Int J Pharm Bio Sci 2013 Oct; 4(4): (B) 845 - 860

Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

EXPRESSION-BASED NETWORK BIOLOGY IDENTIFIES HIGH SCORE SUB-NETWORKS AND THEIR ROLES IN RETT SYNDROME

SARAVANAKUMAR SELVARAJ AND JEYAKUMAR NATARAJAN*

Data Mining and Text Mining Laboratory, Department of Bioinformatics, Bharathiar University, Coimbatore, Tamilnadu-641 046, India.

ABSTRACT

Rett syndrome (RTT) is a childhood neurodevelopmental disorder and one of the most common causes of mental retardation that primarily affects girls. Hence, there is a constant need to discover new and efficient treatment against the disease by seeking to uncover various novel alternate signaling mechanisms that can lead to Rett syndrome and its associated complications. In this present work, we used Rett syndrome microarray data to identify the significant genes by gene expression data analysis and also constructed the protein-protein interaction (PPI) networks of genes/proteins involved in the pathophysiology of RTT. We identified three high score sub-networks from the large PPI networks and these three sub-networks have scale-freeness topology. The functional enrichment analysis for all the genes of these three sub-networks; we found genes from the third sub-network have biological processes such as *neurological disorders* and *nervous system development and function* related to RTT. An experimental research on this sub- network and their associated genes and pathways may further provide suitable drug target identification and better understanding of the pathophysiology of RTT.

KEYWORDS: **Rett syndrome; Microarray data; Protein interaction networks.**

JEYAKUMAR NATARAJAN
Data Mining and Text Mining Laboratory, Department of Bioinformatics, Bharathiar University, Coimbatore, Tamilnadu-641 046, India.

INTRODUCTION

Rett syndrome (RTT) is a neurological and developmental disorder that mostly affecting central nervous system exclusively in female children [1]. This condition affects an estimated 1 in 10,000 females. The RTT phenotype typically includes normal development until 6 to 18 months of age, and then there is regression with slowing of head circumference growth, loss of language, development of stereotypical hand movements and gait and truncal apraxia[2]. By the third year of life some girls also have severe electroencephalogram abnormalities, seizures, spasticity and scoliosis [3]. Mutations in the gene encoding methyl-CpG-binding protein 2 (*MeCP2*) on the X chromosome have been identified in 95% of girls with RTT phenotype as a major cause[4]. To date more than 100 different mutations in the transcriptional silencer MeCP2 gene have been identified in patients with RTT[5]. The main pathology of MeCP2 in the brain is neuronal development arrest[6]. Other genes, like CDKL5[7], FOXG1[8] and NTNG1[9] have also been implicated in a limited number of cases of RTT. Males with RTT almost all have Klinefelter's syndrome (XXY karyotype)[10] and they die before birth or in infancy. The signs and symptoms of this disorder are associated with Angelman syndrome, cerebral palsy and autism[11]. Currently, there is no treatment available for RTT cases. The molecular basis and mechanisms involved in RTT phenotype among its subtypes remain unclear. Therefore, understanding the molecular pathogenesis and improve the therapeutic strategies for RTT is essential to improve the quality of life for affected individuals.The combined analysis of expression profiles and protein-protein interaction (PPI) networks have become one of the major and powerful approaches to elucidating the molecular mechanisms underlying the complex diseases[12,13]. The topology and neighborhood of a given protein in the network have been used to functionally characterize the proteins[14]. Further, it has also been shown that the proteins associated with diseases tend to have a high connectivity between them, particularly in inherited diseases and ataxia[15]. This approach is comprehensively being used to identify the candidate genes and associated networks for various complex diseases including cancers and diabetes. For example, Sengupta et al., identified four novel sub-networks and related signaling pathways which could have an impact on the diabetes and other associated dysfunctions using this network biology approach[16]. Similarly, Kar et al., identified the breast cancer associated gene network and also the novel breast cancer risk gene[17]. Further, this approach has already been proved to be effective for the identification of complex disease genes, including those involved in colon cancer[18], Schizophrenia and other psychiatric disorders[19]. In the present study, we have applied the network biology approach which involves the gene expression data and their corresponding interaction network to identify signature sub-networks in Rett syndrome. The aim of this study is to put promote novel biological sub-networks that describe transcriptional alterations in genes which could contribute to the molecular pathogenesis of RTT and its associated problems. We used Rett syndrome microarray data to identify the significantly expressed genes and their co-expressed gene network. Then, we have applied a network biology approach which involves the integration of co-expressed gene network with corresponding protein interaction network to identify the functional sub-networks. We are the first one to attempt the PPI network analysis in Rett syndrome and hope our results depict their potential involvement in RTT progression and various associated complications as well.

MATERIALS AND METHODS

(i) Microarray dataset

The Rett syndrome microarray dataset (GDS2613) have been retrieved from Gene Expression Omnibus (GEO) at National Center for Biotechnology Information (NCBI) [http://www.ncbi.nlm.nih.gov/geo/]. This dataset

contains six samples, three from normal and another three from Rett syndrome conditions. These samples have been collected from the frontal cortex of the brain. Affymetrix HG_U95Av2 platform was used to perform this study. This dataset contains the expressions of 12602 genes in both conditions (normal vs. diseased).

(ii) Identification of differentially expressed genes

Microarray data sets are commonly very large and analytical precision is influenced by a number of variables. So it is extremely useful to reduce the dataset to those genes that are best distinguished between the two cases (e.g. normal vs. diseased)[20]. In order to identify the differentially expressed genes (DEGs) from the Rett syndrome microarray data, SAM (Significance Analysis of Microarray) a modified t-test[21] was used. Differentially expressed genes are identified from the dataset at the cut-off p-value <=0.05 and fold change value >=1.5. The genes that are satisfied those conditions were identified as differentially expressed genes and the heat-map was also generated by Multi Experimental Viewer (MeV) [http://www.tm4.org/mev.html].

(iii) Construction of protein-protein interaction network

Protein–protein interaction (PPI) networks provide valuable information in the understanding of cellular function and biological process and used to understand the molecular mechanism of disease[22]. In such networks, gene or protein is represented as a node and interactions or links that a protein has with other proteins is defined as edges. In this study, the protein-protein interaction (PPI) network for the differentially genes has been constructed by APID2NET plugin in Cytoscape [http://bioinfow.dep.usal.es/apid/apid2net.html]. APID2NET retrieves all the possible information on protein-protein interaction from five interaction databases namely, Database of Interacting Proteins (DIP) [http://dip.doe-mbi.ucla.edu/dip/Main.cgi], Biomolecular Interaction Network Database (BIND) [http://bond.unleashedinformatics.com/], IntAct [http://www.ebi.ac.uk/intact], Molecular Interactions Database [http://mint.bio.uniroma2.it/mint] and Human Protein Reference Database (HPRD) [http://www.hprd.org/]. The Swissprot/Uniprot IDs for each gene have been collected from APID database[23] and it was given as the input to APID2NET plugin to construct the PPI network. The constructed network was visualized by Cytoscape.

