CEMiTool: Co-Expression Modules Identification Tool

Pedro S. T. Russo[1,2]; Gustavo Rodrigues-Ferreira[1]; César A. Prada-Medina[1]; Matheus C. Bürger[1,2]; Lucas Esteves Cardozo[1], Luciane Schons-Fonseca[3]; Thiago D. C. Hirata[1]; Gonzalo Sepúlveda-Hermosilla[4], Vinícius Maracajá-Coutinho[4]; Helder I. Nakaya[1,2]

1 - School of Pharmaceutical Sciences – University of São Paulo, São Paulo, SP, Brasil; 2 - Graduate Program in Bioinformatics - Institute of Mathematics and Statistics - University of São Paulo, SP, Brasil; 3 - Department of Biology - Massachussets Institute of Technology, Cambridge, MA, USA; 4 - University of Talca, Talca, Chile

Systems Biology approaches provide a holistic view of biological processes, integrating the many different molecular cell components through highly complex networks. By analyzing these networks, it is possible to create mathematical and computational models that help understand the mechanisms triggered by different types of stimuli or perturbations. Also, predictive modelling can be applied to reveal gene signatures associated with systemic responses, such as disease outcome or vaccine-induced immunity. However, in addition to the stochasticity of biological processes and the noise associated with high-throughput technologies, gene networks are very dynamic and highly sensitive to changes in conditions and experimental settings. Therefore, given the inherent modularity of biological systems, in this project we developed an easy-to-use web-based application called CEMiTool, which aims to identify the underlying modules in gene co-expression networks and analyze their changes in response to different types of stimuli and perturbations . CEMiTool's code is written completely in R programming language. In order to test our tool, we ran CEMiTool on Juvenile Idiopathic Arthritis, inflammatory bowel diseases (ulcerative colitis and Crohn's disease) and sepsis studies. Our analyses revealed several genes coexpressed in the different inflammatory diseases. Modules included genes associated with inflammatory and anti-viral pathways, such as type I interferons, NOD1/2 signaling pathways and TNFR1-induced NFkappaB signaling pathways. We found that some of the gene modules were unique to a disease whereas some modules were shared by different diseases. We also associate the activity of the modules with the inflammatory score of patients. This analysis provided novel insights to disease mechanisms.