Title

Transcriptomic and genetic profiling in a spontaneous non-human primate model of hypertrophic cardiomyopathy and sudden cardiac death

Journal

Scientific Reports, 14(1)

ISSN

2045-2322

Authors

Rivas, Victor N

Vandewege, Michael W

Ueda, Yu

et al.

Publication Date

2024-12-01

DOI

10.1038/s41598-024-82770-4

Peer reviewed

scientific reports

Check for updates

OPEN

Transcriptomic and genetic profiling in a spontaneous non-human primate model of hypertrophic cardiomyopathy and sudden cardiac death

Victor N. Rivas[1,2], Michael W. Vandewege[1], Yu Ueda[1], Joanna L. Kaplan[2], JRachel Reader[3], Jeffrey A. Roberts[3] & Joshua A. Stern[1,2]✉

Hypertrophic cardiomyopathy (HCM) afflicts humans, cats, pigs, and rhesus macaques. Disease sequelae include congestive heart failure, thromboembolism, and sudden cardiac death (SCD). Sarcomeric mutations explain some human and cat cases, however, the molecular basis in rhesus macaques remains unknown. RNA-Seq of the LV tissues of five HCM-affected and seven healthy control rhesus macaques was employed for differential transcriptomic analyses. DNA from 15 severely HCM-affected and 21 healthy geriatric rhesus macaques were selected for whole-genome sequencing. A genome-wide association study (GWAS) of disease status and SCD outcome was performed. 614 down- and 1,065 upregulated differentially expressed genes (DEGs) were identified between groups. The top DEG (*MAFF*) was overexpressed in affected animals (log2FoldChange = 4.71; $P_{Adjusted}$-value = 1.14E-133). Channelopathy-associated enriched terms were identified in ~ 57% of downregulated DEGs providing transcriptomic evidence of hypertrophic and arrhythmic disease processes. For GWAS, no putative variant withstood segregation. Polygenic modeling analysis resulted in poor prediction power and burden testing could not explain HCM by an association of multiple variants in any gene. Neither single nor compound genetic variant(s), or identified polygenic profile, suggest complex genotype–phenotype interactions in rhesus macaques. Brought forth is an established dataset of robustly phenotyped rhesus macaques as an open-access resource for future cardiovascular disease genetic studies.

Keywords Whole-genome sequencing, RNA-sequencing, Multi-omics, Monkey, Large-animal model, Arrythmia

Hypertrophic cardiomyopathy (HCM) is an adult-onset disease that affects approximately 1 in 500 humans, 1 in 7 cats, up to 1 in 4 landrace pigs of a closed population, and rarely dogs[1–5]. HCM is largely considered a disease of hypercontractile cardiac sarcomeres: the contractile and most basic functional units of the cardiac muscle. The disease is principally characterized by cardiomyocyte hypertrophy and often accompanied by an increase of myocardial interstitial fibrosis, resulting in decreased left ventricular (LV) luminal volume and increased LV wall thickness and mass that contributes to impaired diastolic filling (diastolic dysfunction), increased systolic function, and, in some cases, left ventricular outflow tract obstruction (LVOTO). Patients with suspected HCM are diagnosed antemortem by echocardiographic imaging or postmortem via gross pathology and histology. In humans, disease sequelae range from mild, with many patients experiencing shortness of breath or exercise intolerance due to LVOTO, to more severe outcomes such as malignant cardiac arrhythmias, thromboembolic complications, congestive heart failure (CHF), or sudden cardiac death (SCD); HCM-mediated SCD being one of the most common clinical manifestations in young adults[6]. The disease is heritable and most frequently reported to display an autosomal dominant mode of inheritance (MOI). In a 25-year period, over 1,500 HCM-

[1]Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, 1060 William Moore Dr, Raleigh, NC 27607, USA. [2]Department of Medicine & Epidemiology, School of Veterinary Medicine, University of California-Davis, Davis, CA, USA. [3]California National Primate Research Center, University of California-Davis, Davis, CA, USA. ✉email: jastern@ncsu.edu

associated mutations in sarcomeric, calcium handling, and z-disk genes have been identified in human patients; of these, *Myosin Heavy Chain 7* (*MYH7*), *Myosin Binding Protein-C* (*MYBPC3*), and *Cardiac Troponin T* (*TTNT2*) variants are the most frequently reported[7,8]. Despite a great number of known HCM-associated genetic variants, approximately 40–50% of patients with HCM have no identifiable genetic cause, creating a knowledge gap that limits the treatment and prevention of HCM in human medicine[9,10].

In cats, the disease has been observed in over 30 breeds and remains common in mixed breeds[2,11]. Candidate-gene-led feline genetic studies of HCM have identified the breed specific A31P and R820W *MYBPC3* autosomal dominant variants in Maine Coon and Ragdoll cats, respectively, and the autosomal dominant G3376R *Alström Syndrome-1 Centrosome and Basal Body Associated Protein* (*ALSM1*) mutation in Sphynx cats[12–16]. These mutations remain breed specific and do not explain the disease in other pure or mixed-breed cats. To study the molecular etiology of HCM for pharmacologic therapies, an A31P-harboring cat colony has been established. Studies revealed that the A31P mutation in this cat colony is incompletely penetrant as not all A31P-positive cats have phenotypic evidence of HCM and a small number of A31P-negative cats have presented with HCM, despite their controlled environment[17]. This finding reduces the ability to reliably produce HCM-affected and control cats within the colony and highlights the continued need for defining additional model organisms, particularly large-animal, for human HCM.

Naturally occurring familial HCM at the California National Primate Research Center (CNPRC) has been reported at a prevalence of 1.3% (1 in 100) on post-mortem examinations over a 20-year period in rhesus macaques (*Macaca mulatta*)[18]. During this period, screening within the nuclear families of SCD cases has identified over 160 HCM-confirmed cases in the CNPRC bringing the prevalence to as high as 29% if a familial relationship to HCM has already been identified. Echocardiographic and post-mortem evaluations of HCM-affected rhesus macaques revealed that the disease closely recapitulates human HCM as these animals have largely symmetrical, frequently severe, LV wall thickening with intermittent LVOTO (Fig. 1). Like humans with HCM, the recorded SCD events are often precipitated by exercise or physiologic stress, a disease outcome that no other small or large animal model has yet to reliably represent[19–22]. Pedigree analysis in the rhesus macaque colony demonstrates that the disease is heritable ($h^2 = 0.50$) and most likely manifests in an autosomal recessive MOI, although an autosomal dominant MOI with incomplete penetrance cannot be excluded[23,24]. Electrocardiographic (ECG) examination demonstrated patterns of LV hypertrophy and altered heart rate variability (HRV), as well as intermittent ventricular arrhythmias[25].

Small animal models of HCM present challenges for direct translation due to their failures in recreating the human disease phenotype, particularly in disease outcomes such as SCD. Genetically engineered HCM rodent models exhibit inherent differences in cardiac physiology, hemodynamics, and gene-disease linkage, which prevent their direct translation to human HCM[26–28]. These cumulative findings strongly support that a more directly translatable, naturally occurring, model of human HCM has been identified in the rhesus macaques. Establishing the first non-human primate (NHP) animal model of HCM holds promise in resolving the aforementioned physiologic discrepancies of genetically engineered rodent models and facilitates translational discoveries in this disease. Genetic characterization and further model development are crucial to foster investigations of novel therapies which hold promise for treating this devastating disease. The use of a NHP model for HCM may aid in bridging the knowledge gaps in human and feline HCM prevention, and paves the

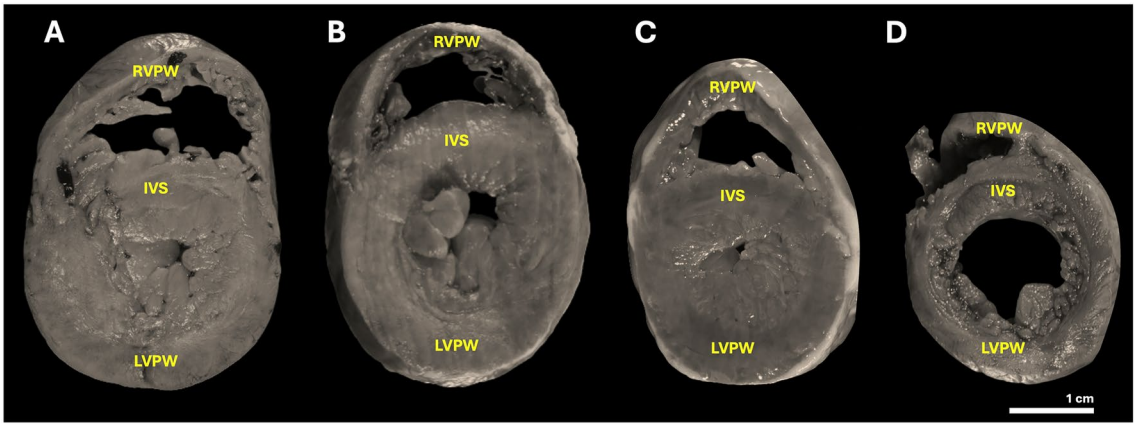

Fig. 1. Gross pathologic images of rhesus macaque left ventricles. Representative gross pathologic images of a four-year-old male (**A**), a six-year-old female (**B**), and a 12-year-old female (**C**) rhesus macaque affected with severe HCM, and a nine-year-old cardiovascularly normal female control (**D**) rhesus macaque are displayed (n = 4). Hearts were cross-sectionally cut midway between the apex and base of the heart at the level of the papillary muscles in the LV. Individuals **A** and **C** were obtained following a spontaneous SCD event. Note severe LVH and complete luminal obliteration in the HCM-affected individuals. Scale bar = 1 cm. RVPW = right ventricular posterior wall, IVS = interventricular septum, LVPW = left ventricular posterior wall, HCM = hypertrophic cardiomyopathy, LV = left ventricle, SCD = sudden cardiac death, LVH = left ventricular hypertrophy.

way for novel drug development, but with the molecular basis and genetic cause of rhesus macaque HCM left unknown, the use of this ideal HCM model remains limited.

To date, no pathogenic mutations causing HCM in rhesus macaques have been reported. However, prior research involving amplicon targeted-sequencing of sarcomeric genes known to harbor human HCM variants in a population of rhesus macaques (n = 94) has been performed, revealing a risk haplotype-block in *MYBPC3* in association with disease status with an implicated downstream autosomal dominant intronic variant postulated to contribute to an alternate isoform of the gene[29]. Despite this finding, expansion of this genotyping in additional animals at the CNPRC failed to confirm segregation with disease. The aim of this present study was to investigate the underlying genetic etiology of rhesus macaque HCM in the CNPRC. RNA-Sequencing (RNA-Seq) of the LV posterior wall (LVPW) tissue samples from five HCM-affected and seven sex- and weight-matched cardiovascularly normal post-pubescent adolescent (3–10 years-of-age) rhesus macaques (n = 12) was employed to evaluate and characterize the transcriptomic profile of rhesus HCM. Additionally, a genome-wide association study (GWAS) involving the use of whole-genome sequencing (WGS) data of 15 severely HCM-affected and 21 geriatric cardiovascularly normal control rhesus macaques (n = 36) was employed to elucidate genetic variants that may explain disease pathogenesis. It was hypothesized that inherited HCM in the CNPRC rhesus macaque colony is caused by a putative variant(s) with an identifiable effect on molecular pathways of cardiac hypertrophy.

