



Systematic Analysis of Integrated Gene Functional Network of Four Chronic Stress-related Lifestyle Disorders

Souvick Roy, Abhik Chakraborty, Chinmoy Ghosh, Birendranath Banerjee

Molecular Stress and Stem Cell Biology Group, School of Biotechnology, KIIT University, Bhubaneswar, Odisha-751024, India

ABSTRACT

Background: Stress is a term used to define factors involved in changes in the physiological balances resulting in disease conditions. Chronic exposure to stress conditions in modern lifestyles has resulted in a group of disorders called lifestyle disorders. Genetic background and environmental factors are interrelated to lifestyle in determining the health status of individuals. Hence, identification of disease-associated genes is the primary step toward explanations of pathogenesis of these diseases. In functional genomics, large-scale molecular and physiological data are used for the identification of causative genes associated with a disease. **Aim:** The objective of our study was to find a common set of genes involved in chronic stress-related lifestyle diseases such as cardiovascular diseases (CVDs), type 2 diabetes (T2D), hypertension (HTN), and obesity.

Materials and Methods: In our study, we have performed a systematic analysis of the functional gene network of four chronic stress-related lifestyle diseases by retrieving genes from published databases. We have tried to systematically construct a functional protein-protein interaction (PPI) network. The goals of establishing this network were the functional enrichment study of interacting partners as well as functional disease ontology annotation (FunDO) of the enriched genes.

Results: This study enabled the identification of key genes involved in these stress-related lifestyle diseases by prioritizing candidate genes based on their degree of involvement. In this systematic analysis, we have found key genes for these diseases based on their involvement and association at the gene network level and PPI.

Conclusion: We have deciphered a group of genes that in combination play a crucial role and may impact the function of the whole genome in the four lifestyle disorders mentioned.

Key words: Chronic stress, cardiovascular disease (CVD), hypertension (HTN), lifestyle, obesity, type 2 diabetes (T2D)

Introduction

Biological stress refers to the consequences of failure of any organism to respond appropriately to physical and chemical cues from the environment.^[1] In the current working-lifestyle scenario of modern living, there has been reported an increase in the incidence of cardiovascular diseases (CVDs).^[2,3] The CVDs are interrelated with diabetes, hypertension (HTN), and obesity.^[4] There is increasing evidence that type 2 diabetes (T2D) and CVDs are partially related with respect to their etiologies and pathophysiological mechanisms. The prediction models presently applied share numerous risk factors: Age; sex; and anthropometric, metabolic, socioeconomic, and lifestyle variables.^[5]

T2D is a complex, heterogeneous group of conditions characterized

by elevated levels of plasma glucose, caused by impairment in both insulin secretion and action. The recent global epidemic of T2D is an indication of environmental factors as well as rapid changes in lifestyle patterns and increasing physical inactivity or sedentary habits.^[6] Studies in Asian Indians have suggested that even a moderate degree of obesity can produce insulin resistance, due to increase in abdominal fat accumulation in this group.^[7] Unusual fat distribution characterized by a high waist-to-hip ratio or a high truncal-to-peripheral skin fold thickness ratio gives a clear indication that an individual is predisposed to developing insulin resistance.^[8]

Chronic HTN, which has become part of the professional world, is also known to cause CVDs.^[9] HTN [according to the World Health Organization (WHO) criteria, the elevation of blood pressure level, exceeding 140/90 mmHg] is a complex disorder, with genetic, environmental, and demographic factors contributing to its occurrence.^[10] It is one of the principal independent risk factors for stroke, myocardial infarction, end-stage renal diseases, and ventricular hypertrophy.^[9]

In the postgenomic era, we have acquired tools to understand common pathophysiological phenomena at the level of genomics. Functional genomics have enabled us to use large-scale molecular and physiological data for not only the identification of causative genes associated with a disease but also the discovery of gene modules that directly respond to genetic and environmental perturbations associated with a disease.^[11-14]

Genes associated with different diseases are correlated based on their functional importance, and a scoring system is generated based on fuzzy logic.^[15] There are numerous methods to study such correlations, such as high-throughput data analysis or protein-protein interactions

Access this article online	
Quick Response Code: 	Website: www.genome-integrity.org
	DOI: 10.4103/2041-9414.155952

Address for correspondence:

Dr. Birendranath Banerjee, Molecular Stress and Stem Cell Biology Group, KIIT School of Biotechnology, KIIT University, Bhubaneswar - 751 024, Odisha, India. E-mail: bnbanerjee@kiitbiotech.ac.in

(PPIs) and constructions of gene regulatory networks.^[16-19] Moreover, methods have also been proposed to integrate multiple data sources for the purpose of achieving highly accurate identification of genes involved in diseases or biological processes.^[20] A general attribute of these methods is the requirement of a set of genes associated with a query disease to obtain novel associations between the query disease and candidate genes. However, according to the recent release of the Online Mendelian Inheritance in Man (OMIM) database,^[21] the genetic bases for a significant proportion of known diseases are completely unknown, and thus applications of these methods are greatly constrained. To overcome this constraint, methods have been proposed to utilize disease phenotype similarity data with PPI network-based data for the prioritization of candidate genes.^[22,23] There is increasing evidence where genome-wide association studies have combined with network analysis to improve the understanding of the molecular basis of complex diseases.^[24,25] The candidate genes can be prioritized based on gene semantic similarities obtained from the gene ontology.^[26]

In this study, we have tried to simplify the complexity of the conventional pathway analysis approaches of protein-protein networks or gene network involvement in pathways. We have tried to correlate the four most common lifestyle and stress-associated disorders (CVDs, T2D, obesity, and HTN) and tried to study their common enriched gene networks based on functionality and interactions.

Materials and Methods

The schematic representation of this study is summarized in Chart 1.

Retrieval of genes and prioritization

In order to build a functional network of the genes involved in four chronic stress-related lifestyle diseases, that is, genes associated with T2D, obesity, HTN, and CVDs were retrieved from the PolySearch database. PolySearch is a comprehensive text mining system that extracts the relationships between diseases, genes, mutations, drugs, and metabolites in humans from different types of biomedical text databases, such as PubMed, OMIM, DrugBank, SwissProt, Human Metabolome Database (HMDB), Human Protein Reference Database (HPRD), and Genetic Association Database (GAD).^[27] Searching was performed using “Disease-Gene/Protein Association” as query type and individual disease name, such as “Type 2 diabetes,” as the query keywords. Searching for the rest of the diseases was carried out in a similar manner. The obtained gene list was prioritized based on a relevancy score, expressed as a z score [Supplementary 1-4]. In our study we have selected genes having z scores >1 for further analysis.