(iv) Searching for high-scoring sub-networks and validation

A sub-network of large PPI network is defined as a set of statistically and functionally significant interacting proteins. The potential sub-networks have been identified by a search method estimating their significance score implemented in BioNet package [http://bionet.bioapps.biozentrum.uni-wuerzburg.de]. First, the significance score has been calculated for all the nodes in the PPI network. The threshold value score has been set as 0.5 and sub-networks which have showing the similar value has been taken into further analysis. The resultant high score sub-networks were further refined by wilcoxon test in R-package [http://www.r-project.org/]. The Wilcoxon signed rank test is a non-parametric test that can be used to test hypotheses about different populations[24]. It calculates the p-values based on their significance score using null hypothesis. The p-values of each sub-network which show <0.05 is more significant and then the high-scoring sub-networks were extracted from the larger PPI networks.

(v) Topological properties of high score sub-networks

The topological and statistical significance of each high score sub-networks, have been calculated using Network Analyzer[25] plugin in Cytoscape. We have used the following topological properties to evaluate the sub-networks. Topological coefficient is a relative measure for the extent to which a node shares neighbours with other nodes. The topological coefficient (T_n) of a node n with k_n neighbours is

computed as, T_n = avg (J (n, m)) / k_n. The value *J* (*n*,*m*) is the number of neighbours shared between the nodes *n* and *m*. Clustering coefficient (Ci) is the average clustering coefficient of all the nodes; where the clustering coefficient of a node i (Ci) is defined as the proportion of links exists between nodes within the i-neighborhood divided by the number of kinks possibly exit between them It is expressed as **C** (n) = $2e_n$ / (k_n (k_n-1)), where k_n is the number of interactions of node n and e_n is the number of connected pairs between all the neighbors of node n. Average degree (k) represents the mean of all degree values of nodes in a network. This is extended by network density, which specifies the compactness of one network distributed through its edges. Centrality measures are used to compute the structure of node in the network and identify the hub proteins. Network Analyzer used least square method[25] and considers only the data points with positive co-ordinate values for fitting the line, where the power-law curve is $y=\beta x^{\alpha}$. It also gives the correlation between the given data points and the corresponding points on the fitted line. The R-squared value (also known as coefficient of determination) measures how well the data points fit to the curve. Betweenness centrality measures how often nodes occur on shortest paths between other nodes. The biological networks are distinguished by topological features which establish their scale-freeness property[26]. The characteristic feature of a scale-free network is the presence of several hubs in the biological networks and a large number of nodes with a few connections. Power law process is used to estimate the parameters and validating the network models with their scale-freeness properties. The R-squared values closer to one indicate higher correlation and a stronger linear relationship between the data variables.

(vi) Functional enrichment analysis of networks

In order to identify how the high-scoring sub-networks are functionally enriched we used BinGO, a cytoscape plug-in[27]. It is based on a hypergeometric test with Benjamini and Hochberg false discovery rate (FDR) and displays the overrepresented functional gene ontology (GO) categories in the given sub-networks.

RESULTS

(i) *Differentially expressed genes for Rett syndrome data*

The differentially expressed genes of the Rett syndrome microarray data were identified by using Significance Analysis of Microarray (SAM) a modified t-test algorithm. A total number of 86 differentially expressed genes were identified and all the genes are positively regulated (up regulated). The heat-map for these 86 differentially expressed genes has created by MeV software and it has shown in Figure 1.

A heat-map of differentially expressed genes in Rett syndrome microarray data

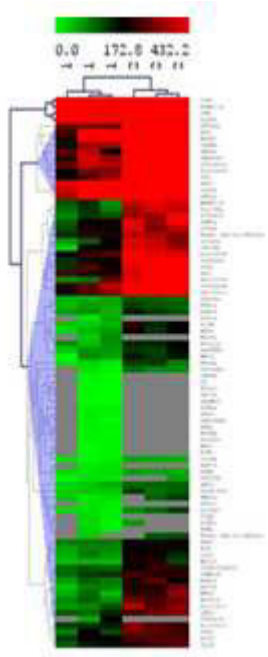

Figure 1
A heat-map showing 86 differentially expressed genes based on their expression pattern (from high to low) of Rett syndrome microarray data

(ii) Protein-Protein Interaction (PPI) network
The PPI network of the differentially expressed genes was constructed based on their interaction data from available interaction databases and the existing literatures by using the APID2NET plug-in. The constructed complex Rett syndrome PPI network showed the interactions of 58 out of 86 significant genes. In this large network, the genes are represented as nodes and the interactions between the genes are referred as edges. The genes which are showing higher interaction with other genes in the network are referred as "hubs" and it has an important biological function among the interacting partners of the PPI network.

(iii)High score Sub-networks
The complex PPI networks are critical to analyze every node in the aspect of biological processes. So these complex networks are likely comprised of several sub-networks or functional modules contributing to various diverse biological processes and identification of novel marker genes for diseases[28]. In this study, we identified three sub-networks which are statistically significant in complex PPI network of Rett syndrome. The p-values of these three sub-networks are 1.98×10^{-6}, 1.30×10^{-5}, and 1.52×10^{-5} respectively. These sub-networks were named as, neuron differentiation, neuron apoptosis and nervous system development and signaling as per their functional enrichment properties and involving genes. The number of nodes, number of edges and the p-values of the sub-networks are shown in Table 1. These three high-score sub-networks have been further investigated to learn the functional relevance pertaining to Rett syndrome by information available from literature and online repositories.

