

Network analysis for the discovery of key biomarkers and potential therapeutic targets in hypertrophic cardiomyopathy

Bin Xie, Zhao Chen, Jian Liu, Jimei Chen, Jian Zhuang, Huimin Guo, Jiyan Chen*

Guangdong Provincial Cardiovascular Institute, Guangdong General Hospital & Guangdong Academy of Medical Sciences, Guangzhou, China

Abstract: Hypertrophic cardiomyopathy (HCM) is a genetic heart disorder that can lead to sudden cardiac death. Current treatment strategies, such as implantable cardioverter-defibrillators, exhibit limitations in impeding disease progression. Despite the identification of several pathogenic genes, the complex mechanisms underlying HCM remain unclear. This study aims to identify novel pathogenic genes and critical biomarkers for HCM using network-based bioinformatics analysis, with the potential to discover therapeutic targets for pharmaceutical interventions. Gene expression data from 119 HCM patients and 55 healthy donors were analyzed using differential gene expression (DEG) analysis, followed by gene ontology (GO), KEGG and Reactome pathway enrichment studies. A protein-protein interaction (PPI) network was constructed to identify hub genes and their interactions, with a focus on potential drug targets. We identified a key gene module, with TNNT1 emerging as a critical hub gene. TNNT1 was found to significantly influence cardiac development and contribute to HCM pathogenesis. Our findings suggest that TNNT1's role in regulating cardiac contractility could provide a foundation for developing new therapeutic agents targeting this pathway. This study provides novel insights into the molecular mechanisms of HCM and identifies TNNT1 as a potential biomarker for early screening and a promising target for pharmaceutical therapies. The identification of TNNT1 offers opportunities for personalized medicine approaches, which could facilitate the development of drugs aimed at modulating its expression and mitigating disease progression in HCM patients.

Keywords: Network analysis; Hypertrophic cardiomyopathy; Protein-protein interaction; Transcription factor-gene interaction.

Submitted on ----- Revised on ----- Accepted on -----

INTRODUCTION

As a common genetic cardiovascular disease, hypertrophic cardiomyopathy (HCM) exhibits a genetic prevalence ranging from 1:500 to 1:200 [1,2]. It is generally considered an inherited condition resulting from mutations in 19 pathogenic genes, including MYH7, MYBPC3, MYL2, MYL3, TPM1, TNNT2, TNNT3, ACTC1, CSRP3, FHL1, PLN, ACTN2, CRYAB, FLNC, MYOZ2, MYH6, TNNC1, TRIM55 and TRIM63 [3-6]. With the advancement and widespread application of gene detection technology, gene screening has been applied to HCM clinical diagnosis and identification of high-risk groups carrying specific gene mutations. Nevertheless, only about 30% of the probands show a sarcomere mutation that is categorized as pathogenic or suspected to be disease-causing in the clinical diagnosis of HCM [7,8]. The pathogenesis of HCM remains enigmatic, particularly with regard to the signal pathway and regulatory network between genes.

In the past decades, microarray technology has been extensively utilized for gene mutation screening. Bioinformatics analysis assists in the identification of differentially expressed genes (DEGs) and functional pathways associated with the mechanism of HCM.

However, independent microarray analysis and the restriction of sample sizes result in elevated false-positive rates, hindering the attainment of accurate outcomes. Thus, in this research, three mRNA microarray datasets were obtained and examined from the Gene Expression Omnibus to identify the DEGs between 115 HCM patients and 55 healthy donors. Subsequently, gene ontology (GO) enrichment, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome pathway were analyzed by constructing a visualization network of DEGs and hub genes. This visualization has contributed to a more profound comprehension of the molecular mechanisms and progression of HCM.

MATERIALS AND METHODS

Transcriptional expression profiling processes

NCBI-GE (<http://www.ncbi.nlm.nih.gov/geo>) is a no-cost public repository that stores transcriptional expression profiles [9]. We collected gene expression profiles of HCM and healthy myocardial tissues from three datasets: GSE36961, GSE32453 and GSE1145. Data from GSE36961 were obtained using the GPL15389 platform (Illumina HumanHT-12 V3.0 expression bead chip) from 106 HCM and 39 donor myocardial tissue samples. Similarly, data from GSE32453 were obtained using the GPL6104 platform (Illumina humanRef-8 v2.0 expression bead chip) from 8 HCM patients and 5 healthy donor

*Corresponding author: e-mail: chenjiyandoctor@163.com

myocardial tissue samples. The GSE1145 data was acquired from the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array), with 5 HCM and 11 healthy myocardial tissues analyzed.

Standardization and elucidation of DEGs

The process of DNA microarray analysis started with preparing and standardizing the raw biological data to remove noise and ensure its accuracy. The robust multi-array average analysis algorithm of the R.limma software package was used for data correction, standardization, and summarization. Identified the DEGs between HCM samples and normal myocardial tissues under experimental conditions [10]. Parameters such as fold change, Benjamini and Hochberg's false discovery rate (FDR), and adjusted P-value (adj.P) were utilized to filter DEGs, aiming to minimize false positives. Removed or averaged the probe sets without corresponding gene symbols or genes with more than one probe set. Genes with absolute values of logarithmic fold change (\log_2FC) >0.5 and P-value <0.05 were deemed to display significant differences between groups. Venn software (<https://bioinfogp.cnb.csic.es/tools/venny/>) was used to analyze raw data in TXT format to identify common DEGs in the three datasets. DEGs with $\log_2FC < 0$ were considered down regulated, while those with $\log_2FC > 0$ were considered upregulated.

GO enrichment and pathway analysis of DEGs

GO is a pivotal bioinformatics instrument for the annotation of genes and the analysis of their biological processes [11]. The KEGG website (<https://www.kegg.jp/>) is a crucial database resource that facilitates comprehension of intricate functions and biological systems by analyzing extensive molecular data generated through high-throughput experimental methods [12]. Reactome (<https://reactome.org/>) is a freely available and open-source database that is manually curated and reviewed by experts. It offers a pathway database that provides intuitive bioinformatics tools for visualizing, interpreting, and analyzing pathway knowledge. This database supports a range of applications, including genome analysis, systems biology, basic and clinical research, modeling, and education [13].

The Database for Annotation, Visualization and Integrated Discovery (DAVID) (<http://david.ncifcrf.gov>) (version 6.8) is a comprehensive online bioinformatics tool that brings together biological data and analysis resources. It offers comprehensive annotation information on gene and protein functions, enabling users to extract significant biological insights [14]. DAVID facilitates the visualization of DEGs enrichment in Biological Process (BP), Cellular Component (CC), Molecular Function (MF), and other biological pathways (P-value <0.05).

Protein-Protein Interaction (PPI) network and module analysis of DEG-encoded proteins and PPIs are performed using the online database STRING (Version

11.0) (<http://STRING-db.org>) [15]. Interactions with a composite score exceeding 0.4 are deemed statistically significant. Cytoscape (version 3.7.1) is a freely available bioinformatics software tool utilized for displaying molecular interaction networks [16,17]. The Molecular Complex Detection (MCODE) plugin in Cytoscape (version 1.5.1) is used to cluster and identify densely connected regions within a network by analyzing its topology. MCODE criteria include an MCODE score ≥ 10 , degree cut-off = 2, node score cut-off = 0.2, maximum depth = 100, and K score = 2 [18]. Biological Networks Gene Ontology tool (BiNGO) (version 3.0.3) is employed to analyze and visualize the GO enrichment of nuclear genes [19].

Hub genes selection and analysis

We employed Cytoscape's plugin cytoHubba (version 0.1) to select and analyze the hub genes [20]. Hub genes with a Density of Maximum Neighborhood Component (DMNC) score greater than 0.95 were identified for further analysis of their co-expressed genes via the STRING online platform. The STRING interactome was analyzed with a confidence score cutoff of 0.900. We collected tissue-specific PPI data from DifferentialNet (<http://netbio.bgu.ac.il/diffnet/>), which delineates the differential PPIs across human tissues [21].

ChIP-seq data of transcription factors or chromatin regulators, along with differential gene expression data, were integrated using the Binding and Expression Target Analysis (BETA) software (<http://cistrome.org>), inferring target genes. Three key functions were designed in BETA: inferring the factor's target genes, predicting whether the factor exerts an activating or repressive role, and identifying the motif of the factor along with its collaborators, which may influence the factor's regulatory function. Target genes and transcription factors were obtained from ENCODE ChIP-seq data. Using the BETA Minus algorithm, we only considered predicted regulatory potential scores less than 1 and peak intensity signals less than 500 [22].

Analysis of transcription factor-gene (TF-gene) interactions was conducted using the BETA online platform. We excluded genes already proven to be hub genes to identify new hub genes. The cloud-based application WebMeV (<http://mev.tm4.org/>) was utilized to analyze, visualize, and categorize extensive genomic data, with a focus on microarray data and RNASeq. The hierarchical clustering of the newly identified hub genes was constructed using WebMEV [23].

Annotation of TNNT1

Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) is a freely accessible scientific tool that allows users to discover data on gene and protein expression. It provides detailed information regarding the levels and distribution of RNA and proteins in various species under different biological circumstances, including cell types, tissues,

diseases, and developmental phases. Expression Atlas was used to investigate the distribution of TNNT1 in human tissues [24]. The Human Protein Atlas (<https://www.proteinatlas.org/>) is a freely accessible database that integrates diverse omics data, including systems biology, transcriptomics, mass spectrometry-based proteomics and antibody-based imaging, mapping all human protein biology in organs, tissues and cells [25]. We applied the Human Protein Atlas to depict the atlas of TNNT1. Reatcome was used to construct the molecular mechanism model of TNNT1 transcripts in myocardial tissue. Integrating the alterations of TNNT1 expression in HCM myocardial tissue samples with the previous studies of TNNT1, we analyzed the possible mechanism of TNNT1 action in HCM. The analysis suggested the prospects of TNNT1 biomarkers in the screening, diagnosis and treatment of HCM.

