## METHOD

# FunMod: A Cytoscape Plugin for Identifying Functional Modules in Undirected Protein–Protein Networks

CrossMark

**Massimo Natale** *, **Alfredo Benso, Stefano Di Carlo, Elisa Ficarra**

*Department of Control and Computer Engineering, Politecnico di Torino, Torino 10128, Italy*

**Abstract**   The characterization of the interacting behaviors of complex biological systems is a primary objective in protein–protein network analysis and computational biology. In this paper we present FunMod, an innovative Cytoscape version 2.8 plugin that is able to mine undirected protein–protein networks and to infer sub-networks of interacting proteins intimately correlated with relevant biological pathways. This plugin may enable the discovery of new pathways involved in diseases. In order to describe the role of each protein within the relevant biological pathways, FunMod computes and scores three topological features of the identified sub-networks. By integrating the results from biological pathway clustering and topological network analysis, FunMod proved to be useful for the data interpretation and the generation of new hypotheses in two case studies.

## Introduction

Systems biology broadly uses networks to model and discover emerging properties among genes, proteins and other relevant biomolecules. Theoretical studies have indicated that biological networks share many features with other types of networks, as computer or social networks [1]. Therefore, biological network analyses allow the application of mathematical and computational methods of the graph theory to biological studies [2]. The computational analysis of biological networks has therefore become increasingly useful to mine the complex

cellular processes and signaling pathways [3]. Many types of biological networks exist, depending on the information associated with their nodes and edges. In general, biological networks can be classified as directed and undirected networks [4]. In directed networks, the nodes are molecules and edges represent causal biological interactions, such as the transcription and translation regulations [5]. In contrast, in undirected networks, an edge indicates a shared property, such as the sequence similarity [6], gene co-expression [7], protein–protein interaction [8], or the term co-occurrence in the scientific literature [9–11]. In order to extract relevant biological implication from undirected networks, which are also called informative networks [12], it is useful to complement the topological information with the independent biological information retrieved from Gene Ontology (GO) and pathway databases. Often the goal is to identify densely-interconnected areas and

---

* Corresponding author.
  E-mail: massimo.natale@polito.it (Natale M).

correlate them with a specific biological function [13,14]. Several algorithms and bioinformatic tools have been proposed for partitioning the network into structural modules, or for clustering sub-network modules within an informative network [15,16]. Researchers in Bioinformatics developed Cytoscape plugins [17] to mine functional modules in a varieties of network types, such as Clust&See [18], clusterMaker [19], CyClus3D [20], GLay [21] and Enrichment Map [22]. These plugins mainly function on the basis of topological properties. The groups of highly-interconnected nodes may form clusters based on a "first topological clustering" strategy, which aimed at partitioning complex networks into modules. The biological functions are then assigned assuming that members within each sub-network shared a similar biological function [23]. However, clusters are identified solely on the basis of the topology. Therefore, the possibility of co-occurrence events cannot be ruled out using these methods. Moreover, these strategies are heavily influenced by the topological structure of the network itself, and the way that the network is constructed [24].

In practice, the connectivity of the informative networks is established by experimental methods, which can lead to sampling of a subnet of the real biological network [25,26]. Often, biases in the sampling strategies lead to apparent scale-free topologies, which do not reflect the actual complete network topology. As an alternate to the "first topological clustering" methods, some authors used a "first pathway enrichment" strategy [27], which enables analyzing gene networks and extracting functional modules starting from the biological enrichment analysis. Several Cytoscape plugins thus have been implemented owning to this strategy, such as BiNGO [28], ClueGO [29], ClusterViz [30], JEPETTO [31] and Reactome [32].

BiNGO and ClueGO are widely used tools to determine which biological functions are statistically over-represented in a list of genes or a network. These plugins offer the possibility to calculate enrichment by using different statistical algorithms. However, they do not evaluate the connectivity between genes, and do not take into account the possibility that nodes spread in the network could still represent a significant biological function.

A recent focus of bioinformatics has been to develop the computational tools that are able to mine the connectivity of gene networks and uncover the sets of molecules that participate in a common biological function [33]. Reactome is established based on an un-weighted human protein functional interaction network and the functional interaction score was calculated with Pearson correlation coefficients among all gene pairs in the biological database. The weighted network was clustered into a series of gene interaction modules using the Markov clustering algorithm. Each module of the Reactome consists of a set of genes that are both connected in the protein functional interaction network and highly-correlated in biological databases [34]. This approach, however, does not consider the topology or the connectivity of gene interaction modules.

JEPETTO identifies functional associations between genes and pathways using protein interaction networks and topological analyses. Although JEPPETTO combines network analysis with functional enrichment, this tool requires a list of genes in input and the selection of a database of interest from which the reference gene sets will be extracted.

Reactome and JEPETTO are based on an internal gene interaction database, in which the users are not able to filter this pre-defined database or to analyze their own protein informative networks. Nowadays there are a great number of tools able to produce undirected protein networks from a user-defined query, such as protein–protein interaction network (STRING, BioGRID, *etc.*), co-expression network (COXPRESdb) [35] and co-occurrence network (Protein-Quest) [36].

To the best of our knowledge, there is no software available to identify sub-networks of genes that are highly connected and belong to the same biological pathways. Moreover, the aforementioned plugins do not allow the analysis of user-defined undirected networks, such as protein–protein interaction, functional association, gene co-expression and literature co-occurrence.

In this work we developed a new approach using a "first biological assignment" strategy and we implemented our method as a Cytoscape plugin, called FunMod. According to the principle that interacting proteins drive common biological processes, FunMod analyzes informative networks combing topological and biological information to identify and extract sub-network modules in proteins that are involved in the same biological pathway. Moreover, in order to describe the shape of the modules and discriminate the proteins' topological proprieties within the single sub-network, FunMod analyzes sub-network features by using three topological scores. The sub-networks that are statistically overrepresented can act as building blocks of complex informative networks and carry out a specific biological function [24]. Assessment of the sub-network topological proprieties and shapes can consequently be used for the gene ranking in the context of a specific research domain, such as a disease [35].

