Cellular and Molecular Biology

Saxena *et al. Cell. Mol. Biol.*2016, 62 (6): 17-21
ISSN: 1165-158X
doi: 10.14715/cmb/2016.62.6.3

Original Research

Mapping of protein-protein interaction network of Alexander disease

A. K. Saxena[1], V. L. Saxena[2], S. Dixit[2*]

[1] Department of Zoology, D.A.V. College, Kanpur (U.P.), India
[2] Bioinformatics Infrastructure Facility, Centre of DBT (Govt. of India), D.G. (P.G.) College, Kanpur (U.P), India

Abstract: Alexander disease (ALXD) is slowly progressive neurodegenerative disorder which affects white matter of the central nervous system. The main cause of disorder is mutation in GFAP gene and mutation in some other genes were also reported. This study was aimed at getting a better insight into ALXD pathogenesis and identifying the important functional and highly interconnected nodes in human protein interaction network, identifying the important sub-networks in the system could be helpful in understanding the underlying molecular mechanism. The topological analysis of human protein interaction network strategy to identify highly interconnected sub-network modules from which six proteins are found i.e. GFAP, PLEC, CRYAB, NDUFV1, CASP3 and MAPK14 plays important role in disease. Further, the enrichment analysis of interaction network identifies crucial pathways in which most of the diseased proteins overlaps. Through system biology approach, the undirected human protein interaction network of ALXD is buildup with the help of Cytoscape tool and its various plugins helps to investigate network further. The systematic approach suggests the finding of previously known proteins, GFAP, PLEC, CRYAB, NDUFV1, CASP3 and MAPK14 can be used as a drug targets and potential treatment discovered also enrichment analysis will provide guidance for the future study on Alexander disease.

Key words: Alexander disease, Neurodegenerative disorder, Topological analysis, Enrichment analysis, Protein interaction network.

Introduction

Alexander disease is an extremely rare, usually progressive and fatal, neurological disorder in which progressive degeneration of the myelin sheath of the white matter accompanied by the accumulation of abnormal deposition of protein in astrocytes take place; so that it is more appropriate to consider ALXD is a disease of astrocytes than a white matter. It is divided into four forms based on the age at onset of disease i.e. neonatal, infantile, juvenile, or adult (1).

GFAP gene plays major role in pathogenesis of ALXD which encodes glial fibrillary acidic protein, the main intermediate filament protein expressed in mature astrocytes (2). It was reported that this gene mutation does not affect protein synthesis but results defective protein (3). Studies also reveal that there are more proteins other than GFAP showed significant change in Alexander disease.

It has been estimated that over 80% of proteins do not operate alone but in complexes. These protein–protein interactions are regulated by several mechanisms. Many protein–protein interactions are part of larger cellular networks of protein– protein interactions. It is believed that the cellular network of protein–protein interactions are built up by highly connected protein nodes and many poorly connected nodes. Each node receives inputs and generates one or more specific outputs in a manner similar to computational units (4).

A class of computational approaches has recently been proposed that exploit these two sources of information, physical interaction networks and linkage intervals, to predict associations between protein and diseases. The study of protein co-expression network is essential to define the molecular networks that contribute to maintain the homeostasis of an organism's body functions. Disruptions in gene or protein interaction networks results in disease in both human and Animals. In effort to understand the underlying mechanism of Alexander disease, Cytoscape and its plugins (21) were used.

Study referred from published literature and it had been found that around 23 genes were causing ALXD and built a protein interaction network using Cytoscape. Various plugins were used to analyze protein association network. This study was aimed at getting a better insight into pathogenesis of disease and identifying the important functional and highly interconnected nodes in the network, as identifying the important sub-networks in the system could be helpful in understanding the underlying molecular mechanism. The objective of this study was to build the protein interaction network by using Cytoscape, analyze statistical significance of network and identification of functional proteins through its various plugins.

Cytoscape is a project dedicated to building open-source network visualization and analysis software. Software "Core" provides basic functionality to layout and query the network and to visually integrate the network with state data. The Core is extensible through a plug-in architecture, allowing rapid development of additional computational analyses and features (5).