Results

Demographic, gross pathology, and group comparisons

While all animals in the RNA-Seq cohort were representatives of rhesus macaques in the adolescent life stage, a significant difference between the ages (years) of HCM-affected 3.87 (± 0.83) and -unaffected 6.33 (± 1.28) rhesus macaques was noted (*P*-value = 0.004). Other tested variables (i.e., Sex, Housing Configuration, SPF Status, and Weight) resulted in non-statistically significant *P*-values (*P*-value = 0.52, *P*-value = 0.20, *P*-value = 0.15, *P*-value = 0.05, respectively). All RNA-Seq HCM-affected animals displayed gross pathologic evidence of left ventricular hypertrophy (LVH) as denoted by an O:I ratio > 3 and pathologist confirmation of an HCM phenotype; the hearts of RNA-Seq controls were unremarkable (Fig. 1; Table 1).

For the GWAS cohort, a significant difference in Housing Configuration was noted between cases and controls (*P*-value = 0.004). Age was statistically significant between the HCM-affected 8.13 (± 4.61) and the geriatric control 23.10 (± 3.02) population of rhesus macaques (*P*-value < 0.0001). Weight (Kg) was significantly increased in geriatric 10.99 (± 2.49) versus HCM-affected 8.06 (± 2.77) rhesus macaques (*P*-value = 0.002). No statistically significant differences between Sex and SPF Status were noted between GWAS groups (*P*-value = 0.50 and *P*-value = 0.17, respectively). All affected animals were confirmed to have a pathologist-confirmed HCM phenotype with severe LVH and LV luminal obliteration noted in the hearts of 11 of 15 HCM-affected GWAS rhesus macaques (73.33%). Ten rhesus macaques in the affected group (66.67%) experienced a SCD outcome; at time of study design and HCM phenotyping, all geriatric rhesus macaques in the control group were alive with cardiologist-confirmed normal echocardiograms, and thus, full necropsies were not performed in this group (Supplemental Table 1).

RNA-sequencing

Principal component analysis of the LVPW tissue samples revealed discrete populations of HCM-affected and -unaffected rhesus macaques; principal component 1 and 2 (PC1 and PC2) explained 52% and 17% of the observed variance within the RNA-Seq sample population, respectively (Fig. 2). A total of 1679 (614 down- and 1065 upregulated) statistically significant DEGs were identified between the LVPW of HCM-affected and -unaffected rhesus macaques. (Supplemental Table 2; Fig. 3). The 20 top-most significant up- and downregulated DEGs are reported in Table 2.

Out of 14,095 nonzero total read counts, the condition (HCM status) + age model outperformed the condition-only model in nine upregulated genes (0.064%) at a log2FoldChange > 0 (i.e., *SLC16A3*, *CD200*, *CX3CL1*, *CXCL10*, *MATCAP1*, *RALGDS*, *RASD2*, *RP1L1*, *SLCO4A1*), where one was previously significant in the condition-only model (i.e., *SLCO4A1*), suggesting age was not an important contributor to DEGs count overall (Supplemental Table 4). No differences were identified in the downregulated DEGs dataset. Given all monkeys in the RNA-Seq HCM-affected and -unaffected group would be considered adolescent (at or beyond the age of puberty, but not yet considered a mature adult) and the observed low number of differences identified in the upregulated DEG dataset when age was added to the model, the condition-only model was opted for data reporting.

Shared differentially expressed genes between rhesus macaques and human HCM

RNA-Seq of rhesus macaque LV samples revealed close molecular relationships to the transcriptomic profile of both pediatric and adult human HCM. Within the list of shared rhesus macaque and pediatric human downregulated DEGs (n = 40), disintegrin and metalloprotease protein (i.e., *ADAM11* and *ADAMTS7*), calcium voltage-gated channel (i.e., *CACNA1G* and *CACNA2D2*), potassium voltage-gated channel (i.e., *KCND3* and *KCNH3*), myosin light chain (i.e., *MYL4* and *MYLK4*), T-box transcription factor (i.e., *TBX20* and *TBX5*), proline-rich transmembrane (i.e., *PRRG3* and *PRRT2*), and zinc finger protein (i.e., *ZNF497* and *ZNF711*) gene families were commonly shared. For shared upregulated DEGs between rhesus and pediatric HCM samples (n = 71), F2R-like trypsin receptor (i.e., *F2RL1*, *F2RL2*, and *F2RL3*), four-and-a-half LIM domain (i.e., *FHL1* and *FHL5*), mitogen-activated protein kinase (i.e., *MAP2K3* and *MAP3K7CL*), pleckstrin homology domain-containing (i.e., *PLEKHA4* and *PLEKHO1*), and solute carrier (i.e., *SLC43A3*, *SLC6A17*, *SLC7A1*, and *SLC7A2*) gene families were identified. Within the list of shared rhesus macaque and adult human downregulated DEGs (n = 64), disintegrin and metalloprotease protein (i.e., *ADAM11* and *ADAMTS7*), ankyrin repeat and SOCS

	Animal	Death Type	Sex	Housing Configuration	SPF Status	Chinese Origin (%)	Indian Origin (%)	Age (yrs)	Weight (Kg)	O:I Ratio	Heart Weight (g)	IVS (mm)	LVPW (mm)
Affected	1	Spontaneous	M	Outdoor	Conventional	87.50	12.5	5.08	8.52	8.67	46.43	9	14
	2	Spontaneous	M	Outdoor	Level 1	100	0	4.08	7.42	3.22	34.85	9	9
	3	Spontaneous	F	Outdoor	Conventional	100	0	3	6.35	4.83	29.66	9	9
	4	Spontaneous	F	Outdoor	Conventional	100	0	3.17	4.2	10.5	–	7	10
	5	Culled	M	Outdoor	Level 1	100	0	4	6.89	3.5	30.48	9	12
	Average (±SD)					97.50(±5.59)	2.5(±5.59)	3.87 (±0.83)	6.68 (±1.60)	6.14 (±3.26)	35.36 (±7.73)	8.60 (±0.89)	10.80 (±2.17)
Unaffected	1	Culled	F	Outdoor	Conventional	100	0	7.83	10.94				
	2	Culled	M	Outdoor	Conventional	100	0	6.92	7.94				
	3	Culled	M	Outdoor	Conventional	100	0	7	9.95				
	4	Culled	M	Outdoor	Conventional	100	0	3.83	5.15				
	5	Culled	M	Mixed Outdoor	Conventional	100	0	6	9.9				
	6	Culled	M	Mixed Outdoor	Conventional	100	0	6.83	13.22				
	7	Culled	M	Mixed Outdoor	Conventional	100	0	5.92	9.37				
	Average (±SD)					100(±0)	0(±0)	6.33 (±1.28)	9.50 (±2.51)				
	P-Value		0.52	0.20	0.15	0.42	0.42	0.004	0.05				

Table 1. Descriptive statistics for phenotypic gross pathologic data of RNA-Seq subjects. Normality testing via Shapiro–Wilk and Kolmogorov–Smirnov tests was performed on all variables. For all categorical variables, two-sided Fisher's exact tests were performed. Unpaired parametric *t*-tests were performed on all normally distributed continuous variables, whereas Mann–Whitney *U* tests were performed for all non-normally distributed variables. *HCM* hypertrophic cardiomyopathy, *LV* left ventricular, *IVS* interventricular septum, *LVPW* left ventricular posterior wall, *F* female, *M* male, *SPF* specific pathogen free, *SD* standard deviation.

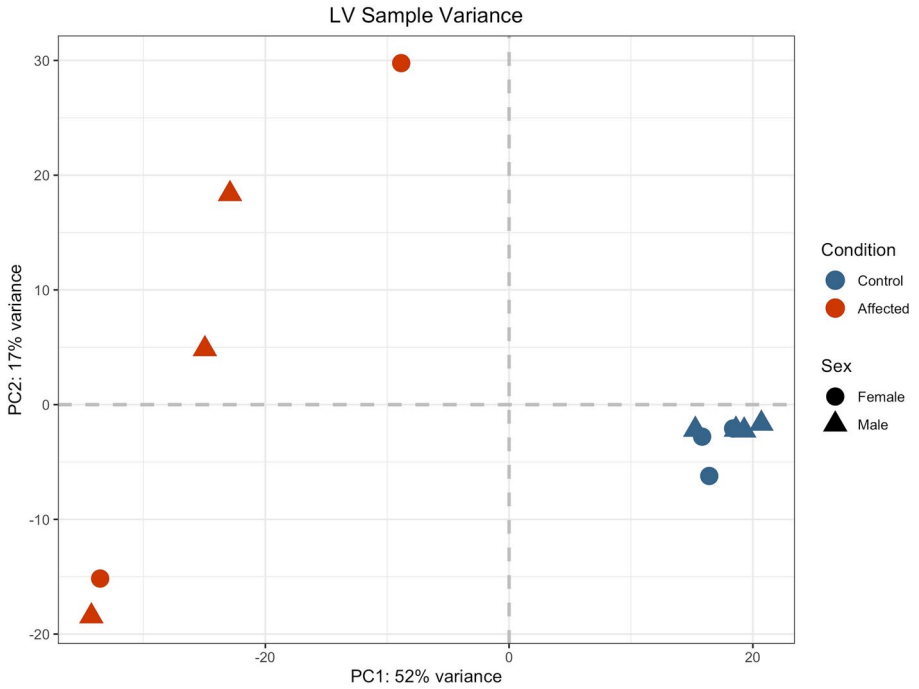

Fig. 2. RNA-Seq principal component analysis plot of rhesus macaque LV tissue. Plotted are the results of similarity within and between groups from a principal component analysis (PCA) of five HCM-affected (red) and seven HCM-unaffected (blue) female (circle) and male (triangle) rhesus macaque gene expression profiles from RNA-Seq (n = 12). PC = principal component.

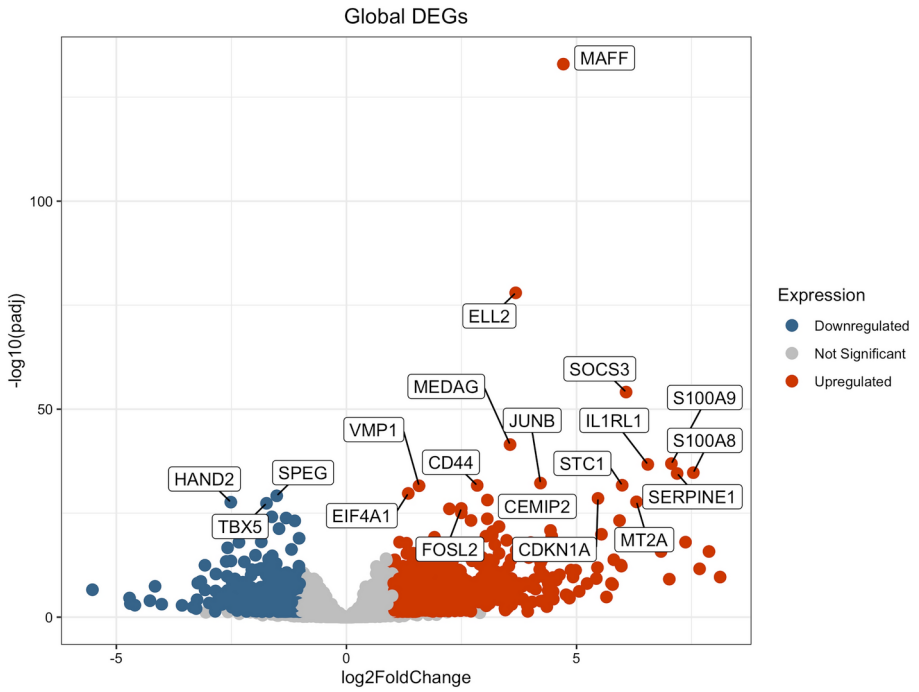

Fig. 3. Volcano plot of DEGs in HCM-affected LV tissue. Illustrated is the global expression profile of upregulated (red), downregulated (blue), and not significant (grey) genes in the LV posterior wall tissues of five HCM-affected and seven HCM-unaffected rhesus macaques (n = 12). Gene names of the 20 top-most statistically significant DEGs are displayed. HCM = hypertrophic cardiomyopathy, DEGs = differentially expressed gene(s).

