Journal of Translational Medicine

RESEARCH

Open Access

Proteomic profiling identifies molecular subtypes and unveils mechanistic insights into clinical features of hypertrophic cardiomyopathy

Xinli Guo[1,2], Changpeng Song[1], Changrong Nie[3], Xinxin Zheng[1], Shuiyun Wang[3,4] and Xiaohong Huang[1*]

Abstract

Background Hypertrophic cardiomyopathy (HCM) is the most common inherited heart disease, shows significant genetic and clinical heterogeneity, yet its molecular basis remains unclear. This study aims to uncover molecular insights behind HCM based on proteomic analysis.

Methods We performed proteomic analysis on formalin-fixed paraffin-embedded septal tissue samples from 105 patients with obstructive HCM undergoing myectomy, including 35 with MYBPC3 and 35 with MYH7 mutations. The primary outcome was major adverse cardiovascular events (MACE), comprising all-cause mortality and other major cardiovascular events.

Results We identified four molecular subtypes: 39 patients with subtype-I (S-I), 38 with S-II, 27 with S-III, and 1 with S-IV. During a median follow-up of 6.8 years, 28.6% of patients developed MACE. Among the subtypes, S-III exhibited the minimal fibrosis and the best clinical outcomes. S-II presented the most severe phenotype and poorest prognosis, characterized by activation of inflammatory and fibrotic pathways, along with downregulation of multiple metabolic processes. Compared to S-II, S-I displayed opposing pathway patterns and better outcomes, with a predominance of MYBPC3 mutation carriers. Beisides, we identified protein modules that were associated with genotypes and clinical features. MYH7-related modules were positively correlated with fibrosis and MACE, primarily enriched in inflammatory and fibrotic pathways. In contrast, MYBPC3-related modules were linked to better clinical outcomes, enriched in energy metabolism pathways.

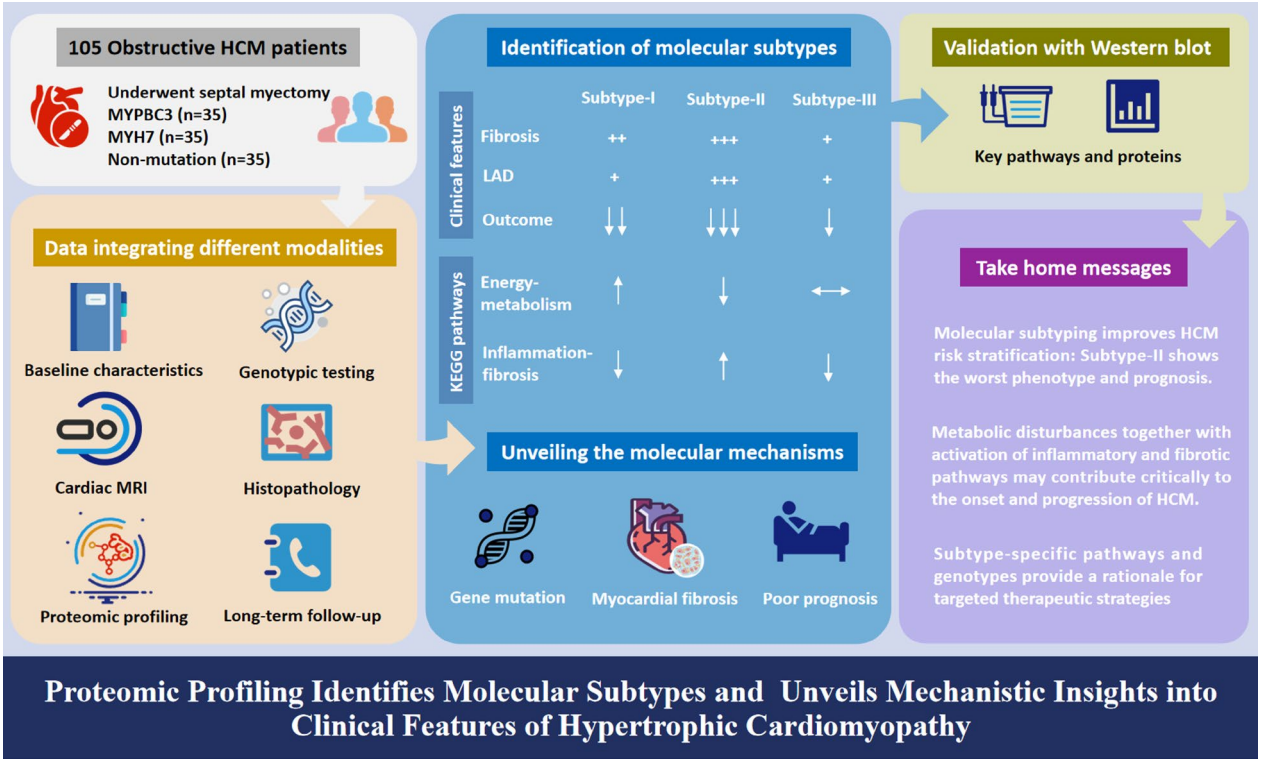
Conclusions This proteomic study revealed molecular mechanisms linking genotype, fibrosis and prognosis in HCM, and identified potential drivers of high-risk subtypes with severe phenotypes. These findings may guide future risk stratification and therapeutic target development.

Keywords Hypertrophic cardiomyopathy, Proteomics, Subtypes, Molecular mechanism, Phenotype, Prognosis

*Correspondence:
Xiaohong Huang
huangxiaohong@fuwaihospital.org

Full list of author information is available at the end of the article

Graphical Abstract



Introduction

Hypertrophic cardiomyopathy (HCM) is the most common inherited cardiovascular disease, affecting approximately 1 in 200 to 500 individuals. Genetic studies have identified over 1,500 sarcomeric gene variants in HCM, which exhibits significant genetic heterogeneity, with MYH7 and MYBPC3 being the most prevalent mutations [1, 2]. Traditionally, research indicates that genotype influences clinical phenotypes. Compared to gene-negative patients, gene-positive individuals tend to show more pronounced left ventricular hypertrophy (LVH) [3], greater myocardial fibrosis (MF) [4, 5], and worse outcomes [6]. Moreover, MYH7 mutations are associated with more severe myocardial energetic dysfunction and poorer clinical profile compared to MYBPC3 mutations [7–9]. However, some studies report no significant phenotypic differences between mutation types [10], suggesting that genetic variants alone do not fully explain the clinical heterogeneity [11]. Notably, 10%-15% of HCM patients experience major adverse cardiovascular events (MACE) annually [12]. Despite the high incidence of MACE, the molecular mechanisms driving disease onset and progression remain poorly understood [13]. A deeper molecular insight of phenotypic variability and disease progression is crucial for developing targeted therapies and improving patient outcomes.

Proteomic analysis enables concurrent quantification of thousands of proteins across both tissue and plasma samples, not only uncovering molecular pathways active in disease progression but also offers innovative strategies for identifying novel biomarkers and therapeutic targets. In HCM research, proteomic analysis plays a crucial role, given the disease's close association with the abnormal expression of key contractile proteins [14]. Recent studies have shown significant myocardial energy metabolism abnormalities in HCM patients, along with dysregulation of the Ras-MAPK and associated signaling and fibrosis-related pathways [15, 16]. However, molecular differences between genotypes remain controversial. Some studies suggest no significant differences in transcriptomic or proteomic profiles across genotypes [17], while others report distinct protein expression patterns [18], warranting further investigation. Moreover, the molecular mechanisms underlying myocardial fibrosis and adverse prognosis remain unclear, limiting our understanding of HCM's clinical heterogeneity.

Therefore, this study aims to perform proteomic profiling of myocardial tissue from patients with obstructive HCM to explore the relationships between proteomic profiles and clinical features. Our investigation will elucidate the molecular mechanisms underlying disease

phenotype, progression, and prognosis while examining genotype-specific proteomic signatures.

