Am J Transl Res 2024;16(2):637-653
www.ajtr.org /ISSN:1943-8141/AJTR0153448

Original Article

Transcriptomics data integration and analysis to uncover hallmark genes in hypertrophic cardiomyopathy

Peng Chen[1], Warda Yawar[2], Ayesha Rida Farooqui[3], Saqib Ali[4], Nida Lathiya[5], Zeeshan Ghous[6], Rizwana Sultan[7], Majid Alhomrani[8,9], Saleh A Alghamdi[8], Abdulraheem Ali Almalki[8], Ahmad A Alghamdi[8], Naif ALSuhaymi[10], Muhammad Razi Ul Islam Hashmi[11], Yasir Hameed[12]

[1]Department of Cardiovascular Medicine, Taiyuan Central Hospital, Taiyuan 030000, Shanxi, China; [2]Department of Emergency, PPHI, Sindh, Karachi 74800, Pakistan; [3]Department of Emergency, Abbasi Shaheed Hospital, Karachi 74800, Pakistan; [4]Department of Computer Science, University of Agriculture, Faisalabad 38040, Pakistan; [5]Department of Physiology, Jinnah Medical and Dental College, Sohail University, Karachi 74800, Pakistan; [6]Department of Cardiology, Punjab Institute of Cardiology, Lahore 54000, Pakistan; [7]Department of Pathology, Faculty of Veterinary and Animal Sciences, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan; [8]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Taif University, Taif 21944, Saudi Arabia; [9]Research Centre for Health Sciences, Taif University, Taif 21944, Saudi Arabia; [10]Department of Emergency Medical Services, Faculty of Health Sciences - AlQunfudah, Umm Al-Qura University, Mekkah, Saudi Arabia; [11]Cardiothoracic ICU, Bahria International Hospital Rawalpindi, Pakistan; [12]Department of Biotechnology, Institute of Biochemistry Biotechnology and Bioinformatics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan

Received September 15, 2023; Accepted January 24, 2024; Epub February 15, 2024; Published February 28, 2024

Abstract: Introduction: Hypertrophic cardiomyopathy (HCM) is a heterogeneous disease that mainly affects the myocardium. In the current study, we aim to explore HCM-related hub genes through the analysis of differentially expressed genes (DEGs) between HCM and normal sample groups. Methods: The GSE68316 and GSE36961 expression profiles were obtained from the Gene Expression Omnibus (GEO) database for the identification of DEGs, to explore hub genes, and to perform their expression analysis. Clinical HCM and control tissue samples were taken for expression and promoter methylation validation analysis via RNA-sequencing (RNA-seq) and targeted bisulfite sequencing (bisulfite-seq) analyses. Then, other different bioinformatics tools were employed to perform STRING, lncRNA-miRNA-mRNA regulatory networks, gene enrichment, and drug prediction analyses. Results: In total, the top 20 DEGs, including 10 up-regulated and 10 down-regulated, were obtained from GSE68316. Out of the 20 DEGs, we subsequently identified the 8 most important hub genes including 5 up-regulated genes (EPB42, UQCRH, CA1, PFDN5, and LSM5) and 3 down-regulated genes (RPS24, TNS1, and RPL26). Expression and promoter methylation dysregulation of these genes were further validated on clinical HCM samples paired with controls. Next, we further investigated hub genes' regulatory 6 miRNAs (has-mir-1-3p, has-mir-129-5p, has-mir-16-5p, has-mir-23b-3p, has-mir-27-3p, and has-mir-182-5p) and miRNAs regulatory 4 lncRNAs (NUTMB2-AS1, NEAT1, XIST, and GABPB1-AS1) in this study via the lncRNA-cricRNA-miRNA-mRNA regulatory network. Later on, gene enrichment analysis revealed that hub genes were enriched in various important pathways including Nitrogen metabolism, Ribosome, RNA degradation, Cardiac muscle contraction, and Coronavirus disease, etc. Finally, the drug prediction analysis highlighted different potential candidate drugs for altering the expression of hub genes in the treatment of HCM. Conclusion: In summary, the identification of key hub genes and their enrichment analysis in the current study may shed light on the mechanisms behind the occurrence and development of HCM.

Keywords: Hypertrophic cardiomyopathy, DEGs, hub gene, miRNA

Introduction

Hypertrophic cardiomyopathy (HCM), a complex cardiovascular genetic disorder, is reported in one out of every 500 individuals worldwide [1, 2]. In recent years, awareness of HCM has significantly improved in clinical settings [3, 4]. However, HCM diagnosis is still a complex process, just like other diseases including asthma, mitral regurgitation, and coronary artery dis-

https://doi.org/10.62347/AXOY3338

ease, resulting in a high mortality rate [5]. From a clinical perspective, the prevalent pathological characteristics of HCM include myocyte hypertrophy, disarray, and interstitial fibrosis [6]. It is important to highlight that approximately 25% of individuals diagnosed with HCM experience left ventricular outflow tract obstruction [7]. The clinical symptoms of HCM vary from patient to patient, due to the natural complex history of the disease [8]. However, if this disease was diagnosed precisely, the patients suffering from HCM could manage this disease more effectively to improve their survival duration.

Currently, echocardiography and some imaging methods such as cardiac resonance imaging are important tools for diagnosing HCM. However, the detection accuracies of these tools are still very much compromised [6]. Moreover, it was also observed that different physical exercise methods have also made great contributions to the management of HCM even without increasing its further risk [9]. The surgical treatment method is a preliminary choice for clinicians to treat HCM. However, due to the increased mortality rate during surgeries, clinicians are still hesitant to choose this method [10].

Recently, a lot of emphasis has been given to research exploring the underlying molecular mechanism of HCM [11, 12]. Genetic testing techniques have now become more accessible at the clinical level. The use of this technique in diagnosing HCM is increasing nowadays [13]. It is important here to understand that HCM has various genotypic and phenotypic variations. HCM was linked with approximately 1400 genetic mutations across more than 10 genes responsible for producing cardiac sarcolemmal protein [2]. Mostly, HCM developed due to a single heterozygous mutation [14, 15]. However, multiple mutations can also lead to a severe form of HCM [16]. According to current research findings, scientists are of the opinion that gene mutations play a significant role in HCM [12]. Various sacromeric genes, namely MYH7, MYBPC3, TPM1, TNNT2, and TNNI3, have been identified as crucial genetic alterations in 34% of individuals with HCM [17-19]. These mutations are typically observed in familial cases with a hereditary pattern. However, it should be noted that gene mutations have not been confirmed in approximately 70% of HCM patients, leaving the underlying genetic causes unclear.

It was also reported in the medical literature that by exploring differentially expressed genes (DEGs) among HCM and normal control samples, potential HCM-associated candidate genes could be figured out [20]. The up-regulation of KRT1 and down-regulation of CYP1A1 at the mRNA level were reported to play a significant role in the development and progression of this disease [21, 22]. Moreover, 3-methylcholanthrene and benzo(a)pyrene-induced cardiac hypertrophy was also associated with the overexpression of CYP1A1 [23].

In this paper, the Gene Expression Omnibus (GEO) datasets related to HCM were analyzed to identify the key hub genes having diagnostic and therapeutic importance for HCM patients. In summary, the results of our study would be helpful to improve the diagnosis and treatment methods for HCM patients.

Methods

Data collection

Normalized gene expression data of HCM patients (n = 7) and normal individuals (n = 5) were obtained from the GSE68316 dataset [24] by using the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Moreover, another GEO dataset GSE36961 consisting of 106 HCM and 39 normal samples was used to obtain the normalized gene expression data for further expression validation. HCM samples in these datasets were collected prior to any clinical intervention.

DEGs identification

For identifying DEGs among HCM and control samples in this study, the R-based limma (Version 3.40.6) package in Bioconductor was utilized [25]. Genes showing *p*-values less than 0.01 and fold changes (FCs) greater than 1.2 were regarded as DEGs. A total of the top 20 genes including 10 up-regulated and 10 down-regulated genes were selected for further study in this manuscript.

Protein-protein interaction (PPI) network construction and hub genes exploration

The PPI network of selected 20 DEGs was constructed via the STRING database [26]. This database was built to provide vital information

about predicted and experimental proven protein-protein interactions. In the present study, the PPI of the DEGs having interaction scores greater than 0.7 was considered for further analysis. Subsequently, the top eight hub genes (highly interacting genes) based on the degree method from the obtained PPI network were explored using Cytohubba [27], which is a plugin application in Cytoscape [28].

The lncRNA-miRNA-mRNA regulatory network

To develop the potential lncRNA-miRNA-mRNA regulatory network of the hub genes we used the following different online databases. The miRNAs were predicted via miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/) [29], and miRDB (http://www.mirdb.org) [30] databases. The MiRcode (http://www.mircode.org/) repository [31] was utilized to predict lncRNAs targeting miRNAs. Ultimately, the lncRNA-miRNA-mRNA regulatory network was incorporated using Cytoscape.

Enrichment analysis

Gene Ontology (GO), BP (biological processes), CC (cellular components), MF (molecular function), and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the hub genes were carried out using DAVID (https://david.ncifcrf.gov/). This is a famous online enrichment annotation tool to explore the biological significance of a given list of genes [32, 33].

Drug prediction analysis

DrugBank (http://www.drugbank.ca) database [34], which contains around seven thousand drug entries and four thousand proteins data, was used in this study to evaluate hub genes-associated potential targeted drugs.

