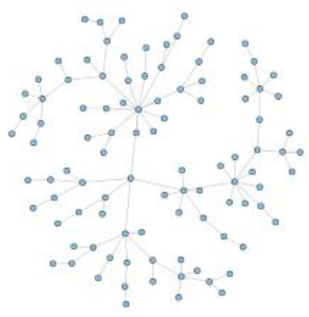
VIII CONGRESO INTERNACIONAL DE INVESTIGACIÓN REDU
UNIVERSIDAD TÉCNICA DE AMBATO

Utility in Genetics of Molecular Interaction Networks and Biological Pathways through the usage of bioinformatic tools.

Cristian Fernando Galarza Galarza[1], Luis Felipe Varela Polit[2], Cinthia Katherine Galarza Galarza[3], Alex Fabián Valencia Silva[1]
Institución: Universidad Técnica de Ambato[1]; Medicina General, Área de Terapia Intensiva, Hospital General Riobamba, Chile 3929 y Av. Unidad Nacional[2]; Instituto CMECH[3]
Autor para correspondencia: Cristian Fernando Galarza G.[1], cf.galarza@uta.edu.ec



INTRODUCTION

Studying protein-protein interactions (PPI) is important, through study it is possible to understand how it affects gene repression in different scenarios or how gene expression processes are affected when a certain drug is applied.

The interconnection of information within the biological network is represented in bioinformatics tools for its study, in which the nodes, axes and their interactions can be analyzed. These have the particularity of being unevenly distributed, within networks some nodes act as concentrators, this means that they have a high degree of connection, which in practice allows us to identify satellite genes.

The structure and dynamics of networks are independent of size, in such a way that the probability that a node in the network connects with others follows an exponential relationship considering the number of nodes in the network..

MATERIALS AND METHODS

For the biological network, the file galFiltered.sif is used, for the attributes the file galExpData.csv is used. These files contain information on yeast GAL gene expression data from transcriptional regulation studies.

The files have eight columns, the first two describe characteristics that identify the gene, the next six represent three different experimental conditions (gal1R, gal4R, gal80R) in each of them the expression of a particular gene has been repressed gal1, gal4 and gal80 respectively. Therefore, each condition has two columns, one for the eigenvalue of expression and the other for the significance value.

Other parameters in the tool allow to visualize the network with characteristics such as color of the nodes to represent repression or overexpression of genes, as well as to observe clusters of genes that have common characteristics, to see the characterization of expression or overexpression.

CONCLUSIONS

Today's tools show a great advance in their design and conception since through the different utilities they bring allow us to solve problems that can be present in the different investigations that we carry out. A very important resource are the graphs, as it is shown in the development of this work, since many details are not as visible by performing a tabular analysis only.

In biological networks that are normally extensive, the application of filters allow us to improve the visualization of certain areas in which the objectives of our study are focused, this is important when limiting the observation of the data.

Detection of PPI or DNA-protein interactions is very important within the scope of proteomics, and in general in genetics since it offers the possibility of finding ways of genetic expression which within these studies is very important for the detection of diseases as well as to determine harmful reactions of the body to medications, which would lead to an improvement in the quality of life of individuals.

REFERENCES

Babu, M. et al (2012) Interaction landscape of membrane protein complexes in Saccharomyces cerevisiae. Nature, 489, 585.
H.W. Mewes, K. Heumann, A. Kaps et al (2006) Yeast Protein Interactome topology provides framework for coordinated functionality, Nucleic Acids Res., 34, 2812.
H.W. Mewes, K. Heumann, A. Kaps et al (2003) Global analysis of protein localization in budding yeast, Nature, 425, 686.
Liu, Q., Song, R., & Li, J. (2015). Inference of gene interaction networks using conserved subsequential patterns from multiple time course gene expression datasets. BMC Genomics, 16(12), S4. doi: 10.1186/1471-2164-16-S12-S4
Montojo, J., Zuberi, K., Rodriguez, H., Bader, G. D., & Morris, Q. (2014). GeneMANIA: Fast gene network construction and function prediction for Cytoscape. F1000Research, 3, 1–7. doi: 10.12688/f1000research.4572.1
Morris, J. S., Ph, D., Kuchinsky, A., Pico, A., & Institutes, G. (2012). Analysis and Visualization of Biological Networks with Cytoscape. UCSF (Workshop), 65. Retrieved from http://www.cgl.ucsf.edu/Outreach/Workshops/NIH-Oct-2012/Cytoscape/Analysis and Visualization of Biological Networks with Cytoscape v6.pdf
Schreiber, F. (2011). Analysis and visualization of biological networks. Informatik-Spektrum, 32(4), 301–309. doi: 10.1007/s00287-009-0351-8
Suresh Mathivanan, Balamurugan Periaswamy, TKB Gandhi et al (2006) An evaluation of human protein-protein interaction data in the public domain, BCM Bioinformatics, 5, 1186.

MAIN RESULTS

We visualize the biological network with its components, the expression network of the GAL4 and GAL11 gene regulation subnetwork resulting from the application of the different experimental cases analyzed by repressing Gal1, Gal4, Gal80 respectively.