RESULTS

Identification of DEGs in HCM

Three micro array datasets, GSE36961, GSE32453, and GSE1145, were downloaded from the Gene Expression Omnibus database. The GSE36961 dataset contained 106 HCM samples and 39 healthy samples. The GSE32453 contained 8 HCM and 5 healthy samples and the GSE1145 contained 5 HCM and 11 healthy samples. The probes were converted to the corresponding gene symbols based on the platform's annotation. The limma software package of R language was used to extract gene expression values from each myocardial sample. Box plots were generated to illustrate the distribution of raw read counts (fig. 1A). A total of 119 HCM myocardial tissues and 55 healthy myocardial tissues were subjected to gene expression analysis, resulting in the extraction of gene expression values for 665,0891 genes.

The sum of reading counts from all features of each sample was plotted (fig. 1B). PCA was used to transform high-dimensional data into lower dimensions (fig. 1C). Density plots against log₂ of reading counts were used to display the distribution of counts in each group (fig. 1D). Diagnostic plots summarized the standard deviation versus mean measures of reads for each gene (fig. 1E). Heatmap visualization analyzed RNA expression in three groups of HCM and healthy samples (fig. 2A). Volcano plots showed differentially expressed genes identified in HCM versus healthy groups, with the light red dots indicating significantly upregulated genes and the light blue indicating significantly down regulated genes (fig. 2B). Venn diagram software identified common DEGs across the three datasets, detecting a total of 316 DEGs, including 141 down regulated and 175 upregulated genes in HCM samples (fig. 2C).

GO enrichment analyses of DEGs in HCM

GO enrichment analysis of 316 DEGs combined with 19 confirmed pathogenic genes was visualized (fig. 3A). The

results showed that significantly enriched BP terms of analyzed genes were mainly associated with anatomical structure morphogenesis, muscle filament sliding, heart process, and regulation of hydrolase activity. The significantly enriched CC terms were primarily linked to the sarcomere, myofibril and I-band cytoplasm. The significantly enriched MF terms of DEGs were associated with cytoskeletal protein binding, protein binding, and actin-binding structural constituent of muscle (fig. 3B).

The most significant module was identified based on the PPI network with 16 nodes and 110 edges (fig. 3C). Light red items represented the upregulated genes, bright blue items represented the down regulated genes, and light green items represented the identified pathogenic genes.

Construction of PPI network and analysis of modules

The PPI network of DEGs was constructed and the pathogenic genes were subsequently confirmed (fig. 3C). The most significant module was identified by MCODE (fig. 4A). Furthermore, the performance and visualization of GO enrichment analysis of the most significant module were conducted by utilizing the BINGO plugin of Cytoscape (fig. 4B-4D).

Selection and analysis of Hub genes

The DMNC score of 16 hub genes is displayed in fig. 4E. We identified 13 genes with DMNC scores exceeding 0.95. Subsequently, these 13 genes - MYOZ2, TNNT3, TNNC1, TNNT2, TPM1, MYBPC3, TNNT1, ACTC1, ACTN2, MYH6, MYL3, MYH7 and MYL2 were subjected to further analysis.

Under non-tissue-specific conditions, the enrichment results of GO, KEGG and Reactome pathways were analyzed using the STRING online platform. The visual interaction network of these genes, along with their co-expressed genes, was constructed and examined (fig. 5A). The degree centrality and betweenness centrality of the interaction network are shown in fig. 5B. In the co-expression network, genes such as ACTN2, TPM1, MYL2, MYH6, MYL3, TNNT3, MYBPC3, TNNT2, TNNT1, TNNC1, ACTC1, MYH7, MYL1, TCAP, MYL4, TPM4, TPM2, MYH8 and TPM3 exhibited a degree greater than 10. The visualization of the results is shown in the bubble plot, where the betweenness centrality score is positively proportional to the bubble size (fig. 5C). The enrichment results of GO, KEGG, and Reactome pathways are illustrated in fig. 5D. The results of GO analysis indicated that significantly enriched BP terms were related to actin filament-based movement, wound healing, blood coagulation, and coagulation. The significantly enriched CC terms were mainly associated with contractile fibers, myofibrils, actin cytoskeleton, and sarcomeres. The significantly enriched MF terms were primarily associated with cytoskeletal protein binding, actin binding, and calmodulin binding.

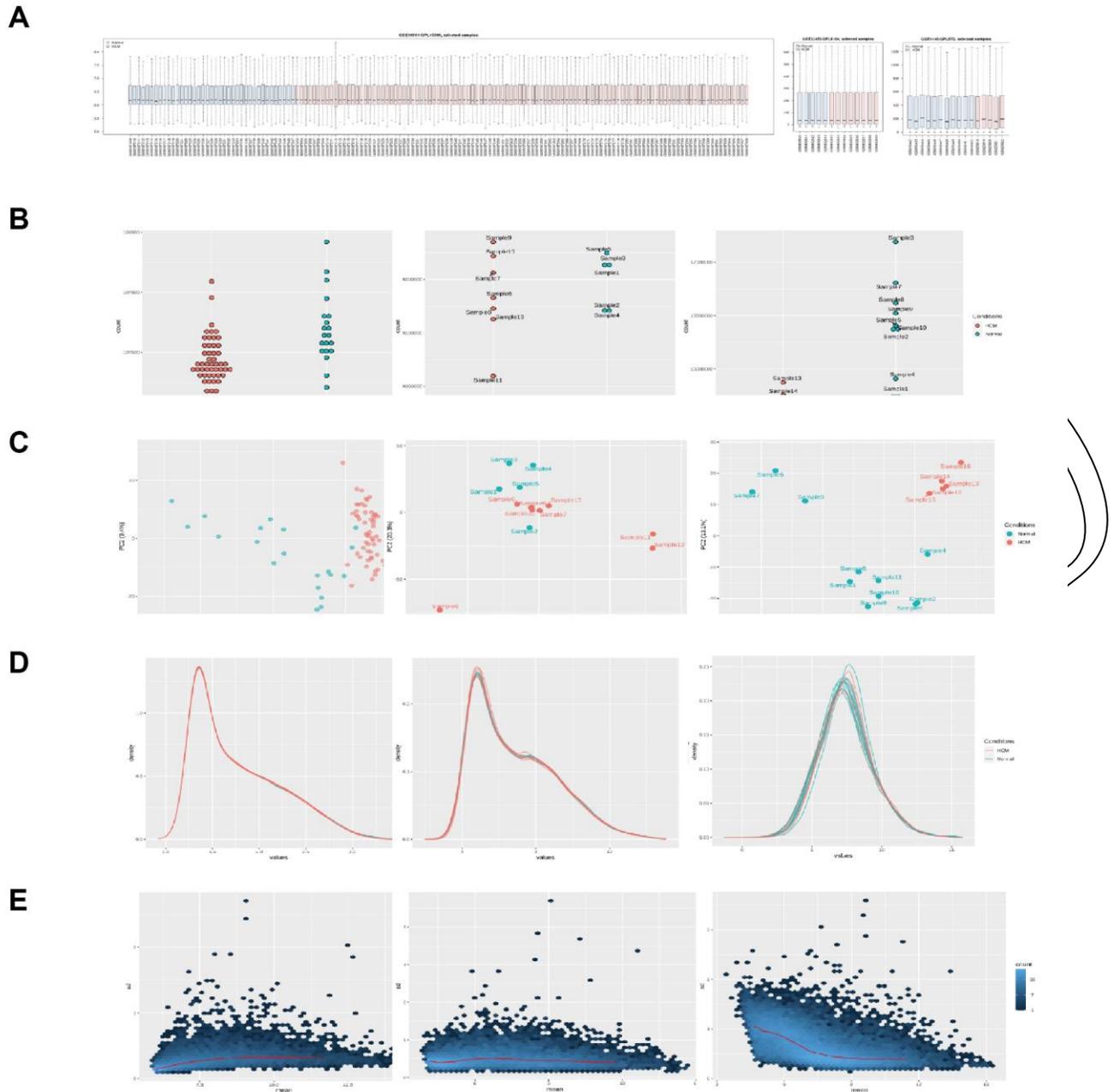


Fig. 1: Overall information of three microarray datasets, GSE36961, GSE32453 and GSE1145, was visualized respectively. (A) Distribution of raw read counts. (B) The sum of the reading counts for all features of each sample in GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets. (C) Principal component analysis (PCA) using all identified genes for normal group (blue) and HCM group (red). The PCA plots of GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets were presented respectively. (D) The distribution curves of estimated insert size for each of GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets. (E) Diagnostic plots of the relationship between standard deviation and mean value of reads in each gene sample. Diagnostic plots of GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets were presented respectively.

KEGG pathway analysis revealed enrichment in focal adhesion, regulation of actin cytoskeleton, HCM, dilated cardiomyopathy and leukocyte transendothelial migration. Reactome pathway analysis indicated enrichment in

muscle contraction, striated muscle contraction, platelet degranulation and platelet activation, aggregation and signaling (fig. 5E).

