Check for updates

OPEN ACCESS

EDITED BY Yeong-Renn Chen, Northeast Ohio Medical University, United States

REVIEWED BY

Chia-Feng Liu, Cleveland Clinic, United States Patrick Kang, Northeast Ohio Medical University, United States

*CORRESPONDENCE Matthew T. Wheeler wheelerm@stanford.edu

RECEIVED 17 December 2024 ACCEPTED 03 March 2025 PUBLISHED 28 July 2025

CITATION

Lindholm ME, Abramowitz S, Waggott DM, Grove ME, Dewey FE, Pan C, Pavlovic A, Shang C, Huang Y, Bensabath L, Goldfeder RL, Cordero P, Erbilgin A, Priest JR, Chaib H, Puckelwartz MJ, Day SM, McNally EM, Cappola T, Dorn GW, Ashley EA and Wheeler MT (2025) Identification of candidate cardiomyopathy modifier genes through genome sequencing and RNA profiling. Front. Cardiovasc. Med. 12:1546493. doi: 10.3389/fcvm.2025.1546493

COPYRIGHT

© 2025 Lindholm, Abramowitz, Waggott, Grove, Dewey, Pan, Pavlovic, Shang, Huang, Bensabath, Goldfeder, Cordero, Erbilgin, Priest, Chaib, Puckelwartz, Day, McNally, Cappola, Dorn, Ashley and Wheeler, This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of candidate cardiomyopathy modifier genes through genome sequencing and RNA profiling

Malene E. Lindholm¹, Sarah Abramowitz^{1,2}, Daryl M. Waggott¹, Megan E. Grove³, Frederick E. Dewey^{1,4}, Cuiping Pan⁵, Aleksandra Pavlovic¹, Ching Shang¹, Yong Huang⁶, Leore Bensabath¹, Rachel L. Goldfeder⁷, Pablo Cordero⁷, Ayca Erbilgin^{1,8}, James R. Priest^{1,9}, Hassan Chaib⁵, Megan J. Puckelwartz¹⁰, Sharlene M. Day¹¹, Elizabeth M. McNally¹⁰, Thomas Cappola¹¹, Gerald W. Dorn¹², Euan A. Ashley^{1,3,5,6} and Matthew T. Wheeler^{1,6*}

¹Stanford Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, United States, ²Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hofstra University, Hempstead, NY, United States, ³Stanford Clinical Genomics Program, Stanford University School of Medicine, Stanford, CA, United States, ⁴Forsite Labs, San Francisco, CA, United States, ⁵Department of Genetics, Stanford University School of Medicine, Stanford, CA, United States, ⁶Center for Undiagnosed Diseases at Stanford, Stanford University School of Medicine, Stanford, CA, United States, ⁷Program in Biomedical Informatics, Stanford University School of Medicine, Stanford, CA, United States, ⁸Department of Medicine, University of California, Los Angeles, CA, United States, ⁹Tenaya Therapeutics, San Francisco, CA, United States, ¹⁰Center for Genetic Medicine, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States, ¹¹Division of Cardiovascular Medicine, Department of Medicine, University of Pennsylvania, Philadelphia, PA, United States, ¹²Center for Pharmacogenomics, Department of Internal Medicine, Washington University School of Medicine, States, ¹⁰Center for Internal Medicine, Washington University

Background: Phenotypic heterogeneity is apparent among individuals with putative monogenic disease, such as familial hypertrophic cardiomyopathy. Genome sequencing (GS) allows interrogation of the full spectrum of inborn genetic variation in an individual and RNA profiling provides a snapshot of the cardiac-specific pathogenic effects on gene expression.

Objectives: Identify candidate genetic modifiers of hypertrophic cardiomyopathy phenotype.

Methods: We performed GS of 48 individuals with variants in *MYH7*, the gene encoding beta myosin heavy chain, and a personal or family history of cardiomyopathy. The genome sequences were annotated with a custom pipeline optimized for cardiovascular gene variant detection. We utilized multiple lines of evidence to prioritize genes together with rare variant gene-based association testing to identify candidate genetic modifiers.

Results: GS identified the *MYH7* variant in all 48 cases. Several variants were reclassified based on best available data. We identified known disease-associated genes (*MYBPC3, FHOD3*), a priori candidate modifiers (*ATP1A2, RYR2*), and novel candidate modifiers of cardiomyopathy including *PACSIN3* and *SORBS2*. We identified regulatory variants and intergenic regions associated with the phenotypes. Using RNA profiling, we show that several genes identified through gene-based association testing are differentially regulated in human hypertrophic cardiomyopathy, and in models of disease.

Conclusion: Evaluation of the whole genome, even in the case of alleged monogenic disease, leads to important new insights. The identified variants, regions, and genes are candidates to modify disease presentation in cardiomyopathy.

KEYWORDS

hypertrophic cardiomyopathy, MYH7, modifier, genome sequencing, left ventricular hypertrophy

Introduction

Familial cardiomyopathy is a major cause of morbidity and mortality. Hypertrophic cardiomyopathy (HCM) was first mapped by positional cloning to the cardiac beta myosin heavy chain (MYH7) (1, 2). Subsequently variants in other sarcomeric genes were found to be pathogenic for HCM. Today, a pathogenic genetic variant is identified in <50% of HCM cases (3), where rare and private variants in myosin binding protein C (MYBPC3) and MYH7 are most common (4). High variability in disease presentation, penetrance and expressivity has caused some to question the single variant nature of these diseases and led to speculation that additional genetic variation may be contributing (5-7). Disease heterogeneity, both among families with identical "disease-causative" variants and between individuals within these families, can span severity and presentation and complicates diagnosis and prognosis (6). Importantly, polygenic risk can explain some variability in disease presentation in patients with HCM (8). Additionally, several groups demonstrated that patients harboring multiple rare variants in disease-associated genes may develop severe phenotype or present earlier (9-11). Genome-wide association studies have identified large effect variants with moderate population frequencies (8, 12), while small to moderate effect variants and variants of rare frequency that could influence disease heterogeneity have been less well-characterized. Our understanding of how additional genetic variants of intermediate effect, commonly referred to as genetic modifiers, influence the severity and manifestation of causal alleles has been limited, due in part to the difficulties encountered in identifying genetic modifiers. Recent evidence also shows contribution of common variants to the risk of developing HCM (12), corroborating its genetic complexity.

Technological advances rapidly reduced the cost of genome sequencing (GS) such that application is a clinical reality. While testing continues to be focused on gene panels for primary variant identification, improved quality of sequencing chemistry and processing algorithms has enabled clinical exome and genome sequencing. GS has identified new pathogenic variants in HCM (13), as well as in diseases previously not considered genetic such as acute myocarditis (14). Data from gnomAD (15) have refined our understanding of gene tolerance to variation. We previously explored the concept of gene tolerance and its utility in classifying cardiovascular disease variants (16, 17). GS uniquely allows interrogation of all genes, regulatory regions, intergenic regions, as well as detection of larger indels and

structural variants. In combination with RNA sequencing, the regulatory implications of different variants can be investigated further to improve our understanding of the genetic complexity of disease.

Here, we utilize patient derived GS and multiple sources of transcriptomic data to identify genetic modifiers of the disease phenotype in cardiomyopathy patients with a previously identified variant in *MYH7*. We develop a pipeline specific to cardiomyopathy, update curation of all putatively causal variants, and examine several aspects of variation uniquely available in genome data. In combination with transcriptional coexpression network data, we identify candidate modifiers of the cardiomyopathy phenotype. The expression profiles of several candidates are assessed in three independent models of cardiac hypertrophy.