	Upregulated		Downregulated		
Gene	log2FoldChange	$P_{Adjusted}$-value	Gene	log2FoldChange	$P_{Adjusted}$-value
MAFF	4.713797587	1.14E-133	*SPEG*	− 1.515402934	6.75E-30
ELL2	3.677862474	1.08E-78	*HAND2*	− 2.513135926	2.28E-28
SOCS3	6.076959329	7.79E-55	*TBX5*	− 1.736627967	4.26E-28
MEDAG	3.555756502	3.28E-42	*SYNPO2L*	− 1.622288513	7.63E-25
S100A9	7.061680893	1.23E-37	*RNF207*	− 1.310953708	1.48E-24
IL1RL1	6.548814619	1.90E-37	*CCDC8*	− 1.123348532	7.06E-24
S100A8	7.538219519	1.83E-35	*UBE2O*	− 1.465443868	5.37E-22
SERPINE1	7.186607877	2.90E-35	*ESRRB*	− 2.045032356	9.98E-22
JUNB	4.216634862	5.91E-33	*NLRX1*	− 1.028690541	1.09E-19
STC1	5.99208802	1.94E-32	*DISP1*	− 1.85081065	7.74E-19
CD44	2.841357478	2.16E-32	*CD38*	− 2.336052931	9.85E-19
VMP1	1.577762369	2.63E-32	*KY*	− 2.583249518	2.14E-17
EIF4A1	1.343120464	1.78E-30	*ZMYM3*	− 1.194510609	5.18E-17
CDKN1A	5.466099767	2.84E-29	*DPF3*	− 1.977408575	1.31E-15
CEMIP2	3.063076193	7.02E-29	*SMO*	− 1.613262449	1.80E-15
MT2A	6.302895349	1.94E-28	*KLHL31*	− 1.691583278	1.12E-14
FOSL2	2.496236194	7.43E-27	*ARHGAP40*	− 2.510534684	3.07E-14
CHIC2	2.237288592	9.14E-27	*CCNJL*	− 2.612943498	4.10E-14
MIDN	2.496776821	8.29E-26	*TPPP*	− 2.218882924	5.11E-14
STK17A	3.060610719	2.24E-24	*ENSMMUG00000045555*	− 1.525647567	1.30E-13

Table 2. Top 20 statistically significant up- and downregulated DEGs in HCM-affected rhesus macaques. Wald testing was performed for generation of $P_{Adjusted}$-values.

box (i.e., *ASB10* and *ASB15*), potassium channel (i.e., *KCND3*, *KCNJ5*, and *KCNT1*), kelch-like (i.e., *KLHL31* and *KLHL32*), and protein phosphatase 1 (i.e., *PPP1R12B* and *PPP1R3A*) gene families were observed. For shared upregulated DEGs between rhesus and adult HCM samples (n = 215), Rho GTPase-activating protein (i.e., *ARHGAP24* and *ARHGEF26*), AT-rich interaction (i.e., *ARID5A* and *ARID5B*), ATPase phospholipid transporting (i.e., *ATP10A* and *ATP8B1*), complement receptor (i.e., *C3AR1* and *C5AR1*), cellular communication network factor (i.e., *CCN2* and *CCN3*), collagen type (i.e., *COL18A1*, *COL4A1*, *COL4A2*, and *COL6A3*), ETS variant transcription factor (i.e., *ETV1* and *ETV5*), F2R-like trypsin receptor (i.e., *F2RL1* and *F2RL2*), insulin-like growth factor (i.e., *IGF1* and *IGF1R*), interleukin receptor (i.e., *IL13RA1* and *IL1R1*), inhibin subunit beta (i.e., *INHBA*, *INHBB*), Krueppel-like factor (i.e., *KLF10* and *KLF7*), kelch-like (i.e., *KLHL2* and *KLHL29*), matrix metallopeptidase (i.e., *MMP14* and *MMP2*), natriuretic peptide (i.e., *NPPA* and *NPPB*), pleckstrin homology domain (i.e., *PLEKHA4* and *PLEKHG1*), plexin (i.e., *PLXDC2* and *PLXNA2*), serpin (i.e., *SERPINE1*, *SERPINE2*, and *SERPINF1*), solute carrier (*SLC20A1*, *SLC43A2*, and *SLCO2A1*), transforming growth factor (i.e., *TGFB2* and *TGFBI*), and thrombospondin (i.e., *THBS1*, *THBS2*, and *THSD7B*) gene families were shared. A total of 10 downregulated HCM DEGs, whereas 36 upregulated HCM DEGs, were shared between the rhesus macaque and human dataset sets regardless of age group (Supplemental Table 3; Fig. 4).

Gene ontology and KEGG pathway enrichment analyses

Gene ontology biological process

Twelve statistically significant BP terms were identified from the list of downregulated DEGs: "cardiac muscle contraction", "striated muscle contraction", "potassium ion transport", "potassium ion transmembrane transport", "regulation of striated muscle contraction", "cardiac muscle cell contraction", "muscle contraction", "cardiac conduction", "photoreceptor cell maintenance", "regulation of heart contraction", "regulation of cardiac muscle contraction", and "heart contraction" (Fig. 5A). From the list of upregulated DEGS, 1035 significant GO BP terms were identified. The top-most significant of upregulated GO BP terms involved commonalities of cell movement, wounding/healing, angiogenesis, taxis, and coagulation terms (Fig. 5B).

Gene ontology molecular function

Eleven GO MF enriched terms ("monoatomic ion gated channel activity", "gated channel activity", "salt transmembrane transporter activity", "potassium ion transmembrane transporter activity", "potassium channel activity", "metal ion transmembrane transporter activity", "voltage-gated potassium channel activity", "voltage-gated monoatomic ion channel activity", "voltage-gated channel activity", "voltage-gated monoatomic cation channel activity", and "G protein-coupled receptor activity") were identified in the downregulated DEGs set (Fig. 5C). In the upregulated DEGs set, 58 significant GO MF enriched terms were identified; terms relative to cell adhesion, receptor/inhibitor activity, and binding were most significant and commonly observed (Fig. 5D).

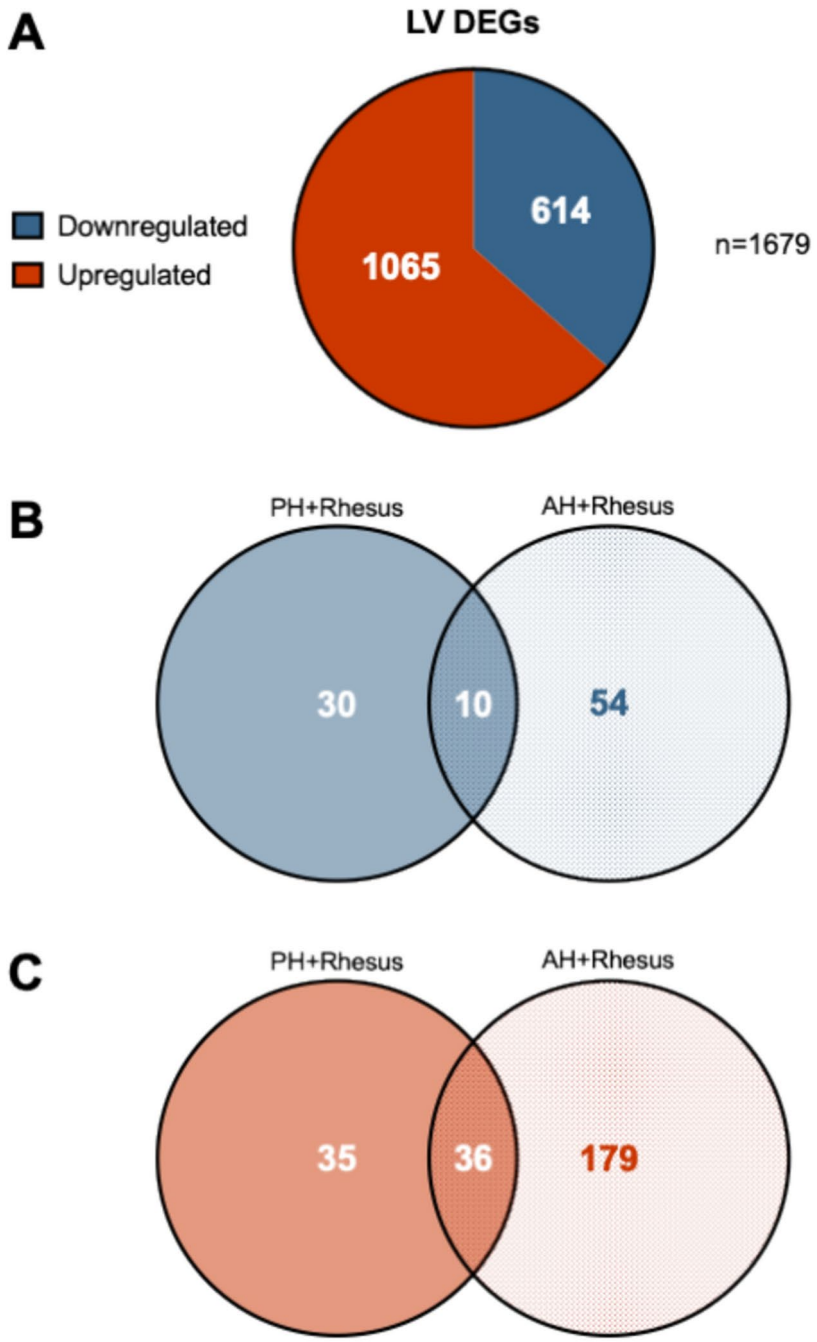

Fig. 4. Summary of total identified and shared rhesus macaque and pediatric and adult human HCM DEGs. The total number of identified down- (blue) and upregulated (red) LV DEGs in HCM-affected rhesus macaques identified via RNA-Seq are presented (**A**). Venn diagrams of shared HCM DEGs between rhesus macaque and pediatric human (PH; solid color) and adult human (AH; dot-patterned) from previously reported datasets are displayed (**B** and **C**, respectively). LV = left ventricle, DEGs = differentially expressed genes, PH = pediatric human, AH = adult human, RNA-Seq = RNA-Sequencing, HCM = hypertrophic cardiomyopathy.

