Volume 5, Issue 10, October-2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Differential Gene Expression Analysis of Hypertrophic and Dilated Cardiomyopathy Signature Genes

Beutline Malgija, Shanmughavel P*
Bharathiar University, Coimbatore,
Tamilnadu, India

Abstract: ***Cardiomyopathy is a heterogeneous disease caused due to the inability in contraction of muscle cells, leading to heart failure and death. Despite understanding of the mechanism behind its pathology, the prognostic and therapeutic strategies become inefficient. Hence we aimed in finding the common genetic signatures among the most prevalent form of cardiomyopathy, namely hypertrophic and dilated by the analysis of gene expression data. We found 54 significant genes to be common among the hypertrophic and dilated cardiomyopathic samples. Further network analysis and functional annotation of the significant genes revealed the importance of these genes and their involvement in the pathogenicity of the disease. The genes OMD, OGN, POSTN, S100A8, CCL2 and SERPINA3 were found to be important by our analysis.***

Keywords: Cardiomyopathy, Myocardium, Hypertrophic, Dilated, Gene expression, Microarray

I. INTRODUCTION

Cardiomyopathy, one of the important cardiovascular diseases is a leading cause for heart transplants and the most common identifiable cause of sudden death in young adults. Cardiomyopathies are diseases of the myocardium associated with cardiac dysfunction [1] due to inflammation and enlargement of the heart muscle. This arises when there is a problem with responses to various genetic and extrinsic forces, when upholding cardiac contraction. Coronary artery disease, inflammatory disease, chronic alcoholism, metabolic and blood disorders and genetic mutations can cause cardiomyopathy and can be acquired or inherited [2]. Cardiomyopathy can be classified as dilated, hypertrophic, restrictive, arythmogenic right ventricular and unclassified and one form can progress to another [3]. Among the wide range of acquired and inherited cardiomyopathies stated by the World Health Organization (WHO), Hypertrophic and Dilated were found to be the most predominant. Hypertrophic results in abnormal thickening of heart muscle, especially the left ventricle whereas dilated arise in the enlargement and weakening of the heart, both resulting in the disruption to the electrical functioning of the heart [4].

The variation in symptoms among individuals, even among family members, affected by the same type of cardiomyopathy made difficulty in diagnosis and progosis of the disease, suggesting the involvement of non sequence variations. This insists the need for identification of genes and the role of epigenetic factors and the molecular mechanism behind them, providing better insight into the pathogenesis and progression of the disease. Also the response to the current ACE (Acetyl choline Esterase) inhibitors, beta blockers, and other drugs using nowadays, vary among individuals [5], emphasize the need for novel biomarkers. Even though several cytoskeletal and sarcoplasmic genes and their coordinated genes have been reported for their association, still the clear molecular mechanism behind it is not clear. This is due to the epigenetic factors and the lack of knowledge on other proteins that might correlate with the pathogenesis of cardiomyopathy.

Hypertrophic may begin with a hypertrophic pattern and become dilated [6], which requires the identification of common genetic factors underlying the disease in order to develop better gene therapy procedures. Thus finding the overlapping genes that were differentially expressed in DCM and HCM might provide new insights into their molecular mechanism. Hence this study aims in identifying novel Hypertrophic and Dilated cardiomyopathy signature genes based on a twofold procedure. First, the genes were screened by performing single gene based expression profile analysis of two different datasets of Dilated cardiomyopathy (DCM) and Hypertrophic cardiomyopathy (HCM) in comparison with non failing hearts. Secondly, Gene set enrichment analysis was performed. Further network analysis identified the interactive neighbors of the screened genes.

II. MATERIALS AND METHODS

Screening of DEGs

Gene expression data for Dilated (DCM, GSE3585) and Hypertrophic (HCM, GSE1145) cardiomyopathy were retrieved from Gene Expression Omnibus [7]. The raw datasets were normalized separately using the Robust Multi Array (RMA) method by the affy package of Bioconductor using R language. Further these normalized values were utilized for the identification of Differentially Expressed Genes (DEGs) in various datasets using limma package. The threshold with P-value <0.05 and a fold change of 1.5 were considered to be differentially expressed. The significant probe IDs were converted to gene symbols by DAVID gene ID conversion tool [8] and annotate package. The significant genes were

further screened based on preranking by GSEA. We searched for the transcription factor (TF) activity of DEGS through literature and TFcheckpt.