Network-Analyzer is a Java plugin for Cytoscape, a software platform for the analysis and visualization of molecular interaction networks. It performs analysis of biological networks and calculates network topo-

Received November 7, 2015; **Accepted** April 30, 2016; **Published** May 30, 2016

*** Corresponding author:** Supriya Dixit, Bioinformatics Infrastructure Facility, Centre of DBT (Govt. of India), D.G. (P.G.) College, Kanpur (U.P), India. Email: supriya.dixit.28@gmail.com

logy parameters including the diameter of a network, the average number of neighbors, and the number of connected pairs of nodes. It also computes the distributions of more complex network parameters such as node degrees, average clustering coefficients, topological coefficients, and shortest path lengths.

Clustering of biological networks, such as protein-protein interaction network and metabolic network, is one of the most common approaches for identifying functional modules and protein complex, predicting protein/gene functions. Based on MCODE (21), a Cytoscape plug-in named ClusterViz (7) was developed, which features with various clustering algorithms, intuitive visualization of clustering results, easy-to-use topological and Gene Ontology enrichment analysis (6).

JEPETTO (Java Enrichment of Pathways Extended To TOpology) Cytoscape plugin performing integrative human gene set analysis (7). It identifies functional associations between genes or proteins and known cellular pathways and processes using protein interaction networks and topological analysis.

Materials and Methods

Data source

In an effort to understand the interaction of various proteins and protein pathways which are involved or getting affected in Alexander disease, the protein-protein interaction network map was built and analyzed through Cytoscape. Proteins responsible for ALXD were searched from the GeneCards database and published literatures also reports on the Alexander disease, were also referred (Table 1).

Table 1. Protein name with their respective GC id given in GeneCards Database.

Protein Name	GC id
GFAP	GC17M042982
RPS27A	GC02P055459
HSPB1	GC07P075931
CRYAB	GC11M111813
MAPK14	GC06P035995
HTT	GC04P003076
CASP3	GC04M185548
BLZF1	GC01P169337
MTOR	GC01M011166
NDUFV1	GC11P067374
DES	GC02P220283
MAP3K10	GC19P040697
PLEC	GC08M144989
HSH2D	GC19P016244
BFSP2	GC03P133118
SYNM	GC15P099645
IL2RA	GC10M006041
DSC3	GC18M028593
HDAC6	GC0XP048659
ABCB1	GC07MO87132
R79L	12581808 (PMID)
R239C	12581808
R79C	12581808

#GC id is unique id which is given to each gene. It starts from GC stand for GeneCards then followed by chromosome no (where "00" means unknown chromosome no and "MT" means mitochondria), "P"and"M" stands for Plus and Minus strand and then approx. kilobase start coordinate.

Methodology

The first aim of this study is to make protein interaction network of ALXD where interactions are gathering into network in form of node and edges where nodes are defined as proteins and their relationship i.e. edges with other proteins then it followed by analysis of network topology where network's architecture and performance can be calculated and in enrichment analysis nodes are characterized in term of biological attribute. Both steps are important steps and plays vital role in analysis done through NetworkAnalyser and JEPETTO plugin (Figure 1).

Results

The genes involved in Alexander disease find through GeneCards database and publish literature. All genes that are involved in this disease are protein coding genes hence; protein-protein interaction network was imported through Cytoscape's web services option.

Interactions were imported as a merged network, which contains total 7796 nodes and 9310 edges. This protein interaction network contains all protein of Homo sapiens and other organisms also like Mus Musculus, Mycobacterium, Canisfamiliaris, Gallus and other species.

The study was aimed to study proteins involved in ALXD, only humans protein were selected and if network contains any self-loops, duplicate edges and any singletons then it was removed. Protein interaction network contains 3490 proteins and 4691 interactions. More analysis was done in order to know which causative proteins will be used as drug targets which are discussed further.