Methods

Cohort

The study was approved by local IRB (GAP 4237) and conformed to the Declaration of Helsinki. We identified 48 individuals with prior genetic testing for cardiomyopathy associated variants in *MYH7*. Patients were recruited from the inherited cardiomyopathy clinics of Stanford University, University of Michigan, and the University of Chicago Medical Centers. Prior to study inclusion, all participants underwent informed consent to research-based genome sequencing. Clinical details including variant classification from the original genetic testing report were obtained. Clinical measurements and demographics are found in Tables 1, 2 and Supplementary Table S1 (echocardiography) and Data S1 (individual clinical characteristics).

| TABLE 1 | Baseline | characteristics. |
|---------|----------|------------------|
|---------|----------|------------------|

| НСМ | | | | | | |
|-----------------|----|--------|-------|------|-----|-----|
| Trait | N | Median | Mean | SD | Min | Max |
| Age (yrs) | 42 | 49.5 | 47.2 | 16.1 | 7 | 79 |
| Weight (lbs) | 40 | 182.5 | 177.8 | 47.7 | 47 | 272 |
| Height (inches) | 40 | 67 | 66.7 | 5.4 | 48 | 80 |
| Other | | | | | | |
| Trait | Ν | Median | Mean | SD | Min | Max |
| Age (yrs) | 6 | 46.0 | 39.3 | 21.2 | 8 | 61 |
| Weight (lbs) | 5 | 143 | 117.8 | 47.8 | 42 | 155 |
| Height (inches) | 5 | 68 | 61.6 | 10.6 | 44 | 69 |

| TABLE 2 | Demographic | s. |
|---------|-------------|----|
|---------|-------------|----|

| Name | Level | N | % |
|---------------|-------------------------------|----|------|
| Sample origin | Site 1 | 26 | 54 |
| | Site 2 | 13 | 27 |
| | Site 3 | 9 | 19 |
| Sex | F | 24 | 50 |
| | М | 24 | 50 |
| Race | Caucasian | 34 | 71 |
| | Hispanic | 2 | 4 |
| | Asian | 3 | 6 |
| | Middle Eastern | 2 | 4 |
| | African American | 7 | 15 |
| Morphology | Apical hypertrophy | 2 | 4.2 |
| | Asymmetric septal hypertrophy | 30 | 62.5 |
| | Burnt-out hypertrophic | 1 | 2.1 |
| | Concentric hypertrophy | 6 | 12.5 |
| | Dilated | 1 | 2.1 |
| | ECG+/LVH- | 3 | 6.2 |
| | Inferior hypertrophy | 2 | 4.2 |
| | Isolated noncompaction | 3 | 6.2 |

Clinical curation

Our curation process was two tiered. First, we manually curated any variant found within MYH7 to confirm the presence of the original diagnosis and reclassified pathogenicity according to additional lines of evidence. Second, we curated a broader set of candidate myocardial genes (n = 452; Supplementary Table S2) to identify alternate pathogenic variants, most of which were not interrogated in the initial gene panel screening. The clinical curation protocol was based on American College of Medical Genetics (ACMG) guidelines (18, 19) as described (16). The curation classes were: (1) Very likely disease causing (equivalent to "published, disease causing mutation"; "known disease causing mutation"; "pathogenic"; "disease causing mutation"). (2) Likely disease causing (Equivalent to "presumed pathogenic mutation"; "likely pathogenic"; "probably associated"). (3) VUS, likely disease causing (equivalent to "novel variant of uncertain significance, likely disease causing"). (4) VUS (equivalent to "class II possible deleterious mutation"; "VUS"; "novel variant of uncertain significance"). (5) VUS, likely benign. Primary criteria for reclassification included population allele frequency (GnomAD greater than 0.1%), pedigree segregation, supporting publication evidence and clinical laboratory assertions available through ClinVar (March 2021).

Sequencing, variant calling, filtering

Genome sequencing was performed on isolated genomic DNA acquired from blood samples from each participant. DNA library preparation and 2×100 paired-end sequencing was completed by Illumina. Coverage was on average $30 \times$ and sufficient for high quality germline variant genotyping.

Alignment and variant calling were completed using Real Time Genomics (20). Filtering was applied based on an AVR (adaptive

variant rescoring) score of 0.02. Variant imputation, filtering and structural variation calling were performed using a custom pipeline as previously described (21, 22) with modification. Variants were only called at positions that had evidence for an alternate allele. Missing variant imputation was completed to avoid bias during merging samples with different sets of variants. First, the union of variants across all samples was selected. Next, each sample was interrogated for missing genotypes at these positions. Missing genotypes were set to homozygous reference if there was in excess of 10 supporting reads and no evidence of an alternative allele.

The called variants were then limited to those that were novel or had an allele frequency of less than 1% (GnomAD v2 Spring 2021). The adopted 1% minor allele frequency filter is different than what was implemented for pathogenicity as we are considering modifiers of intermediate effect and thus intermediate population frequency. Regions enriched for false positives were also removed (Undiagnosed disease program blacklist (23), ENCODE Mappability (24), Heng Li low complexity (25), GIAB low confidence (26)). To protect against overzealous filtering, variants in ClinVar (Sep 2014), GWAS NHGRI Catalog (Fall 2014) (27), PharmGKB (Fall 2014) (28), Cosmic(v70) (29) were whitelisted. ClinVar variant classifications and disease associations are current as of July 2021 (Supplementary Dataset S1).

The remaining rare high confidence variants were further limited based on functional evidence. Specifically, variants were selected that were predicted by snpEff (30) to have moderate to high impact on protein function or have a normalized CADD score (31) greater than 20. The resulting VCF was subdivided into three sets depending on the required analysis. Set 1 includes all functional candidate variants, both genic and intergenic. Set 2 includes the genic region (refFlat Fall 2014) including introns and 5 kb up and downstream of the transcription start site. Set 3 is a digital exome using a 10 bp extension of the Agilent SureSelect exome capture v2. Each of these sets was selected for burden testing in order to compare power of discovery for a gene agnostic sliding window, whole gene region and coding exome.

We used a custom pipeline to call structural variants (SVs) across the genome. As structural variant callers are known to have lower confidence particularly with short read sequences, we evaluated for SVs that were shared among multiple algorithms and restricted to no more than 20% of samples. The tools included—breakdancer v1.1.2 (32), breakseqlite v1.0, cnvnator v0.2.7 (33), delly v0.0.9 (34) and freec v1.0 (35). High confidence copy number calls are reported in the analysis. They include anything called by Breakseqlite (36) (which is considered conservative) or having greater than 50% reciprocal overlap between at least two algorithms.

Intergenic and pathway burden testing

Intergenic regions were tested through a 50 kb sliding window with 25 kb overlap (~5 variants). Pathways were based primarily on candidates gene sets of mitochondrial and sarcomeric function. Additionally, a large number gene sets from Reactome (37) and MsigDB (38) were used for exploratory hypothesis generation. A key challenge with burden testing is appropriate selection, tuning of the number of variants (or genes) and subsequent weighting. The default, internally calculated, frequency based weighting parameters were used with SKAT.

Aggressive post burden test filtering was applied to the case control analysis to protect against false positives due to platform bias (genome vs. exome sequencing). Genes were removed if they were previously implicated as spurious in the UDP (23) or had a prefix consistent with large paralogous gene families (OR, PRAME, MUC, POTE, BAGE, ANK, NBF, DNA, DYN, CCDC, PCD, DUX, D3X, KRT, FOX, PRSS, KIR, OPN, USP). Preliminary analysis indicated comparison of genome to exome sequencing results in spurious batch effects enriched for gene families due to capture specificity or lack thereof. Burden testing results are available in Supplementary Data S4.