GSEA [9] was performed to find the gene sets enriched in the genes that are significantly associated with the phenotype (HCM or DCM) using the kolmogarov smirnov (KS) statistic [10] and the genes are ranked based on signal-to-noise ratio. we used curated gene sets (c2) from pathway databases (KEGG, biocarta), publications in pubmed, and knowledge from the domain experts.

Identification of Transcription factor Binding sites

The significant genes were clustered using different clustering algorithms namely Hierarchical, K-means and Fuzzy C means (FCM) using R. The expression matrix was in the dimension of 242 x 12 and 756 x 16 for DCM and HCM respectively. We took the average of all the probes for the same mRNA for dealing with interrogated genes [11]. We input the expression matrix of both DCM and HCM to calculate distance and the distance matrix was constructed based on euclidean distance, which computes the distance between two points in euclidean space. Among the three algorithms, K-means and FCM requires the number of clusters to be predefined, since it varies based on the input dataset. We calculated the number of clusters based on different methods like silhoutte, and calinski criterion for each of the datasets separately. The fuzziness parameter for FCM was set to 2, which is the default value in most cases. The generated clusters were analyzed and further validated using the clValid package [12]. Each of the clusters were further analysed for their transcription factor binding sites (TFBS) using oPOSSUM server

Network Analysis of significant genes

Neighbour interactive partners of the common significant genes (hypertrophic and dilated samples) were obtained from STRING [13], a database of predicted functional association between proteins. Based on these interactive partners and their confidence scores, network was constructed using Cytoscape [14] for both the commonly significant genes. The highly interconnected sub graphs were analyzed using Molecular Complex Detection (MCODE) module.

III. RESULTS

Analysis of Differential Gene Expression

In the current study, the differentially expressed genes among HCM and DCM samples were identified by gene and gene set enrichment based analysis. The median values of all the samples in each of the datasets were found to be unbiased after normalization and the DEGs were screened . The total number of significant genes in both datasets (HCM and DCM) are shown in Table I. The replicate genes in each condition were removed, resulting in 820 significant genes for HCM and 434 for DCM. Totally 18833 genes from the HCM datasets and 13321 genes of dilated were pre-ranked based on their expression profile using GSEA preranking. Further manual inspection of these genes revealed 242 genes to be commonly significant in both single gene and gene set analysis for DCM and 758 genes for HCM datasets. We found 15 TFs (including 8 up and 7 down regulated) and 48 TFs (23 up and 25 down) to be significant in DCM and HCM respectively.

Table I: Total number of significant genes among datasets

Disease & Dataset	Up regulated	Down regulated	Total significant genes
HCM (GSE1145)	310	448	758
DCM (GSE3585)	148	94	242

The pairwise combination of the datasets, reported the overlap of 54 DEGs and further geneset enrichment analysis also confirmed these genes to be significant based on ranking and these common genes are presented in Supplementary file I.

Geneset Enrichment Analysis

Gene set enrichment was further carried out for the screened DEGs to find their enriched pathways. 1437 among 3643 and 505 among 3643 of curated (c2) genesets from MsigDB were found to be upregulated in HCM and DCM respectively. Among which 35 gene sets in HCM were significantly enriched with a nominal p-value less than 0.05 and normalized enrichment score (NES) greater than 1. Whereas 20 gene sets were significantly enriched in DCM with the above preferred cutoff. This showed the enrichment of pathways dependent on calcium, potassium and degradation of extracellular matrix.

Analysis of Co-expressed genes

Clustering techniques group the genes with similar expression pattern which tends to share a common biological pathway.The number of clusters for K-means and FCM were set to 2 for DCM and 3 for HCM based on their estimation using silhoutte and calinski criterion.

