Bioinformatics Analysis for Identifying Micro-RNAs, Long Noncoding RNAs, Transcription Factors, Immune Genes Regulatory Networks, and Potential Therapeutic Agents for Diabetic Cardiomyopathy Using an Integrated Bioinformatics Analysis

Jiefang Zhou
Shaoxing Hospital of Traditional Chinese Medicine
Xiaowei Ji
Ruian People's Hospital
Xiuwei Shen
Ruian People's Hospital
Kefeng Yan
Shaoxing Hospital of Traditional Chinese Medicine
Peng Huang
Ruian People's Hospital
Chunyan Huang (skies165@163.com)
Ruian People's Hospital

Research Article

Keywords: Diabetic cardiomyopathy, bioinformatics, differentially expressed genes, protein–protein interaction network

Posted Date: December 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1118638/v1

Abstract

Objectives

We identified functional genes and studied the underlying molecular mechanisms of diabetic cardiomyopathy (DCM) using bioinformatics tools.

Methods

Original gene expression profiles were obtained from the GSE21610 and GSE112556 datasets. We used GEO2R to screen the differentially expressed genes (DEGs). Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses were performed on DEGs. Protein–protein interaction (PPI) networks of DEGs were constructed using STRING and hub genes of signaling pathways were identified using Cytoscape. Aberrant hub gene expression was verified using The Cancer Genome Atlas dataset. Connectivity Map was used to predict the drugs that could treat DCM.

Results

The DEGs in DCM were mainly enriched in the nuclei and cytoplasm and involved in DCM- and chemokine-related signaling pathways. In the PPI network, 32 nodes were chosen as hub nodes and an RNA interaction network was constructed with 517 interactions. The expression of key genes (JPIK3R1, CCR9, XIST, WDFY3.AS2, hsa-miR-144-5p, and hsa-miR-146b-5p) was significantly different between DCM and normal tissues. Danazol, ikarugamycin, and semustine were identified as therapeutic agents against DCM using CMAP.

Conclusion

The identified hub genes could be associated with DCM pathogenesis and the above drugs could be used for treating DCM.

Introduction

Diabetic cardiomyopathy (DCM) is a specific cardiac condition in patients with diabetes, and is caused by hyperglycemia and insulin resistance. DCM development is independent of coronary artery disease, hypertension and other usual cardiac risk factors. It is clinically characterized by left ventricular hypertrophy, diastolic function damage, cardiomyocyte hypertrophy, myocardial fibrosis, and cardiomyocyte apoptosis, eventually leading to heart failure [1]. The pathophysiological mechanism of DCM could be related to molecular conditions such as metabolic disorders, calcium imbalance, mitochondrial damage, oxidative stress, increased inflammation, vascular dysfunction, cardiac

autonomic neuropathy, micro-RNAs (miRNAs) and long noncoding RNAs (lncRNAs) dysregulation [2, 3]. Due to the lack of clear diagnostic criteria, the specific pathogenic mechanisms of DCM remain unclear and specific histological characteristics, clinical manifestations and biomarkers for the definitive diagnosis of DCM still need more investigation [4–6]. Therefore, screening the potential molecules and pathways implicated with DCM and further studying the molecular mechanisms of DCM is important for preventing it and developing effective treatments.

Bioinformatic analysis and gene expression profiling analysis showed endogenous noncoding RNAs, miRNAs are widely present in eukaryotes and play an important role in the development of cardiovascular diseases [7]. miRNAs participate in cell proliferation and differentiation, metastasis, apoptosis, immune response, and other biological processes. They also play an important role in the development of DCM; more than 300 miRNAs have been confirmed to be expressed in the cardiomyocytes. Abnormal expression of miRNAs has been confirmed to be closely related to lncRNAs and transcription factors (TFs). For instance, lncRNAs can regulate miRNA expression by binding to target miRNAs and participating in the regulation of mRNA expression [8, 9]. TFs play an important role in gene transcription and post-transcriptional regulation and participate in the control of miRNA signaling pathways [10, 11]. However, further research is required to determine the regulatory mechanisms of miRNAs, lncRNAs, TFs, and mRNAs in DCM. Microarray analysis can rapidly identify all genes expressed at a given time; thus, future research can benefit from the integration and analysis of microarray data [12, 13]. In this study, we aimed to conduct an in-depth investigation of the miRNAs, lncRNAs, target genes, and pathways involved in the development of DCM.

Materials And Methods

Obtaining raw data

The datasets used in the present study were downloaded from the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) [14]. The GSE112556 dataset includes microRNA expression profiles. The original gene expression profiles were obtained from the GSE21610 and GSE17800 datasets. The GSE21610 dataset includes the gene expression profile of DCM tissues from 42 patients with DCM and normal tissues from 8 patients (Table 1). GPL570 Affymetrix HumanGenome U133 Plus 2.0 Array was used to analyze these data. Theplatform used for these data is also specified in Table 1.

Table 1
Gene Expression Omnibus (GEO) dataset

GEO accession	Platforms	Normal samples	DCM	Organism	Experiment type
GSE21610	GPL570	8	42	*Homo sapiens*	Expression profiling by array
GSE17800	GPL570	8	40	*Homo sapiens*	Expression profiling by array
GSE112556	GPL18402	3	3	*Homo sapiens*	Noncoding RNA profiling by array

Identification of differentially expressed genes

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) is an interactive web tool on the basis of R language [15] that can be used to compare two or more groups of samples for identifying differential expression in a GEO series. In the present study, GEO2R was used to filter differentially expressed mRNAs, and miRNAs between normal and DCM samples separately in each of the datasets. The false discovery rate (FDR) is a method of concept ualizing the type I error rate in a null hypothesis test when multiple comparisons are made. GEO2R automatically calculates FDR. Multiple t-tests were used to simultaneously detect genes that were FDR-corrected and statistically significant. A fold change of >2 and a *P* value of <0.05 were set as the cut-off criteria. Then filter probes were set without corresponding gene symbols.

Prediction of mRNA–miRNA–lncRNA interactions

We used miRWalk 3.0 (http://mirwalk.umm.uni-heidelberg.de/) to predict the interaction between differentially expressed miRNAs and mRNAs, which integrated the prediction results of TargetScan and miRDB [16, 17]. A score ≥0.95 was considered as the cut-off criterion for predictive analysis in miRWalk. We selectedonlythe target mRNA identified in all the above databases for further analysis. DIANA-LncBase v2.0 [18] was used to predict the interaction between miRNA and lncRNA, and a score of ≥0.4 was considered as the cut-off criterion for predictive analysis in the LncBase experimental module. After determining the number of predicted DEGs, the miRNAs, lncRNAs, and mRNAs were selected for further analysis. The Cytoscape software (version 3.40) was used to visualize the supervision network.

Gene function analysis

The DAVID database (http://david.abcc.ncifcrf.gov/) was used to perform gene ontology (GO)and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of significantDEGs and target genes of miRNAs with differential expression. The species was limited to *Homo sapiens*, and the “adjusted *P* value (from the Benjamini–Hochberg method), 0.05” was considered statistically significant [19]. The GO term includedthe following three criteria: molecular function (MF), cell composition (CC), and biological process (BP).

Protein–protein interaction (PPI) network construction and analysis

The STRING online database and PPI pairs with a combined score of ≥0.4 were used to construct a PPI network. TFs were annotated with TF checkpoints [20]. Then, the Cytoscape software (version 3.4.0) was used to visualize the regulatory relationship among genes and analyze the topological properties of the network, including the network degree distribution using the CentiScaPe application [21]. In addition, according to previous studies, genes with a degreeof>5 were defined as pivotal genes in the regulatory network [22]. Proteins with high degrees tend to be key factors.

TCGA dataset analysis

TCGA is a platform for researchers to download and evaluate free public datasets (https://cancergenome.nih.gov/) [23]. In this study, we used the online tool UALCAN (http://ualcan.path.uab.edu/) [24] to verify the expression of the hub genes in the TCGA dataset in order to improve the reliability of our analysis. UALCAN uses TCGA Level 3 RNA-seq and clinical data. As recommended by Li and Dewey [25], the Practical Extraction and Reporting Languageprogramwas used to multiply “scaledestimate” by 10^6 to obtain the transcripts per million (TPM) expression value. Since TPM is considered more comparable than reads per kilobase transcripts per million mapped reads (FPKM) and reads per kilobase transcript reads per million mapped reads (RPKM) [26], it was used as a measurement of gene expression in our study. Box and whisker plots were generated to represent the hub gene expression between normal samples and DCM samples in pipeline for RNA sequencing data analysis.