In the present study, FunMod proves to be a useful method for identifying functional sub-networks in an informative protein network, exploring biomedical information and inferring automated functional hypotheses between an user defined protein–protein network [36]. FunMod is unique at its capability of analyzing user-defined undirected networks, in order to provide more realistic models that incorporate information from certain cellular types, developmental states and/or disease conditions.

## Methods

FunMod analyzes the informative protein network displayed in the Cytoscape Main Network View window. The plugin supports many standard protein dictionaries, the protein nodes can be identified (node ID) by six different dictionaries: Entrez Gene ID, Ensembl, Official Symbol (HGNC symbol), NCBI Reference Sequence, UniGene and Uniprot Entry Name.

FunMod iteratively selects all edges of the network and assigns a functional annotation to an edge when two linked nodes are annotated in the same biological group or pathway in the ConsensusPathDB (DB) database [37]. In other words, FunMod considers an order pair network $G = (V, E)$, where vertices ($V$, nodes) join with edges ($E$), and collects, for each ConsensusPathDB pathway, pairs of linked nodes to model a functional sub-network $G_p = (V_p, E_p)$, where $V_p \subseteq$ ends on the pathways and $E_p \in E$. FunMod performs a global enrichment analysis screening the pathways whose proteins are co-annotated with their neighbors. Accordingly, all the pathways identified by FunMod are also significant in a global enrichment analysis, since the connected nodes are a fraction of nodes in the network. So pathways enriched in a sub-network

are also enriched in the global network, but only a few pathways enriched in the global network are enriched in a cluster. Our algorithm thus could find the pathways enriched in the global network and whose proteins are densely connected within a sub-network.

FunMod extracts all pairs of nodes annotated for the same pathway in a new sub-network. Subsequently, FunMod tests the statistical significance and calculates the topological properties of the sub-network to identify the sub-networks that are statistically enriched in biological functions and that exhibiting interesting topological features. The statistical significance of the sub-network is determined by performing a hypergeometric test, a well-established method used in gene enrichment analyses [27]. The hypergeometric probability ($h$) is based on the following formula:

$$h\,(x; X, n, N) = \frac{\binom{X}{x}\binom{N-X}{n-x}}{\binom{N}{n}} \tag{1}$$

where $x$ is the number of nodes of the sub-network (the items in the sample that are classified as success), $n$ is the number of genes in the network (items in the sample); $X$ is the number of genes annotated in the DB with that pathway (items in the population that are classified as success) and $N$ is the number of all genes annotated in DB (items in the population). FunMod preserves the sub-networks with a $P$ value $< 0.05$.

For a better understanding of the systemic functions and the cooperative interactions between genes within the functional modules, FunMod checks whether the sub-network topology fits into a specific module. Network modules represent patterns occurring significantly more often than random in the complex networks. They consist of sub-graphs of local interconnections between network elements. FunMod calculates a fitting score of each sub-network for three types of modules: clique, star and path [38].

A clique is a sub-network in which all nodes are connected to each other. Cliques are the most widely-used modules for assigning a biological function to a topological sub-network. FunMod calculates the tendency to be a clique by graph density (GD), a score that can also be defined as the local clustering coefficient, using the formula:

$$GD = \frac{2E}{n \times (n-1)} \tag{2}$$

where $E$ is the number of edges in the sub-network and n is the number of genes in the sub-network.

The star module is particularly interesting for identifying drug targets. It is characterized by a central gene with a high degree connection to a set of first-degree neighbors, which are loosely connected between each other. In a star sub-network, the central gene (the hub gene) has influence on its neighbor genes and possibly on the whole network. To identify a star module, FunMod calculates the sub-network centralization (CE) using the formula:

$$CE = \frac{n}{(n-2)} \times \left[\frac{\max(k)}{n-1} - GD\right] \tag{3}$$

where $\max(k)$ is the highest degree of central gene connection in the sub-network and $n$ is the number of genes in the sub-network.

The path module corresponds to a real pathway where the genes contribute to a signal transduction. The path score is evaluated by the sub-network diameter ($D$), which uses the maximum length of all shortest paths between any two connected nodes, using the formula:

$$D = \max_{i,j} \delta_{\min}(i,j) \tag{4}$$

where $\delta_{\min}$ is the minimum path between two nodes $i$ and $j$ of the network.

Pathways identified using FunMod were displayed in the Cytoscape Results Panel and ranked based on their $P$ values. For each pathway, FunMod displays its clique, star and path coefficients. By clicking the "Pathway" button, FunMod selects the corresponding nodes in the network. And by using the "View subnet" function, it creates a new network containing only those genes and edges annotated within that pathway. Moreover, FunMod enables saving the results in a tab-delimited file. FunMod plugin, user's guide, screenshot and demo networks can be freely downloaded from the SourceForge project page at: http://sourceforge.net/projects/funmodnetwork/. Developed in Java, FunMod is a platform independent plugin for Cytoscape 2.8.4, which is freely available without charge for non-commercial purposes.

## Results and discussion

Gene Ontology (GO) provides information on the location of genes and gene products in a cell or the extracellular environment and also on the molecular function they carry out. However, GO does not provide information about the interaction of proteins in the same biological context. For example, GO does not allow us to describe genes in terms of which cells or tissues they are expressed in, which developmental stages they are expressed at, or their involvement in disease (http://www.geneontology.org/GO.doc.shtml). We thus chose ConsensusPathDB [39] to identify proteins that are strictly involved in the same pathways. We also assess the topological shape of each sub-network in order to reveal evidence of its biological function and the function of its components. Three topological scores are calculated to describe the global features of the sub-network: graph density, network centralization and shortest path. Other topological scores, such as centrality and degree, describe the relative importance of a single node within the sub-network and thus left out in FunMod.

To demonstrate the usage and performance of FunMod, we showed here case studies to analyze two different informative networks: a bibliometric network of proteins related to the Budd–Chiari syndrome and a co-expression network of proteins related to spinal muscular atrophy. Identification of the sub-network functional modules in the undirected protein networks allows the reduction in network complexity, clustering of proteins on the basis of common biological functions and discovery of the mechanisms underlying a disease [40].