Gene ontology cellular compartment

No significant GO CC terms were identified in the set of downregulated DEGs between HCM-affected and -unaffected rhesus; however, 66 significant GO CC terms were identified in the upregulated DEGs set. The top-most upregulated GO CC enriched terms consisted of extracellular matrix, membrane, structure, lumen, and granule terms (Fig. 5E).

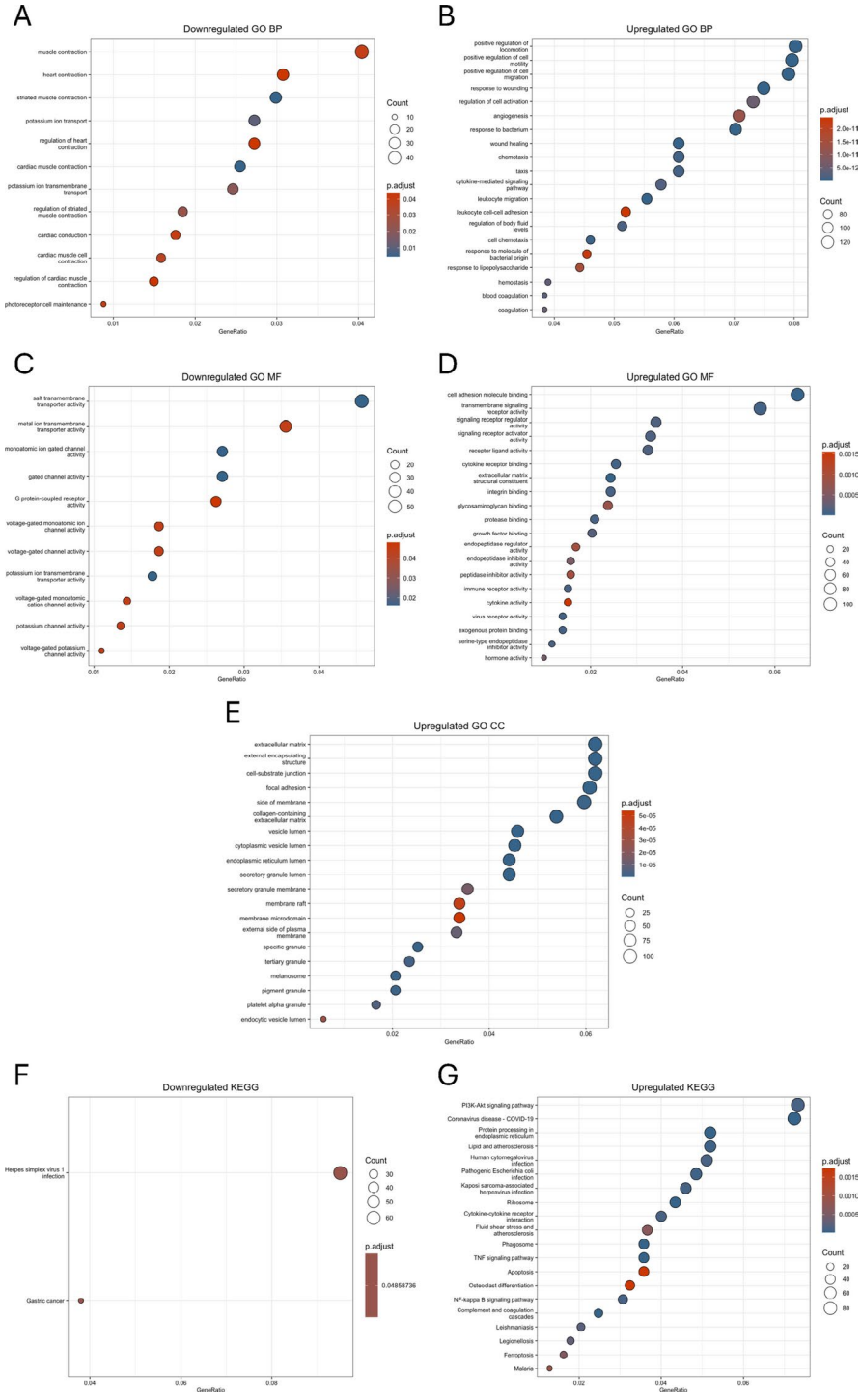

Fig. 5. Top 20 enriched GO and KEGG pathway terms for LVPW DEGs between HCM-affected and -unaffected rhesus macaques. Enrichment plots for gene ontology (GO) biological process (BP) term analyses are presented for down- and upregulated DEGs (**A** and **B**, respectively). Enrichment plots for GO molecular function (MF) term analyses are presented for down- (**C**) and upregulated (**D**) DEGs. Enrichment plot for GO cellular component (CC) terms for upregulated DEGs is presented (**E**). KEGG[30–32] pathway enrichment plots are provided for down- and upregulated DEGs (**F** and **G**, respectively). The y-axis on each graph lists identified enriched terms, whereas the gene ratio is denoted on the x-axis. The size of each individual point depicts the total number of genes binned to a given enriched term; high or low P_{Adjusted}-values are represented by red or blue coloring, respectively. Enrichment terms were obtained from the gene expression profile data of five and seven HCM-affected and -unaffected rhesus macaques, respectively (n = 12). GO = gene ontology, LVPW = left ventricular posterior wall, DEG = differentially expressed gene, HCM = hypertrophic cardiomyopathy, BP = biological process, MF = molecular function, CC = cellular component.

KEGG pathways

Two significant KEGG[30–32] terms ("herpes simplex virus 1 infection", and "gastric cancer") were identified in the downregulated DEGs set (Fig. 5F). Sixty-one significant KEGG pathway[30–32] enriched terms were identified for upregulated DEGs; the top-most significant terms involved pathways related to parasitic/bacterial/viral infection, ribosome, phagosome, TNF signaling, apoptosis, and complement and coagulation cascades (Fig. 5G).

Whole-genome sequencing

All genomic DNA (gDNA) samples were successfully sequenced and mapped at an average of 24.20X across all positions including those with zero depth. A total of 64,035,993 Freebayes-called variants were identified in the 36 rhesus macaques.

Autosomal recessive MOI

Of the 64,035,993 Freebayes-called variants, 7360 variants with a $P_{Allelic}$-value < 0.00011 were filtered and used for subsequent filtering by predicted variant impact on protein function. A total of five 'HIGH' and 14 'MODERATE' predicted effect variants were identified (Supplemental File 1). The filtered VCF files were manually inspected for variant segregation validation according to animal disease status. All filtered statistically significant autosomal recessive variants failed to segregate according to HCM status, and thus, were deemed of unknown significance in the pathogenesis of rhesus macaque HCM and were excluded from consideration as sole causative variants.

Autosomal dominant MOI

A total of 18,456 variants with a $P_{Allelic}$-value < 0.00031 were successfully filtered from the 64,035,993 Freebayes-called variants and used for subsequent filtering by predicted variant impact on protein function. Eight 'HIGH' and 55 'MODERATE' effect variants were identified (Supplemental File 1). Manual interrogation of the resulting VCF file revealed no segregating genetic variant(s) to HCM either due to 1) more than four alternate alleles present in the control population and/or 2) the alternate allele missing in one or more of HCM-affected rhesus macaque. As such, these variants were deemed of unknown significance in the pathogenesis of rhesus macaque HCM and were excluded from consideration as sole causative variants.

Manhattan plot analysis of whole-genome variants

Inability to detect a single segregating 'HIGH'/'MODERATE' variant in either autosomal dominant or recessive MOI scenarios prompted continued efforts in identifying candidate segregating variants regardless of predicted variant effect. This is visualized by Manhattan plot (Fig. 6). Twelve variants that withstood Bonferroni genome-wide significance ($P_{Allelic}$-value = 2.49E-8; − log10[2.49E-8] = 7.60) were identified. However, when additional neighboring variants positioned 5 kilobases (Kb) up- and downstream (i.e., 10 Kb span) of the Bonferroni-significant variant positions were manually inspected, the chromosome-level Manhattan plots revealed absence of leading trails in these variants, suggestive of sporadic false-positive variants, and were subsequently excluded from the study. Instead, additional variants with true trailing peak regions surpassing suggestive genome-wide significance ($P_{Allelic}$-value = 0.00001; − log10[0.00001] = 5) were subjectively evaluated. A total of 10 candidate trailing regions were visible on Chr6 (1), Chr11 (1), Chr12 (1), Chr13 (1), Chr10 (2), Chr16 (2), and ChrX (2) (Supplemental Fig. 1). The positions of the top-most significant variants of each associated peak were obtained and manually inspected to verify the relative proximity to other trailing variants; top-most significant peak variants with < 3 trailing variants within a 10 Kb genomic span (not including the top-most significant variant) that did not meet suggestive genome-wide threshold were excluded from the study. Of the ten candidate trailing loci, two harbored at least three trailing variants following the most statistically significant variant within its respective region (Chr6:128,273,728 and ChrX:63,610,473) and were considered as candidates for expanded filtering analyses of genomic variants.

Raw variant filtering for HCM status

To investigate segregation of candidate variants harbored within postulated trailing regions, all variants with a $P_{Allelic}$-value < 0.00031 (n = 18,456) were filtered and subsequently used for genotypic filtering by HCM-affection status; this $P_{Allelic}$-value threshold was used as it also allows for variants with an autosomal recessive MOI segregation ($P_{Allelic}$-value < 0.00011) to be included in the analysis. Genotypes from animals in the affected group

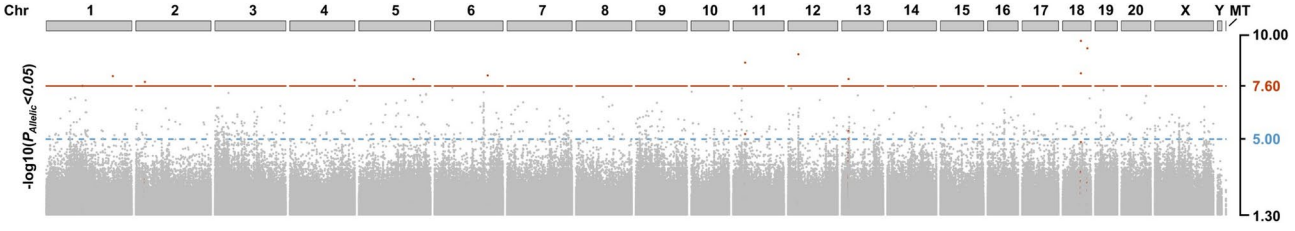

Fig. 6. Manhattan plot of global genomic $P_{Allelic}$-value < 0.05 variants. The total $P_{Allelic}$-value < 0.05 variants (n = 2,005,816) of 15 HCM-affected and 21 cardiovascularly normal geriatric rhesus macaques are plotted along their corresponding genomic position (x-axis; n = 36). The − log10($P_{Allelic}$-value) are plotted along the y-axis, where the blue- and red-dashed lines represent $P_{Allelic}$-value = 0.00001 (− log10[0.00001] = 5) and $P_{Allelic}$-value = 2.49E-8 (− log10[2.49E-8] = 7.60) significance thresholds, respectively.

were filtered as heterozygous or homozygous for the alternate allele using the 'isVariant(GEN[VCF_Header_ID])' SnpSift filtering function (v5.2). The remaining filtered variants (1092) were manually analyzed to identify segregating variants regardless of SnpEff predicted effects and genomic position (e.g., gene body, intronic, or intergenic). Filtered variants, including those harbored within candidate trailing regions (Chr6:128,273,728 and ChrX:63,610,473), failed to associate according to HCM-status regardless of MOI and were excluded from the study.

