

# Identification and Visualization of Expression Patterns by the Integration of Pathways, Transcriptome and Proteome profiles

Henrique Cursino Vieira<sup>1</sup>, Bruno Ferreira de Souza<sup>2</sup>, Hugo A. Armelin<sup>3</sup>, Milton Yutaka Nishiyama Junior<sup>4</sup>,

*1 LECC-CeTICS, Instituto Butantan*

*2 ECC-CeTICS, Instituto Butantan*

*3 Instituto Butantan*

*4 LETA-CeTICS, Instituto Butantan*

## Abstract

The integration of multi-omics data is a great challenge and is necessary for the understanding of the complexity of biological systems and signaling network interactions. Discovering the gene or protein expression signature associated with a specific treatment or condition is the basic question. However, the studies based in only on molecular level (e.g. genome, transcriptome, or proteome) may be incomplete and fail to reveal the multi-layer interactions in the different molecular levels. The computational approaches for good predictions and efficient data integration are not well established yet and can be dependent on the specific experimental design. Many scientists are looking for integrated and graphical tools for visualization of networks, pathways and generate pictures and tables presenting their findings. This work is part of The Center of Toxins, Immune-response and Cell Signaling (CeTICS), which aims to understand the behavior of biological systems based on analysis of -omics data and signaling networks; The studies and research in CeTICS are intrinsically interdisciplinary, which is coupled to the -omics data from genomics, transcriptomics and proteomics, heterogeneous knowledge. The necessities to the integrative analysis of such multiple layers, is especially focused on the most preferred methods for understanding global gene regulation: high-throughput RNA sequencing and mass spectrometry (MS) expression profiles, coupled to the integration with the relative abundance of metabolic and signaling pathways, which can be increased with the application of machine learning methods and viewers for the inspection and analysis on-the-fly. Data visualization tools can be categorized into two types, although they can overlap: tools focused on automated methods for the interpretation and exploitation of large biological networks; and tools for assembly and validation of tracks. Thus, we present a Shiny application for an interactive viewer with automated analysis tools, which will enable novel conclusions to be drawn from transcriptomic and proteomic integrative analysis. The researcher will can interact with the application, so it can view the content in a clean and fluid way, making possible aggregate data from different biological sources. The goal is beyond visualization, and will allow performing multivariate analysis for systems-level, understanding from multi-dimensional data and be optimized to show the results through the graphs. Finally, our mid-term objective is to develop an integrative approach, which will be aggregated to the CeTICSdb platform, which will be available to the scientific community.

Funding: #2013/07467-1, São Paulo Research Foundation (FAPESP)