### Case study 1: Analysis of a bibliometric network

Budd–Chiari syndrome is an uncommon condition characterized by obstruction of the hepatic venous outflow tract [41]. Budd–Chiari syndrome is difficult for diagnosis and management and often fatal if not treated optimally, since none of the current medical therapies are based on the clinical evidence due to the difficulty in performing appropriate clinical studies

on a sufficient number of patients with this rare disease [42]. In this circumstance, a meta-analysis can be useful to provide valuable information for researchers to understand the mechanism of this disease [43].

In this case study, we used ProteinQuest (PQ), an advanced text-mining tool that resorts to the web services offered by PubMed to identify all proteins co-cited in the same abstract or figure legend [44] and to export the literature co-occurrence information as a network file that can be managed in Cytoscape. We submitted a query to PQ using the starting search terms ''Budd–Chiari syndrome'' and its NCBI medical subject heading (MeSH) term alias. PQ retrieved more than 2200 papers. The network obtained by PQ was an undirected protein–protein network consisting of 76 nodes that represent 76 proteins cited in at least one document and 174 edges indicating co-occurrence of two proteins in at least one document. A screenshot of the plugin is shown in **Figure 1**, whereas a quick guide of the Budd–Chiari syndrome network can be downloaded at the SourceForge project page (http://source-forge.net/projects/funmodnetwork/).

Using FunMod we identified 33 different biological pathways that are significantly enriched in the Budd–Chiari syndrome network. **Table 1** shows 13 most relevant pathways. The most significant one is related to the coagulation cascades, which is in agreement with the previous reports that Budd–Chiari is a disorder frequently characterized by the thrombotic

obstruction of hepatic venous outflow [45,46]. This result indicated that FunMod, relying on the ABC principle (A and C have no direct connection but are connected via shared B intermediates) [47,48] to establish new relationships between proteins from literature co-occurrence, can be effectively used to combine information from the relevant scientific literature and from pathway databases, supporting the discovery of new knowledge. Moreover, we calculated the three types of topological coefficients of each pathway sub-network with FunMod and selected the most relevant pathways using three types of coefficients. As shown in **Figure 2**, fibrinolysis pathway was exemplified as a clique module (Figure 2A), the JAK-STAT pathway as an example of a star module (Figure 2B), and the platelet activation, signaling and aggregation as an example of path module (Figure 2C).

By examining the topological feature of a sub-network functional module, we can extract information of pathways functioning in Budd–Chiari syndrome. The cliques represent a highly connected sub-network, where the proteins of the pathway are highly co-studied and co-cited. The fibrinolysis pathway is the central pathway in Budd–Chiari syndrome, consistent with the fact that prothrombotic tendency is caused by abnormalities in the coagulation of fibrinolysis pathways [49]. These data may explain why the fibrinolysis pathway functional sub-network is represented by a clique. A star module represents a functional sub-network where a protein plays
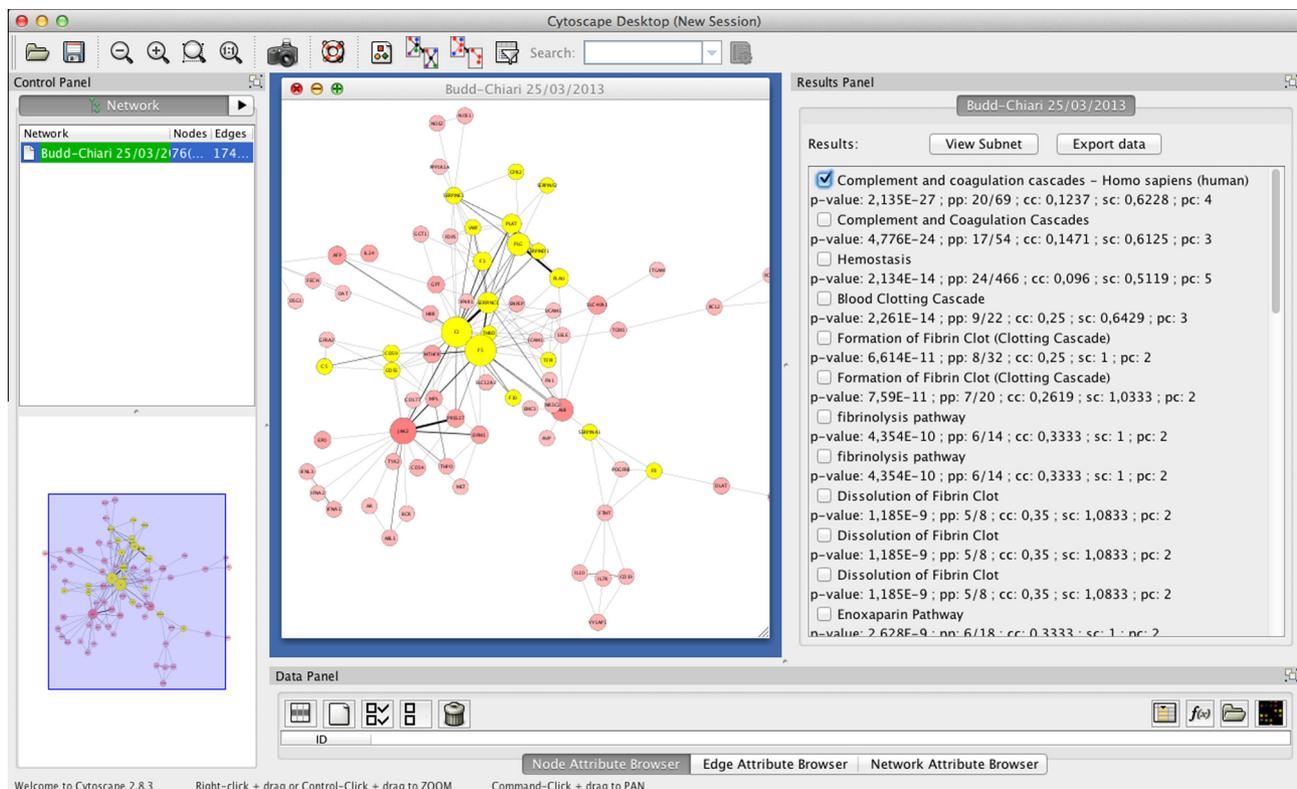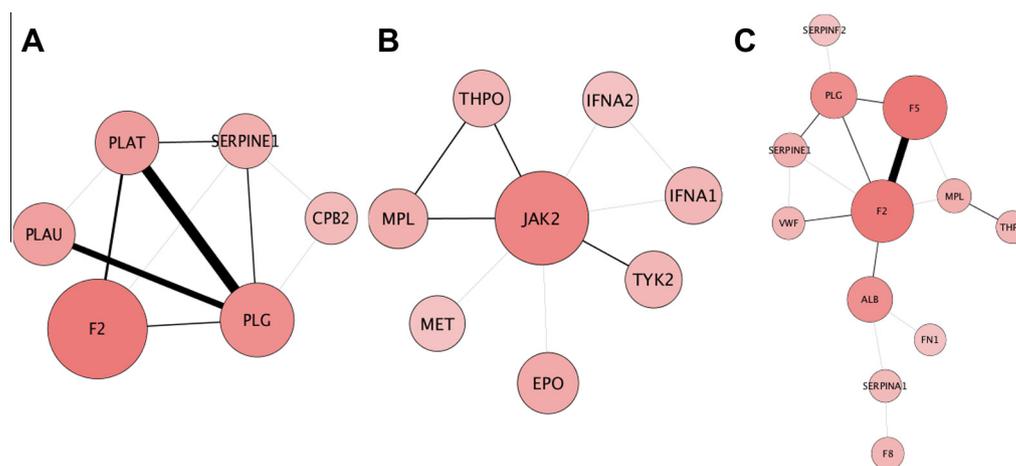