Raw variant filtering for SCD outcome

Additional variant filtering steps were performed to explain, at minimum, the sudden cardiac death outcome in the HCM-affected group (n = 10). Re-calculation of $P_{Allelic}$-value thresholds were performed via Fisher's exact testing from the 2 × 2 contingency table for both autosomal dominant and recessive MOIs ($P_{Allelic}$-value = 0.0008 and $P_{Allelic}$-value = 0.0001, respectively). CaseControl was performed on the ten SCD (spontaneous) cases and on the cardiovascularly healthy geriatric control population (n = 21), while excluding animals that were HCM-affected but did not experience a SCD event from variant $P_{Allelic}$-value calculations. Total Freebayes-called variants (n = 64,035,993) were filtered according to a $P_{Allelic}$-value < 0.00081 (n = 27,003) and 'isVariant' status (n = 2,022), and manually analyzed for autosomal dominant and recessive MOI segregation as previously described. Expanded segregation analysis of the filtered variants failed segregation according to SCD outcome and were also excluded from the study.

Large structural variant analysis

A total of 73,843 concordant large structural variants (LSVs) were identified in the total WGS cohort of rhesus macaques (n = 36). After $P_{Allelic}$-value < 0.00031 filtering, 21 associated LSVs were identified, however, manually inspection of LSVs genotypes across all animals revealed no variant with segregation to HCM status. Similar to Freebayes-called variants, an additional analysis for the SCD outcome alone was performed; 73,456 concordant LSVs were present in the 31 animals included in this analysis. $P_{Allelic}$-value < 0.00081 filtering resulted in 58 associated LSVs, of which none successfully segregated according to SCD outcome. As such, these LSVs were deemed of unknown significance in the pathogenesis of rhesus macaque HCM and were excluded from consideration as sole causative variants.

Polygenic modeling analysis

Following single nucleotide polymorphism (SNP) pruning, a total of 18.1 million SNPs remained. A summary of the hyperparameters suggested a high percentage of variance explained (pve > 0.99), high variance explained by sparse effect variants (pge > 0.99), and a small number of sparse effect variants (2–8) (Fig. 7A; Table 3), suggesting all samples could be segregated by phenotype using a small number of variants. Evidence that phenotypes could be accurately partitioned by the PGS were noted (Fig. 7B), however, leave-one-out cross validation had poor prediction power. Roughly 50% of the HCM-phenotypes were accurately called (Fig. 7C), ultimately driven by low power due to a small sample size. These model site variants were not observed in previously associated human HCM genes. However, a 'LOW' effect synonymous mutation (c.837G > A; p.Arg279Arg) in *COX10*, a gene previously associated with isolated COX deficiency and early onset clinical phenotypes, including SCD and infantile HCM, was identified[33,34]. The genotypes of the six 'MODIFIER' and one 'LOW' effect gene body model sites were filtered from the VCF file and inspected across all rhesus macaques; model sites failed to sufficiently segregate regardless of deleterious or protective relationship to HCM, including the *COX10* variant as it was observed in all HCM-control (14 homozygous and 7 heterozygous) and in 10 HCM-affected (one homozygous and nine heterozygous) rhesus macaques.

Burden testing

A total of 17,409 variants were included and grouped by gene for burden testing. Burden testing identified 201 HCM-associated genes with raw *P*-values < 0.05. Following multiple testing correction, no initially identified genes had significant associations to disease. HCM in this cohort could not be explained by an association of multiple variants in any gene.

Methods

Animal care and husbandry

All procedures were approved by the CNPRC Institutional Animal Care and Use Committee of the University of California-Davis (Davis, CA, USA) as part of ongoing base grant #P51OD011107 and carried out in accordance with the Animal Welfare Act, current NIH Guide for the Care and Use of Laboratory Animals, and the ARRIVE2.0 guidelines[35–37]. Standard CNPRC rhesus macaque care and husbandry were performed as previously described[22]. Briefly, rhesus macaques were exposed to the same environmental conditions, including diet (LabDiet Monkey Diet 5047 [Purina Mills International, St Louis, MO] and fresh mixed vegetables/fruits), lighting exposure, and environmental enrichments. Routine wellness evaluations, clinicopathological assessments, and disease testing (tuberculosis, herpes B virus, simian Type D retrovirus, simian immunodeficiency virus, and simian T-lymphotropic virus) were performed. Echocardiographic evaluations under a single ketamine hydrochloride sedation (10 mg/kg; intramuscular [IM]) were performed for ease of animal handling. No animals were euthanized for the inclusion of the RNA-Seq aspect of this study. The heart tissues from cardiovascularly normal and HCM-affected animals that did not succumb to a SCD event were obtained from ongoing CNPRC studies following humane euthanasia with an overdose of pentobarbital (≥ 10 mg/kg; intravenous [IV]).

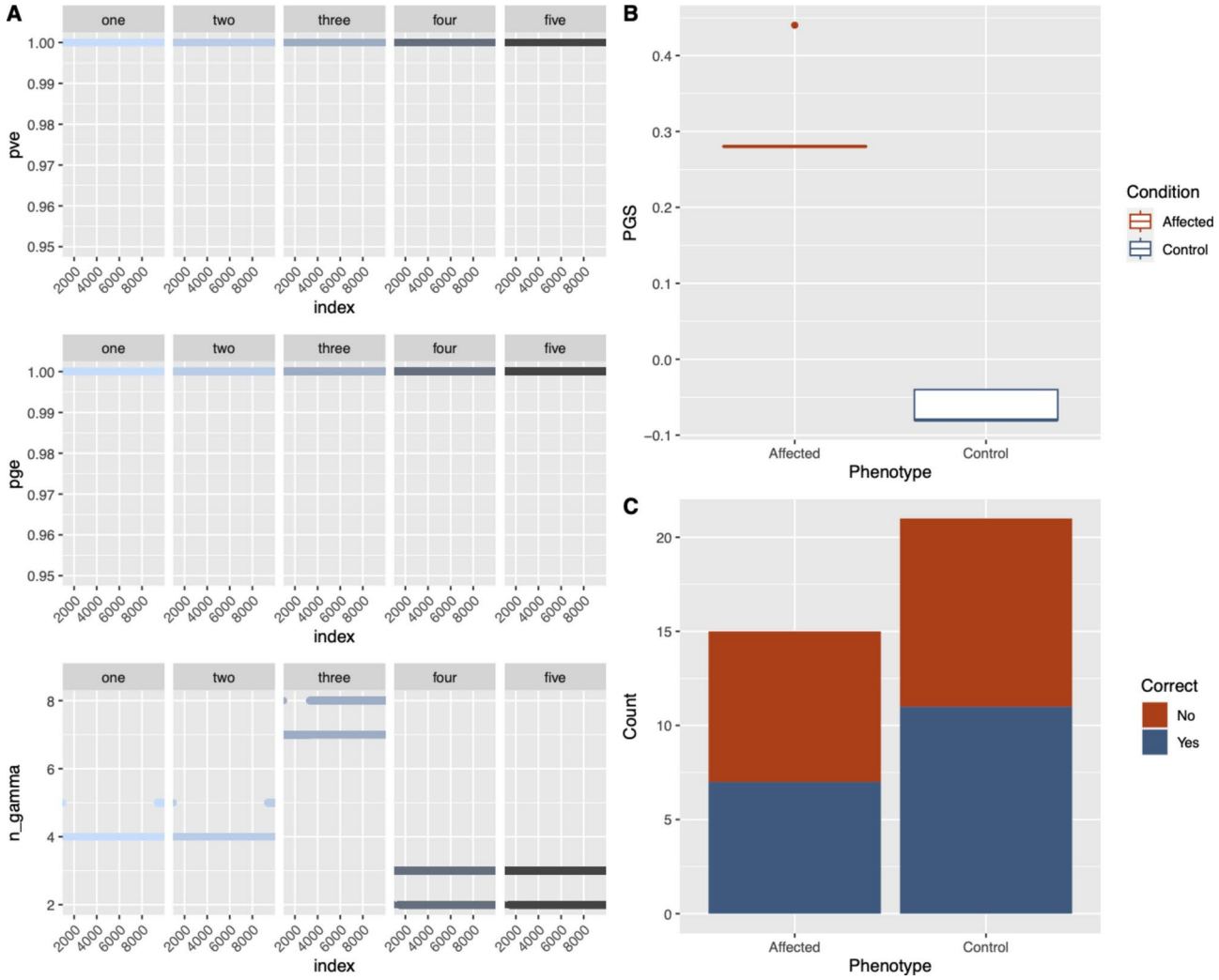

Fig. 7. Polygenic modeling of HCM status in rhesus macaques. Summary of hyper parameters (pve, pge and n_gamma) of five independent runs of BSLMM indicating model convergence among runs (**A**). Polygenic score distributions of the control (blue) and affected (red) groups (**B**). Accuracy summary of leave-one-out cross validation (**C**). Bar graphs represent the number of individuals in each group, and colors indicate the proportion of individuals correctly or incorrectly (blue or red, respectively) predicted to be affected or unaffected by HCM. Polygenic modeling was conducted utilizing the whole-genome sequencing data of 15 and 21 HCM-affected and -unaffected rhesus macaques (n = 36). Random (alpha), fixed (beta), and sparse effect (gamma/PIP) from the five runs were averaged, then polygenic scores (PGS) from the identified eight intergenic and seven gene variants with PIP > 0 (*QSOX1* [1], *ENSMMUG00000055166* [1], *SLC26A3* [1], *MYO3A* [1], *TMEM135* [1], *COX10* [1], *MAPK4* [1]; were estimated for each animal. Polygenic scores were calculated as: the effect size for each variant: beta × gamma + alpha, multiplied by the number of variant alleles, then summed among sites. pve = percent variation explained, pge = percent variation explained by sparse effects, n_gamma = number of sparse effect sites, PGS = polygenic scores, HCM = hypertrophic cardiomyopathy.

Sample collection and assessment of HCM

For the purpose of this study, the HCM phenotype was diagnosed by gross pathological assessment confirming significant LVH (Fig. 1). When available, measures of the O:I ratio were obtained and required to exceed the reported cut-off value for severe LVH (i.e., O:I ratio > 3)[18]. Pathology confirmation of an HCM phenotype included severe LVH with frequent luminal obliteration for HCM-affected animals. Extensive interrogation of the hearts of the control RNA-Seq group was performed, and required confirmed absence of LVH, no evidence of lumen obliteration, or histological evidence of myocytolysis, fibrosis, or infarction. Control animals for the GWAS arm of the study were confirmed to be normal based on an echocardiogram with no noted congenital heart disease and no structural LV changes; LV diastolic wall measures were confirmed to be normal based upon diastolic interventricular septal (IVSd) and/or diastolic LVPW (LVPWd) wall measurements < 8.8 and < 7.4 mm, respectfully[22].