Functional disease ontology annotations (FunDOs) of the retrieved genes

All the genes retrieved from the PolySearch database that were involved in the four chronic stress-related lifestyle diseases under discussion were annotated using FunDO (<http://django.nubic.northwestern.edu/fundo>). Disease ontology annotations of a human gene describe unique roles for genes in the context of disease and are complementary to gene ontology annotations. This process enabled us to find the relevant diseases with respect to the entered list of genes, based on statistical analysis from the disease ontology annotation database.

Construction of PPI networks

The genes obtained from the PolySearch database (having z score >1) were further used for construction of the PPI networks by using the STRING 9.1 (<http://string-db.org>) database.^[28] This database provides known and predicted PPIs from different sources on the basis of their neighborhood, gene fusions, co-occurrence, coexpression, experiments, and literature sources. An extended network was constructed for each

individual disease based on a high confidence score (>0.7) as selection parameter, which implies that only interactions with high level of confidence were extracted from the database and considered as valid links for the PPI networks.

Topological analysis of the PPI networks

The PPI networks obtained by using STRING 9.1 were analyzed based on topological parameters such as betweenness centrality (BC) and node degree using a Cytoscape (<http://www.cytoscape.org>) plug-in named “Network Analyzer.”^[29] In these networks, each gene corresponded to a node and the interaction between the nodes were defined as edges. “Degree” signifies the number of links or edges the node has to the other nodes in a network. The nodes that possess high degree represented the genes having important biological functions. BC reflects the importance of the node based on the number of shortest paths that pass through each node. The network for each disease was visualized based on these parameters, where we mapped the node degree to the node size and betweenness to the node color in the network view. The genes having high BC value and node degree (BC ≥ 0.05 and node degree ≥ 5) for each individual disease were taken into consideration for further analysis.

Functional enrichment analysis

In the present study, after setting up the selection parameter based on the BC values and node degrees (BC ≥ 0.05 and node degree ≥ 5) of the genes involved in these four diseases, functional enrichment analysis was performed using ClueGO.^[30] ClueGO, a Cytoscape plug-in, was used to decipher functionally grouped gene ontology and pathway annotation networks. It facilitates the visual interpretation of functionally related genes as a clustered network and chart. The statistical test used for the enrichment analysis was a right-sided hypergeometric test with a Benjamini-Hochberg correction and a kappa score of 0.5 (medium). It enabled us to elucidate the biological functions of the genes, their regulation, their involvement in pathways, and their relevance to diseases.

FunDOs of enriched genes involved in the top five diseases

The enriched gene list was obtained using the stringent selection criteria of BC ≥ 0.05 and node degree ≥ 5 , and they were further annotated using FunDO. This enabled us to know about the involvement of enriched genes in the top five diseases.

Results

Gene retrieval and prioritization

The genes retrieved from the PolySearch database source involved in these four chronic stress-related lifestyle diseases were prioritized based on a z score >1. In order to exclude outcomes due to chance, the genes having positive z scores (>1) were considered. The results showed 51, 82, 70, and 71 genes involved in CVDs, HTN, obesity, and T2D, respectively, on the basis of their z scores.

Network of retrieved genes involved in four chronic stress-related lifestyle diseases

A list of 141 genes was analyzed by FunDO, out of which 131 genes were found to be associated with the diseases. It was observed that these enriched genes were involved in diabetes mellitus; obesity; HTN; atherosclerosis, which is one of the primary reasons of CVD; and anorexia nervosa, which is associated with obesity [Figure 1]. The sizes of the disease nodes are proportional to the number of edges. The numbers of genes involved in these top five diseases are shown in Table 1.

PPI networks

PPIs are the physical contacts established between two or more proteins

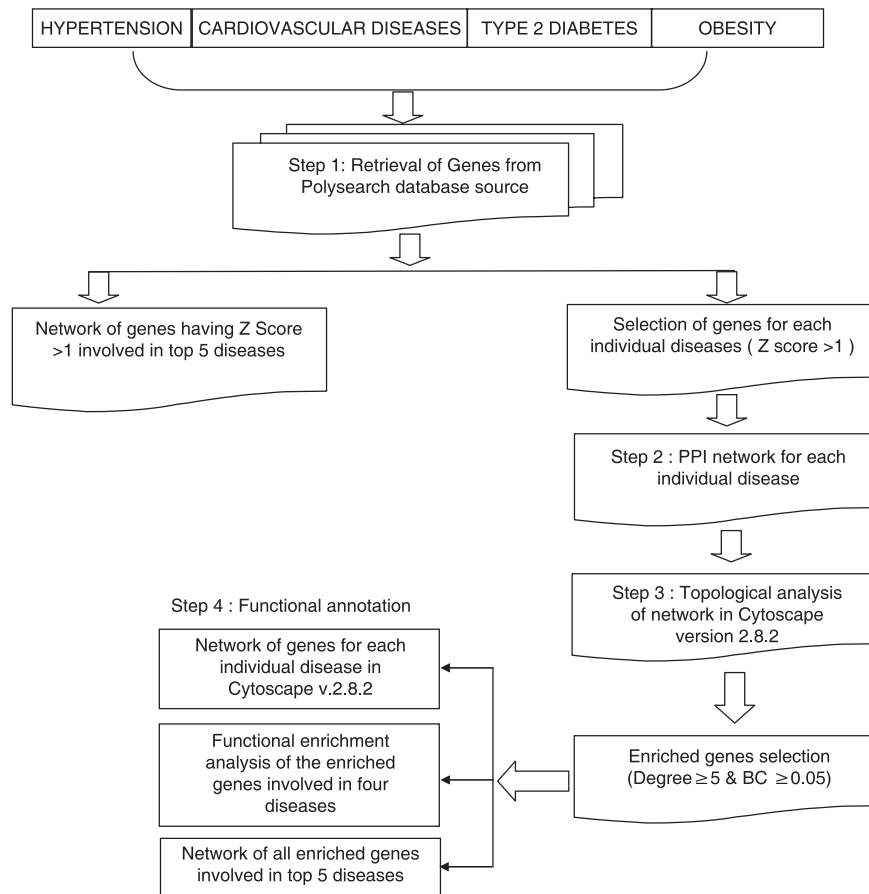


Chart 1: A pictorial representation of the methodology has been summarized as follows: Step 1- Retrieval and prioritization of genes involved in four chronic stress-related lifestyle diseases from the PolySearch text mining tool; Step 2- Analysis of PPI network by using STRING database; Step 3- Topological analysis of network using Cytoscape v 2.8.2, based on BC and node degree, leading to the identification of enriched genes involved in these four diseases; Step 4- Functional annotation of the enriched genes involved in four chronic stress-related lifestyle diseases

due to biochemical events. In our study, the generation of PPI networks with the retrieved genes resulted in 130 interactions among 26 nodes for genes involved in CVD, 142 interactions among 42 nodes for HTN, 160 interactions among 36 nodes for obesity, and 204 interactions among 39 nodes for T2D [Figure 2].

