Bioinformatics method identifies potential biomarkers of dilated cardiomyopathy in a human induced pluripotent stem cell-derived cardiomyocyte model

YU ZHUANG[1], YU-JIA GONG[2], BEI-FEN ZHONG[3], YI ZHOU[1] and LI GONG[4]

[1]Department of Cardiovascular Surgery, Shanghai First People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200080; [2]Stomatology Faculty, School of Medicine, Nantong University, Nantong, Jiangsu 226000; [3]Department of Urology, Shanghai First People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200080; [4]Department of Cardiothoracic Surgery, The Affiliated Huai'an Hospital of Xuzhou Medical University, Huai'an, Jiangsu 223002, P.R. China

Received February 19, 2016; Accepted February 10, 2017

DOI: 10.3892/etm.2017.4850

Abstract. Dilated cardiomyopathy (DCM) is the most common type of cardiomyopathy that account for the majority of heart failure cases. The present study aimed to reveal the underlying molecular mechanisms of DCM and provide potential biomarkers for detection of this condition. The public dataset of GSE35108 was downloaded, and 4 normal induced pluripotent stem cell (iPSC)-derived cardiomyocytes (N samples) and 4 DCM iPSC-derived cardiomyocytes (DCM samples) were utilized. Raw data were preprocessed, followed by identification of differentially expressed genes (DEGs) between N and DCM samples. Crucial functions and pathway enrichment analysis of DEGs were investigated, and protein-protein interaction (PPI) network analysis was conducted. Furthermore, a module network was extracted from the PPI network, followed by enrichment analysis. A set of 363 DEGs were identified, including 253 upregulated and 110 downregulated genes. Several biological processes (BPs), such as blood vessel development and vasculature development (*FLT1* and *MMP2*), cell adhesion (*CDH1*, *ITGB6*, *COL6A3*, *COL6A1* and *LAMC2*) and extracellular matrix (ECM)-receptor interaction pathway (*CDH1*, *ITGB6*, *COL6A3*, *COL6A1* and *LAMC2*), were significantly enriched by these DEGs. Among them, MMP2, CDH1 and FLT1 were hub nodes in the PPI network, while COL6A3, COL6A1, LAMC2 and ITGB6 were highlighted in module 3 network. In addition, PENK and APLNR were two crucial nodes in module 2, which were linked to each other. In conclusion, several potential biomarkers for DCM were identified, such as *MMP2*, *FLT1*, *CDH1*, *ITGB6*, *COL6A3*, *COL6A1*, *LAMC2*, *PENK* and *APLNR*. These genes may serve significant roles in DCM via involvement of various BPs, such as blood vessel and vasculature development and cell adhesion, and the ECM-receptor interaction pathway.

Correspondence to: Dr Li Gong, Department of Cardiothoracic Surgery, The Affiliated Huai'an Hospital of Xuzhou Medical University, 62 South Huaihai Road, Huai'an, Jiangsu 223002, P.R. China
E-mail: 1740385802@qq.com

Key words: dilated cardiomyopathy, cell adhesion, extracellular matrix-receptor interaction, induced pluripotent stem cell, enrichment analysis, protein-protein interaction, bioinformatics

Introduction

Cardiomyopathy is the leading cause of heart failure and mainly consists of four types: Dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), restrictive cardiomyopathy and arrhythmogenic right ventricular cardiomyopathy (ARVC) (1). Characterized by enlarged ventricular dimensions and systolic dysfunction (2), DCM is the most frequent type of cardiomyopathy, accounting for 10% of the newly diagnosed heart failure cases (3). Management of DCM is an economic burden worldwide (4). Therefore, it is urgent to identify the genetic etiologies of this disease.

Genetic mutations are known to influence the majority of cardiomyopathies, and DCM is triggered by mutations in giant protein titin (5). Reportedly, DCM is associated with mutations in several genes, such as *DES*, *LMNA*, *SCN5A* and *TNNT2* (6-8). Among them, *TNNT2* is the gene that encodes for cardiac troponin T, one of the subunits of the troponin complex that regulates muscle contraction and sarcomere assembly (9). Mouse models carrying *TNNT2* mutations that denote the human DCM phenotype have been established and used to determine the underlying mechanisms of DCM (10). However, due to the differences between mouse and human models, the interpretation of such findings may be limited.

In mice, four transcription factors, namely *Oct3/4*, *Sox2*, *Klf4* and *c-Myc*, are essential to generate the induced pluripotent stem cells (iPSCs). By contrast, in humans, the four indispensable factors for iPSCs generation from fibroblasts are *OCT4*, *SOX2*, *NANOG* and *LIN28* (11). This enables the establishment of disease-specific cellular models in humans, which contributes to the understanding of mechanisms and

suitable drug screening for certain diseases. In humans, iPSCs derived from DCM patients have been used to obtain cardiomyocytes (12). A previous study created a human DCM iPSC-derived cardiomyocyte model, which was generated from DCM patients carrying the *TNNT2* mutation and from healthy controls without this mutation, in order to investigate relevant DCM mechanisms and potential drug therapy (13). As a result, the study discovered that, in comparison with the healthy control group, the iPSC-derived cardiomyocytes from DCM patients showed altered regulation of Ca^{2+} and abnormal distribution of sarcomeric α-actin (13). Despite these valuable findings, gene alterations in DCM cardiomyocytes, as well as regulatory correlations of these genes were not considered.