Clinical HCM and normal control tissue sample collection

Following the approval of the ethics committee, we conducted a prospective collection of 25 pairs of HCM tissue and corresponding normal samples from patients who visited the Institute of Nishtar Hospital, Multan between August 2022 and May 2023. Prior to their participation, all individuals provided informed consent by signing consent forms. All patients included in the study were diagnosed with HCM and had not undergone any type of therapy.

Total RNA, DNA extraction, RNA sequencing (RNA-seq) and targeted bisulfite sequencing (targeted bisulfite-seq) analyses

Total RNA extraction from all these HCM and control tissues was done using the TRIzol® reagent method [35], and DNA extraction was done via the organic method [36-38]. RNA and DNA samples were sent to the Beijing Genomics Institute (BGI) Company for RNA-sequencing and targeted bisulfite-sequencing analysis.

Following RNA-seq and targeted bisulfite-seq analyses, the gene expression values of the hub genes were normalized using reads per kilo base million reads (RPKM) and fragments per kilo base million reads (FPKM). While methylation values were normalized as beta values. The obtained FPKM and beta values against hub genes in HCM and normal control samples were compared to identify differences in the expression and methylation levels.

Statistics of bioinformatics analysis

DEGs were identified using a t-test [39, 40]. For GO and KEGG enrichment analysis, we used Fisher's Exact test for computing statistical difference [41]. All the analyses were carried out in R version 3.6.3 software.

Results

The overall study design is given in **Figure 1**.

DEGs and hub genes between HCM patients and normal individuals

The variations in the gene expression pattern between normal individuals and patient group in disease may be closely associated with the occurrence of that disease [42]. The normalized gene expression data were taken from the GSE68316 dataset via the GEO database and analyzed to discover DEGs specifically associated with HCM. By selecting *p*-values less than 0.01 and FC greater than 1.2 as the screening thresholds, a total of 850 DEGs were explored between HCM patients and healthy controls (**Figure 2A**). From the obtained DEGs, there were a total of 636 genes that were up-regulated and 270 genes that were down-regulated in HCM samples relative to healthy controls (**Figure 2B**). We selected the top 20 DEGs

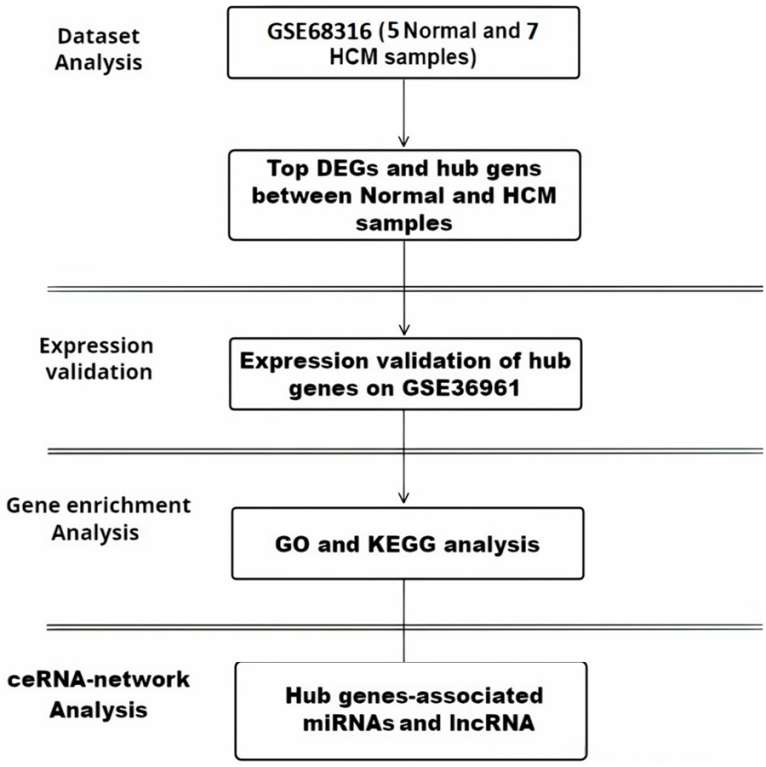

Figure 1. The overall workflow sheet diagram of our study.

including 10 up-regulated and 10 down-regulated genes for further study in this manuscript (**Table 1**). After constructing a PPI of these 20 DEGs (**Figure 2C**), we further shortlisted the eight most important genes as the hub genes including Erythrocyte Membrane Protein Band 42 (EPB42), Ribosomal Protein S24 (RPS24), Ubiquinol-Cytochrome C Reductase Hinge Protein (UQCRH), Tensin 1 (TNS1), Ribosomal Protein L26 (RPL26), Carbonic Anhydrase 1 (CA1), Prefoldin Subunit 5 (PFDN5), and U6 SnRNA-Associated Sm-Like Protein (LSM5) (**Figure 2D**).

Validation of hub genes using additional GEO dataset

To validate the expression levels of hub genes in another GEO expression dataset (GSE36961), the normalized gene expression data from 106 HCM patients and 39 healthy samples were obtained from the GEO database. Expression analysis of the hub genes using GSE36961 revealed that EPB42, UQCRH, CA1, PFDN5, and LSM5 hub genes were significantly up-regulated while RPS24, TNS1, and RPL26 hub genes were significantly down-regulated in HCM samples relative to healthy controls (**Figures 3** and **4**). Therefore, the re-analysis of hub gene expression further proves the important roles of these genes in HCM occurrence and progression. In summary, it is suggested that these hub genes between HCM and healthy individuals may be aberrantly expressed genes related to HCM.

lncRNA-miRNA-mRNA regulatory network analysis

We constructed a lncRNA-miRNA-mRNA regulatory network of the hub genes using different online databases (referred to methodology) in the current study. We noticed in the current study that there were 189 miRNAs that target all hub genes including EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, PFDN5, and LSM5 (**Figure 5A**). Further interaction analysis showed that there were 8 miRNAs (has-mir-23b-3p, has-mir-124-3p, has-mir-129-5p, has-mir-182-5p, has-mir-16-5p, has-mir-1-3p, has-mir-7b-5p, and has-mir-27a-23p) in total predicted 189 miRNAs which predicted to be targeting all the hub genes collectively (**Figure 5B**). Additional analysis revealed that, out of the 8 miRNAs, 6 miRNAs (has-mir-1-3p, has-mir-129-5p, has-mir-16-5p, has-mir-23b-3p, has-mir-27-3p, and has-mir-182-5p) were also found targeted by 246 lncRNAs (**Figure 6A**) and out of which only 4 lncRNAs including NUTMB2-AS1, NEAT1, XIST, and GABPB1-AS1 revealed to be collectively targeting all 6 miRNAs (**Figure 6B**).

GO enrichment analysis

GO analysis for predicting BP, CC, and MF of the hub genes was conducted via the DAVID tool. The identified hub genes were highly enriched in Reg. of DNA damage response, signal transduction, by p53 class mediator resulting, Pos. reg. of intrinsic apoptotic signaling pathway by p53 class mediator, Pos. reg. of intrinsic apoptotic signaling pathway in response to DNA damage, SRP-dependent cotranslational pro-