Table 1
Number of nodes, edges and p-values of three potential sub-networks

Name of the sub-network	No of nodes	No of edges	p-values
Neuron differentiation sub-network	108	359	1.98×10^{-6}
Neuron apoptosis sub-network	93	302	1.30×10^{-5}
Nervous system development and signaling sub-network	77	188	1.52×10^{-5}

(a) Neuron differentiation sub-network

The five of our significant genes such as NTRK2, WASF3, ADDG, EZRI, and SASH1 has been found in statistically significant sub-network of neuronal development. From the identified sub-network the above five genes are act as 'hub genes' because it has a large number of interaction and these genes are linked through the CDK5 gene. This CDK5 gene is essential for the development of the central nervous system through regulation of neuronal migration[29]. In mature neurons, CDK5 has been implicated in various signaling transduction pathways, which contribute to functional neuronal activity. It might be a potential drug target for the treatment of neurodegenerative diseases, drug abuse and diabetes mellitus[30]. The gene NTRK2 is involved in the development and the maturation of central and peripheral nervous systems through regulation of neuron survival, proliferation, migration, differentiation, and synapse formation and plasticity[31]. Phosphorylation of EZR proteins by LRRK2 promotes the rearrangement of actin cytoskeleton in neuronal morphogenesis[32]. NGF and BDNF proteins are involved in various synaptotrophic effects and might also be related to neuropsychiatric diseases such as dementia, depression, schizophrenia, autism, Rett syndrome, anorexia nervosa and bulimia nervosa[33]. Most of the interacting genes/proteins in this network shared common pathways such as, MAPK signaling pathway (hsa04010) and neurotrophin signaling pathway (hsa04722). The significant genes and its interacting partners of this sub-network are shown in Figure 2. The hub genes (SASH1, NTRK2, EZR, ADDG and WASF3) are represented in red colour and the genes which are directly connected to hub genes (FYN, CDK5, EGFR, etc) are shown in green colour. The other interacting genes are shown in violet colour.

Neuron differentiation sub-network

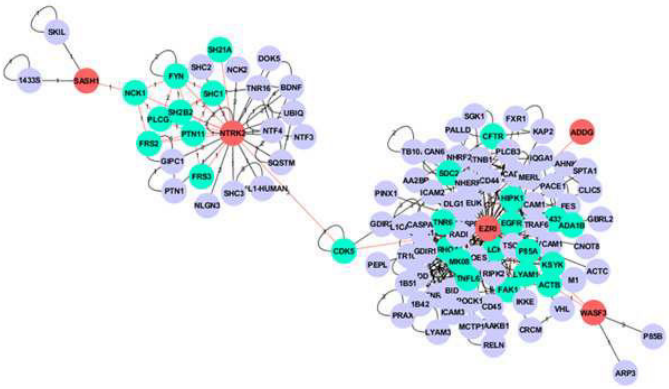

Figure 2

Neuron differentiation sub-network: In this sub-network, the significant hub genes SASH1, NTRK2, EZR, ADDG and WASF3 are marked in red colour. The green colour genes are the primary interacting partners of hub genes. The violet colour genes are the other interacting genes with our hub genes and primary interacting genes.

(b) Neuron apoptosis sub-network

The sub-network of neuron apoptosis has two hub genes such as BCLX and PIK3CD. These two hub genes are primarily connected by the genes such as, BCL2, IRS2, P85A and P85B. In this sub-network, the BCL-2 family proteins are mostly involved in programmed cell death or apoptosis. It interacts with DMN1L to stimulate the GTPase activity in synapses[34]. The gene HRK also interacts with BCL2 and activates the neuron apoptosis. PIK3CD are involved in the process of immune response and it is play an important role in both protein sorting and nutrient sensing in concert with the mTOR pathway[35]. IRS2 is mainly involved in the condition of type II diabetes. This protein regulates the glucose and lipid metabolism mainly through PI-3 kinase-Akt pathway and through activation of the mitogen-activated protein kinase (MAPK) cascade gene expression, cell growth and differentiation[36]. The most interacting genes of this network shares the biological roles such as signal transduction and neuron apoptosis. The genes involved in this neuron apoptosis sub-network are shown in Figure 3. In this sub-network hub genes (BCLX and PIK3CD) are showed in red colour. The green colour indicates the genes (IRS2, BCL2, P85A and P85B) which are connected primarily with hub genes and the violet colour genes are the other interacting partners of this sub-network.

Neuron apoptosis sub-network

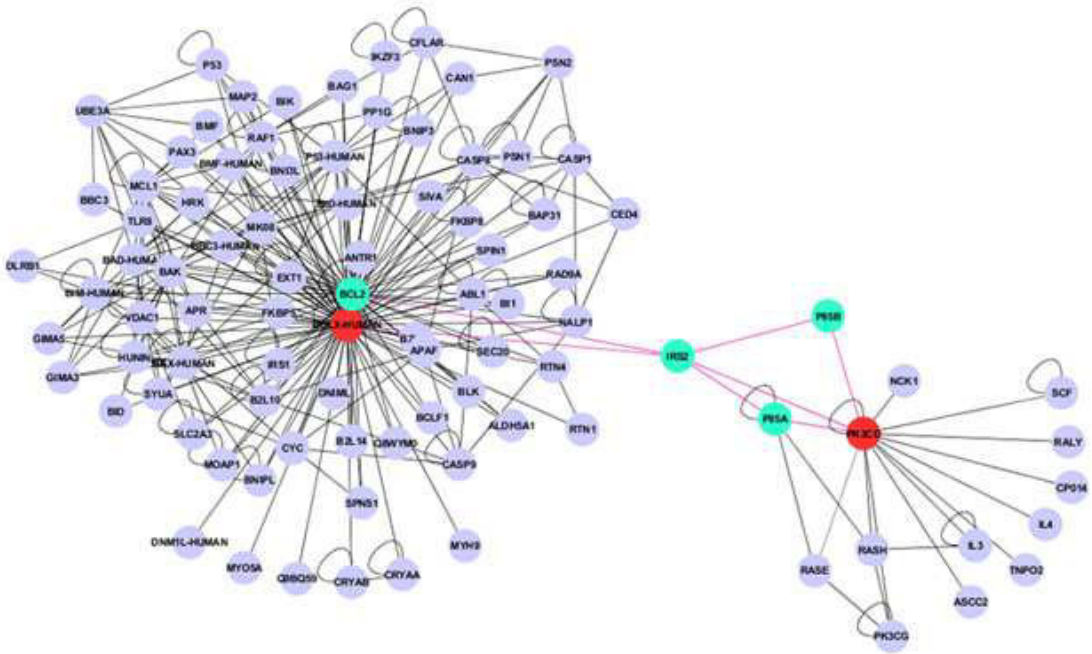

Figure 3

Neuron apoptosis sub-network: In this sub-network the co-expressed genes (hub genes) are colored as red and the green colour genes are the primary genes which connects all of our co-expressed genes. The violet colour genes are next level interacting partners of our hub genes.