Therefore, the present study re-analyzed the dataset GSE35108, mainly focusing on the gene alterations, as well as their biological functions and potential correlations in human DCM. Differentially expressed genes (DEGs) were identified between DCM iPSC-derived cardiomyocyte samples and healthy controls, followed by enrichment analysis. In addition, a protein-protein interaction (PPI) network was established to explore potential correlations between these DEGs, and three module networks were further extracted to explore detailed crucial genes and their functions. Through these comprehensive bioinformatics methods, the current study aimed to reveal underlying molecular mechanisms of DCM and to provide potential biomarkers for the detection and determination of the prognosis of the disease.

Materials and methods

Data resource. Gene expression profile with the accession number of GSE35108 (13) was obtained from the public Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/). The dataset contained a total of 17 samples; among them, 4 normal iPSC-derived cardiomyocytes (N samples) and 4 DCM iPSC-derived cardiomyocytes (DCM samples) were utilized in the present study. The platform for the gene expression profiles was Affymetrix Human Gene 1.0 ST Array (Affymetrix Inc., Santa Clara, CA, USA).

Data preprocessing and differential analysis. The Bioconductor oligo package (http://www.bioconductor.org/packages/release/bioc/html/oligo.html) was used to perform the pretreatment analysis (14). Background correction and quantile normalization were performed for the raw data using the robust multichip average (RMA) algorithm (15). Next, expression values at probe level (probe ID) were converted to gene level (gene symbols). After removing the redundant probes, the gene expression matrix was finally obtained.

The Linear Models for Microarray Analysis (Limma) package (16) in R software (http://www.bioconductor.org/packages/release/bioc/html/limma.html) was applied to identify DEGs between the N and DCM samples, based on Student's t-test. Adjusted P-values were calculated using the Benjamini-Hochberg method (17). The thresholds for significant DEG selection were an adjusted P value of <0.05 and |log2 fold change (FC)|of >2.

Functional enrichment analysis. To explore potential functions and pathways that may be altered by the DEGs, the online tool Database for Annotation, Visualization and Int egration Discovery (DAVID; https://david.ncifcrf.gov/) (18) was selected to perform functional and pathway enrichment analyses of the identified DEGs. The Gene Ontology (GO; http://www.geneontology.org/) database was used to identify the biological processes (BPs) that the DEGs may be involved. In addition, the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/pathway.html) database was used to perform pathway enrichment analysis for the identified DEGs, according to the modified Fisher exact test. The selection criteria for significant GO terms and pathways were P<0.05 and an enriched gene number (also known as count) of >2.

PPI network construction. To further investigate potential associations of these DEGs at the protein level, the DEGs were mapped into the Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org/) database. Subsequently, a PPI network was constructed, with the criterion of combined score of >0.4 between two interplayed proteins.

The Cytoscape software (http://cytoscape.org/) was used to perform topological structure analysis of the PPI network. A node in the PPI network was deemed as a protein product of the DEG. Based on the Connectivity Degree analysis (19), crucial nodes in the PPI network with high degree (where degree refers to the number of edges for a DEG in the PPI network) were highlighted and named as hub nodes.

Module analysis of the PPI network. Module networks were extracted from the PPI network by the ClusterONE software (http://www.paccanarolab.org/clusterone/) in the Cytoscape, with the default parameter. The sub-modules network with P<0.05 was considered as significant functional sub-modules. Thereafter, functional enrichment analysis was performed for genes in the module networks, according to aforementioned method.

Results

DEGs between N and DCM samples. According to the criteria of adjusted P-value of <0.05 and |log2 FC| of >2, a set of 363 DEGs were identified in the DCM samples, in comparison with the N samples. Among them, 253 genes were upregulated, while 110 were downregulated.

Altered function and pathway by these DEGs. As indicated in Table I, upregulated DEGs were significantly correlated with various BPs, including ectoderm or epidermis development (e.g., *KLF4*), blood vessel and vasculature development (e.g., *FLT1*, *MMP2* and *COL1A2*), as well as cell and biological adhesion (e.g., *CDH1*, *ITGB6*, *COL6A3*, *COL6A1* and *LAMC2*). By contrast, the downregulated DEGs were markedly enriched in BPs, such as muscle contraction and muscle system processes (e.g., *FXYD1*, *SLC8A1*, *CRYAB*), and muscle organ development (e.g., *ITGA11*) and heart development (e.g., *XIRP1*, *MYL2*, *MYBPC3*).

Based on the KEGG pathway enrichment analysis, five predominant pathways for upregulated DEGs were identified, as follows: Extracellular matrix (ECM)-receptor interaction (e.g., *COL6A3*, *ITGB6*, *COL1A2*, *COL6A1*), focal adhesion

Table I. Significantly enriched biological processes of the differentially expressed genes.