Table II: Number of clusters generated by various methods

Algorithms	DCM		HCM		
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3
HC	235	7	232	523	1

KM	142	100	177	235	344
FCM	104	138	185	231	340

HC- Hierarchical, KM- K-Means, FCM- Fuzzy C Means

The significant genes were clustered based on their expression profiles with Euclidean distance. Based on the distance matrix constructed by the expression profile of both the datasets, the genes occupy the same clusters based on K-means and FCM. Validation of the generated clusters by clValid package confirms the hierarchical cluster with two and five clusters to be the best one and thus the clusters generated by hierarchical algorithm was considered. TFBS analysis of the clusters predicted the enrichment of different transcription factors and their binding sites. The transcription factors enriched in different clusters of DEGs are depicted in Table III. We found the overlap of transcription factors Foxd1, Nkx2-5, Hoxa5, SRY, Gfi, Foxa1, Foxo3, Nfatc2, Arid3a, Sox5, TBP, ELF5, Pdx1, Sox17, CEBPA, Foxi1, Ap1, Nobox in DCM and HCM.

Table III: Transcription factors Enriched in various clusters

Disease	**Cluster 1**	**Cluster 2**	**Cluster3**
DCM	SRF, FOXD1, Nkx2-5, HOXA5, SRY, Gfi, Foxa2, FEV, FOXA1, FOXO3, NFATC2, ARID3A, Sox5, TBP, ELF5, Pdx1, Sox17, CEBPA, STAT1, RUNX1, FOXI1, AP1, Nobox	NHLH1, TAL1::TCF3, REST, MZF1_1-4, Stat3, HIF1A::ARNT, IRF2, EBF1, NFIL3	-
HCM	Zfx, Klf4, SP1, Pax5, CTCF	Nkx2-5, Foxd3, SRY, ARID3A, HOXA5, FOXA1, FOXI1, Foxa2, NKX3-1, Pdx1, Sox5, FOXO3, CEBPA, Prrx2, FOXD1, TBP, NFATC2, Gfi, Nobox, Foxq1, Gata1, SOX9, Lhx3, Sox17, NFIL3, SPIB, AP1, FOXF2, NR3C1, RORA_2, MEF2A, ELF5, HLF, IRF1, Sox2, Pou5f1, Ddit3::Cebpa, Tal1::Gata1, TAL1::TCF3, PBX1	RELA, NF-1, GR, Olf-1

Analysis of regulatory Networks

The interactive neighbors of the common significant genes i.e., 54 genes including 40 up and 14 down regulated with their confidence score was downloaded from the STRING database for construction of the network. Network of the significant genes for the common genes with their interactions were constructed individually and then merged and visualized using Cytoscape. The merged network contains 282 nodes with a network density of 0.008 and clustering coefficient 0.111. The important hub nodes were identified based on their degree and cluster coefficient. The highly interconnected sub graphs were identified using the MCODE algorithm, which predicted 4 clusters from the network and the results are depicted in figure 2.

