

Talk

Use of bioinformatics tools to find new genes involved in rare diseases



Cristian Jesús Bazo Ramos(1), Gualberto Asencio Cortés(2), Antonio Jesús Pérez Pulido (1)

(1)Genetics Department, Pablo de Olavide University of Seville, Spain, (2)Computer Science Department, Pablo de Olavide University of Seville, Spain.

Keywords: Bioinformatics; Rare disease; Transcriptomic; Gene Ontology.

ABSTRACT

Motivation: Rare diseases are a big challenge of our community and it is important to find answers and put forward a cure in the medicine field. Thanks to the huge amount of data that transcriptomic researches provide to public databases, we can use bioinformatics tools to analyse and seek new paths to understand better their molecular mechanisms and find new molecules that are related to the disease in order to make a future drug discovery process to this kind of research.

Methods: We use transcriptomic data from the Expression Atlas repository, searching for experiments where the gene related to rare diseases is differently expressed.. We use the fold change data to choose those proteins that the expression are correlated to the expression of our gene of interest ($R^2 \geq 0.95$). Using enrichment tools from Reactome database, or DAVID computational tool, we can establish a Gene Ontology (GO) study among which we we can choose those that belong to the same biological process and path. This first step means an approach to select from thousands of genes a few gene cluster that may be highly related with the gene that cause the disease. The use of analysis tool R with bioinformatics packages, such as Bioconductor, CompGo, RDavidWebService or ClusterProfiler, allow us to keep improving the methodology making a deep analysis of Gene Ontology of our gene cluster, crafting relationships between them.

Results: The current status of this research consists in the analysis of all GO terms that are belonged to our genes of interest that were crossed with the terms of the gene related to the studied disease. This step is crucial in order to find genes that are also affected by rare diseases in their metabolic path. This methodology could discover new biomarkers or, in another case, new strategies to understand the correct operation of the biological process of rare diseases and most importantly, the possibility to find a possible cure for these conditions.

REFERENCES

- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc.* 2009;4(1):44-57
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37(1):1-13. [PubMed] Crenshaw, B.,III, and Jones, W.B.,Jr (2003) The future of clinical cancer management: one tumor, one chip. *Bioinformatics*, doi:10.1093/bioinformatics/btn000.
- Fabregat et al. 2018 PMID: 29145629
- Sam D. Bassett and Ashley J. Waardenberg (2018). CompGO: An R pipeline for .bed file annotation, comparing GO term enrichment between gene sets and data visualisation. R package version 1.18.0.
- Cristobal Fresno, Elmer A. Fernandez (2013). RDAVIDWebService: a versatile R interface to DAVID Bioinformatics, 29(21), 2810-2811. URL <http://bioinformatics.oxfordjournals.org/content/29/21/2810>.

