BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(12)

Systems biological approach to investigate the lack of familial link between Down's Syndrome & Neural Tube Disorders

PK Ragunath* & PA Abhinand

Department of Bioinformatics, Sri Ramachandra University, Porur, Chennai – 600 116, India; PK Ragunath – Email: sru.bioinformatics.research2@gmail.com; Phone: +919841351069, +91-44-45928603; *Corresponding author

Received May 29, 2013; Accepted June 06, 2013; Published July 12, 2013

Abstract:
Systems Biology involves the study of the interactions of biological systems and ultimately their functions. Down's syndrome (DS) is one of the most common genetic disorders which are caused by complete, or occasionally partial, triplication of chromosome 21, characterized by cognitive and language dysfunction coupled with sensory and neuromotor deficits. Neural Tube Disorders (NTDs) are a group of congenital malformations of the central nervous system and neighboring structures related to defective neural tube closure during the first trimester of pregnancy usually occurring between days 18-29 of gestation. Several studies in the past have provided considerable evidence that abnormal folate and methyl metabolism are associated with onset of DS & NTDs. There is a possible common etiological pathway for both NTDs and Down's syndrome. But, various research studies over the years have indicated very little evidence for familial link between the two disorders. Our research aimed at the gene expression profiling of microarray datasets pertaining to the two disorders to identify genes whose expression levels are significantly altered in these conditions. The genes which were 1.5 fold unregulated and having a p-value <0.05 were filtered out and gene interaction network were constructed for both NTDs and DS. The top ranked dense clique for both the disorders were recognized and over representation analysis was carried out for each of the constituent genes. The comprehensive manual analysis of these genes yields a hypothetical understanding of the lack of familial link between DS and NTDs. There were no genes involved with folic acid present in the dense cliques. Only – CBL, EGFR genes were commonly present, which makes the allelic variants of these genes – good candidates for future studies regarding the familial link between DS and NTDs.

Keywords: Bioinformatics, Systems Biology, Down's syndrome, Neural Tube Disorders, Folate metabolism.

Abbreviations: NTD - Neural Tube Disorders, DS - Down's Syndrome, MTHFR- Methylenetetrahydrofolate reductase, MTRR– 5-methyltetrahydrofolate-homocysteine methyltransferase reductase.

Background:
Systems biology involves understanding biological systems in a holistic way. It deals with interactions between biological systems and consequently their functions. A thorough understanding of systems structure is critical in such a study. Quantitative and qualitative methods such as modeling gene regulatory, biochemical networks, flux analysis etc are key constituents of systems approach **[1].** Systems biology provides a framework for assembling models of biological systems from systematic measurements **[2, 3].**

Down's syndrome (DS) is one of the most common genetic disorders which affect about 1 in 800 live births across the globe. The disorder is caused by a total, or occasionally partial, triplication of chromosome 21 resulting in a multifarious and capricious phenotype **[4].** The disorder is primarily

characterized by cognitive and dysfunction of verbal communication along with neuromotor and sensory deficit. The neuropathology of the disorrder is chiefly characterized by reduced brain size and weight along with abnormal gyrification and neurogenesis **[5].** Two main theories have been hypothesized to explain the mechanism by which trisomy 21 leads to the DS phenotype. The 'developmental instability' theory hypothesizes a dosage disparity on the entire chromosome 21, which interrupts various developmental pathways **[6].** The other theory - 'gene-dosage' theory suggests increased dosage for certain genes on 21st chromosome adds more directly to different manifestations of the disorder **[7].**

Neural Tube Disorders (NTDs) are a collection of inborn malformations of the central nervous system and adjacent structures related to flawed neural tube closure during the first trimester of pregnancy occurring usually between 18-29 days of gestation **[8].** Principally Ectodermal and mesodermal malformations concerning the skull and vertebrae can arise as an effect of faults in neural tube closure **[9].** NTDs: are classified into open, and closed types.