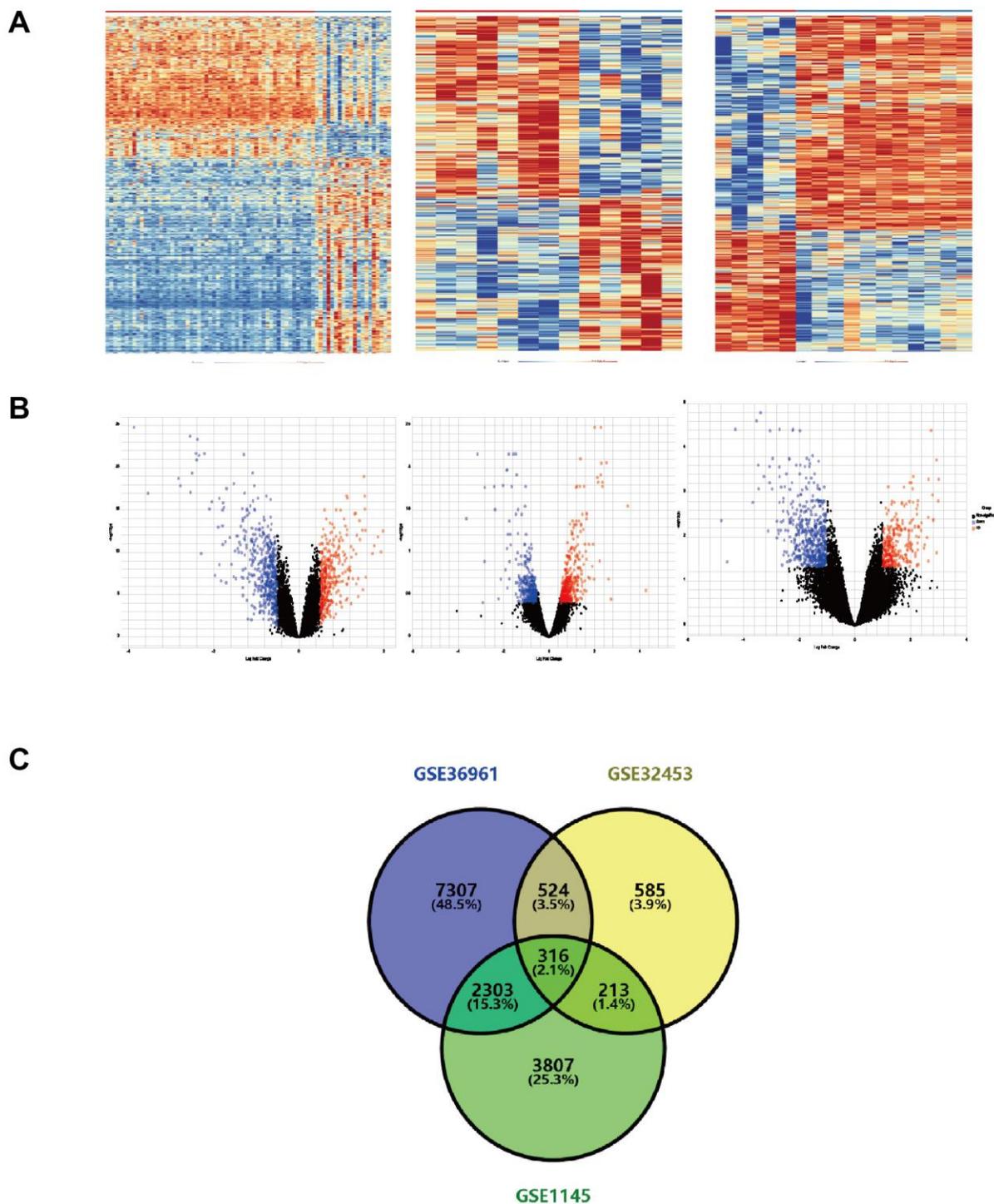
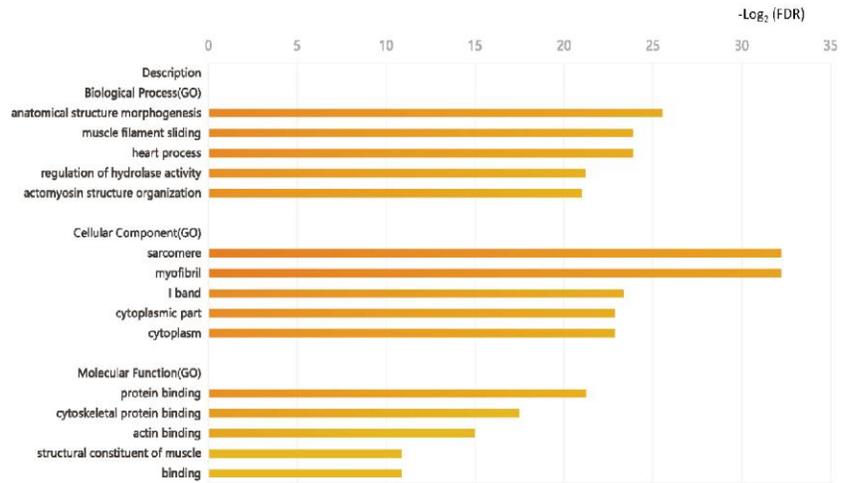


Fig. 2: The expression profiles of the coding genes of GSE36961, GSE32453 and GSE1145 were shown respectively. (A) Heatmaps of the expression levels of coding genes expressed in HCM and healthy group from GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets. Each row represents a gene, and each column represents a sample. (B) Volcano plots illustrating differentially expressed genes between HCM and healthy group from GSE36961 (left panel), GSE32453 (middle panel) and GSE1145 (right panel) datasets. The red dots denote significantly upregulated gene expression, and the blue dots denote significantly down regulated gene expression. The point size reflects absolute fold change. (C) The Venn diagram of DEGs shows the selected DEGs in the gene expression profiles of three datasets, GSE36961, GSE32453 and GSE1145. Three datasets show 316 gene overlaps.

A

Description	False Discovery Rate (FDR)
Biological Process (BP)	
anatomical structure morphogenesis	2.08E-08
muscle filament sliding	0.00000065
heart process	0.00000065
regulation of hydrolase activity	0.000000417
actomyosin structure organization	0.00000048
Cellular Component (CC)	
sarcomere	2.01E-10
myofibril	2.01E-10
I band	9.48E-08
cytoplasmic part	1.32E-07
cytoplasm	1.32E-07
Molecular Function (MF)	
protein binding	0.000000412
cytoskeletal protein binding	0.00000547
actin binding	0.00000599
structural constituent of muscle	0.00054
binding	0.00054

B



C

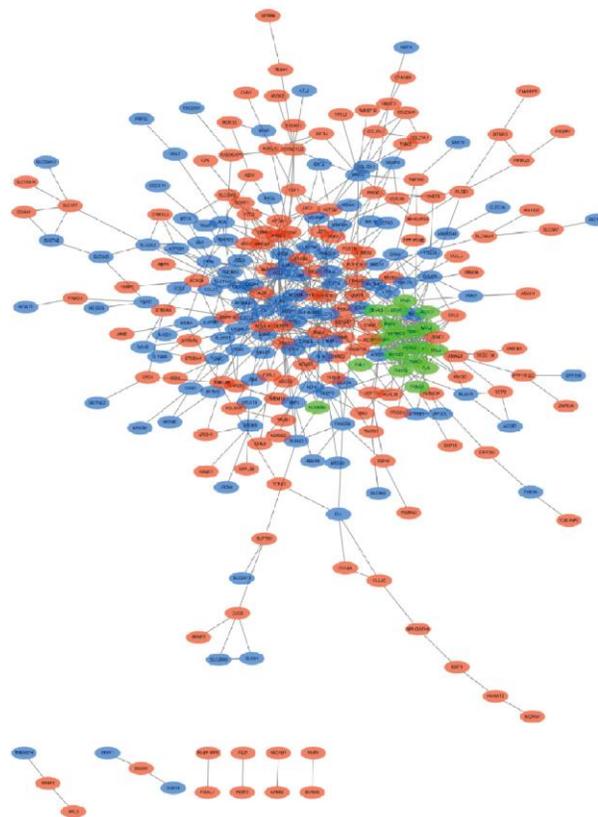


Fig. 3: Differential genes were subject to GO enrichment analysis and network visualization analysis. (A) GO enrichment analysis of 316 DEGs combined with 19 confirmed pathogenic genes. (B) Visualization of GO enrichment analysis of 316 DEGs combined with 19 confirmed pathogenic genes. (C) The PPI network of DEGs constructed using Cytoscape. The most significant module obtained from the PPI network with 16 nodes and 110 sides was shown. Light red items represented the upregulated genes, bright blue items represented the down regulated genes, and light green items represented the identified pathogenic genes.