Methods

Study design and sample

This study retrospectively enrolled patients with obstructive HCM who underwent septal myectomy at Fuwai Hospital, Beijing, between January 2016 and July 2019. Through preoperative genetic screening, three genetically defined subgroups were identified: 35 patients carrying pathogenic MYH7 mutations, 35 with MYBPC3 mutations, and 35 without detectable sarcomeric gene mutations. All participants were over 18 years of age and underwent preoperative comprehensive cardiac imaging, and postoperative evaluation for pathological MF. HCM was diagnosed based on the presence of unexplained LVH (excluding cardiac or systemic diseases) with imaging evidence of a septal or maximum left ventricular (LV) wall thickness ≥ 15 mm or ≥ 13 mm in patients with a family history. Patients exhibiting a left ventricular outflow tract (LVOT) pressure gradient exceeding 30 mmHg at rest or during provocation were classified as having obstructive HCM. This study received approval from the Research Ethics Committee of Fuwai Hospital, conducted in accordance with the principles of the Declaration of Helsinki, and written informed consent was obtained from all patients.

Comprehensive baseline data were systematically collected for all participants through the hospital's electronic medical record system, including demographic characteristics, preoperative medication history (documented use for ≥ 2 weeks), and clinical comorbidities.

Cardiac imaging

Echocardiographic images were obtained using the E9 ultrasound system (GE Healthcare, Horten, Norway), and all echocardiographic parameters were measured according to standard protocols. Cardiovascular magnetic resonance (CMR) imaging was performed at Fuwai Hospital using 1.5T or 3.0T MRI scanners to assess cardiac structure, function, and tissue characteristics. All images were acquired with electrocardiogram (ECG) gating and respiratory control, following internationally recognized scanning protocols. LV volume, mass, and function were assessed using a commercial imaging workstation from Siemens Healthineers. Late gadolinium enhancement (LGE) image analysis was performed using a commercial imaging workstation (CVI 42; Circle Cardiovascular Imaging). Quantitative LGE was defined as myocardial enhancement exceeding six standard deviations above the signal intensity of a remote myocardium reference region without LGE in the same slice [19].

Outcome measures

All patients underwent annual telephone follow-ups by clinical physicians, documenting postoperative survival status, event occurrence, and corresponding dates. The last follow-up was conducted in October 2024. The primary outcome, MACE, was a composite of as a composite of all-cause death, new-onset heart failure (HF), atrial fibrillation (AF), acute coronary syndrome, or stroke. If a patient experienced multiple events, the time of the first event was recorded as the MACE time.

Tissue sample processing and proteomics profiling

Tissue Collection: Myocardial specimens were obtained from hypertrophic septal myocardium protruding 1–3 cm below the aortic valve during septal myectomy. The resected myocardial tissues were immediately fixed in 10% formalin solution and embedded in paraffin.

Pathological MF detection: Masson's trichrome stain provides clear color contrast to highlight structures such as collagen and muscle fibers under the microscope. The degree of interstitial myocardial fibrosis was evaluated with Image-Pro Plus 6.0, measuring the collagen fiber percentage as the ratio of collagen area to total field area.

Protein Extraction: Following dewaxing treatment, samples were transferred to 1.5 ml centrifuge tubes and mixed with 4× volume of lysis buffer (1% sodium deoxycholate, 1% protease inhibitor cocktail) for ultrasonic disruption. After centrifugation at 12,000×g for 10 min at 4 °C, cellular debris was removed and the supernatant was collected for protein quantification using the bicinchoninic acid assay kit.

Trypsin Digestion: Equal amounts of protein from each sample were subjected to enzymatic digestion. Initial digestion was performed with trypsin at 1:50 (enzyme: protein, w/w) ratio overnight, followed by secondary digestion at 1:100 ratio for 4 h.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis: Peptides were reconstituted in mobile phase A (0.1% formic acid in water) and loaded onto Evotip cartridges according to manufacturer's protocol. Separation was achieved using the Evosep One UHPLC system with a 60SPD (samples per day) gradient program. Mobile phase B consisted of 0.1% formic acid in 100% acetonitrile. Eluted peptides were ionized through a Capillary ion source (1.6 kV voltage) and analyzed by timsTOF Pro2 mass spectrometer operated in data-independent parallel accumulation serial fragmentation (diaPASEF) mode. MS1 scans covered 100–1700 m/z range followed by 10 PASEF acquisitions per cycle. MS2 spectra were collected from 400 to 1200 m/z using 35 m/z isolation windows.

Proteomics analysis

Pathway Enrichment Analysis: Differentially expressed proteins (DEPs) were defined by fold change (FC) > 1.5 (upregulated) or < 1/1.5 (downregulated) with $p < 0.05$. KEGG functional enrichment and cluster analyses were performed on DEPs.

Molecular Subtyping: Proteins with > 50% missing values were filtered, followed by K-nearest neighbors imputation. Features were selected by median absolute deviation, retaining the top 50% for consensus clustering (ConsensusClusterPlus R package; Euclidean distance, k-means, 1,000 iterations). Subtype-specific signatures (top 250 upregulated proteins per subtype) were visualized by heatmap.

Single-Sample Gene Set Enrichment Analysis (ssGSEA): Gene set activity scores were calculated per sample by ranking expression values and computing enrichment scores for predefined gene sets, reflecting pathway activation in individual specimens.

Weighted Gene Co-expression Network Analysis (WGCNA): A systematic methodology for identifying coordinated biological patterns through weighted correlation networks. In the current investigation, we implemented this algorithm using the R package to analyze proteomics datasets through four sequential phases: (1) Network construction by calculating soft-thresholded adjacency matrices based on pairwise gene expression correlations; (2) Module detection through hybrid clustering integrating dynamic tree-cutting with hierarchical clustering, where distinct modules were demarcated by color assignments in cluster dendrograms; (3) Module-trait association analysis using eigengene-based connectivity measures to pinpoint phenotype-relevant modules; (4) Functional characterization of critical modules via integrated KEGG/GO enrichment profiling and hub gene identification through intramodular connectivity ranking.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics 21.0 (Version 21.0, NY, USA) and R software (Version 4.1.1, Vienna, Austria). Continuous variables were presented as either mean ± standard deviation or median (interquartile range) depending on their distribution, and categorical variables as frequency (percentage). For comparisons among the three genotype groups in HCM patients, one-way analysis of variance (ANOVA) was used for continuous variables, while categorical variables were compared using the Chi-square test or Fisher's exact test. For comparisons between patients with and without MACE events, independent-samples t-tests were used for continuous variables, and categorical variables were analyzed with the Chi-square test or Fisher's exact test. Kaplan–Meier survival curves were generated, and the Log-rank test was employed to assess differences in survival between groups.

Validation of protein quantification consistency

The consistency of protein expression results among three molecular subtypes were validated using Western blotting (WB) and immunohistochemistry (IHC), with detailed antibody information provided in Supplementary Table 1.

WB analysis was performed to validate protein quantification consistency. A 10% separation gel and 5% stacking gel were prepared according to protein molecular weight. Samples containing 10 μg protein were loaded with 1×loading buffer to equalize total volume per well. Electrophoresis was conducted at 80 V (stacking gel, 40 min) followed by 120 V (separation gel) until bromophenol blue reached the gel bottom. Proteins were transferred to 0.45 μm PVDF membranes at 200 mA constant current. Membranes were blocked with 5% BSA-TBST (1 h, RT) and incubated with primary antibodies diluted in 5% BSA-TBST overnight at 4 °C. After three 10-min PBST washes, membranes were incubated with HRP-conjugated secondary antibodies (1 h, RT) followed by four 10-min PBST washes. ECL detection was performed using equal volumes of solutions A/B, with exposure parameters optimized based on signal intensity. Protein band quantification was conducted using Image J software for grayscale analysis.