(c) Nervous system development and signaling sub-network

The Nervous system development and signaling sub-network has seven of our significant genes such as UBE3A, CDK5R2, MYO5A, FBLN1, MDF1, HOXA1 and DTNA. In this sub-network, the CDK5R2 gene acts as a critical regulator of neuronal migration in the developing central nervous system. The cross talk between CDK5R2 and CaMKII signal transduction pathways may be a component of the complex molecular mechanisms contributing to synaptic

plasticity, memory and learning[37]. HOXA1 gene is critical for hindbrain development and it has a phenotypic features frequently observed in autism[38]. CHN1 gene plays an important role in neuronal signal-transduction mechanisms. MDF1 binds to the AXIN1 complex and it affects the regulation of Wnt and JNK signaling pathways[39]. On the basis of these observations supported by the literatures, these sub-networks have the significant genes which are mostly involved in the nervous system developmental processes and signal transduction. The interacting genes of this sub-network are shown in Figure 4.

Nervous system development and signaling sub-network

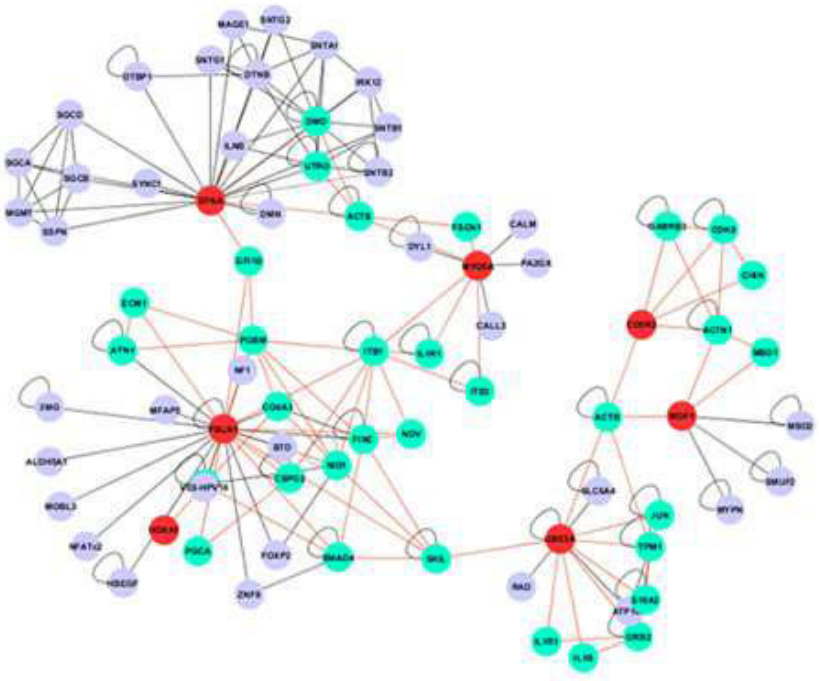

Figure 4

Nervous system development and signaling sub-network: In this sub-network the significant hub genes (UBE3A, CDK5R2, MYO5A, FBLN1, MDF1, HOXA1 and DTNA) are marked in red colour and the green colour genes are the primary interacting genes with all of our significant genes. The violet colour genes are next level of interacting partners.

(iv) Scale-freeness topology of three high-score sub-networks

Network topology is to provide the significance of a node in communicating with other genes or proteins of interest in the network. The parameters such as average clustering coefficient, topological coefficient, average degree and power law distribution of degrees and betweenness centrality are used to predict the topology of the networks. Additionally, the distribution of clustering coefficient is also an important characteristic of scale-free networks. The clustering coefficient of a node is always lies between 0 and 1. The average clustering coefficient describes the 'cliquishness' of networks. The average clustering coefficient, average degree and network density of high score sub-networks were shown in Table-2. The log-log graph of average clustering coefficient, power law node degree distribution, topological coefficient and betweenness centrality of three sub-networks are shown in Figure-5, Figure-6 and Figure-7 respectively. R-square also known as coefficient of determination is a commonly used statistic to evaluate model fit. The R-square value is an indicator of how well the model fits

the data. When the fit is good, the R-squared value is very close to 1.0. All of our three high-score sub-networks are satisfied the above property and the scale-freeness topology is maintained in all the three sub-networks. Therefore, the information on network topology provides high confidence to their scale–freeness property supporting the fact that the underlying models are linear.

Table 2
Topological parameters of three high score sub-networks

Sub-networks	Average degree	Network density	Clustering coefficient
Neuron differentiation sub-network	5.34	0.053	0.542
Neuron apoptosis sub-network	5.63	0.061	0.555
Nervous system development and signaling sub-network	4.23	0.050	0.514

(v) Functional properties of sub-networks

In general, similarly functioning proteins tend to form clusters of PPI and share common biological processes. The biological functions assigned to the dataset are ranked according to the significance of that biological function to the dataset. Indeed, the hub genes and its associated genes of three high-score sub-networks were enriched with functional categories such as cellular development, immune cell trafficking, nervous system development and function, cell death and cellular movement. They shared significantly altered canonical pathways including glucocorticoid receptor signaling, IL-4 signaling, neurotrophin signaling and NF-kB pathways. The genes/proteins which are involved in the two specific functional categories related to RTT (e.g. neurological diseases and nervous system development) are listed in the Table 3 with their functional properties. Interestingly most of these genes are present in sub-network 3 (nervous system development and signaling sub-network). The genes and their associated network may have an important role in Rett syndrome and interesting to further investigate.