Topological analysis of networks and identification of key genes

The PPI network obtained from STRING 9.1 for each individual stress-related lifestyle disease was further analyzed as described in the Materials and Methods section. Here, each gene was represented as a node. Two topological parameters, that is, node degree and BC were used for identification of key genes. The gene with high node degree and high BC value represented the key gene involved in a particular disease. The selection criteria of node degree ≥ 5 and BC ≥ 0.05 were used for identification of key genes involved in these four chronic stress-related lifestyle diseases. In this study, as we have considered five diseases, a node degree ≥ 5 was taken as the cut-off value. In case of BC, a minimum value of ≥ 0.05 was considered for its effective significance in screening the optimum number of genes associated with five diseases. Moreover, a higher BC value would have led to the minimum criteria of gene association. The key genes having a node degree ≥ 5 and BC ≥ 0.05 for each individual disease are represented in Table 2. The networks of key genes for each disease were visualized in “Network Analyzer.” Nodes having high degree were displayed as big circles, while shades of red to green colors represented high to low BC values for the node. The

networks of key genes involved in these stress-related lifestyle diseases are shown in Figure 3.

Molecular function and pathway analysis of key genes

After functional enrichment of the genes, it was observed that these genes are involved in insulin signaling pathways, adipocytokine signaling pathways, glucose homeostasis of the body, the sleep-wakefulness cycle, fat cell differentiation, DNA damage response, PI3K/AKT-mediated signaling pathways, regulation of insulin hormone secretion, tissue homeostasis, the circadian clock, regulation of stem cell proliferation, glycogen metabolism, regulation of lipid biosynthesis process, lipid metabolism, several cancers, etc. [Figure 4].

Functional annotation of the enriched genes involved in four chronic stress-related lifestyle diseases

The enriched or key genes obtained based on the selection criteria (node degree ≥ 5 and BC ≥ 0.05) were annotated by using FunDO. Thirty-six genes were analyzed by the web version of FunDO. It was observed that these enriched genes are involved in diabetes mellitus, obesity, hyperinsulinism, metabolic disease, hyperglycemia, and atherosclerosis, which is one of the primary CVDs [Figure 5]. The sizes of the disease nodes are proportional to the number of edges. The number of genes involved in these diseases was shown in Table 3.

Discussion

Lifestyle disorders are always associated with chronic stress-induced

physiological imbalances. The imbalances generated in the form of hormonal dysregulation or enzyme dysfunction start as low-key events. It was reported earlier that inflammatory cytokines such as interleukin-6 (IL-6) were involved in chronic and systemic inflammation, termed as metaflammation, which was originally ascribed to obesity.^[31] Subsequent research has shown that metaflammation is not limited to obesity but associated with other lifestyle and environmental inducers, which have been linked, either directly or indirectly, to certain chronic diseases and conditions such as heart disease, and T2D.^[32] It appears to be part of a metabolic cascade, including cellular oxidative stress and insulin resistance, which induces allostatic overload, dysmetabolism, and ultimately chronic diseases.^[33]

In our study, with the help of published database searching and other bioinformatics tools, we have derived key genes involved in the four chronic stress-related lifestyle diseases CVD, obesity, T2D, and HTN, which are known to be interlinked in pathophysiological terms. After database mining for each disease, we considered the genes having a z score >1 for each disease, and then by topological analysis we considered the enriched genes involved in each disease, on the basis of node degree ≥5 and BC ≥0.05.

Functional gene enrichment of the diseases, showed involvement of common pathways and biological functions. The schematic representation of the interconnectedness of these diseases, along with the genes involved, is given in Figure 6. We have considered common genes based on their interconnectedness and biological involvement in all the four lifestyle disorders mentioned. The figure shows that *IL6*,

LEP, *CLOCK*, *AGRP*, and *BDNF* were the key genes involved in all these stress-related lifestyle diseases.

In our study, it was observed that one of the key genes, *IL6* (having node degree 16 and BC 0.18641026) [Table 2], which encodes the inflammatory cytokine IL-6, plays a crucial role in atherosclerosis.^[34] It is one of the principal complications found in cardiac disease victims. It is characterized by the deposition of lipids underneath the tunica intima of blood vessels. The primary reasons for this are hyperlipidemia and hypercholesterolemia. Due to the deposition of lipids underneath blood vessels, atherosclerotic plaque is produced,^[35] which can lead to several CVDs such as myocardial infarction, stroke, coronary artery blockage, as well as chronic HTN.^[36] Studies have been reported that suggested that polymorphism in the promoter region of the *IL6* gene leads to T2D.^[37]

T2D is caused by complex interactions between adverse environmental factors and certain genetic factors. Hyperglycemia is one of the primary features of T2D. It is characterized by glycosuria, loss of electrolytes, increased caloric loss, polyphagia, as well as loss of body weight due to loss of calories through the urine, and mobilization of fats and proteins for energy production.^[13] Hyperinsulinemia is another symptom of the early onset of T2D and can occur due to metabolic dysfunction. It is associated with HTN, obesity, dyslipidemia, glucose intolerance, low high-density lipoprotein (HDL) cholesterol levels, elevated triglyceride levels, and impaired fasting glucose, as well as parental history of diabetes.^[38] The *AgRP* (Agouti-related neuro peptide)-coding gene *AGRP* (having node degree 16 and BC 0.08615385) [Table 2], another key gene obtained from our study, is involved in all these four stress-related lifestyle diseases. *AGRP* encodes an antagonist of the melanocortin-3-receptor and the melanocortin-4 receptor. The hypothalamic control of feeding behavior is regulated by melanocortin receptors, and they also help in the regulation of intracellular calcium levels. Mutation in this gene leads to the late onset of obesity.^[39,40]