IHC was performed on paraffin-embedded tissue sections. Sections were deparaffinized in xylene, rehydrated through graded ethanol, and rinsed in distilled water. Heat-induced antigen retrieval was carried out in EDTA buffer using a microwave protocol (8 min at medium power, an 8-min interval, followed by 7 min at medium–low power). After cooling, the sections were washed three times with phosphate-buffered saline (PBS) and blocked with bovine serum albumin (BSA) for 30 min. Primary antibodies were applied and incubated overnight at 4 ℃. Following PBS washes, sections were incubated with the corresponding secondary antibodies for 50 min at room temperature in the dark. Signal detection was achieved using 3,3′-diaminobenzidine (DAB), after which sections were counterstained with hematoxylin, differentiated in acid alcohol, blued, and rinsed. Finally, the sections were dehydrated in absolute ethanol, cleared in xylene, and mounted with neutral resin.

Results

Baseline characteristics

Supplementary Table 2 summarizes the baseline characteristics of the study cohort. Among the 105 enrolled patients, the mean age was 41.7 years old, with 65.9% male. Echocardiography revealed enlarged left atrial diameter (LAD) (45.3 ± 6.8 mm), ventricular septal

thickening (20.7 ± 4.2 mm), and elevated resting LVOT gradient (70.6 ± 37.3 mmHg). Moderate-to-severe mitral regurgitation was observed in 44.8% ($n = 47$) of patients, and 94.3% exhibited systolic anterior motion (SAM). Genotype-stratified analysis revealed significant age differences at diagnosis, with non-carriers presenting the oldest and MYH7 mutation carriers the youngest. The patients with MYH7 mutation exhibited the highest prevalence of baseline AF and the largest LAD. Fibrosis quantification revealed MYH7 carriers had the most pronounced myocardial remodeling, with LGE of 11.2 ± 10.3% and pathological MF of 19.8 ± 10.9%, followed by MYBPC3 mutation carriers (LGE: 10.8 ± 7.9%; MF: 13.4 ± 5.4%) and non-carriers (LGE: 9.8 ± 5.7%; MF: 11.4 ± 5.8%). Over a median follow-up of 6.8 years (IQR: 5.6–7.5y), 28.6% ($n = 30$) experienced MACE, comprising new-onset AF (12.4%, $n = 13$), HF rehospitalization (10.5%, $n = 11$), all-cause mortality (5.7%, $n = 6$), and cardiovascular events (11.4%, $n = 12$).

Proteomics-based molecular subtypes of HCM

Supplementary Fig. 1 presents the quality control results of the proteomic analysis. A total of 4,332 proteins were identified, among which 2,385 proteins had complete KEGG functional annotations and 4,188 proteins had complete GO annotations. Molecular weight distribution analysis revealed that the identified proteins ranged from 5 to 3,816 kDa. Principal component analysis is shown in Fig. 1a. To further explore the data structure and distribution, unsupervised clustering was performed to group samples with high intra-group similarity and large inter-group differences (supplementary Fig. 2). Supplementary Fig. 3 illustrates the criteria for model selection in molecular subtyping and the resulting classification of the 105 patients into four molecular subtypes: Subtype-I (S-I) ($n = 39$), S-II ($n = 38$), S-III ($n = 27$), and S-IV ($n = 1$). Detailed information on the differentially expressed proteins among the subtypes is provided in Supplementary Table 3. The S-IV subtype was excluded from subsequent analyses because of its extremely small sample size ($n = 1$), which limited its representativeness and precluded meaningful interpretation.

Functional enrichment analysis of each molecular subtype (Supplementary Tables 4 and Fig. 1b) revealed distinct pathway profiles: S-I was primarily enriched in pathways related to oxidative phosphorylation, thermogenesis, cardiac muscle contraction, and glycine, serine, and threonine metabolism. S-II showed significant enrichment in the complement and coagulation cascades, ECM-receptor interaction, PI3K-Akt signaling pathway, and TGF-β signaling pathway. S-III was mainly enriched in pathways associated with the complement and coagulation cascades. No KEGG-related pathways were enriched in S-IV. The ssGSEA results (Fig. 1c) revealed distinct pathway activities across subtypes. In S-II, inflammatory and fibrotic pathways such as complement and coagulation cascades, ECM-receptor interaction, PI3K-Akt signaling, and TGF-β signaling were significantly upregulated, while pathways related to cardiac muscle contraction and various metabolic processes were markedly downregulated. In contrast, S-I showed increased metabolic activity and reduced activation of inflammatory and fibrotic pathways. S-III was mainly characterized by the downregulation of inflammatory and fibrotic pathways, with no notable changes observed in metabolic pathways.

Molecular subtypes and its clinical associations

Clinical characteristics across molecular subtypes are summarized in Table 1. No significant differences were observed in demographic characteristics, comorbidities, or medication use among the three molecular subtypes. S-I was mainly composed of MYBPC3 mutation carriers (84.6%, $n = 33$), while 15.4% had MYH7 mutations. In the S-II, half of the patients carried MYH7 mutations and only one had a MYBPC3 mutation. S-III primarily included mutation-negative patients (63.0%, $n = 17$), with the remainder carrying MYH7 mutations (37.0%, $n = 10$). S-II showed the most impaired diastolic function, with a mean LAD of 45.8 mm and LVEDVi of 90.9 ml/m² on CMR. It also had the highest levels of fibrosis (LGE: 13.0 ± 8.2%; pathological MF: 19.9 ± 9.6%), followed by S-I, then S-III. In terms of prognosis, S-II exhibited the poorest outcomes, with the highest incidences of AF (23.7%, 9/38) and MACE (42.1%, 16/38), compared to S-I (AF: 10.3%, MACE: 25.6%) and S-III (AF: 0%, MACE: 14.8%) ($p < 0.05$ for both). Furthermore, Kaplan-Meier analysis (Fig. 1d) showed significant differences in prognosis among subtypes, with S-II having the worst outcomes, S-I intermediate, and S-III the most favorable.

Proteomic clustering and clinical correlation

This study used WGCNA to cluster genes with similar expression patterns and analyze their associations with clinical traits. Figure 2a shows ten modules, with sizes ranging from 51 to 790 genes. The brown module was positively correlated with MYH7 mutations ($r = 0.21$), MF (pathological fibrosis: $r = 0.54$; LGE: $r = 0.42$), and poor prognosis (MACE: $r = 0.22$; AF: $r = 0.18$) (Fig. 2b). The magenta module correlated with MYH7 mutations ($r = 0.32$) and MF (pathological fibrosis: $r = 0.45$; LGE: $r = 0.21$), but not prognosis. The blue module was positively correlated with MYBPC3 mutations ($r = 0.60$) and negatively with MYH7 mutations (r = -0.29) and MF (pathological fibrosis: r = -0.39; LGE: r = -0.30). The turquoise module strongly correlated with MYBPC3 mutations ($r = 0.85$) and negatively with MYH7 mutations (r = -0.24).