CMAP analysis

We screened the effective drugs used for treating DCM by using the CMAP database. The DEG file was uploaded onto the CMAP website. The screening criteria were set to a mean valueof ≤0.4 and a *P* value of<0.05.

Results

Identification of differentially expressed miRNAs, lncRNAs, TFs, and immune gene mRNAs

Analysis of theGSE112556 dataset led to the identification of 13 miRNAs in DCM samples compared with normal tissues, named DmiRNA, including 4 upregulated miRNAs and 9 downregulated miRNAs. On the other hand, analysis of the GSE21610 dataset led to the identification of 629 immune gene mRNAs (named DimGene, including 331 upregulated and 298 downregulated mRNAs), 260 TF gene mRNAs (named DTFGene, including 154 upregulated and 106 downregulated mRNAs), and 1044 lncRNAs (named DlncRNA, including 759 upregulated and 285 downregulated lncRNAs). We selected all

significantly upregulated and downregulated miRNAs (Fig. 1), mRNAs (Fig. 2), and lncRNAs (Fig. 3) to plot their expression patterns on heat-maps and volcano plots.

Function enrichment analysis of DEGs

GO enrichment analyses for DEGs were performed. The top 10 most significant GO terms in each group are presented in Fig. 4 and detailed in Table 2. Considering the BP criterion, the DEGs were mainly enriched in BPs related to the regulation of biological and cellular processes. Considering the MF criterion, the DEGs were mainly enriched in nucleic acid binding, transcription regulator activity, and signaling receptor activity. Considering the CC criterion, the DEGs were mainly enriched in cell, membrane-bounded organelle, and intracellular membrane-bounded organelle. The top 20 most significantKEGG pathway terms are shown in Fig. 4D. The genes were mainly enriched in the cytokine–cytokine receptor interaction, PI3K-Akt signaling pathway, and chemokine signaling pathway.

Table 2
Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway

KEGG_ID	Pathway_Name	Observed gene count	Background gene count	*P*value	Richnessfactor
hsa04060	Cytokine–cytokine receptor interaction	30	263	1.8E-14	0.114068441
hsa04062	Chemokine signaling pathway	16	181	0.00000474	0.08839779
hsa04630	Jak-STAT signaling pathway	15	160	0.00000474	0.09375
hsa04350	TGF-beta signaling pathway	11	83	0.00000599	0.13253012
hsa04640	Hematopoietic cell lineage	11	94	0.0000143	0.117021277
hsa04390	Hippo signaling pathway	13	152	0.0000313	0.085526316
hsa04657	IL-17 signaling pathway	10	92	0.0000616	0.108695652
hsa04933	AGE-RAGE signaling pathway in diabetic complications	10	98	0.0000942	0.102040816
hsa04668	TNF signaling pathway	10	108	0.00018	0.092592593
hsa04917	Prolactin signaling pathway	8	69	0.00029	0.115942029
hsa05134	Legionellosis	7	54	0.00047	0.12962963
hsa04620	Toll-like receptor signaling pathway	9	102	0.00056	0.088235294
hsa05418	Fluid shear stress and atherosclerosis	10	133	0.00071	0.07518797
hsa04621	NOD-like receptor signaling pathway	11	166	0.00085	0.06626506
hsa04650	Natural killer cell mediated cytotoxicity	9	124	0.0017	0.072580645
hsa01521	EGFR tyrosine kinase inhibitor resistance	7	78	0.0024	0.08974359
hsa04930	Type II diabetes mellitus	5	46	0.0065	0.108695652

KEGG_ID	Pathway_Name	Observed gene count	Background gene count	*P*value	Richnessfactor
hsa04660	T cell receptor signaling pathway	7	99	0.0069	0.070707071
hsa04151	PI3K-Akt signaling pathway	14	348	0.007	0.040229885
hsa04510	Focal adhesion	10	197	0.007	0.050761421

PPI analysis

To further explore the most significant clusters of DEGs,we performed a PPI network analysis of DimGene and DTFGene. A confidence score of ≥0.9 was set as the threshold and protein nodes that did not interact with other proteins were removed. A total of 162 of 889 DEGs were mapped into the PPI network complex and the PPI network data file of STRING was imported into Cytoscape. In this network, with a degree of >5, 32 nodes were chosen as hub nodes, including 3 TFs and 22 immune genes mRNAs. The results are presented in Fig. 5A **and B**. In the PPI networks, we identified the following top 5 DEGs with the highest node degrees: phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1; degree, 24), pro-melanin concentrating hormone (PMCH; degree, 19), proenkephalin (degree, 19), pro-platelet basic protein (PPBP; degree,16), and epidermal growth factor receptor (EGFR; degree, 16). The most significant hub genes whose protein levels were upregulated were PIK3R1, JAK2, EGFR, IL7R, CD2, CXCR4, KLRD1, IGLL5, DAPP1, HTR2B, GNRHR, and CYSLTR1, and the most significant hub genes whose protein levels were downregulated were ARTN, IL5RA, NGFR, LRRFIP1, ESR2, LTB4R, GCG, PMCH, GHSR, PTPN1, BCAR1, CDH1, and IL1RAP (Fig. 5C). These findings suggest that these nodes may play important roles in the development of DCM.

mRNA–miRNA–lncRNA network analysis

The miRNA–mRNA regulatory network was constructed, and interaction analysis indicated that 13 differentially expressed miRNAs targeted 26 mRNAs. In detail, the miRNAs downregulated 10 mRNAs but upregulated 16 mRNAs. We also constructed the miRNA–lncRNA regulatory network. Interaction analysis indicated that differentially expressed miRNAs targeted 13 lncRNAs, of which 6 were upregulated and 7 were downregulated. Based on the ceRNA theory, we used the sharedmiRNA as a junction and analyzed the lncRNA–miRNA–mRNA ceRNA regulatory relationship (Table 3). In the miRNA–mRNA and miRNA–lncRNA interaction pairs, the downregulated miRNAs accompanied the upregulated lncRNAs and mRNAs, and upregulated miRNAs accompanied the downregulated lncRNAs and mRNAs. Finally, the lncRNA–miRNA–TF–immune gene pathway ceRNA network was constructed using 517 interactions, including 10 miRNAs, 2 lncRNAs, and 161 mRNAs (123 immune genes and 38 TFs) (Fig. 6A). The lncRNA–miRNA–Hub gene pathway ceRNA network of key immune genes was constructed through CytoHubba and MCODE, and two important pathways, PIK3R1 and CCR9, were obtained as shown in Fig. 6B **and C**.

Table 3
lncRNA–miRNA–mRNA ceRNA regulatory relationship

lncRNA	Expression	Mature miRNA	Expression	Immune gene	Expression
WDFY3-AS2	Up	hsa-miR-144-3p	Down	SMARCE1	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	SMARCE1	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	KMT2A	Up
WDFY3-AS2	Up	hsa-miR-144-5p	Down	KLRD1	Up
WDFY3-AS2	Up	hsa-miR-144-3p	Down	LIFR	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	FZD1	Up
WDFY3-AS2	Up	hsa-miR-144-3p	Down	FST	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	FST	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	TGFB2	Up
WDFY3-AS2	Up	hsa-miR-144-3p	Down	KLRC4	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	CYSLTR1	Up
WDFY3-AS2	Up	hsa-miR-144-3p	Down	EGFR	Up
WDFY3-AS2	Up	hsa-miR-144-5p	Down	PIK3R1	Up
WDFY3-AS2	Up	hsa-miR-9-3p	Down	PIK3R1	Up
XIST	Down	hsa-miR-21-3p	Up	BCAR1	Down
XIST	Down	hsa-miR-21-3p	Up	TRIM36	Down
XIST	Down	hsa-miR-146b-5p	Up	CCR9	Down
XIST	Down	hsa-miR-21-5p	Up	THBS1	Down
XIST	Down	hsa-miR-21-3p	Up	NOD2	Down
XIST	Down	hsa-miR-21-5p	Up	C7	Down

lncRNA: Long noncoding RNA; **miRNA**: micro RNA; **ceRNA**: competing endogenous RNA

TCGA dataset analysis

TCGA dataset analyses were performed to demonstrate that the aberrant expression of the hub genes, including PIK3R1, CCR9, WDFY3.AS2, XIST, hsa-miR-144-5p, andhsa-miR-146b-5p, was significantly different between DCM and normal tissue samples (Fig. 7).