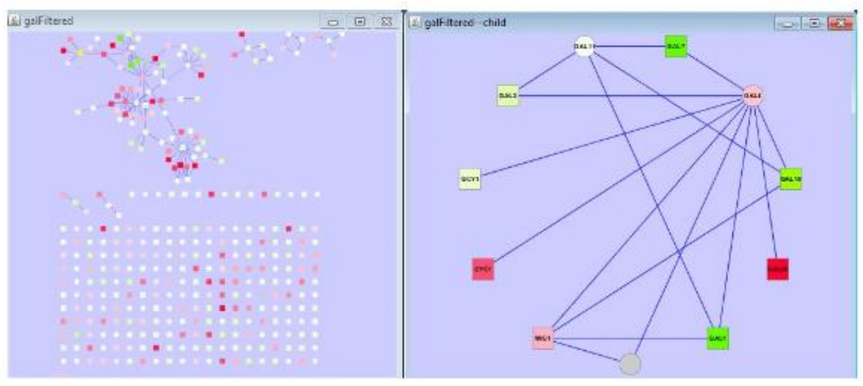


Figure 1. Expression of the GAL4 and GAL11 Gene Regulation Subnetwork result of the application of the different experimental cases analyzed repressing Gal1, Gal4, Gal80 respectively.

We visualize differentially expressed genes by applying the Hierarchical Cluster algorithm as well as the K-means Cluster algorithm, with these results we can verify that the differentially expressed genes coincide.

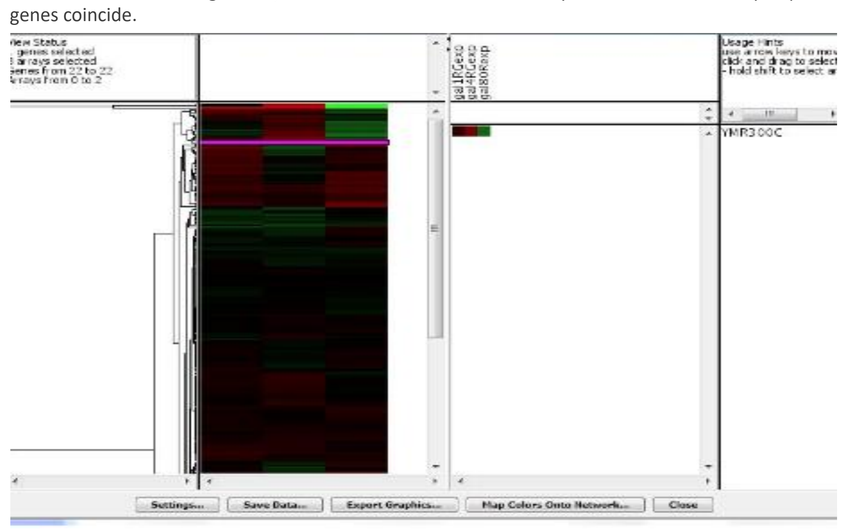


Figure 2. Hierarchical Cluster Application, to view Genes Differentially Expressed.

DISCUSSION

Proteins rarely work in isolation, but rather it is the supramolecular organization of proteins in complexes and machines molecular factors that are crucial to carry out biological function. The detection of protein complexes on a global scale by therefore, it is one of the most important missions in proteomics.
We can verify if there is an existing connection between two nodes thanks to the options offered by the Cytoscape tool, in this case the Clustering Coefficient or transitivity allows us to quantify the probability that adjacent nodes are connected or not, if there is a value of 1 their nodes are connected and if we obtain a value of 0 there is no connection between them.
In Biological Networks, Small World Networks are very frequent, this is a type of graph where most nodes are not neighbors to each other, however, these can be reached from any source node through a relatively short number jumping between them. We have applied the Watts and Strogatz model, the same as with the initial parameters that is defined, creates a network with a very high value for the minimum mean path, however, with some modifications made to it, can be verified that this minimum path takes a value more adjusted to reality without affecting the clustering coefficient.
However biological networks are not scale-free networks, in a graph, the Degree of a node is defined as the number of connections that enter that node. In Free-Scale Networks, there are many little connected nodes, while highly connected nodes (hubs) are very rare. It is said that the distribution of nodes in free of scale networks follows a power law P (k).

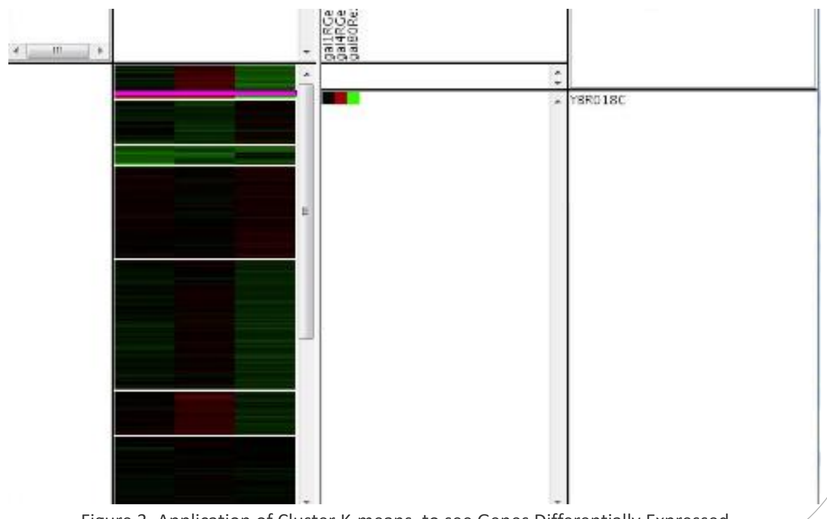


Figure 3. Application of Cluster K-means, to see Genes Differentially Expressed.

Con el aval de:
CACES Consejo de Aseguramiento de la Calidad de la Educación Superior
Ministerio de Turismo
Ministerio de Cultura y Patrimonio
Con el auspicio de:
Secretaría de Educación Superior, Ciencia, Tecnología e Innovación
República del Ecuador
Gobierno del Encuentro
Juntos lo logramos