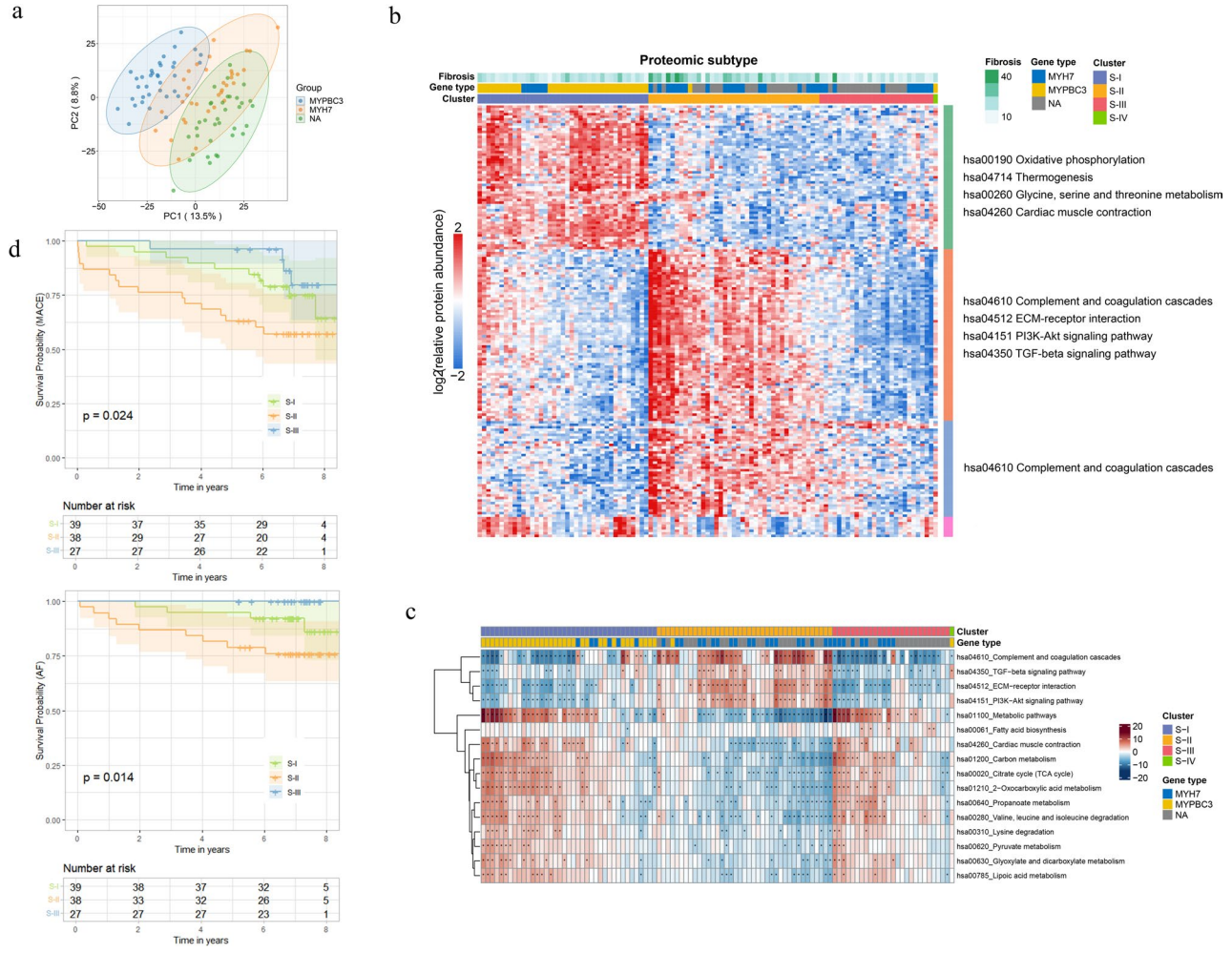


Fig. 1 Proteomic profiling delineates molecular subtypes of HCM with distinct functional pathways and clinical outcomes. **(a)** PCA of quantitative proteomics analysis for three sample groups. Blue points represent MYPBC3 mutation samples, orange points represent MYH7 mutation samples, and green points represent non-mutated samples. The groups form distinct clusters, with partial overlap between MYH7 and the other two, while MYPBC3 and non-mutated groups do not overlap. PCA: Principal Component Analysis. **(b)** Heatmap of Genotype-Molecular Subtype Correlation and Functional Features. Proteomic analysis identified four subtypes ($n = 105$): S-I (blue, $n = 39$), S-II (yellow, $n = 38$), S-III (red, $n = 27$), and S-IV (green, $n = 1$). The genotype and fibrosis status are annotated above. The heatmap shows the relative abundance of characteristic proteins, with associated key KEGG biological functions displayed on the right. **(c)** KEGG-ssGSEA Functional Enrichment of Protein Molecular Features and Genotype Background. Each column represents a sample, with the color gradient indicating gene set enrichment scores. Red shows upregulation of a KEGG pathway, and blue shows downregulation in the corresponding sample. The gradient from blue to red reflects the relative activity or suppression of pathways in different samples. The genotype and protein features are annotated at the top, and the enriched pathways are shown on the right. **(d)** Kaplan-Meier Curves of MACE and AF Incidence Across Three Molecular Subtypes. The colored curves represent the three molecular subtypes. S-II has the worst prognosis, followed by S-I, with S-III showing the best prognosis. Number of patients at risk for each time point is shown in the table below. MACE: Major Adverse Cardiovascular Events; AF * Atrial Fibrillation

KEGG enrichment analysis (Fig. 2c) revealed that the brown module was enriched in complement and coagulation cascades (log2Fold change [FC] = 7.02), ECM-receptor interaction (log2FC = 4.47), TGF-β and PI3K-Akt signaling (log2FC = 4.47 and 1.97, respectively). The magenta module was enriched in metabolic pathways, including lysine degradation (log2FC = 8.53) and motor proteins (log2FC = 3.84). The blue module focused on oxidative phosphorylation (log2FC = 3.32), thermogenesis (log2FC = 2.92), and cardiac muscle contraction (log2FC = 3.67), along with multiple metabolic pathways. The turquoise module was enriched in MAPK and RAS signaling (log2FC = 1.79 and 1.63, respectively).

Protein quantification validation

In the proteomic analysis, candidate proteins were selected based on molecular subtyping and pathway enrichment. For the S-I, associated with oxidative phosphorylation, we selected NDUFB11, CYCS, and ATP5PF. For the S-II, linked to PI3K-Akt and ECM-receptor interaction pathways, we selected VTN, COL1A1, and COL6A2. WB results (Supplementary Fig. 4 and Fig. 3)

Table 1 Clinical Characteristics by molecular subtype

	S-I (*n* = 39)	S-II (*n* = 38)	S-III (*n* = 27)	*p**
Demographic characteristics				
Age, y	42.4 ± 11.5	42.4 ± 13.9	40.1 ± 13.1	0.739
Male, n (%)	28 (71.8)	24 (63.2)	17 (63.0)	0.660
Duration, y	3.9 ± 4.9	6.0 ± 6.7	3.1 ± 2.6	0.052[c]
Smoking, n (%)	21 (53.8)	15 (39.5)	7 (25.9)	0.074[b]
Drinking, n (%)	9 (23.1)	8 (21.1)	2 (7.4)	0.231
BMI, kg/m^2	25.3 ± 3.6	25.4 ± 4.0	24.7 ± 3.3	0.707
BSA, m^2	1.8 ± 0.2	1.8 ± 0.2	1.8 ± 0.2	0.478
SBP, mmHg	122.8 ± 18.0	121.0 ± 12.2	121.7 ± 14.4	0.865
DBP, mmHg	73.3 ± 12.1	71.7 ± 9.3	72.7 ± 12.8	0.826
HR, beats	72.2 ± 9.0	73.9 ± 9.6	71.0 ± 8.3	0.422
NYHA III-IV, n	11 (28.2)	16 (42.1)	12 (44.4)	0.311
Family history, n	14 (35.9)	16 (42.1)	5 (18.5)	0.130[c]
Comorbidity, n				
Hypertension	8 (20.5)	6 (31.6)	5 (18.5)	0.865
Diabetes	0 (0.0)	0 (0.0)	0 (0.0)	1.000
Hyperlipidemia	9 (23.1)	15 (39.5)	10 (37.0)	0.264
Coronary heart disease	1 (2.6)	0 (0.0)	1 (3.7)	0.526
Atrial fibrillation	2 (5.1)	9 (23.7)	4 (14.8)	0.068
Medication, n				
β-Blocker	37 (94.9)	37 (97.4)	26 (96.3)	0.849
CCB	11 (28.2)	9 (23.7)	6 (22.2)	0.835
Diuretic	7 (17.9)	5 (13.2)	3 (11.1)	0.711
ACEI/ARB	5 (12.8)	4 (10.5)	2 (7.4)	0.781
Amiodarone	1 (2.6)	3 (7.9)	0 (0.0)	0.230
Echocardiographic indicators				
Rest LVOTG, mmHg	66.4 ± 33.0	65.7 ± 29.1	82.6 ± 50.4	0.143
Moderate to severe MR, n (%)	17 (43.6)	16 (42.1)	13 (48.1)	0.885
SAM, n (%)	35 (89.7)	37 (97.4)	26 (96.3)	0.310
CMR parameters				
IVS, mm	28.1 ± 5.3	25.7 ± 5.5	26.7 ± 5.6	0.168
LAD, mm	40.6 ± 8.6	45.8 ± 10.3	41.5 ± 7.3	0.031[a]
LVEDD, mm	46.4 ± 5.3	46.7 ± 5.7	44.4 ± 3.4	0.177
LVESVi, ml/m^2	29.5 ± 8.7	33.6 ± 13.1	30.0 ± 9.3	0.197
LVEDVi, ml/m^2	82.2 ± 16.2	90.9 ± 20.3	81.9 ± 13.5	0.046[a, c]
LVEF, %	64.2 ± 8.0	63.8 ± 8.5	64.0 ± 6.5	0.980
CO, L/min	6.5 ± 1.9	6.8 ± 2.0	6.3 ± 1.4	0.474
LVMassi, g/m^2	99.1 ± 43.1	104.2 ± 37.4	107.4 ± 39.8	0.686
LGE extent, %	10.6 ± 8.5	13.0 ± 8.2	6.5 ± 5.2	0.004[b, c]
Pathological MF, %	13.1 ± 5.3	19.9 ± 9.6	10.3 ± 7.3	< 0.001[a, c]
AF events	4 (10.3)	9 (23.7)	0 (0)	0.015[c]
MACE events	10 (25.6)	16 (42.1)	4 (14.8)	0.049
Genotype				< 0.001
G-	0 (0.0)	18 (47.7)	17 (63.0)	
MYPBC3	33 (84.6)	1 (2.6)	0 (0.0)	
MYH7	6 (15.4)	19 (50.0)	10 (37.0)	