The *LEP* (leptin) gene (having node degree 37 and BC 0.48282051) [Table 2] was involved in all the four chronic stress-related lifestyle disorders. It encodes an adipocyte-derived hormone, leptin, which regulates feeding behavior and energy expenditure.^[41] In animal models,

Table 1: Number of genes involved in top 5 diseases from the list of key genes

Disease Ontology (DO) term	Number of genes in list
Diabetes mellitus	42
Obesity	27
Atherosclerosis	21
Hypertension	20
Anorexia nervosa	11

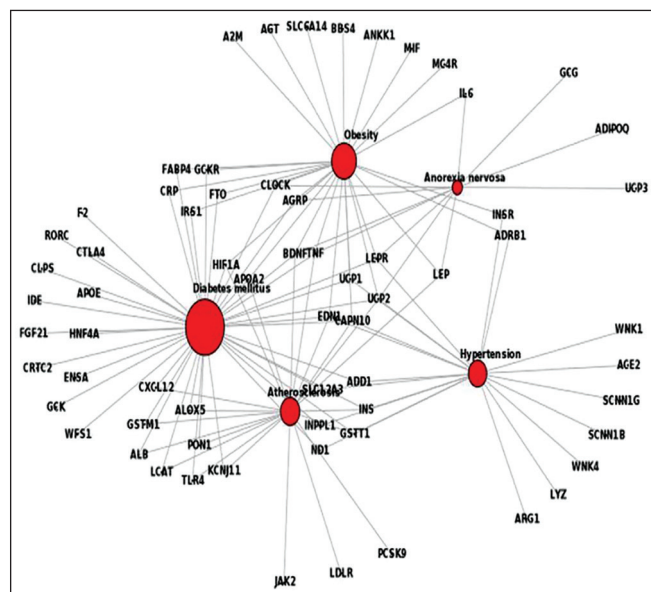


Figure 1: Figure showing network of genes having z score >1 involved in top five diseases using FunDO

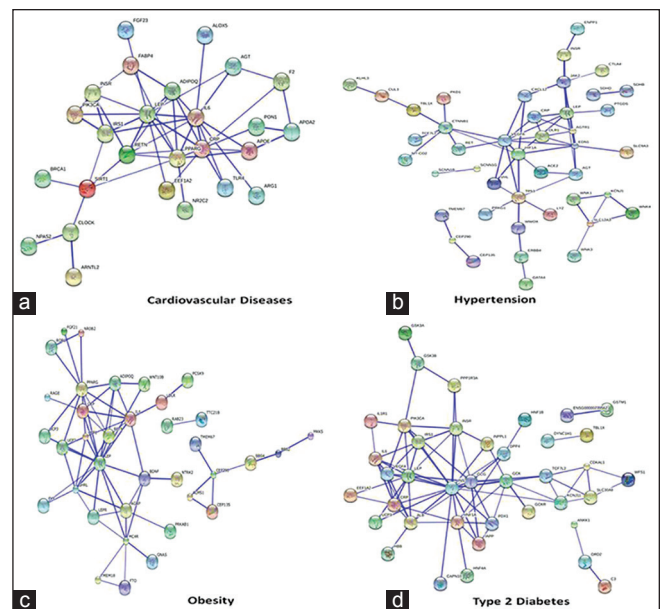


Figure 2: Figure showing protein-protein interaction network based on confidence score (0.7) for each chronic stress related lifestyle diseases. (a) CVDs (b) HTN (c) obesity (d) T2D

Table 2: Key genes involved in chronic stress related lifestyle diseases

Genes involved in cardiovascular diseases (CVD)		
Gene	Betweenness centrality**	Degree*
<i>CLOCK</i>	0.15666667	10
<i>IRS1</i>	0.07874206	15
<i>CRP</i>	0.07692308	11
<i>LEP</i>	0.48282051	37
<i>INSR</i>	0.06666667	8
<i>SIRT1</i>	0.29416667	14
<i>AGT</i>	0.06590038	11
<i>PPARG</i>	0.23666667	24
<i>IL6</i>	0.18641026	16

Genes involved in obesity		
Gene	Betweenness centrality**	Degree*
<i>CEP290</i>	0.73333333	14
<i>BDNF</i>	0.10461538	12
<i>CRP</i>	0.07692308	11
<i>LEP</i>	0.48282051	37
<i>MC4R</i>	0.22051282	17
<i>PPARG</i>	0.23666667	24
<i>IL6</i>	0.18641026	16
<i>LDLR</i>	0.07692308	5
<i>AGRP</i>	0.08615385	16
<i>BBS2</i>	0.33333333	8
<i>BBS4</i>	0.53333333	8

Genes involved in hypertension (HTN)		
Gene	Betweenness centrality**	Degree*
<i>CUL3</i>	0.06666667	6
<i>TBL1X</i>	0.12873563	5
<i>ERBB4</i>	0.13103448	8
<i>CTNNB1</i>	0.36091954	21
<i>VEGFA</i>	0.33889436	22
<i>LEP</i>	0.48282051	37
<i>INSR</i>	0.06666667	8
<i>AGT</i>	0.06590038	11
<i>SLC12A3</i>	0.58333333	9
<i>KCNJ1</i>	0.08333333	7
<i>JAK2</i>	0.08640394	18
<i>TP53</i>	0.36464149	19
<i>HIF1A</i>	0.25064039	27
<i>WVWX</i>	0.1862069	6

Genes involved in type 2 diabetes (T2D)		
Gene	Betweenness centrality**	Degree*
<i>KCNJ11</i>	0.08379928	10
<i>LEP</i>	0.48282051	37
<i>INSR</i>	0.06666667	8
<i>DPP4</i>	0.06523297	10
<i>DRD2</i>	1	5
<i>HNF1A</i>	0.08960573	19
<i>TBL1X</i>	0.12873563	5
<i>IRS</i>	0.07874206	15
<i>INS</i>	0.5116129	49
<i>CRP</i>	0.07692308	11
<i>VEGFA</i>	0.33889436	22
<i>GSK3B</i>	0.06577061	11

Table 2: Continued

<i>GCG</i>	0.09236559	26
<i>ALB</i>	0.07258065	18
<i>IL6</i>	0.18641026	16
<i>GCK</i>	0.08225806	16

Node degree* and betweenness centrality (BC)** are topological parameters used for gene prioritization in the network; a cut-off of BC ≥ 0.05 and/or node degree ≥ 5 were considered for gene prioritization

