Comparative RNA profile analysis of idiopathic dilated cardiomyopathy and ischemic cardiomyopathy


Department of Cardiology, Danyang People’s Hospital of Jiangsu Province, Danyang, Jiangsu Province, China

Corresponding author: X.Q. Liu
E-mail: liuxueqingjs@163.com

Received August 24, 2015
Accepted November 26, 2015
Published March 18, 2016
DOI http://dx.doi.org/10.4238/gmr.15016910

ABSTRACT. Previous research has focused on revealing the functions of each individual gene and/or pathway in idiopathic dilated cardiomyopathy (DCM) or ischemic cardiomyopathy (IC). However, the common or specific pathways of the initiation and processes of DCM and IC are still unclear. Here, we attempted to uncover the critical genes and potential molecular networks that play important roles in DCM and IC progression commonly or specifically. The transcriptional profiles from normal and DCM or IC patient samples were analyzed and compared using bioinformatic methods. Initially, the normal and DCM or IC sample data were processed and the most notable differentially expressed genes (DEGs) from DCM or IC were identified. By comparing the DEGs from DCM with those from IC, the DCM- and IC-specific DEGs were identified. The gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses indicated the significance of multiple biological processes as well as signaling pathways that affect heart function and DCM or IC progression. Protein-protein interaction network analysis identified the relationships between different genes, and some important genes such as MYC and FN1 were found to be hubs, which master each individual module of DCM-specific and IC-specific DEGs, respectively. We discovered commonalities and differences...
of gene expression profiles and molecular pathways between different cardiomyopathies. The gene discovery and molecular signature analysis in this study could offer insights into disease mechanisms and also identify markers useful for diagnostic, prognostic, and therapeutic purposes.

**Key words:** Idiopathic dilated cardiomyopathy; Ischemic cardiomyopathy; Transcriptional profiles; Molecular network

### INTRODUCTION

Despite recent advances in the management of patients with heart failure, morbidity and mortality rates remain high. Common causes of heart failure are ischemic heart disease, uncontrolled hypertension, and valvular disease. Initially, however, in up to 50% of cases, the cause of heart disease remains unknown; this condition is called idiopathic dilated cardiomyopathy (DCM) (Felker et al., 2000). Idiopathic DCM may result from a wide variety of causes, including virus-mediated disease; immune dysregulation; toxic, metabolic, and inherited factors; and tachycardia-induced conditions. DCM affects the heart’s ventricles and atria (Kelly and Strauss, 1994; Hazebroek et al., 2012).

Recently, a variety of evidence has implicated diverse pathways in the mechanism of progression from DCM to overt heart failure. Those pathways include the vascular renin-angiotensin system (Escobales and Crespo, 2008); Gi-coupled receptors (McCloskey et al., 2008); the TGFβ-activin-A/Smad signaling pathway (Mahmoudabady et al., 2008); SH2-containing cytoplasmic tyrosine phosphatase (Princen et al., 2009); apoptotic signaling (Birks et al., 2008; Escobales and Crespo, 2008; Mahmoudabady et al., 2008; Princen et al., 2009; Zheng et al., 2009); and tricarboxylic acid cycle and angiogenesis (Colak et al., 2009).

Ischemic cardiomyopathy (IC) is a condition that occurs when the heart muscle is weakened. In this condition, the left ventricle, which is the main heart muscle, is usually enlarged and dilated. This condition can result from a heart attack, coronary artery disease, or a narrowing of the arteries. Ischemic heart disease is associated with inflammation (Toldo et al., 2014), interstitial fibrosis, and ventricular dysfunction prior to the development of heart failure. It has been claimed that endocannabinoids and the cannabinoid receptor CB2 are involved (Duerr et al., 2014). The endoplasmic reticulum has been linked to several processes involved in cardiovascular diseases including IC and DCM (Ortega et al., 2014).

Despite great efforts to understand the mechanisms involved in the progression from DCM or IC to overt heart failure, the underlying triggering factors for the disease remain unclear. Here, we compared these two kinds of cardiomyopathy to determine commonalities and differences of gene expression profiles between them. The gene discovery and molecular signature analysis in this study could offer insights into disease mechanisms and also identify markers useful for diagnostic, prognostic, and therapeutic purposes.

### MATERIAL AND METHODS

**Data collection**

We searched Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) and obtained one microarray expression dataset (GSE 1145). Three groups of left ventricle biopsy samples were available. One group comprised 27 idiopathic dilated cardiomyopathy samples, a second had 31...
IC samples, and the third had 11 normal samples. Unprocessed data (.cel files) were collected for further analysis. The Affymetrix Human Genome U133 Plus 2.0 Array was used in this experiment. The probe annotation files were downloaded for further research.