Numerous studies have yielded considerable evidence that abnormal folate and methyl metabolism are associated with inception of Down's syndrome. The abnormalities in folate metabolism are implicated with DNA hypomethylation, which in turn is associated with chromosomal instability, improper chromosomal segregation and consequently aneuploidy **[10] [11].** At the same time studies also point to impaired folate status in mothers of children born with neural tube disorders. Several genetic investigations have revealed more than the expected frequency of certain mutations in the genes coding for Methylenetetrahydrofolate reductase and 5-methyltetrahydrofolate-homocysteine methyltransferase reductase proteins **[12, 13].** At the molecular level MTHFR 677C-T polymorphism is recognized as the foremost genetic risk factor for NTDs. Homozygosity for this allele has been recognized to be very prevalent in NTD parents and their off springs, in comparison to controls **[14].**

There is a possible common etiological pathway for both NTDs and Down's syndrome. Numerous epidemiological characteristics are very common between NTDs and DS such as large maternal contribution to the risk of occurrence, differences between ethnic groups, reliance on maternal age at the time of pregnancy, high probability of miscarriages. If there were a causal link, the two conditions should arise more often in affected families than in the population in general. But there is little epidemiological evidence showing familial link between the diseases **[13].**

Gene expression microarrays can provide quantitative information on the status of a cell in a particular condition and point in time. Gene regulatory networks based on microarrays can be pervasive and can be instrumental in getting an insight into pathological roots of a given clinical conditions. This kind of network data extends and compliments a great deal of other information available in the biomedical sciences. The gene network provides ample knowledge on not only the physical interaction between two genes but also about indirect regulation via proteins, metabolites and ncRNA **[15].**

The study will involve gene expression profiling of micro array data sets available on public domain databases for NTDs and Down's syndrome and identify genes whose expression levels are significantly altered in these clinical conditions in comparison with the control. Consequently gene regulatory networks can be built on basis of the basis of the differential expression profile and can be used to understand the complex interactions underlying the pathogenesis of the two disorders. The major goal of the study was to understand the complex molecular interactions which are central to the pathogenesis of both NTDs and Down's syndrome and thereby draw inferences regarding the lack of familial co morbidity between the two disorders.

Methodology:

Dataset Collection

A comprehensive and thorough survey for all differential gene expression studies on Down's syndrome and neural tube defects was carried out. Only those studies conducted on samples from human subjects were considered (Till August 2012). Only one study for each of the disorders was found to fulfill the selection criteria. The platform files – GPL570 & GPL90 for NTDs and DS was downloaded and the expression datasets for each of the above mentioned studies were downloaded in order to be subjected to gene expression profiling analysis. Furthermore, all literature pertaining to studies on familial links between neural tube defects and Down's syndrome was searched for, across various populations.

Gene expression analysis & Filtering

The microarray analysis was carried out using R/Bioconductor **[16],** open source software for the analysis of genomic data. The datasets were normalized to standardize microarray data to facilitate demarcation between real variations in gene expression levels and variations due to the measurement procedure. All microarrays CEL file involved in our study was processed using RMA algorithm and normalized based on quartile array. Further the gene expressions were log transformed to determine the fold change and their significance was measured by standard t-test. The genes were filtered based on fold changes. The fold changes in gene expression levels between the disease samples control samples to check for the differential expression **[17].** Genes which ere differentially up regulated by 1.5 fold were filtered out and their gene ontology was identified.

Gene Network Construction

Two gene networks were constructed using BisoGenet plugin **[18]** for Cytoscape for the two disorders – neutral tube defects and Down's syndrome to explore the molecular factors involved in the etiological pathway underlying the pathogenesis of the disorders and thereby derive plausible reasons for the lack of familial link between the two disorders .The Networks were generated taking as input an initial list of identifiers of genes filtered out on basis of fold change.