Statistical analysis

All analysis was completed using the R statistical environment and Bioconductor (39). Burden testing was performed using SKAT with adaptive weighting and adjustment for small sample sizes (40). In addition, for each region Fisher exact tests were used to compare the number of samples with rare variants among high and low groups. The high group was defined by median dichotomization for continuous traits or cases with binary trait analysis. The frequency of observed rare functional alleles was reported for each group (including a proportion test) to estimate the direction and magnitude of burden effect.

Burden testing was applied to four traits: maximum left ventricular wall thickness (LVWT), ejection fraction (EF), left ventricular outflow tract gradient (LVOTG) and case control status. The first three traits were analyzed within the cohort while case control analysis compared HCM patients against external controls (ARIC, n = 100) (41). For within cohort analysis a multivariate model using age, sex and BMI was used for adjustment with SKAT.

Modifier gene prioritization

Modifier identification was performed using a gene-ranking algorithm based on a custom point system rewarding different analytical and annotation lines of evidence (Table 3). Gene-based association testing using both SKAT and SKAT-O was applied to four traits: maximum LVWT, EF, LVOTG and case control status. All gene-based association testing was restricted to unrelated Caucasian patients with clinically diagnosed HCM (N = 28). For within cohort analysis a multivariate model using age, sex and BMI was used for adjustment. The modifier algorithm also incorporated extensive myocardial gene expression data from heart failure and hypertrophy (42), GnomAD missense and loss of function scores, cardiac gene expression level, as well as the curated candidate myocardial gene list (Supplementary

TABLE 3 Modifier gene scoring algorithm.

| Trait | Criteria | Points |
|-----------------------------------------------|----------------------|--------|
| Left Ventricular Wall Thickness (LVWT) | Gene-based ass. test | 2 |
| | <i>p</i> < 0.05 | |
| Ejection Fraction (EF) | Gene-based ass test | 1 |
| | <i>p</i> < 0.05 | |
| Maximal Left Ventricular Outflow Tract | Gene-based ass test | 1 |
| Gradient (LVOTG) | <i>p</i> < 0.05 | |
| Case Control ARIC (WES, N = 100), (Fu et al., | Gene-based ass test | 1 |
| 2013) | <i>p</i> < 0.05 | |
| GTEx Gene Expression | Cardiac RPKM > 10 | 1 |
| GnomAD LOF | >3 SD | 1 |
| GnomAD MIS tolerance | >3 SD | 1 |
| Hypertrophy vs. Normal (coexpression | Q < 0.01 | 1 |
| networks) | | |
| Heart Failure vs. Normal (coexpression | Q < 0.01 | 1 |
| networks) | | |
| Candidate Cardiac Gene | N = 452 genes | 2 |
| Flagged Prefix or UDP Blacklist | | -2 |

Table S2). The primary analysis was LVWT and therefore any gene reaching a nominal burden testing p-value < 0.05 was assigned 2 points. EF, LVOTG and Case-Control were secondary analyses and assigned 1 point each. The lowest burden test p-value for LVWT, EF, LVOTG and Case-Control per gene was used in the modifier scoring algorithm. Heart gene expression and significant differential expression within hypertrophy or heart failure coexpression networks (42) were each assigned 1 point. RNA expression was assessed based on the mean RPKM (reads per kilobase of exon per million reads mapped) values for 82 left ventricle GTEx (43) samples (>10 RPKM gave 1 point). GnomAD loss of function (LOF) z-score >3 and missense z-score >3 (positive z-scores indicate fewer variants than expected) each gave 1 point. Candidate genes (Supplementary Table S2), manually curated by a group of cardiologists specializing in inherited cardiovascular disease, were given 2 points. Genes previously flagged as prone to false positive signals (23) lost 2 points.

Human HCM validation

Cardiac gene expression of selected top candidate modifiers were investigated in a separate cohort of 39 HCM patients and 13 healthy control individuals (IRB approval GAP 4237). The HCM cohort (excluding one individual without available data) was 34% female and composed of individuals with a mean age of 57.9 years (standard deviation, SD = 15.7), mean height of 169.6 cm (SD = 12.8) and mean weight of 86.7 kg (SD = 34.0). RNA was isolated from cardiac left ventricle tissue from cardiac transplants or myectomies using the mirVana miRNA isolation kit (ThermoFisher Scientific) according to the manufacturer's specifications. Total RNA was reverse transcribed to cDNA using the High-Capacity cDNA Reverse Transcription kit (ThermoFisher Scientific). Quantitative real-time PCR was performed on a ViiA 7 Real-Time PCR System (ThermoFisher Scientific) using pre-designed Taqman gene expression assays

from ThermoFisher Scientific (FHOD3 Hs00400902_m1, MSRB2 Hs00255292 m1, MYH7B Hs00293096_m1, PACSIN3 Hs00367625_m1, SORBS2 Hs01125197_m1) or predesigned probes from IDT (NPPB HS.PT.58.19450190). The eukaryotic probe elongation factor EEF1 (IDT predesigned Hs.PT.58.3514123) was used as a housekeeping gene, and expression was quantified using the $\Delta \Delta C^{T}$ method followed by an unpaired *t*-test with BH-correction.

Functional RNA-seq verification

All animal procedures were in keeping with all federal and state regulations governing the humane care and use of laboratory animals, including the USDA Animal Welfare Act, and our Assurance of Compliance with PHS Policy on Humane Care and Use of Laboratory Animals, Animal Welfare Assurance Number: A3213-01, in accordance with the NIH Guide for the Care and Use of Laboratory Animals. The laboratory animal care program at Stanford is accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC International, Accredited Unit Number 000679). All animals were handled under protocols 22920 and 22922 approved by the Stanford Administrative Panel on Laboratory Animal Care. Mice were anesthetized with 3% isofluorane (inhalation), euthanized using cervical dislocation, after which the hearts were removed. The methods used to generate and analyze cardiac RNA sequencing data from the MYL2 transgenic mouse model have been described in detail elsewhere (44).

Neonatal rats were anesthetized on a pad on ice for 20 min and euthanized by decapitation. Neonatal rat ventricular myocytes (NRVMs) were isolated using standard methods on postpartum day 3 using the Worthington Neonatal Cardiomyocytes isolation system. After removal of non-cardiomyocyte cells, myocytes were plated on collagen I coated plates in DMEM 7.5% FBS, 7.5% horse serum, and penicillin/streptomycin. Forty-eight hours after isolation, cells were treated with 50 uM phenylephrine in low glucose DMEM with ITS and 20 mM AraC. Phenylephrine treatment (2-50uM) for 24-48 h is known to induce cardiomyocyte hypertrophy (increased cell volume, cell area and total protein content) (45-48). A subset of cells were harvested prior to addition of phenylephrine; half of remaining wells were treated with media alone. After 24 h and 48 h of treatment, acute phenylephrine induced cardiac cells were harvested. RNA was extracted using the Qiagen RNeasy kit according to the manufacturer's instructions and was DNase-treated using the DNA-free RNA kit from Zymo Research. RNA integrity was verified using a 2100 BioAnalyzer (Agilent) and all samples had an RIN score of 7.0 or higher. RNAseq libraries were prepared using the TrueSeq Stranded mRNA kit (Illumina), according to the manufacturers' instructions. Libraries were barcoded, qualitychecked and run in rapid run flow cells in a HiSeq 2,500 (Illumina), producing at least 30 million paired-end reads. Sequencing reads were aligned to the Rattus Norvegicus rn5 UCSC reference genome using the STAR (49) and Cufflinks was used to quantify and perform differential expression (50). Reads were normalized using Cufflinks and FPKM (Fragments Per Kilobase Of Exon Per Million Fragments Mapped) was calculated on a per gene basis. We evaluated the FPKM count for top candidate modifiers and canonical markers of the acute hypertrophic program and evaluated the differential expression after phenylephrine vs. pre-treatment as a ratio vs. sham vs. pretreatment. All genes with an FPKM>1 were considered expressed and included in the analysis.