Data processing and filtering

The GC Robust Multi-array Average algorithm (Wu et al., 2005) was applied in this research. The normalization process consisted of three major steps: model-based background correction, quantile normalization, and summarizing. The nsFilter function of the genefilter package (Gentleman et al., 2007) in R language was used to filter out uninformative data, such as control probesets and internal controls, as well as to remove genes that were expressed uniformly at background levels. However, the filter does not remove probesets without Entrez Gene identifiers or with identical Entrez Gene identifiers.

Differentially expressed genes (DEGs) analysis

Limma (Smyth, 2005) in the R language was applied to identify significant differences in gene expression level. We performed two statistical comparisons. Comparison 1 was conducted between idiopathic DCM samples and normal samples; Comparison 2 was performed between IC and normal samples. For probes that had identical Entrez Gene identifiers, only the probe that occupied the biggest variance was kept for further DEG analysis. Only those genes with $|\log_2(\text{fold change})| > 1.5$ and adjusted $P$ value $< 0.01$ were recognized as being statistically differentially expressed between two groups. The adjusted $P$ value was obtained by applying Benjamini and Hochberg's false discovery rate correction of the original $P$ value.

The obtained DEGs from two comparisons were compared with each other, and then commonly changed genes in two comparisons were identified.

Hierarchical clustering

Hierarchical clustering was carried out using DEGs to observe the global gene expression patterns (Tavazoie et al., 1999). The DEGs that were classified in specific biological processes [gene ontology (GO) terms] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were further extracted and the expression pattern of those DEGs was characterized; heat maps for the DEGs were classified in targeted biological processes or KEGG pathways using the R package.

GO and KEGG pathway analysis

R packages - GO.db (Carlson et al., 2009), KEGG.db (Carlson et al., 2009), and KEGGREST (Tenenbaum, 2013) - were used to detect GO categories and KEGG pathways. The significantly enriched biological processes were identified as having $P$ values less than the threshold value of 0.01. For the KEGG pathway, the $P$ value threshold was set to less than 0.05.

Construction of biological network

We downloaded protein-protein interaction (PPI) databases from HPRD (Keshava Prasad et
al., 2009), BIOGRID (Chatr-Aryamontri et al., 2013), and PIP (McDowall et al., 2009) databases. Pair interactions were chosen for inclusion in our curated PPI database. As a result, 561,405 pair interactions were found in our database. Cytoscape (Smoot et al., 2011) was utilized to construct an interaction network. Interacting gene pairs in our curated PPI database were imported as a stored network. After functional enrichment analysis, the DEGs specified in dramatically altered biological processes (GO terms) and KEGG pathways were mapped to corresponding networks to analyze interaction.

RESULTS

Identification of DEGs

DEGs analysis was performed as previously described (Smyth, 2005). At fold changes of 1.5 or more and adjusted P values of 0.01 or less, 227 DEGs were identified in Comparison 1, in which 76 genes were upregulated, and 151 genes were downregulated (Table 1). Three-hundred and seventy-five DEGs were obtained in Comparison 2, among which 231 genes were upregulated, and 144 genes were downregulated (Table 1).

Table 1. Statistical distribution of differentially expressed genes (DEGs) in Comparison 1 or Comparison 2.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>GO-BP-ID</th>
<th>Number of DEGs</th>
<th>Fold change</th>
<th>Adjusted P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Upregulated</td>
<td>Downregulated</td>
</tr>
<tr>
<td>Comparison 1</td>
<td>227</td>
<td>76</td>
<td>151</td>
<td>&gt;1.5</td>
</tr>
<tr>
<td>Comparison 2</td>
<td>375</td>
<td>231</td>
<td>144</td>
<td>&gt;1.5</td>
</tr>
</tbody>
</table>

In Comparison 1, the up- or downregulation was determined by comparing gene expression in samples of idiopathic dilated cardiomyopathy (DCM) to that of normal samples. In Comparison 2, the up- or downregulation was determined by comparing gene expression in samples of ischemic cardiomyopathy (IC) with that of normal samples. Comparison 1 gave DEGs for DCM and Comparison 2 gave DEGs for IC.

