OPEN ACCESS Freely available online

Journal of Theoretical & Computational Science

Commentary

Correlation Networks using Modular Structure of Network Communities and Gene Co-Expression Data

Herbert Frank*

Department of Networks and Communications, Utah State University, Logan,USA

DESCRIPTION

Large-scale networks like gene expression networks, Protein-Protein Interaction (PPI) networks, and metabolic networks can be used to study biological systems. Understanding the connections between biological components across various phenotypes and during evolution requires comparing these networks (e.g., diseased *vs.* healthy, good prognosis *vs.* bad prognosis, mouse *vs.* human, etc). Studying how these networks are "re-wired" can shed more light than looking at biological entities as separate, unconnected parts. PPI network comparison and analysis are possible using a variety of techniques. For non-model organisms, it is difficult to build a specialized PPI network, which is necessary for drawing conclusions about evolution. However, because gene expression profiles for model and non-model organisms are readily available, creating gene expression networks is a simple operation.

The transcriptome of an organism, which formerly referred to all messenger RNA (mRNA) molecules expressed but now describes the whole range of RNA transcripts expressed by an organism, can be used to infer the interactions between genes. As a result of the intimate relationship between the transcriptome and an organism's phenotype, such as its morphological structure, transcriptomic activity can influence how well an organism works. Comparative transcriptomics has been valuable for tracking gene expression variations that may underlie biological mechanisms of evolution with the development of high-throughput technologies like RNA-seq and single-cell RNA-seq. Studying coordinated gene expression patterns across multiple traits and creatures is made possible by gene expression networks.

Gene-gene interactions are represented by Gene Co-expression Networks (GCNs), which are undirected graphs with nodes for genes and edges for co-expression intensity between nodes. These networks nonetheless enable the simultaneous investigation of several genes and the potential interactions between them, despite the fact that they do not contain information about regulatory direction. To comprehend the coordinated changes in gene-gene interactions, GCNs can be compared across various tissues, cell types, or species. Cross-species GCN comparisons are now performed using a variety of methods, such as functional annotation transfer, inter- and intra-modular hub discovery, and differential co-expression network analysis methods.

By comparing and contrasting the biological interactions across various species, comparative analysis of GCNs can be a useful method for developing hypotheses and gaining understanding of how biological processes have evolved. For instance, co-expression connection alterations are more common in genes with a relatively recent evolutionary history and low connectivity. Homologous genes also tend to be negatively connected with molecular evolution rates. Low connectivity genes those with fewer edges connecting them to other genes in the network also frequently co-express young genes. Depending on how vital they are to functional activities, these young genes may eventually grow more interconnected and even turn into network hubs. The mapping of orthologs between species and the comparison of modules linked to certain functional processes have also been used to compare GCNs. Differential co-expression analysis, which commonly compares samples from sick and healthy individuals, also finds changes in the co-expressed genes between the two circumstances. It can also compare two species.

Co-expression network representation

Biological networks can be graphically depicted in a variety of ways, using various techniques to depict the connections between the nodes. PPI networks frequently have edges without a corresponding weight. Another option is to utilize a weighted graph, where the edge weight can also represent how certain one can be that an edge exists based on the data at hand or through experimentation. Typically, this is expressed as a number between 0 and 1, with 1 being the highest level of confidence and 0 the lowest. The links between proteins can be predicted based on computer analysis of other known biological facts, or they can be direct physical interactions discovered *via* an experimental method. The socio-affinity index, which offers a measurement of the relationship between a pair of

Correspondence to: Dr. Herbert Frank, Department of Networks and Communications, Utah State University, Logan, USA, E-mail: herbertfrank@ads.edu

Received: 28-Oct-2022, Manuscript No. JTCO-22-20376; **Editor assigned:** 02-Nov-2022, Pre QC No. JTCO-22-20376 (PQ); **Reviewed:** 17-Nov-2022, QC No. JTCO-22-20376; **Revised:** 23-Nov-2022, Manuscript No. JTCO-22-20376 (R); **Published:** 30-Nov-2022, DOI: 10.35248/2376-130X. 22.8.154.

Citation: Frank H (2022) Correlation Networks Using Modular Structure of Network Communities and Gene Co-Expression Data. J Theor Comput Sci. 8:154.

proteins based on a full affinity purification-mass spectrometry dataset, is one example of an edge weight value.

The metric is based on the frequency of each protein in the dataset, the ability of a protein to retrieve another when tagged, and the ability of a protein to retrieve two other proteins. It is also possible to use known interactions from primary databases, such as pathway knowledge from sources like KEGG, similarity measurements between protein structures, or gene information like conserved relationships across multiple genomes to infer the possibility of functional relationships between the proteins encoded by related genes.

High-throughput measures like microarray and/or RNA-seq are used to build GCNs. RNA-seq has the additional advantage of not being restricted to only model animals with preexisting genomic resources, which may allow for comparisons of gene expression across a large number of species at once in the context of evolutionary studies. Building transcriptomes without a reliable genetic resource, however, can result in a less precise assembly.