Topological parameters of neuron differentiation sub-network

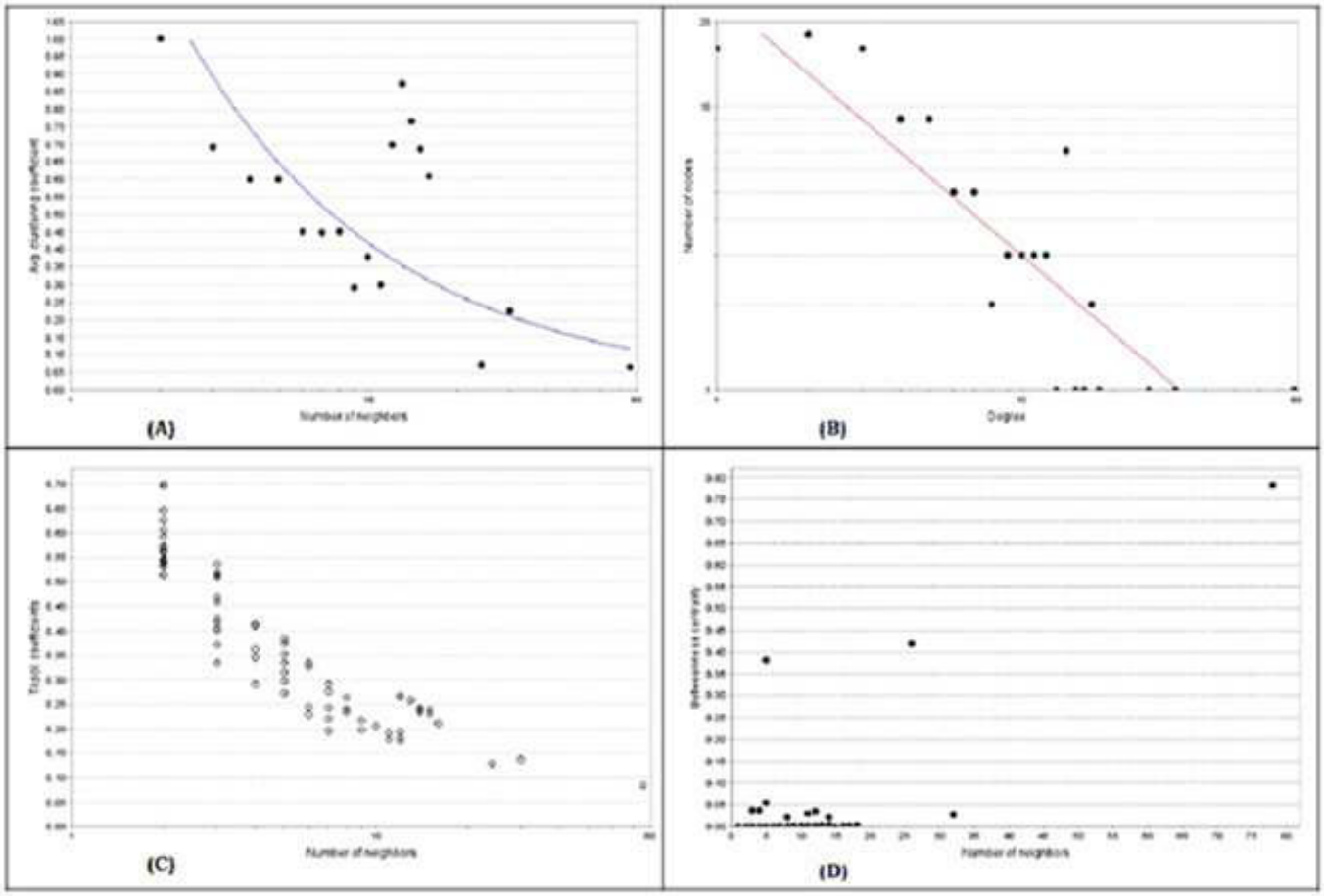

Figure 5

(A) Average clustering coefficient (R-squared value 0.583),
(B) Power law node degree distribution (R-squared value 0.746),
(C) Topological coefficients and
(D) Betweenness centrality of the interacting proteins.

Topological parameters of neuron apoptosis sub-network

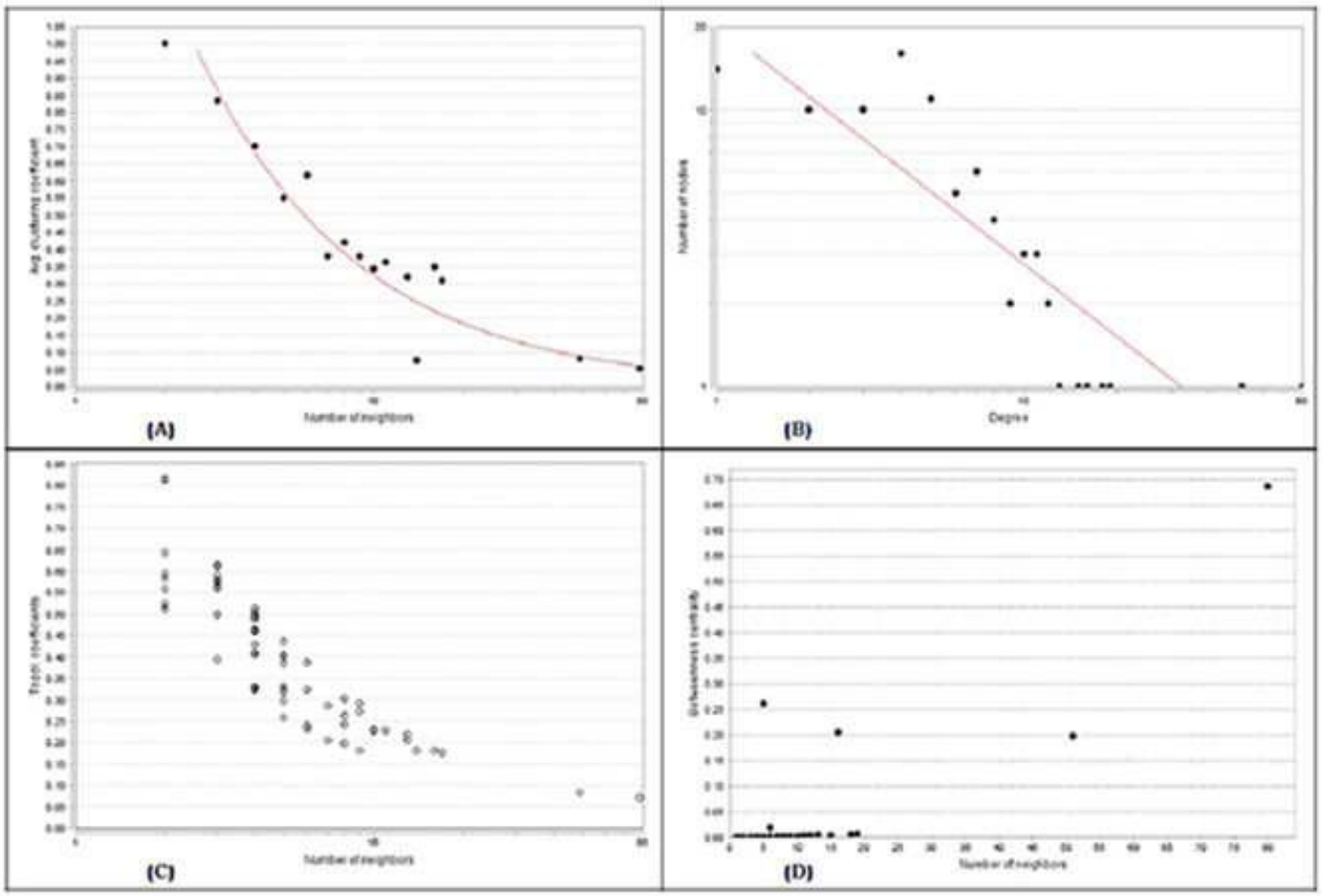

Figure 6

(A) Average clustering coefficient (R-squared value 0.820),
(B) Power law node degree distribution (R-squared value 0.756),
(C) Topological coefficients and
(D) Betweenness centrality of the interacting proteins