Chromosome	Position	Gene	Alpha	Beta	Gamma	Effect
1	37,381,921	*QSOX1*	2.28E-10	− 1.89E-04	0.4	− 7.55E-05
1	120,054,697	–	− 4.25E-10	4.00E-01	0.4	1.60E-01
1	136,994,338	*ENSMMUG00000055166*	1.09E-10	− 9.99E-02	0.2	− 2.00E-02
3	134,084,043	*SLC26A3*	2.26E-10	− 4.00E-04	0.05704	− 2.28E-05
5	167,177,931	–	9.00E-11	− 1.00E-01	0.2	− 2.00E-02
6	170,689,099	–	1.05E-10	− 1.00E-01	0.2	− 2.00E-02
9	27,192,112	*MYO3A*	5.10E-10	− 1.00E-01	0.2	− 2.00E-02
10	72,601,711	–	− 3.20E-10	5.89E-05	0.13336	7.85E-06
12	1,741,963	–	− 2.50E-10	4.01E-01	0.4	1.60E-01
13	102,405,968	–	2.20E-10	− 1.00E-01	0.2	− 2.00E-02
14	80,243,573	*TMEM135*	− 2.53E-10	4.01E-01	0.4	1.60E-01
16	13,986,607	*COX10*	− 3.22E-10	− 4.64E-05	0.02452	− 1.14E-06
18	31,834,798	*MAPK4*	− 3.69E-12	4.00E-01	0.4	1.60E-01
X	72,686,562	–	6.02E-12	4.99E-02	0.2	9.98E-03
X	130,303,364	–	− 9.34E-11	− 2.50E-02	0.2	− 5.00E-03

Table 3. Total 15 polygenic variants with posterior inclusion probability greater than zero. A Bayesian sparse linear mixed model (BSLMM) was performed for polygenic modelling.

Echocardiography

Complete echocardiographic evaluations on apparently healthy geriatric (> 18 years-of-age) rhesus macaques were performed by a board-certified veterinary cardiologist (J.A.S.) or an experienced sonographer (Y.U.) under the direct supervision of a board-certified veterinary cardiologist (J.A.S.). Echocardiographic evaluations were performed with the use a 4- to 12mHZ or 1- to 5mHZ sector array transducer on an Affinity 50 echocardiographic ultrasound machine (Phillips, Amsterdam, Netherlands). Two-dimensional, M-mode, color, and spectral Doppler echocardiographic images were acquired in right and left lateral recumbency and later analyzed with the use of the Syngo Dynamics offline software (Siemens, Erlangen, Germany) according to adult comprehensive transthoracic echocardiographic measurement guidelines as previously described[38]. Geriatric control rhesus macaques were echocardiographically screened and confirmed to be free from: 1) apparent congenital disease (i.e., aortic or subaortic stenosis), 2) presence of ventricular or mid-ventricular septal bulges, and 3) LVH defined as an IVSd > 8.8 mm and/or LVPWd > 7.4 mm[22]. Banked from previously obtained whole-blood EDTA samples (2 mL) acquired via standard venipuncture were obtained for all the rhesus macaques in the GWAS arm of this study. Descriptive statistics (mean ± standard deviation) for all available demographic and phenotypically relevant echocardiographic data were performed for the HCM-unaffected geriatric rhesus macaque group. Statistical comparisons between the HCM-affected and -unaffected groups were carried out. Briefly, normality testing via Shapiro–Wilk and Kolmogorov–Smirnov tests was performed on all variables. For all categorical variables, two-sided Fisher's exact tests were performed. Unpaired parametric *t*-tests were performed on all normally distributed continuous variables, whereas Mann–Whitney *U* tests were performed for all non-normally distributed variables. A P-value < 0.05 was considered statistically significant for all demographic and phenotypic variable analyses.

Gross pathology/necropsy

Gross pathologic assessments on all rhesus macaques in the GWAS disease group and all animals for RNA-Seq (HCM-affected and -unaffected animals) were performed by a board-certified veterinary pathologist (J.R.R.) to confirm HCM status. Hearts were cut cross-sectionally midway between the apex and base of the heart at the level of the papillary muscles; severe HCM status was diagnosed in the event that an animal's LV outer-to-inner ratio was greater than three (O:I ratio > 3). Prior behavioral and physical exam data (including oscillometric blood pressure measurements and skin turgor) and routine clinical pathology results (glucose, fructosamine, T4, BUN, creatinine, and aforementioned disease testing) in the "in-life" phase of all HCM-affected rhesus macaques were examined for ruling-out potential phenocopies of HCM, including acute LV wall thickening (i.e., transient myocardial thickening and severe dehydration), metabolic disease (i.e., hypersomatotropism, diabetes, and hyperthyroidism), and systemic hypertension as a results from kidney disease. Attention to identifying evidence of left atrial enlargement, significant infarction, thromboembolism, and pulmonary edema in the HCM group was also given during necropsy. During gross pathologic examination, LVPW tissue samples from the control and affected groups were acquired and immediately flash-frozen in liquid nitrogen and later stored in − 80 °C for subsequent RNA isolation. Necropsy on the control rhesus macaques were performed as part of other ongoing terminal CNPRC projects; the tissue samples in this study's control population were deemed cardiovascularly healthy representatives of their species.

Descriptive statistics (mean ± standard deviation) for all available demographic and phenotypically relevant gross pathological data for all HCM-affected rhesus macaques used in this study, as well as for the cardiovascularly healthy RNA-Seq group, were performed; statistical testing between RNA-Seq HCM-affected and -unaffected groups was carried out as previously described.

RNA-sequencing

RNA-Seq was employed to study the gene expression profiling of HCM and identify DEGs between HCM-affected and -unaffected rhesus macaques (Table 1). From the LVPW tissue samples, total RNA was extracted using the RNeasy Plus Universal Mini Kit (QIAGEN, Hilden, Germany) following manufacturer's protocol. Mature mRNA transcripts were selected using the polyA selection method. High quality mature mRNA (RNA Integrity Number; [RIN] > 7) were converted to cDNA and NEBNext Ultra II RNA library preparations were made. Samples were later used for paired-end 150 base-pair (bp) RNA-Seq on the Illumina HiSeq platform with a targeted read depth of ~ 50 million reads per sample (GENEWIZ, South Plainfield, NJ). Sequenced reads underwent quality control to remove Illumina adapter sequences and low-quality reads with the Trimmomatic (v0.39) software. High quality reads were mapped to the annotated ENSEMBL *Macaca mulatta* reference genome (Mmul_10/rheMac10) where BAM files were generated and GeneCounts obtained using the STAR aligner software (v2.7.10b); normalized unique gene hit counts were later used for downstream differential expression analysis using the DESeq2 package (v1.42.0) in R (v4.3.2). Genes with low read counts were removed. Similarity within and between groups was assessed by performing a principal component analysis (PCA) (Fig. 2). Genes with a P_{Adjusted}-value < 0.05 and |log2FoldChange| > 1 were called as DEGs for each comparison. A volcano plot was constructed to illustrate the global transcriptional change across the groups in comparison (Fig. 3), and the top 20 up- and downregulated DEGs (n = 40) were reported (Table 2). As an exploratory analysis to highlight similarities between DEGs in rhesus macaques and human HCM, rhesus macaque down-/upregulated HCM DEGs identified were compared to a dataset of previously identified DEGs in both pediatric and adult HCM human heart samples[39].

Age was subsequently tested as a possible covariate in an additional RNA-Seq analysis. A design model of condition (HCM status) + age was compared to a condition-only model using a likelihood ratio test (LRT) implemented in DESeq2 package (v1.42.0).

DEGs enrichment analyses

Global statistically significant DEGs were divided between down- and upregulated (log2FoldChange > − 1 or > 1, respectively) DEGs. Gene enrichment analyses were performed with the clusterProfiler package (v4.10.0) in R (v4.3.2), where KEGG pathway[30–32] and Gene Ontology (GO; biological process [BP], molecular function [MF], and cellular compartment [CC]) enriched terms were interrogated. The human annotation file (org.Hs.eg.db) was used as a superior annotation file as it was observed to outperform the available rhesus macaque annotation file (org.Mmu.eg.db) and provide a larger number of observed DEGs. False discovery rate (FDR) *P*-value adjustment was performed. A P_{Adjusted}-value < 0.05 was considered statistically significant for all enrichment analyses.

Whole-genome sequencing and association analyses

The previously acquired peripheral EDTA whole-blood samples were used for gDNA extraction according to protocol (QIAGEN, Hilden, Germany). The gDNA from the 36 sex-matched rhesus macaques were selected for WGS (Supplemental Table 1). NanoDrop One/One Spectrophotometer (ThermoFisher) was used to quantify [DNA] and 260/280 ratios. To ensure DNA integrity, samples were run on a 1% agarose gel at 80 V for one hour. Paired-end 150 bp DNA libraries were created using a TruSeq Nano library kit for each sample[40]. Libraries were pooled and sequenced on the Illumina NovaSeq6000 sequencing platform at ~ 30X coverage (Psomagen, Rockville, MD). Briefly, all adapter contamination and low-quality sequences of a Phred quality score ≤ 28 were removed using the Trimmomatic software (v0.39) and evaluated using the FastQC program (v0.12.1)[41]. High quality reads were aligned to the ENSEMBL rhesus macaque reference genome (Mmul_10/rheMac10) using BWA-MEM (v0.7.17)[42,43]. The resulting bam files were sorted and indexed with Samtools (v1.17)[44]. Samtools was also used to calculate average sequencing depth for all successfully mapped and sorted bam files across all positions including those with zero depth. PCR duplicates were removed with Picard tools (v2.26.11). Freebayes (v1.3.6) was employed for variant calling of single nucleotide polymorphisms, small insertions/deletions (indels), and small structural variants[45]. Generated VCF files were used to assign CaseControl status on individuals and predict variant effects on protein function with the use of the SnpSift and SnpEff packages (v5.2)[46,47]. Segregation analyses were performed for both autosomal dominant and autosomal recessive MOI scenarios. $P_{Allelic}$-value thresholds for both analyses were calculated via 2 × 2 contingency table and Fisher's exact testing. The $P_{Allelic}$-value thresholds conditions for an autosomal recessive MOI permitted all HCM-affected animals to bear two copies and all control animals to be heterozygous for the alternate allele, resulting in a $P_{Allelic}$-value = 0.0001. $P_{Allelic}$-value threshold conditions for an autosomal dominant MOI permitted all HCM-affected animals to harbor at least one copy of the alternate allele, while allowing a total of four alternate alleles in the control group (i.e., four misphenotyped animals in the heterozygous state), resulting in a $P_{Allelic}$-value = 0.0003. The $P_{Allelic}$-value thresholds were selected as worst-case scenarios to ensure all variants of interest were retained. Annotated variants were subsequently filtered according to previously determined $P_{Allelic}$-value thresholds ($P_{Allelic}$-value < 0.00011 and $P_{Allelic}$-value < 0.00031 for the recessive and dominant MOI, respectively) and predicted 'HIGH' and/or 'MODERATE' variant effect on protein function.

