Computational Biology and Chemistry 73 (2018) 85–94

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/compbiolchem

Collective transcriptomic deregulation of hypertrophic and dilated cardiomyopathy – Importance of fibrotic mechanism in heart failure

Beutline Malgija[a], Nachimuthu Senthil kumar[b], Shanmughavel Piramanayagam[a,*]

[a] Department of Bioinformatics, Bharathiar University, Coimbatore - 641 046, India
[b] Department of Biotechnology, Mizoram University, Aizawl - 796 004, India

ARTICLE INFO

Article history:
Received 8 June 2016
Received in revised form 24 January 2018
Accepted 25 January 2018
Available online 10 February 2018

Keywords:
Cardiomyopathy
Extracellular matrix
Microarray
Network analysis
Fibrosis

ABSTRACT

Myocardial fibrosis reside a common pathological feature in hypertrophic and dilated cardiomyopathy that results in ventricular dysfunction leading to heart failure. Though several studies reported the role of fibrosis in cardiac diseases, their pathologic mechanisms leading to heart failure remains unclear. A few studies have proposed integrated analysis of microarray information and protein–protein interaction (PPI) systems to discover subnetwork markers related to diagnosis and prognosis of the disease. In addition to PPI networks, we incorporated miRNAs and transcription factors to find the putative miRNAs and transcription factors that might regulate the pathological process and progression of cardiomyopathy and their further progression to heart failure. The important submodules from network revealed the significance of Small Leucine Rich Proteoglycans (SLRPs), Extracellular matrix (ECM) related proteins and complement system in fibrosis. Sequence analysis of different SLRPs suggest that Keratocan and Fibromodulin possesses the same collagen binding site. A predicted mechanism of TGFβ1 shows the involvement of different pathway of HCM and DCM in progression of heart failure.

1. Introduction

Cardiomyopathies constitute a group of diseases, often with unclear etiology, in which the established component is the contribution of the heart muscle itself (Wynne and Braunwald, 2001). They can be classified as dilated, hypertrophic, restrictive, arythmogenic right ventricular and unclassified by which one form can move forward to another (National Institute of Health, 2016). Among the extensive variety of inherited and acquired cardiomyopathies described by the World Health Organization (WHO), Hypertrophic and Dilated were observed to be the most significant disease states. Hypertrophic cardiomyopathy (HCM) results in irregular thickening of heart muscle, particularly the left ventricle. Though dilated cardiomyopathy (DCM) leads to the enlargement and weakening of the heart, both bring about interruption to the electrical functioning of the heart (Towbin and Lorts, 2011). Both HCM and DCM are the reasons for sudden cardiovascular deaths and an essential reason for heart failure (Gajewski and Saul, 2010; Elliott and McKenna, 2004).

Hypertrophic may start as hypertrophic and get to be dilated (Kim et al., 2016), which requires the detection of common hereditary components responsible for the disease to develop better quality treatment methods. In spite of the fact that progression from HCM to ventricular dilation with systolic and diastolic dysfunction has been seen in a subset of HCM population, HCM and DCM are clinically perceived as distinct diseases (Freeman et al., 2001). A key mechanism for poor outcomes is accepted to be myocardial fibrosis, an obsessive hallmark of the condition. Assessing the mechanism in this manner by finding the overlapping genes that are deregulated in DCM and HCM may give new knowledge into their regulatory mechanism. To infer the significance of the deregulated genes, we searched for the network modules that play a role in progression of heart failure.

Network analysis offers a platform to explore efficiently not only the molecular complexity of a particular disease leading to the detection of disease modules and pathways, but also additionally helps in detecting the molecular associations among apparently distinct phenotypes. Advances in this direction are important for identifying novel disease genes, for revealing the biological significance of disease-associated mutations identified by genome-wide association studies (GWAS) and full-genome

* Corresponding author.

https://doi.org/10.1016/j.compbiolchem.2018.01.011
1476-9271/

sequencing, and also for identifying drug targets and biomarkers for complex diseases (Barabasi et al., 2011).

2. Materials and methods

We used the differentially expressed genes (DEGs) obtained from our previous analysis (Malgija and Shanmughavel, 2015) for network construction. Different datasets from Gene expression Omnibus (GEO) for DCM (GSE3585) and HCM (GSE1145) were used. Both the datasets were compared against a same set of normal samples. We observed the altered expression of 758 genes in HCM and 242 in DCM. 54 DEGs including 40 up and 14 down regulated genes were commonly deregulated. Here, we performed a detailed network based analysis of our previous study inorder to find the putative modules associated with the pathologic behavior in HCM and DCM.

2.1. Functional annotation

The functionally enriched Gene Ontology (GO) terms such as biological process (BP), molecular function (MF) and cellular component were determined using various tools like Blast2go (Conesa et al., 2005), DAVID (Huang et al., 2009) and GO consortium. Significant cut-off with P-value < 0.05 was considered.