Discussion

Although it was large protein interaction network it is important to create sub networks to analyze network completely. NetworkAnalyser was used to create subnetwork and analyze the protein interaction network and discover its topological properties. In this study two

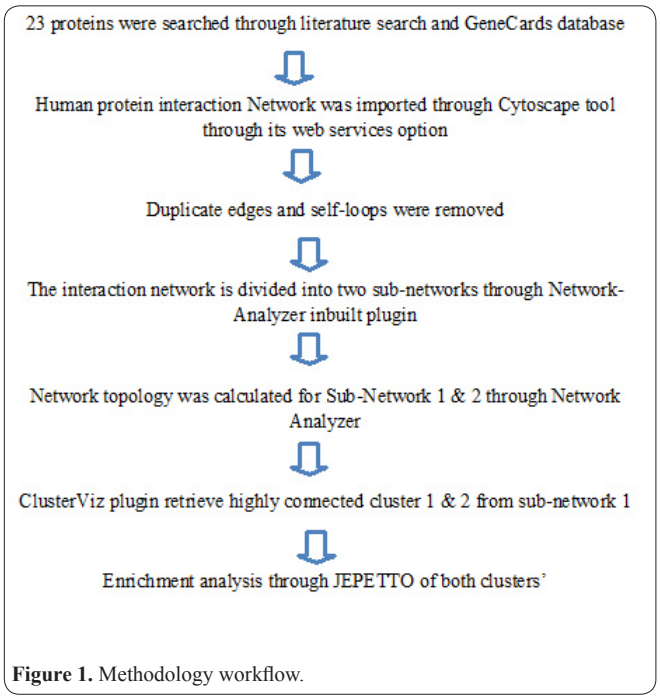

Figure 1. Methodology workflow.

sub-networks was created which contains 2532 proteins in component 1 and 924 proteins in component 2.

Topological parameters were calculated for both sub-networks and comparative analysis of both sub-networks was carried out to obtain significant sub-network. The analysis of the network was done assuming the network as undirected graph.

In undirected networks, two nodes are connected if there is a path of edges between them. Within a network, all nodes that are pairwise connected form a connected component. The number of connected components indicates the connectivity of a network a lower number of connected components suggest a stronger connectivity. A value of 1 for the obtained both sub-network indicates that the connected components have a stronger connectivity.

The shortest path calculated was 100% for both sub-networks. This shows that the minimum requirement of a gene for the transduction of a response from one molecule to another should be atleast 100%.

The network heterogeneity is 6.708 (for sub-network 1) and 4.647 (for sub-network 2) which reflect the good tendency of a network to contain hub (essential) nodes.

After visualizing and calculating all the parameter through NetworkAnalyzer, sub-network 1 is selected for further analysis of study because published literature reported which proteins are responsible for disease are all present in sub-network 1 with interaction of other proteins.

Clusters in the network can be considered as gene complexes and functional modules, which can be identified as highly interconnected sub-graphs. A cluster in the network can be considered as a community which has strong inter-relationship among their members.

The analysis of the network was done using the ClusterViz plugin which uses MCODE algorithm in Cytoscape to find out highly inter-connecting groups and parameter for cluster analysis was set as default values (Table 2), 2 clusters were obtained with high score in descending order.

The first cluster with the highest score of 5.333 and second cluster with 3.842 score (Table 3) was considered in which all the interacting genes could be the potential genes which may cause Alexander disease and depicts that these clusters are functionally linked and possibly can play major role in causing Alexander disease or could be involved in the pathogenesis/molecular mechanism of the disorder which is still unknown.

The enrichment analysis was done with the help of JEPETTO plugin, which gives list of significant pathway in which proteins of Alexander disease overlaps (in terms of XD-score). Enrichment analysis was done to know the potential pathways/process which can be play important role in disease and for this analysis threshold values were taken as default and reactome is an open source curated bioinformatics database of human pathways and reactions hence, it was chosen as reference database. Statistically significant results were highlighted in result table. Table 4 and Table 5 show the list of significant pathway/process with their respective XD-score.