Topological parameters of nervous system development and signaling sub-network

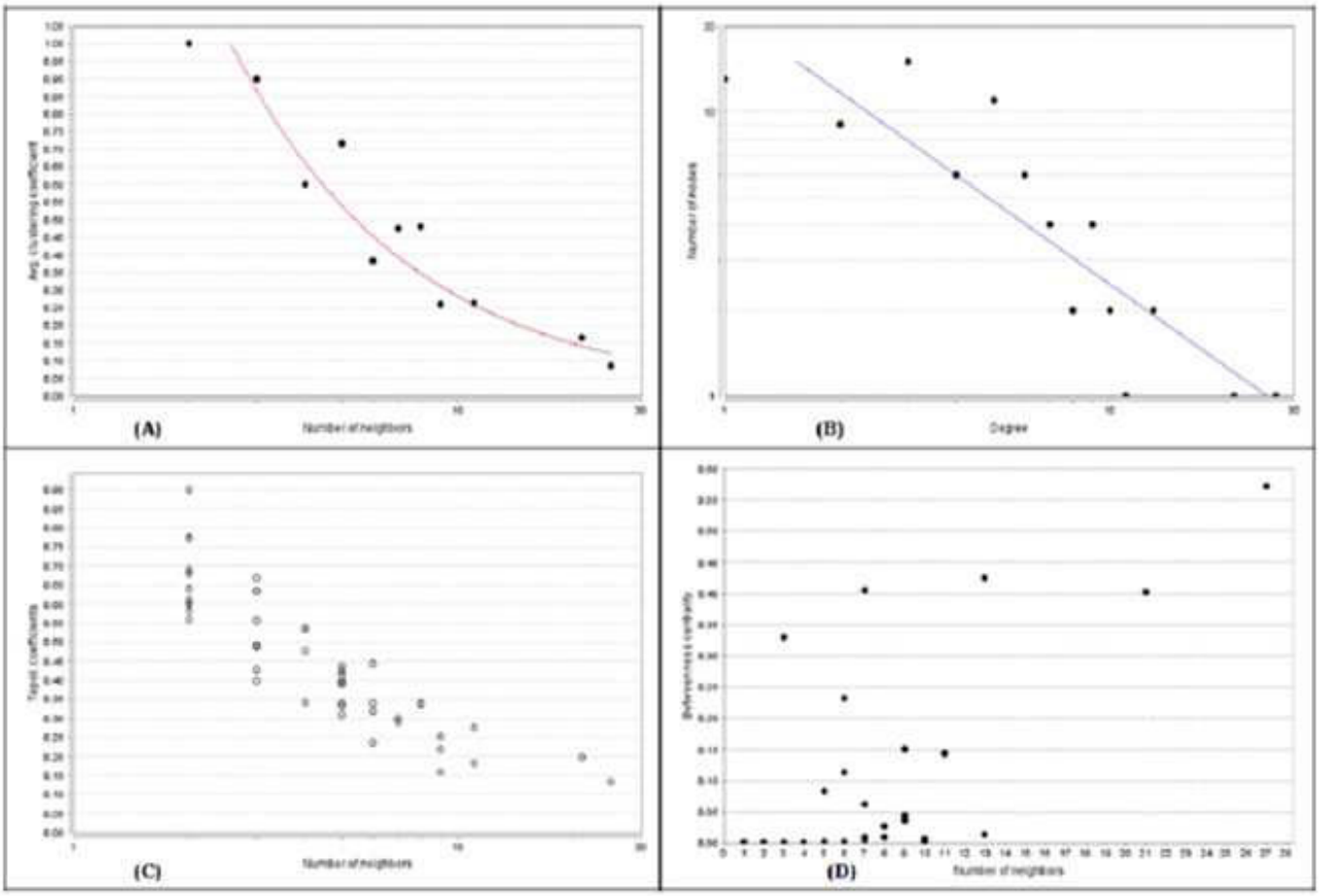

Figure 7

(A) Average clustering coefficient (R-squared value 0.912),
(B) Power law node degree distribution (R-squared value 0.768),
(C) Topological coefficients and
(D) Betweenness centrality of the interacting proteins

Table 3

The genes involved in neurological diseases and nervous system development in three high-score sub-networks

Functional annotation	p-value	Genes
Nervous System Development and Function		
Activation of microglia	1.58E-04	F2RL1, IGHG1, IL1B, IL1R1, PTGS2
Activation of astrocytes	6.53E-04	IL1B, IL1R1, PTGS2
Excitation of neurons	1.02E-03	IL1B, IL1R1, KCNJ2
Outgrowth of sensory axons	2.24E-03	NFATC2, NFATC3
Outgrowth of axons	7.55E-03	ADRB2, NFATC2, NFATC3
Neurological disorders		
Astrocytoma	4.72E-03	AKAP13, GNA12, IL13RA1, PTGS2
Glioma	4.93E-03	AKAP13, GNA12, IL13RA1, MGMT, PTGS2
Ataxia	2.80E-03	CACNA2D2, IL1B, IL1R1, RUNX3