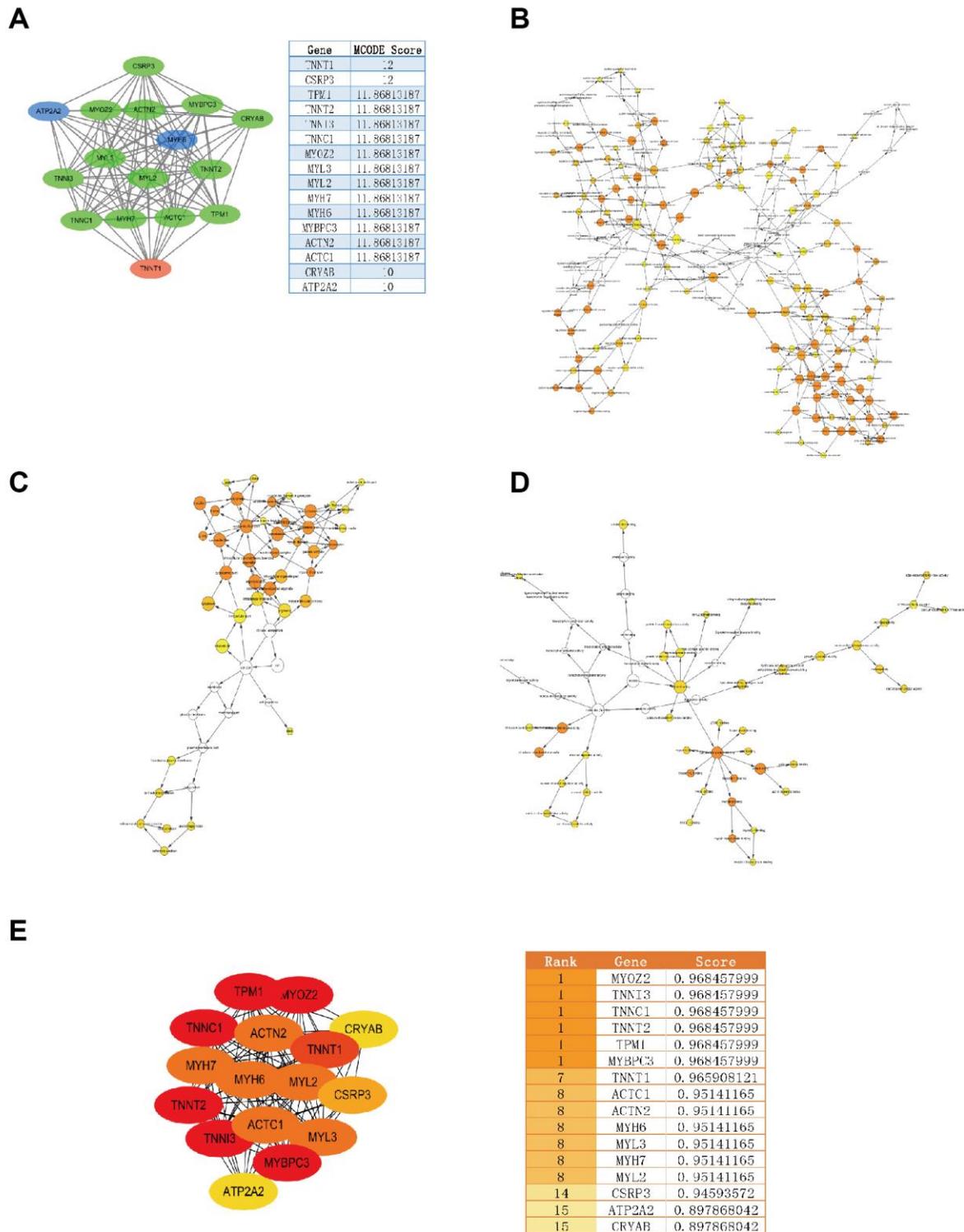


Fig. 4: Identification of the most significant module and analysis of its gene network. (A) Using MCODE to identify the PPI network with 16 nodes and 110 sides as the essential module (left panel). The MCODE score of genes from the most significant module is listed (right panel). (B) The biological process analysis of the genes involved in the most significant module constructed using BiNGO. The color depth of the node refers to the adjusted p-value of the ontology. The size of the node refers to genes related to ontology. P<0.01 is considered statistically significant. (C) The cellular component analysis of the genes involved in the most significant module constructed using BiNGO. (D) The molecular function analysis of the genes involved in the most significant module constructed using BiNGO. (E) Screening out the 13 most crucial hub genes using the Cytoscape software plugin cytoHubba.

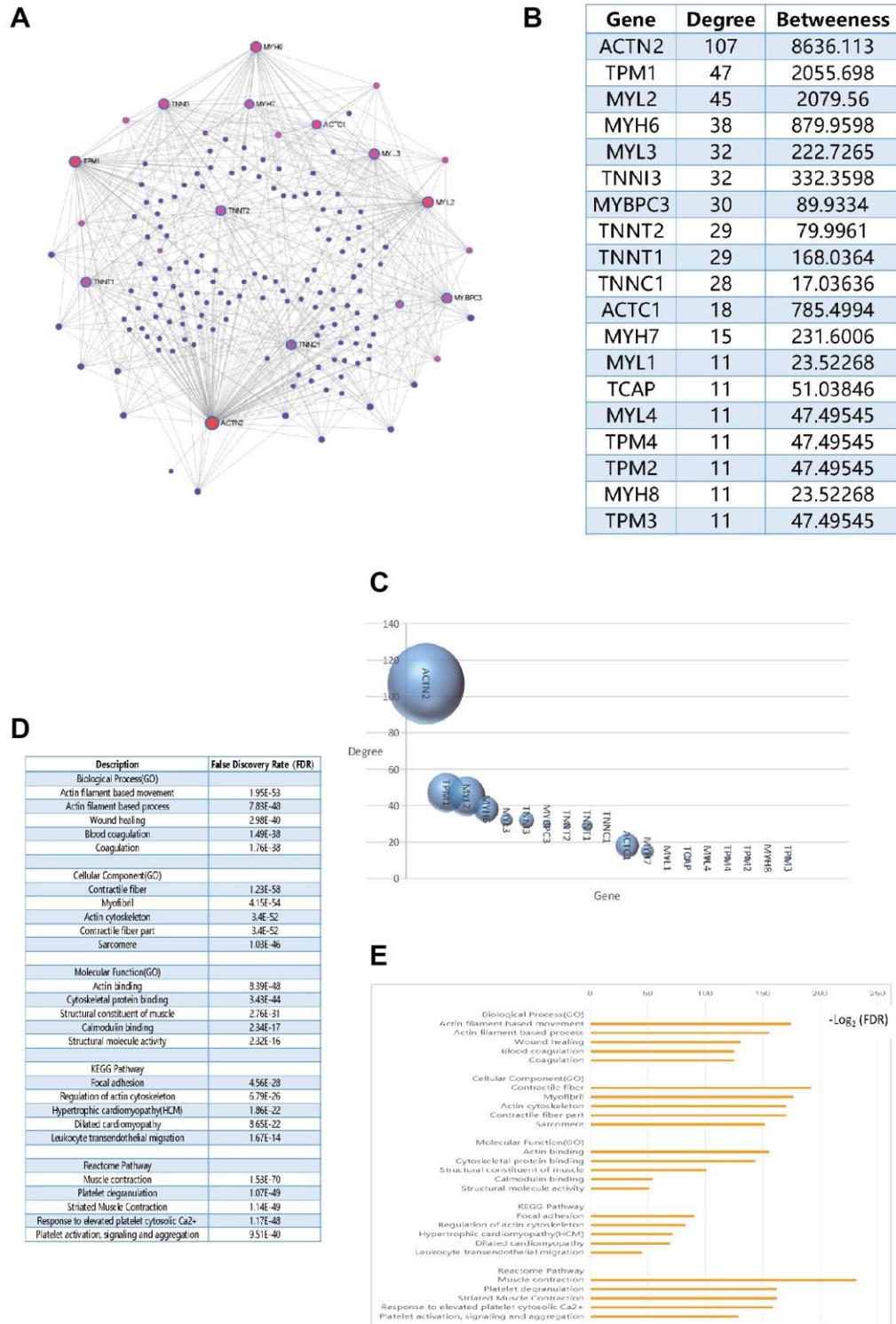


Fig. 5: The 13 most critical hub genes were subjected to co-expression network and enrichment analysis. (A) The hub genes and its co-expressed genes analyzed by STRING. The nodes in the bold blue outline represent the hub genes. Purple nodes represent co-expressed genes. (B) Genes with node degree >10 and their betweenness centrality scores in the co-expression network of fig. 5A were listed. (C) The bubble chart showing the degree centrality score and betweenness centrality score of genes with node degree >10 in the co-expression network of fig. 5A. The abscissa represents a gene, the ordinate represents degree centrality score and bubble size represents the betweenness centrality score. (D) The enrichment analysis of GO, KEGG pathway and Reactome pathway of the 13 hub genes. (E) FDR of the enrichment after negative logarithm transformation.