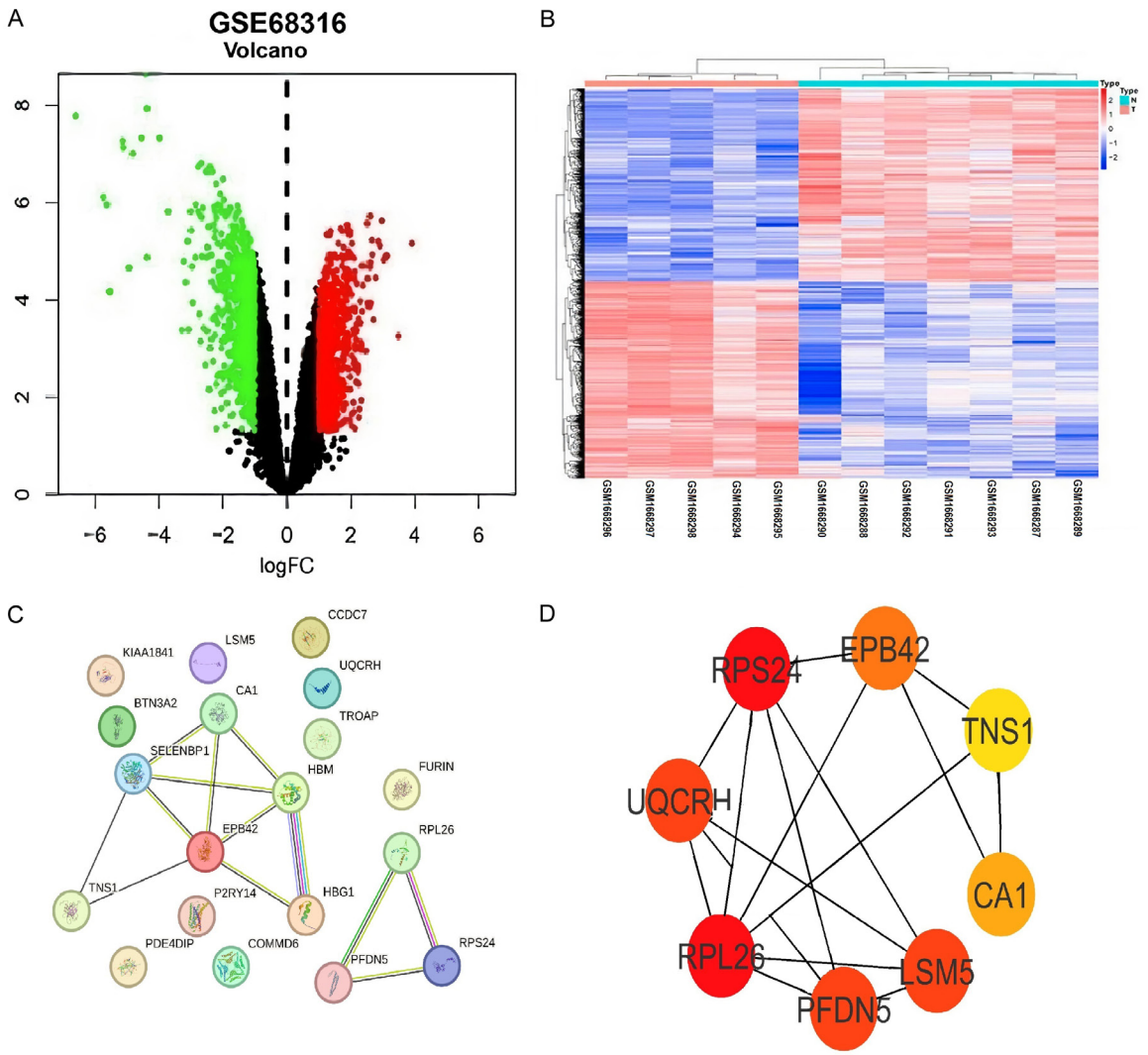

Figure 2. This figure illustrates key aspects of the analysis, incorporating visual representations of differentially expressed genes (DEGs) and hub genes. (A) showcases a volcano graph depicting the DEGs identified in the GSE68316 dataset. (B) presents an overarching heatmap displaying the expression patterns of DEGs identified in the GSE133054 dataset. (C and D) depict protein-protein interaction (PPI) networks for the DEGs and hub genes, respectively.

tein targeting to membrane, Protein targeting to ER, Establishment of protein localization to endoplasmic reticulum, Erythrocyte homeostasis, Nuclear-transcribed mRNA catabolic processes, nonsense-mediated decay, and Protein targeting to endoplasmic reticulum, etc. BP terms (**Figure 7A**).

Concerning CC, the identified hub genes were mainly enriched in Lsm-7-Pat complex, Prefoldin complex, Lsm2-8 complex, Mitochondrial respiratory chain complex III, Respiratory chain complex III, U6 snRNP, Cytosolic ribosome, Ribosomal subunit, Ribosome, Ribonucleoprotein complex, etc. CC terms (**Figure 7B**).

Regarding MF, the identified hub genes were mainly involved in Arylestrease activity, Ubiquinol-cytochrome-c reductase activity, Oxidoreductase activity, acting on diphenois and related substances as donors, Protei glutamine gamma-glutamyltransferase activity, Carbonate dehydratase activity, Structural constituent of ribosome, etc. MF terms (**Figure 7C**).

KEGG pathway analysis

To further clarify the roles of hub genes (EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, PFDN5, and LSM5) in the development of HCM, the KEGG pathway analysis was also performed by

Table 1. The top 20 dysregulated genes from the GSE68316 dataset

Sr. No	Gene symbol	Expression Status	Log FC	Adjust *P*-value
1	SELENBP1	Up-regulation	6.481	2.22E-17
2	LSM5	Up-regulation	6.360	2.91E-13
3	HBG1	Up-regulation	5.848	3.36E-06
4	TROAP	Up-regulation	5.734	6.80E-17
5	BTN3A2	Up-regulation	5.246	4.26E-17
6	UQCRH	Up-regulation	5.117	1.72E-18
7	CA1	Up-regulation	5.066	1.24E-13
8	SLCA4A1	Up-regulation	4.797	8.42E-15
9	PFDN5	Up-regulation	4.557	2.08E-18
10	PDE4DIP	Up-regulation	4.386	7.35E-19
11	COMMD6	Down-regulation	3.843	2.02E-19
12	P2RY14	Down-regulation	3.299	7.18E-13
13	CCDC7	Down-regulation	3.080	4.90E-17
14	KIAA1841	Down-regulation	3.010	1.26E-19
15	TNS1	Down-regulation	2.997	2.31E-17
15	RPL26	Down-regulation	2.995	6.68E-20
17	RPS24	Down-regulation	2.902	6.91E-19
18	FURIN	Down-regulation	2.793	1.51E-13
19	EPB42	Down-regulation	2.764	5.80E-17
20	HBM	Down-regulation	2.682	2.17E-17

DAVID. The selected hub genes were mainly involved in Nitrogen metabolism, Ribosome, RNA degradation, Cardiac muscle contraction, and Coronavirus disease, etc. KEGG terms (**Figure 8A-C**).

Drug prediction analysis

Medical treatment is the preliminary choice to handle disease for patients who are suffering from HCM. Therefore, a selection of suitable candidate potential drugs is necessary. In the current study, with respect to identified hub genes, we explored some suitable therapeutic drugs for the treatment of HCM via the DrugBank database. For example, Rofecoxib drug was identified as the negative expression regulator of EPB42 mRNA expression (**Table 2**) while Acetaminophen was identified as the negative expression regulator of RPS24 mRNA expression (**Table 2**).

Hub genes expression and promoter methylation analysis using clinical HCM and control samples

In the current study, using RNA-seq and targeted bisulfite-seq data of 25 HCM and control samples, the expression and promoter methylation levels of identified hub genes were validated. The expression and promoter methylation levels of these genes were validated using FPKM and beta values, which are quantitative values with widespread use in the RNA-seq and targeted bisulfite-seq analyses.

As shown in **Figure 8A**, it was noticed that EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, PFDN5, and LSM5 hub genes were expressed in both HCM and normal control samples and FPKM values of EPB42, UQCRH, CA1, PFDN5, and LSM5 were higher while FPKM values of RPS24, TNS1, and RPL26 were higher in HCM samples as compared to normal controls (**Figure 9A**). Moreover, the beta values of the EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, and PFDN5 were lower, while beta values of RPS24, TNS1, and RPL26 were higher in HCM samples as compared to normal controls (**Figure 9B**).

Discussion

Exploring the exact underlying molecular mechanisms of HCM can help to improve the clinical management of this disease. The current study

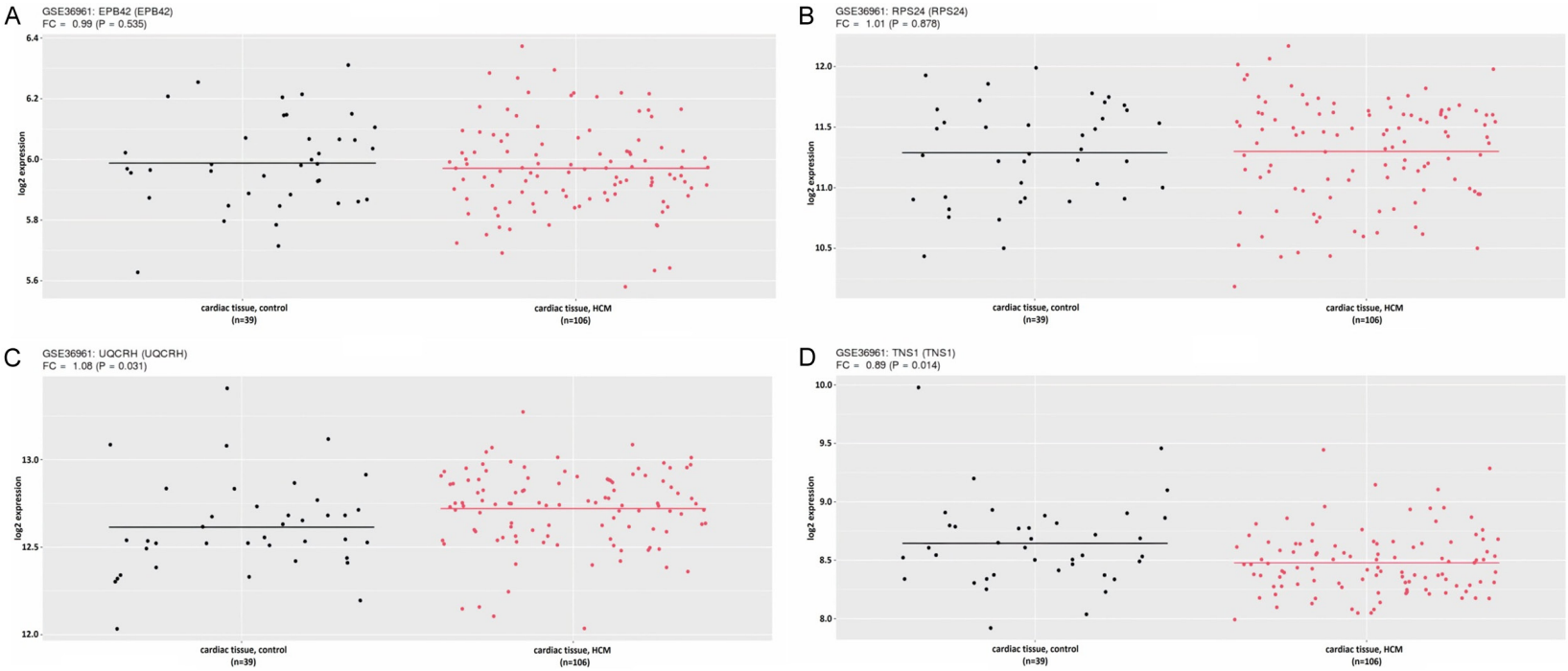

Figure 3. The mRNA expression levels of the shortlisted 4 hub genes were validated using the GSE36961 dataset. (A) EPB42, (B) RPS24, (C) UQCRH, and (D) TNS1. A $P < 0.01$ was regarded as the selection criteria. EPB42 = Erythrocyte Membrane Protein Band 42, RPS24 = Ribosomal protein S24, UQCRH = Ubiquinol-cytochrome c reductase hinge protein, TNS1 = Tensin 1.