Drug prediction analysis using CMAP

Using the CMAP database, we used 19 hub immune genes (KMT2A, PTPN1, TFAP2C, TLR7, SMARCE1, KLRD1, LIFR, FZD1, FST, TGFB2, KLRC4, CYSLTR1, PIK3R1, BCAR1, TRIM36, CCR9, THBS1, NOD2, and C7) to predict the potential therapeutic drugs for DCM. We also performed the enrichment analysis of PIK3R1 and indicated that PIK3R1 was significantly enriched in the JAK-STAT signaling pathway, T Toll-like receptor signaling pathway, TNF signaling pathway, fluid shear stress, and atherosclerosis AGE−RAGE signaling pathway (which are associated with diabetes-related complications), and the chemokine signaling pathway (Fig. 8A). A total of 5 drugs were identified: danazol, ikarugamycin, semustine, cefamandole, and molindone (Table 4). 3D conformers of the drug molecules obtained from PubChem [27] are shown in Fig. 8B-F.

Table 4
Identification of the potential therapeutic drugs for DCM by using the CMAP database

CMAP name	Mean	*n*	Enrichment	*P*	Specificity	Percent non-null
Danazol	0.512	4	0.86	0.00052	0	75
Cefamandole	−0.391	4	−0.792	0.00378	0.0083	50
Semustine	−0.271	4	−0.805	0.0028	0.0148	50
Ikarugamycin	−0.524	3	−0.89	0.00256	0.0152	66
Molindone	0.342	4	0.738	0.00923	0.025	50

CMAP: Connectivity Map; **DCM**: diabetic cardiomyopathy

Discussion

DCM is a serious complication associated with diabetesand increasing incidence rate; however, the molecular mechanism underlying the development of DCM is unclear. Therefore, studying its mechanism and determining molecular targets for diagnosis and treatment are essential. Some studies have found that mRNAs, miRNAs, lncRNAs, and TFs play important regulatory roles in the development of DCM [2, 3]. In this study, we firstly performed a comprehensive bioinformatics analysis to identify the miRNAs, lncRNAs, immune genes, and TFs in the interaction network and reported the key genes related to DCM.

Noncoding RNAs (ncRNAs) have been proved to be of great significance in human life, health, and disease diagnosis. According to their functions, ncRNAs are classified as miRNAs, lncRNAs, and circular RNAs (circRNAs) [28]. lncRNAs act as miRNA sponges that can regulate miRNA abundance and compete with mRNA for miRNA binding. miRNAs are considered to be potential tissue-specific biomarkers and play an important role in the pathogenesis of myocardial remodeling [29]. They also contribute to the development of DCM by regulating oxidative stress, inflammation, cardiomyocyte apoptosis, and mitochondrial function [2]. Previous study has reported that miRNA-195 inhibitors can reduce diabetic myocardial hypertrophy and improve cardiac function by reducing oxidative damage, inhibiting cell apoptosis, and promoting angiogenesis [30]. Moreover, miRNAs are involved in the regulation of multiple

biological functions in the mitochondria by binding to their target genes [31]. The miR-30 family, including miR-30a, miR-30b, miR-30c, and miR-30d, is highly expressed in cardiomyocytes [32] and inhibits the expression of P53 to prevent mitochondrial division and cell apoptosis. The mechanisms by which miRNAs regulate the expression of genes involved in myocardial diseases and their effect on the development of diabetes have not yet been clearlyunderstood.

In this study, we analyzed the DEGs identified from the GSE21610 and GSE112556 datasets, which consist of the gene expression profiles of patients with DCM and healthy individuals. The results indicated that 13 miRNAs, 629 immune genes, 260 TFs, and 1044 lncRNAs were differentially expressed in patients with DCM. The DEGs were mainly enriched in the nuclei and cytoplasm; involved in the regulation of BPs and nucleic acid binding; and participated in the regulation of the cytokine–cytokine receptor interaction and the PI3K-Akt signaling pathway. These pathways are related to the development of DCM and chemokine signaling pathways. Our enrichment analysis results help in further studying the role of DEGs in DCM. To further analyze the key genes related to DCM, we constructed a PPI network, wherein 32 nodes were chosen as the hub nodes that consisted of 3 TFs and 22 immuGenes mRNAs. The lncRNA–miRNA–TF–immune gene pathway ceRNA network was constructed with 517 interactions, including 10 miRNAs, 2 lncRNAs, and 161 mRNAs (123 immune genes and 38 TFs).

TCGA dataset analyses indicated that aberrant expression of the hubgenes, including WDFY3.AS2, XIST, hsa-miR-144-5p, and hsa-miR-146b-5p, was significantly different between DCM and normal tissuesamples. Previous studies have shown that MiR-144 can target Nrf2 directly and down-regulated miR-144 is found to regulate oxidative stress in diabetic cardiomyocytes [33]. MiR-146a plays a regulatory role with the NF-κB signaling pathway components, which is the key mediator of inflammation and hyperglycemia [34, 35]. This is consistent with our finding. The roles of different miRNAs and lncRNAs, especially WDFY3.AS2, XIST, hsa-miR-144-5p, and hsa-miR-146b-5p, in the context of DCM should be further studied, including type 1 and type 2 diabetes, the contribution of pathophysiological mechanisms including inflammation, apoptosis, hypertrophy and fibrosis, and oxidative stress to the development of DCM.

Our dataset analyses also indicated the aberrant expression of the hub genes PIK3R1 and CCR9 between DCM and normal tissue. The enrichment results of PIK3R1 indicated that it was significantly enriched in the JAK-STAT signaling pathway, T Toll-like receptor signaling pathway, TNF signaling pathway, fluid shear stress, and atherosclerosis AGE-RAGE signaling pathway (which cause diabetic complications), and the chemokine signaling pathway. The JAK/STAT signaling pathway is an intracellular signaling pathway closely related to cardiac hypertrophy, which plays a key role in cell growth, survival and differentiation, and regulation of gene expression. It has been shown that high glucose activates the JAK/STAT signaling of vascular endothelial cells in vitro through phosphorylation of JAK2 and subsequently STAT3, leading to the proliferation of endothelial cells. Hyperglycemia, hyperlipidemia, and insulin resistance are the main causes of chronic low-grade inflammation of the diabetic heart. Cardiac Toll-like receptors and inflammasome complexes, probably through NF-B activation and ROS

overproduction, may be key inducers for inflammation [36]. Therefore, what role of PIK3R1 plays in regulation of JAK/STAT signaling pathway, T Toll-like receptor signaling pathway and others represents an important strategy for the treatment of DCM and needs to be further explored.

We also performed CMAP analysis to examine the potential small-molecule drugs that can be potentially used for treating DCM. The three most significant drugs identified were danazol, ikarugamycin, and semustine. Danazol is a medicine used to treat endometriosis, fibrocystic breast disease, hereditary angioedema and other diseases. Ikarugamycin is a previously discovered antibiotic which has been shown to inhibit clathrin-mediated endocytosis and the uptake of oxidized low-density lipoprotein in macrophages [37]. Semustine is a drug used in chemotherapy [38]. Although there is no relevant literature regarding the effect of these compounds on DCM treatment, we speculate that these small-molecule drugs may have the potential for effectively treating DCM by interfering with the expression of the hub genes.