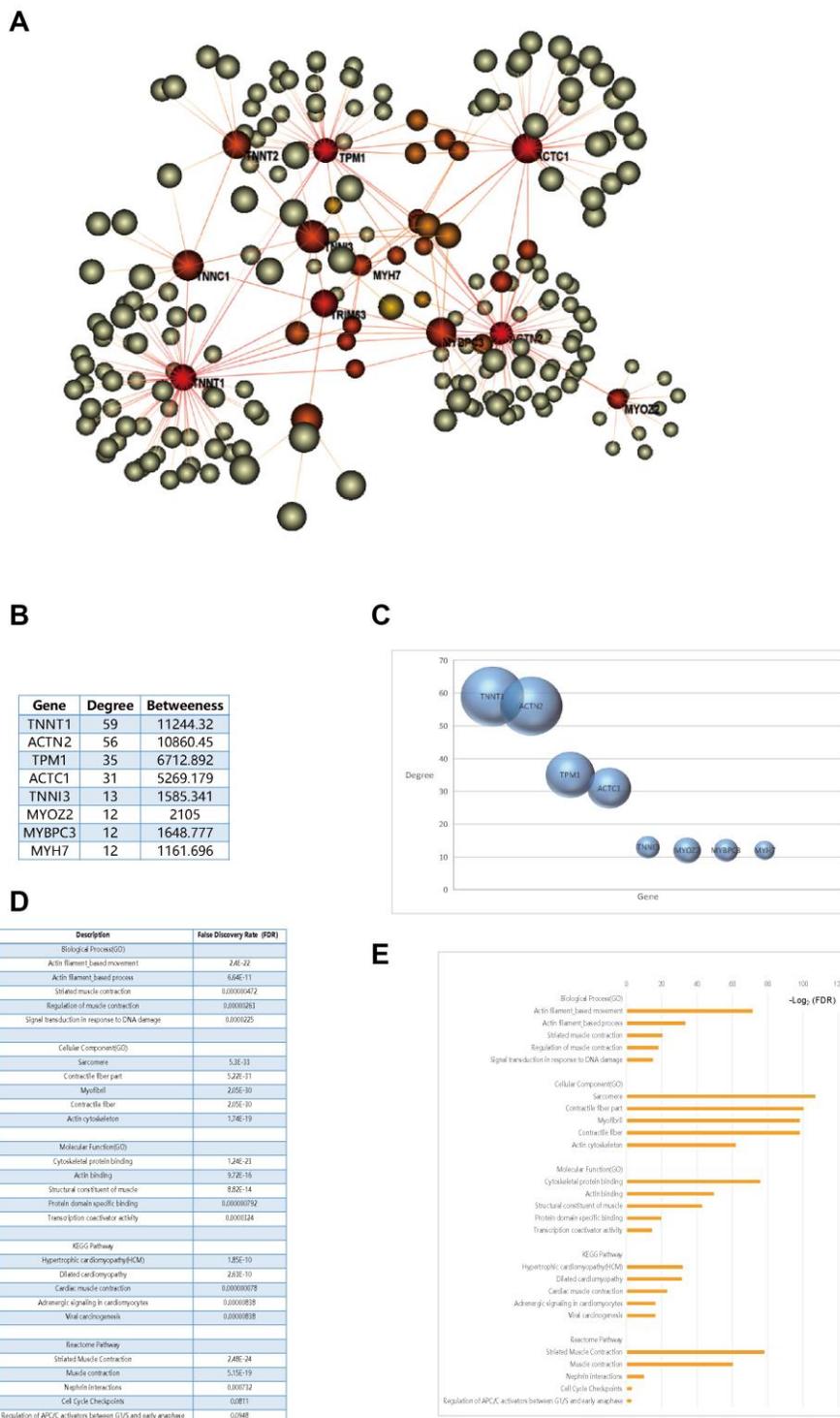


Fig. 6: Tissue-specific (left ventricle) gene co-expression network and enrichment analysis of the 13 most critical hub genes are shown. (A) The 13 hub genes and their co-expression genes analysis based on the Differential NET database. Each sphere represents a gene. The red sphere represents the nuclear gene, and the yellow-brown sphere represents the co-expression gene. (B) List of the genes with node degree >10 and their betweenness centrality scores in the co-expression network of fig. 6A. (C) The bubble chart showing the degree centrality score and betweenness centrality score of genes with node degree >10 in fig. 6A. The abscissa represents a gene, the ordinate represents degree centrality score, and bubble size represents the betweenness centrality score. (D) The tissue-specific enrichment analysis of GO, KEGG pathway and Reactome pathway of the 13 hub genes. (E) FDR of the tissue-specific enrichment after negative logarithm transformation.

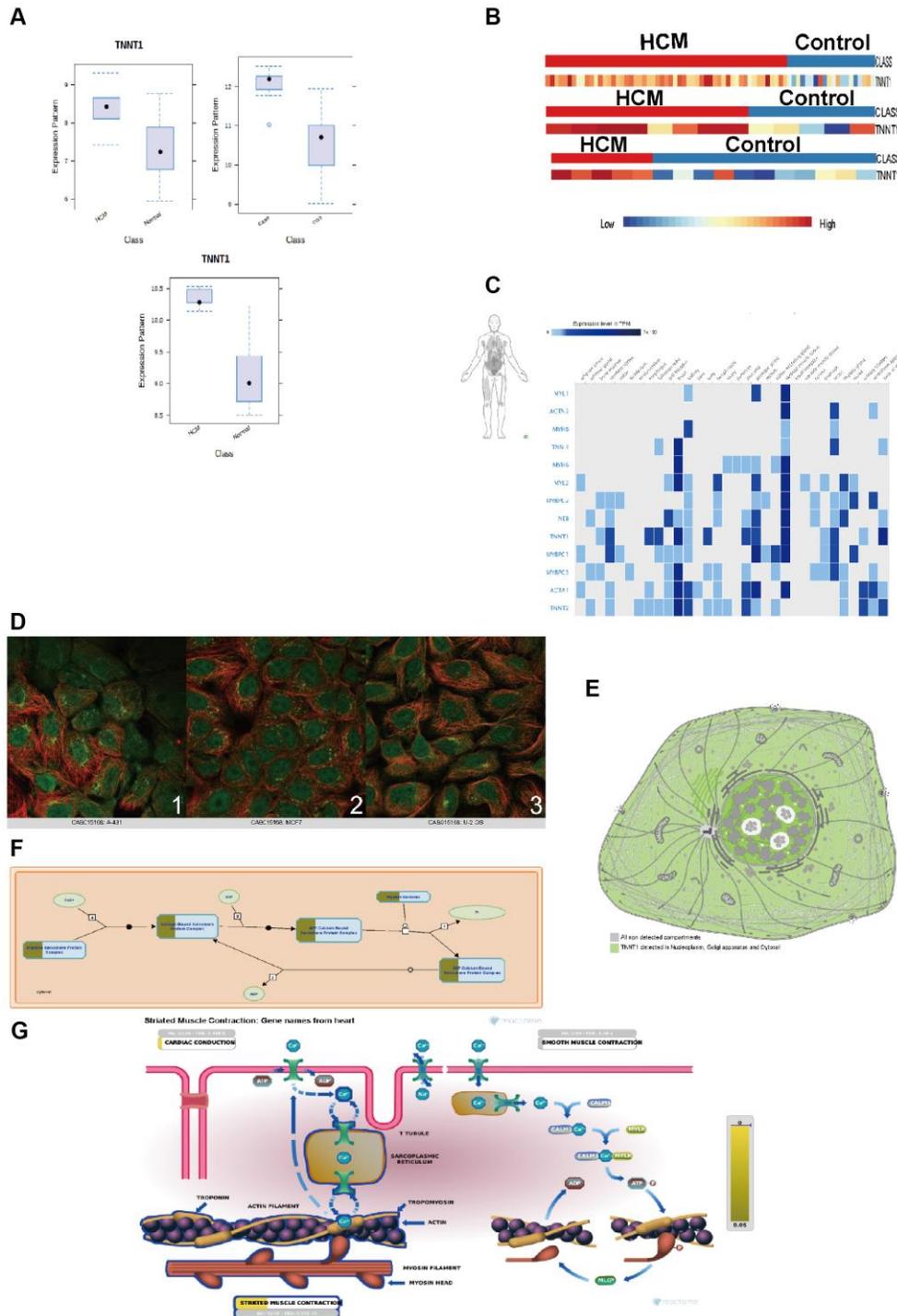


Fig. 8: Information of TNNT1 was further annotated. (A) The expression level of TNNT1 in HCM and normal myocardium. (B) Relative expression of TNNT1 in HCM and normal group from GSE36961, GSE32453, and GSE1145 datasets shown in heatmaps. Expression level for TNNT1 is displayed by a color range from blue (low) to orange (high). (C) The expression profile of the hub genes and its co-expression gene in human tissue based on The Human Protein Atlas (only partial genes shown). (D) Immunofluorescence staining showing that TNNT1 transcripts localized in nucleoplasm, cytoplasm and Golgi body. (E) The TNNT1 expression pattern in cells according to immunofluorescence staining. Light green region indicates the expression site of TNNT1. (F) Schematic diagram showing the role of TNNT1 products in myocardial contraction and the molecular model of energy conversion. (G) Mechanism diagram of TNNT1 products participating in the biological model of myocardial contraction. Blue bold region indicates the site of TNNT1 products involved in the process of myocardial contraction.

Moreover, the Differential NET database was utilized to analyze the co-expression network of hub genes in the left ventricular myocardium (fig. 6A). The degree centrality and betweenness centrality of the interaction network are shown in fig. 6B. In the co-expression network, genes such as TNNT1, ACTN2, TPM1, ACTC1, TNNT3, MYOZ2, MYBPC3 and MYH7 with a degree greater than 10, were represented in the bubble plot where the betweenness centrality score is positively proportional to the bubble size (fig. 6C). The enrichment results of GO, KEGG and Reactome pathways are presented in fig. 6D. GO analysis showed that significantly enriched BP terms of analyzed genes were related to actin filament-based movement, striated muscle contraction and signal transduction in response to DNA damage. The significantly enriched CC terms were primarily linked to sarcomeres, myofibrils, and the actin cytoskeleton. The significantly enriched MF terms were associated with transcription coactivator activity, actin binding and cytoskeletal protein binding. KEGG pathway analysis indicated enrichment in HCM, adrenergic signaling in cardiomyocytes, cardiac muscle contraction, dilated cardiomyopathy and viral carcinogenesis. Reactome pathway analysis revealed enrichment in striated muscle contraction, muscle contraction, cell cycle checkpoints, and regulation of APC/C activators between G1/S and early anaphase (fig. 6E).

The Gene Regulatory Network demonstrated a broad regulatory relationship between hub genes in TF-gene interaction (fig. 7A). The degree centrality (node degree >10) and betweenness centrality score of each gene were shown in fig. 7B. Genes such as TNNT3, TNNT1, TPM1, MYBPC3 and TNNT1, with a degree greater than 10, are represented in the bubble plot, where the betweenness centrality score is positively proportional to the bubble size (fig. 7C). The enrichment results of GO, KEGG, and Reactome pathways are illustrated in fig. 7D. GO analysis indicated that significantly enriched BP terms of analyzed genes were related to the regulation of transcription DNA-dependent and RNA metabolic processes. The significantly enriched CC terms were mainly associated with the nucleus, nuclear chromatin and nucleoplasm. The significantly enriched MF terms were primarily associated with DNA binding, sequence-specific DNA binding, and transcription regulation. KEGG pathway analysis revealed enrichment in transcriptional misregulation in cancer, adrenergic signaling in cardiomyocytes, Huntington's disease, HCM and dilated cardiomyopathy. Reactome pathway analysis indicated enrichment in striated muscle contraction, generic transcription pathway, and transcriptional activity of the SMAD2/SMAD3: SMAD4 heterotrimer (fig. 7E).

Identification of new genes

Among the 13 genes identified in the most significant module, TNNT1 was notable since it has not been

previously reported in association with HCM. We analyzed the significance of TNNT1 within this key module. In the MCODE analysis, TNNT1 exhibited a node degree of 12. Its degree centrality and betweenness centrality scores in non-tissue-specific co-expression networks were 29 and 168.0364, respectively. In the gene co-expression network specific to left ventricular tissue, TNNT1 demonstrated the highest scores in the network, with a degree centrality of 59 and a betweenness centrality of 11244.32. In the network of transcription factor gene interactions, the degree centrality and betweenness centrality scores of TNNT1 were 57 and 2912.306, respectively, ranking it second only to TNNT3.