To understand the common or specific molecular mechanisms of idiopathic DCM and IC, we comparatively analyzed DEGs from the two comparisons. One hundred and thirty-six DEGs were identified as commonly changed genes between DCM and IC, of which 70 genes were upregulated and 66 genes were downregulated. This suggests that the 136 DEGs may represent common pathways for cardiomyopathy. Ninety-one DEGs were only found in Comparison 1, in which six genes were upregulated and 85 genes were downregulated. This indicates that these 91 genes are specifically related to DCM. Meanwhile, 239 DEGs were only found in Comparison 2, in which 161 genes were upregulated and 78 genes were downregulated. This indicates that these 161 genes are specifically related to IC. However, none was upregulated in Comparison 1 or downregulated in Comparison 2 and vice versa (Table 2, Tables S1-S6).

Table 2. Differentially expressed genes (DEGs) in Comparisons 1 and 2 with different directions of regulation. (|log2(fold change)| < -1.5 and adjusted P value < 0.01) (Tables S1 to S6).

<table>
<thead>
<tr>
<th>Changes in Comparisons 1 and 2</th>
<th>Direction of DEGs</th>
<th>Gene count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonly changed</td>
<td>Up</td>
<td>70</td>
</tr>
<tr>
<td>Down</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Changed only in Comparison 1</td>
<td>No</td>
<td>85</td>
</tr>
<tr>
<td>Up</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>Changed only in Comparison 2</td>
<td>No</td>
<td>78</td>
</tr>
<tr>
<td>No</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>Opposite direction of change</td>
<td>Up</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Comparative RNA profile analysis of DCM and IC

Construction of biological network

After identifying genes whose transcript level was significantly altered in idiopathic DCM samples or IC samples compared with normal samples, we investigated the relationship among those DEGs to understand how those genes coordinate as molecular groups to guide the development of idiopathic DCM or IC.

Significantly regulated genes identified in the two comparisons were used to construct corresponding biological networks. The network for Comparison 2 was more complicated and focused than that for Comparison 1. In Comparison 1, DEGs formed simple and disrupted network and genes such as MYC and CDK2, SYNCRIP and SNRPA1, THBS1, and AR constituted hubs of the small sub-networks (Figure 1A); while in Comparison 2, FN1 was obviously the hub of this huge network (Figure 1B).

Figure 1. Biological network of differentially expressed genes in dilated cardiomyopathy or ischemic cardiomyopathy.
Because commonly changed genes in both diseases and those only changed in one disease were identified, corresponding biological networks were also constructed. In the biological network of commonly changed genes, the downregulated SNRPA1 and THBS1 genes constituted two individual network hubs (Figure 2A), which indicates significant roles for SNRPA1 and THBS1 in the development of cardiomyopathies. In the biological network of genes only changed in DCM (Figure 2B), MYC was downregulated and formed its hub, whereas in that only changed in IC, FN1 was downregulated and formed its hub (Figure 2C).

Figure 2. Biological network of commonly changed differentially expressed genes (DEGs) and dilated cardiomyopathy- or ischemic cardiomyopathy-specific DEGs.
GO and KEGG pathway analysis

GO analysis was performed to determine the gene set enrichment in biological processes and molecular functions. We found that 86 biological processes and 14 KEGG pathways were over-represented by commonly changed DEGs. The top four significant biological processes included regulation of multicellular organism processes, immune system processes, extracellular matrix organization, and single-multicellular organism processes (Table 3 and Table S7), whereas the top four significant KEGG pathways were hypertrophic cardiomyopathy, dilated cardiomyopathy, the p53 signaling pathway, and phagosome and arrhythmogenic right ventricular cardiomyopathy (Table 4 and Table S8). These biological processes and genes of the KEGG pathway may be involved in common mechanisms of cardiomyopathic development.

Table 3. Top significantly changed gene ontology (GO) biological processes of differentially expressed genes (DEGs) changed in different datasets (Tables S7, S9, and S11).

