurent@acobiom.com

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