Furthermore, we observed that TNNT1 expression in HCM patients was significantly higher compared to that in donors (fig. 8A). Hierarchical clustering analysis conducted using WebMEX indicated that TNNT1 expression could differentiate HCM samples from healthy samples (fig. 8B). Tissue-specific enrichment analysis revealed that TNNT1 is predominantly expressed in myocardial and skeletal muscle tissues (fig. 8C). The presence of TNNT1 transcripts in the human cell line MCF7 was confirmed through immunofluorescence staining (fig. 8D), showing localization in the nucleoplasm, cytoplasm and Golgi apparatus (fig. 8E). TNNT1 encodes troponin T type 1, which forms a part of the troponin complex and is situated on sarcomere filaments (fig. 8F). This complex is crucial in regulating striated muscle contraction in response to intracellular calcium levels (fig. 8G).

All of these findings collectively suggested that TNNT1 may play a significant role in the development and progression of HCM.

DISCUSSION

HCM is a prevalent genetic heart condition worldwide. It is widely acknowledged that a comprehensive and systematic early screening program for HCM remains to be developed. As we all know, there is still no established early screening program for HCM. The efficacy of pharmacologic interventions remains uncertain, and the surgical intervention continues to be the only effective treatment [26]. Since the initial report on HCM decades ago, the research on the formation and progression of HCM have been extensively developed, including the most recent multi-omics studies. However, the incidence of sudden cardiac death caused by HCM has not decreased, and the positive rate of gene screening has remained low [27,28]. As is the case with other genetic diseases, HCM is characterized by abnormal gene and protein expression [29]. A substantial body of research has demonstrated that HCM results from the accumulation of cellular and molecular aberrations, including epigenetics, transcriptome, micro RNA,

proteomics, and metabolomics [30]. The primary purpose of multi-omics research is to identify biomarkers capable of diagnosing HCM in its early stages. It aims not only to identify the heterogeneity of HCM but also to discover the potential molecular commonalities at various points in time. HCM exhibits notable heterogeneity at the molecular level, involving many protein-level and genetic changes. Consequently, this disease can be represented by a carefully selected set of candidates [31,32].

With the advancements in bioinformatics, numerous molecular markers have been discovered for HCM, which have the potential to serve as diagnostic and prognostic indicators. However, many of these markers have not yet been independently validated. Furthermore, there has been a lack of comparative analysis among these biomarkers to pinpoint candidates for further research and screening. While the molecular characteristics of HCM have the potential to serve as indicators for detecting and tracking disease progression in an early stage, a comprehensive understanding of these characteristics remains to be studied.

Diverging from traditional single genetic or cohort studies, we undertook a comprehensive analysis of three micro array data sets encompassing 115 HCM and 55 normal myocardial tissue samples. A total of 316 differential genes in HCM were identified in our study, including 175 upregulated genes and 141 down regulated genes. The GO analysis of DEGs, performed using DAVID, revealed that HCM-regulated genes were primarily involved in processes such as anatomical structure morphogenesis, muscle filament sliding, cardiac processes, regulation of hydrolase activity and actomyosin structure organization.

We further conducted pathway enrichment analysis using the KEGG database to assess the functional significance of these DEGs and to identify pathogenic genes. We employed the STRING database and Cytoscape software to construct a visual gene interaction network, which comprised 334 nodes and 682 edges. Following MCODE modeling, the most significant module was identified, characterized by 16 nodes (16 genes), 110 edges, an average node degree of 13.8, an average local clustering coefficient of 0.939, an expected edge number of 2, and a PPI enrichment P-value of less than 1×10^{-16} . This gene cluster is a promising target for the screening and early diagnosis of HCM.

To identify the hub genes within this significant module, we initially screened for genes with a node degree greater than 10. Subsequently, Cytoscape's plugin Hubba was employed to filter these genes, and those with DMNC scores greater than 0.95 were designated as hub genes for this module. Ultimately, thirteen genes were identified through this process.

We analyzed three interaction networks of hub genes: a non-tissue-specific PPI network, a left ventricular-specific PPI network, and a TF-gene interaction network. The enrichment results from the STRING and DifferentialNet databases were compared, revealing several similarities. These included two BP terms, the top five CC terms, and three MF terms, which were actin binding, cytoskeletal protein binding and structural constituent of muscle. Two enriched KEGG pathways were HCM and Dilated Cardiomyopathy, and two enriched Reactome pathways were Striated Muscle Contraction and Muscle Contraction.

BP pathway enrichment in the left ventricular myocardium indicated that proteins encoded by hub genes were involved in regulation of muscle contraction, striated muscle contraction, and signal transduction in response to DNA damage. MF pathway enrichment analysis revealed that hub genes also participated in protein domain-specific binding and transcription coactivator activity in the left ventricular myocardium. These results highlighted the diverse functions of hub gene transcripts, underscoring the functional diversity resulting from differential gene expression.

Transcription factors are critical regulators of gene expression, playing pivotal roles in almost all cellular regulatory processes, including differentiation, proliferation, survival and apoptosis. Moreover, genetic diseases and complex conditions such as HCM are often associated with transcription factor regulation, involving feed-forward loops and other network motifs that regulate cellular transcription in an interconnected manner. Hence, utilizing regulatory data on transcription factors and hub genes can facilitate the elucidation of key driver genes in HCM, potentially leading to novel therapeutic strategies for its treatment.

A gene regulatory network was constructed to investigate the factors driving differential expression. Within the TF-gene regulatory network, five genes with a node degree greater than 10 were identified: TNNI3, TNNT1, TPM1, MYBPC3 and TNNC1. The differential expression of these genes during transcription is likely responsible for the various clinical manifestations observed in HCM. Additionally, the diversity of clinical phenotypes in HCM, despite similar genetic backgrounds, is likely attributable to differences in the transcription of these genes.

Enrichment analysis of BP pathways suggested that mutations in hub genes could lead to abnormal regulation of nucleobase-containing compound metabolic processes, subsequently affecting DNA translation, RNA transcription, and protein biosynthesis. Targeting these specific BP pathways in drug development may help

manage the incidence of pathogenic mutations in HCM carriers. Enrichment analysis of CC pathways indicated that HCM-related variations predominantly occur in the nucleus, suggesting that focusing on RNA variation in the nucleus could be crucial in HCM etiology studies. Enrichment analysis of MF pathways revealed that transcription factors closely associated with hub genes primarily involved DNA binding and RNA polymerase II promoter activity. Further investigation of these MF pathways could elucidate the molecular mechanisms underlying HCM. KEGG pathway analysis indicated that dysregulation may impact adrenergic signal transduction in cardiomyocytes. The Reactome pathway analysis suggested a possible relation of HCM to the SMAD4 heterotrimer.

After excluding the currently known pathogenic genes, we identified a novel gene, TNNT1, that is closely associated with hypertrophic cardiomyopathy (HCM) in the core module of our analysis. In the non-tissue-specific PPI network, TNNT1 exhibited a centrality score of 29 and a betweenness score of 168.0364. Notably, in the left ventricular tissue-specific PPI network, TNNT1 emerged as the most critical gene, with a centrality score of 59 and a betweenness score of 1244.32. These results indicate that TNNT1 has a significant impact on HCM, potentially aiding in the screening and early diagnosis of the disease.

As a subunit of the troponin complex, troponin T (TnT) along with tropomyosin, troponin I (TnI), and troponin C (TnC), attaches the troponin complex to actin filaments. This is essential for controlling calcium (Ca^{2+}) levels during muscle contraction [33]. This precise interaction between Ca^{2+} and the thin filament regulatory complex is vital for coordinating the normal systolic rhythm essential for cardiac function. Research indicates that the composition of isoforms within the myofilament regulatory complex undergoes modifications during the heart development and in response to heart disease [34].

Vertebrates possess three distinct TnT genes: slow skeletal muscle TnT (ssTnT), fast skeletal muscle TnT, and cardiac TnT. Each of these three genes encodes specific isoforms with unique expression patterns and functional roles. The TNNT1 gene is responsible for encoding muscle-specific TnI and TnT subtypes, which are arranged in three tandem pairs within the invertebrate genome. Notably, TNNT1 expression is subject to alteration during mammalian cardiac development [35,36]. Whole-mount in situ hybridization (WISH) studies utilizing anti-TNNT1 RNA probes have shown critical TNNT1 expression in the cardiac outflow tract and the developing interventricular groove of mouse embryos at embryonic day 9.5 (E9.5). The expression of TNNT1 extends to the cardiac apex at E10.5, notably within the left ventricle. Nevertheless, expression diminishes in the outflow tract at E11.5, becoming

predominantly localized to the ventral side of the interventricular groove by E12, with diminished presence in the left ventricle and complete absence in outflow tract cells from approximately E12.0. At E12.5, the expression of TNNT1 in the interventricular septum expands laterally until E14.5, with notable expression in the ventral aspect of the left ventricle and enhanced expression in the apical region of the right ventricle, while cells in the outflow tract remain negative. Furthermore, at E12.5, the expression of TNNT1 is prominent in the atrium, especially in the left atrium.

Previous research has investigated the function of TNNT1 in mammalian development and cardiogenesis through targeted disruption of its cardiac homolog TNNT2, revealing that loss of TNNT1 results in embryonic lethality by approximately E10, typically with absent cardiac activity. Nevertheless, in the outflow tract of all E10 and approximately half of E9 TNNT2 $-/-$ (cTnT $-/-$) embryos, pulsatile cells were observed, underscoring the critical role of TNNT1 expression in this region for normal cardiac morphology and function [36]. These findings, combined with our network analysis, suggest that myocardial hypertrophy in HCM, particularly the level of blockage in the left ventricular outflow tract in individuals with hypertrophic obstructive cardiomyopathy could potentially be associated with TNNT1 over expression.