Results

Genome sequencing of patients harboring *MYH7* variants

Patients were classified based on cardiac morphology and *MYH7* variant identified (primary variants identified between 2007 and 2011, n = 33, Tables 1-2, Supplementary Table S1 and Supplementary Data S1). Of the 48 patients (7 to 79 years old), 50% were female. The cohort included 6 families for a total of 40 unrelated individuals. Cardiomyopathy phenotypes were classified as hypertrophic in 42, dilated without known prior hypertrophy in 1, and noncompaction in 3. The overall study design is illustrated in Figure 1.

Multiple genetic variants of different classifications in *MYH7* have been reported in GnomAD (15) and ClinVar (51) (Figure 2A). Using our automated genome annotation pipeline, we re-identified all unique *MYH7* variants detected on initial panel-based sequencing (population minor allele frequency <0.1%, Figure 2B bottom panels, Supplementary Dataset S1). They were distributed across the actin-binding domain and the rod region, similar to the majority of reported variants in ClinVar (Figure 2B, top panels). All *MYH7* variants reported on prior clinical testing were identified and called at high quality and adequate depth to confidently call heterozygosity. All other genomic variants identified on panel testing were confirmed.

MYH7 variant reclassification

Secondary findings-coding

We next searched for secondary likely pathogenic genetic variants that may contribute to the cardiomyopathy phenotype (Supplementary Data S2). In addition to potentially disease-associated variants in *MYH7* from panel based testing, we identified one *MYBPC3* VUS (p.Val757Met), and rare *PRKAG2* variants (p.Val535Gly and p.Val534Gly respectively). We identified an early termination (p. Cys1013X) and rare missense (p.Lys953Gln) variant in *FLNC*, a gene first implicated in hypertrophic and dilated cardiomyopathy (DCM) years after the patients' clinical tests were reported (52, 53). Interestingly, the former *FLNC* truncation variant was found in a patient with a family history of DCM whose *MYH7* variant was reclassified to benign. Another patient whose *MYH7* variant was initially classified as VUS also had a rare nonsynonymous variant in RNA binding motif protein 20 (*RBM20* p.Ile921Val), a titin-



Overall study design and workflow. Main patient cohort (N = 48) all harbored *MYH7* variants. 42 patients had been diagnosed with HCM, and 6 had other diagnoses. Genome sequencing (GS) was performed with an Illumina HiSeq on all patients, followed by variant calling and curation. HCM candidate modifiers were identified through a scoring system that utilized a combination of RNA expression data, gene-based association testing and previous candidate hypertrophy-associated genes. Molecular follow-up of top modifier candidates was performed in three separate models of cardiac hypertrophy.

splicing gene implicated in cardiomyopathy (54). Two rare variants in BCL2-associated athanogene-3 (*BAG3* p.Glu471Gly and p.Arg473Gly), a stress responsive protein implicated in familial DCM were also identified (55–57). The *MYH7* Variant Reclassification is summarized in Supplementary Figure S1.

Review of rare coding variants in a cardiovascular subset of the 78 ACMG gene list (58) identified additional rare variants (Supplementary Data S2). These were found in genes associated with arrhythmogenic right ventricular cardiomyopathy (DSC2, DSG2, TMEM43, PKP2, DSP); Catecholaminergic polymorphic ventricular tachycardia (RYR2, CASQ2, TRDN); DCM (LMNA, FLNC, TTN, BAG3, RBM20); Vascular Ehlers-Danlos Syndrome (COL3A1); Long QT Syndrome 3/Brugada Syndrome (SCN5A); Familial Hypercholesterolemia (LDLR, APOB) and aortopathies (TGFBR1, SMAD3, MYH11).

Secondary findings—non-coding

Next, we interrogated miRNA binding sites and consensus promoter regions for potentially interesting variants (Supplementary Data S3). We identified two variants in conserved miRNA binding sites of genes previously implicated in cardiomyopathy: SCN5A (chr3:38589677) RAF1 and (chr3:12625903). The RAF1 miRNA binding site variant alters position 5 of a consensus seed recognition sequence. This variant is predicted to disrupt a permissive consensus binding site recognized by multiple miRNAs and found in more than 290 genes. We also identified a variant in a miRNA binding site of ACTN2 (chr1:236927207), a gene for which we and others recently demonstrated that protein-truncating variants cause hypertrophy and restrictive cardiomyopathy (RCM) in humans (59). We found variants of uncertain significance in non-coding

RNA, promoter, UTR, enhancers and nonsynonymous variants in alternative isoforms of tropomyosin found in long-read heart RNAseq data (Supplementary Table S3). Published work demonstrates a protective role of long non coding RNAs (*mhrt*) that reside in the 3' myosin tail-coding domain and downstream of *MYH7* (60). We identified 5 rare nonsynonymous variants that intersect with exons of these transcripts (Supplementary Data S3 and Table S3). As these variants affect both coding and non-coding transcripts, the impact of these variants is currently unknown. Reclassification (61) of these variants without knowledge of the role of their impact on *Mhrt* transcripts resulted in downgraded pathogenicity.

Scoring algorithm for modifier identification

We developed a multi-omic cardiomyopathy-data rich metric of modifying potential to evaluate for genetic modifiers of the HCM phenotype. To prioritize genes, we developed a 12-point scale (Table 3; Methods) that combined rare variant gene-based association testing with multiple external priors, including extensive myocardial gene expression data from heart failure and hypertrophy. Inclusion of parameters such as LVWT and disease-specific gene expression data enhanced sensitivity to an HCM-specific phenotype and cardiac function. There were 165 genes that reached a modifier score of 6 or better (Supplementary Table S4, see Supplementary Data S4 for full table) and 51 genes had scores ≥7. CACNA1C scored 10, while RYR2, TTN, ATP1A2, FHOD3, TJP2, CACNA1D and DYNC1H1 had modifier scores of 9. Several were identified as a priori candidates. The sodium-potassium ATPase subunit alpha 2 gene (ATP1A2) was highly significant in the case-control comparison $(p < 4 \times 10^{-19}).$ The cardiac specific ryanodine receptor encoded



pathogenic or likely pathogenic in ClinVar, including patient variants classified as pathogenic or likely pathogenic. At amino acid residues where multiple variants of different classification are present, the residue is colored based on the variant with maximum severity. Residues in both A and B are colored based on ClinVar classification with pathogenic variants in red, likely pathogenic variants in orange, variants with conflicting interpretations of pathogenicity in blue, variants of uncertain significance (VUS) in dark blue, likely benign variants in green and benign variants in yellow. GnomAD variants without a ClinVar classification are shown in dark green.

by *RYR2* was associated with increased wall thickness (p < 0.01), a trend towards LVOT gradient (p = 0.07) and with HCM in the case-control comparison (p < 0.05). Several variants in the L-type calcium channel component *CACNA1C* have also been demonstrated to predispose for arrhythmic phenotypes, especially long-QT syndromes (62), while the gap junction protein encoded by *TJP2* has been associated with hypertrophy in other cell types (63).

Validating our approach, the known modifier Formin homology 2 domain containing 3 (*FHOD3*) was a top scoring modifier. *FHOD3* has been implicated as disease predisposing in DCM and HCM (64, 65) and as an HCM modifier in casecontrol common variant analysis and a genome-wide association study (12, 66). Previously cited high frequency variants (rs516514, rs2303510) were not included in the rare functional variants used for gene-based association testing; despite this, variation in *FHOD3* was associated with increased LVWT (p < 0.05). *SORBS2* (sorbin and SH3 domaincontaining 2), modifier score 6, is an adhesion junction protein that was recently associated with arrhythmogenic cardiomyopathy (ACM). Knockdown of *SORBS2* in mice resulted in an ACM-like phenotype and two patients with ACM have been identified to carry likely pathogenic variants in SORBS2 (67).