**Figure 1    Screenshot of FunMod plugin in action**
Screenshot shows the bibliometric network of proteins related to Budd–Chiari syndrome. Proteins are displayed as pink nodes and the size and color density are proportional to the occurrence of the protein term in the literature of Budd–Chiari syndrome. The line between nodes represents the co-occurrence, and its thickness is proportional to the number of times that the proteins are co-cited in the same abstract or figure caption. Cytoscape ''Results Panel'' displays the results of FunMod analysis. When a particular pathway is selected, all the proteins on this pathway are selected in the network as yellow nodes.

**Table 1    The most relevant biological pathways identified in the Budd−Chiari syndrome network**

| Pathway | Pathway coverage | P value | Clique coefficient | Star coefficient | Path coefficient |
|---|---|---|---|---|---|
| Complement and coagulation cascades | 20/69 | 2.13E−27 | 0.1237 | 0.6228 | 4 |
| Homeostasis | 24/466 | 2.13E−14 | 0.096 | 0.5119 | 5 |
| Fibrinolysis pathway | 6/14 | 4.35E−10 | 0.3333 | 1 | 2 |
| Heparin pathway | 6/18 | 2.63E−9 | 0.3333 | 1 | 2 |
| Vitamin $B_{12}$ metabolism | 8/51 | 3.57E−9 | 0.1786 | 1.0952 | 2 |
| Platelet degranulation | 9/83 | 1.03E−8 | 0.0972 | 0.3571 | 3 |
| Response to elevated platelet cytosolic $Ca^{2+}$ | 9/88 | 1.75E−8 | 0.0972 | 0.3571 | 3 |
| Folate metabolism | 8/66 | 2.94E−8 | 0.1786 | 1.0952 | 2 |
| JAK/STAT molecular variation 1 | 9/99 | 4.95E−8 | 0.125 | 0.8036 | 2 |
| Selenium pathway | 8/77 | 1.01E−7 | 0.1786 | 1.0952 | 2 |
| Lepirudin pathway | 5/17 | 1.24E−7 | 0.3 | 1.1667 | 2 |
| Dicumarol pathway | 5/18 | 1.71E−7 | 0.3 | 1.1667 | 2 |
| PI3K−Akt signaling pathway | 7/334 | 1.29E−2 | 0.1429 | 0.7333 | 2 |



**Figure 2    Three sub-networks extracted from Budd–Chiari network**
The three sub-network functional modules selected according to the highest coefficient for each type of topological module—cluster module (**A**), star module (**B**) and pathway module (**C**). The three sub-networks were extracted from the Budd–Chiari network by clicking on the "View Subnet" button.

a central role in the pathway. Indeed JAK2 is the hub (the central node) in the JAK-STAT pathway sub-network. Moreover, a mutation in JAK2 tyrosine kinase (JAK2 V617F) is frequently found in patients with Budd–Chiari syndrome [50,51]. Finally, the path module represents a significant pathway in the context of the disease, which is not studied comprehensively. The platelet activation, signaling and aggregation pathway is a therapeutic target for the treatment of Budd–Chiari syndrome [52]. von Willebrand factor (vWF) [53] and SerpinA1 [54], two peripheral sub-network proteins that were recently described as markers of abnormalities of coagulation play roles in provoking thrombosis [55]. The information extracted by FunMod will provide guidance for the future study on Budd–Chiari syndrome.

## Case study 2: Analysis of a co-expression network

Spinal muscular atrophy (SMA) is an inheritable neuromuscular disorder, leading to degeneration of α-motor neurons, due to recessive mutations of the *SMN1* gene [56]. No treatment is currently available for curing SMA patients, and the development of a computational approach could facilitate a better

understanding of this disease and identification of potential therapeutic targets. In the second case study, we applied FunMod to analyze a co-expression network obtained using PQ and COXPRESdb. The protein–protein interactions within this co-expression network reveal proteins examined in the context of SMA and that are simultaneously documented to be co-regulated at mRNA level. The analysis of gene co-expression network provides powerful searching ability that allowed us to identify more than 400 functional clusters related to SMA.