<table>
<thead>
<tr>
<th>Datasets</th>
<th>GO-BP-ID</th>
<th>Term</th>
<th>Count</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>In commonly changed DEGs</td>
<td>GO:0051239</td>
<td>Regulation of multicellular organism process</td>
<td>31</td>
<td>2.04E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0003012</td>
<td>Muscle system process</td>
<td>12</td>
<td>3.91E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0022376</td>
<td>Immune system process</td>
<td>60</td>
<td>8.28E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0010810</td>
<td>Regulation of cell-substrate adhesion</td>
<td>31</td>
<td>8.60E-07</td>
</tr>
<tr>
<td>Only in Comparison 1</td>
<td>GO:006954</td>
<td>Regulation of multicellular organism process</td>
<td>11</td>
<td>9.42E-05</td>
</tr>
<tr>
<td>Only in Comparison 2</td>
<td>GO:003198</td>
<td>Extracellular matrix organization</td>
<td>20</td>
<td>6.86E-08</td>
</tr>
<tr>
<td></td>
<td>GO:0043062</td>
<td>Extracellular structure organization</td>
<td>20</td>
<td>7.17E-08</td>
</tr>
<tr>
<td></td>
<td>GO:0071229</td>
<td>Cellular response to acid chemical</td>
<td>10</td>
<td>1.91E-05</td>
</tr>
<tr>
<td></td>
<td>GO:0067044</td>
<td>Cell-substrate junction assembly</td>
<td>7</td>
<td>4.14E-05</td>
</tr>
</tbody>
</table>

Furthermore, we used DCM-specific DEGs (only in Comparison 1) or IC-specific DEGs (only in Comparison 2) to analyze specific biological processes and KEGG pathways in DCM or IC.

In DCM-specific DEGs (only changed in Comparison 1), one biological process and 19 KEGG pathways were found. The one biological process was inflammatory response, which was over-represented by a total of 11 genes, ALOX5, CD59, CHI3L1, S1PR3, CXCL8, PTX3, S100A8, TNFAIP6, CD163, NFKBIZ, and ITCH (Table 3 and Table S9). It should be noted that all these 11 genes were downregulated. Immunoregulatory disorders were seen in DCM. This suggests the important role of these genes in maintaining normal cellular immunity to prevent viral or autoimmune myocarditis-induced DCM. The top four significant KEGG pathways were bladder cancer, one-carbon pool by folate, thyroid cancer, and progesterone-mediated oocyte maturation, progesterone-mediated oocyte maturation, progesterone-mediated oocyte maturation, progesterone-mediated oocyte maturation.
which were over-represented by CXCL8, KRAS, MYC, GART, SHMT2, CDK2, and ANAPC5 (Table 4 and Table S10). Interestingly, all these genes were also downregulated. The oncogenes KRAS and MYC and the cell cycle control gene CDK2 play important roles in cell proliferation. Inhibition of these genes often causes apoptosis.

In IC-specific DEGs (only changed in Comparison 2), 30 biological processes and eight KEGG pathways were found. The top four significant biological processes included extracellular matrix organization, extracellular structure organization, cellular response to acid chemical, and cell-substrate junction assembly, totaling 31 genes involved in these biological processes (Table 3 and Table S11). The top four significant KEGG pathways were focal adhesion, leukocyte transendothelial migration, extracellular matrix-receptor interaction, and Fc gamma R-mediated phagocytosis (Table 4 and Table S12). This suggests the important roles of these pathways and genes in the pathogenesis of IC.

These biological processes and genes of the KEGG pathway may determine the specific characteristics of DCM or IC and the development of specific cardiomyopathies with other common pathways (Figure 3).

Figure 3. Proposed mechanism of dilated cardiomyopathy (DCM) and ischemic cardiomyopathy (IC) pathogenesis.

**DISCUSSION**

Microarray experiments are a suitable method for analyzing the global expression of genes involved in human diseases, such as cardiomyopathy, and show alterations in gene expression profiles. Moreover, several studies have made use of this method to examine expression levels of genes related to DCM (Yung et al., 2004; Barth et al., 2006; Colak et al., 2009). In this paper, we
first investigated the DEGs of DCM (Comparison 1) and IC (Comparison 2) by comparing their RNA profiles with that of normal tissues. We then compared the DEGs of DCM with those of IC; thus, common DEGs and DEGs found only in DCM or IC were obtained.

GO analysis revealed the common and disease-specific biological processes of the two different cardiomyopathies. Furthermore, KEGG pathway analysis revealed candidate genes and their over-represented pathways (KEGG term).

In the DCM-specific biological process analysis, only the inflammatory response pathway was found, which was represented by 11 genes. Inflammation is an important component in the pathogenesis of many common cardiovascular diseases. In most cases, the role of inflammation is a natural response to injury, and an important mechanism for healing and tissue repair. Accumulating data have revealed the important role of inflammatory response in the pathogenesis of DCM (Pankuweit et al., 2004). The 11 genes of the inflammatory response pathway were all downregulated. This suggests a role of immunoregulatory disorders in the development of DCM.