We identified several genes as novel candidate modifiers of the HCM phenotype (Supplementary Table S4). Rare variant genebased association testing of DYNC1H1 (score 9), the heavy chain component of dynein, was associated with increased LVWT (p < 0.01) and reduced EF (p < 0.01). Another top modifier, HNRNPC (score 8), was recently identified as a regulator of sarcomeric protein translation with higher expression in failing hearts (68). ERBB2 (score 8) and ERBB4 (score 7), both Neuregulin-1 receptors, are involved in cardiac regeneration after injury (69, 70), and Neuregulin-1 itself is a therapeutic target for heart failure (71). Rare variants in CAMSAP1 (score 7) were found in 8 participants with severely increased LVWT but not in participants with milder hypertrophy (p < 0.001). ASH1l (score 7), a histone H3 methyltransferase, was most significant in the case-control comparison $(p < 1 \times 10^{-35})$. The transcription cofactor LIM domain binding 2 (LDB2, score 6) was associated with LVWT (p < 0.01) and with case status in case-control analysis ($p < 2 \times 10^{-15}$). PACSIN3 (score 6) is an adapter protein that regulates membrane dynamics. Cardiomyocytes in PACSIN3

knockout mice lack caveolae (72), while its potential modifying function in cardiac hypertrophy remains to be elucidated.

Independent experimental RNA profiling of identified candidate modifiers

To further evaluate if the identified candidate modifier genes were involved in cardiac hypertrophy, we analyzed RNA sequencing data from a mouse model of RCM and a cellular model of acute hypertrophy in neonatal rat ventricular myocytes (NRVMs). We specifically investigated the relative change in expression for all modifier genes with a score of 6 or above (n = 165) in response to the chronic and acute hypertrophic stimulus (n = 106 detected in both, Figures 3A,B, Supplementary Table S5, with low-scoring genes shown in Supplementary Figure S2A). While a change in expression suggests a role in development of cardiac hypertrophy, lack of change does not invalidate a potential modifier. We confirmed several genes as highly potentiated by a hypertrophic signal, including the known modifier FHOD3, and UCHL1 (Supplementary Figure S2B), for which knockdown in cardiomyocytes was recently shown to prevent hypertrophy (73). In contrast, other modifiers were modestly changed or expressed at low levels in cardiomyocytes, such as CAMSAP1 and ASH11. The most downregulated modifier in response to both acute and longterm hypertrophy was the potassium channel gene KCNJ3 (score 7, Supplementary Figure S2C). Genetic variation in KCNJ3 is associated with arrhythmias, and blocking the channel in zebrafish improved the arrhythmogenic phenotype (74).

Next, we analyzed gene expression differences in human left ventricular myocardial tissue from an independent cohort of 39 HCM patients compared to healthy controls (N = 13). Expression of the hypertrophy marker NPPB and five candidate modifiers (FHOD3, MSRB2, MYH7B, PACSIN3 and SORBS2) in all three experiments are shown in Figure 3C. The expression changes in SORBS2 follows the effect of NPPB, suggesting a role in hypertrophic remodeling. To further explore the utility of gene expression data in identifying and explaining the roles of modifiers, we assessed the expression of genes with recently-identified HCM susceptibility loci (8, 12) in our cardiomyopathy models. Genes with known cardiac functions, including BAG3, PLN, TBX3, TRDN, and MTSS1, all had lower expression in both the acute and chronic cardiomyopathy compared to controls (Supplementary Figure S3). SLC6A6, a taurine transporter, was higher in both models. ADPRHL1, a gene implicated in cardiac development, was upregulated in response to acute hypertrophy and downregulated in the chronic hypertrophy model (Supplementary Figure S3, Table S5).

Discussion

Genome sequencing offers unique and untapped power for understanding the complexity of genetic disease. Variability in severity, penetrance, and phenotype found within and among families harboring putatively disease-causative *MYH7* variants is presumed to be caused by both genetic and environmental modifiers. We evaluated the utility of GS in combination with cardiac expression data to identify candidate genetic modifiers in a patient cohort initially diagnosed with monogenic cardiomyopathy attributed to variants in *MYH7*.

Whereas prior studies of GS in HCM have focused on its utility in providing a genetic diagnosis in cardiomyopathy (13, 75), we showed how comprehensive data from GS of genetically diagnosed HCM patients identified potential genetic modifiers of the phenotype. Increasing evidence suggests that HCM is not a wholly monogenic disease. We searched each genome for evidence of additional rare variants causative of or substantially contributory to the observed cardiomyopathy phenotype and found rare protein altering variants in multiple genes and individuals, including an early termination variant in the cardiomyopathy associated gene FLNC. A VUS in FLNC in the presence of a VUS in MYBPC3 has previously been suggested as a potential modifier of HCM (75). These findings support previously identified advantages of GS in HCM, including the potential to improve yield in testing for primary pathogenic variants (13), and identify a range of secondary findings throughout the genome (75). Further exploration of secondary genetic variants, where gene expression is one potential tool, is necessary to validate their significance in disease pathophysiology and translate such findings into clinically actionable insight.

Novel potential genetic contributors to disease were discovered in the non-coding genome, including microRNA sites and lncRNAs. The discovery of non-coding DNA variants that may act as disease modifiers is the first step in developing a new field of biology aimed at understanding how regulatory gene variants impact Mendelian disease; this is fodder for applied genetic research. Further mechanistic research is needed to understand the impact of non-coding variants in the genetic basis for HCM.

We further explored the genomes for modifiers of disease phenotype. With limited sample size to afford true statistical power for genome wide testing we undertook two strategies to maximize the robustness of findings. We used a gene scoring algorithm that boosted nominally significant modifier phenotype associations with specialized annotation of hypertrophic cardiomyopathy and heart failure molecular phenotypes. We prioritized variants based on cardiac expression and differential expression in cardiac hypertrophy and heart failure. Several candidates were investigated for cardiac expression changes in a separate cohort of human HCM, a transgenic mouse model of chronic hypertrophy and an acute hypertrophy model in NRVMs. Our top candidates included previously known modifier genes (FHOD3), a gene known to cause DCM and HCM when mutated (TTN) and genes associated with related cardiovascular phenotypes such as heart failure (HNRNPC) and arrhythmia (CACNA1C, RYR2). The plausibility, and in the case of for example FHOD3, SORBS2 and UCHL1, replication of prior published results, lends credibility to our method and strengthen the discovery of novel candidates such as CAMSAP1, MYH7B and PACSIN3. Importantly, rare genetic variation does not inform directionality of effects, and utilizing several disease models for investigation of potential modifier expression changes in response to acute or chronic hypertrophy emphasizes the