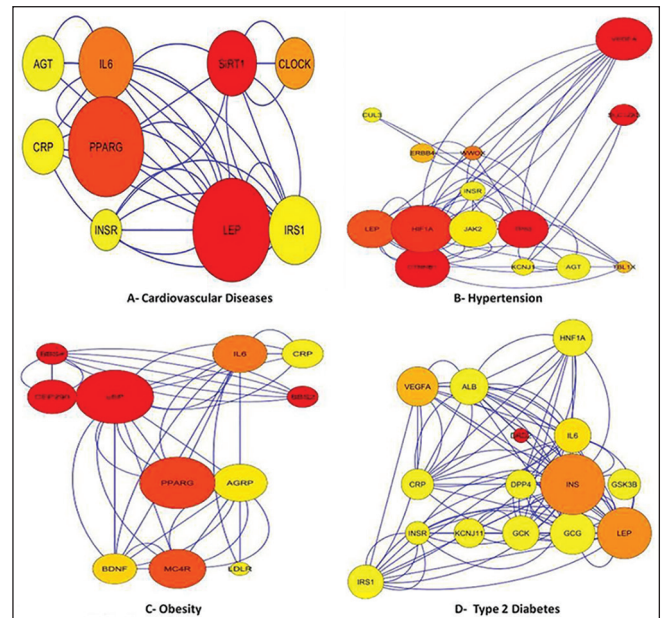


Figure 3: Figure showing network of key genes involved in four chronic stress-related lifestyle diseases based on node degree ≥ 5 and BC ≥ 0.05 in Network Analyzer, a Cytoscape plug-in

mutations in the leptin gene cause severe obesity.^[42] A polymorphic tetranucleotide repeat in the 30-flanking region of the LEP gene (LEP-tet) serves as a useful marker for linkage and association studies between this gene and a number of phenotypes other than obesity, such as noninsulin-dependent diabetes mellitus or T2D, HTN, and insulin resistance syndrome.^[43] Recent evidence predicts the role of leptin plasma level as a risk factor for the development of CVDs. It has been reported that a strong association exists between leptin plasma levels and increased risk of myocardial infarction and stroke.^[44] Brain-derived neurotrophic factor (BDNF)-coding gene *BDNF*, another key gene (having node degree 12 and BC 0.10461538) [Table 2], was obtained from our study; it plays a role in feeding behavior as well in maintaining homeostasis of the body. Haploinsufficiency of *BDNF* gene and polymorphism in *BDNF* val66met causes impairment in *BDNF* secretion and cellular signaling,^[45] which subsequently leads to obesity. The onset of obesity is associated with T2D and other medical complications.^[46]

Another key gene, *CLOCK* (Circadian Locomotor Output Cycles Kaput), having node degree 10 and BC 0.15666667 [Table 2], was obtained from our study that also plays an important role in maintaining the circadian rhythm of our body. The circadian clocks are transcriptionally regulated cell-autonomous molecular mechanisms that directly regulate cellular and biological function at multiple sequential levels: Daily, seasonal, and throughout the lifespan.^[47] A cell-autonomous circadian clock allows the anticipation of changes in extracellular or environmental stimuli. This

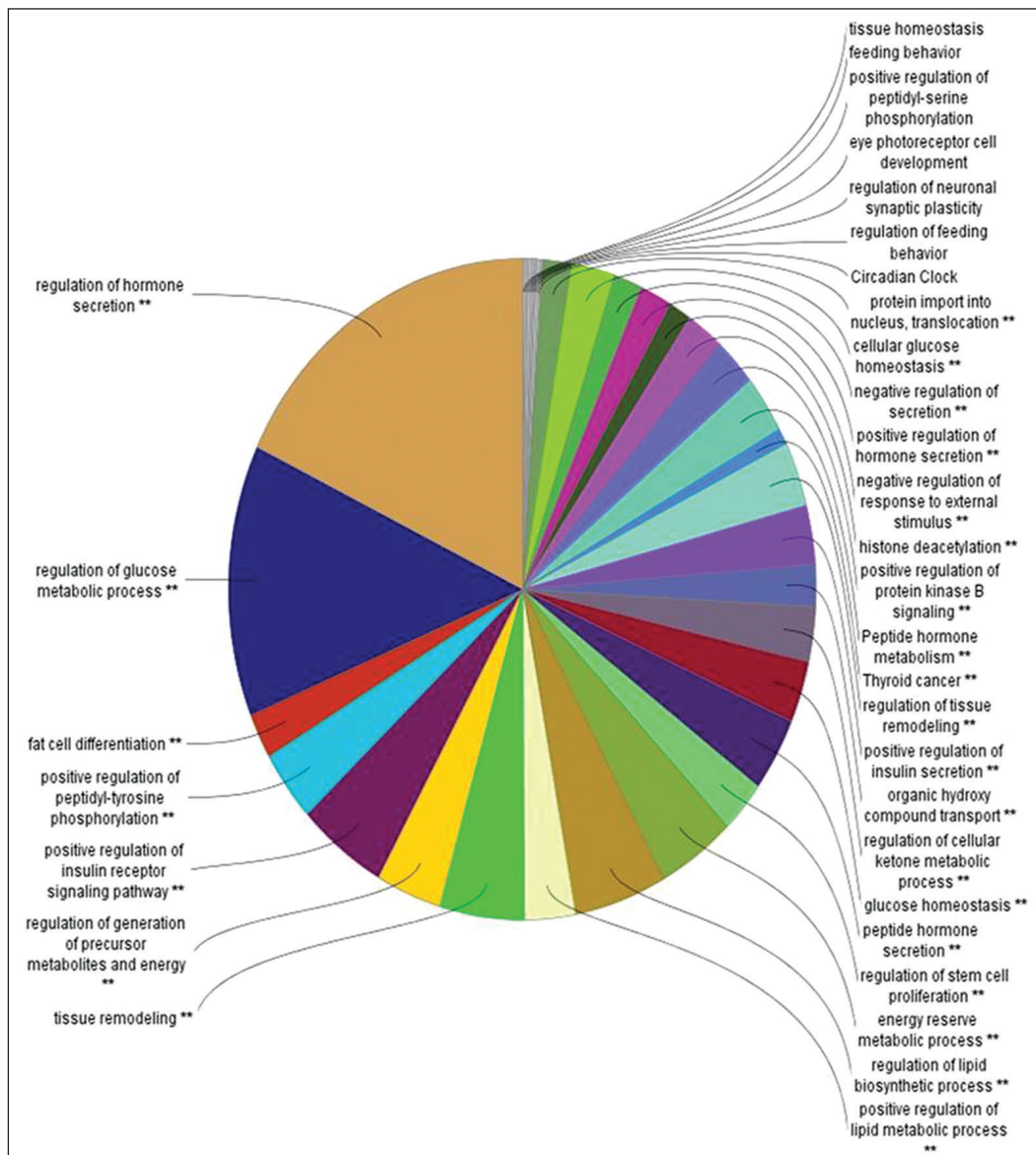


Figure 4: Figure showing molecular function and pathway analysis of genes involved in four chronic stress-related lifestyle disorders