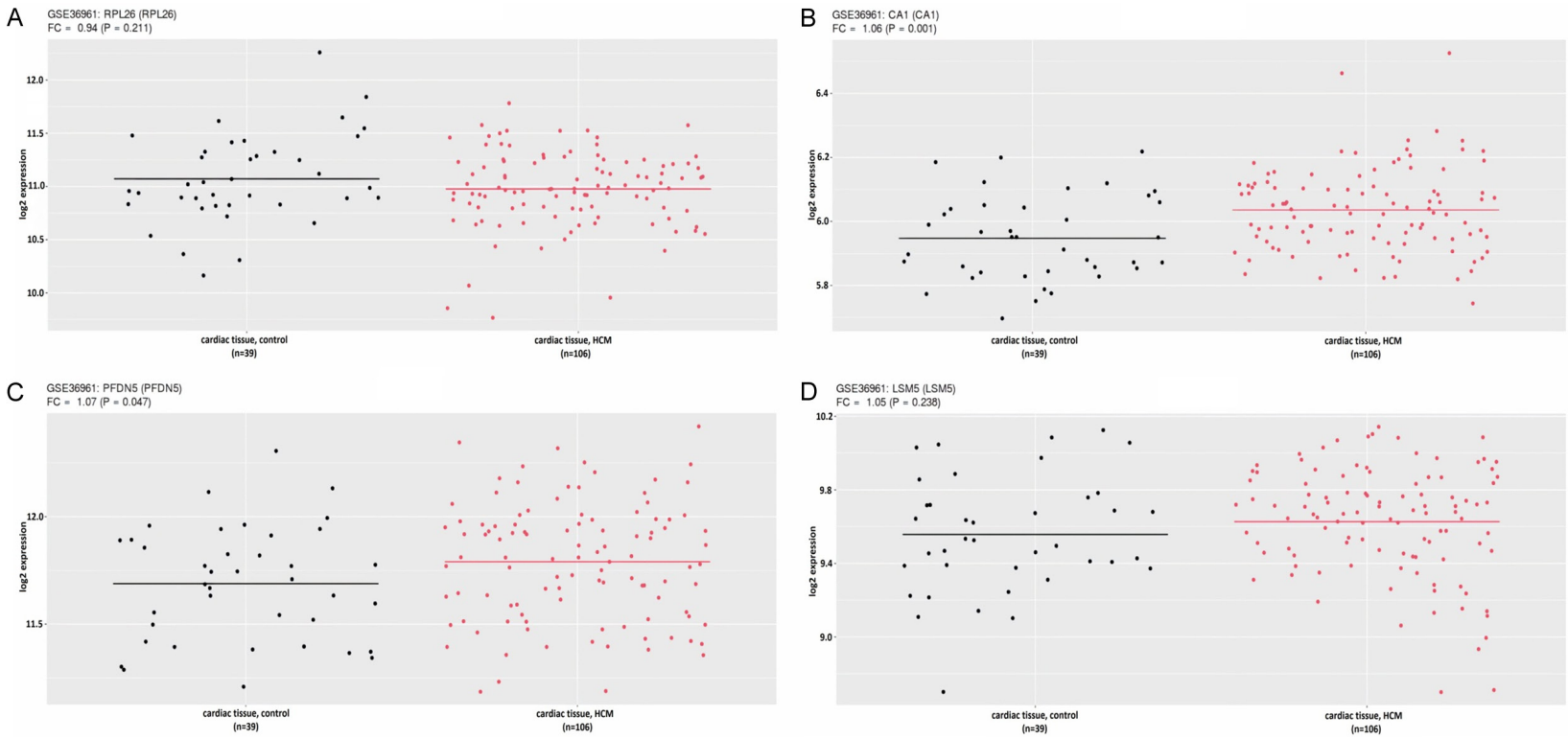

Figure 4. The mRNA expression levels of the shortlisted 4 hub genes were validated using the GSE36961 dataset. (A) RPL25, (B) CA1, (C) PFDN5, and (D) LSM5. A $P < 0.01$ was regarded as the selection criteria. RPL25 = Ribosomal Protein L26, CA1 = CA1 carbonic anhydrase 1, PFDN5 = Prefoldin subunit 5, LSM5 = U6 SnRNA-Associated Sm-Like Protein.

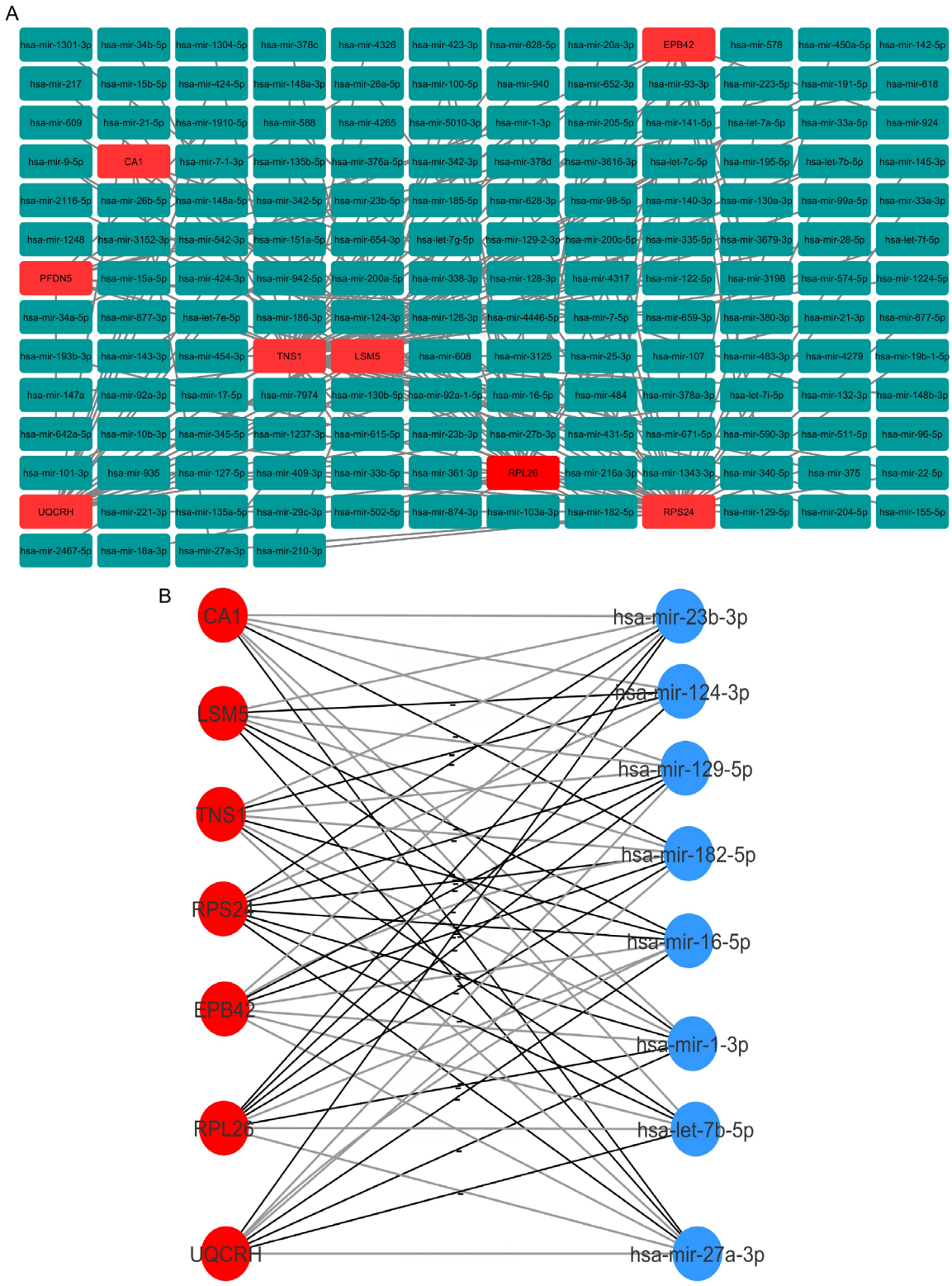

Figure 5. PPI networks highlighting associations between miRNAs and identified hub genes. A. A network of overall predicted miRNAs targeting hub genes, and B. A PPI network between meaningful 8 miRNAs and hub genes. The red nodes are hub genes, green nodes are the miRNAs, and blue nodes are the meaningful miRNAs. PPI = Protein-protein interaction.