A. Upregulated genes

Description	Count	Genes	P-value
GO:0007398-ectoderm development	15	*KRT6A*, *KRT6B*, *PTGS2*, *SPRR2F*, *GRHL3*, *SPRR2E*, *SCEL*, *PTHLH*, *FLG*, *COL1A2*, *SPRR3*, *LAMC2*, *KLK14*, *KLF4*, *EMP1*	$7.18x10^{-7}$
GO:0008544-epidermis development	13	*PTGS2*, *SPRR2F*, *GRHL3*, *SPRR2E*, *SCEL*, *PTHLH*, *FLG*, *COL1A2*, *SPRR3*, *LAMC2*, *KLF4*, *EMP1*, *KLK14*	$9.95x10^{-6}$
GO:0001568-blood vessel development	13	*PLAT*, *APOB*, *FLT1*, *EPAS1*, *HAND1*, *TDGF1*, *COL1A2*, *SEMA3C*, *TGFBR3*, *ANPEP*, *PLXND1*, *MMP2*, *CEACAM1*	$1.62x10^{-4}$
GO:0001944-vasculature development	13	*PLAT*, *APOB*, *FLT1*, *EPAS1*, *HAND1*, *TDGF1*, *COL1A2*, *SEMA3C*, *TGFBR3*, *ANPEP*, *PLXND1*, *MMP2*, *CEACAM1*	$2.03x10^{-4}$
GO:0010035-response to inorganic substance	11	*APOB*, *PTGS2*, *FGA*, *OLR1*, *FGB*, *DUOX2*, *MGP*, *CA2*, *CAT*, *ITPR3*, *AQP3*	$5.90x10^{-4}$
GO:0007155-cell adhesion	22	*CLDN7*, *CLDN18*, *CLCA2*, *MPZL2*, *TNXB*, *OLR1*, *MGP*, *CDH1*, *VTN*, *PCDH18*, *CD9*, *NLGN4Y*, *CD34*, *DSG3*, *TEK*, *ITGB6*, *COL6A3*, *DSC3*, *COL6A1*, *LAMC2*, *SPON2*, *CEACAM1*	$7.33x10^{-4}$
GO:0022610-biological adhesion	22	*CLDN7*, *CLDN18*, *CLCA2*, *MPZL2*, *TNXB*, *OLR1*, *MGP*, *CDH1*, *VTN*, *PCDH18*, *CD9*, *NLGN4Y*, *CD34*, *DSG3*, *TEK*, *ITGB6*, *COL6A3*, *DSC3*, *COL6A1*, *LAMC2*, *SPON2*, *CEACAM1*	$7.46x10^{-4}$
GO:0060541-respiratory system development	8	*WNT2*, *PTHLH*, *RBP4*, *LIPA*, *EPAS1*, *CRISPLD2*, *ANO1*, *MGP*	$7.66x10^{-4}$
GO:0035295-tube development	11	*WNT2, PTHLH, RBP4, FLT1, LIPA, GPC3, HAND1, EPAS1, CRISPLD2, MGP, PLXND1*	$1.02x10^{-3}$
GO:0016337-cell-cell adhesion	12	*CLDN7*, *CLDN18*, *MPZL2*, *OLR1*, *CD34*, *DSG3*, *TEK*, *DSC3*, *MGP*, *CDH1*, *CEACAM1*, *PCDH18*	$1.67x10^{-3}$

B. Downregulated genes

Description	Count	Genes	P-value
GO:0006936-muscle contraction	11	*FXYD1*, *SLC8A1*, *CRYAB*, *CKMT2*, *MYBPC3*, *RYR2*, *PGAM2*, *SMPX*, *CHRNB1*, *CASQ2*, *MB*	$1.46x10^{-8}$
GO:0003012-muscle system process	11	*FXYD1*, *SLC8A1*, *CRYAB*, *CKMT2*, *MYBPC3*, *RYR2*, *PGAM2*, *SMPX*, *CHRNB1*, *CASQ2*, *MB*	$3.57x10^{-8}$
GO:0006941-striated muscle contraction	6	*MYBPC3*, *RYR2*, *PGAM2*, *SMPX*, *CASQ2*, *MB*	$5.98x10^{-6}$
GO:0007517-muscle organ development	8	*XIRP1*, *MYL2*, *CRYAB*, *DMD*, *MYBPC3*, *ITGB1BP2*, *ITGA11*, *CSRP3*	$1.94x10^{-4}$
GO:0007507-heart development	8	*XIRP1*, *MYL2*, *MYBPC3*, *PDLIM3*, *SMYD1*, *CSRP3*, *CASQ2*, *MB*	$2.18x10^{-4}$
GO:0044057-regulation of system process	9	*SLC8A1*, *CYP2J2*, *MYL2*, *BCHE*, *MYBPC3*, *NLGN1*, *RYR2*, *CNN1*, *CSRP3*	$3.63x10^{-4}$
GO:0031032-actomyosin structure organization	4	*XIRP1*, *MYL2*, *MYOZ1*, *CNN1*	$5.24x10^{-4}$
GO:0030036-actin cytoskeleton organization	7	*XIRP1*, *MYL2*, *PDLIM3*, *FHOD3*, *OPHN1*, *MYOZ1*, *CNN1*	$1.80x10^{-3}$
GO:0030029-actin filament-based process	7	*XIRP1*, *MYL2*, *PDLIM3*, *FHOD3*, *OPHN1*, *MYOZ1*, *CNN1*	$2.48x10^{-3}$

Table I. Continued.

Description	Count	Genes	P-value
GO:0006575-cellular amino acid derivative metabolic process	6	*COLQ*, *BCHE*, *CKMT2*, *GSTT1*, *MOXD1*, *GHR*	2.54×10^{-3}

GO, Gene Ontology; count, gene numbers enriched in a specific GO term.

(e.g., *FLT1*, *COL6A3*, *ITGB6*, *COL1A2*, *COL6A1*), glutathione metabolism (e.g., *GPX2*, *GSTA2*, *ANPEP*), complement and coagulation cascades (e.g., *PLAT*, *FGA*, *FGB*) and O-Glycan biosynthesis (e.g., *GCNT4*, *GALNT12*, *GCNT1*). By contrast, only three significant pathways were enriched for the downregulated DEGs, including HCM, DCM and ARVC (e.g., *ITGA11* in all these pathways; Table II).