2.2. Transcription factor (TF) and miRNA prediction

Identification of TFs intervening the co-expressed genes provides possible insight into the structure of regulatory networks. DEGs are searched for their transcription factor activity using TFcheckpt (Chawla et al., 2013). The transcription factors enriched by the DEGs were identified by their search for transcription factor binding sites (TFBS) using oPOSSUM (Sui et al., 2005) and Geniouspro (Kearse et al., 2012). The whole set of miRNAs and their respective target genes were downloaded from miR2Base (Griffiths-Jones et al., 2008). The screened DEGs from both HCM and DCM were extracted with their respective miRNAs using Java script.

2.3. Network construction using cytoscape

The neighbourhood interactive partners of the common DEGs were obtained from STRING (Mering et al., 2003), a database of predicted functional association between proteins. The partners with confidence score above 0.8 and with experimental proof were considered for further construction of protein–protein interaction networks. Proteins with no such interactive partners and those beyond this cut-off were not considered. Based on these interactive partners and their confidence scores, protein–protein interaction network was constructed using Cytoscape for the DEGs. Similarly, the TF-miRNA-gene network was also constructed based on the above prediction results. Cytoscape is an open source platform for the analysis and visualization of complex biomolecular interaction networks with high throughput gene expression data and other molecular states (Shannon et al., 2003).

3. Results

The principal finding of this study is that cardiomyopathies of different etiologies exhibit both shared and diverse changes in gene expression pattern compared with healthy hearts in which only the shared genes are concentrated. Two hundred and fourty two transcripts, including 148 upregulated and 94 down regulated were found to be differentially expressed in DCM comparable with NF heart. Similarly, 758 genes were significantly differentially expressed (310 up regulated and 448 down regulated) in HCM. Fifty four genes were shared in both cases and the integration of these gene expression data with network features revealed their importance in fibrotic mechanism. Instead of identifying a gene expression profile as a diagnostic marker, the current study focuses on gene discovery that identify differentially expressed genes to gain more insights into the similarities and differences between HCM and DCM.The heatmap showing the common significant genes from both disease cases is shown in Fig. 1.

3.1. Functional significance of the common DEGs

The significant genes were analysed based on their enriched GO terms using Gene Ontology consortium. Majority of the 54 common genes fell into functional classes of extracellular matrix organization, immune response and cell adhesion. GO cellular component predicts most of the genes in the extracellular region (Fig. 2). Also alteration of genes related to various important processes such as cardiac system growth and development was observed. GO molecular function showed the importance of heparin binding, glycosaminoglycan binding and sulphur compound binding (Fig. 2). Apart from gene based GO annotations, we performed sequence based GO enrichment analysis using Blast2GO.

Genes involving various cellular processes reported in coronary artery disease namely endothelin integrity (THBS2) (Mannarino and Pirro, 2008), Lipid metabolism (PLA2G2A) (Breslow, 2000), immune and inflammation (CFH, CXCR4, CCL2; Galkina and Ley, 2009) and collagen formation (COL1A2, COL1A1, COL5A1; Menon et al., 2009; Khan and Sheppard, 2006) were commonly altered.

3.2. Prediction of TF and miRNA targets

Gene expression in humans that regulate specific cell states are controlled by thousands of transcription factors and cofactors and their mis-regulation can lead to a wide range of diseases, including cardiovascular diseases (Roger et al., 2012). Hence, we searched for the differentially expressed transcription factors using TFcheckpt

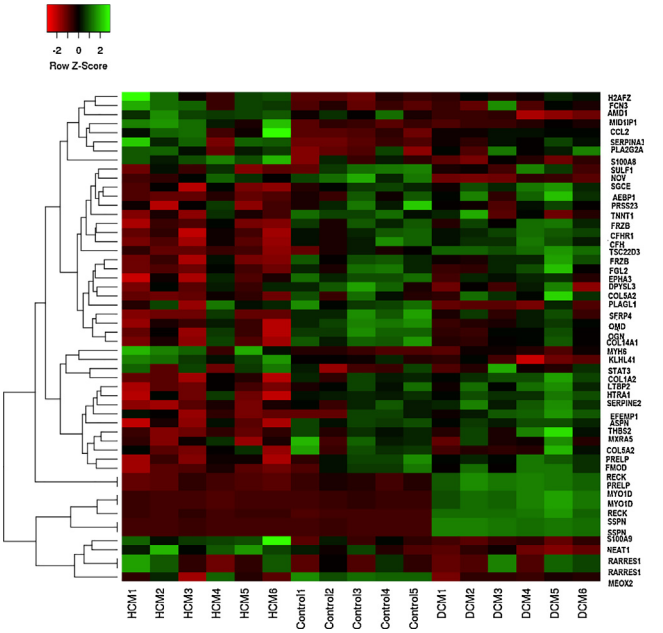

Fig. 1. Heatmap showing the common deregulated genes.
Samples are given horizontally and the genes vertical. Each field is colored based on scores.