The mapped bam files from the 36 rhesus macaques were used for LSV calling with the use of four programs: 1) Smoove (v0.2.8), a wrapper for Lumpy (v0.2.13), 2) GRIDSS (v2.13.2–2), 3) DELLY (v1.1.6), and 4) Manta (v1.6.0)[48–51]. Each tool discovers LSVs based on a combination of coverage, split-reads, and assembly and has unique discrete working pipelines. Briefly, population level LSV discovery and downstream merging was conducted independently with each tool. SURVIVOR (v1.0.7) was then used to merge LSVs longer than 50 bp, of the same type, less than 100 bp apart, and identify concordant calls[52]. The output file from SURVIVOR was then filtered to include LSVs that were discovered by two or more callers, except in the case of insertions. Insertions were largely discordant among callers, partially due to length discrepancies. Interim testing revealed that Manta

called more expected insertions than DELLY or GRIDSS, therefore, insertions that were called by Manta, or had support from two LSVs callers, were included. Samples were then genotyped for each LSV using GraphTyper (v2.7.2)[53]; 'PASS'-filtered calls with a 'SVMODEL = AGGREGATED' were retained for subsequent segregation analysis.

Polygenic modeling analysis

Due to high false positive rates and high SNP densities, indels were removed from the Freebayes-called VCF file. SNP calls with quality (QUAL) < 20, overall read depth < 10, and minor allele frequency (MAF) < 0.05 were pruned. The Bayesian sparse linear mixed model (BSLMM) from the GEMMA (v0.98.5) package was employed to estimate the genetic architecture and heritability of HCM in the study population of rhesus macaques, and test a polygenic model to explain and predict the HCM phenotype[54,55]. The polygenic modeling analysis comprised of five independent Monte Chain Monte Carlo (MCMC) chains, each running for 100,000,000 steps, with a burn-in and sampling of 50,000,000 and 10,000 steps, respectively. The hyperparameters' total phenotypic variation explained (pve), variation explained by sparse effect variants (those with measurable effects; pge), and number sparse effect variants (n_gamma) were examined among runs to assess convergence of MCMC chains. BSLMM was used to infer the posterior inclusion probability (PIP) for each SNP, indicating the likelihood of a non-zero effect size, conditional on non-zero presence. Model-averaged breeding values (PGS) for all animals were estimated based on the summed effect size of sparse effect variants. To gauge the predictive capacity of the BSLMM models, leave-one-out cross-validation was executed utilizing the predict option, systematically withholding phenotypic data for one individual at a time and employing the remaining samples to forecast its phenotype based on genotype through BSLMM. To determine the biological relevance of gene body model sites, the gene symbols of these model sites were cross-checked across a comprehensive human gene list of ClinVar human HCM-associated genes (n = 246) curated from a "hypertrophic cardiomyopathy"[dis] search (accessed 31 May 2024).

Burden testing

Burden testing of rare potentially deleterious variants was performed by extracting variants with MAF < 0.05 and QUAL > 20 with either 'MODERATE' or 'HIGH' SnpEff-predicted variant impacts on protein. These categories included missense, frameshift, in-frame indels, splice site variants, start lost, and stop gained variants. Variants were subsequently grouped by gene and set-testing was conducted in PLINK (v1.9) with 10,000 permutations and multiple test correction (i.e,. –assoc –adjust) parameters.

Discussion

The incidence of a spontaneously occurring HCM phenotype with SCD in the CNPRC rhesus macaque colony provides a unique opportunity for translational research aiming to better explain the pathogenesis, diagnosis, and successful therapeutic strategies for this disease. In cats, the disease has been observed in over 30 breeds and remains common in mixed breeds[2,11]. Candidate-gene-led feline genetic studies of HCM have identified the breed specific A31P and R820W *MYBPC3* autosomal dominant variants in Maine Coon and Ragdoll cats, respectively, and the autosomal dominant G3376R *Alström Syndrome-1 Centrosome and Basal Body Associated Protein* (*ALSM1*) mutation in Sphynx cats[12–16]. These mutations remain breed specific and do not explain the disease in other pure or mixed-breed cats. To study the molecular etiology of HCM for pharmacologic therapies, an A31P-harboring cat colony has been established. Studies revealed that the A31P mutation in this cat colony is incompletely penetrant as not all A31P-positive cats have phenotypic evidence of HCM and a small number of A31P-negative cats have presented with HCM, despite their controlled environment[17]. This finding reduces the ability to reliably produce HCM-affected and control cats within the colony and highlights the continued need for defining additional model organisms, particularly large-animal, for human HCM. Given the unique applicability of NHP models and the significant knowledge gaps in human and feline HCM, the genetic etiology of these rhesus macaque cases was studied extensively. Broadly, this study aimed to identify the molecular signature of rhesus macaque HCM and SCD via WGS identification of causative variants and transcriptomics of affected heart tissues.

This study first sought to perform RNA-Seq profiling on the LVPW tissues of HCM-affected and -unaffected rhesus macaques to characterize the disease at the transcriptomic level. A total of 614 downregulated and 1,065 upregulated DEGs were identified in this study's population of monkeys; the 20 top-most statistically significant DEGs included 17 upregulated (*MAFF*, *ELL2*, *SOCS3*, *MEDAG*, *S100A9*, *IL1RL1*, *S100A8*, *SERPINE1*, *JUNB*, *STC1*, *CD44*, *VMP1*, *EIF4A1*, *CDKN1A*, *CEMIP2*, *MT2A*, and *FOSL2*) and three downregulated (*SPEG*, *HAND2*, and *TBX5*) genes.

The top-most significant DEG identified in the RNA-Seq dataset was *MAF BZIP Transcription Factor F* (*MAFF*), a gene encoding for a basic leucine zipper (bZIP) transcription factor (TF). MAFF proteins lack a transactivation domain, allowing both enhanced gene expression via heterodimerization with other bZIP TFs or gene repression via homodimer formation[56]. The role of *MAFF* as a central regulator of networks leading to atherosclerosis and coronary artery disease has previously been reported; *MAFF* modulates cholesterol levels and downstream cascade processes for the regulation of inflammation and coronary artery disease risk[57]. A 26-fold increase in the transcript expression of *MAFF* was present in the tissues of HCM-affected monkeys. Following manual STRING protein–protein analysis of the 20 top-most DEGs, MAFF was found to interact with eight (*JUNB, FOSL2, SOCS3, CDKN1A, SERPINE1, CD44, S100A8, and S100A9*) of the 17 upregulated DEGs. JunB Proto-Oncogene, AP-1 Transcription Factor Subunit (JUNB) is a TF that allows for sequence-specific double-stranded DNA binding and is involved in the positive regulation of other TFs by RNA pol II. Increased JunB levels were reported in the HCM-affected hearts of *Mekk1*-null genetically modified mice and is thought to mediate LV hypertrophy via muscle size maintenance, rapid hypertrophy induction, and block atrophy[58,59]. In

the same JUN protein family is *FOS-Like 2, AP-1 Transcription Factor Subunit* (*FOSL2*), a leucine zipper TF, when dimerized, forms the TF AP-1 complex. In a *Fosl2* knock-out HCM mouse model receiving continuous infusion of angiotensin II, the hearts of the Fosl2-null mice are protected from cardiac fibrosis and aberrant autophagocytosis of cardiac fibroblasts[60]. Overexpression of *Fosl2* in mice has shown increased interstitial cardiac fibrosis and deranged expression of genes involving conduction system maintenance; interestingly, *FOSL2* overexpression is associated with cardiac fibrosis, arrythmias, and aberrant response to stress, under immunofibrotic conditions[61]. Upregulation of *Suppressor of Cytokine Signaling 3* (*SOCS3*), a STAT-induced STAT inhibitor (SSI) responsible for negative regulation of cytokine signaling, is thought to be protective against cardiac hypertrophy[62,63]. The observed increase in *SOCS3* expression in the LV of HCM-affected monkeys might be reflective of an as of yet uncharacterized compensatory pathway of hypertrophy and may not be involved in the disease pathogenesis. *Cyclin Dependent Kinase Inhibitor 1A* (*CDKN1A*) is a robust cyclin-dependent kinase inhibitor known for its involvement in the p53 DNA repair and cell senescence pathways; genetic loci and elevated CDKN1A levels have previously been associated with myocardial hypertrophy in humans[64–68]. *Serpin Family E Member 1* (*SERPINE1*) is a chief serine proteinase inhibitor (serpin) for fibrinolysis via tissue and urokinase plasminogen activator (tPA and uPA, respectively) inhibition. Interestingly, null SERPINE1 expression in humans bearing homozygous copies of a loss-of-function (LoF) mutation experience age-dependent cardiac fibrosis without LV dilation and increased risk of cardiomyocyte injury. Increased transcript expression of *SERPINE1* in the myocardial tissues of rhesus macaques could explain a mechanism by which fibrosis is not a commonly observed pathophysiologic or histopathologic feature of HCM in this species[69,70]. CD44 is a glycoprotein molecule localized to cellular surfaces and is heavily involved in cell-to-cell interactions, cellular migration, ligand-interaction, lymphocyte activation, and circulation and homing[71]. Inhibition of CD44 in mouse models of LV pressure overload via transverse aortic constriction attenuates cardiac inflammation, fibrosis, and progression to heart failure[72]. *S100 Calcium Binding Protein A8* and *A9* (*S1A008* and *S1A009*, respectively) are members of the S100 protein family that are localized to the cytoplasm and nucleus in a number of cell types, and are involved in cellular differentiation and proliferation. *S1A008* and *S1A009* form heterodimers and are currently promising candidate molecular biomarkers for myocardial infarction; increased serum levels of S100A proteins following acute ischemic events in humans have been reported[73,74]. *S1A008* and *S1A009* may exert prohypertrophic effects, particularly in the case of SCD which are often precipitated by exercise or physiologic stress in rhesus macaques. Surprisingly, ischemic preconditioning with *S1A008* and *S1A009* in a norepinephrine-induced myocyte hypertrophy mouse model demonstrated antihypertrophic effects and attenuates progression of heart failure; collectively this represents a unique opportunity for a targeted approach at mitigating SCD in rhesus macaques[75].

Considering the translational applicability of the rhesus macaque model to human HCM, a total of 40 and 64 rhesus downregulated DEGs were found to be shared between pediatric and adult human HCM-affected patients, respectively. A total of 71 shared upregulated rhesus DEGs were identified in the human pediatric cohort, whereas 215 shared upregulated DEGs were identified in the adult human cohort. The 10 downregulated and 36 upregulated HCM DEGs shared between the rhesus macaque and human dataset sets, regardless of age group, collectively support the close molecular similarity as well as biologically relevant transcriptomic signatures of rhesus macaque HCM that can be directly translated to human disease.

Shared molecular pathways as identified by GO and KEGG pathway[30–32] term enrichment analyses were identified between rhesus macaques and humans affected by HCM. For example, top rhesus enriched GO terms corresponding to "heart contraction", "regulation of heart contraction", "cardiac conduction", "cardiac muscle cell contraction", and "response to lipopolysaccharide" are shared with human HCM[76–78]. While the list of down- and upregulated DEGs in the LV tissues of HCM-affected rhesus macaques driving these term enrichments are not the same DEGs driving the terms in human HCM, these shared terms call attention to fixed pathophysiologic processes inherent to the disease pathogenesis. These ubiquitously shared processes are promising translational pharmacologic targets and hold promise at treating HCM in both human and veterinary patients, particularly in the asymptomatic period of the disease; for instance, sarcomere inhibitors developed to target the hypercontractile state of cardiac sarcomeres[79–86].