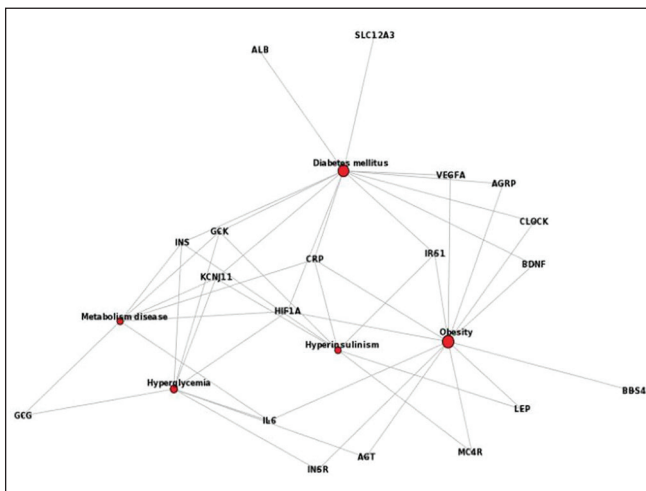


Figure 5: Figure showing network of enriched genes involved in top five diseases using FunDO

mechanism enables the cell/organ/organism to react to a stimulus with appropriate timing and response.^[48] Hence, impairment of the clock mechanism may result in responses to extracellular cues outside of the normal physiological range.^[47] The circadian clock regulates heart rate, glycogen metabolism, and triglyceride levels, as well as adapting the responsiveness of the myocardium to extracellular stimuli, such as fatty acids and β -adrenergic signaling.^[49] In multiple animal models of CVD, these mechanisms have been altered and are said to modulate the severity of myocardial damage in response to stresses. It has been considered that dyssynchrony of the cardiomyocyte circadian clock in shift workers, as well as individuals with diabetes and obesity, may lead to cardiovascular complications.^[50] It was observed from network analysis that the common driver genes involved in all four diseases under discussion have identical node degree and BC values for each of these diseases. It reflects that these genes have common interacting partners, and the interconnectedness of the genes. It signifies the fact that there is a set of common genes responsible for the generation of these lifestyle diseases due to chronic lifestyle stress in the form of sleep deprivation, abdominal fat accumulation, dietary irregularity, heightened awareness,

Table 3: Enriched genes involved in the stress-related lifestyle diseases, by FunDO

Disease ontology (DO) term	Number of genes in list	Names of the genes involved
Obesity	13	<i>BDNF, MC4R, CLOCK, BBS4, LEP, IL6, IRS1, HIF1A, AGRP, AGT, CRP, INSR, VEGFA</i>
Diabetes mellitus	18	<i>BDNF, ALB, CLOCK, INS, IRS1, HIF1A, AGRP, KCNJ11, CRP, SLC12A3, VEGFA, GCK, GCG, IL6, AGT, INSR, MC4R, LEP</i>
Atherosclerosis and other CVDs	15	<i>BDNF, JAK2, LDLR, INS, LEP, HIF1A, ALB, CRP, KCNJ11, VEGFA, IL6, CLOCK, GCG, AGRP, AGT</i>
Hypertension	9	<i>IL6, CLOCK, LEP, GCG, BDNF, AGRP, INS, INSR, SLC12A3</i>

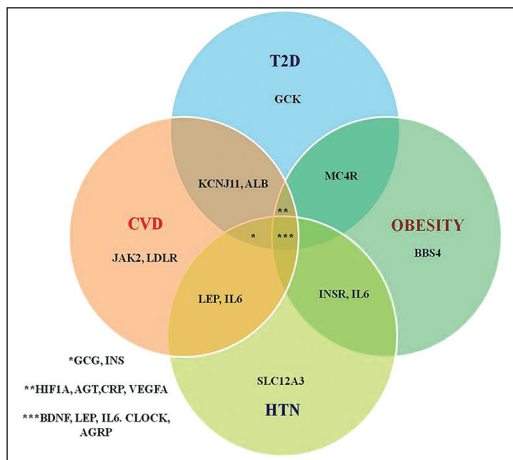


Figure 6: Figure showing the interconnectedness of the diseases along with their genes involved as represented by a basic Venn diagram

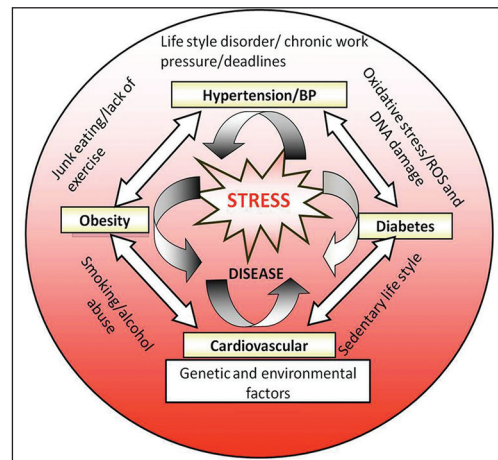


Figure 7: Figure showing factors responsible for interlinked chronic stress-related lifestyle diseases like CVDs, HTN, obesity, and T2D

and spikes of hypertensive episodes.