Conclusion

We constructed and analyzed mRNAs, miRNAs, lncRNAs, and TF interaction networks to identify the key genes related to the development of DCM. Our findings suggest that DEGs were mainly involved in the cytokine–cytokine receptor interaction, chemokine signaling pathway, and PI3K-Akt signaling pathway, and the hub genes JPIK3R1, CCR9, XIST, WDFY3.AS2, hsa-miR-144-5p, and hsa-miR-146b-5p may be the key factors associated with the pathogenesis of DCM. Danazol, ikarugamycin, and semustine could be potential targets for performing the follow-up study of the molecular mechanism underlying the development of DCM and the therapeutic drugs that can effectively treat this condition. However, further studies are required to confirm the functions of the DEGs we found and the effectiveness of these small-molecule drugs against DCM.

Declarations

Disclosure statement

The author did not declare any conflicts of interest with the current research.

Funding statement

This research was supported by the Basic Research Projects (Y20211015) organized by the Wenzhou Municipal Science and Technology Bureau.

Declaration of Competing Interest

The authors declare they have no conflict of interest.

Acknowledgements

We thank the GEO network for generously sharing amounts of data, which benefited this study.

AUTHOR CONTRIBUTIONS

CH, JZ, and XS conceived the ideas. CH and XS designed the study. JZ and XJ collected research data. CH, JZ, and XS performed data analysis. CH and JZ interpreted the data and wrote the main manuscript text. PH and KY supervised this work and provided financial support for the project. XS revised the manuscript. All authors read and approved the final manuscript.

References

1. Tan, Y., et al., *Mechanisms of diabetic cardiomyopathy and potential therapeutic strategies: preclinical and clinical evidence.* Nature Reviews Cardiology, 2020. **17**(9): p. 585-607.
2. Jakubik, D., et al., *MicroRNAs and long non-coding RNAs in the pathophysiological processes of diabetic cardiomyopathy: emerging biomarkers and potential therapeutics.* Cardiovascular Diabetology, 2021. **20**(1): p. 1-29.
3. Djordjevic, D.B., et al., *Diabetic Cardiomyopathy: Clinical and Metabolic Approach.* Current Vascular Pharmacology, 2021.
4. Paolillo, S., et al., *Diabetic cardiomyopathy: definition, diagnosis, and therapeutic implications.* Heart failure clinics, 2019. **15**(3): p. 341-347.
5. Lee, M.M.Y., et al., *Diabetic cardiomyopathy.* Heart, 2019. **105**(4): p. 337-345.
6. Jia, G., M.A. Hill, and J.R. Sowers, *Diabetic cardiomyopathy: an update of mechanisms contributing to this clinical entity.* Circulation research, 2018. **122**(4): p. 624-638.
7. Águila, S., et al., *MicroRNAs as New Regulators of Neutrophil Extracellular Trap Formation.* International Journal of Molecular Sciences, 2021. **22**(4): p. 2116.
8. Chen, L., et al., *Long non-coding RNA MALAT1 regulates ZEB1 expression by sponging miR-143-3p and promotes hepatocellular carcinoma progression.* Journal of Cellular Biochemistry, 2017. **118**(12): p. 4836-4843.
9. Guttman, M. and J.L. Rinn, *Modular regulatory principles of large non-coding RNAs.* Nature, 2012. **482**(7385): p. 339-346.
10. Afshar, A.S., J. Xu, and J. Goutsias, *Integrative identification of deregulated miRNA/TF-mediated gene regulatory loops and networks in prostate cancer.* PloS one, 2014. **9**(6): p. e100806.
11. Xu, H., et al., *Liver-enriched transcription factors regulate microRNA-122 that targets CUTL1 during liver development.* Hepatology, 2010. **52**(4): p. 1431-1442.
12. Guo, Y., et al., *Identification of key candidate genes and pathways in colorectal cancer by integrated bioinformatical analysis.* International journal of molecular sciences, 2017. **18**(4): p. 722.
13. Vogelstein, B., et al., *Cancer genome landscapes.* science, 2013. **339**(6127): p. 1546-1558.
14. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.* Nucleic acids research, 2002. **30**(1): p. 207-210.

15. Diboun, I., et al., *Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma.* BMC genomics, 2006. **7**(1): p. 1-14.
16. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* cell, 2005. **120**(1): p. 15-20.
17. Wong, N. and X. Wang, *miRDB: an online resource for microRNA target prediction and functional annotations.* Nucleic acids research, 2015. **43**(D1): p. D146-D152.
18. Paraskevopoulou, M.D., et al., *DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts.* Nucleic acids research, 2016. **44**(D1): p. D231-D238.
19. Ashburner, M., et al., *Gene ontology: tool for the unification of biology.* Nature genetics, 2000. **25**(1): p. 25-29.
20. Chawla, K., et al., *TFcheckpoint: a curated compendium of specific DNA-binding RNA polymerase II transcription factors.* Bioinformatics, 2013. **29**(19): p. 2519-2520.
21. Scardoni, G., M. Petterlini, and C. Laudanna, *Analyzing biological network parameters with CentiScaPe.* Bioinformatics, 2009. **25**(21): p. 2857-2859.
22. Han, J.-D.J., et al., *Evidence for dynamically organized modularity in the yeast protein–protein interaction network.* Nature, 2004. **430**(6995): p. 88-93.
23. Tomczak, K., P. Czerwińska, and M. Wiznerowicz, *The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge.* Contemporary oncology, 2015. **19**(1A): p. A68.
24. Chandrashekar, D.S., et al., *UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses.* Neoplasia, 2017. **19**(8): p. 649-658.
25. Li, B. and C.N. Dewey, *RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.* BMC bioinformatics, 2011. **12**(1): p. 1-16.
26. Li, B., et al., *RNA-Seq gene expression estimation with read mapping uncertainty.* Bioinformatics, 2010. **26**(4): p. 493-500.
27. Wang, Y., et al., *Pubchem bioassay: 2017 update.* Nucleic acids research, 2017. **45**(D1): p. D955-D963.
28. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions.* Nature Reviews Molecular Cell Biology, 2021. **22**(2): p. 96-118.
29. Mathur, P. and V. Rani, *MicroRNAs: A critical regulator and a promising therapeutic and diagnostic molecule for diabetic cardiomyopathy.* Current Gene Therapy, 2021.
30. Giordano, M., et al., *Circulating MiRNA-195-5p and-451a in Diabetic Patients with Transient and Acute Ischemic Stroke in the Emergency Department.* International Journal of Molecular Sciences, 2020. **21**(20): p. 7615.
31. Jusic, A. and Y. Devaux, *Mitochondrial noncoding RNA-regulatory network in cardiovascular disease.* Basic research in cardiology, 2020. **115**(3): p. 1-17.
32. Li, J., et al., *miR-30 regulates mitochondrial fission through targeting p53 and the dynamin-related protein-1 pathway.* PLoS genetics, 2010. **6**(1): p. e1000795.

33. Yu, M., et al., *Inhibiting microRNA-144 abates oxidative stress and reduces apoptosis in hearts of streptozotocin-induced diabetic mice.* Cardiovascular Pathology, 2015. **24**(6): p. 375-381.
34. Emadi, S.S., et al., *MicroRNA-146a expression and its intervention in NF-κB signaling pathway in diabetic rat aorta.* Endocrine regulations, 2014.
35. Wang, Q., et al., *Regulation of retinal inflammation by rhythmic expression of MiR-146a in diabetic retina.* Investigative ophthalmology & visual science, 2014. **55**(6): p. 3986-3994.
36. Fuentes-Antrás, J., et al., *Activation of toll-like receptors and inflammasome complexes in the diabetic cardiomyopathy-associated inflammation.* International journal of endocrinology, 2014. **2014**.
37. Elkin, S.R., et al., *Ikarugamycin: a natural product inhibitor of clathrin-mediated endocytosis.* Traffic, 2016. **17**(10): p. 1139-1149.
38. Group, G.T.S., *Radiation therapy and fluorouracil with or without semustine for the treatment of patients with surgical adjuvant adenocarcinoma of the rectum.* J Clin Oncol, 1992. **10**: p. 549-557.