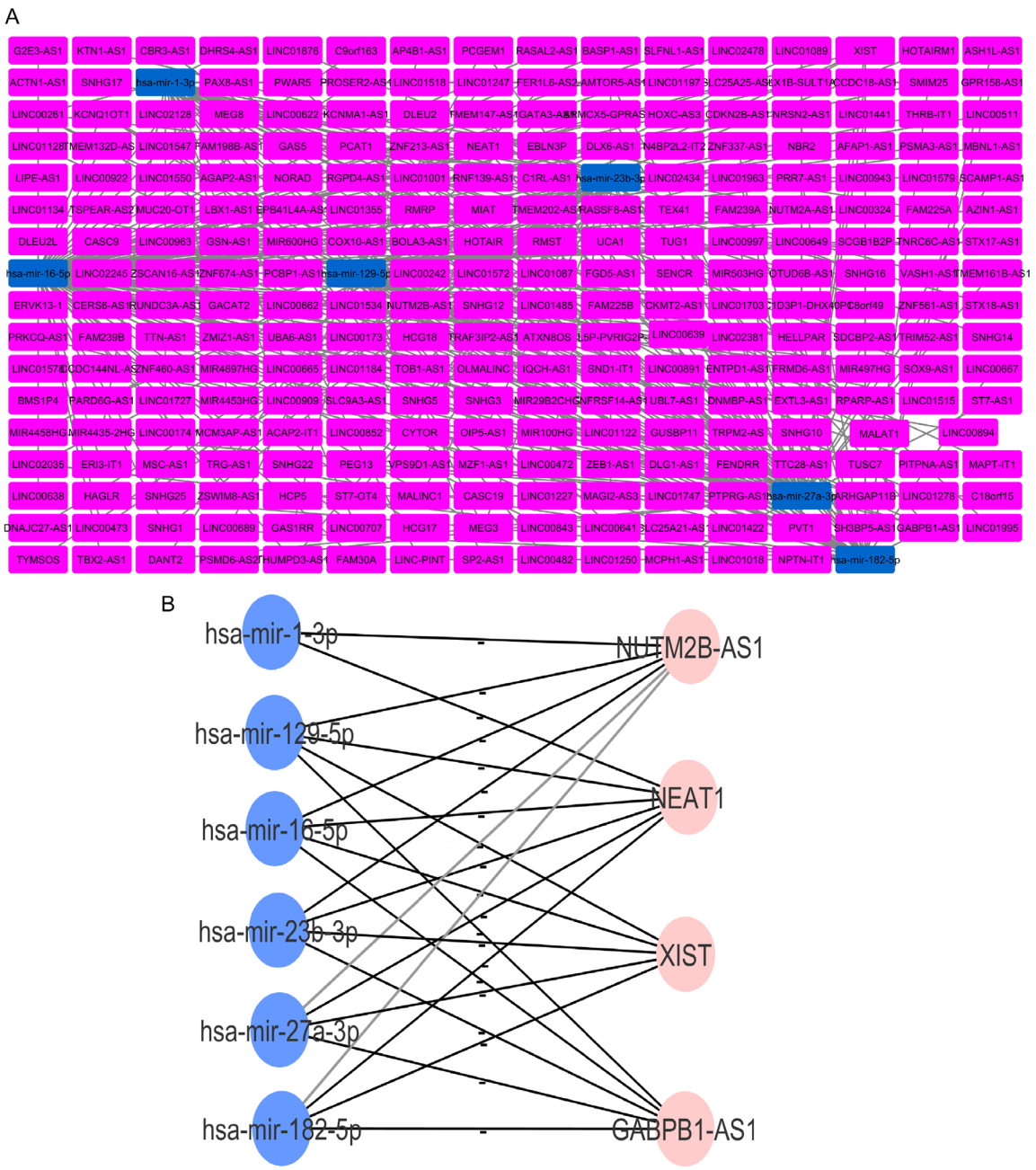

Figure 6. PPI networks highlighting associations between lncRNAs and identified miRNAs. (A) A network of overall predicted lncRNAs targeting miRNAs, and (B) A PPI network between meaningful 4 lncRNAs and miRNAs. Pink nodes are lncRNAs and blue nodes are the miRNAs. PPI = Protein-protein interaction.

was based on DEGs in HCM patients relative to healthy individuals, combined with protein-protein interaction analysis to discover a few key hub genes associated with HCM. We analyzed HCM GSE68316 and GSE36961 datasets from the GEO database to identify the top 20 DEGs, hub genes, and explore the molecular mechanisms of this disease.

By constructing a PPI network of the identified DEGs, we found 8 hub genes in HCM patients, namely EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, PFDN5, and LSM5. The expression verification of these hub genes via the HCM GSE36961 dataset revealed that EPB42, UQCRH, CA1, PFDN5, and LSM5 hub genes were significantly up-regulated while RPS24,

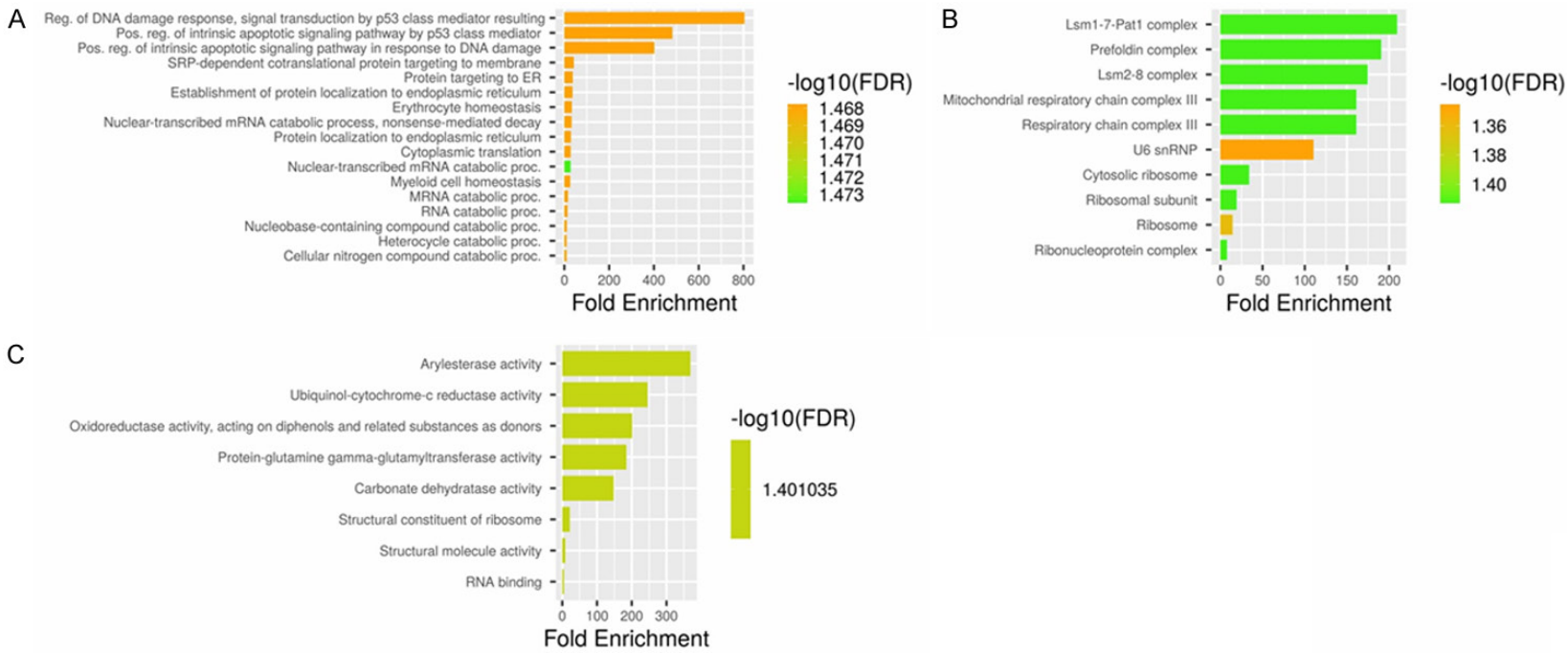

Figure 7. GO enrichment analysis of the hub genes via the DAVID. (A) BP terms, (B) CC terms, and (C) MF terms. A $P < 0.01$ was regarded as the selection criteria. GO = Gene Ontology, BP = Biological process, CC = Cellular components, MF = Molecular function.