FIGURE 3

UCHL1

USP43

COL1A1

RAB15

SLC7A5

I AMB3

MSRB2

OGDHL

OPA1

PDE4D

NDRG1

GRAMD1B

6

6

6

6

6

6

6

6

6

6

6

6

1.12

1 02

-1.06

0.81

-0.34

-0.02

1.78

-0.55

0.29

0.09

0.52

-1.25

Cardiac candidate modifier gene expression changes. (A) Differential expression analysis of all candidate modifiers with a score 6 or higher that were expressed in both a chronic model of hypertrophy (*MYL2* transgenic mice compared to WT mice, *x*-axis) and an acute model of hypertrophy (phenylephrine-treated (PE) compared to non-treated (Ctrl) neonatal rat ventricular myocytes (*NRVMs*), *y*-axis]. Expression changes are shown as log2 of the fold difference and genes are colored based on modifier score. (B) Table of all labeled candidate modifiers in (A), with score and log2FD listed for the acute and chronic hypertrophy models. Blue indicates negative change (log2FD -1 or lower), orange indicates positive change (log2FD 1 or more). (C) Differential expression of selected modifiers and the hypertrophy marker *NPPB* was investigated in a separate cohort of human HCM compared to healthy control left ventricle cardiac tissue, shown in the top left panel for each gene. Differential expression is also reported for the chronic hypertrophy model (*MYL2* transgenic mice compared to WT, top right panel for each gene) and in the acute model of cardiomyocyte hypertrophy (PE, red lines) of NRVMs. Candidate modifier expression in human cardiac left ventricle samples was investigated with qRT-PCR (*N* = 34-39 HCM patients for all factors, and *N* = 5-13 control individuals for each factor). Expression in *MYL2* transgenic mice (Mut, *N* = 3 mice) vs. littermate WT controls (WT, *N* = 3 mice), and in PE treated (*N* = 3 wells per time point) and non-treated (*N* = 3 wells per time point) and non-treated (*N* = 4 wells per time point). NRVMs was investigated using RNA sequencing. For NRVMs, results are shown for 3 time points: 0 h when PE was added, after 24 h of PE treatment and after 48 h of PE treatment. NRVM data is shown as mean \pm 95% CI (# indicates *p* < 0.05 ANOVA treatment effect), human and mouse data as mean \pm SEM, * *p* < 0.01 (BH-corrected unpaired t-test).

complexity of the observed effects. For *FHOD3*, for example, expression was substantially lower in RCM mice, but showed a trend towards higher levels in the HCM validation cohort and in response to acute hypertrophy.

Our study has several limitations. Our sample size is limited and includes mostly Caucasian individuals. The modifier scoring algorithm was developed to include cardiomyopathy-relevant metrics, but is limited by its scope and selection of scoring cutoff. While we further evaluated some of the modifiers using cardiac hypertrophy models and a separate HCM cohort, larger cohort studies are needed to examine their potential clinical relevance and future mechanistic studies to establish causal evidence. Importantly, the cell and animal models are limited with regards to cross-species concordance and the potential impact by acute or chronic hypertrophy at two time points. They are therefore unlikely to capture all differential signaling that influence human HCM development over time. Lack of validation thus does not invalidate a putative modifier. In addition, identified phenotypic associations were concurrent with the point of assessment. Hence, the phenotypic association is contingent upon the gene's direct impact but also on potential gene-mediated influences on ascertainment timing, treatment history etc. The test statistics and p-values will be inflated/deflated due to multiple tests using sparse data and should therefore only be used for ranking purposes and not to determine actual statistical significance. Furthermore, no validation was performed at the protein level, which will be an important avenue for future research as proteomic and phosphoproteomic signatures also impact disease severity in sarcomeric HCM (76, 77).

In summary, our results provide new insights into genetic modifiers of HCM, including disease-associated genes, proximal regulatory regions and non-coding variants. The complexity of secondary genetic variants with respect to modifying effects is currently minimally evaluated. The initial cataloging of these gene-gene interactions are preliminary and will require refinement and extension, just as the initial discovery of causal *MYH7* variants have been. We hypothesize that these secondary findings provide greater understanding in the interpretation of primary variants, and could eventually provide utility in predicting disease severity. True modifiers of disease severity represent promising targets for drug therapy irrespective of causative variants and warrant further mechanistic investigation.

Data availability statement

The genome data presented in the current publication have been deposited in and are available from the dbGaP database under dbGaP Study Accession phs004024.v1.p1.

Ethics statement

The studies involving humans were approved by Stanford University Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. The animal study was approved by Laboratory Animal Care (APLAC) at Stanford University. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

ML: Formal analysis, Visualization, Writing - original draft, Writing - review & editing. SA: Formal analysis, Visualization, Writing - review & editing. DW: Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. MG: Data curation, Writing - review & editing. FD: Methodology, Writing - review & editing. CP: Formal analysis, Writing - review & editing. AP: Data curation, Project administration, Writing - review & editing. CS: Investigation, Writing - review & editing. YH: Investigation, Writing review & editing. LB: Formal analysis, Writing - review & editing. RG: Formal analysis, Writing - review & editing. PC: Methodology, Writing - review & editing. AE: Data curation, Writing - review & editing. JP: Methodology, Writing review & editing. HC: Data curation, Writing - review & editing. MP: Conceptualization, Investigation, Writing - review & editing. SD: Conceptualization, Investigation, Writing - review & editing. EM: Conceptualization, Investigation, Writing - review & editing. TC: Conceptualization, Investigation, Writing review & editing. GD: Conceptualization, Investigation, Writing review & editing. EA: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing review & editing. MW: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. Parts of this work were supported by NIH grants R01 HL108943 (GWDII), R01 HL128075 (EMM), T32 HL094274 (FED, AE), K12 HD000850 (JRP), U01 HG007708 (MTW, EAA), the Knut and Alice Wallenberg Foundation (MEL), Graciano Family Fund (MW), and unrestricted grants from the Li Ka-shing Foundation, Breetwor Foundation, and SP (DW, AP, EA).

Acknowledgments

We thank the subjects for their participation.

Conflict of interest

EA is a founder of Personalis, Inc, DeepCell, Inc, Svexa Inc., a founding advisor of Nuevocor, a non-executive director at AstraZeneca, an advisor to SequenceBio, Novartis, Medical Excellence Capital and Foresite Capital. MW, PC, FD hold equity in Personalis. MW has past and current clinical research support from Novartis, Cytokinetics, BMS, Alexion, and Salubris Bio and has been a consultant to BMS and Cytokinetics. Current company employees; DW -Svexa; RG -AstraZeneca; PC -Stripe; JP -Tenaya Therapeutics; AE -ZS; FD -Versant Ventures. EMM has been a consultant to Amgen, AstraZeneca, Avidity, Cytokinetics, PepGen, Pfizer and Tenaya Therapeutics, and is founder of Ikaika Therapeutics and Kardigan Therapeutics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N Engl J Med.* (1989) 321:1372–8. doi: 10.1056/NEJM198911163212005

2. Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, et al. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell.* (1990) 62:999–1006. doi: 10.1016/0092-8674(90)90274-I

3. Walsh R, Mazzarotto F, Whiffin N, Buchan R, Midwinter W, Wilk A, et al. Quantitative approaches to variant classification increase the yield and precision of genetic testing in Mendelian diseases: the case of hypertrophic cardiomyopathy. *Genome Med.* (2019) 11:5. doi: 10.1186/s13073-019-0616-z

 Maron BJ, Maron MS. Hypertrophic cardiomyopathy. Lancet. (2013) 381:242–55. doi: 10.1016/S0140-6736(12)60397-3

5. Lopes LR, Zekavati A, Syrris P, Hubank M, Giambartolomei C, Dalageorgou C, et al. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. *J Med Genet.* (2013) 50:228–39. doi: 10.1136/jmedgenet-2012-101270

6. Harper AR, Parikh VN, Goldfeder RL, Caleshu C, Ashley EA. Delivering clinical grade sequencing and genetic test interpretation for cardiovascular medicine. *Circ Cardiovasc Genet.* (2017) 10(2):e001221. doi: 10.1161/CIRCGENETICS.116.001221

7. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. N Engl J Med. (2011) 364:1643–56. doi: 10.1056/NEJMra0902923

8. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, et al. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nat Genet.* (2021) 53(2):128–34. doi: 10.1038/s41588-020-00762-2

9. Maron BJ, Maron MS, Semsarian C. Double or compound sarcomere mutations in hypertrophic cardiomyopathy: a potential link to sudden death in the absence of conventional risk factors. *Heart Rhythm.* (2012) 9:57–63. doi: 10.1016/j.hrthm.2011. 08.009