The cardiac Troponin I (cTnI, TNNT3) and slow skeletal muscle Troponin T (ssTnT, TNNT1) genes are closely linked. In vitro promoter analyses have revealed that the KB-2.2 to -5.4 upstream region of the TNNT1 gene overlaps with the 3' region of the TNNT3 gene, encompassing the primary enhancer activity essential for high-level transcription of TNNT1 [37]. A study involving genetically engineered mouse lines demonstrated that targeted deletion in embryonic stem cells not only excises the entire TNNT3 gene but also a significant part of the 5' enhancer region of TNNT1, confirming the genomic integration of the cTnI and ssTnT genes.

Further assessments of troponin I isoform gene expression, transfer, and function in adult cardiac myocytes transduced with vectors containing rat cTnI or ssTnI expression cassettes have been conducted. By the sixth day post-gene transfer, the substitution rate of endogenous cTnI by ssTnI exceeded 90% and there are no detectable changes in other sarcomere protein expression patterns or chemometric properties. Moreover, neither the expression levels of ssTnI nor cTnI had any effect on the total troponin I content, indicating a stoichiometric of TnI in the sarcomere was replaced. The epitope-tagged TnI confirmed the efficiency and precise subcellular localization of the newly expressed TnI within the cardiac myocytes, with over 95% of cardiomyocytes

demonstrating specific localization to the myocardial filaments.

At lower replacement levels, the Ca^{2+} sensitivity of tension significantly increased ($P < 0.01$), underscoring the dominant influence of ssTnI in regulating myofilament contraction. SDS-PAGE analysis indicated that the ratio of ssTnI to cTnI was 27.5% on the second day post-transduction, increasing to 67% by the fourth day. Additionally, testing with a second troponin I antibody (Fitzgerald) showed that ssTnI replaced approximately 29% of cTnI by the second day post-transduction. The use of two distinct troponin I antibodies (chemical and Fitzgerald) along with SDS-PAGE further substantiated that the range of TnI replacement varied between 14% and 29% on the second day, closely aligning with previous reports of a 90% replacement in terms of Ca^{2+} sensitivity of tension.

Intriguingly, it has been demonstrated that the transfer of ssTnT into adult cardiomyocytes induces two notable alterations in the mechanical function of Ca^{2+} activation. Firstly, the threshold for Ca^{2+} activation and the molecular synergy in these cardiomyocytes are diminished. Specifically, in myocardial cells expressing ssTnT, tension development is evident at PCA levels between 7.0 and 6.5, whereas, at similar Ca^{2+} levels, cardiomyocytes expressing cardiac TnI (cTnI) are fully relaxed. Secondly, ssTnT-expressing cardiomyocytes show significantly reduced desensitization to Ca^{2+} -activated tension under acidic pH conditions. Furthermore, no differences were observed in the standardized maximum Ca^{2+} activation tension between cardiomyocytes expressing ssTnT and those expressing cTnI, indicating that ssTnT gene transfer does not alter the peak mechanical properties of the myocardium. Notably, the expression of the ssTnT subtype is also crucial in reducing the pH sensitivity of the fetal rat myocardium [38,39]. These modifications align with the pathophysiological changes observed in HCM.

However, the mechanisms of transcriptional regulation across the TnIc-TnTs locus remain elusive. Since the discovery of the globin gene cluster, the notion of controlling gene expression through a unified set of regulatory elements has served as a paradigm in the study of gene regulation. Classical models include the concept of a single chromatin domain or gene clusters forming loops, wherein shared distal enhancers target multiple promoters, supplemented by flanking insulator elements [40,41]. The TnIc-TnTs region is closely aligned with differentially expressed genes, exhibiting mutually exclusive expression patterns among adjacent genes. Nonetheless, detailed information on transcriptional regulation in this region is scant, with only the promoter of the cTnI gene being thoroughly examined [42,43]. A notable characteristic of this region is its compact

architecture, which necessitates only proximal sequences to recapitulate many facets of endogenous gene expression. Studies on TnT genes in humans and mice have suggested that intergenic regions can drive expression in skeletal muscle cells in vitro. The inclusion of additional 5' sequences led to enhanced activity, hinting that elements regulating TnT expression in mice might also be embedded in the 3' end of the TnIc gene [44,45].

Despite these findings, it remains unclear which elements are essential for achieving the correct tissue specificity or developmental regulation required to drive substantial expression in cardiomyocytes (data not shown). Further studies have revealed that the human TnTs gene exhibits transient regional expression in the fetal heart and markedly reduced expression in the adult human heart, complicating the understanding of gene regulation in this area [46,47]. Therefore, introns or distal regulatory elements might be necessary for appropriate regulation. This finding challenges the traditional model previously outlined, suggesting that genes in this region may function as independent transcription units, which contain all necessary tissue-specific and developmental response elements within a compact proximal promoter, possibly alongside other elements situated within an intron.

A number of theories have been postulated, and research has indicated that the promoter region of the TNNT1 harbors numerous transcription factor binding elements common among genes expressed in myocardial tissues. Deletion analysis of the human gene promoter revealed that the sequence from -98 bp to +67 bp is sufficient to drive activity in myocardial cells. Animal studies further demonstrated that TNNT1 expression in the heart requires the GATA4-FOG2 transcription complex, corroborating our previous TF-gene interaction network analysis which linked the differential expression of TNNT1 in cardiomyocytes primarily to TNNT1 DNA binding and transcription initiation from RNA polymerase II promoters. It was also reported that the cis-regulatory element of the GATA4-FOG2 complex significantly influences heart-specific regulation of the TNNT1 gene. This finding was reached through the construction of a bacterial β -galactosidase (lacZ) fusion transgenic structure, revealing direct expression of TNNT1 regulatory elements in bone and myocardial tissues. Consequently, we hypothesize that the GATA4, FOG2, and the GATA4-FOG2 transcription complex may be pivotal targets for gene editing and targeted drug development in HCM.

In HCM, mutations or small molecules could potentially alter contractility via modifications to the Ca^{2+} sensitivity of myofilaments. However, subcellular biomarkers for myofilaments remain undeveloped. Although the link between TNNT1 and HCM is not fully elucidated, our

findings suggest that TNNT1 transcripts, which are restricted to myofilaments, reflect an increased expression pattern in HCM, indicative of the myofilaments' sensitivity to calcium. This finding suggests that TNNT1 transcripts could serve as a biomarker for the progression and severity of HCM, representing a promising target for future therapeutic interventions.

The objective of this study was to construct a gene network involving DEGs identified in both healthy and HCM tissues, initiating functional annotation of shared hub genes. However, this study is not without its limitations. Firstly, the sample size remains inadequate and can be increased in the future with the collection of additional samples. Secondly, while a variety of enrichment analyses and functional annotation were performed to elucidate the regulatory mechanisms of TNNT1 in HCM, these findings have yet to be confirmed in human myocardial tissues. Furthermore, although this study identified potential target gene TNNT1 based on three datasets containing a total of 115 HCM and 55 healthy samples, the specific molecular mechanism of its effect on HCM needs to be verified by more experiments in the future. Meanwhile, the correlation between TNNT1 and HCM needs to be supported by correlation analysis of more clinical data. In addition, given the absence of available inhibitors or activators for TNNT1, TNNT1-targeted therapies for HCM should be explored through the lens of its downstream mechanisms. Therefore, future studies should concentrate on detailing the interactions between TNNT1 and HCM to clarify these mechanisms further. The subsequent research plan will be carried out in the following manner. The first step will be the detection of the frequency of TNNT1 mutation in the HCM population and the analysis of the clinical relevance between TNNT1 and HCM based on the collection of clinical data. Secondly, we will construct TNNT1-overexpressed induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) to detect the expression changes of cardiac hypertrophy-related genes. Thirdly, mice with cardiomyocyte-specific TNNT1 overexpression will be constructed to detect the physiological phenotype of cardiac hypertrophy and investigate the underlying mechanism. TNNT1 cardiomyocyte-specific knockout mice will be generated and induced cardiac hypertrophy by TCA to test whether TNNT1 can be used as a target for the treatment of pathological cardiac hypertrophy.

In the future, we may face the following challenges in the clinical translation process. Firstly, currently there are no available inhibitors or activators of TNNT1. Therefore, to identify TNNT1-targeted drugs, it is necessary to start with drugs that target to downstream of TNNT1 or compete with TNNT1 at the site of action. Secondly, current models of drug research in the field of HCM are relatively limited. For example, induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) is closely

resemble human tissue. However, it is difficult to obtain and the sample size is limited [48]. Thirdly, the translation of TNNT1-targeted therapies to clinical applications is a time-consuming process. Given the absence of drugs that target TNNT1, the potential interactions with different classes of cardiovascular drugs are not considered.

Our study not only identified TNNT1 as a novel biomarker for hypertrophic cardiomyopathy (HCM) but also highlighted its potential as a therapeutic target. The up-regulation of TNNT1 in HCM offers promising avenues for the development of targeted pharmacological interventions aimed at modulating the troponin complex, which could potentially prevent or mitigate the pathogenesis of HCM. Current treatment options for HCM, such as implantable devices and surgical intervention, carry significant risks and limitations. Consequently, there is a growing need for novel therapeutic strategies. At present, pharmacological mechanisms for the drug treatment of HCM mainly include the following aspects: (1) the regulation of myocardial cell metabolism; (2) ion channel inhibition; (3) antioxidant; (4) oxygen free radical scavenging; (5) induction of autophagy; (6) allosteric inhibition of cardiac myosin ATPase; (7) reduction of left ventricular fibrosis [49]. The identification of TNNT1 and its involvement in cardiac contractility present a novel opportunity to design drugs that specifically target the molecular pathways associated with this biomarker. Moreover, further studies into the development of TNNT1 pathway inhibitors or modulators could provide a new class of pharmacological treatments for managing HCM and reducing the risk of sudden cardiac death in affected individuals. This discovery also aligns with the broader goal of personalized medicine, as it offers potential for patient-specific therapeutic approaches based on TNNT1 expression profiles.