In this case study, we examined all proteins involved in SMA using PQ and found that 1482 proteins were cited in PubMed articles describing SMA. This list of proteins was submitted to COXPRESdb [35], a publicly-available co-expression database, which provides a tool to export the networks in Cytoscape. We used a co-expression network to enhance the predictive power of the informative network. We submitted the list of SMA proteins to COXPRESdb and obtained a co-expression protein network with 2022 nodes and 9166 edges (Figure 3). Further analyzing the SMA co-expression network using FunMod, we identified 429 functional modules and the 13 most relevant modules were shown in Table 2.
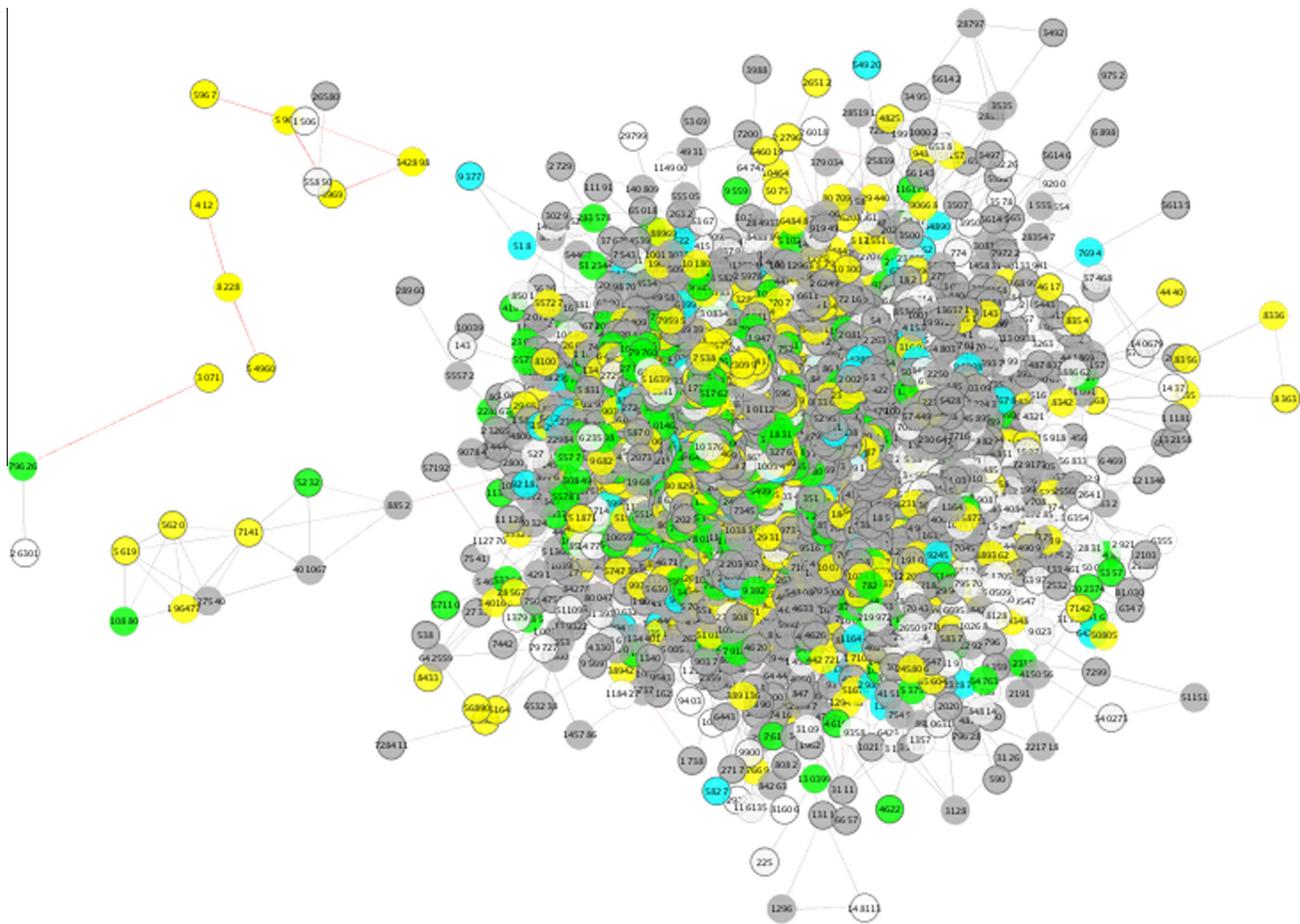
**Figure 3    The SMA co-expression network**
COXPRESdb provides NetworkDrawer, a tool for drawing and visualizing the protein co-expression network in Cytoscape Web. The network can be downloaded in the XGMML format, the standard format for saving and importing network into Cytoscape, and imported into a stand-alone Cytoscape distribution for advanced analyses. NetworkDrawer performs cluster detection and Gene Ontology enrichment analysis, providing a node color scale to indicate protein localization: cytosol (green); cytoskeleton, endoplasmic reticulum, Golgi body, lysosome, peroxisome and plasma membrane (gray); mitochondrion (cyan); nucleus (yellow); and extracellular matrix (white). The edge represents the type of interaction between two nodes: gene co-expression (solid in black) and indicates known protein–protein interaction (dotted in red). SMA, spinal muscular atrophy; XGMML, extensible graph markup and modeling language.