It has long been believed that apoptosis does not occur in terminally differentiated cells such as adult cardiomyocytes. However, all mechanisms responsible for the induction of apoptosis are operative in myocytes and are particularly activated during heart failure (Narula et al., 2000). Using the KEGG pathway analysis program, we identified 18 DCM-specific KEGG pathways. A great deal of evidence has already shown that some genes in DCM-specific KEGG pathways, such as IL8, KRAS, MYC, GART, and CDK2, are associated with some types of heart diseases and tissue abnormalities (Verhaar et al., 2002; Berthet et al., 2006; Niibori et al., 2006; Kerkela et al., 2008; Velásquez et al., 2014). Oncogenes KRAS and MYC, and the cell cycle control gene CDK2, were found in the top DCM-specific KEGG pathways. Decreased expression of these genes in myocytes may increase susceptibility to apoptosis.

The immune system process was common in both DCM and IC. In contrast, only the inflammatory response pathway was found to be DCM specific, and 30 biological processes were found in IC-specific DEGs. The top processes were related to the extracellular matrix and structure. This suggests that the abnormal extracellular matrix and structure and its metabolism may play important roles in the pathogenesis of IC. The top four IC-specific KEGG pathways were over-represented by 17 genes. Among these genes, FLNC, FN1, and VASP have already been reported to be associated with different cardiomyopathies (Pontén et al., 2005; Kong et al., 2010; Tsoutsman et al., 2013).

To obtain a deeper insight into the interactions of the DEGs, the DEGs of DCM and DCM-specific DEGs or the DEGs of IC and IC-specific DEGs were mapped to the gene networks. To date, in the networks constructed by commonly changed genes, the association of SNRPA1 and THBS1 - the hubs of the networks - with cardiomyopathy or other heart diseases remains unknown. MYC, the hub of the network of DCM-specific genes, was found to play an important role in regulating cardiac metabolism (Ahuja et al., 2010). MYC is upregulated rapidly in response to virtually all hypertrophic stimuli and leads to hypertrophy (Lee et al., 2009). MYC mRNA is increased in hypertrophic cardiomyopathy patients (Kai et al., 1998). Here, we found that the RNA expression of MYC was decreased in idiopathic DCM samples compared with normal samples. This suggests that MYC may play different roles in DCM and hypertrophic cardiomyopathy. Fibronectin is a high-molecular weight (~440 kDa) glycoprotein of the extracellular matrix. It is involved in cell adhesion, growth, migration, and differentiation (Pankov and Yamada, 2002). Our findings suggest that FN1 may play a key role in IC.

In summary, evaluation of global gene expression patterns between DCM and IC provides a molecular depiction specific to the two different heart diseases, yields insights into the
pathophysiological aspects of heart muscle disease, and identifies novel genes and pathways whose cardiac-related functions have yet to be deciphered. The present study reveals not only the common genes and pathways of the two heart muscle diseases, but also the specific genes and pathways of each cardiomyopathy.

CONCLUSION

Our study discovered commonalities and differences of gene expression profiles and molecular pathways between different cardiomyopathies. The gene discovery and pathway analysis in this study could offer insights into disease mechanisms and also identify markers useful for diagnostic, prognostic, and therapeutic purposes.

Conflicts of interest

The authors declare no conflict of interest.

REFERENCES


Supplementary material

Table S1. Commonly up-regulated genes in comparison 1 and comparison 2 (log2(fold change) > 1.5 && adjusted P value < 0.01).

Table S2. Commonly down-regulated genes in comparison 1 and comparison 2 (log2(fold change) < -1.5 & adjusted P value < 0.01).

Table S3. Genes that were up-regulated in comparison 1 and have no change in comparison 2 (log2(fold change) > 1.5 && adjusted P value < 0.01).

Table S4. Genes that were down-regulated in comparison 1 and have no change in comparison 2 (log2(fold change) < -1.5 && adjusted P value < 0.01).

Table S5. Genes that were up-regulated in comparison 2 and have no change in comparison 1 (log2(fold change) > 1.5 && adjusted P value < 0.01).

Table S6. Genes that were down-regulated in comparison 2 and of no change in comparison 1 (log2(fold change) < -1.5 && adjusted P value < 0.01).

Table S7. Significantly changed GO biological processes in commonly changed DEGs.

Table S8. Significantly changed KEGG pathways in commonly changed DEGs.

Table S9. Significantly changed GO biological processes in DEGs only changed in comparison 1.

Table S10. Significantly changed KEGG pathways in DEGs only changed in comparison 1.

Table S11. Significantly changed GO biological processes in DEGs only changed in comparison 2.

Table S12. Significantly changed KEGG pathways in DEGs changed in comparison 2.