* [a]$p < 0.05$ for S-I vs. S-II; [b]$p < 0.05$ for S-I vs. S-III; [c]$p < 0.05$ for S-II vs. S-III. The S-IV subtype was excluded from analysis due to insufficient sample size ($n = 1$)

Abbreviations: BMI, body mass index; BSA, body surface area; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; NYHA, New York heart association; CCB, calcium channel blocker; ACEI, indicates angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; LAD, left atrial diameter; LVEDD, left ventricular end-diastolic diameter; IVS, interventricular septum; LVOTG, left ventricular outflow tract gradient; MR, mitral regurgitation; SAM, systolic anterior motion; CMR, cardiac magnetic resonance; LVESVi, left ventricular end-systolic volume index; LVEDVi, left ventricular end-diastolic volume index; LVEF, left ventricular ejection fraction; CO, cardiac output; LGE, late gadolinium enhancement; MF, myocardial fibrosis; AF, atrial fibrillation; MACE, major adverse cardiovascular events

showed that the expression of NDUFB11, CYCS, and ATP5PF was significantly higher in the S-I group compared to the S-II and S-III groups ($p < 0.05$). VTN expression was significantly higher in the S-II group, while COL1A1 and COL6A2 levels were lower in the S-I group and showed a trend toward lower expression in the S-III group, although the p-value was not significant. Besides, the IHC results were shown in Supplementary Fig. 5, illustrating staining for the six target proteins in myocardial tissue from three representative individuals corresponding to the three subgroups. NDUFB11, CYCS, and ATP5PF exhibited stronger staining in the S-I group compared with S-II and S-III, indicating heightened mitochondrial activity. In contrast, VTN, COL1A1, and COL6A2 showed more intense staining in the S-II group, reflecting pronounced fibrotic remodeling.

Discussion

Summary of findings

This study performed proteomic profiling of myocardial tissue, integrating protein expression with genotype, clinical phenotype, and outcomes. We finally identified four molecular subtypes, among which S-II was characterized by the most severe phenotype and poorest prognosis. This subtype exhibited significant upregulation of inflammatory and fibrotic pathways, along with marked downregulation of myocardial contraction and multiple metabolic processes, including those related to energy, amino acids, and lipids. Notably, patients with MYBPC3 mutations were almost exclusively classified into the S-I subtype, displaying remarkable molecular homogeneity, which may inform future targeted therapeutic strategies. Further WGCNA analysis identified protein modules strongly associated with clinical features and genotypes, shedding light on the potential molecular mechanisms and pathological characteristics of HCM.

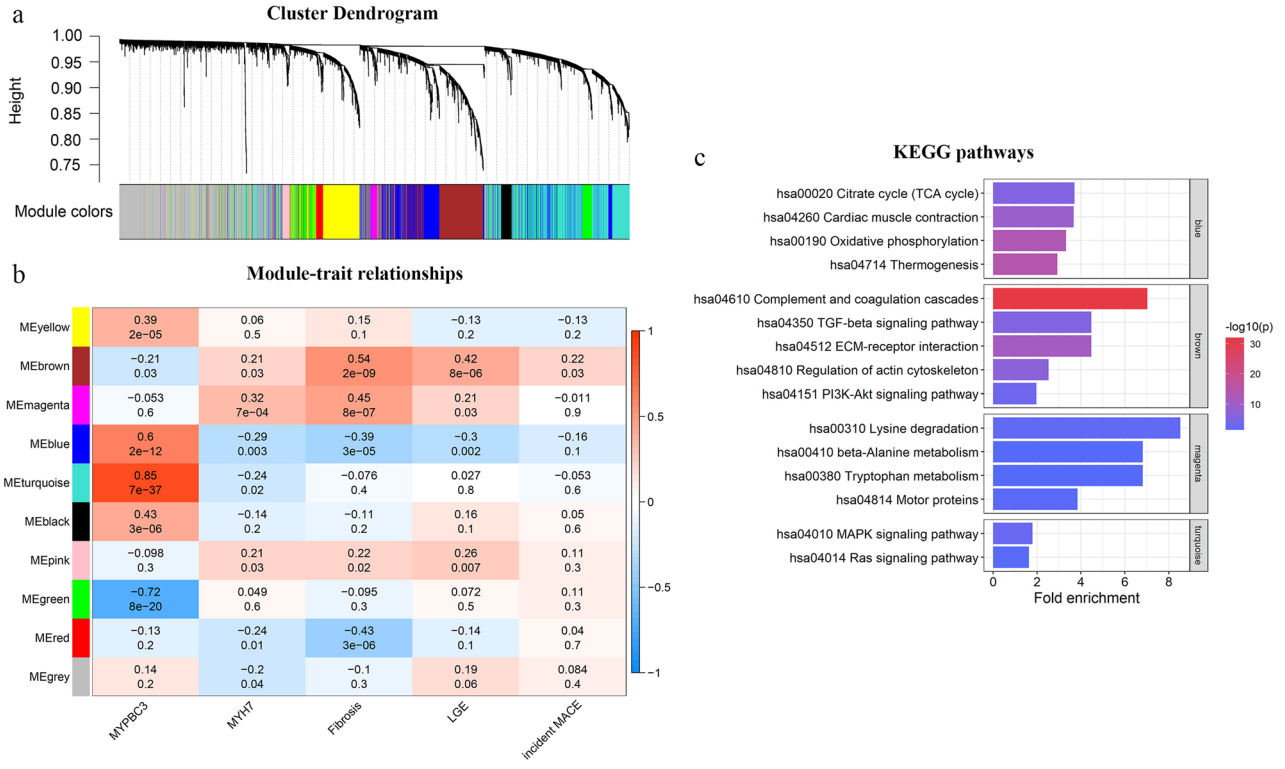


Fig. 2 Correlation heatmap of modules with key clinical features and KEGG pathway analysis. **(a)** Cluster dendrogram of gene modules. **(b)** Heatmap showing the correlation between 10 gene modules and clinical features. Colors indicate the strength and direction of correlations: red for positive, blue for negative. The numbers above each cell represent correlation coefficients, and those below represent p-values. **(c)** KEGG pathways enriched in important gene modules from WGCNA analysis. The length of each horizontal bar represents the enrichment fold value, while the bar color indicates the corresponding p-value.MACE: Major Adverse Cardiovascular Events; LGE: Late Gadolinium Enhancement

Interpretation of the results

To our knowledge, this is one of the largest proteomic studies of myocardial tissue in HCM patients to date, with the longest follow-up. Molecular subtyping identified S-II as the most severe subtype, characterized by extensive fibrosis, impaired diastolic function, and the highest risk of MACE, with half of the patients carrying MYH7 mutations. Consistent with recent proteomic studies linking AF and HF to dysregulation of Ras-MAPK related pathways and MACE to additional involvement of TGF-β signaling [20–23], we observed dysregulation of inflammatory, fibrotic, and Ras-MAPK associated pathways in the S-II subtype. Further ssGSEA analysis confirmed upregulation of these pathways and downregulation of metabolic and contractile pathways, suggesting impaired energy metabolism and worse outcomes.