Predicted PPI network and module network analysis. According to the predefined criterion, a PPI network was constructed, containing 520 interplayed PPIs. As shown in Fig. 1, the top ten hub nodes in the PPI network were MMP2 (degree, 26), CDH1 (degree, 25), CAT (degree, 21), NANOG (degree, 18), PROM1 (degree, 18), KLF4 (degree, 15), FGF13 (degree, 13), PDE5A (degree, 12), FLT1 (degree, 12) and PDE2A (degree, 12).

Three significant module networks were extracted, including module 1 ($P=5.80 \times 10^{-4}$), module 2 ($P=2.92 \times 10^{-3}$) and module 3 ($P=1.34 \times 10^{-2}$). In module 1, nodes such as CAT, PDE5A, PDE2A and LRRN1 were highlighted; in module 2, PENK, APLNR, BDKRB1, OXGR1 and P2RY14 were predominant and linked to each other; and in module 3, the highlighted nodes were ITGB6, COL6A3, COL1A2, COL6A1 and ITGA11 (Fig. 2). No significant BP terms and pathways were enriched for the nodes in module 1. Enrichment analysis indicated that genes in module 2 were significantly associated with the following BPs (Table III): Cell surface receptor linked signal transduction (*PLAT, APLNR*, *PENK*, *P2RY14*, *BDKRB1* and *OXGR1*), G-protein coupled receptor protein signaling pathway (*APLNR*, *PENK*, *P2RY14*, *BDKRB1* and *OXGR1*), sensory perception of pain (*PENK* and *BDKRB1*), neuroactive ligand-receptor interaction (*APLNR, P2RY14* and *BDKRB1*), and complement and coagulation cascades (*PLAT* and *BDKRB1*). Furthermore, as shown in Table IV, genes in module 3 were significantly associated with BPs such as cell adhesion (*COL6A3*, *ITGB6*, *ITGA11, COL6A1* and *LAMC2*), biological adhesion (*COL6A3*, *ITGB6*, *ITGA11, COL6A1* and *LAMC2*), integrin-mediated signaling pathway (*ITGB6* and *ITGA11*), cell-matrix adhesion (*ITGB6* and *ITGA11*), cell-substrate adhesion (*ITGB6* and *ITGA11*), ECM-receptor interaction (e.g., *COL6A3*, *ITGB6*, *COL1A2, ITGA11*, *COL6A1*, *LAMC2* and *SV2C*) and focal adhesion (e.g., *COL6A3*, *ITGB6*, *COL1A2, ITGA11*, *COL6A1* and *LAMC2*). However, no significant BP terms were enriched based on the set criterion for enrichment analysis.

Discussion

By comparing gene expression profiles in iPSC-derived cardiomyocytes from DCM patients and healthy individuals, the present study identified a set of critical DEGs, such as *FLT1* and *MMP2*, which were upregulated and significantly enriched in the blood vessel development and vasculature development BPs. In addition, *CDH1*, *ITGB6*, *COL6A3*, *COL6A1* and *LAMC2* were upregulated and markedly correlated with the cell adhesion BP and the ECM-receptor interaction pathway. Notably, MMP2, CDH1 and FLT1 were three hub nodes in the PPI network, while COL6A3, COL6A1, LAMC2 and ITGB6 were the crucial nodes in

Table II. Significantly enriched pathways of the differentially expressed genes.

A. Upregulated genes

Description	Count	Genes	P-value
hsa04512: ECM-receptor interaction	7	*TNXB*, *COL6A3*, *ITGB6*, *COL1A2*, *COL6A1*, *LAMC2*, *VTN*	$1.47x10^{-3}$
hsa04510: Focal adhesion	10	*FLT1*, *TNXB*, *COL6A3*, *ITGB6*, *COL1A2*, *COL6A1*, *LAMC2*, *VTN*, *HGF*, *FLNB*	$2.82x10^{-3}$
hsa00512: O-Glycan biosynthesis	4	*GCNT4*, *GALNT12*, *GCNT1*, *ST6GALNAC1*	$9.75x10^{-3}$
hsa04610: Complement and coagulation cascades	5	*PLAT*, *FGA*, *FGB*, *F3*, *BDKRB1*	$1.92x10^{-2}$
hsa00480: Glutathione metabolism	4	*GPX2*, *GSTA2*, *ANPEP*, *GSTO1*	$3.81x10^{-2}$

B. Downregulated genes

Description	Count	Genes	P-value
hsa05410: Hypertrophic cardiomyopathy	7	*SLC8A1*, *MYL2*, *DMD*, *MYBPC3*, *ITGA11*, *RYR2*, *PRKAA2*	$3.22x10^{-6}$
hsa05414: Dilated cardiomyopathy	6	*SLC8A1*, *MYL2*, *DMD*, *MYBPC3*, *ITGA11*, *RYR2*	$8.47x10^{-5}$
hsa05412: Arrhythmogenic right ventricular cardiomyopathy	4	*SLC8A1*, *DMD*, *ITGA11*, *RYR2*	$6.52x10^{-3}$

Count, gene numbers enriched in a specific pathway term; ECM, extracellular matrix.