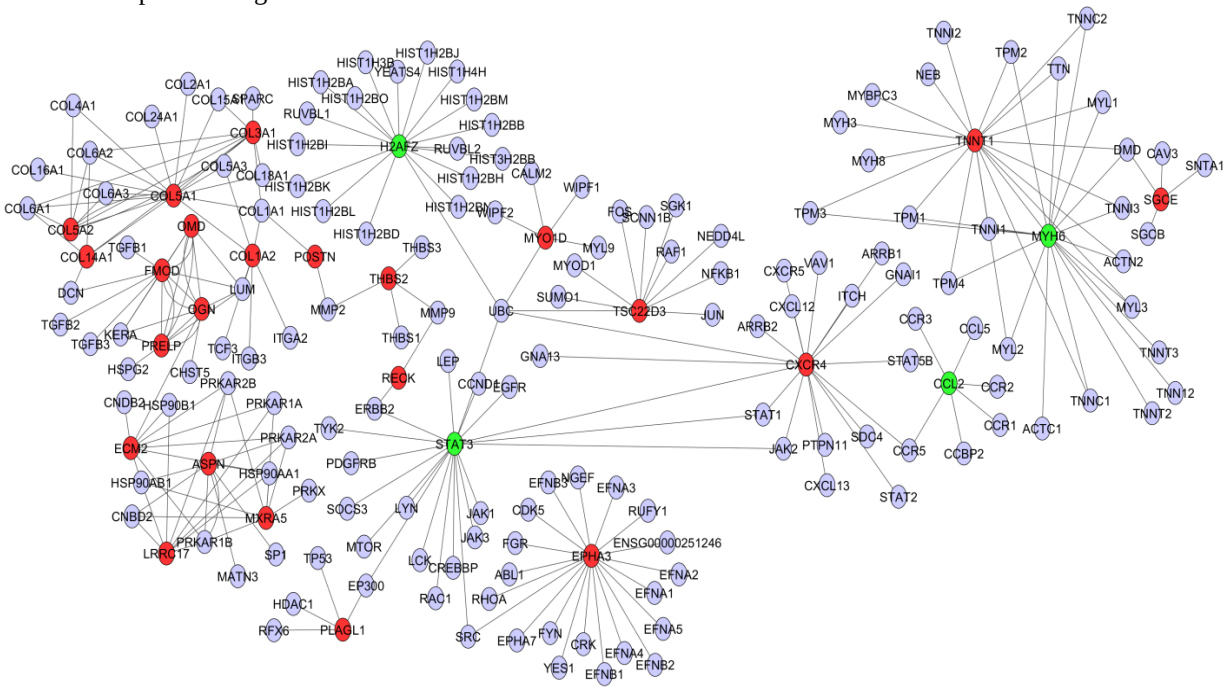

Figure 1: Protein-protein interaction Network for common significant genes

Grey circle- Interactive partners, Red circle- Up regualted genes, Green circle – Down regulated genes

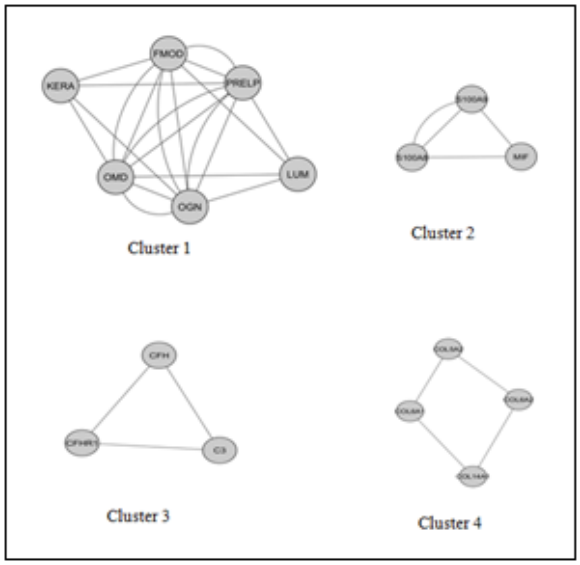

Figure 2: Clusters generated from protein-protein interaction Network

Functional Enrichment

The differentially expressed genes were analyzed in the context of Gene Ontology (GO). The up regulated genes in hypertrophic cardiomyopathy were found to be localized in the extracellular region and were found to have receptor activity and involved in immune system process. Whereas most of the down regulated genes were found to be in the intracellular region and involved in nuclear receptor binding and various metabolic process. On annotation with Gene Ontology (GO), the majority of the common differentially expressed genes were found to be involved in cell growth, maintenance, signal transduction and regulation of metabolic process.

IV. DISCUSSION

In the current study, we integrate independent microarray datasets involving HCM, DCM and non failing hearts, aimed at finding the common signature genes and the biological factors involved in the pathogenesis of the disease. The functional and pathway based enrichment analysis of the significant genes provides insights into the molecular mechanism of cardiomyopathy. GO analysis revealed the association of the identified significant genes in various biological processes.

Altered Gene Expression in DCM

242 transcripts, including 148 upregulated and 94 down regulated were found to be differentially expressed in DCM comparable with NF heart. This shows an increase in the up regulation of gene expression compared to down regulation, suggesting a net transcriptional activation in dilated cardiomyopathy. Some of the genes, including Natriuretic peptide A (NPPA) [15], Connective tissue Growth factor (CTGF), Nephroblastoma over expressed (NOV) which were already reported to be associated with heart failure were identified as differentially expressed by the study. GO cellular component revealed that most of the significantly up regulated genes were localized in the extracellular region and most of the downregulated were found to be in the intracellular region. Up regulation of genes involved in the organization of the extracellular matrix specifies the reason behind the deposition of extracellular matrix components. Similarly the down regulated genes involve various immune related processes. Genes related to respiration and lung development was found to be down regulated. The sarcomeric gene vinculin was up regulated.