DISCUSSION

Network–based analysis gives the system level relationship of molecules across different levels of regulatory controls of biological processes by integrating functional interactions with co-expression information. The present study focused on expression profiles of the interactions, which existed most consistently in maximum number of Rett syndrome microarray dataset. The analysis and the literature of different biological processes offer an insight into the identification of several facts associated with RTT. Our approach may suggest as a predictive tool for identifying high score sub-networks and its roles in the development of Rett syndrome and its complications. A distinct scale–freeness has been predicted for three high-score sub-networks. The topological parameters such as average degree, average clustering co-efficient, topological coefficient and betweenness centrality values were identified. The betweenness centrality value is higher in these sub-networks, because a gene with higher degree is more likely to be essential since it has more links. All these topological properties showed the scale-free nature of these sub-networks. In neuron differentiation sub-network, the genes such as CDK5, EGFR, DLG1, and LRFN1 are involved in the biological processes such as neuron migration, signal transduction and nervous system development. The other genes such as ABL1, LCK, JUN and FYN are identified as proto-oncogenes[40]. FBX31, IL4, PK3CD, IL3, IRS2, NCK1 and HRAS of neuron apoptosis sub-network genes involved in the biological processes like cell cycle arrest, cellular defense response, signal transduction and mental retardation. In nervous system development and signaling sub-network, most of the genes (GABRA5, NOV, DTNA, RELN, MAPK8, PAX3 etc) are involved in nervous system development, cell cycle arrest, regulation of cell growth, neuron apoptosis and cell-cell signaling. From these three sub-networks, most of the genes from nervous system development and signaling sub-network have significant roles in related to RTT. The functional pathways like glucocorticoid receptor signaling, IL-4 signaling, neurotrophin signaling and NF-kB pathways are also significantly present in this sub-network genes. The genes involved in neurological diseases and nervous system development in these three sub-networks are shown in table-3. The another interacting genes of these three sub-networks are also play an important role in nervous system developmental process in RTT and its related cases are also described. NF-kB plays an important role in nervous system development and function, particularly in synaptic transmission, plasticity, cognition and behavior[41]. IL1B and IL4 are expressed in the developing central nervous system and showed a key role in the regulation of cognitive function[42]. The non-imprinted GABRB3 gene also showed significantly reduced expression in multiple Rett, Angelman and autism brain samples. An overlapping pathway of gene dysregulation within chromosome15q11-q13 in Rett syndrome, Angelman syndrome and autism and implicated MECP2 in the regulation of UBE3A and GABRB3 expression in the postnatal mammalian brain[43]. EGR2 plays an important role in the transient formation of hindbrain developmental compartments and is also an important factor in peripheral myelination, maintenance of synaptic plasticity and long-term potentiation[44]. Severity of autism and / or language impairment suggesting that EGR2 might play a role in the development of autism[45]. Brain-derived neurotrophic factor (BDNF) is a neurotrophic factor playing a major role in neuronal survival, neurogenesis, and plasticity, and it has been shown that BDNF expression is regulated by MeCP2 through a complex interaction. BDNF function may play a significant role in the pathogenesis of RTT[46].

CONCLUSION

In summary, though Rett syndrome was first described in 1966 by Andreas Rett, still to date there is no much work was done. We are the first one to attempt the PPI network analysis on Rett

syndrome gene expression data and identified three high score sub-networks. We also validated these sub-networks based on its scale-freeness. On further, functional enrichment analysis of all the genes of the three sub-networks reveals, the third sub-network named as nervous system development and signaling genes are involved in nervous system developmental disorders and its associate functions. This network and their associated genes are interesting to further investigate. An intensive biochemical analysis on these interactions can bring more insight in to the understanding of their relationships in Rett syndrome. The development of new gene/protein targets leading to the development of therapeutic drugs are the interesting one for further analysis of these PPI interactions.

ACKNOWLEDGEMENT

We thank our lab members Madhuvanthi Kalyanaraman, Suresh Subramani, Kalpana Raja and Mahalaksami Sivamani for their valuable comments.

REFERENCES

1. Amir R.E, Van den Veyver I.B, Schultz R, Malicki D.M, Tran C.Q, Dahle E.J, Philippi A, Timar L, Percy A.K, Motil K.J, Lichtarge O, Smith E.O, Glaze D.G and Zoghbi H.Y, Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann Neurol,* 47:670–679, (2000).
2. Renieri A, Meloni I, Longo I, Ariani F, Mari F and Pescucci C, Rett syndrome: the complex nature of a monogenic disease. *J. Mol. Med,* 81:346–354, (2003).
3. Moog U, Smeets E.J, van Roozendaal K.E.P, Schoenmakers S, Herbergs J, Schoonbrood-Lenssen A.M.J and Schrander-Stumpel C.T.R.M, Neurodevelopmental disorders in males related to the gene causing Rett syndrome in females (MECP2). *Europ. J. Paediat. Neurol*., 7:5-12, (2003).
4. Gonzales L and LaSalle J.M, The role of MeCP2 in brain development and neurodevelopmental disorders. *Curr Psychiatry Rep*., 12(2): 127-34, (2010).
5. LaSalle J.M and Yasui D.H, Evolving role of MeCP2 in Rett syndrome and autism. *Epigenomics,* 1(1): 119–130, (2009).
6. Noruzinia M, Akbari M.T, Ghofrani M and Sheikhha H, Rett syndrome molecular diagnosis and implications in genetic counseling. *Indian J Hum. Genet*, 13(3): 119-121, (2007).
7. Scala E, Ariani F, Mari F, Caselli R, Pescucci C, Longo I, Meloni I, Giachino D, Bruttini M, Hayek G, Zappella M, and Renieri A, CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *J. Med. Genet*, 42:103–107, (2005).
8. Ariani F, Hayek G, Rondinella D, Artuso R, Mencarelli M.A, Spanhol-Rosseto A, Pollazzon M, Buoni S, Spigan O, Ricciardi S, Meloni I, Longo I, Mari F, Broccoli V, Zappella M and Renieri A, FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet,* 83: 89–93, (2008).
9. Borg I, Freude K, Kubart S, Hoffmann K, Menzel C, Laccone F, Firth H, Ferguson-Smith M.A, Tommerup N, Ropers H.H, Sargan D and Kalscheuer V.M, Disruption of Netrin G1 by a balanced chromosome translocation in a girl with Rett syndrome. *Eur. J. Hum. Genet*, 13: 921–927, (2005).
10. Schwartzman J.S, Bernardino A, Nishimura A, Gomes R.R, Zatz M, Rett Syndrome in a boy with 47, XXY karyotype confirmed by a rare mutation on the MECP2 Gene. *Neuropediatrics,* 32: 162-164, (2001).
11. Chahrour M and Zoghbi H.Y, The Story of Rett Syndrome: From Clinic to Neurobiology. *Neuron,* 56: 422-437, (2007).
12. Goh K.I, Cusick M.E, Valle D, Childs B, Vidal M and Barabasi A.L, The human