**Table 2    The most relevant biological pathways identified in spinal muscular atrophy network**

| Pathway | Pathway coverage | $P$ value | Clique coefficient | Star coefficient | Path coefficient |
|---|---|---|---|---|---|
| AGE−RAGE pathway | 39/66 | 3.75E−12 | 0.0789 | 0.3051 | 6 |
| BDNF signaling pathway | 61/140 | 1.93E−10 | 0.0484 | 0.2602 | 6 |
| Prolactin signaling pathway | 40/76 | 2.74E−10 | 0.0814 | 0.2922 | 5 |
| Inhibition of cellular proliferation by Gleevec | 17/22 | 1.25E−08 | 0.136 | 0.3417 | 4 |
| Leptin signaling pathway | 32/61 | 1.86E−08 | 0.1008 | 0.3742 | 5 |
| Androgen receptor | 59/148 | 2.22E−08 | 0.0649 | 0.6289 | 4 |
| ATF-2 transcription factor network | 31/61 | 8.13E−08 | 0.0753 | 0.3471 | 5 |
| Glucocorticoid receptor regulatory network | 37/80 | 1.11E−07 | 0.0691 | 0.3675 | 4 |
| MAPK signaling pathway | 61/161 | 1.16E−07 | 0.0522 | 0.2562 | 5 |
| Chronic myeloid leukemia | 34/73 | 2.94E−07 | 0.066 | 0.3163 | 5 |
| mRNA processing | 50/126 | 3.07E−07 | 0.0861 | 0.3142 | 5 |
| Striated muscle contraction | 22/38 | 3.26E−07 | 0.0736 | 0.3905 | 6 |
| ErbB2/ErbB3 signaling events | 23/42 | 6.89E−07 | 0.1206 | 0.3658 | 4 |

FunMod identified a few pathways known to be related to systemic aspects of neurodegenerative conditions, such as the AGE-RAGE pathway [57], striated muscle contraction [58]

and leptin signaling pathway [59]. FunMod identified the brain-derived neurotropic factor (BDNF) signaling among the most relevant pathway (**Figure 4A**). The BDNF promotes
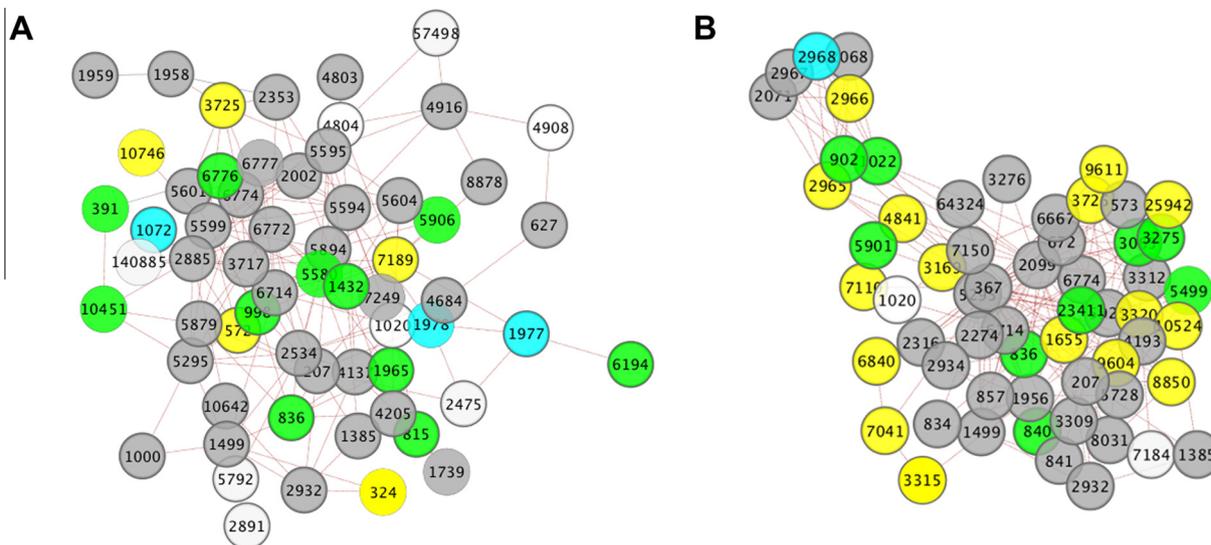
**Figure 4    Two sub-networks identified within the SMA co-expression network**
**A**. The BDNF signaling pathway. **B**. The androgen receptor pathway. The Entrez Gene IDs were used for the node label as required by COXPRESdb for input. The node color scale is inherited from COXPRESdb. Larger nodes are the query genes and smaller gray nodes are additional nodes with one or more edges to at least one query node. Solid lines and red dotted lines indicate gene co-expression and protein–protein interactions from the Human Protein Reference Database (HPRD) and IntAct databases, respectively.

neurogenesis, neuronal survival and synaptic plasticity [60]. Therefore, this finding validates the ability of our software to reveal new relevant functional modules. Interestingly, FunMod identified the androgen receptor signaling pathway in SMA co-expression network as well (**Figure 4**B). Androgens interact with BDNF during development to regulate non-pathological death of motor neurons. Androgens regulate BDNF levels in the target musculature and moreover, androgenic actions at the muscle regulate BDNF levels in motoneurons [61]. Therefore, these interactions have important implications for the maintenance of motoneuron morphology. These findings provide further insights into the pathogenic role of BDNF and androgens in SMA, which might be a preferential route for targeting motor neurons [62].

## Conclusion

Using a "first topological assignment" strategy to identify sub-network functional modules, such as stars and cliques, can be tricky because informative networks are known to have a huge number of edges that are not always pertinent to biological functions. In this work we presented FunMod, a new Cytoscape 2.8 plugin, which can analyze undirected protein networks, such as co-occurrence and co-expression networks, and guide the discovery of sub-network functional modules. A functional module can be considered as a distinct group of interacting proteins [32] within a pathway relevant to a condition of interest.

FunMod identifies within an informative network, pairs of nodes belonging to the same biological pathways and assesses their statistical significance. It then analyzes the topology of the identified sub-network to infer the topological relations (motifs) of its nodes. In this work, the network topology is influenced by the biomedical knowledge since the link between two proteins was established when two gene symbols appear in the same MEDLINE record. The study

of the connection between biomedical concepts by using co-occurrence network extracted from MEDLINE proved capable of guiding the discovery of novel knowledge from scientific literature [63]. This sub-network profiling combined with information from the biological database will help us to better understand the biological significance of the protein–protein network.

FunMod was tested using the co-occurrence network of proteins cited in Budd–Chiari syndrome papers, identifying 33 different biological pathways that are significantly enriched; and using the co-expression network of proteins discussed in publications on SMA. FunMod proves to be a useful tool for a better understanding of the cooperative interactions between proteins and discriminating the biological role played by each protein within a functional module.