The extracellular matrix proteins (COMP, ELN), collagen genes (COL1A1, COL1A2, COL3A1, COL5A1, COL5A2, COL14A1 and COL16A1), growth factors (CTGF, LTBP1) and member of the kinase family (JAK2) were found to be up regulated. The cytoskeletal proteins and the other proteins which connect the sarcomere with the extracellular matrix (ECM) allow cardiomyocytes to sense and respond to changes in contraction. Mutations in cytoskeletal proteins cause disruption of the response [16] and Mutation in a cytoskeletal protein, dystrophin (DMD) leads to development of DCM. This is because the hinge (H1) region of mutated dystrophin, having sequence similarity to the region of troponin I, enables the association of this region with calcium binding proteins like calmodulin (CALM1), which modulates the interaction of dystrophin with regulatory proteins causing DCM [17]. Up regulation of the genes involved in calmodulin binding, RGS4 and AEBP1 suggest their involvement in dystrophin disruption.

TGF-β plays a major role in the pathogenesis of both hypertrophic and dilated cardiomyopathy by stimulating the growth in cardiomyocyte and by inducing interstitial fibrosis [18]. Genes coding for TGF-β binding proteins LTBP1 (Latent transforming growth factor beta binding protein 1) and LTBP2 (Latent transforming growth factor beta binding protein 2) playing structural role in ECM were found to be up regulated. TGF-β can also induce CTGF and the abnormal expression of CTGF will lead to cardiac [19].

The most significant pathologic characteristic of DCM is the accumulation of ECM proteins, mainly collagen and they vary from other type of cardiomyopathy by lack of inflammatory response [20]. Up regulation of collagen genes

COL1A1, COL1A2, COL5A1 and other ECM proteins COMP, ELN implies their importance in ECM accumulation. Down regulation of genes related to chemokine activity (CCL11, CCL8, CCR3) and immune response (VSIG4), which correlates with the findings of Wang *et al*., [11] suggest their deficiency in inflammatory mediated response.

Differential Gene Expression in HCM

We found 758 genes to be significantly differentially expressed (310 up regulated and 448 down regulated) compared to non failing heart in HCM. The extracellular matrix protein (ECM2), members of small leucine rich proteoglycan family (LUM, ASPN), protein kinase (NRK) and collagen genes (COL1A2, COL5A2, COL3A1, COL5A1, COL8A1, COL12A1 and COL14A1) were found to be up regulated. Members of chemokine family (CCL11, CCL8, CCR1) and leukocyte antigens (CD14, CD53) were found to be down regulated, suggesting their fall in the immune system and inflammatory response mediated biological process.

The up regulation of the Extracellular matrix (ECM) proteins, namely ECM2 and COL8A1 will lead to excessive accumulation of ECM proteins, a well-known characteristic of HCM. Lumican (LUM), a small leucine rich proteoglycan (SLPR) can interact with collagens, especially of type VI, XII, XIV, fibronectin, elastin and growth factors. Hence its up regulation might leads to over expression of type XIV collagen (COL14A1) leading to disruption in ECM framework. Their interactions with growth factor TGF-β give attention to the importance of TGF-β signalling pathway in HCM. ASPN, the third class I SLPR differs from other proteoglycans by their contiguous polyaspartate sequence (14 aspartate residue). The difference in number of these aspartate residues will alter the TGF-β driven chondrogenesis [21]. Thrombospondin (THBS2), a TGF-β activator, which disrupt the non covalent interaction between latency associated peptide (LAP) and transcription growth factor (TGF) molecule [22] was found to be up regulated by the present study.

Unlike DCM which mainly involves cytoskeletal proteins, HCM is mainly associated with sarcomeric proteins. The sarcomere genes Cardiac myosin binding protein C (MYBPC3), α-tropomyosin (TPM1), Titin (TTN) and Obscurin (OBSCN) were found to be up regulated, whereas α-Myocin heavy chain (MYH6), playing structural and functional role in the sarcomere, already reported in the pathogenesis of HCM [23] was found to be down regulated. Carniel *et al*., [24] reported the down regulation of the same in correlation with systolic dysfunction, a characteristic of DCM which is also reported in some patients with HCM. We also noticed the down regulation of nuclear membrane protein Lamin A (LAMA).