Figures

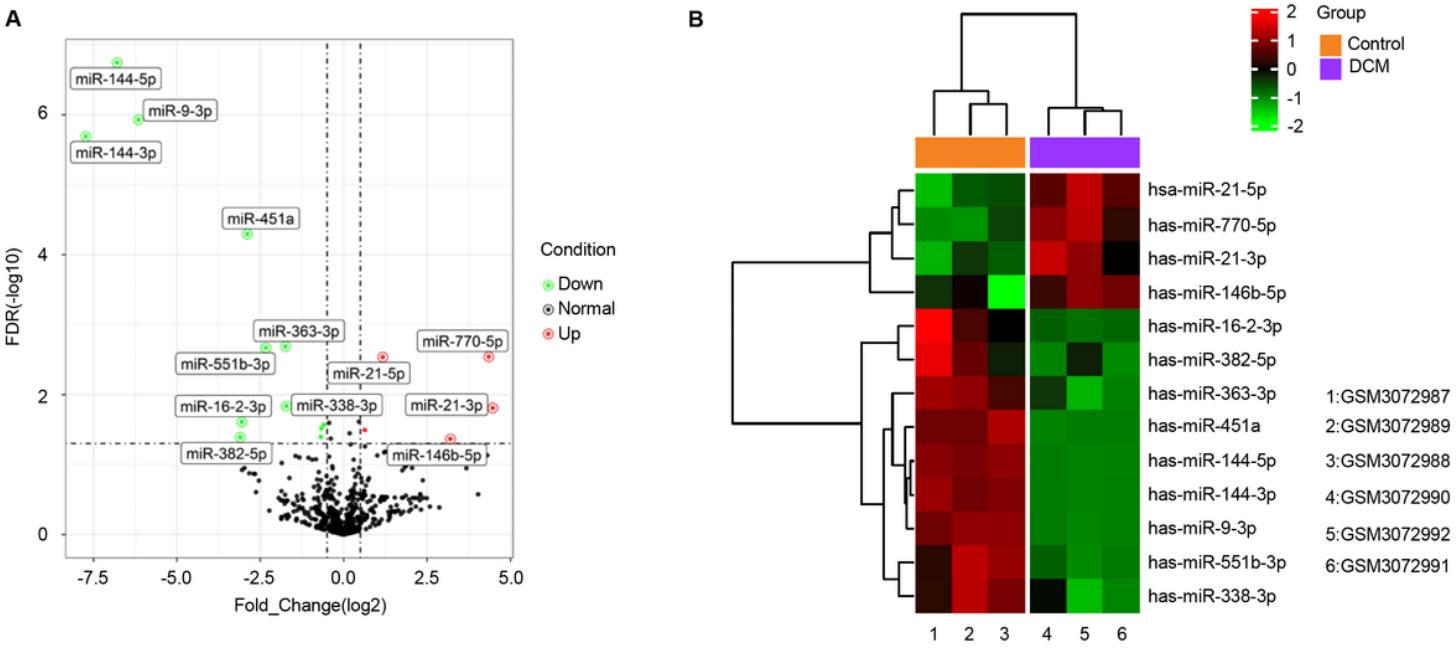

Figure 1

Differentially expressed miRNAs in DCM and normal samples. (A) Avolcano plot of differentially expressed miRNAs. (B) A heat-map of differentially expressed miRNAs. For Volcano plots, red and green points correspond to log2fold change (log2fold change>0.5 and log2fold change<-0.5) up/down, respectively, and indicate FDR <0.05. For the heatmap, each column represents a sample, and every row represents a miRNA. The expression value for each row was normalized by the z-score. Red indicates high relative expression and green indicates low relative expression.

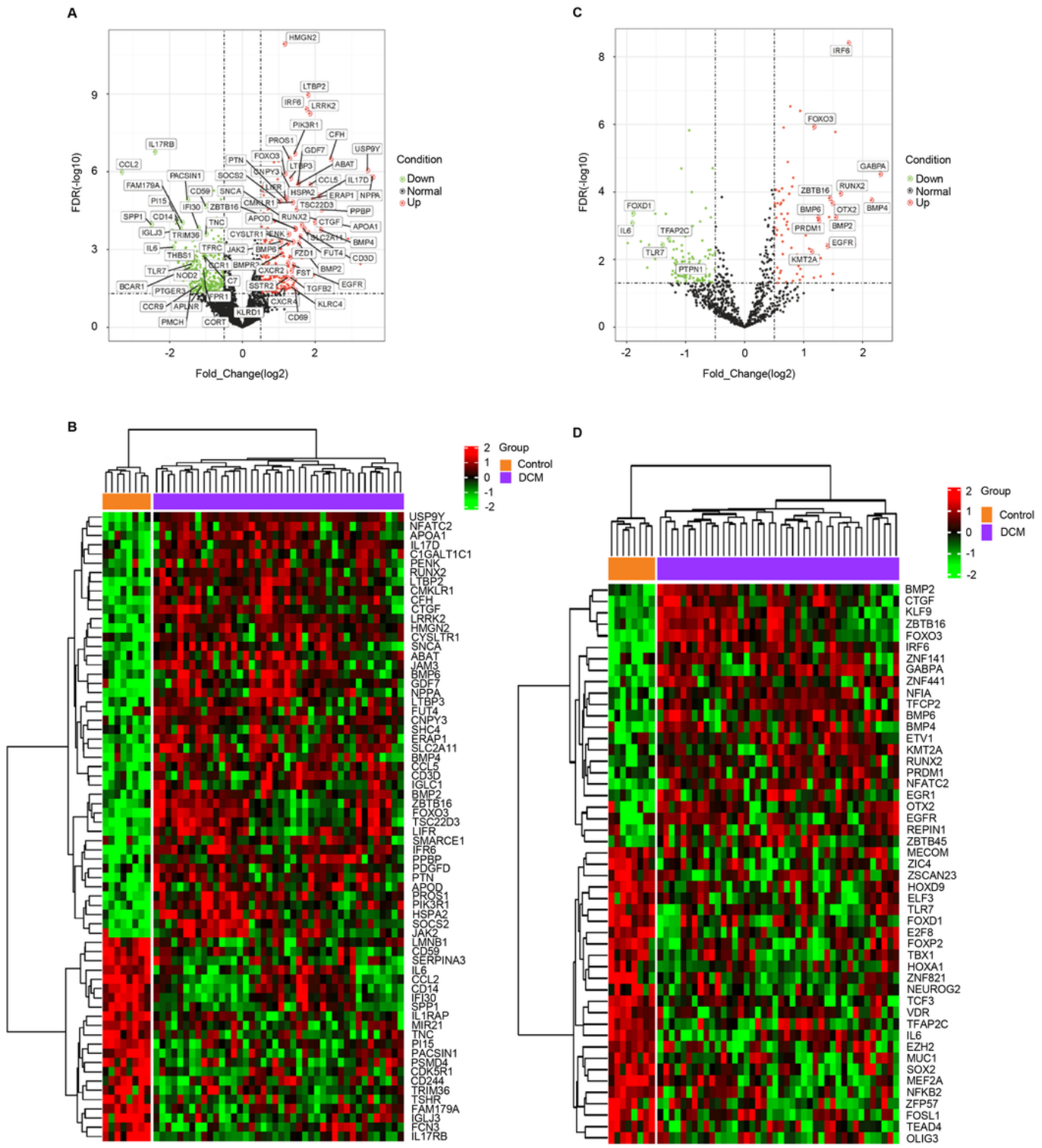

Figure 2

Differentially expressed mRNAs in DCM and normal samples. (A) A volcano plot of differentially expressed immune gene mRNAs. (B) A heat-map of differentially expressed immune gene mRNAs. Two datasets were selected by using GSE21610as the training dataset and GSE17800 as the validation dataset. (C) A volcano plot of differentially expressed TF mRNAs. (D) A heat-map of differentially expressed TFmRNAs. For Volcano plots, red and green points correspond to log2fold change (log2fold

change>0.5 and log2fold change<-0.5) up/down, respectively, and indicate FDR <0.05. For the heatmap, each column represents a sample, and every row represents an Immune gene or TF. The expression value for each row was normalized by the z-score. Red indicates high relative expression and green indicates low relative expression.