In contrast, nearly all patients with MYBPC3 mutations were classified as the S-I subtype, suggesting greater molecular homogeneity and more favorable clinical outcomes. This group showed downregulation of inflammatory and fibrosis-related pathways and upregulation of metabolic and contractile pathways, possibly reflecting compensatory mechanisms that support better prognosis, consistent with previous studies [24]. This may be due to MYBPC3 mutations preserving contractile function by maintaining MyBP-C protein levels [25] or causing less disruption to ATPase activity and calcium sensitivity compared to MYH7 mutations [7, 26]. WGCNA further validated these associations. MYH7-related modules were enriched for dysregulated metabolic and fibrotic pathways and linked to worse outcomes, while the MYBPC3-related module was positively associated with energy metabolism and negatively correlated with fibrosis and adverse outcomes.

Notably, most proteomic and transcriptomic studies in HCM have focused on comparisons with healthy controls or genotype-positive versus genotype-negative patients, while direct comparisons between specific sarcomeric genotypes, particularly MYBPC3 and MYH7, remain limited. A metabolomic study of HCM demonstrated substantial overlap in PCA and largely comparable metabolite levels across genotypes, despite marked differences from healthy controls [16]. Similarly, another study directly comparing the proteomes and phosphoproteomes of HCM identified only minimal genotype-specific differences [27]. In addition, a multi-omics study reported that the ventricular proteome, largely independent of genotype, exhibited widespread

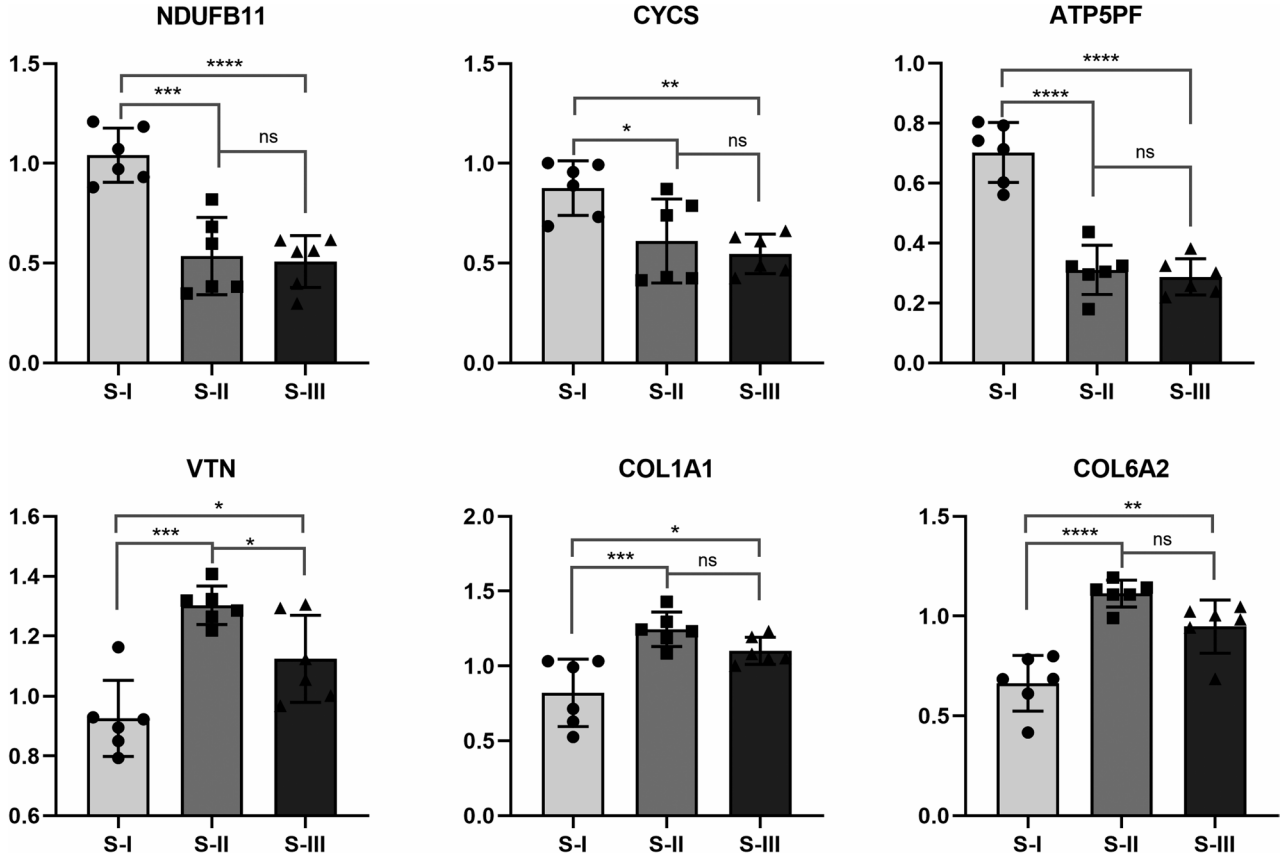


Fig. 3 Expression levels of candidate proteins across molecular subtypes. WB analysis of NDUFB11, CYCS, ATP5PF, VTN, COL1A1, and COL6A2 across the three molecular subtypes. Bar graphs show relative protein levels normalized to the internal control (mean ± SD). Statistical significance was assessed by one-way ANOVA: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$; ns indicates no significant difference

activation of hypertrophic pathways, predominantly involving RAS–MAPK signaling [17]. However, the relatively small sample sizes for each genotype in both studies limit the strength of these conclusions. Our findings are not inconsistent with these observations but provide a complementary perspective. By applying unsupervised molecular subtyping, WGCNA and integrating clinical parameters and long-term outcomes, we identified genotype-associated tendencies within clinically relevant molecular subtypes. Our findings above imply that global analyses may mask meaningful genotype-related signals, whereas incorporating pathway-level information together with disease severity and prognosis allows gene-specific differences and molecular subtypes to emerge more clearly.

Molecular mechanisms in HCM

In this study, we identified the S-II subtype as having the most severe clinical presentation and the highest risk of MACE. Pathway analysis showed activation of inflammatory and fibrotic pathways in this subtype, but suppression in S-I and S-III, suggesting that these pathways may drive myocardial fibrosis and MACE events. Similarly, further WGCNA analysis revealed a strong association between MYH7 gene mutations, MACE risk, fibrosis, and these pathways. Previous studies have shown that various components of the complement and coagulation systems (e.g., complement C3, thrombin, fibrinogen) act as mediators in inflammation, inflammation control, and tissue damage [28, 29]. The PI3K/Akt signaling pathway is activated in HCM patients, inducing cardiac hypertrophy, apoptosis, oxidative stress, and fibrosis [30–32]. TGF-β, a key pro-fibrotic cytokine involved in cardiac fibroblast and myocyte maintenance, is a major factor in myocardial fibrosis in HCM [33, 34], and angiotensin receptor blockade in HCM patients may alleviate myocardial hypertrophy and fibrosis by modulating the TGF-β pathway [35]. Ayca et al. found that TGF-β levels were higher in HCM patients than in healthy controls, with higher levels correlating with worse phenotypes and more clinical events [36]. Additionally, our previous research found that elevated preoperative TGF-β levels were independent predictors of postoperative AF in obstructive HCM patients following myectomy [37]. These suggest that the development of fibrosis in the heart adversely affects cardiac function and may contribute to poor prognosis.