Network Analysis

54 common significant genes (40 up regulated and 14 down regulated) were further analysed to understand their interactive partners. Network analysis identified the following genes OMD, OGN, POSTN, ELN, CCL2, SERPINA3 and NR4A3 as the hub genes as these have more interactions and the clustering coefficient of these nodes were found to be 1, revealing the nodes are at the core of a fully interlinked cluster. Further investigation of the MCODE generated clusters revealed their involvement in various pathways, including ECM-receptor interaction, TGF-beta signaling pathway, calcium signaling pathway, cytokine-cytokine receptor interaction, Toll-like receptor signaling pathway, focal adhesion, glycerophospho-lipid metabolism and MAPK signaling pathway. Among which ECM-receptor interaction and TGF-beta signaling pathway were shared by most of the clusters (cluster 1, 2, 3), focussing these pathways to be important in cardiomyopathy research. TGF-beta activated by inflammatory response stimulates the activation and proliferation of fibroblasts (mediated by type I (COL1A1, COL1A2) and III collagen), resulting in myocardial fibrosis and associated hypertrophy, which deposits connective tissue.

Importance of the Common Significant genes

Among the significant genes identified from both datasets (dilated and hypertrophic), 54 genes, including 40 up regulated and 14 down regulated were found to overlap, which were believed to participate in a common pathogenic mechanism behind both cases.

The commonly up regulated genes POSTN, OGN, OMD and LRRC17 shares a functional similarity, involving cell adhesion and localized in proteinaceous extracellular matrix. Activation of POSTN (periostin, osteoblast specific factor) will enhance the incorporation of BMP-1 in the fibronectin matrix of connective tissues, which increase fibrosis, a prominent feature in both types of cardiomyopathy. SSPN (Sarcospan), encodes a member of dystrophin-glycoprotein complex (DGC), provides a structural linkage connecting the sub-sarcolemmal cytoskeleton and the extracellular matrix of the muscle cells. This might interact with ECM proteins, allowing cardiomyocytes to sense and respond to contraction. The over expression of SSPN, disrupts cell adhesion and signaling pathways, which affects the level of reactive oxygen species [25], resulting in extracellular matrix accumulation, cardiac dyshomeostasis and apoptosis (loss of cardiomyocyte in case of DCM) which further leads to cardiomyopathy [26].

The sarcomeric genes MYH6 was found to be commonly down regulated and TPM1 was down regulated in HCM and up regulated in DCM.

Osteoglycin (OGN), a small leucine rich proteoglycan, regulates growth of the left ventricle [27]. Disruption in OGN regulation might leads to increase in left ventricular mass (LVM), an independent risk factor of cardiovascular disease. They can also modify the extracellular milieu by their interaction with collagen and soluble growth factors. Hence up regulation of OGN might trigger the activation of collagens including COL14A1 in HCM and DCM, COL1A1, COL1A2 and COL5A1 in DCM and COL8A1 in HCM which will further results in ECM accumulation, a common feature in both cases of cardiomyopathy.

Sulfatase I (SULF1), a cell surface polypeptide which modifies the sulfation status of heparin sulphate proteoglycans (HSPGs) thereby creating changes in HSPG-mediated cell signaling pathways. Down regulation of the

gene has been reported in several cancers [28,29] and also in normal development including neural, muscular, vascular and skeletal development. In addition, they have been reported to involve in epigenetic mechanism such as methylation [11]. However, no direct link between SULF1 and cardiomyopathy has been reported yet. Complement factor H (CFH), involved in heparin binding and HSPGs binding was also up regulated in common.

The genes RARRES1, S100A8, FCN3, CCL2 and SERPINA3 were found to be down regulated in both the cases (HCM and DCM). Among which calgranulin (S100A8), Chemokine ligand 2 (CCL2) and Ficolin 3 (FCN3), involved in stress and immune response leads to deficiency in inflammatory mediated response in cardiomyopathy. SERPINA3 is not yet reported to have a role in cardiomyopathy, but their up regulation is playing a part in Alzheimer's disease, cystic fibrosis, stroke and cerebral haemorrhage. Our study identified the down regulation of the gene in cardiomyopathy. Differential gene expression analysis by kittleson *et al*., [30] also reported the down regulation of RARRES1 gene when comparing gene expression profiles of Ischemic and non-ischemic cardiomyopathy. The deregulation of the same was also reported in cancer [31]. The S100 calcium binding protein A8 (S100A8), which functions as intracellular Ca^{2+} sensors and extracellular factors was found significantly down regulated commonly. The down regulation of the same was also reported earlier by Kittelson *et al*., [30]. Alterations in the level of calcium handling genes alter the regulation of calcium signaling by the sarcoplasmic reticulum [32], causes change in cardiac function and remodeling.