From our study, it was observed that *CRP*, *VEGFA*, *HIF1A*, and *AGT* are involved in CVDs, T2D, and obesity. Several earlier reports suggested that C-reactive protein encoded by the *CRP* gene plays an important role in metabolic syndrome. It is also regarded as an indicative biomarker for CVDs,^[51] whereas vascular endothelial growth factor A (VEGF-A), encoded by the *VEGFA* gene, is a key member of the family of growth factors. Reports have established that patients with CVDs have higher levels of VEGF-A in their serum.^[52] *HIF1A*, another gene obtained from our study encodes the alpha subunit of the transcription factor hypoxia-inducible factor-1 (HIF-1), indicating its role in obesity, T2D, and CVDs by disease ontology annotation. Reports have also established that, due to loss of HIF-1, fasting blood glucose levels were significantly increased and insulin response was impaired, with delayed glucose clearance from the blood and significantly decreased glucose uptake into the brain and heart^[53] The *AGT* gene, which encodes angiotensinogen precursor, has an important role in maintaining blood pressure. Reports indicated that polymorphism in the *AGT* gene leads to cardiovascular complications such as atherosclerosis.^[54]

Network analysis can serve as a powerful tool for gene prioritization.^[22] Hence, using the network analysis approach in our study we obtained five key genes that play significant roles in disease development due to stress-related lifestyle patterns and behaviors.

Conclusion

Studies on the human disease network indicate a common genetic origin for many diseases, hence suggesting the interconnectedness of the genes.^[55] Among the key genes obtained from our study, most are involved in maintaining homeostasis of the body. The homeostasis of the body is primarily influenced by lifestyle. Therefore it is imperative that we study the genome as a whole. The purpose of our study was to find

out the common driver genes, abnormalities in which lead to a disease condition. A schematic diagram is presented in Figure 7 in order to show the causative agents that lead to these disease conditions among people. It is evident that lack of exercise, a sedentary lifestyle, consumption of tobacco, consumption of alcohol, work pressure, exposure to mutagens, and the environment work in synergy in the development of stress-related lifestyle disorders. Hence, gaining deeper knowledge of the genes involved as well as their pathophysiological relevance to the disease will enable us to take the necessary preventive measures. The experimental validation of all the candidate genes is not a feasible proposition, due to the high costs involved; getting biological samples from apparently healthy individuals is challenging, as well. It has been used not only to visualize complex interactions among the individual components, but also to comprehend their relative importance in the network based on well-defined topological parameters. In our study, by database mining and using a network-based functional annotation approach, we have derived the primary set of genes associated with chronic stress-related lifestyle disorders and their involvement in the key pathways and biological processes by functional enrichment analysis. Such visualization and identification can promote better understanding of the underlying disease process and also identify specific gene targets for therapy. Nonetheless, additional studies are required to confirm these initial findings so that their true potential may eventually be realized in a clinical setting.

Acknowledgments

We would like to acknowledge extramural funding from the Department of Biotechnology (DBT), Govt. of India.