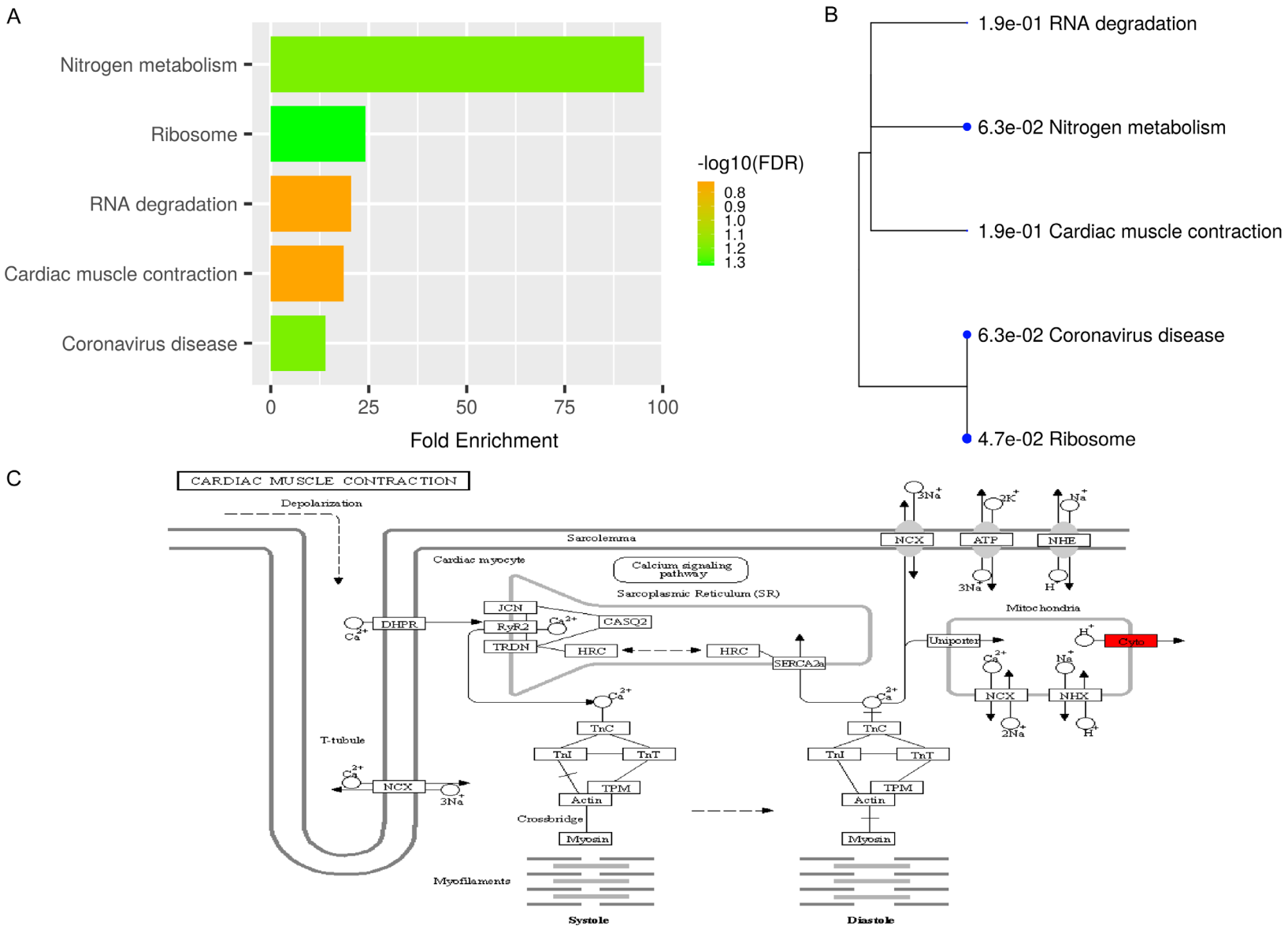

Figure 8. KEGG enrichment analysis of the hub genes via the DAVID. (A) KEGG terms, (B) KEGG terms phylogram, and (C) Most significant KEGG term. A $P < 0.01$ was regarded as the selection criteria. KEGG = Kyoto Encyclopedia of Genes and Genomes.

Table 2. Hub genes associated drugs from DrugBank database

Sr. No	Hub gene	Drug name	Effect	Reference	Group
1	EPB42	Rofecoxib	Increased expression of EPB42 mRNA	A22107	Approved
2	RPS24	Acetaminophen	Decreased expression of RPS24 mRNA	A20420	Approved
		Vincristine		A22706	
3	UQCRH	Disulfiram	Decreased expression of UQCRH mRNA	A21750	Approved
4	TNS1	Ciglitazone	Increased expression of TNS1 mRNA	A21647	Approved
		Resveratrol		A24017	
5	RPL26	Bortezomib	Increased expression of RPL26 mRNA	A21434	Approved
6	CA1	Rifampicin	Decreased expression of RPS24 mRNA	A24040	Approved
7	PFDN5	Estradiol	Decreased expression of PFDN5 mRNA	A21098	Approved
		Valproic acid		A24690	
8	LSM5	Tretinoin	Decreased expression of LSM5 mRNA	A24454	Approved

TNS1, and RPL26 hub genes were significantly down-regulated in HCM samples relative to healthy controls.

EPB42 is an ATP-binding protein, which is reported to regulate the association between protein 3 and ankyrin [43]. This protein may also have a key role in maintaining erythrocyte shape and providing mechanical properties [43]. Mutations in this gene or its dysregulation are linked to the recessive spherocytic elliptocytosis [44]. However, we are the first to report the dysregulation of this gene in HCM. RPS24 and RPL26 genes encode ribosomal proteins, which are part of the ribosome machinery [45]. Ribosomes are the organelles in the cells which carry out the synthesis of proteins [46]. Mutations in these genes are known to be associated with Diamond Blackfan Anemia and HCM [47]. UQCRH protein plays an important role to affect the mitochondrial function [48]. Recent data suggested that this protein also plays an important role in cardioprotection and its dysregulation may disturb normal heart functioning [49]. The TNS1 gene encodes for the TNS1 protein, which plays a key role in attaching the membrane of a cell to the extracellular matrix [50]. In recent research, the dysregulation of TNS1 resulted in myocardial infraction in zebra fish model [50]. CA1 gene in humans encodes for carbonic anhydrase 1 enzyme, which catalyzes the conversion of carbon dioxide to carbonic acid [51]. The dysregulation of this gene was previously correlated with heart failure and other diseases [52]. PFDN5 gene codes for a prefoldin alpha subunit family member involved in the stabilization of newly synthesized polypeptide [53]. Dysregulation of the PFDN5 gene was earlier linked to the development and progression of cancer [54]. However, we are the first to report the dysregulation of this gene in HCM. LSM5 encodes a protein that plays a key role in the assembly of spliceosome [55]. Abnormal expression of this gene was associated with circadian rhythms and HCM [56].

In recent times, the participation of miRNAs, and lncRNAs in HCM development is a hot topic for researchers. In this regard, different studies have discovered the involvement of a variety of miRNAs and lncRNAs in HCM. For example, the lncRNA Rp5-833A20.1 was involved in the prevention of miR-382-5p to target NFIA, which is associated with HCM [57]. The lncRNA MIAT was verified to promote myocardial fibrosis after acute ischemia by targeting miRNA24 to target two important mRNAs including Furin and TGF-β1 [58]. The lncRNA ROR is also reported to sponge miR-133 to dysregulate two important mRNAs (ANP and BNP), associated with HCM [59]. In the current manuscript, concerning lncRNA-miRNA-mRNA regulatory network, we explored 6 miRNAs (has-mir-1-3p, has-mir-129-5p, has-mir-16-5p, has-mir-23b-3p, has-mir-27-3p, and has-mir-182-5p) targeting all the shortlisted hub genes and getting targeted by one only 4 lncRNAs including NUTMB2-AS1, NEAT1, XIST, and GABPB1-AS1. These critical data may expand our knowledge about the regulatory mechanisms of HCM. In addition, functional analysis revealed that HCM-associated dysregulated hub genes were majorly involved in a variety of GO and KEGG terms (**Figures 6** and **7**). Lastly, we have also suggested a few hub gene-related therapeutic