V. CONCLUSION

We identified the common genes that were up and down regulated in both cases of cardiomyopathy including hypertrophic and dilated. Most of the significant genes identified, shows either direct or indirect link to TGF-β, which plays an important role in cardiac remodelling, and also their interaction with ECM, emphasize the importance of TGF-β signalling pathway and ECM-receptor interaction in the pathogenic mechanism of cardiomyopathy. These genes and their further analysis in regard to the pathway may provide new avenues for the therapeutic strategies of cardiomyopathy (hypertrophic, dilated).

REFERENCES

[1] McKenna W.J., 1996, Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of Cardiomyopathies, *Circulation*, 93, 841-842.

[2] Maron B.J., Thiene G., 2011, Chapter 31. Classification of cardiomyopathies, New York, NY: McGraw Hill. Hurst's the Heart, 13e.

[3] Weerakkody Y., *et al.* 1995, Cardiomyopathy (classification), WHO/ISFC.

[4] Towbin J.A., and Lorts A, 2011, Arrhythmias and Dilated Cardiomyopathy - Common Pathogenetic Pathways?, *JACC*, 57, 21.

[5] Boer R.A and Veldhuisen D.J, 2002, Polypharmacy in chronic heart failure: practical issues regarding the use of angiotensin-converting enzyme inhibitors, beta-blockers and other drugs, *European Heart Journal Supplements*, 4, D11 l-D116.

[6] Richardson P., McKenna W., Bristow M., Maisch B., Mautner B., O'Connell J., Olsen E., Thiene G., Goodwin J., Gyarfas I., Martin I and Nordet, 1996, Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the definition and classification of cardiomyopathies, *Circulation*, 93(5), 841-842.

[7] Edgar R., Domrachev M., and Alex E.L, 2002, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, *Nucl. Acids Res*. 30 (1): 207-210.

[8] Huang D.W., Sherman T.B., Qina T., Collins J.R., Alvord W.G., Roayaei J., Stephens R., Michael W.B., Clifford L.H., and Richard A.L, 2007, The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists, *Genome Biology*, 8, R183.

[9] Subramanian A., Tamayo P., Mootha K.V., Mukherjee S., Ebert L.B., Gillette M.A., Paulovich A., Pomeroy S.L., Golub T.R., Lander E.S and Mesirov J.P, 2005, Gene Set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, *PNAS*, 102 (43), 15545-15550.

[10] Sheskin and David J, 2011, In Handb. Parametr. Nonparametric Stat. Proced., Chapman & Hall, 261–276.

[11] Wang X., Falkner B., Zhu H., Shi H., Su S., Xu X., Sharma A.K., Dong Y., Treiber F., Gutin B., Harshfield G., Snieder H, 2013, A Genome-Wide Methylation Study on Essential Hypertension in Young African American Males, *PLOS ONE*, 8 (1), e53938.

[12] Brock G, Pihur V, Susmita D, Somnath D, 2008, clValid, an R package for cluster validation, *Journal of Statistical software*, 1-14

[13] Mering V.C., Huynen M., Jaeggi D., Schmidt S., Bork P., Snel B, 2003, STRING: a database of predicted functional associations between proteins, *Nucleic Acids Res*, 31(1), 258-61.

[14] Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B., Ideker T., 2003, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res*, 13(11), 2498-504.

[15] Gopal D.M., Sam, F, 2013, New and emerging biomarkers in left ventricular systolic dysfunction-insight into dilated cardiomyopathy, *J Cardiovasc Transl Res*, 6(4), 516-527.

[16] Harvey P.A., and Leinwand A.L, 2011, Cellular mechanisms of cardiomyopathy, *JCB*, 194 (3), 355-365.

[17] Ahn A.H., Kunkel L.M, 1993, The structure and functional diversity of dystrophin, *Nat genet*, 3(4), 283-291.

[18] Bujak M and Frangogiannis N.G, 2007, The role of TGF-ß signaling in myocardial infarction and cardiac remodeling, *Cardiovascular Research*, 74, 184–195.

