We invent nothing,*
we reach to make sense to biological data.

Bioinformatics and Biostatistical Treatments

"***RNA transcripts from biological samples can be easily sequenced, but making sense of those that fail to map exactly to a reference genome is tough***" (Nature, May 2013).

The volume and complexity of data from High Throughput Sequencing (HTS) and Real-Time PCR (qPCR) experiments need scalable, fast and mathematically principled analysis software.
Acobiom technical team provides support to its partners and customers for designing the study and the experiment plan to ensure the best treatment strategy which will be applied to the project needs and objectives. Thus, according to requirements and specifications of *in silico* treatments, the protocol will begin with raw sequencing reads analysis (FastaQ for example). Then, a transcriptome assembly will be performed, a list of differentially expressed and regulated genes, transcripts or exons, will be edited, and publication-quality visualizations of analysis results will be established. Finally, Acobiom is able to develop a personalized database with a "user-friendly" web interface to increase integration of data and accelerate partners' cooperation.

Application #1: For example, in RNAseq data treatment with adapted and specific parameters, TopHat (http://tophat.cbcb.umd.edu/) and Cufflinks (http://cufflinks.cbcb.umd.edu/) tools allow to identify new genes and new splice variants of known genes, as well as to compare gene and transcript expression data under two or more conditions. Output data are compatible with several accessory tools and utilities helping in managing data, including IGV (http://www.broadinstitute.org/igv/), a tool for visualizing RNAseq results. Then, R package CummeRBund is used to complete the statistical analysis. Differential expression exon analysis of count data can be implemented using the Python scripts **htseq-qa** (http://www-huber.embl.de/users/anders/HTSeq) and DexSeq (http://www.bioconductor.org/).

Application #2: Newly, the method CRAC (Philippe N and coll., Genome Biol. 2013) allows a combined integration of Digital Gene Expression (DGE) and RNAseq data to quantify. This method uses a double *k*-mer profiling approach to **detect candidate mutations, indels, and splice or fusion junctions in each single read**. In addition, output files are compatible with tools for investigators to understand biological meaning behind large list of genes, like DAVID (http://david.abcc.ncifcrf.gov/) or Blast2Go (*www.blast2go.com/*). The latter is a suite as a comprehensive bioinformatic tool for functional annotation of sequences and data mining on the resulting annotations, primarily based on the gene ontology (GO) vocabulary. The tool includes numerous functions for the visualization, management, and statistical analysis of annotated results, including gene set enrichment analysis.

Application #3: *Other case with high throughput quantitative real-time PCR benefits from* extensive miniaturization with more than 30,000 PCR reactions in a single day from 300ng of total RNA. RNA levels quantification could be assessed by calculating $2^{-\Delta\Delta CT}$ (Schmittgen et Livak, Nat Protoc, 2008). For comparison of the quantitative variables the normality is tested using Shapiro-Wilk, the normality was rejected and a non-parametrical method was used to assert of significant differential gene expression. A Kruskall-Wallis test and a Nemenyi test are used to identify the difference (R software v2.14.0).

**'When Picasso first saw the ancient cave paintings at Lascaux he declared: "We have invented nothing…"'*

Acobiom
Offices & Laboratories: 1682, rue de la Valsière - CS77394 - Cap Delta - Biopole Euromedecine II - 34184 Montpellier Cedex 4 - France
Tel.: +33 (0)467 419 748 – Fax: +33 (0)467 457 726 - Email: info@acobiom.com - Web: www.acobiom.com