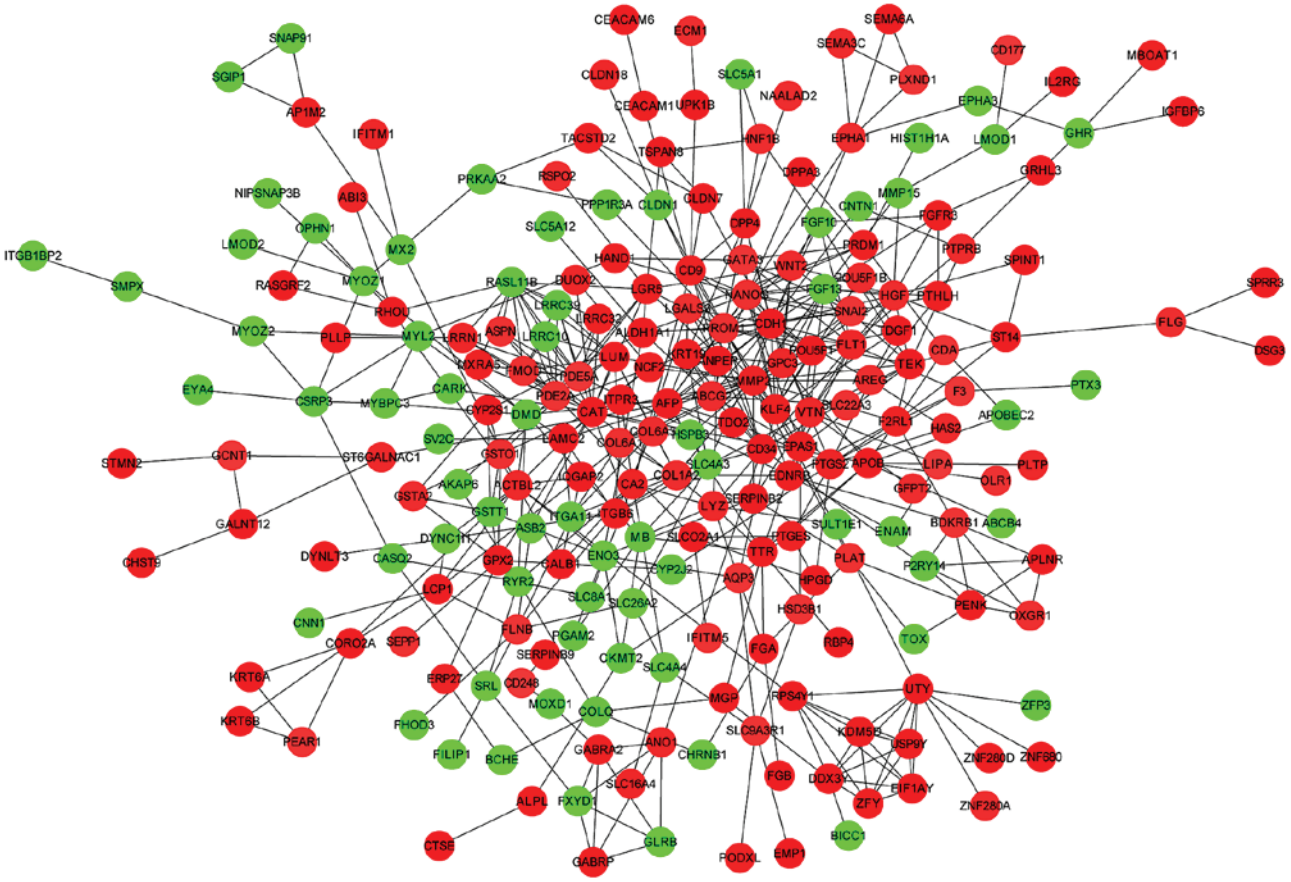

Figure 1. Protein-protein interaction network of the differentially expressed genes. Red nodes indicate upregulated genes, and green nodes represent downregulated genes.

Table III. Significantly enriched biological process and pathway for genes in Module 2.

Category	Description	Count	Genes	P-value
GOBP	GO:0007166-cell surface receptor linked signal transduction	6	*PLAT*, *APLNR*, *PENK*, *P2RY14*, *BDKRB1*, *OXGR1*	$4.84x10^{-5}$
GOBP	GO:0007186-G-protein coupled receptor protein signaling pathway	5	*APLNR*, *PENK*, *P2RY14*, *BDKRB1*, *OXGR1*	$2.21x10^{-4}$
GOBP	GO:0019233-sensory perception of pain	2	*PENK*, *BDKRB1*	$1.18x10^{-2}$
KEGG	hsa04080:Neuroactive ligand-receptor interaction	3	*APLNR*, *P2RY14*, *BDKRB1*	$7.32x10^{-3}$
KEGG	hsa04610:Complement and coagulation cascades	2	*PLAT*, *BDKRB1*	$4.02x10^{-2}$

GO, Gene Ontology; BP, biological process; KEGG, Kyoto Encyclopedia of Genes and Genomes; Count, gene numbers enriched in a specific GO term or pathway term.

Table IV. Significantly enriched biological process and pathway for genes in Module 3.

Category	Description	Count	Genes	P-value
GOBP	GO:0007155-cell adhesion	5	*COL6A3*, *ITGB6*, *ITGA11*, *COL6A1*, *LAMC2*	$9.81x10^{-5}$
GOBP	GO:0022610-biological adhesion	5	*COL6A3*, *ITGB6*, *ITGA11*, *COL6A1*, *LAMC2*	$9.86x10^{0}$
GOBP	GO:0007229-integrin-mediated signaling pathway	2	*ITGB6*, *ITGA11*	$3.07x10^{-2}$
GOBP	GO:0007160-cell-matrix adhesion	2	*ITGB6*, *ITGA11*	$3.88x10^{-2}$
GOBP	GO:0031589-cell-substrate adhesion	2	*ITGB6*, *ITGA11*	$4.27x10^{-2}$
KEGG	hsa04512:ECM-receptor interaction	7	*COL6A3*, *ITGB6*, *COL1A2*, *ITGA11*, *COL6A1*, *LAMC2*, *SV2C*	$1.70x10^{-11}$
KEGG	hsa04510:Focal adhesion	6	*COL6A3*, *ITGB6*, *COL1A2*, *ITGA11*, *COL6A1*, *LAMC2*	$5.34x10^{-7}$

GO, Gene Ontology; BP, biological process; KEGG, Kyoto Encyclopedia of Genes and Genomes; Count, gene numbers enriched in a specific GO term or pathway term; ECM, extracellular matrix.