disease network. *Proc. Natl. Acad. Sci. U S A,* 104: 8685-8690, (2007).
13. Ideker T and Sharan R, Protein networks in disease. *Genome Res,* 18: 644-65, (2008).
14. Sharan R and Ideker T, Modeling cellular machinery through biological network comparison. *Nat. Biotechnol*, 24: 427-433, (2006).
15. Aragues R, Sander C and Oliva B, Predicting cancer involvement of genes from heterogeneous data. *BMC Bioinformatics*, 9: 172, (2008).
16. Sengupta U, Ukil S, Dimitrova N and Agrawal S, Expression-Based Network Biology Identifies Alteration in Key Regulatory Pathways of Type 2 Diabetes and Associated Risk/Complications. *PLoS ONE*, 4(12): e8100, (2009).
17. Kar G, Gursoy A and Keskin O, Human Cancer Protein-Protein Interaction Network: A Structural Perspective. *PLoS Comput Biol*, 5(12): e1000601, (2009).
18. Nibbe R.K, Markowitz S, Myeroff L, Ewing R and Chance M.R**,** Discovery and scoring of protein interaction subnetworks discriminative of late stage human colon cancer. *Mol. Cell Proteomics*, 8(4): 827-45, (2009).
19. Sun J, Jia P, Fanous A.H, van den Oord E, Chen X, Riley B.P, Amdur R.L, Kendler K.S and Zhao Z, Schizophrenia gene networks and pathways and their applications for novel candidate gene selection. *PLoS ONE*, 5(6): e11351, (2010).
20. Mutch D.M, Berger A, Mansourian R, Rytz A and Roberts M.A, Microarray data analysis: a practical approach for selecting differentially expressed genes. *Genome Biol*, 2(12), (2001).
21. Tusher V.G, Tibshirani R and Chu G, Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci.USA*, 98: 5116-21, (2001).
22. Ma H.W and Zeng A.P, Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics*, 19: 270–277, (2003).
23. Hernandez-Toro J, Prieto C and De Las Rivas J, APID2NET: unified interactome graphic analyzer. *Bioinformatics,* 23(18): 2495-2497, (2007).
24. McDonald J.H, Handbook of Biological Statistics, second ed., Sparky House Publishing, Baltimore, Maryland, 2009.
25. Assenov Y, Ramı´rez F, Schelhorn S.E, Lengauer T and Albrecht M, Computing topological parameters of biological networks. *Bioinformatics,* 24: 282–284, (2008).
26. Barabasi A.L and Albert R, Emergence of scaling in random networks. *Science,* 286: 509–512, (1999).
27. Maere S, Heymans K and Kuiper M, BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics*, 21: 3448-3449, (2005).
28. Muff S, Rao F and Caflisch A, Local modularity measure for network clusterizations. *Phys Rev E Stat Nonlin Soft Matter Phys*, 72(5 Pt 2), (2005).
29. Zhang J and Herrup K, Cdk5 and the non-catalytic arrest of the neuronal cell cycle, *Cell Cycle,* 7(22): 3487-90, (2008).
30. Wei F.Y and Tomizawa K, Cyclin-dependent kinase 5 (Cdk5): a potential therapeutic target for the treatment of neurodegenerative diseases and diabetes mellitus. *Mini Rev Med Chem*, 10: 1070-4, (2007).
31. Chen Z, Simmons M.S, Perry R.T, Wiener H.W, Harrell L.E, Go R.C.P, Genetic Association of Neurotrophic Tyrosine Kinase Receptor Type 2 (*NTRK2*) With Alzheimer's Disease. *Am J Med Genet*, 147B: 363–369, (2008).
32. Parisiadou L, Xie C, Cho H.J, Lin X, Gu X.L, Long C.X, Lobbestael E, Baekelandt V, Taymans J.M, Sun L and Cai H, Phosphorylation of ERM Proteins by LRRK2 Promotes the Rearrangement of Actin Cytoskeleton in Neuronal Morphogenesis. *J Neurosci.*, 29(44): 13971–13980, (2009).
33. Chaldakov G.N, Tonchev A.B and Aloe L, NGF and BDNF: from nerves to adipose

tissue, from neurokines to metabokines. *Riv Psichiatr,* 44(2): 79-87, (2009).
34. Hockenbery D, Nuñez G, Milliman C, Schreiber R.D and Korsmeyer S.J, Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature*, 348(6299): 334-6, (1990).
35. Berrie C.P, Phosphoinositide 3-kinase inhibition in cancer treatment. *Expert Opin Investig Drugs,* 10(6): 1085-98, (2001).
36. Fritsche L, Weigert C, Häring H.U and Lehmann R, How Insulin Receptor Substrate Proteins Regulate the Metabolic Capacity of the Liver - Implications for Health and Disease. *Curr. Med. Chem*,15:1316-1329, (2008).
37. Dhavan R, Greer P.L, Morabito M.A, Orlando L.R and Tsai L.H, The cyclin-dependent kinase 5 activators p35 and p39 interact with the alpha-subunit of Ca2+/calmodulin-dependent protein kinase II and alpha-actinin-1 in a calcium-dependent manner. *J Neurosci,* 22(18): 7879-91, (2002).
38. Ingram J.L, Stodgell C.J, Hyman S.L, Figlewicz D.A, Weitkamp L.R and Rodier P.M, Discovery of allelic variants of HOXA1 and HOXB1: genetic susceptibility to autism spectrum disorders. *Teratology,* 62(6): 393-405, (2000).
39. Ikeda S, Kishida S, Yamamoto H, Murai H, Koyama S and Kikuchi A, Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. *EMBO J,* 17(5): 1371-84, (1998).
40. Liu Q, Huang J, Liu H, Wan P, Ye X and Xu X**,** Analyses of domains and domain fusions in human proto-oncogenes. *BMC Bioinformatics,* 10: 88, (2009).
41. Memet S, NF-kappaB functions in the nervous system: from development to disease. *Biochem. Pharmacol*, 72: 1180–1195, (2006).
42. Derecki N.C, Cardani A.N, Yang C.H, Quinnies K.M, Crihfield A, Lynch K.R and Kipnis J, Regulation of learning and memory by meningeal immunity: a key role for IL-4. *J. Exp. Med,* 207: 1067–1080, (2010).
43. Samaco R.C, Hogart A and LaSalle J.M, Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. *Hum Mol Genet,* 14(4): 483-92, (2005).
44. Swanberg S.E, Nagarajan R.P, Peddada S, Yasui D.H and LaSalle J.M, Reciprocal co-regulation of *EGR2* and *MECP2* is disrupted in Rett syndrome and autism. *Hum. Mol. Genet,* 18 (3): 525-534, (2009).
45. Hu V.W, Frank B.C, Heine S, Lee N.H and Quackenbush J, Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. *BMC Genomics,* 7: 118, (2006).
46. Zeev B.B, Bebbington A, Ho G, Leonard H, de Klerk N, Gak E, Vecsler M and Christodoulou J, The common BDNF polymorphism may be a modifier of disease severity in Rett syndrome. *Neurology,* 72(14): 1242-7, (2009).