The molecular interaction study of Alexander disease was done using Cytoscape tool and its various plugins. This study focuses on building and analyzing the disease causing human protein interaction network of Alexander disease. Human protein interaction network was retrieve from Cytoscape web services and two sub-networks were created and analyzed through NetworkAnalyser inbuilt plugin. The idea behind this study was to identify probable drug targets for Alexander disease, through cluster analysis it was found that four proteins i.e. GFAP, PLEC, CRYAB and NDUFV1 in cluster 1 and in cluster 2 two proteins i.e. CASP3 and MAPK14 proteins plays major role in causing Alexander disease and used as a drug targets. Through enrichment analysis with JEPETTO plugin suggests significant signaling pathways and by triggering those pathways different therapeutic strategies will develop further in order to prevent disease.

Table 2. Parameter of ClusterViz plugin used to find out highly connected clusters.

Algorithm: MCODE
Scoring: IncludeLoop: false DegreeThreshold: 2
Clustering: NodeScoreThreshold: 0.2 Haircut: true Fluff: false K-Core: 2 Max.DepthFromSeed: 100

Table 3. List of proteins found in clusters through ClusterViz plugin.

Cluster	Score	(Nodes * Edges)	Nodes Name
1	5.333	10*25	VIM,UBB, UBA52, PLEC, GFAP, SLC1A2, HSPA4, SOD1, CRYAB, NDUFV1
2	3.842	39*73	CCNB1, VEGFA, CYCS, CTNNB1, KRT18, TNF, KIAA0683, FAM48A, TTI1, IL1B, CDKN1A, CASP3, MAP2K4, IL6, IFNG, FOXO3, MAPK1, STAT1, YWHAZ, CCND1, OXA1L, MAP2K3, MAP2K1, MAPK11, PIK3CA, MAPK14, BCL2, NFKBIA, ATF2, BAX, CDKN1B, HYRC, MMP9, TIP49a, PPARG, FBXO32, RELA, PPARGC1A, JUN

Table 4. Result of enrichment analysis of cluster 1 through JEPETTO plugin. Table shows pathways/process with their respective XD-score and number of overlap proteins.

Pathway or Process	XD-score	q-value	Overlap/Size
CASPASE MEDIATED CLEAVAGE OF CYTOSKELETAL PROTEINS	0.72123	0.04759	2/11

Table 5. Result of enrichment analysis of cluster 2 through JEPETTO plugin. Table shows pathways/process with their respective high XD-score and number of overlap proteins.

Pathway or Process	XD-score	q-value	Overlap/Size
ACTIVATION OF THE AP1 FAMILY OF TRANSCRIPTION FACTORS	1.94157	0.00000	5/10
ACTIVATED TAK1 MEDIATES P38 MAPK ACTIVATION	0.86465	0.00334	3/13
TRAF6 MEDIATED INDUCTION OF THE ANTIVIRAL CYTOKINE IFN ALPHA BETA CASCADE	0.71080	0.00000	10/52
APOPTOTIC CLEAVAGE OF CELL ADHESION PROTEINS	0.66885	0.05030	2/11
SHC MEDIATED SIGNALLING	0.66885	0.05030	2/11
MAP KINASES ACTIVATION IN TLR CASCADE	0.66885	0.00000	8/44
TOLL LIKE RECEPTOR 3 CASCADE	0.63123	0.00000	10/58
GRB2 EVENTS IN EGFR SIGNALING	0.60824	0.05030	2/12
P38MAPK EVENTS	0.60824	0.05030	2/12
P75NTR SIGNALS VIA NFKB	0.60824	0.05030	2/12
SHC RELATED EVENTS	0.60824	0.05030	2/12
SOS MEDIATED SIGNALLING	0.60824	0.05030	2/12
VIRAL DSRNA TLR3 TRIF COMPLEX ACTIVATES RIP1	0.60824	0.05030	2/12
MAPK TARGETS NUCLEAR EVENTS MEDIATED BY MAP KINASES	0.60824	0.00012	5/30
SIGNALLING TO RAS	0.58157	0.00130	4/25
SIGNALLING TO P38 VIA RIT AND RIN	0.55696	0.05725	2/13

Acknowledgements

I like to put my sincere acknowledgements to DBT for providing us such platform and financial assistance. And my sincere thanks to BIF Center at D.G.P.G. College, Kanpur and all staff members there.