[19] Chen M.M, Lam A, Abraham J.A, Schreiner F.G and Joly A.H, 2000, CTGF Expression is Induced by TGF- in Cardiac Fibroblasts and Cardiac Myocytes: a Potential Role in Heart Fibrosis, *J Mol Cell Cardiol*, 32, 1805–1819.

[20] Ares-Carrasco S., Picatoste B., Benito-Martin, A., *et al*., 2009, Myocardial fibrosis and apoptosis, but not inflammation, are present in long-term experimental diabetes, *Am J Physiol Heart Circ Physiol*, 297, H2109–19.

[21] Kizawa H., Kou I., Iida A., Sudo A., Miyamoto Y., Fukuda A., Mabuchi A., Kotani A., Kawakami A. and Yamamoto S, 2005, An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis, *Nat. Genet*, 37, 138-144.

[22] Murphy-Ullrich J.E., Poczatek M, 2000, Activation of latent TGF-beta by thrombospondin-1: mechanisms and physiology, *Cytokine Growth Factor Rev*,11(1-2), 59-69.

[23] Niimura H., Bachinski L.L., Sangwatanaroj S., Watkins H., Chudley A.E., McKenna W., Kristinsson A., Roberts R., Sole M., Maron B.J., Seidman J.G., Seidman C.E, 1998, Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy, *N. Engl. J. Med*, 338(18), 1248-1257.

[24] Carniel E., Matthew R.G., Taylor., Sinagra G., Lenarda A.D., Ku L, 2005, α-Myosin Heavy Chain- A Sarcomeric Gene Associated With Dilated and Hypertrophic Phenotypes of Cardiomyopathy, *Circulation*, 112, 54-59

[25] Santillo M., Mondola P., Seru R., Annella T., Cassano S., Ciullo I., Tecce, M.F., Iacomino G., Damiano S., Cuda G., Paterno R., Martignetti V., Mele E., Feliciello A and Avvedimento E.V, 2001, Opposing functions of Ki-and Ha-Ras genes in the regulation of redox signals, *Curr. Biol*, 11(15), 614–619.

[26] Chen J., Zhang Z., and Ca L, 2014, Diabetic Cardiomyopathy and Its Prevention by Nrf2: Current Status, *Diabetes Metab J*, 38(5), 337-345.

[27] Petretto E., Sarwar R., Grieve I., Lu H., Kumaran M.K., Muckett P.J., Mangion J., Schroen B., Benson M., Punjabi P.P., Prasad S.K., Pennell D.J., Kiesewetter C., Tasheva E.S., Corpuz L.M., Webb M.D, Conrad G.W., Kurtz T.W., Kren V., Fischer J., Hubner N., Pinto Y.M., Pravenec M., Aitman T.J., and Cook S.A, 2008 Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass, *Nat Genet*, 40 (5), 546-552, 2008.

[28] Li J., Kleeff J., Abiatari I., Kayed H., Giese N.A., Felix K., Giese T., Buchler M.W, Friess H, 2005, Enhanced levels of Hsulf-1 interfere with heparin-binding growth factor signaling in pancreatic cancer, *Mol Cancer*, 4(1), 14.

[29] Lai J., Chien J., Staub J., Avula R., Greene E.L, Matthews T.A, Smith D.I, Kaufmann S.H, Roberts L.R and Shridhar V, 2003, Loss of HSulf-1 upregulates heparin-binding growth factor signaling in cancer, *J Biol Chem*, 278, 23107–23117.

[30] Kittleson M., Minhas K.M., Irizarry R.A., Ye S.Q., Edness G., Breton E., Conte J.V., Tomaselli G., Garcia J., Hare M.J, 2005, Gene expression analysis of ischemic and nonischemic cardiomyopathy: shared and distinct genes in the development of heart failure, *Physiol. Genomics*, 21, 299-307.

[31] Wu C.C., Shyu R.Y., Chou J.M., Jao S.W., Chao P.C., Kang J.C., Wu S.T., Huang S.L, Jiang S.Y, 2006, RARRES1 expression is significantly related to tumour differentiation and staging in colorectal adenocarcinoma, *Eur J Cancer*, 42(4), 557-65.

[32] Kranias E.G., Bers D.M, 2007, Calcium and cardiomyopathies, *Subcell Biochem*, 45, 523-37.