Network Analysis & Recognizing Dense Cliques

The network obtained from the BisoGenet Server is analyzed using the plugin Network Analyzer, which computes the degree - its clustering coefficient, the number of self-loops and a

variety of other parameters for every node in a gene regulatory network. Topological parameters such as - the number of nodes, edges, and connected components, the network diameter, radius, density, centralization, heterogeneity, and clustering coefficient, the characteristic path length, and the distributions of node degrees, neighborhood connectivity, average clustering coefficients, and shortest path lengths are calculated & exhibited by this plug-in **[19].**

A complex gene network that emerged as a resultant of the interaction between significantly upregulated or down regulated genes was broken down into smaller sub networks using mcode module of cytoscape. mcode module of cytoscape. Allegro MCODE plugin finds densely connected regions of a network Graph theoretic based clustering algorithm. It works 3 stages - Network Weighting, Complex Detection, and Optional Post-processing **[20].** The over representation analysis of the genes present in the top ranked dense clique for both NTDs and DS was performed in order to map it to their gene ontology.

Results & Discussion:

The etiological connection between NTDs and DS, especially pertaining to their link to in born errors in folic acid metabolism, makes it imperative that familial link would exist between these disorders. But contrary to this hypothesis, studies carried out in the past have indicated otherwise. Several studies in the past have provided considerable evidence that abnormal folate and methyl metabolism are associated with onset of Down's syndrome. At the molecular level - MTHFR 677C-T polymorphism of the gene was the first genetic risk factor for neural tube defects identified at the molecular level. But, various research studies over the years have indicated very little evidence for familial link between the two disorders. *Barkai* and colleagues (2003) reported significantly high frequency of Down's syndrome in pregnancies at high risk of NTD; even this has been attributed to biased selection of study participants due to incomplete ascertainment of individuals **[21].** Marcia *et al.* found no association occurred between families at risk of neural tube defects and those at risk of Down's syndrome with their studies on Latin American countries. Källén *et al*. found no association with anencephaly, spina bifida, cephalocele, or hydrocephalus.in their studies on 5581 cases of Down's syndrome **[22]**

For long, it has been hypothesized that "if there were a causal link, the two conditions [neural tube defects and Down's syndrome] should arise more often in affected families than in the population in general" **[21].** But several research projects have only yielded inconsistent information about the actual frequency of functional mutations in genes associated with folate metabolism in the mothers of individuals with neural tube defect or Down's syndrome. The gene expression profiling analysis on microarray datasets for both disorders clearly showed significant alteration with the expression level of genes involved in folic acid metabolism.

The research focused on identifying the genes and their interactions which are central to the pathogenesis of NTDs and Down's syndrome and thereby recognize the reasons for the limited familial link between the two clinical conditions.