References

1. Brenner M, Johnson AB, Boespflug-Tanguy O, *et al*. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. Nat Genet 2001; 27: 117-20.
2. Chen YS, Lim SC, Chen MH, *et al*. Alexander disease causing mutations in the C-terminal domain of GFAP are deleterious both to assembly and network formation with the potential to both activate caspase 3 and decrease cell viability. Exp Cell Res 2011; 16: 2252-66.
3. Rodriguez D, Gauthier F, Bertini E, *et al*. Infantile Alexander Disease: Spectrum of GFAP Mutations and Genotype-Phenotype Correlation. Am J Hum Genet 2001; 69:1134-40.
4. Berggård T, Linse S, James P. Methods for the detection and analysis of protein-protein interactions. Proteomics 2007; 7:2833-42.
5. Raman K. Construction and analysis of protein-protein interaction networks. Autom Exp 2010; 2:2.
1. Cai J, Chen G, Wan J. ClusterViz: a Cytoscape plugin for graph clustering and visualization. School of Information Science and Engineering, Central South University, Changsha, China, 1.01 edition, Feb 2010.
2. Winterhalter C, Widera P, Krasnogor N. JEPETTO: a Cytoscape plugin for gene set enrichment and topological analysis based on interaction networks. Bioinformatics 2014; 30:1.
3. Head MW, Corbin E, Goldman JE. Overexpression and Abnormal Modification of the Stress Proteins aB-Crystallin and HSP27 in Alexander Disease. Am J Pathol 1993; 143:1743-53.
4. Huang daW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: path towards the comprehensive functional analysis of large gene list. Nucleic Acids Research 2009; 37:1-13.
5. Melchionda L, Fang M, Wang H, et al. Adult onset Alexander disease, associated with a mutation in an alternative GFAP transcript, may be phenotypically modulated by non-neutral HDAC6 variant. Orphanet J Rare Dis. 2013; 8:66.
6. Najafi A, Masoudi-Nejad A, Ghanei M, et al. Pathway Reconstruction of Airway Remodeling in Chronic Lung Diseases: A Systems Biology Approach. PLoS One 2014; 6, e100094.

7. Natale M, Benso A, Di Carlo S, *et al*. FunMod: a Cytoscape plugin for identifying functional modules in undirected protein-protein networks. Genomics Proteomics Bioinformatics 2014; 12: 178-86.
8. Pekny T, Faiz M, Wilhelmsson U, *et al*. Synemin is expressed in reactive astrocytes and Rosenthal fibers in Alexander disease. APMIS 2014; 122: 76-80.
9. Rebhan M, Chalifa-Caspi V, Prilusky J, *et al*. GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. Bioinformatics 1998; 14:656-64.
10. Stumpf E, Masson H, Duquette A, *et al*. Adult Alexander disease with autosomal dominant transmission: a distinct entity caused by mutation in the glial fibrillary acid protein gene. Arch Neurol. 2003; 9, 1307-12.
11. Tang G, Yue Z, Talloczy Z, *et al*. Autophagy induced by Alexander disease-mutant GFAP accumulation is regulated by p38/MAPK and mTOR signaling pathways. Hum Mol Genet 2008; 17:1540-55.
12. Tang G, Yue Z, Talloczy Z, Goldman JE. Adaptive autophagy in Alexander disease-affected astrocytes. Autophagy 2008; 4:701-3.
13. Tian R, Gregor M, Wiche G, *et al*. Plectin Regulates the Organization of Glial Fibrillary Acidic Protein in Alexander Disease. Am J Pathol 2006; 168:888-97.
14. Zhu X, Gerstein M, Snyder M. Getting connected: analysis and principles of biological networks. Genes Dev 2007; 21:1010-24.
15. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 2003; 13(11):2498-504.
16. Bader GD and Hogue WVC. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003; 4:2.