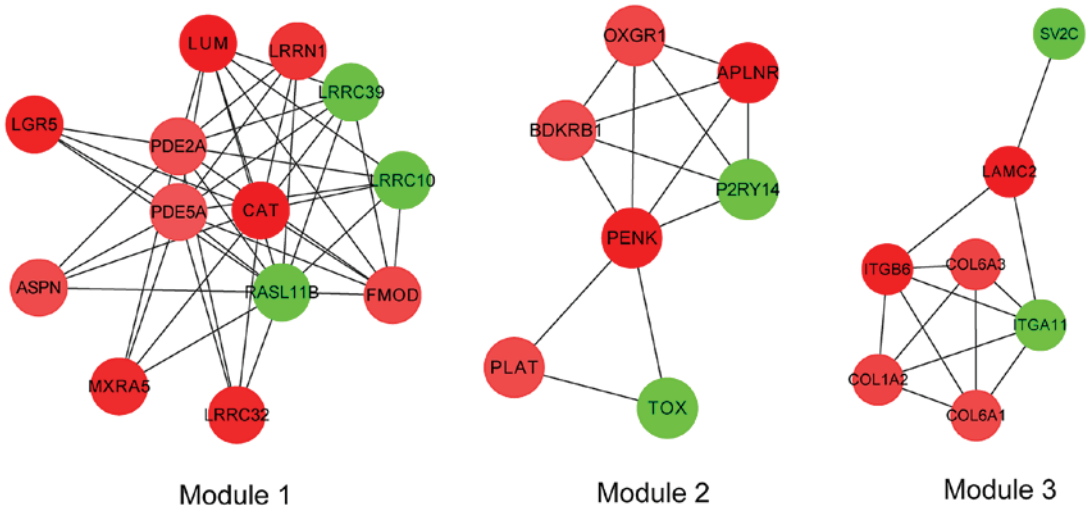

Figure 2. Three module networks of the differentially expressed genes. Red nodes indicate upregulated genes, and green nodes represent downregulated genes.

module 3 network. Furthermore, other crucial nodes were identified in module 2, such as PENK and APLNR, which were linked to each other.

A structural feature of DCM is chamber dilation. Dilation requires myocyte and ECM alteration, which are demonstrated as causative factors for cardiac enlargement (20).

Matrix metalloproteinases (MMPs) are family members of Ca-dependent and Zn-containing endopeptidases that are essential for ECM degradation and remodeling (21). MMP2 is a member of the MMP family that has been identified as a crucial protease in myocardial ischemia (22). Furthermore, MMP2 serves significant roles in infarction, which is associated with fibrous tissue enlargement (23). Reportedly, normal blood vessel development is regulated by signals between ECM and endothelial cells, while MMP2 is mainly secreted by endothelial cells (24). Additionally, MMP2 can enter the vasculature (25). The *FLT1* encoded protein is a family member of the vascular endothelial growth factor receptor. During blood vessel morphogenesis, FLT1 is identified as a crucial molecule for the control of endothelial cell activity (26). Alteration of *FLT1* has been detected in cardiomyopathies by the NHLBI-Go Exome Sequencing Project (27); however, the association between this gene and HCM or DCM needs to be re-evaluated (28). Blood vessels serve important roles in the heart during cardiac disease and remodeling, and several vasculature-associate genes are involved in cardiac hypertrophy (29). DCM has the characteristics of cardiac enlargement and hypertrophy. Therefore, it is not surprising that several upregulated genes in DCM, such as *MMP2* and *FLT1*, were enriched in the blood vessel and vasculature development BPs in the present study. These collectively suggest that the two genes may have significant roles in human DCM, via regulation of the blood vessel and vasculature development.

Two hallmarks of DCM are abnormalities of cellular function and immune system. Overexpression of cell adhesion molecules is a common event in DCM (30). *CDH1* encodes an E-cadherin protein, which is a cell-cell adhesion molecule. Epithelial-to-mesenchymal transition (EMT) is proposed as the key mechanism for the production of cardiovascular progenitor cells, which finally differentiate into endothelial cells and cardiomyocytes, and *CDH1* is one of the prominent regulators of EMT (31). Increased expression of *CDH1* in DCM samples with the enrichment in cell adhesion BP in the current study suggests that *CDH1* may be a crucial gene in DCM that controls the regulation of cell adhesion. *ITGB6*, *LAMC2*, as well as collagens *COL6A1* and *COL6A3*, are other cell adhesion-associated molecules (32). All these genes were enriched in the cell adhesion BP in the present study, suggesting that altered cell adhesion may serve a significant role in DCM.

The *ITGB6* encoded protein is a family member of adhesion receptors that have significant roles in signaling from the ECM to the cell. Intergrins are known as the major receptors for substantial ECM-mediated cellular activities, including cell adhesion, cell proliferation and differentiation (33). Enrichment analysis in a mouse model has indicated that integrins, such as *ITGB6,* can participate in the DCM pathway (34). The findings of the present study indicated that the DEG *ITGB6* in DCM was significantly enriched in the ECM-receptor interaction pathway. Notably, ECM-receptor interaction is also involved in DCM pathway (35). Thus, it can be inferred that ECM-receptor interaction may be an important pathway for DCM, which may be altered by the integrin gene *ITGB6*.