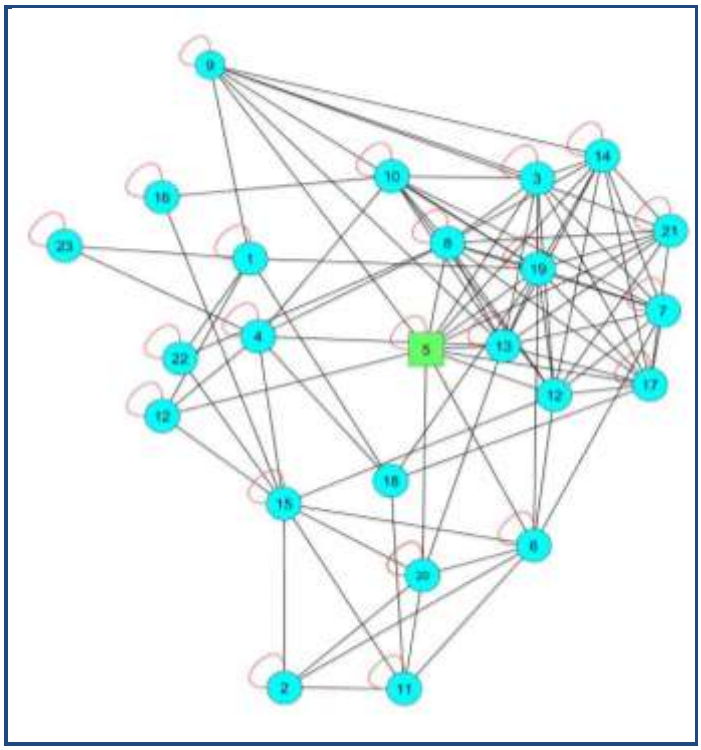

Figure 1: Illustration of the top ranked gene dense clique for DS. Nodes are represented in cyan blue circle, depicting a single gene; the edges which depict interactions between the genes are represented as black lines. The edges represented in red denote possible self interaction between the gene products**.** The node represented as green square represents the gene which is commonly present in the dense cliques for both NTDs and DS. The numbers on each node corresponds to the S.no: given in the **Table 1 (see supplementary material).**

Gene Expression Profiling Studies

The microarray datasets were subjected to gene expression profiling. Microarray technology is highly useful in recognizing the co-regulated genes, pathways, and systems facilitating a deep insight about the transcriptome. Various research have indicated that changes in the significance level of differential expressed gene products along with the fold change cut-offs, can give very different results that imply different signaling pathways and functions involved **[23].** T-tests were used to identify deviation from the mean, large sampling sizes can have an impact on the number of false positives and may yield little information, if anything about the biology. Fold change on the other hand lends itself to a more biologically meaningful assessment, **[24].**

Initial filtering of the genes was performed on the basis of fold change in the expression levels and the p-value. Only genes which were up regulated by least 1.5 fold with a p-value of lesser than 0.05 were chosen for analysis.

Gene Interaction Network

The gene network for Down's syndrome was much bigger than the one for NTDs owing to the distinct difference in the number of genes which were obtained as a result of the initial filtering process. But, interestingly this gene remained isolated, with no notable interaction with any other gene or their protein

products. The gene network constructed for Down's syndrome consisted of 539 nodes and 4547 edges & for neural tube defects totally 29 nodes and 80 edges. The gene network obtained for Down's syndrome was considerably larger than the one obtained for Neural Tube Defects. The smaller size of NTD gene network can be attributed to the smaller number initial genes (obtained from gene expression profiling and subsequent filtering based). The graphical representation of the top ranked dense cliques for DS & NTDs are displayed in **(Figure 1 & 2)** respectively.

Allgro. Mcode module was employed for identifying the top ranked dense clique. The genes in the dense clique are the most interconnected and therefore, must be center of the etiological pathway underlying the two diseases. The genes which are part of the top ranked dense cliques for DS & NTDs are displayed in **Table 1 & 2 (se supplementary material)** respectively. Over representation analysis, based on their respective ontology was carried out for all the genes which constitute the top ranked dense clique for both NTDs and Down's syndrome.

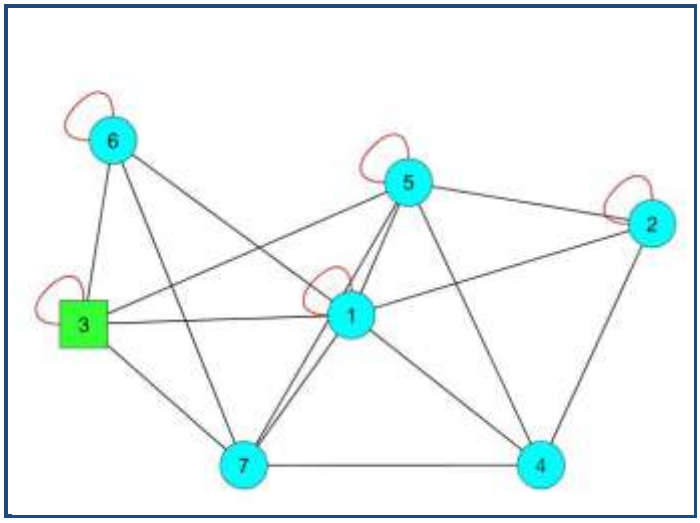

Figure 2: Illustration of the top ranked gene dense clique for NTDs. The nodes are represented in cyan blue circle, depicting a single gene; the edges which depict interactions between the genes are represented as black lines. The edges represented in red denote possible self interaction between the gene products. The node represented as green square represents the gene which is commonly present in the dense cliques for both NTDs and DS. The numbers on each node corresponds to the S.no: given in the **Table 1 (se supplementary material).**