References

1. Epel ES. Psychological and metabolic stress: A recipe for accelerated

- cellular aging? Hormones (Athens) 2009;8:7-22.
2. Banerjee B, Hande MP. Age-independent telomere shortening and ion-channel defects in SCD. *Nat Rev Cardiol* 2013;10:362.
 3. Taylor AE, Ebrahim S, Ben-Shlomo Y, Martin RM, Whincup PH, Yarnell JW, *et al.* Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: A study using data from 4 UK cohorts. *Am J Clin Nutr* 2010;91:547-56.
 4. Egger G, Dixon J. Beyond obesity and lifestyle: A review of 21st century chronic disease determinants. *Biomed Res Int* 2014;2014:731685.
 5. Herder C, Karakas M, Koenig W. Biomarkers for the prediction of type 2 diabetes and cardiovascular disease. *Clin Pharmacol Ther* 2011;90:52-66.
 6. Radha V, Mohan V. Genetic predisposition to type 2 diabetes among Asian Indians. *Indian J Med Res* 2007;125:259-74.
 7. Raji A, Seely EW, Arky RA, Simonson DC. Body fat distribution and insulin resistance in healthy Asian Indians and Caucasians. *J Clin Endocrinol Metab* 2001;86:5366-71.
 8. Abate N, Garg A, Peshock RM, Stray-Gundersen J, Grundy SM. Relationships of generalized and regional adiposity to insulin sensitivity in men. *J Clin Invest* 1995;96:88-98.
 9. Tanira MO, Al Balushi KA. Genetic variations related to hypertension: A review. *J Hum Hypertens* 2005;19:7-19.
 10. Lifton RP. Molecular genetics of human blood pressure variation. *Science* 1996;272:676-80.
 11. Lim J, Hao T, Shaw C, Patel AJ, Szabó G, Rual JF, *et al.* A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* 2006;125:801-14.
 12. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, *et al.* Genetics of gene expression and its effect on disease. *Nature* 2008;452:423-8.
 13. Capes SE, Hunt D, Malmberg K, Pathak P, Gerstein HC. Stress hyperglycemia and prognosis of stroke in nondiabetic and diabetic patients: A systematic overview. *Stroke* 2001;32:2426-32.
 14. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461:218-23.
 15. Perez-Iratxeta C, Bork P, Andrade MA. Association of genes to genetically inherited diseases using data mining. *Nat Genet* 2002;31:316-9.
 16. Turner FS, Clutterbuck DR, Semple CA. POCUS: Mining genomic sequence annotation to predict disease genes. *Genome Biol* 2003;4:R75.
 17. López-Bigas N, Ouzounis CA. Genome-wide identification of genes likely to be involved in human genetic disease. *Nucleic Acids Res* 2004;32:3108-14.
 18. Tiffin N, Kelso JF, Powell AR, Pan H, Bajic VB, Hide WA. Integration of text- and data-mining using ontologies successfully selects disease gene candidates. *Nucleic Acids Res* 2005;33:1544-52.
 19. Adie EA, Adams RR, Evans KL, Porteous DJ, Pickard BS. Speeding disease gene discovery by sequence based candidate prioritization. *BMC Bioinformatics* 2005;6:55.
 20. Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, *et al.* Gene prioritization through genomic data fusion. *Nat Biotechnol* 2006;24:537-44.
 21. McKusick VA. Mendelian inheritance in man and its online version, OMIM. *Am J Hum Genet* 2007;80:588-604.
 22. Wu X, Jiang R, Zhang MQ, Li S. Network-based global inference of human disease genes. *Mol Syst Biol* 2008;4:189.
 23. Lage K, Karlberg EO, Störling ZM, Olason PI, Pedersen AG, Rigina O, *et al.* A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol* 2007;25:309-16.
 24. Galhardo M, Sinkkonen L, Berninger P, Lin J, Sauter T, Heinänen M. Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network. *Nucleic Acids Res* 2014;42:1474-96.
 25. Valcárcel B, Ebbels TM, Kangas AJ, Soininen P, Elliot P, Ala-Korpela M, *et al.* Genome metabolome integrated network analysis to uncover connections between genetic variants and complex traits: An application to obesity. *J R Soc Interface* 2014;11:20130908.
 26. Freudenberg J, Propping P. A similarity-based method for genome-wide prediction of disease-relevant human genes. *Bioinformatics* 2002;18 (Suppl 2):S110-5.
 27. Cheng D, Knox C, Young N, Stothard P, Damaraju S, Wishart DS. PolySearch: A web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. *Nucleic Acids Res* 2008;36:W399-405.
 28. Snel B, Lehmann G, Bork P, Huynen MA. STRING: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* 2000;28:3442-4.
 29. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, *et al.* A travel guide to cytoscape plugins. *Nat Methods* 2012;9:1069-76.
 30. 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.
 31. Hotamisligil GS, Shargill NS, Spiegelman BM. Adipose expression of tumor necrosis factor-alpha: Direct role in obesity-linked insulin resistance. *Science* 1993;259:87-91.
 32. Libby P. Inflammatory mechanisms: The molecular basis of inflammation and disease. *Nutr Rev* 2007;65:S140-6.
 33. Gregor MF, Hotamisligil GS. Inflammatory mechanisms in obesity. *Annu Rev Immunol* 2011;29:415-45.
 34. Whelton SP, Narla V, Blaha MJ, Nasir K, Blumenthal RS, Jenny NS, *et al.* Association between resting heart rate and inflammatory biomarkers (high-sensitivity C-reactive protein, interleukin-6, and fibrinogen) (from the Multi-Ethnic Study of Atherosclerosis). *Am J Cardiol* 2014;113:644-9.
 35. Finn AV, Nakano M, Narula J, Kolodgie FD, Virmani R. Concept of vulnerable/unstable plaque. *Arterioscler Thromb Vasc Biol* 2010;30:1282-92.
 36. Ross R. Atherosclerosis — An inflammatory disease. *N Engl J Med* 1999;340:115-26.
 37. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, *et al.* IL6 gene promoter polymorphisms and type 2 diabetes: Joint analysis of individual participants' data from 21 studies. *Diabetes* 2006;55:2915-21.
 38. Modan M, Halkin H, Almog S, Lusky A, Eshkol A, Shefi M, *et al.* Hyperinsulinemia. A link between hypertension obesity and glucose intolerance. *J Clin Invest* 1985;75:809-17.
 39. Chen AS, Marsh DJ, Trumbauer ME, Frazier EG, Guan XM, Yu H, *et al.* Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. *Nat Genet* 2000;26:97-102.
 40. Haskell-Luevano C, Cone RD, Monck EK, Wan YP. Structure activity studies of the melanocortin-4 receptor by *in vitro* mutagenesis: Identification of agouti-related protein (AGRP), melanocortin agonist and synthetic peptide antagonist interaction determinants. *Biochemistry* 2001;40:6164-79.
 41. Havel PJ. Role of adipose tissue in body-weight regulation: Mechanisms regulating leptin production and energy balance. *Proc Nutr Soc* 2000;59:359-71.
 42. Friedman JM, Halaas JL. Leptin and the regulation of body weight in mammals. *Nature* 1998;395:763-70.
 43. Porreca E, Di Febbo C, Pintor S, Baccante G, Gatta V, Moretta V, *et al.* Microsatellite polymorphism of the human leptin gene (LEP) and risk of cardiovascular disease. *Int J Obes (Lond)* 2006;30:209-13.
 44. Wallace AM, McMahon AD, Packard CJ, Kelly A, Shepherd J, GawA, *et al.* Plasma leptin and the risk of cardiovascular disease in the west of Scotland coronary prevention study (WOSCOPS). *Circulation* 2001;104:3052-6.
 45. Chen ZY, Jing D, Bath KG, Ieraci A, Khan T, Siao CJ, *et al.* Genetic variant BDNF (Val66Met) polymorphism alters anxiety-related behavior. *Science* 2006;314:140-3.
 46. Wynne K, Stanley S, McGowan B, Bloom S. Appetite control. *J Endocrinol* 2005;184:291-318.
 47. Ederly I. Circadian rhythms in a nutshell. *Physiol Genomics* 2000;3:59-74.
 48. Wilsbacher LD, Takahashi JS. Circadian rhythms: Molecular basis of the clock. *Curr Opin Genet Dev* 1998;8:595-602.
 49. Durgan DJ, Young ME. The cardiomyocyte circadian clock: Emerging roles in health and disease. *Circ Res* 2010;106:647-58.

50. Härmä MI, Ilmarinen JE. Towards the 24-hour society — new approaches for aging shift workers? *Scand J Work Environ Health* 1999;25:610-5.
51. Lee K, Song YM, Sung J. Genetic and environmental associations between C-reactive protein and components of the metabolic syndrome. *Metab Syndr Relat Disord* 2013;11:136-42.
52. Kranz A, Rau C, Kochs M, Waltenberger J. Elevation of vascular endothelial growth factor-A serum levels following acute myocardial infarction. Evidence for its origin and functional significance. *J Mol Cell Cardiol* 2000;32:65-72.
53. Huang Y, Lei L, Liu D, Jovin I, Russell R, Johnson RS, *et al.* Normal glucose uptake in the brain and heart requires an endothelial cell-specific HIF-1 α -dependent function. *Proc Natl Acad Sci U S A* 2012;109:17478-83.
54. Al-Najai M, Muiya P, Tahir AI, Elhawari S, Gueco D, Andres E, *et al.* Association of the angiotensinogen gene polymorphism with atherosclerosis and its risk traits in the Saudi population. *BMC Cardiovasc Disord* 2013;13:17.
55. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. *Proc Natl Acad Sci U S A* 2007;104:8685-90.

How to cite this article: Roy S, Chakraborty A, Ghosh C, Banerjee B. Systematic analysis of integrated gene functional network of four chronic stress-related lifestyle disorders. *Genome Integr* 2015;6:1.
Source of Support: Grant from Department of Biotechnology, Ministry of Science and Technology, Government of India, Grant number: BT/PR14659/31/108/2010. **Conflict of Interest:** None declared.

Received: 27th December, 2014; **Accepted:** 5th March, 2015;
Publish: 28th April, 2015

- Supplementary 1:** List of genes retrieved from PolySearch database source for CVD
Supplementary 2: List of genes retrieved from PolySearch database source for HTN
Supplementary 3: List of genes retrieved from PolySearch database source for obesity
Supplementary 4: List of genes retrieved from PolySearch database source for T2D