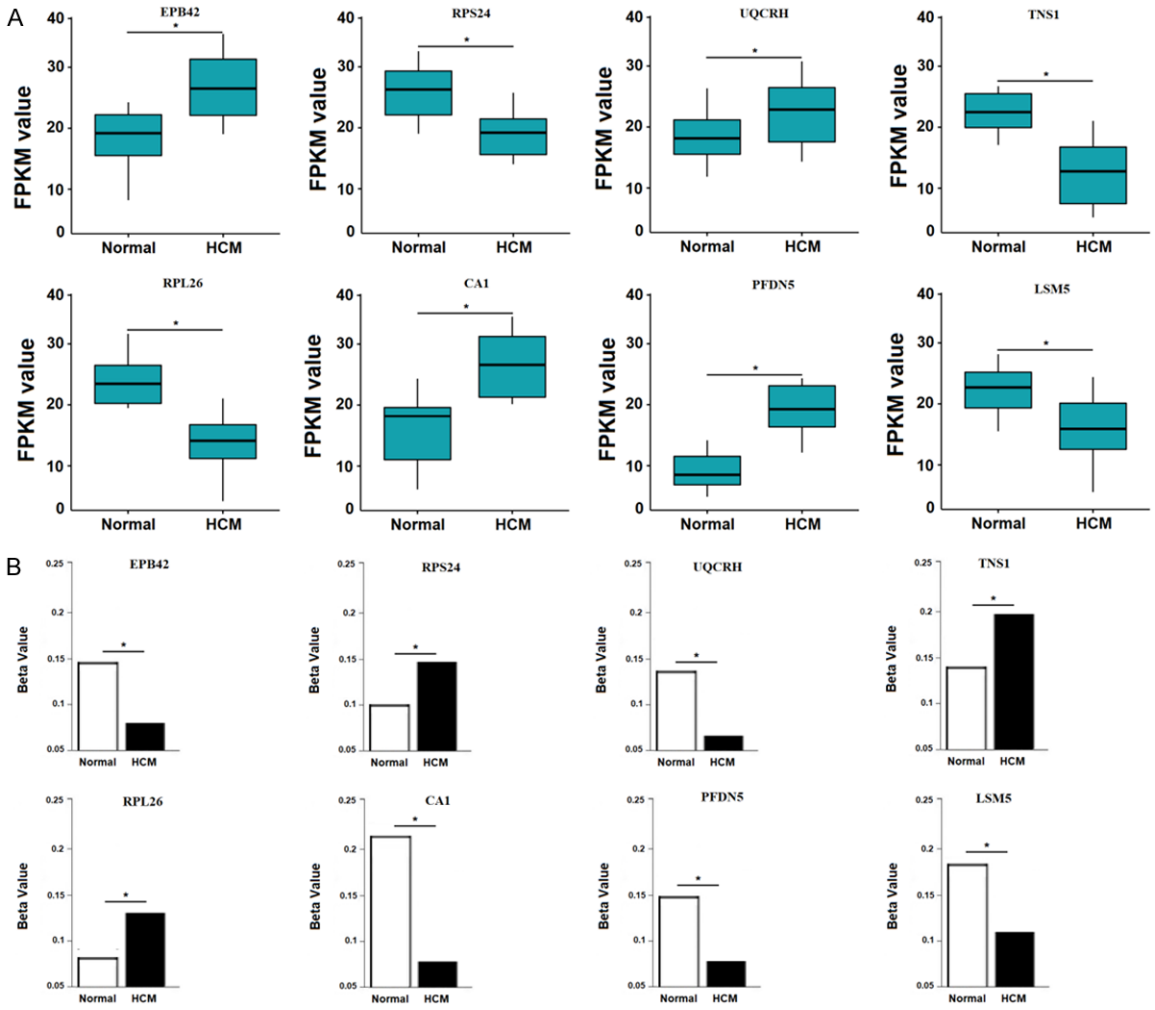

Figure 9. Validating hub gene expression and promoter methylation using clinical HCM and control samples via RNA sequencing and targeted bisulfite sequencing analysis. (A) FPKM values based expression validation of the hub genes, and (B) Beta values based promoter methylation validation of the hub genes. HCM = Hypertrophic cardiomyopathy, FPKM = Fragments Per Kilobase of transcript per Million mapped reads.

drugs from the DrugBank database whose use might help to treat HCM patients.

Previously, it was reported in the medical literature that HCM is the outcome of genetic mutations across the oncoprotein coding gene in Mendel's autosomal dominant genetic pattern [60]. The discovery of novel dysregulated hub genes in the current study may be helpful for clinically determining the diagnosis of HCM patients. The outcomes of the present study may also guide clinicians to design and use novel treatment therapies for HCM patients.

Conclusion

In the present study, 8 key dysfunctional hub genes (EPB42, RPS24, UQCRH, TNS1, RPL26, CA1, PFDN5, and LSM5) were discovered via detailed analysis between HCM and control samples. Via further enrichment analysis, we revealed that underlying molecular mechanisms behind HCM were associated with Nitrogen metabolism, Ribosome, RNA degradation, Cardiac muscle contraction, and Coronavirus disease signaling pathways. However, this study requires further experimentation to validate the analysis results. One limitation of this study is the focus on developmental diseases as the research subject. As a result, there is no availability of clinical data to accurately determine the patients' origins, ages, and genders. This restricts the generalizability and applicability of the findings to specific groups. Additionally, there is a scarcity of clinical samples to validate crucial analysis results.

Disclosure of conflict of interest

None.

Address correspondence to: Yasir Hameed, Department of Biotechnology, Institute of Biochemistry Biotechnology and Bioinformatics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan. E-mail: Yasirhameed2011@gmail.com; Saqib Ali, Department of Computer Science, University of Agriculture, Faisalabad 38040, Pakistan. E-mail: saqib@uaf.edu.pk

References

[1] Calderon Martinez E, Ortiz-Garcia NY, Herrera Hernandez DA, Arriaga Escamilla D, Diaz Mendoza DL, Othon Martinez D, Ramirez LM, Reyes-Rivera J, Choudhari J and Michel G. Hypertrophic cardiomyopathy diagnosis and treatment in high- and low-income countries: a narrative review. Cureus 2023; 15: e46330.

[2] Glavaški M, Velicki L and Vučinić N. Hypertrophic cardiomyopathy: genetic foundations, outcomes, interconnections, and their modifiers. Medicina (Kaunas) 2023; 59: 1424.

[3] Gartzonikas IK, Naka KK and Anastasakis A. Current and emerging perspectives on pathophysiology, diagnosis, and management of hypertrophic cardiomyopathy. Hellenic J Cardiol 2023; 70: 65-74.

[4] Desai MY, Owens A, Wolski K, Geske JB, Saberi S, Wang A, Sherrid M, Cremer PC, Lakdawala NK, Tower-Rader A, Fermin D, Naidu SS, Smedira NG, Schaff H, McErlean E, Sewell C, Mudarris L, Gong Z, Lampl K, Sehnert AJ and Nissen SE. Mavacamten in patients with hypertrophic cardiomyopathy referred for septal reduction: week 56 results from the VALOR-HCM randomized clinical trial. JAMA Cardiol 2023; 8: 968-977.

[5] Picano E, Pierard L, Peteiro J, Djordjevic-Dikic A, Sade LE, Cortigiani L, Van De Heyning CM, Celutkiene J, Gaibazzi N, Ciampi Q, Senior R, Neskovic AN and Henein M. The clinical use of stress echocardiography in chronic coronary syndromes and beyond coronary artery disease: a clinical consensus statement from the European Association of Cardiovascular Imaging of the ESC. Eur Heart J Cardiovasc Imaging 2024; 25: e65-e90.

[6] Ottaviani A, Mansour D, Molinari LV, Galanti K, Mantini C, Khanji MY, Chahal AA, Zimarino M, Renda G, Sciarra L, Pelliccia F, Gallina S and Ricci F. Revisiting diagnosis and treatment of hypertrophic cardiomyopathy: current practice and novel perspectives. J Clin Med 2023; 12: 5710.

[7] Maron MS, Olivotto I, Betocchi S, Casey SA, Lesser JR, Losi MA, Cecchi F and Maron BJ. Effect of left ventricular outflow tract obstruction on clinical outcome in hypertrophic cardiomyopathy. N Engl J Med 2003; 348: 295-303.

[8] Field E, Norrish G, Acquaah V, Dady K, Cicerchia MN, Ochoa JP, Syrris P, McLeod K, McGowan R, Fell H, Lopes LR, Cervi E and Kaski JPP. Cardiac myosin binding protein-C variants in paediatric-onset hypertrophic cardiomyopathy: natural history and clinical outcomes. J Med Genet 2022; 59: 768-775.

[9] Rowin EJ, Maron BJ, Olivotto I and Maron MS. Role of exercise testing in hypertrophic cardiomyopathy. JACC Cardiovasc Imaging 2017; 10: 1374-1386.

[10] Desai MY, Owens A and Wang A. Medical therapies for hypertrophic cardiomyopathy: current state of the art. Prog Cardiovasc Dis 2023; 80: 32-37.

[11] Melas M, Beltsios ET, Adamou A, Koumarelas K and McBride KL. Molecular diagnosis of hypertrophic cardiomyopathy (HCM): in the heart of cardiac disease. J Clin Med 2022; 12: 225.

[12] Maron BA, Wang RS, Carnethon MR, Rowin EJ, Loscalzo J, Maron BJ and Maron MS. What causes hypertrophic cardiomyopathy? Am J Cardiol 2022; 179: 74-82.

[13] Cirino AL, Seidman CE and Ho CY. Genetic testing and counseling for hypertrophic cardiomyopathy. Cardiol Clin 2019; 37: 35-43.

[14] Marwaha S, Knowles JW and Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. Genome Med 2022; 14: 23.

[15] de Boer RA, Heymans S, Backs J, Carrier L, Coats AJS, Dimmeler S, Eschenhagen T, Filippatos G, Gepstein L, Hulot JS, Knöll R, Kupatt C, Linke WA, Seidman CE, Tocchetti CG, van der Velden J, Walsh R, Seferovic PM and Thum T. Targeted therapies in genetic dilated and hypertrophic cardiomyopathies: from molecular mechanisms to therapeutic targets. A position paper from the Heart Failure Association (HFA) and the Working Group on Myocardial Function of the European Society of Cardiology (ESC). Eur J Heart Fail 2022; 24: 406-420.

[16] Popa-Fotea NM, Micheu MM, Bataila V, Scafa-Udriste A, Dorobantu L, Scarlatescu AI, Zamfir D, Stoian M, Onciul S and Dorobantu M. Exploring the continuum of hypertrophic cardiomyopathy-from DNA to clinical expression. Medicina (Kaunas) 2019; 55: 299.

[17] Erdmann J, Daehmlow S, Wischke S, Senyuva M, Werner U, Raible J, Tanis N, Dyachenko S, Hummel M, Hetzer R and Regitz-Zagrosek V. Mutation spectrum in a large cohort of unrelated consecutive patients with hypertrophic cardiomyopathy. Clin Genet 2003; 64: 339-349.