Inferences from the Gene Networks

The gene networks constructed on the basis of gene express profiling provides us a hypothetical insight into the pathology of both the disorders. No genes involved in folic acid metabolism were a part of the top ranked dense clique for both NTDs and Down's syndrome.Notably XIST was the top up regulated gene for both neural tube defects and Down's syndrome, with more than 4 fold up regulation in both cases. Xist (X-inactive specific transcript) is a RNA gene, present on the X chromosome of the placental mammals, acts as major effector of the X inactivation process. It is a component of the *Xic* - X-chromosome inactivation centre **[25].** But,uniquely the gene Xist remained unconnected to other nodes in the network-thereby indicating lack of interaction with other members of the gene network.

DS individuals present higher cerebral cortex and cerebellum protein levels of the proapoptotic genes Fas and p53. Altered apoptosis has been suggested as one of the mechanisms responsible for different DS phenotypes. The most prominent feature of DS is cognitive disability, which is likely to be partially due to widespread brain hypo-cellularity. Although neuronal cultures from human fetal and mouse models of DS brains show enhanced apoptosis, different studies have demonstrated that apoptosis has a prominent role in other important DS phenotypes, such as neurodegeneration in later life stages, impaired retinal development, heart anomalies, immunological alterations and predisposition to the development of different types of cancers **[26].**

The construction of separate gene regulatory networks for NTDs and DS yield an hypothetical understanding of the pathogenisis of the diseases and the lack of familial link between the two disorders.

Conclusion:

The neural tube defects and Down's syndrome defects are amongst the most birth defects across the globe. Though, it has been hypothesized that there is a common etiological pathway underlying the two disorders, especially related to the folic acid metabolism - there has been little evidence that points to familial link between the two disorders. Our study aimed at investigating the gene gene interactions involved in the pathogenesis neither two disorders and thereby draw inferences regarding the lack of familial link between the two. We generated gene regulatory networks for NTDs and DS based on their gene expression profiling and consequently, recognizing top ranked dense clique from the gene regulatory networks. Only the genes - EGFR, CBL were found to be common between NTDs and DS. In future studies can be carried out to investigate the allelic variants of these genes and Meta analysis can be carried out to study their association with NTDs and DS. The gene regulatory network gives us a picture of all the interactions at molecular level which conspire and combine to bring about the pathogenesis of the two disorders.

References:

[1] http://www.systems-biology.org/
[2] Vidal M, *FEBS letters.* 2009 **583**: 3891 [PMID: 19913537]
[3] Chuang HY *et al. Annu Rev Cell Dev Biol.* 2010 **26**: 721 [PMID: 20604711]
[4] Lockstone HE *et al. Genomics.* 2007 **90**: 647 [PMID: 17950572]
[5] Reeves RH *et al. Trends Genet.* 2001 **17**: 83 [PMID: 11173117].
[6] Shapiro BL, *Hum Genet.* 1997 **99**: 421 [PMID: 9050934].
[7] Korenberg JR *et al. Am J Hum Genet.* 1990 **47**: 236 [PMID: 2143053]
[8] http://www.nlm.nih.gov/cgi/mesh/2013/MB_cgi?field=uid&term=D009436
[9] http://www.diseasesdatabase.com/ddb8926.htm
[10] Van der Put NM *et al. QJM.* 1997 **90**: 505 [PMID: 9327028]
[11] Jacob RA *et al. J nutr.* 1998 **128**: 1204 [PMID: 9649607]
[12] Fenech M, *Mutat Res.* 2001 **480**: 51 [PMID: 11506798]