CONCLUSIONS

In conclusion, our study has achieved several notable objectives. We constructed an interaction network between HCM pathogenic genes and found the most significant module of pathogenic genes. Furthermore, we analyzed the mode of interaction between hub genes from bioinformatic visualization and found that TNNT1 may be a potential key factor in the pathogenesis and progression of HCM.

REFERENCES

- Diagnosis and Evaluation of Hypertrophic Cardiomyopathy: JACC State-of-the-Art Review. *J Am Coll Cardiol.* 2022 Feb 1; **79**(4): 372-389.
- Thomson KL, Ormondroyd E and Harper AR *et al* (2019). Analysis of 51 proposed hypertrophic

- cardiomyopathy genes from genome sequencing data in sarcomere negative cases has negligible diagnostic yield. *Genet Med.*, **21**(7): 1576-1584.
- Base editing correction of hypertrophic cardiomyopathy in human cardiomyocytes and humanized mice *Nat Med.* 2023 Feb; **29**(2): 401-411.
- Lopes LR, Brito D, Belo A and Cardim N (2019). Genetic characterization and genotype-phenotype associations in a large cohort of patients with hypertrophic cardiomyopathy - An ancillary study of the Portuguese registry of hypertrophic cardiomyopathy. *Int J Cardiol.*, **278**: 173-179.
- Rupp S, Felimban M and Schanzer A *et al* (2019). Genetic basis of hypertrophic cardiomyopathy in children. *Clin Res Cardiol.*, **108**(3): 282-289.
- Minor hypertrophic cardiomyopathy genes, major insights into the genetics of cardiomyopathies. *Nat Rev Cardiol.* 2022 Mar; **19**(3): 151-167.
- Molecular Genetic Basis of Hypertrophic Cardiomyopathy. *Circ Res.* 2021 May 14; **128**(10): 1533-1553
- Recent Findings Related to Cardiomyopathy and Genetics. *Int J Mol Sci.*, 2021 Nov 20; **22**(22): 12522.
- Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 2021 Jan 8; **49**(D1): D10-D17.
- Identification of key biomarkers and immune infiltration in systemic lupus erythematosus by integrated bioinformatics analysis. *J Transl Med.* 2021 Jan 19; **19**(1):35.
- Tanabe M, Kanehisa M (2012). Using the KEGG database resource. *Curr Protoc Bioinformatics*, 2012; Chapter 1: 1-12.
- Haw RA, Croft D, Yung CK *et al* (2011). The Reactome BioMart. Database-Oxford bar031.
- Huang DW, Sherman BT and Tan Q *et al* (2007). The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.*, **8**(9): R183.
- Szklarczyk D, Gable AL and Lyon D *et al* (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**(D1): D607-D613.
- Kohl M, Wiese S and Warscheid B (2011). Cytoscape: Software for visualization and analysis of biological networks. *Methods Mol Biol.*, **696**: 291-303.
- Wang J, Zhong J, Chen G, Li M, Wu FX and Pan Y (2015). ClusterViz: A Cytoscape APP for Cluster Analysis of Biological Network. *Ieee Acm T Comput Bi.*, **12**(4): 815-822.
- Hong C and Hauskrecht M (2016). Multivariate Conditional Outlier Detection and Its Clinical Application. *Proc AAAI Conf Artif Intell.*, pp.4216-4217.
- Maere S, Heymans K and Kuiper M (2005). BiNGO: A Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics.*, **21**(16): 3448-3449.
- Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.*, **8** 4(Suppl 4): S11.
- Basha O, Shpringer R, Argov CM and Yeger-Lotem E (2018). The DifferentialNet database of differential protein-protein interactions in human tissues. *Nucleic Acids Res.*, **46**(D1): D522-D526.
- Wang S, Sun H, Ma J *et al* (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat Protoc.*, **8**(12): 2502-2515.
- Howe EA, Sinha R, Schlauch D, Quackenbush J (2011). RNA-Seq analysis in MeV. *Bioinformatics.*, **27**(22): 3209-3210.
- Papatheodorou I, Moreno P and Manning J *et al* (2020). Expression Atlas update: From tissues to single cells. *Nucleic Acids Res.*, **48**(D1): D77-D83.
- Pontier F, Jirstrom K and Uhlen M (2008). The Human Protein Atlas--a tool for pathology. *J Pathol.*, **216**(4): 387-393.
- Prondzynski M, Mearini G and Carrier L (2019). Gene therapy strategies in the treatment of hypertrophic cardiomyopathy. *Pflug Arch Eur J Phy.*, **471**(5): 807-815.
- Douglas JJ (2020). Current state of the roles of alcohol septal ablation and surgical myectomy in the treatment of hypertrophic obstructive cardiomyopathy. *Cardiovasc Diagnostics*, **10**(1): 36-44.
- Liu J, Wu G and Zhang C *et al* (2020). Improvement in sudden cardiac death risk prediction by the enhanced American College of Cardiology/American Heart Association strategy in Chinese patients with hypertrophic cardiomyopathy. *Heart Rhythm.*, **17**(10): 1658-1663.
- Rowin EJ, Maron MS, Bhatt V, Gillam L and Maron BJ (2020). Hypertrophic Cardiomyopathy in "Real-World" Community Cardiology Practice. *Am J Cardiol.*, **125**(9): 1398-1403.
- Mazzarotto F, Olivetto I and Boschi B *et al* (2020). Contemporary Insights Into the Genetics of Hypertrophic Cardiomyopathy: Toward a New Era in Clinical Testing? *J Am Heart Assoc.*, **9**(8): e015473.
- Hayashi T (2020). Hypertrophic Cardiomyopathy: Diverse Pathophysiology Revealed by Genetic Research, Toward Future Therapy. *Keijo J Med.*, **69**(4): 77-87.
- Glavaski M, Velicki L and Vucinic N (2023). Hypertrophic Cardiomyopathy: Genetic Foundations, Outcomes, Interconnections, and Their Modifiers. *Medicina-Lithuania.*, **59**(8): 1424
- Bueno MM, Celeghin R and Cason M *et al* (2020). A microRNA Expression Profile as Non-Invasive Biomarker in a Large Arrhythmogenic Cardiomyopathy Cohort. *Int J Mol Sci.*, **21**(4): pp.--? **Hypertrophic Cardiomyopathy Mutations of Troponin**

- Reveal Details of Striated Muscle Regulation. *Front Physiol.* 2022 May 26;13: 902079.
- Cardiomyocyte contractile impairment in heart failure results from reduced BAG3-mediated sarcomeric protein turnover. *Nat Commun.* 2021 May 19;12(1):2942.
- A Revised Perspective on the Evolution of Troponin I and Troponin T Gene Families in Vertebrates. *Genome Biol Evol.* 2023 Jan 4;15(1):evac173.
- Manuylov NL and Tevosian SG (2009). Cardiac expression of Tnnt1 requires the GATA4-FOG2 transcription complex. *Scientific World Journal*, **9**: 575-587.
- Feng HZ, Wei B and Jin JP (2009). Deletion of a genomic segment containing the cardiac troponin I gene knocks down expression of the slow troponin T gene and impairs fatigue tolerance of diaphragm muscle. *J Biol Chem.*, **284**(46): 31798-31806.
- The loss of slow skeletal muscle isoform of troponin T in spindle intrafusal fibres explains the pathophysiology of Amish nemaline myopathy. *J Physiol.* 2019 Aug;597(15):3999-4012.
- Pinto JR, Gomes AV and Jones MA *et al* (2012). The functional properties of human slow skeletal troponin T isoforms in cardiac muscle regulation. *J Biol Chem.*, **287**(44): 37362-37370.
- Tandem CTCF sites function as insulators to balance spatial chromatin contacts and topological enhancer-promoter selection. *Genome Biol.*, **21**(1): 75.
- Oki M and Kamakaka RT (2002). Blockers and barriers to transcription: Competing activities? *Curr. Opin. Cell Biol.*, **14**(3): 299-304.
- Sheng JJ and Jin JP (2016). TNNI1, TNNI2 and TNNI3: Evolution, regulation, and protein structure-function relationships. *Gene*, **576**(1 Pt 3): 385-394.
- Bhavsar PK, Dellow KA, Yacoub MH, Brand NJ and Barton PJ (2000). Identification of cis-acting DNA elements required for expression of the human cardiac troponin I gene promoter. *J Mol Cell Cardiol.*, **32**(1): 95-108.
- Rajib SA, Sharif SM (2020). Characterization and Analysis of Mammalian AKR7A Gene Promoters: Implications for Transcriptional Regulation. *Biochem Genet.*, **58**(1): 171-188.
- Mondal A and Jin JP (2016). Protein Structure-Function Relationship at Work: Learning from Myopathy Mutations of the Slow Skeletal Muscle Isoform of Troponin T. *Front Physiol.*, **7**: 449.
- Feng HZ, Chen X, Hossain MM and Jin JP (2012). Toad heart utilizes exclusively slow skeletal muscle troponin T: An evolutionary adaptation with potential functional benefits. *J. Biol. Chem.*, **287**(35): 29753-29764.
- Xu Z, Feng X and Dong J *et al* (2017). Cardiac troponin T and fast skeletal muscle denervation in ageing. *J Cachexia Sarcopeni*, **8**(5): 808-823.
- Santini L, Palandri C, Nediani C, Cerbai E and Coppini R (2020). Modelling genetic diseases for drug development: Hypertrophic cardiomyopathy. *Pharmacol. Res.*, **160**: 105176.