PENK encodes for the proenkephalin protein, and has been reported as one of the DEGs identified in skeletal muscle, while the muscle development-associated DEGs are implicated in the DCM pathway (36). *APLNR* encodes a family member of the G protein-coupled receptors. Several studies have investigated polymorphisms of *APLNR* and the susceptibility of DCM (37-39). In addition, the 212A allele of the *APLNR* G212A polymorphism is considered to be significantly associated with reduced risk of idiopathic DCM (39), suggesting that alteration of this gene may influence the DCM. In the present study, *PENK* and *APLNR* were identified as two crucial genes in module 2, suggesting that they may exert more important functions in DCM than other genes. In addition, the linkage of these two genes in the module 2 network implies potential co-regulation of DCM.

Although a relatively comprehensive bioinformatics analysis was conducted in the present study, there remained several limitations. As the dataset was obtained from the public database, the sample size was small. Furthermore, there lacked validations of DEG expressions, as well as correlations between them. However, the obtained predictive results are also of great value to provide novel insights into the underlying DCM mechanisms and potential biomarkers for DCM prediction.

In conclusion, several potential biomarkers for DCM were identified based on the bioinformatics results of the current study, including *MMP2*, *FLT1*, *CDH1*, *ITGB6*, *COL6A3*, *COL6A1*, *LAMC2*, *PENK* and *APLNR*. These genes may serve significant roles in DCM via the involvement of various BPs, including blood vessel and vasculature development, and cell adhesion, as well as through the pathway of ECM-receptor interaction. However, further experimental validations are required to confirm these findings.

References

1. McNally EM, Golbus JR and Puckelwartz MJ: Genetic mutations and mechanisms in dilated cardiomyopathy. J Clin Invest 123: 19-26, 2013.
2. Maron BJ, Towbin JA, Thiene G, Antzelevitch C, Corrado D, Arnett D, Moss AJ, Seidman CE, Young JB; American Heart Association; Council on Clinical Cardiology, Heart Failure and Transplantation Committee, *et al*: Contemporary definitions and classification of the cardiomyopathies: An American heart association scientific statement from the council on clinical cardiology, heart failure and transplantation committee; quality of care and outcomes research and functional genomics and translational biology interdisciplinary working groups; and council on epidemiology and prevention. Circulation 113: 1807-1816, 2006.
3. Li X, Luo R, Hua W, Li L, Kwong JSW, Chan CP and CM Y: Cardiac resynchronization therapy for dilated cardiomyopathy. Cochrane Libr, 2013.
4. Vergaro G, Del Franco A, Giannoni A, Prontera C, Ripoli A, Barison A, Masci PG, Aquaro GD, Cohen Solal A, Padeletti L, *et al*: Galectin-3 and myocardial fibrosis in nonischemic dilated cardiomyopathy. Int J Cardiol 184: 96-100, 2015.
5. McNally EM, Barefield DY and Puckelwartz MJ: The genetic landscape of cardiomyopathy and its role in heart failure. Cell Metab 21: 174-182, 2015.
6. Hershberger RE and Morales A: LMNA-related dilated cardiomyopathy. In: GeneReviews®. Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH and Stephens K (eds). University of Washington, Seattle, 2008 [updated July 7, 2016].
7. Robyns T, Nuyens D, Van Casteren L, Corveleyn A, De Ravel T, Heidbuchel H and Willems R: Reduced penetrance and variable expression of SCN5A mutations and the importance of co-inherited genetic variants: Case report and review of the literature. Indian Pacing Electrophysiol J 14: 133-149, 2014.