[13] Steegers-Theunissen RP *et al. Metabolism.* 1994 **43**: 1475 [PMID: 7990699]
[14] Smoot ME *et al. Bioinformatics.* 2011 **27**: 431 [PMID: 21149340].
[15] Bansal M *et al. Mol syst Biol.* 2007 **3**: 78 [PMID: 17299415]
[16] Gentleman RC *et al. Genome Biol.* 2004 **5**: R80 [PMID: 15461798]
[17] Smoot ME *et al. Bioinformatics.* 2011 **27**: 431 [PMID: 21149340]
[18] Martin A *et al. BMC bioinformatics.* 2010 **11**: 91 [PMID: 20163717]
[19] http://med.bioinf.mpi-inf.mpg.de/netanalyzer/help/2.7/index.html
[20] Bader GD & Hogue CWV, *BMC bioinformatics.* 2003 **4**: 2 [PMID: 12525261]
[21] Barkai G *et al.* 2003 **361**: 1331 [PMID: 12711468]
[22] Amorim MR *et al. BMJ.* 2004 **328**: 84 [PMID: 14662523]
[23] Dalman MR *et al. BMC bioinformatics.* 2012 **13**: S11 [PMID: 22536862]
[24] Jeffery IB *et al. BMC bioinformatics.* 2006 **7**: 359 [PMID: 16872483]
[25] Chow JC *et al. Annu Rev Genomics Hum Genet.* 2005 **6**: 69 [PMID: 16124854]
[26] Rueda N *et al. Apoptosis.* 2013 **18**: 121 [PMID: 23224708]

Edited by P Kangueane
Citation: **Ragunath & Abhinand,** Bioinformation 9(12): 610-616 (2013)

BIOINFORMATION

open access

Supplementary material:

Table 1: Gene list: Top-ranked Dense clique from Allegro MCODE for Down's syndrome

S.No: Node Id	Gene Symbol	Gene Name	Gene Ontology
1	ARRB2	Arrestin, Beta 2	Sensory Perception, Signal Transduction
2	BID	Bh3 Interacting Domain Death Agonist	Death Receptor Binding, Apoptotic Mitochondrial Changes, Induction Of Apoptosis Via Death Domain Receptors
3	CBL	Cas-Br-M (Murine) Ecotropic Retroviral Transformin	Calcium Ion Binding, Ligase Activity, Protein Binding, Signal Transducer Activity, Transcription Factor Activity,Ubiquitin-Protein Ligase Activity, Zinc Ion Binding, Cell Surface Receptor Linked Signal Transduction.
4	CUL3	Cullin 3	G1/S Transition Of Mitotic Cell Cycle,Cell Cycle Arrest, Induction Of Apoptosis By Intracellular Signals, Positive Regulation of Cell Proliferation
5	EGFR	Epidermal Growth Factor Receptor (Erythroblastic L	ATP Binding, Epidermal Growth Factor Receptor Activity, Receptor Activity, Cell Proliferation, Epidermal Growth Factor Receptor Signaling Pathway, Negative Regulation of Cell Cycle, Protein Amino Acid Phosphorylation
6	FAS	Fas (TNF Receptor Superfamily, Member 6)	Protein Binding,Transmembrane Receptor Activity, Anti-Apoptosis
7	FYN	FYN Oncogene Related To SRC, FGR, YES	ATP Binding, Non-Membrane Spanning Protein Tyrosine Kinase Activity, Protein Binding, Calcium Ion Transport, Intracellular Signaling Cascade, Protein Amino Acid Phosphorylation,Protein Kinase Cascade
8	GRB2	Growth Factor Receptor-Bound Protein 2	SH3/SH2 adaptor activity,Ras protein signal transduction
9	IGF1R	Insulin-Like Growth Factor 1 Receptor	Insulin Receptor Signaling Pathway,Positive Regulation of Cell Proliferation,Protein Amino Acid Phosphorylation
10	KHDRBS1	KH Domain Containing, RNA Binding, Signal Transduction	DNA/RNA binding,SH3/SH2 adaptor activity,transcriptional repressor activity,G2/M transition of mitotic cell cycle arrest
11	MYC	V-Myc Myelocytomatosis Viral Oncogene Homolog (Avi	protein binding,transcription factor activity,iron ion homeostasis,regulation of transcription from RNA polymerase II promoter
12	PLCG1	Phosphoinositide Phospholipase C-Gamma	Epstein-Barr virus infection,Metabolic pathways,VEGF signaling pathway,NF-kappa B signaling pathway,Leukocyte transendothelial migration
13	PIK3R1	Phosphoinositide-3-Kinase, Regulatory Subunit 1 (P	kinase activity, phosphatidylinositol 3-kinase activity
14	PTPN11	Protein Tyrosine Phosphatase, Non-Receptor Type 11	hydrolase activity,non-membrane spanning protein tyrosine phosphatase activity,intracellular signaling cascade,protein amino acid dephosphorylation
15	RHOA	Ras Homolog Gene Family, Member A	GTP binding, magnesium ion binding,signal transducer activity,Rho protein signal transduction,actin cytoskeleton organization and biogenesis,positive regulation of I-kappaB kinase/NF-kappaB cascade
16	SMAD2	SMAD, Mothers Against DPP Homolog 2 (Drosophila)	protein binding,regulation of transcription, DNA-dependent signal transduction