Early inhibition of TGF-β signaling could therefore be a therapeutic strategy to prevent HCM-related fibrosis [33]. Besides, a study analyzing the GEO database identified key differential genes in HCM patients with HF, highlighting the potential roles of ECM-receptor interaction, TGF-β signaling, and PI3K/Akt signaling in HCM-related HF risk [38]. In conclusion, our findings suggest that the synergistic effect of these signaling pathways may exacerbate fibrosis, leading to worse prognosis.

The S-II subtype exhibits downregulation in key metabolic pathways, including energy, amino acid, and lipid metabolism, reflecting significant impairment in myocardial energy production and metabolism, leading to myocardial dysfunction and poor prognosis [16, 39]. A recent study suggests that in cardiac hypertrophy mice, increasing glucose consumption by reducing fatty acid oxidation can prevent hypertrophy and improve energy status [40]. Thus, therapies targeting these metabolic defects may help prevent or alleviate adverse cardiac remodeling. In contrast, the upregulation of metabolic pathways in the S-I subtype likely represents a compensatory mechanism, where enhanced glycolysis maintains energy supply, partially alleviating energy deficits and delaying myocardial dysfunction [41]. This may explain the slower disease progression and lower MACE risk in S-I patients. Additionally, we found that the S-I subtype is primarily composed of MYBPC3 gene mutations. WGCNA revealed that MYBPC3-related modules are enriched in Ras-MAPK and associated pathways and oxidative phosphorylation pathways, suggesting these mechanisms influence disease phenotype and prognosis. Previous studies have shown that Mavacamten, by inhibiting cMyBPC activity, restores diastolic and contractile function in MYBPC3-mutant cardiomyocytes, benefiting most patients [42]. Overall, these findings provide important insights for understanding HCM heterogeneity and developing targeted therapies.

Clinical implications

The cumulative burden of major adverse cardiovascular events in patients with HCM is substantial [14], especially in early-onset cases and those with sarcomeric gene mutations, highlighting the need for lifelong monitoring and prevention of adverse remodeling. Traditional risk prediction for HCM progression mainly relies on clinical data and cardiac imaging, but its predictive ability is limited. This study integrates genotype, clinical data, and proteomics to reveal the heterogeneity of HCM patients across different genetic mutations and molecular subtypes, providing molecular insights into disease progression and poor prognosis, and offering valuable support for clinical prognosis assessment and precision medicine.

First, we identified the molecular subtype S-II, associated with severe phenotypes and poor prognosis, followed by S-I and the best prognosis in S-III. Through myocardial tissue proteomics analysis, this study enhances risk stratification for HCM patients and supports the development of more targeted follow-up and intervention strategies. From a clinical perspective, planned prospective studies with paired baseline blood samples and postoperative tissue specimens aim to identify circulating biomarkers for these subtypes, which may facilitate blood-based disease classification and personalized management. Additionally, we identified key molecular pathways related to myocardial fibrosis and poor prognosis, including inflammatory and fibrotic pathways. Targeting these could slow disease progression. S-II subtype patients may benefit from interventions targeting the PI3K-Akt and TGF-β pathways, while S-I subtype patients with metabolic changes may require treatments for metabolic disorders. Finally, different genotypes significantly impact clinical manifestations and outcomes. The MYBPC3 genotype shows high homogeneity, suggesting its importance for clinical management and targeted therapies. In contrast, MYH7 mutations are linked to severe progression and worse prognosis, necessitating enhanced management for these patients.

Strengths and potential limitations

This study has several notable strengths. First, it represents one of the largest myocardial tissue-based proteomic analyses of HCM, integrating genetic, pathological, clinical, and imaging data for comprehensive investigation. Second, by leveraging robust and well-validated analytical approaches, we objectively uncovered molecular features and underlying mechanisms from large-scale data, reducing false discoveries and strengthening the scientific rigor of the study. Third, all patients were followed for over five years, providing valuable long-term outcome data. Fourth, using myocardial tissue rather than plasma allowed us to capture cardiac-specific protein alterations, particularly in intracellular and membrane-associated pathways, offering a more precise depiction of the pathological and functional changes in HCM. Finally, Key pathways and core proteins were further validated by WB and IHC in relevant patient subsets, strengthening the credibility of the findings and supporting mechanistic interpretation.

However, several limitations should be acknowledged. All patients were recruited from a single national center, potentially introducing selection bias toward more severe cases. Moreover, since all myocardial samples were obtained from patients undergoing septal myectomy, the cohort included only hospitalized patients with obstructive HCM, limiting the generalizability of findings to other subtypes or early-stage disease. Understanding early disease stages and dynamic changes remains crucial, though myocardial tissue sampling in early HCM is

challenging. In addition, fibrosis assessment was based on resected septal tissue, which may not fully represent the global myocardial fibrosis burden. Lastly, despite potential protein degradation or cross-linking in formalin-fixed paraffin-embedded (FFPE) samples, optimized proteomic workflows enabled reliable identification of over 4,000 proteins. Key findings were further supported by orthogonal validation using WB and IHC, reinforcing the robustness and credibility of the proteomic data.

Conclusion

This study explores the molecular mechanisms underlying the correlation between genotype and clinical features in HCM through myocardial proteomic profiling. We identified four distinct molecular subtypes, each with unique mechanisms and prognostic implications. Our findings highlight inflammatory and fibrotic pathway dysregulation together with metabolic disturbances, as key drivers of HCM progression. Additionally, we demonstrate significant molecular heterogeneity across genetic variants. These results provide a foundation for risk stratification and the development of targeted therapies, advancing precision medicine in the management of HCM.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-026-08206-x.

Supplementary Material 1

Supplementary Material 2

Author Contribution

XG, CS, SW, and XH contributed to the conception or design of the study. XG and CS were responsible for data acquisition, analysis, or interpretation; XG drafted the initial manuscript. XZ and CN verified the integrity and validity of the data analysis. SW and XH critically revised the manuscript. All authors have reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work, ensuring its integrity and accuracy.

Funding

Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0504500); the National Natural Science Foundation of China (Grant No. 82570466); and CAMS Innovation Fund for Medical Sciences (2024-I2MC&T-B-046).

Data availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest

All authors declare no conflict of interest.