## Authors' contributions

AB and SDC conceived the idea and supervised the study. MN designed and implemented the network analysis method, and performed data analysis. AB and SDC analyzed the results and drafted the manuscript. EF edited the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors have declared that no competing interests exist.

## Acknowledgements

## References

[1] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13.

[2] Huber W, Carey VJ, Long L, Falcon S, Gentleman R. Graphs in molecular biology. BMC Bioinformatics 2007;8:S8.

[3] Spirin V, Mirny LA. Protein complexes and functional modules in molecular networks. Proc Natl Acad Sci U S A 2003;100:12123–8.

[4] Pieroni E, de la Fuente van Bentem S, Mancosu G, Capobianco E, Hirt H, de la Fuente A. Protein networking: insights into global functional organization of proteomes. Proteomics 2008;8:799–816.

[5] Li J, Hua X, Haubrock M, Wang J, Wingender E. The architecture of the gene regulatory networks of different tissues. Bioinformatics 2012;28:i509–14.

[6] Kuchaiev O, Przulj N. Integrative network alignment reveals large regions of global network similarity in yeast and human. Bioinformatics 2011;27:1390–6.

[7] Prifti E, Zucker JD, Clément K, Henegar C. Interactional and functional centrality in transcriptional coexpression networks. Bioinformatics 2010;26:3083–9.

[8] Chen L, Xuan J, Riggins RB, Wang Y, Clarke R. Identifying protein interaction sub-networks by a bagging Markov random field-based method. Nucleic Acids Res 2013;41:e42.

[9] Gatti S, Leo C, Gallo S, Sala V, Bucci E, Natale M, et al. Gene expression profiling of HGF/Met activation in neonatal mouse heart. Transgenic Res 2012;3:579–93.

[10] Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY. Rational association of genes with traits using a genome-scale gene network for *Arabidopsis thaliana*. Nat Biotechnol 2010;28:149–56.

[11] Gabow AP, Leach SM, Baumgartner WA, Hunter LE, Goldberg DS. Improving protein function prediction methods with integrated literature data. BMC Bioinformatics 2008;9:198.

[12] Lysenko A, Defoin-Platel M, Hassani-Pak K, Taubert J, Hodgman C, Rawlings CJ, et al. Assessing the functional coherence of modules found in multiple-evidence networks from *Arabidopsis*. BMC Bioinformatics 2011;12:203.

[13] Bu D, Zhao Y, Cai L, Xue H, Zhu X, Lu H, et al. Topological structure analysis of the protein–protein interaction network in budding yeast. Nucleic Acids Res 2003;31:2443–50.

[14] Weatheritt RJ, Jehl P, Dinkel H, Gibson TJ. IELM-a web server to explore short linear motif-mediated interactions. Nucleic Acids Res 2012;40:W364–9.

[15] Thomas S, Bonchev D. A survey of current software for network analysis in molecular biology. Hum Genomics 2010;4:353–60.

[16] Shen R, Goonesekere NC, Guda C. Mining functional subgraphs from cancer protein–protein interaction networks. BMC Syst Biol 2012;6:S2.

[17] Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011;27:431–2.

[18] Spinelli L, Gambette P, Chapple CE, Robisson B, Baudot A, Garreta H, et al. Clust&See: a Cytoscape plugin for the identification, visualization and manipulation of network clusters. Biosystems 2013;113:91–5.

[19] Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, et al. ClusterMaker: a multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics 2011;12:436.

[20] Audenaert P, Van Parys T, Brondel F, Pickavet M, Demeester P, Van de Peer Y, et al. CyClus3D: a Cytoscape plugin for clustering network motifs in integrated networks. Bioinformatics 2011;27:1587–8.

[21] Su G, Kuchinsky A, Morris JH, States DJ, Meng F. GLay: community structure analysis of biological networks. Bioinformatics 2010;26:3135–7.

[22] Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment Map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One 2010;5:e13984.

[23] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science 2002;298:824–7.

[24] Dong J, Horvath S. Understanding network concepts in modules. BMC Syst Biol 2007;4:1–24.

[25] Aittokallio T, Schwikowski B. Graph-based methods for analysing networks in cell biology. Brief Bioinform 2006;7:243–55.

[26] Heatha AP, Kavrakia LE. Computational challenges in systems biology. Comput Sci Rev 2009;3:117.

[27] Griswold AJ, Ma D, Cukier HN, Nations LD, Schmidt MA, Chung RH, et al. Evaluation of copy number variations reveals novel candidate genes in autism spectrum disorder-associated pathways. Hum Mol Genet 2012;21:3513–23.

[28] Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 2005;21:3448–9.

[29] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091–3.

[30] Cai J, Chen G, Wang J. ClusterViz: a Cytoscape plugin for graph clustering and visualization. Changsha, Hunan: School of Information Science and Engineering, < http://chianti.ucsd.edu/cyto_web/plugins/displayplugininfo.php?name = ClusterViz > ; 2010 [accessed 22.03.10].

[31] Glaab E, Baudot A, Krasnogor N, Schneider R, Valencia A. EnrichNet: network-based gene set enrichment analysis. Bioinformatics 2012;28:i451–7.

[32] Jupe S, Akkerman JW, Soranzo N, Ouwehand WH. Reactome – A curated knowledgebase of biological pathways: megakaryocytes and platelets. J Thromb Haemost 2012;10:2399–402.

[33] Mitra K, Carvunis AR, Ramesh SK, Ideker T. Integrative approaches for finding modular structure in biological networks. Nat Rev Genet 2013;14:719–32.

[34] Wu G, Stein L. A network module-based method for identifying cancer prognostic signatures. Genome Biol 2012;13:R112.

[35] Obayashi T, Okamura Y, Ito S, Tadaka S, Motoike IN, Kinoshita K. COXPRESdb: a database of comparative gene coexpression networks of eleven species for mammals. Nucleic Acids Res 2013;41:D1014–20.