17	SOS1	Son Of Sevenless Homolog 1 (Drosophila)	DNA binding,Ras guanyl-nucleotide exchange factor activity,Rho GTPase activator activity
18	SPTAN1	Spectrin, Alpha, Non-Erythrocytic 1 (Alpha-Fodrin)	actin binding,calcium ion binding,calmodulin binding,structural constituent of cytoskeleton,barbed-end actin filament capping
19	SRC	V-Src Sarcoma (Schmidt-Ruppin A-2) Viral Oncogene	ATP binding,SH3/SH2 adaptor activity,protein binding,protein-tyrosine kinase activity
20	VIL2	Villin 2 (Ezrin)	cytoskeletal protein binding,structural molecule activity,cytoskeletal anchoring,regulation of cell shape
21	ERBB2	Erythroblastic Leukemia Viral Oncogene Homolog	Stabilizing ligand binding and enhancing kinase-mediated activation of downstream signalling pathways
22	MAP3K1	Mitogen-Activated Protein Kinase Kinase Kinase	Signal transduction cascade
23	HNRPD	Heterogeneous Nuclear Ribonucleoprotein D	Nucleic acid binding proteins and they complex with heterogeneous nuclear RNA

Table 2: Gene list: Top-ranked Dense clique from Allegro MCODE for Neural Tube Defects

S.No:/ Node Id	Gene Symbol	Gene Name	Gene Ontology
1	CBL	Cas-Br-M (Murine) Ecotropic Retroviral Transformin	Calcium Ion Binding, Ligase Activity, Protein Binding, Signal Transducer Activity, Transcription Factor Activity,Ubiquitin-Protein Ligase Activity, Zinc Ion Binding, Cell Surface Receptor Linked Signal Transduction.
2	CD3Z	CD3Z antigen, zeta polypeptide (TiT3 complex)	Protein Binding, Protein Homodimerization Activity, Receptor Signaling Protein Activity
3	EGFR	Epidermal Growth Factor Receptor (Erythroblastic L	ATP Binding, Epidermal Growth Factor receptor Activity, Receptor Activity, Cell Proliferation, Epidermal Growth Factor Receptor Signaling Pathway, Negative Regulation of Cell Cycle,
4	LAT	Linker for activation of T cells	SH3/SH2 Adaptor Activity,Ras Protein Signal Transduction,Calcium-Mediated Signaling,Integrin-Mediated Signaling Pathway
5	ZAP70	Zeta-chain (TCR) associated protein kinase 70kda	ATP Binding,Protein-Tyrosine Kinase Activity,Protein Amino Acid Phosphorylation
6	PGFRB	Platelet-Derived Growth Factor Receptor Beta	Regulation of embryonic development, cell proliferation, survival, differentiation, chemotaxis and migration
7	VAV	Guanine Nucleotide Exchange Factor	Role in transcriptional control, angiogenesis