Author details

[1]Department of Cardiology, Fuwai Hospital, National Center for Cardiovascular Disease, Chinese Academy of Medical Sciences, Peking Union Medical College, 167 Beilishilu Road, Beijing 100037, China
[2]Department of Cardiology, Anzhen Hospital, Beijing Anzhen Hospital, Capital Medical University, Beijing, China
[3]Department of Cardiovascular Surgery, Fuwai Hospital, National Center for Cardiovascular Disease, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100037, China
[4]Department of Cardiovascular Surgery, Fuwai Hospital, Chinese Academy of Medical Sciences, Shenzhen 510000, China

Received: 24 September 2025 / Accepted: 23 April 2026
Published online: 11 May 2026

References

1. Ingles J, et al. Application of Genetic Testing in Hypertrophic Cardiomyopathy for Preclinical Disease Detection. Circ Cardiovasc Genet. 2015;8(6):852–9.
2. Elliott PM, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy. Eur Heart J. 2014;35(39):2733–79. The Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC).
3. Guo X, et al. MYH7 mutation is associated with mitral valve leaflet elongation in patients with obstructive hypertrophic cardiomyopathy. Heliyon. 2024;10(14):e34727.
4. Haq IU, et al. Atrial Fibrillation Substrate and Catheter Ablation Outcomes in MYBPC3- and MYH7-Mediated Hypertrophic Cardiomyopathy. JACC Clin Electrophysiol. 2024;10(7 Pt 1):1380–91.
5. Huang M, et al. SARC Gene Mutation Is Associated With Myocardial Fibrosis Measured by Histopathology and Cardiac Magnetic Resonance in Patients With Hypertrophic Cardiomyopathy. J Am Heart Assoc. 2023;12(5):e028293.
6. Gwak SY, et al. Role of Genetic Testing in Diagnosis and Prognosis Prediction in Hypertrophic Cardiomyopathy in Korea. J Korean Med Sci. 2024;39(50):e313.
7. Witjas-Paalberends ER, et al. Gene-specific increase in the energetic cost of contraction in hypertrophic cardiomyopathy caused by thick filament mutations. Cardiovasc Res. 2014;103(2):248–57.
8. Sabater-Molina M, et al. Genetics of hypertrophic cardiomyopathy: A review of current state. Clin Genet. 2018;93(1):3–14.
9. Marsiglia JD, et al. Screening of MYH7, MYBPC3, and TNNT2 genes in Brazilian patients with hypertrophic cardiomyopathy. Am Heart J. 2013;166(4):775–82.
10. Weissler-Snir A, et al. Lack of phenotypic differences by cardiovascular magnetic resonance imaging in MYH7 (β-myosin heavy chain)- versus MYBPC3 (myosin-binding protein C)-related hypertrophic cardiomyopathy. Circ Cardiovasc Imaging. 2017;10(2).
11. Landstrom AP, Ackerman MJ. Mutation type is not clinically useful in predicting prognosis in hypertrophic cardiomyopathy. Circulation. 2010;122(23):2441–9. discussion 2450.
12. Maron BJ. Clinical Course and Management of Hypertrophic Cardiomyopathy. N Engl J Med. 2018;379(7):655–68.
13. Zampieri M, et al. Pathophysiology and Treatment of Hypertrophic Cardiomyopathy: New Perspectives. Curr Heart Fail Rep. 2021;18(4):169–79.
14. Ho CY, et al. Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy: Insights from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). Circulation. 2018;138(14):1387–98.
15. Lumish HS, et al. Comprehensive Proteomic Profiling of Human Myocardium Reveals Signaling Pathways Dysregulated in Hypertrophic Cardiomyopathy. J Am Coll Cardiol. 2024;84(20):1999–2011.
16. Previs MJ, et al. Defects in the Proteome and Metabolome in Human Hypertrophic Cardiomyopathy. Circ Heart Fail. 2022;15(6):e009521.
17. Garmany R, et al. Multi-Omic Architecture of Obstructive Hypertrophic Cardiomyopathy. Circ Genom Precis Med. 2023;16(2):e003756.
18. Schuldt M, et al. Proteomic and Functional Studies Reveal Detyrosinated Tubulin as Treatment Target in Sarcomere Mutation-Induced Hypertrophic Cardiomyopathy. Circ Heart Fail. 2021;14(1):e007022.
19. Harrigan CJ, et al. Hypertrophic cardiomyopathy: quantification of late gadolinium enhancement with contrast-enhanced cardiovascular MR imaging. Radiology. 2011;258(1):128–33.
20. Lee C, et al. Signaling Pathways Associated With Prior Cardiovascular Events in Hypertrophic Cardiomyopathy. J Card Fail. 2024;30(3):462–72.
21. Lumish HS, et al. Prediction of new-onset atrial fibrillation in patients with hypertrophic cardiomyopathy using plasma proteomics profiling. Europace. 2024;26(11).
22. Lumish HS, et al. Prediction of worsening heart failure in hypertrophic cardiomyopathy using plasma proteomics. Heart. 2023;109(24):1837–43.

23. Liang LW, et al. Proteomics profiling reveals a distinct high-risk molecular subtype of hypertrophic cardiomyopathy. Heart. 2022;108(22):1807–14.
24. Sedaghat-Hamedani F, et al. Clinical outcomes associated with sarcomere mutations in hypertrophic cardiomyopathy: a meta-analysis on 7675 individuals. Clin Res Cardiol. 2018;107(1):30–41.
25. Helms AS, et al. Effects of MYBPC3 loss-of-function mutations preceding hypertrophic cardiomyopathy. JCI Insight. 2020;5(2).
26. Witjas-Paalberends ER, et al. Mutations in MYH7 reduce the force generating capacity of sarcomeres in human familial hypertrophic cardiomyopathy. Cardiovasc Res. 2013;99(3):432–41.
27. Garmany R, et al. Proteomic and phosphoproteomic analyses of myectomy tissue reveals difference between sarcomeric and genotype-negative hypertrophic cardiomyopathy. Sci Rep. 2023;13(1):14341.
28. Hoppe B, Dörner T. Coagulation and the fibrin network in rheumatic disease: a role beyond haemostasis. Nat Rev Rheumatol. 2012;8(12):738–46.
29. Lappegård KT, et al. A vital role for complement in heart disease. Mol Immunol. 2014;61(2):126–34.
30. Worou ME, et al. Hemin decreases cardiac oxidative stress and fibrosis in a rat model of systemic hypertension via PI3K/Akt signalling. Cardiovasc Res. 2011;91(2):320–9.
31. Cheng Y, et al. Qingda granule attenuates angiotensin II-induced cardiac hypertrophy and apoptosis and modulates the PI3K/AKT pathway. Biomed Pharmacother. 2021;133:111022.
32. Wang L, et al. MYH7 R453C induced cardiac remodelling via activating TGF-β/Smad2/3, ERK1/2 and Nox4/ROS/NF-κB signalling pathways. Open Biol. 2024;14(6):230427.
33. Teekakirikul P, et al. Cardiac fibrosis in mice with hypertrophic cardiomyopathy is mediated by non-myocyte proliferation and requires Tgf-β. J Clin Invest. 2010;120(10):3520–9.
34. Schlittler M, et al. Myocardial fibrosis in hypertrophic cardiomyopathy: a perspective from fibroblasts. Int J Mol Sci. 2023;24(19).
35. Shimada YJ, et al. Effects of losartan on left ventricular hypertrophy and fibrosis in patients with nonobstructive hypertrophic cardiomyopathy. JACC Heart Fail. 2013;1(6):480–7.
36. Ayça B, et al. Increased Transforming Growth Factor-β Levels Associated With Cardiac Adverse Events in Hypertrophic Cardiomyopathy. Clin Cardiol. 2015;38(6):371–7.
37. Guo Y, et al. Usefulness of Preoperative Transforming Growth Factor-Beta to Predict New Onset Atrial Fibrillation After Surgical Ventricular Septal Myectomy in Patients With Obstructive Hypertrophic Cardiomyopathy. Am J Cardiol. 2017;120(1):118–23.
38. Dong H, et al. Identification of differentially expressed genes to predict the risk of heart failure in older patients with hypertrophic cardiomyopathy. Aging. 2024;16(13):10860–7.
39. Ranjbarvaziri S, et al. Altered Cardiac Energetics and Mitochondrial Dysfunction in Hypertrophic Cardiomyopathy. Circulation. 2021;144(21):1714–31.
40. Ritterhoff J, et al. Metabolic Remodeling Promotes Cardiac Hypertrophy by Directing Glucose to Aspartate Biosynthesis. Circ Res. 2020;126(2):182–96.
41. Toepfer CN, et al. Myosin Sequestration Regulates Sarcomere Function, Cardiomyocyte Energetics, and Metabolism, Informing the Pathogenesis of Hypertrophic Cardiomyopathy. Circulation. 2020;141(10):828–42.
42. Toepfer CN, et al. Hypertrophic cardiomyopathy mutations in MYBPC3 dysregulate myosin. Sci Transl Med. 2019;11(476).

Publisher's note