[18] Burns C, Bagnall RD, Lam L, Semsarian C and Ingles J. Multiple gene variants in hypertrophic cardiomyopathy in the era of next-generation sequencing. Circ Cardiovasc Genet 2017; 10: e001666.
[19] Maron BJ, Bonow RO, Cannon RO 3rd, Leon MB and Epstein SE. Hypertrophic cardiomyopathy. Interrelations of clinical manifestations, pathophysiology, and therapy (2). N Engl J Med 1987; 316: 844-852.
[20] Wu S, Ud Din I, Sadiq FM, Abdel-Maksoud MA, Haris M, Mubarak A, Farrag MA, Alghamdi S, Almekhlafi S, Akram M and Li G. Dysfunctional network of hub genes in hypertrophic cardiomyopathy patients. Am J Transl Res 2022; 14: 8918-8933.
[21] Koo SJ, Spratt HM, Soman KV, Stafford S, Gupta S, Petersen JR, Zago MP, Kuyumcu-Martinez MN, Brasier AR, Wiktorowicz JE and Garg NJ. S-Nitrosylation proteome profile of peripheral blood mononuclear cells in human heart failure. Int J Proteomics 2016; 2016: 1384523.
[22] Zhao M, Fajardo G, Urashima T, Spin JM, Poorfarahani S, Rajagopalan V, Huynh D, Connolly A, Quertermous T and Bernstein D. Cardiac pressure overload hypertrophy is differentially regulated by β-adrenergic receptor subtypes. Am J Physiol Heart Circ Physiol 2011; 301: H1461-H1470.
[23] Aboutabl ME, Zordoky BN and El-Kadi AO. 3-methylcholanthrene and benzo(a)pyrene modulate cardiac cytochrome P450 gene expression and arachidonic acid metabolism in male Sprague Dawley rats. Br J Pharmacol 2009; 158: 1808-1819.
[24] Yang W, Li Y, He F and Wu H. Microarray profiling of long non-coding RNA (lncRNA) associated with hypertrophic cardiomyopathy. BMC Cardiovasc Disord 2015; 15: 62.
[25] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
[26] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.
[27] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014; 8 Suppl 4: S11.
[28] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[29] Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, Tang Y, Chen YG, Jin CN, Yu Y, Xu JT, Li YM, Cai XX, Zhou ZY, Chen XH, Pei YY, Hu L, Su JJ, Cui SD, Wang F, Xie YY, Ding SY, Luo MF, Chou CH, Chang NW, Chen KW, Cheng YH, Wan XH, Hsu WL, Lee TY, Wei FX and Huang HD. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 2020; 48: D148-D154.
[30] Wang X. miRDB: a microRNA target prediction and functional annotation database with a wiki interface. RNA 2008; 14: 1012-1017.
[31] Jeggari A, Marks DS and Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics 2012; 28: 2062-2063.
[32] Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007; 8: R183.
[33] Zhang L, Sahar AM, Li C, Chaudhary A, Yousaf I, Saeedah MA, Mubarak A, Haris M, Nawaz M, Reem MA, Ramadan FA, Mostafa AAM, Feng W and Hameed Y. A detailed multi-omics analysis of GNB2 gene in human cancers. Braz J Biol 2022; 84: e260169.
[34] Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, Gautam B and Hassanali M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 2008; 36: D901-906.
[35] Rio DC, Ares M Jr, Hannon GJ and Nilsen TW. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb Protoc 2010; 2010: pdb.prot5439.
[36] Gupta N. DNA extraction and polymerase chain reaction. J Cytol 2019; 36: 116-117.
[37] Hameed Y and Ejaz S. TP53 lacks tetramerization and N-terminal domains due to novel inactivating mutations detected in leukemia patients. J Cancer Res Ther 2021; 17: 931-937.
[38] Ullah L, Hameed Y, Ejaz S, Raashid A, Iqbal J, Ullah I and Ejaz SA. Detection of novel infiltrating ductal carcinoma-associated BReast CAncer gene 2 mutations which alter the deoxyribonucleic acid-binding ability of BReast CAncer gene 2 protein. J Cancer Res Ther 2020; 16: 1402-1407.
[39] Kim TK. T test as a parametric statistic. Korean J Anesthesiol 2015; 68: 540-546.
[40] Khalil T, Okla MK, Al-Qahtani WH, Ali F, Zahra M, Shakeela Q, Ahmed S, Akhtar N, AbdElgawad H, Asif R, Hameed Y, Adetunji CO, Farid A and

Ghazanfar S. Tracing probiotic producing bacterial species from gut of buffalo (Bubalus bubalis), South-East-Asia. Braz J Biol 2022; 84: e259094.
[41] Kim HY. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod 2017; 42: 152-155.
[42] Eichler EE. Genetic variation, comparative genomics, and the diagnosis of disease. N Engl J Med 2019; 381: 64-74.
[43] Li W, Zhao Y, Wang D, Ding Z, Li C, Wang B, Xue X, Ma J, Deng Y, Liu Q, Zhang G, Zhang Y, Wang K and Yuan B. Transcriptome research identifies four hub genes related to primary myelofibrosis: a holistic research by weighted gene co-expression network analysis. Aging (Albany NY) 2021; 13: 23284-23307.
[44] He BJ, Liao L, Deng ZF, Tao YF, Xu YC and Lin FQ. Molecular genetic mechanisms of hereditary spherocytosis: current perspectives. Acta Haematol 2018; 139: 60-66.
[45] Gopanenko AV, Kolobova AV, Tupikin AE, Kabilov MR, Malygin AA and Karpova GG. Knockdown of the ribosomal protein eL38 in HEK293 cells changes the translational efficiency of specific genes. Int J Mol Sci 2021; 22: 4531.
[46] Pecoraro A, Pagano M, Russo G and Russo A. Ribosome biogenesis and cancer: overview on ribosomal proteins. Int J Mol Sci 2021; 22: 5496.
[47] Gazda HT, Preti M, Sheen MR, O'Donohue MF, Vlachos A, Davies SM, Kattamis A, Doherty L, Landowski M, Buros C, Ghazvinian R, Sieff CA, Newburger PE, Niewiadomska E, Matysiak M, Glader B, Atsidaftos E, Lipton JM, Gleizes PE and Beggs AH. Frameshift mutation in p53 regulator RPL26 is associated with multiple physical abnormalities and a specific pre-ribosomal RNA processing defect in diamond-blackfan anemia. Hum Mutat 2012; 33: 1037-1044.
[48] Wang Q, Li M, Gan Y, Jiang S, Qiao J, Zhang W, Fan Y, Shen Y, Song Y, Meng Z, Yao M, Gu J, Zhang Z and Tu H. Mitochondrial protein UQCRC1 is oncogenic and a potential therapeutic target for pancreatic cancer. Theranostics 2020; 10: 2141-2157.
[49] Yi T, Chen H, Zhan J, Li Y, Long Z, Wu Z, Yang M, Peng T and Li H. Ubiquinol-cytochrome c reductase core protein 1 contributes to cardiac tolerance to acute exhaustive exercise. Exp Biol Med (Maywood) 2022; 247: 165-173.
[50] Iske J, Roesel MJ, Cesarovic N, Pitts L, Steiner A, Knoedler L, Nazari-Shafti TZ, Akansel S, Jacobs S, Falk V, Kempfert J and Kofler M. The potential of intertwining gene diagnostics and surgery for mitral valve prolapse. J Clin Med 2023; 12: 7441.
[51] Aspatwar A, Tolvanen MEE, Barker H, Syrjänen L, Valanne S, Purmonen S, Waheed A, Sly WS and Parkkila S. Carbonic anhydrases in metazoan model organisms: molecules, mechanisms, and physiology. Physiol Rev 2022; 102: 1327-1383.
[52] Torella D, Ellison GM, Torella M, Vicinanza C, Aquila I, Iaconetti C, Scalise M, Marino F, Henning BJ, Lewis FC, Gareri C, Lascar N, Cuda G, Salvatore T, Nappi G, Indolfi C, Torella R, Cozzolino D and Sasso FC. Carbonic anhydrase activation is associated with worsened pathological remodeling in human ischemic diabetic cardiomyopathy. J Am Heart Assoc 2014; 3: e000434.
[53] St-Germain J, Khan MR, Bavykina V, Desmarais R, Scott M, Boissonneault G, Brunet MA and Laurent B. Functional characterization of a PHF8 processed pseudogene in the mouse genome. Genes (Basel) 2023; 14: 172.
[54] Herranz-Montoya I, Park S and Djouder N. A comprehensive analysis of prefoldins and their implication in cancer. iScience 2021; 24: 103273.
[55] Koncz C, Dejong F, Villacorta N, Szakonyi D and Koncz Z. The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator. Front Plant Sci 2012; 3: 9.
[56] Hsieh MC, Yang SC, Tseng HL, Hwang LL, Chen CT and Shieh KR. Abnormal expressions of circadian-clock and circadian clock-controlled genes in the livers and kidneys of long-term, high-fat-diet-treated mice. Int J Obes (Lond) 2010; 34: 227-239.
[57] Hu YW, Zhao JY, Li SF, Huang JL, Qiu YR, Ma X, Wu SG, Chen ZP, Hu YR, Yang JY, Wang YC, Gao JJ, Sha YH, Zheng L and Wang Q. RP5-833A20.1/miR-382-5p/NFIA-dependent signal transduction pathway contributes to the regulation of cholesterol homeostasis and inflammatory reaction. Arterioscler Thromb Vasc Biol 2015; 35: 87-101.
[58] Qu X, Du Y, Shu Y, Gao M, Sun F, Luo S, Yang T, Zhan L, Yuan Y, Chu W, Pan Z, Wang Z, Yang B and Lu Y. MIAT is a pro-fibrotic long non-coding RNA governing cardiac fibrosis in post-infarct myocardium. Sci Rep 2017; 7: 42657.
[59] Jiang F, Zhou X and Huang J. Long non-coding RNA-ROR mediates the reprogramming in cardiac hypertrophy. PLoS One 2016; 11: e0152767.
[60] Cui Y, Liu C, Luo J and Liang J. Dysfunctional network and mutation genes of hypertrophic cardiomyopathy. J Healthc Eng 2022; 2022: 8680178.