8. Hershberger RE, Pinto JR, Parks SB, Kushner JD, Li D, Ludwigsen S, Cowan J, Morales A, Parvatiyar MS and Potter JD: Clinical and functional characterization of TNNT2 mutations identified in patients with dilated cardiomyopathy. Circ Cardiovasc Genet 2: 306-313, 2009.
9. Sehnert AJ, Huq A, Weinstein BM, Walker C, Fishman M and Stainier DY: Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. Nat Genet 31: 106-110, 2002.
10. Ahmad F, Banerjee SK, Lage ML, Huang XN, Smith SH, Saba S, Rager J, Conner DA, Janczewski AM, Tobita K, *et al*: The role of cardiac troponin T quantity and function in cardiac development and dilated cardiomyopathy. PLoS One 3: e2642, 2008.
11. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, *et al*: Induced pluripotent stem cell lines derived from human somatic cells. Science 318: 1917-1920, 2007.
12. Zhang J, Wilson GF, Soerens AG, Koonce CH, Yu J, Palecek SP, Thomson JA and Kamp TJ: Functional cardiomyocytes derived from human induced pluripotent stem cells. Circ Res 104: e30-e41, 2009.
13. Sun N, Yazawa M, Liu J, Han L, Sanchez-Freire V, Abilez OJ, Navarrete EG, Hu S, Wang L, Lee A, *et al*: Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. Sci Transl Med 4: 130ra47, 2012.
14. Schafer E, Irizarry R, Negi S, McIntyre E, Small D, Figueroa ME, Melnick A and Brown P: Promoter hypermethylation in MLL-r infant acute lymphoblastic leukemia: Biology and therapeutic targeting. Blood 115: 4798-4809, 2010.
15. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur MH, Diebel ME, Monville F, Dutcher J, *et al*: Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. Cancer Res 69: 1302-1313, 2009.
16. Smyth GK: Limma: Linear models for microarray data. In: Bioinformatics and computational biology solutions using R and Bioconductor. Springer, pp397-420, 2005.
17. Ferreira JA and Zwinderman AH: On the Benjamini-Hochberg method. Ann Stat 34: 1827-1849, 2006.
18. Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA: DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4: P3, 2003.
19. Greicius MD, Krasnow B, Reiss AL and Menon V: Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100: 253-258, 2003.
20. Chegeni S, Khaki Z, Shirani D, Vajhi A, Taheri M, Tamrchi Y and Rostami A: Investigation of MMP-2 and MMP-9 activities in canine sera with dilated cardiomyopathy. Iran J Vet Res 16: 182-187, 2015.
21. Stamenkovic I: Extracellular matrix remodelling: The role of matrix metalloproteinases. J Pathol 200: 448-464, 2003.
22. Cheng Y and Zhang C: MicroRNA-21 in cardiovascular disease. J Cardiovasc Transl Res 3: 251-255, 2010.
23. Hayashidani S, Tsutsui H, Ikeuchi M, Shiomi T, Matsusaka H, Kubota T, Imanaka-Yoshida K, Itoh T and Takeshita A: Targeted deletion of MMP-2 attenuates early LV rupture and late remodeling after experimental myocardial infarction. Am J Physiol Heart Circ Physiol 285: H1229-H1235, 2003.
24. Ali S, Saik JE, Gould DJ, Dickinson ME and West JL: Immobilization of cell-adhesive laminin peptides in degradable PEGDA hydrogels influences endothelial cell tubulogenesis. Biores Open Access 2: 241-249, 2013.
25. Lebel R and Lepage M: A comprehensive review on controls in molecular imaging: Lessons from MMP-2 imaging. Contrast Media Mol Imaging 9: 187-210, 2014.
26. Herbert SP and Stainier DY: Molecular control of endothelial cell behaviour during blood vessel morphogenesis. Nat Rev Mol Cell Biol 12: 551-564, 2011.
27. Punetha J and Hoffman EP: Short read (next-generation) sequencing: A tutorial with cardiomyopathy diagnostics as an exemplar. Circ Cardiovasc Genet 6: 427-434, 2013.
28. Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, Sajadieh A, Haunsø S, Svendsen JH and Olesen MS: New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. Eur J Hum Genet 21: 918-928, 2013.
29. Rajan S, Pena JR, Jegga AG, Aronow BJ, Wolska BM and Wieczorek DF: Microarray analysis of active cardiac remodeling genes in a familial hypertrophic cardiomyopathy mouse model rescued by a phospholamban knockout. Physiol Genomics 45: 764-773, 2013.
30. Felix SB and Staudt A: Non-specific immunoadsorption in patients with dilated cardiomyopathy: Mechanisms and clinical effects. Int J Cardiol 112: 30-33, 2006.
31. Tang YL, Wang YJ, Chen LJ, Pan YH, Zhang L and Weintraub NL: Cardiac-derived stem cell-based therapy for heart failure: Progress and clinical applications. Exp Biol Med (Maywood) 238: 294-300, 2013.
32. Huang Y, De Reyniès A, De Leval L, Ghazi B, Martin-Garcia N, Travert M, Bosq J, Brière J, Petit B, Thomas E, *et al*: Gene expression profiling identifies emerging oncogenic pathways operating in extranodal NK/T-cell lymphoma, nasal type. Blood 115: 1226-1237, 2010.
33. Kim TH, Shin SW, Park JS and Park CS: Genome wide identification and expression profile in epithelial cells exposed to TiO_2 particles. Environ Toxicol 30: 293-300, 2015.
34. Diamanti D, Mori E, Incarnato D, Malusa F, Fondelli C, Magnoni L and Pollio G: Whole gene expression profile in blood reveals multiple pathways deregulation in R6/2 mouse model. Biomark Res 1: 28, 2013.
35. Lee HJ, Jang M, Kim H, Kwak W, Park W, Hwang JY, Lee CK, Jang GW, Park MN, Kim HC, *et al*: Comparative transcriptome analysis of adipose tissues reveals that ECM-receptor interaction is involved in the depot-specific adipogenesis in cattle. PLoS One 8: e66267, 2013.
36. Zhang RP, Liu HH, Liu JY, Hu JW, Yan XP, Wang DM, Li L and Wang JW: Transcriptional profiling identifies location-specific and breed-specific differentially expressed genes in embryonic myogenesis in anas platyrhynchos. PLoS One 10: e0143378, 2015.
37. Stark K: Candidate gene approach on DCM causing or susceptibility genes, 2010.
38. Esteban-Martínez RL, Pérez-Razo JC, Vargas-Alarcón G, Martínez-Rodríguez N, Cano-Martínez LJ, López-Hernández LB, Rojano-Mejía D, Canto P and Coral-Vazquez RM: Polymorphisms of APLN-APLNR system are associated with essential hypertension in Mexican-Mestizo individuals. Exp Mol Pathol 101: 105-109, 2016.
39. Kotanidou EP, Kalinderi K, Kyrgios I, Efraimidou S, Fidani L, Papadopoulou-Alataki E, Eboriadou-Petikopoulou M and Galli-Tsinopoulou A: Apelin and G212A apelin receptor gene polymorphism in obese and diabese youth. Pediatr Obes 10: 213-219, 2015.