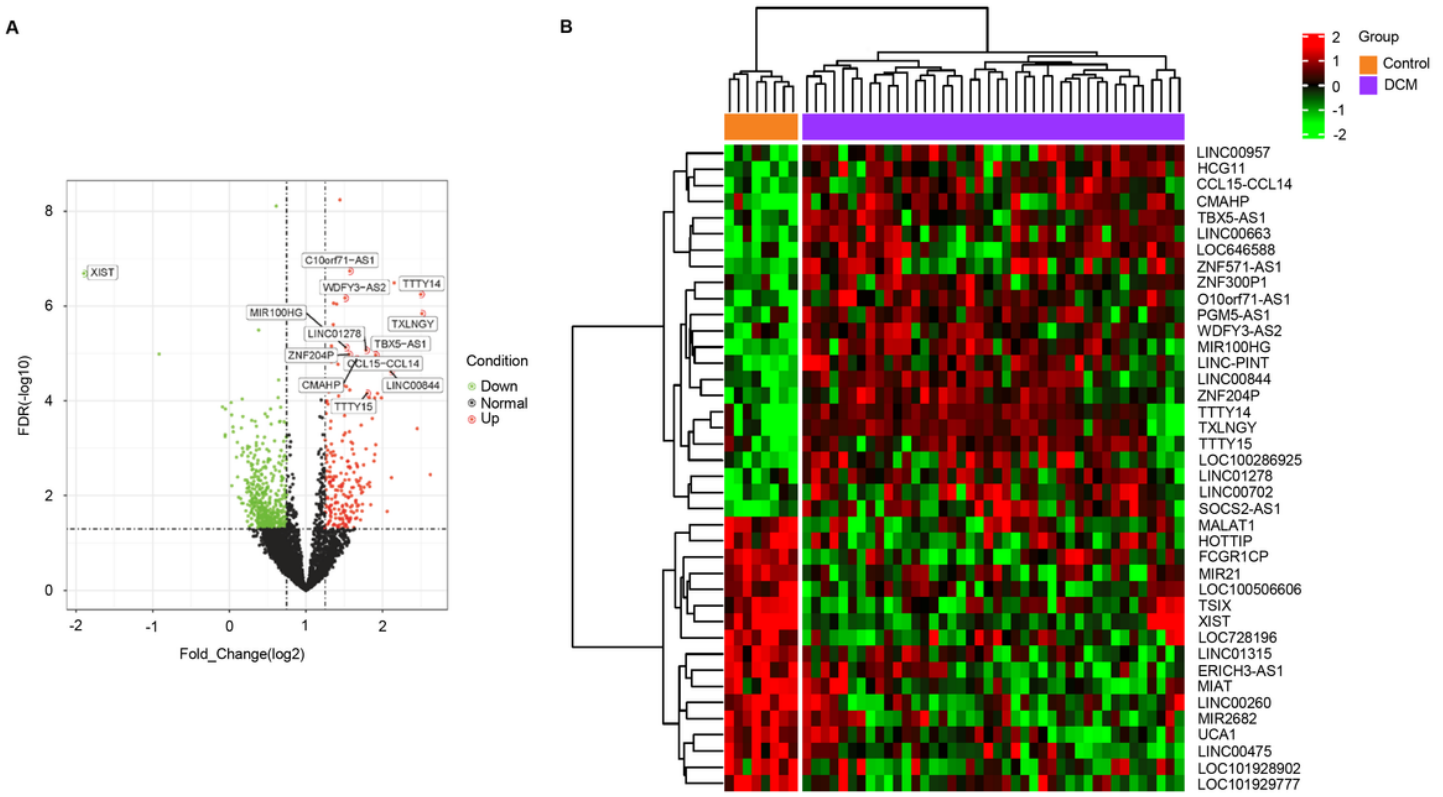

Figure 3

Differentially expressed lncRNAs in DCM and normal samples. (A) Avolcanoplot of differentially expressed lncRNAs. (B) Aheat-map of differentially expressed lncRNAs. For Volcano plots, red and green points correspond to log2fold change (log2fold change>0.5 and log2fold change<-0.5) up/down, respectively, and indicate FDR <0.05. For the heatmap, each column represents a sample, and every row represents an lncRNA. The expression value for each row was normalized by the z-score. Red indicates high relative expression and green indicates low relative expression.

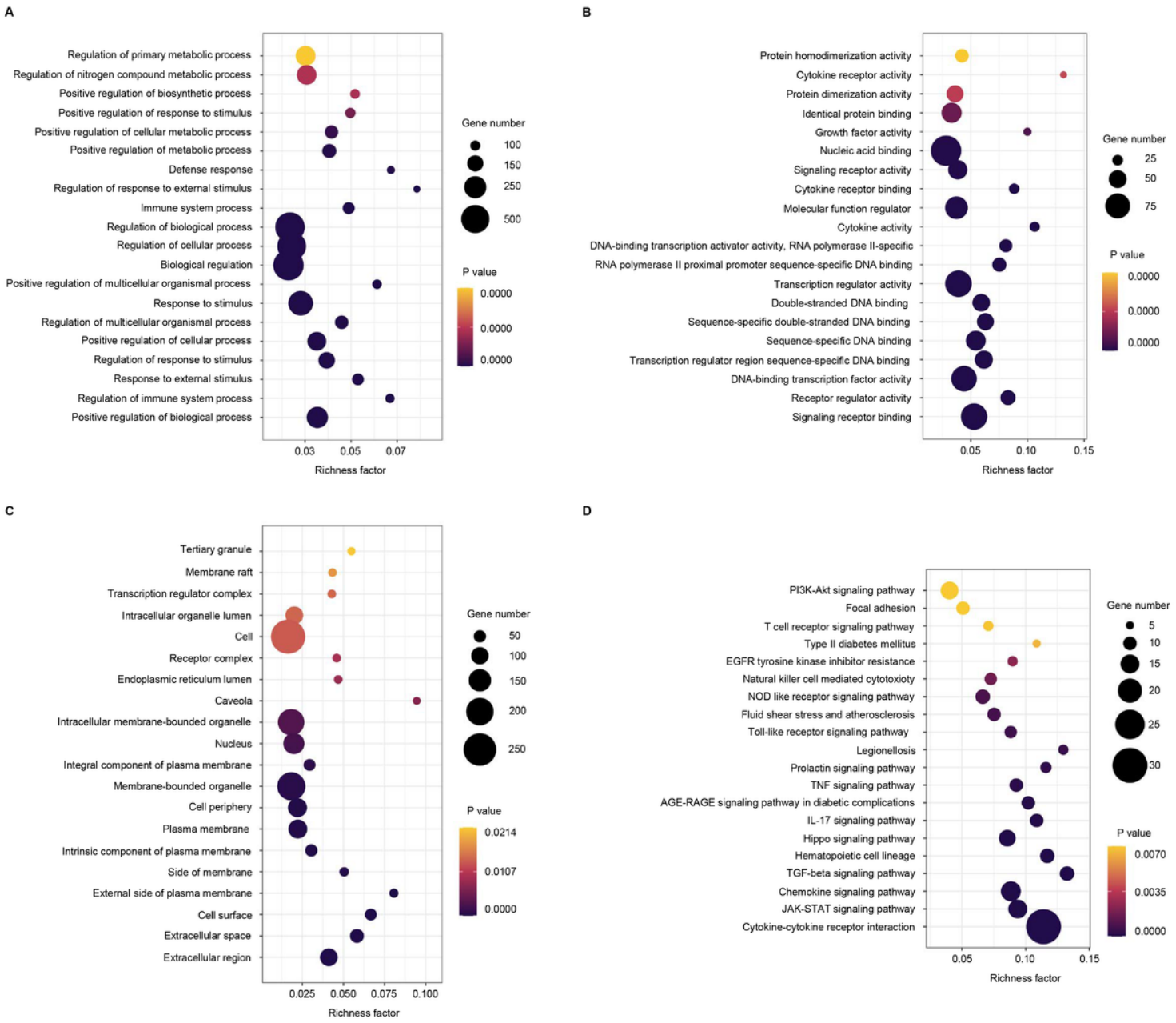

Figure 4

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genome (KEGG) enrichment analysis. (A) The top 20 most significant changes in the GO biological process. (B) The top 20 most significant changes in the GO molecular function. (C) The top 20 most significant changes in the GO cellular component. (D) The top 20 most significant KEGG pathway terms.