[36] Benso A, Cornale P, Di Carlo S, Politano G, Savino A. Reducing the complexity of complex gene coexpression networks by coupling multiweighted labeling with topological analysis. Biomed Res Int 2013;2013:676328.

[37] Liekens AM, De Knijf J, Daelemans W, Goethals B, De Rijk P, Del-Favero J. BioGraph: unsupervised biomedical knowledge discovery via automated hypothesis generation. Genome Biol 2011;12:R57.

[38] Dolinski K, Chatr-Aryamontri A, Tyers M. Systematic curation of protein and genetic interaction data for computable biology. BMC Biol 2013;11:43.

[39] Kamburov A, Stelzl U, Lehrach H, Herwig R. The Consensus-PathDB interaction database: 2013 update. Nucleic Acids Res 2013;41:D793–800.

[40] Royer L, Reimann M, Andreopoulos B, Schroeder M. Unraveling protein networks with power graph analysis. PLoS Comput Biol 2008;4:e1000108.

[41] MacNicholas R, Olliff S, Elias E, Tripathi D. An update on the diagnosis and management of Budd–Chiari syndrome. Expert Rev Gastroenterol Hepatol 2012;6:731–44.

[42] Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths and limitations. Cleve Clin J Med 2008;75:431–9.

[43] Jaqua NT, Stratton A, Yaccobe L, Tahir U, Kenny P, Kerns T. A review of the literature on three extraintestinal complications of ulcerative colitis: an ulcerative colitis flare complicated by Budd–Chiari syndrome, cerebral venous thrombosis and idiopathic thrombocytopenia. Acta Gastroenterol 2013;76:311–6.

[44] Giordano M, Natale M, Cornaz M, Ruffino A, Bonino D, Bucci EM. IMole, a web based image retrieval system from biomedical literature. Electrophoresis 2013;34:1965–8.

[45] Wang R, Meng Q, Qu L, Wu X, Sun N, Jin X. Treatment of Budd–Chiari syndrome with inferior vena cava thrombosis. Exp Ther Med 2013;5:1254–8.

[46] Akbulut S, Yilmaz M, Kahraman A, Yilmaz S. Budd–Chiari syndrome due to giant hydatid cyst: a case report and brief literature review. J Infect Dev Ctries 2013;7:489–93.

[47] Frijters R, van Vugt M, Smeets R, van Schaik R, de Vlieg J, Alkema W. Literature mining for the discovery of hidden connections between drugs, genes and diseases. PLoS Comput Biol 2010;6:e1000943.

[48] Li C, Liakata M, Rebholz-Schuhmann D. Biological network extraction from scientific literature: state of the art and challenges. Brief Bioinform 2013;23:e237–44.

[49] Smalberg JH, Kruip MJ, Janssen HL, Rijken DC, Leebeek FW, de Maat MP. Hypercoagulability and hypofibrinolysis and risk of deep vein thrombosis and splanchnic vein thrombosis: similarities and differences. Arterioscler Thromb Vasc Biol 2011;31:485–93.

[50] Qi X, Yang Z, Bai M, Shi X, Han G, Fan D. Meta-analysis: the significance of screening for JAK2V617F mutation in Budd–Chiari syndrome and portal venous system thrombosis. Aliment Pharmacol Ther 2011;33:1087–103.

[51] Patel RK, Lea NC, Heneghan MA, Westwood NB, Milojkovic D, Thanigaikumar M, et al. Prevalence of the activating JAK2 tyrosine kinase mutation V617F in the Budd–Chiari syndrome. Gastroenterology 2006;130:2031–8.

[52] Chaudhuri M, Jayaranganath M, Chandra VS. Percutaneous recanalization of an occluded hepatic vein in a difficult subset of pediatric Budd–Chiari syndrome. Pediatr Cardiol 2012;33:806–10.

[53] Alkim H, Ayaz S, Sasmaz N, Oguz P, Sahin B. Hemostatic abnormalities in cirrhosis and tumor-related portal vein thrombosis. Clin Appl Thromb Hemost 2012;18:409–15.

[54] Hourigan SK, Anders RA, Mitchell SE, Schwarz KB, Lau H, Karnsakul W. Chronic diarrhea, ascites, and protein-losing enteropathy in an infant with hepatic venous outflow obstruction after liver transplantation. Pediatr Transplant 2012;16:E328–31.

[55] Ding PX, Li Z, Han XW, Zhang WG, Zhou PL, Wang ZG. Spontaneous intrahepatic portosystemic shunt in Budd–Chiari syndrome. Ann Vasc Surg 2014;28:742.e1–4.

[56] Van Meerbeke JP, Sumner CJ. Progress and promise: the current status of spinal muscular atrophy therapeutics. Discov Med 2011;12:291–305.

[57] Li J, Liu D, Sun L, Lu Y, Zhang Z. Advanced glycation end products and neurodegenerative diseases: mechanisms and perspective. J Neurol Sci 2012;317:1–5.

[58] Katsuno M, Adachi H, Tanaka F, Sobue G. Spinal and bulbar muscular atrophy: ligand-dependent pathogenesis and therapeutic perspectives. J Mol Med 2004;82:298–307.

[59] Paz-Filho G, Wong ML, Licinio J. The procognitive effects of leptin in the brain and their clinical implications. Int J Clin Pract 2010;64:1808–12.

[60] Teixeira AL, Barbosa IG, Diniz BS, Kummer A. Circulating levels of brain-derived neurotrophic factor: correlation with mood, cognition and motor function. Biomark Med 2010;4:871–87.

[61] Verhovshek T, Rudolph LM, Sengelaub DR. Brain-derived neurotrophic factor and androgen interactions in spinal neuromuscular systems. Neuroscience 2013;239:103–14.

[62] Dupuis L, Echaniz-Laguna A. Skeletal muscle in motor neuron diseases: therapeutic target and delivery route for potential treatments. Curr Drug Targets 2010;11:1250–61.

[63] Narayanasamy V, Mukhopadhyay S, Palakal M, Potter DA. TransMiner: mining transitive associations among biological objects from text. J Biomed Sci 2004;11:864–73.