10. Girolami F, Ho CY, Semsarian C, Baldi M, Will ML, Baldini K, et al. Clinical features and outcome of hypertrophic cardiomyopathy associated with triple sarcomere protein gene mutations. *J Am Coll Cardiol.* (2010) 55:1444–53. doi: 10. 1016/j.jacc.2009.11.062

11. Zou Y, Wang J, Liu X, Wang Y, Chen Y, Sun K, et al. Multiple gene mutations, not the type of mutation, are the modifier of left ventricle hypertrophy in patients with hypertrophic cardiomyopathy. *Mol Biol Rep.* (2013) 40:3969–76. doi: 10.1007/s11033-012-2474-2

12. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, et al. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet.* (2021) 53(2):135–42. doi: 10.1038/s41588-020-00764-0

13. Bagnall RD, Ingles J, Dinger ME, Cowley MJ, Ross SB, Minoche AE, et al. Whole genome sequencing improves outcomes of genetic testing in patients with hypertrophic cardiomyopathy. *J Am Coll Cardiol.* (2018) 72:419–29. doi: 10.1016/j. jacc.2018.04.078

14. Kontorovich AR, Patel N, Moscati A, Richter F, Peter I, Purevjav E, et al. Myopathic cardiac genotypes increase risk for myocarditis. *JACC Basic Transl Sci.* (2021) 6:584–92. doi: 10.1016/j.jacbts.2021.06.001

15. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. (2020) 581:434–43. doi: 10.1038/s41586-020-2308-7

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2025. 1546493/full#supplementary-material

16. Pan S, Caleshu CA, Dunn KE, Foti MJ, Moran MK, Soyinka O, et al. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. *Circ Cardiovasc Genet.* (2012) 5:602–10. doi: 10. 1161/CIRCGENETICS.112.963421

17. Golbus JR, Puckelwartz MJ, Fahrenbach JP, Dellefave-Castillo LM, Wolfgeher D, McNally EM. Population-based variation in cardiomyopathy genes. *Circ Cardiovasc Genet.* (2012) 5:391–9. doi: 10.1161/CIRCGENETICS.112.962928

18. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, et al. ACMG Recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med.* (2008) 10:294–300. doi: 10.1097/GIM. 0b013e31816b5cae

19. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–24. doi: 10.1038/gim.2015.30

20. Cleary JG, Braithwaite R, Gaastra K, Hilbush BS, Inglis S, Irvine SA, et al. Joint variant and *de novo* mutation identification on pedigrees from high-throughput sequencing data. *J Comput Biol.* (2014) 21:405–19. doi: 10.1089/cmb.2014.0029

21. Priest JR, Gawad C, Kahlig KM, Yu JK, O'Hara T, Boyle PM, et al. Early somatic mosaicism is a rare cause of long-QT syndrome. *Proc Natl Acad Sci U S A*. (2016) 113:11555–60. doi: 10.1073/pnas.1607187113

22. Priest JR, Osoegawa K, Mohammed N, Nanda V, Kundu R, Schultz K, et al. *de novo* and rare variants at multiple loci support the oligogenic origins of atrioventricular septal heart defects. *PLoS Genet.* (2016) 12:e1005963. doi: 10.1371/ journal.pgen.1005963

23. Fuentes Fajardo KV, Adams D, NISC Comparative Sequencing Program, Mason CE, Sincan M, Tifft C, et al. Detecting false-positive signals in exome sequencing. *Hum Mutat.* (2012) 33:609–13. doi: 10.1002/humu.22033

24. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. (2012) 489:57–74. doi: 10.1038/nature11247

25. Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*. (2014) 30:2843–51. doi: 10.1093/bioinformatics/ btu356

26. Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, et al. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nat Biotechnol.* (2014) 32:246–51. doi: 10.1038/nbt.2835

27. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* (2014) 42:D1001-6. doi: 10.1093/nar/gkt1229

28. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther.* (2012) 92:414–7. doi: 10.1038/clpt.2012.96

29. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, et al. The catalogue of somatic mutations in cancer (COSMIC). *Curr Protoc Hum Genet.* (2008). Chapter 10:Unit 10.11 57(1):10.11.1-26. doi: 10.1002/0471142905. hg1011s57

30. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* (*Austin*). (2012) 6:80–92. doi: 10.4161/fly.19695

31. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* (2014) 46:310–5. doi: 10.1038/ng.2892

32. Fan X, Abbott TE, Larson D, Chen K. Breakdancer: identification of genomic structural variation from paired-end read mapping. *Curr Protoc Bioinformatics*. (2014) 45:15.6.1–11. doi: 10.1002/0471250953.bi1506s45

33. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* (2011) 21:974–84. doi: 10.1101/gr. 114876.110

34. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. (2012) 28:i333–9. doi: 10.1093/bioinformatics/bts378

35. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, et al. Control-FREEC: a tool for assessing copy number and allelic content using nextgeneration sequencing data. *Bioinformatics*. (2012) 28:423–5. doi: 10.1093/ bioinformatics/btr670

36. Lam HYK, Mu XJ, Stütz AM, Tanzer A, Cayting PD, Snyder M, et al. Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library. *Nat Biotechnol.* (2010) 28:47–55. doi: 10.1038/nbt.1600

37. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* (2019) 48:D498–503. doi: 10. 1093/nar/gkz1031

38. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

39. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* (2004) 5:R80. doi: 10.1186/gb-2004-5-10-r80

40. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet.* (2012) 91:224–37. doi: 10.1016/j.ajhg.2012.06.007

41. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature*. (2013) 493:216–20. doi: 10.1038/nature11690

42. Dewey FE, Perez MV, Wheeler MT, Watt C, Spin J, Langfelder P, et al. Gene coexpression network topology of cardiac development, hypertrophy, and failure. *Circ Cardiovasc Genet*. (2011) 4:26–35. doi: 10.1161/CIRCGENETICS.110.941757

43. GTEx Consortium. The genotype-tissue expression (GTEx) project. Nat Genet. (2013) 45:580-5. doi: 10.1038/ng.2653

44. Zaleta-Rivera K, Dainis A, Ribeiro AJS, Cordero P, Rubio G, Shang C, et al. Allele-specific silencing ameliorates restrictive cardiomyopathy attributable to a human myosin regulatory light chain mutation. *Circulation*. (2019) 140:765–78. doi: 10.1161/CIRCULATIONAHA.118.036965

45. Simpson P. Norepinephrine-stimulated hypertrophy of cultured rat myocardial cells is an alpha 1 adrenergic response. *J Clin Invest.* (1983) 72:732–8. doi: 10.1172/JCI111023

46. Eble DM, Qi M, Waldschmidt S, Lucchesi PA, Byron KL, Samarel AM. Contractile activity is required for sarcomeric assembly in phenylephrine-induced cardiac myocyte hypertrophy. *Am J Physiol Cell Physiol.* (1998) 274:C1226–37. doi: 10.1152/ajpcell.1998.274.5.C1226

47. Lin Y-H, Warren CM, Li J, McKinsey TA, Russell B. Myofibril growth during cardiac hypertrophy is regulated through dual phosphorylation and acetylation of the actin capping protein CapZ. *Cell Signal.* (2016) 28:1015–24. doi: 10.1016/j. cellsig.2016.05.011

48. Hartman TJ, Martin JL, Solaro RJ, Samarel AM, Russell B. Capz dynamics are altered by endothelin-1 and phenylephrine via PIP2- and PKC-dependent mechanisms. *Am J Physiol Cell Physiol*. (2009) 296:C1034–9. doi: 10.1152/ajpcell.00544.2008

49. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635

50. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat Protoc.* (2012) 7:562–78. doi: 10.1038/nprot.2012.016

51. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, et al. Clinvar: improvements to accessing data. *Nucleic Acids Res.* (2020) 48:D835-44. doi: 10. 1093/nar/gkz972

52. Valdés-Mas R, Gutiérrez-Fernández A, Gómez J, Coto E, Astudillo A, Puente DA, et al. Mutations in filamin C cause a new form of familial hypertrophic cardiomyopathy. *Nat Commun.* (2014) 5:5326. doi: 10.1038/ncomms6326

53. Deo RC, Musso G, Tasan M, Tang P, Poon A, Yuan C, et al. Prioritizing causal disease genes using unbiased genomic features. *Genome Biol.* (2014) 15:534. doi: 10. 1186/s13059-014-0534-8

54. Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, et al. RBM20, A gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med.* (2012) 18:766–73. doi: 10.1038/nm.2693

55. Schoenauer R, Lange S, Hirschy A, Ehler E, Perriard J-C, Agarkova I. Myomesin 3, a novel structural component of the M-band in striated muscle. *J Mol Biol.* (2008) 376:338–51. doi: 10.1016/j.jmb.2007.11.048

56. Brauch KM, Karst ML, Herron KJ, de Andrade M, Pellikka PA, Rodeheffer RJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. J Am Coll Cardiol. (2009) 54:930–41. doi: 10.1016/j.jacc.2009.05.038

57. Norton N, Li D, Rieder MJ, Siegfried JD, Rampersaud E, Züchner S, et al. Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. *Am J Hum Genet.* (2011) 88:273–82. doi: 10.1016/j.ajhg.2011.01.016

58. Miller DT, Lee K, Abul-Husn NS, Amendola LM, Brothers K, Chung WK, et al. ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* (2022) 24:1407–14. doi: 10.1016/j. gim.2022.04.006

59. Lindholm ME, Jimenez-Morales D, Zhu H, Seo K, Amar D, Zhao C, et al. Monoand biallelic protein-truncating variants in alpha-actinin 2 cause cardiomyopathy through distinct mechanisms. *Circ Genom Precis Med.* (2021) 14:e003419. doi: 10. 1161/CIRCGEN.121.003419

60. Han P, Li W, Lin C-H, Yang J, Shang C, Nuernberg ST, et al. A long noncoding RNA protects the heart from pathological hypertrophy. *Nature*. (2014) 514:102–6. doi: 10.1038/nature13596

61. Kelly MA, Caleshu C, Morales A, Buchan J, Wolf Z, Harrison SM, et al. Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: recommendations by ClinGen's inherited cardiomyopathy expert panel. *Genet Med.* (2018) 20:351–9. doi: 10.1038/gim.2017.218

62. Zhang Q, Chen J, Qin Y, Wang J, Zhou L. Mutations in voltage-gated L-type calcium channel: implications in cardiac arrhythmia. *Channels*. (2018) 12:201–18. doi: 10.1080/19336950.2018.1499368

63. González-Mariscal L, Gallego-Gutiérrez H, González-González L, Hernández-Guzmán C. ZO-2 is a master regulator of gene expression, cell proliferation, cytoarchitecture, and cell size. *Int J Mol Sci.* (2019) 20(17):4128. doi: 10.3390/ ijms20174128

64. Arimura T, Takeya R, Ishikawa T, Yamano T, Matsuo A, Tatsumi T, et al. Dilated cardiomyopathy-associated FHOD3 variant impairs the ability to induce activation of transcription factor serum response factor. *Circ J.* (2013) 77:2990–6. doi: 10.1253/circj.CJ-13-0255

65. Ochoa JP, Sabater-Molina M, García-Pinilla JM, Mogensen J, Restrepo-Córdoba A, Palomino-Doza J, et al. Formin homology 2 domain containing 3 (FHOD3) is a genetic basis for hypertrophic cardiomyopathy. *J Am Coll Cardiol.* (2018) 72:2457–67. doi: 10.1016/j.jacc.2018.10.001

66. Wooten EC, Hebl VB, Wolf MJ, Greytak SR, Orr NM, Draper I, et al. Formin homology 2 domain containing 3 variants associated with hypertrophic cardiomyopathy. *Circ Cardiovasc Genet.* (2013) 6:10–8. doi: 10.1161/ CIRCGENETICS.112.965277

67. Ding Y, Yang J, Chen P, Lu T, Jiao K, Tester DJ, et al. Knockout of SORBS2 protein disrupts the structural integrity of intercalated disc and manifests features of arrhythmogenic cardiomyopathy. *J Am Heart Assoc.* (2020) 9:e017055. doi: 10.1161/JAHA.119.017055

68. Martino F, Varadarajan NM, Perestrelo AR, Hejret V, Durikova H, Vukic D, et al. The mechanical regulation of RNA binding protein hnRNPC in the failing heart. *Sci Transl Med.* (2022) 14:eabo5715. doi: 10.1126/scitranslmed. abo5715

69. D'Uva G, Aharonov A, Lauriola M, Kain D, Yahalom-Ronen Y, Carvalho S, et al. ERBB2 Triggers mammalian heart regeneration by promoting cardiomyocyte dedifferentiation and proliferation. *Nat Cell Biol.* (2015) 17:627–38. doi: 10.1038/ncb3149

70. Wadugu B, Kühn B. The role of neuregulin/ErbB2/ErbB4 signaling in the heart with special focus on effects on cardiomyocyte proliferation. *Am J Physiol Heart Circ Physiol.* (2012) 302:H2139–47. doi: 10.1152/ajpheart.00063.2012

71. Odiete O, Hill MF, Sawyer DB. Neuregulin in cardiovascular development and disease. *Circ Res.* (2012) 111:1376–85. doi: 10.1161/CIRCRESAHA.112. 267286

72. Seemann E, Sun M, Krueger S, Tröger J, Hou W, Haag N, et al. Deciphering caveolar functions by syndapin III KO-mediated impairment of caveolar invagination. *Elife*. (2017) 6:e29854. doi: 10.7554/eLife.29854

73. Bi H-L, Zhang X-L, Zhang Y-L, Xie X, Xia Y-L, Du J, et al. The deubiquitinase UCHL1 regulates cardiac hypertrophy by stabilizing epidermal growth factor receptor. *Sci Adv.* (2020) 6(16):eaax4826. doi: 10.1126/sciadv.aax4826

74. Yamada N, Asano Y, Fujita M, Yamazaki S, Inanobe A, Matsuura N, et al. Mutant KCNJ3 and KCNJ5 potassium channels as novel molecular targets in bradyarrhythmias and atrial fibrillation. *Circulation*. (2019) 139:2157-69. doi: 10. 1161/CIRCULATIONAHA.118.036761 75. Cirino AL, Lakdawala NK, McDonough B, Conner L, Adler D, Weinfeld M, et al. A comparison of whole genome sequencing to multigene panel testing in hypertrophic cardiomyopathy patients. *Circ Cardiovasc Genet.* (2017) 10:e001768. doi: 10.1161/CIRCGENETICS.117.001768

76. Tonry C, Linden K, Collier P, Ledwidge M, McDonald K, Collins BC, et al. Proteomic characterisation of heart failure reveals a unique molecular phenotype for hypertrophic cardiomyopathy. *Biomedicines.* (2024) 12:1712. doi: 10.3390/biomedicines12081712

77. Garmany R, Bos JM, Dasari S, Johnson KL, Tester DJ, Giudicessi JR, et al. Proteomic and phosphoproteomic analyses of myectomy tissue reveals difference between sarcomeric and genotype-negative hypertrophic cardiomyopathy. *Sci Rep.* (2023) 13:14341. doi: 10.1038/s41598-023-40795-1