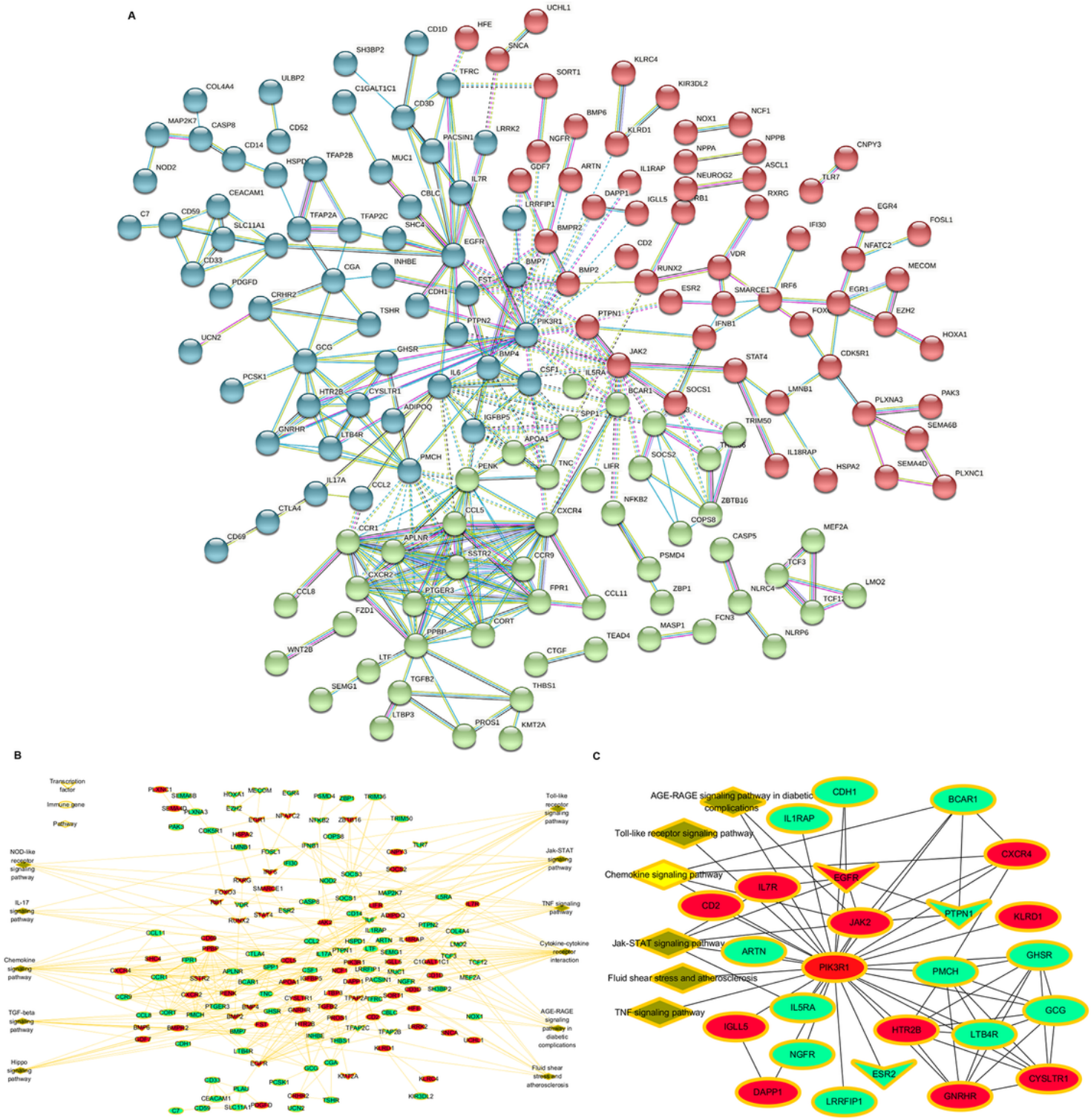

Figure 5

Protein–protein interaction (PPI) networks of the common differentially expressed genes (DEGs) and module analysis. (A) The resultant protein–protein interaction (PPI) network of DEGs identified from the samples of patients with DCM and healthy controls. The edge between the nodes represents the interaction between two connected proteins. (B) TF–immune gene pathway network. Transcription factors, immune genes, and pathways are indicated with arrows, circles, and oval shapes, respectively.

The orange line indicates the relationship of the protein interactions. The color red represents high expression and green represents low expression. (C) Most densely connected network components based on CytoHubba.

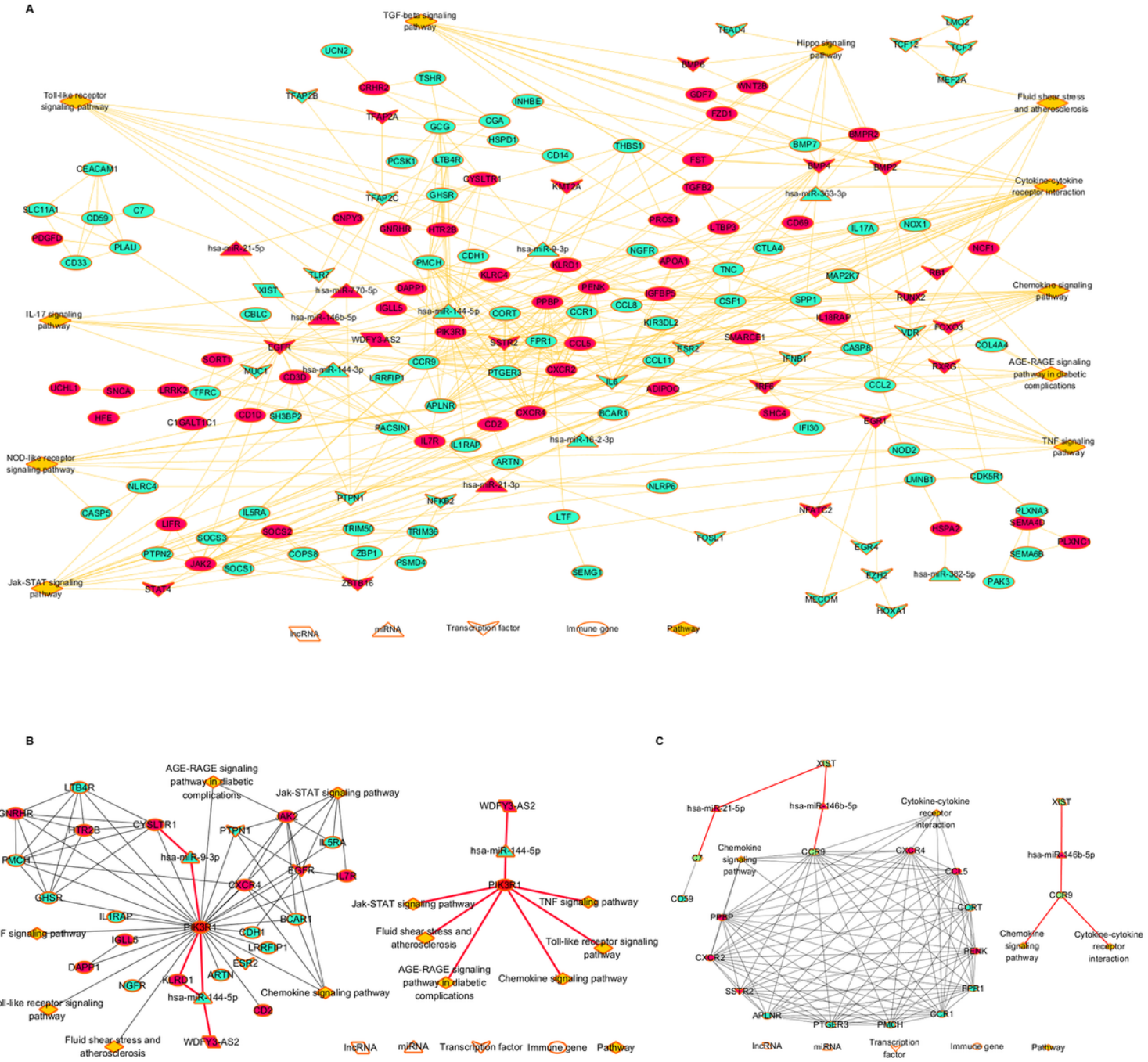

Figure 6

The lncRNA–miRNA–TF–immune gene pathway ceRNA network. Every node represents one gene and each edge represents the interaction between genes. (A) ThelncRNA–miRNA–immune gene pathway ceRNA network. (B) The hub lncRNA–miRNA–mRNA ceRNA pathway according to MCODE. (C) The hub lncRNA–miRNA–mRNA ceRNA pathway according to CytoHubba. Parallelograms, triangles, V-shapes,

ellipses, and orange rhombuses represent lncRNAs, miRNAs, TFs, immune genes, and pathways, respectively. Red and green colors represent upregulated and downregulated biomarkers, respectively.

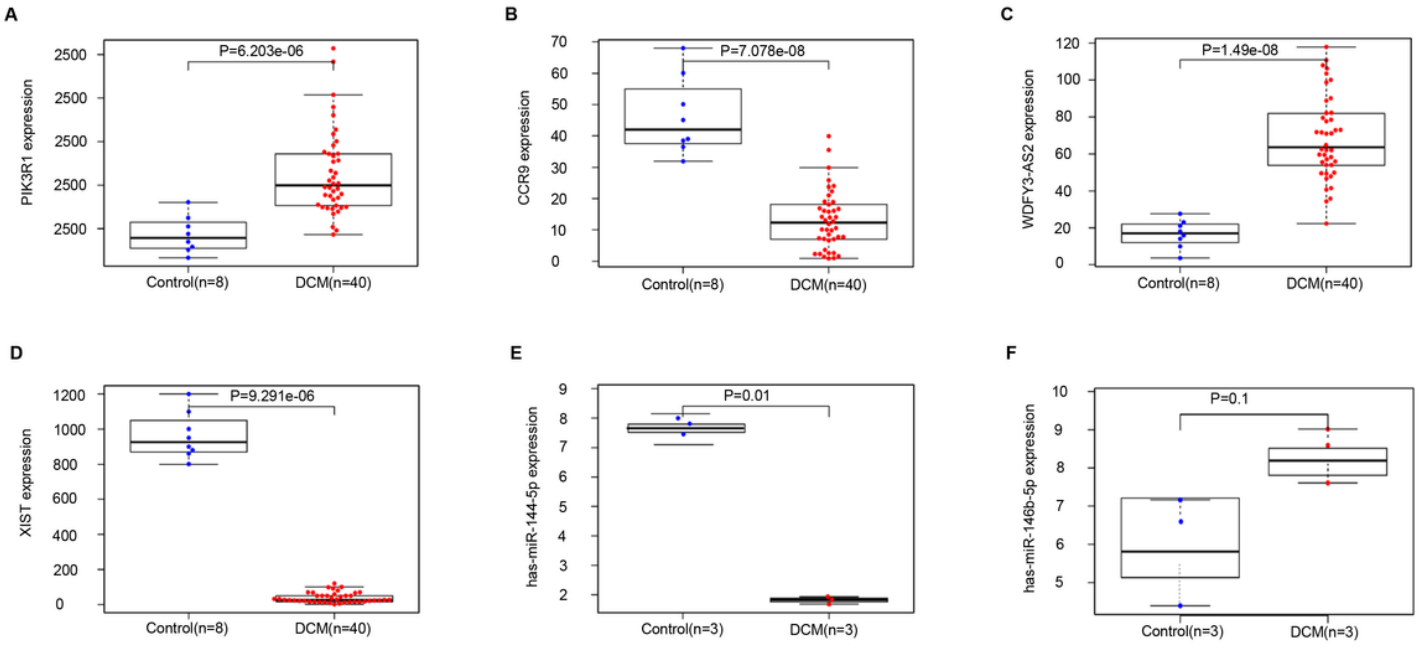

Figure 7

TCGA dataset analysis. (A) PIK3R1 expression in DCM. (B) CCR9 expression in DCM. (C) WDFY3.AS2 expression in DCM. (D) XIST expression in DCM. (E) hsa-miR-144-5p expression in DCM. (F) hsa-miR-146b-5p expression in DCM.

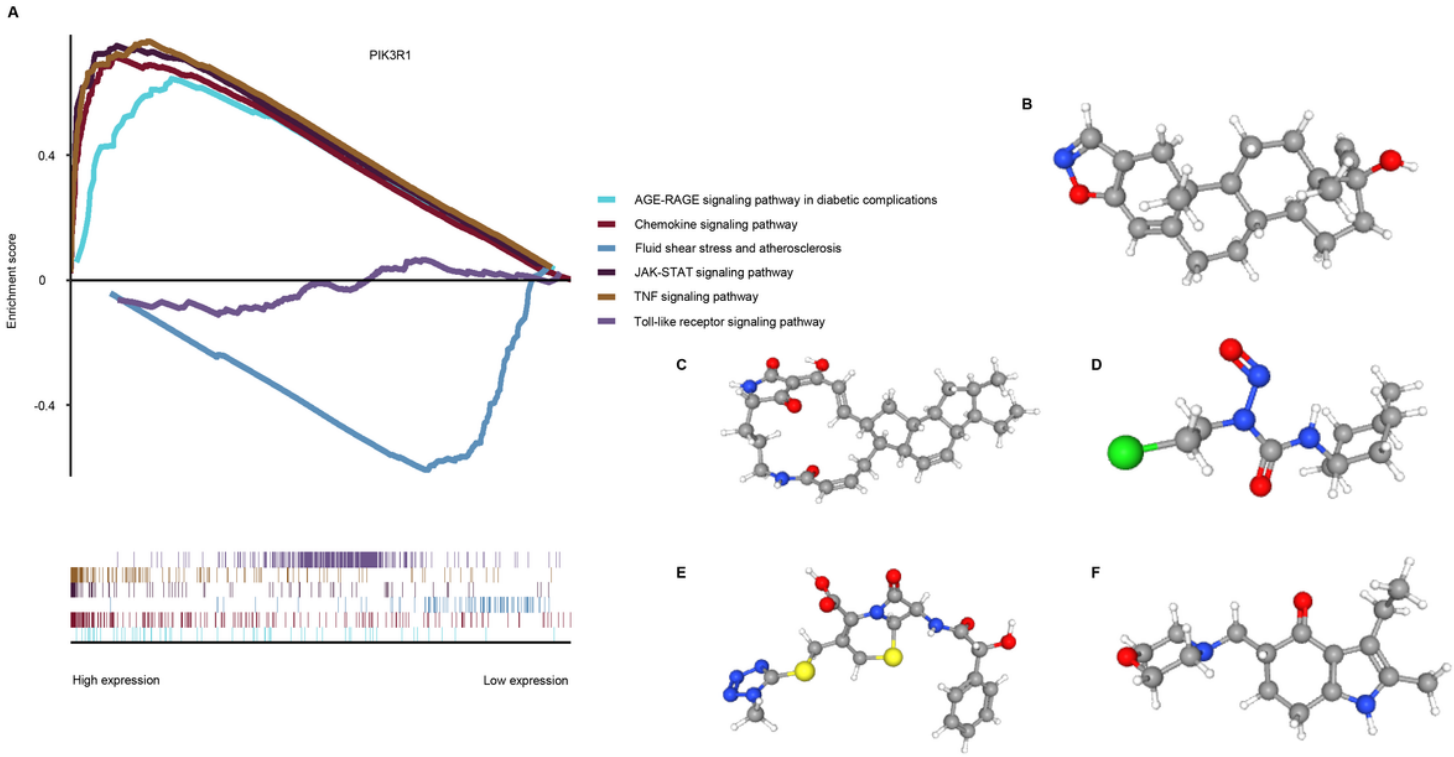

Figure 8

Predict potential therapeutic drugs for DCM. (A) Enrichment plots from Gene Set Enrichment Analyses for PIK3R1. (B-F) 3D conformers of the 5 drugs: danazol, ikarugamycin, semustine, cefamandole, and molindone.