Of the top enriched 32 GO BP, 31 GO MF, and 22 KEGG terms[30–32] reported here (n = 85), 13 enriched terms associated with channelopathies were identified for GO BP ("potassium ion transport", "potassium ion transmembrane transport", and "cardiac conduction") and GO MF ("salt transmembrane transporter activity", "metal ion transmembrane transporter activity", "monoatomic ion gated channel activity", "gated channel activity", "voltage-gated monoatomic ion channel activity", "voltage-gated channel activity", "potassium ion transmembrane transporter activity", "voltage-gated monoatomic cation channel activity", "potassium channel activity", and "voltage-gated potassium channel activity") analyses. Remarkably, channelopathy-associated terms were identified in ~ 57% of GO BP and GO MF of downregulated DEGs, while none were present in the upregulated dataset. In addition to similar hemodynamic and structural changes, the observed overrepresentation of channelopathy-associated terms in HCM downregulated genes provides transcriptomic evidence on how rhesus macaque HCM model recapitulates human disease, distinctly in the SCD outcome thought to be a result of arrythmias following episodes of increased exercise or physiologic stress. Taken all together, the rhesus macaque gene-expression profiling data presented in this study validates the phenotype parameters of the study and confirms and strengthens 1) the utility of rhesus macaques as a model of HCM and 2) the close molecular relationship to other species affected by the disease, paving the way for future expanded molecular and clinical studies.

Despite best attempts at identifying HCM-associated variants in a strictly phenotyped population of rhesus macaques (necropsy confirmed cases and cardiovascularly healthy geriatric controls greater than 18 years-of-age), all statistically significant putative variants were ultimately excluded as they failed to segregate within both HCM status and SCD outcome segregation analyses. Polygenic statistical modeling analyses were performed in attempts to elucidate compound variants that would explain the HCM phenotypes, however, such complex

analysis resulted in leave-one-out cross validation with poor prediction power (50%) explained by a low population size. Continued polygenic modeling of HCM in the CNPRC in an expanded population of animals is needed. Burden testing was performed as an additional attempt at identify HCM-associated variants. A caveat of this study's burden test relies on analyzing rare variants, but given our WGS sample size, the smallest allele frequency was 0.014 (i.e., 1/72 alleles), therefore robust sites with a MAF < 0.05 were included. Burden testing identified 201 HCM-associated genes with raw *P*-values < 0.05. However, burden testing failed to identify associated variants following multiple test correction, thus, HCM could not be clearly explained by an association of multiple variants in any gene. This failure to identify a single causative variant is perhaps not surprising giving the heterogenous genetic landscape of HCM that is now well described in humans and cats. In humans with HCM, approximately half of cases are left with an unidentified causative variant; this has been postulated to be due to complex genetic architectures, shared common variants, modifying genes, and environmental impacts on disease expression[7]. When the 'HIGH' and 'MODERATE' effect variants identified through our rhesus macaque GWAS were reviewed, a number of recessive variants can each explain 25–40% of the cases (Supplemental File 1). Perhaps the rhesus macaque population more closely mimics the human population than previously expected, where many genetic variants (some yet to be discovered) are required to explain the totality of disease. With further population-based testing and use of future validation cohorts, the role of multiple explanatory variants may be elucidated.

The population of rhesus macaques included in the RNA-Seq and WGAS aspect of the study represent possible limitations of this study. HCM-affected rhesus macaques were significantly younger (3.87yrs [± 0.83]) than their control counterparts (6.33yrs [± 1.28]; *P*-value = 0.004). Given this age difference, the possibility of segregating pre-pubescent versus post-pubescent rhesus macaques between the control and affected cohorts is possible. However, in the authors' experience, captive rhesus macaques can have a shortened time to sexual maturity which has been observed across prior studies and reports[87,88]. Broadly, the age of puberty in males is frequently described between 3–4 years, whereas 2.5–3.5 years in females. Using these definitions, all monkeys in the RNA-Seq HCM-affected and -unaffected group would be considered adolescent and at, or beyond, the age of puberty, but not yet considered a mature adult. It is believed, at minimum, the influence of sexual hormones was present within both study groups; however, neither sexual organ examination at necropsy, nor hormone-level assessments, were included in this work which represents a limitation. Despite the inherent numerical difference in ages between groups, RNA-Seq results show biologically expected and appropriate data that goes in the opposite direction of what is to be expected with the relatively higher age of the control group. Future continued research efforts at investigating expression patterns of the most compelling DEGs reported here (i.e., top 20 significant and/or human-shared HCM DEGs genes) involving more targeted methods (e.g., RT-qPCR, western blot, and/or immunohistochemistry) are warranted. Secondly, the majority of the rhesus macaques used for the RNA-Seq and GWAS aspects of this study were of pure Chinese origin, albeit, a small number of rhesus macaques of pure Indian origin and of admixed make-up were also included. Importantly, all rhesus macaque HCM cases included in this study were of pure or admixed Chinese origin. A single Chinese/Indian-admixed rhesus was included in the HCM-affected group for the RNA-Seq arm of the study. For the GWAS aspect of the study, 3/15 Chinese/Indian-admixed rhesus macaques were included in the HCM-affected cohort, whereas 4/21 pure Indian-descent and 5/21 Chinese/Indian-admixed rhesus macaques were included in the HCM-unaffected group. Chinese and Indian percent genetic make-up were not statistically significant between HCM-affected and -unaffected groups within the RNA-Seq and GWAS population of rhesus macaques (*P*-value = 0.42 and *P*-value = 0.093, respectively). While the addition of Chinese and Indian rhesus macaque HCM-affected individuals could impede the ability to identify a single genetic variant of interest, the authors were careful to only include animals of varied sole origin (i.e., 100% Chinese versus 100% Indian) in the larger control group to reduce identification of non-causal associated variants.

While unsuccessful identification of a single causative genomic variant may further delay the use of this ideal NHP model of HCM, reported are key transcriptomic signatures of HCM and candidate molecular pathways for the development of novel drug therapies targeted at mitigating characteristic features of the disease in rhesus macaques that can be translated to other species. It is further worth noting that KEGG[30–32] analysis revealed an upregulation of terms related to the immune system and infectious disease. While these have intermittently been reported in human and other HCM transcriptomic studies of various species, this finding may suggest the role of environment and possible non-genetic modification of rhesus macaque HCM expression.

Furthermore, provided are the raw WGS files of a valuable established cohort of robustly phenotyped geriatric rhesus macaques as well as RNA-Seq from necropsy-confirmed normal LV tissue samples as an open-access resource for future robust genetic studies of cardiovascular disease.

Data availability

Raw RNA-Seq and WGS reads are available from NCBI's SRA under BioProject PRJNA1185182 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1185182). Additional data is available upon reasonable request from the corresponding author.

Received: 13 September 2024; Accepted: 9 December 2024
Published online: 28 December 2024

References

1. Maron, B. J. et al. Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. *Circulation.* **92**(4), 785–789. https://doi.org/10.1161/01.cir.92.4.785 (1995).

2. Payne, J. R., Brodbelt, D. C. & Luis, F. V. Cardiomyopathy prevalence in 780 apparently healthy cats in rehoming centres (the CatScan study). *J. Vet. Cardiol.* **17**(Suppl 1), S244–S257. https://doi.org/10.1016/j.jvc.2015.03.008 (2015).
3. Ueda, Y. & Stern, J. A. A one health approach to hypertrophic cardiomyopathy. *Yale J. Biol. Med.* **90**(3), 433–448 (2017).
4. Schober, K. E. et al. Retrospective evaluation of hypertrophic cardiomyopathy in 68 dogs. *J. Vet. Intern. Med.* **36**(3), 865–876. https://doi.org/10.1111/jvim.16402 (2022).
5. Huang, S. Y. et al. Heritability estimate of hypertrophic cardiomyopathy in pigs (Sus scrofa domestica). *Lab. Anim. Sci.* **46**(3), 310–314 (1996).
6. Maron, M. S. et al. Effect of left ventricular outflow tract obstruction on clinical outcome in hypertrophic cardiomyopathy. *N. Engl. J. Med.* **348**(4), 295–303. https://doi.org/10.1056/NEJMoa021332 (2003).
7. Marian, A. J. & Braunwald, E. Hypertrophic cardiomyopathy: Genetics, pathogenesis, clinical manifestations, diagnosis, and therapy. *Circ. Res.* **121**(7), 749–770. https://doi.org/10.1161/CIRCRESAHA.117.311059 (2017).
8. Topriceanu, C. C., Pereira, A. C., Moon, J. C., Captur, G. & Ho, C. Y. Meta-analysis of penetrance and systematic review on transition to disease in genetic hypertrophic cardiomyopathy. *Circulation.* **149**(2), 107–123. https://doi.org/10.1161/CIRCULATIONAHA.123.065987 (2024).
9. Maron, B. J. & Fox, P. R. Hypertrophic cardiomyopathy in man and cats. *J. Vet. Cardiol.* **17**(Suppl 1), S6-9. https://doi.org/10.1016/j.jvc.2015.03.007 (2015).
10. Freeman, L. M., Rush, J. E., Stern, J. A., Huggins, G. S. & Maron, M. S. Feline hypertrophic cardiomyopathy: A spontaneous large animal model of human HCM. *Cardiol Res.* **8**(4), 139–142. https://doi.org/10.14740/cr578w (2017).
11. Fox, P. R. et al. International collaborative study to assess cardiovascular risk and evaluate long-term health in cats with preclinical hypertrophic cardiomyopathy and apparently healthy cats: The REVEAL Study. *J. Vet. Intern. Med.* **32**(3), 930–943. https://doi.org/10.1111/jvim.15122 (2018).
12. Meurs, K. M. et al. A cardiac myosin binding protein C mutation in the Maine Coon cat with familial hypertrophic cardiomyopathy. *Hum. Mol. Genet.* **14**(23), 3587–3593. https://doi.org/10.1093/hmg/ddi386 (2005).
13. Meurs, K. M., Norgard, M. M., Ederer, M. M., Hendrix, K. P. & Kittleson, M. D. A substitution mutation in the myosin binding protein C gene in ragdoll hypertrophic cardiomyopathy. *Genomics.* **90**(2), 261–264. https://doi.org/10.1016/j.ygeno.2007.04.007 (2007).
14. Mary, J. et al. Prevalence of the MYBPC3-A31P mutation in a large European feline population and association with hypertrophic cardiomyopathy in the Maine Coon breed. *J. Vet. Cardiol.* **12**(3), 155–161. https://doi.org/10.1016/j.jvc.2010.06.004 (2010).
15. Longeri, M. et al. Myosin-binding protein C DNA variants in domestic cats (A31P, A74T, R820W) and their association with hypertrophic cardiomyopathy. *J. Vet. Intern. Med.* **27**(2), 275–285. https://doi.org/10.1111/jvim.12031 (2013).
16. Meurs, K. M. et al. A deleterious mutation in the ALMS1 gene in a naturally occurring model of hypertrophic cardiomyopathy in the Sphynx cat. *Orphanet. J. Rare Dis.* **16**(1), 108. https://doi.org/10.1186/s13023-021-01740-5 (2021).
17. Stern, J. A. et al. Hypertrophic cardiomyopathy in purpose-bred cats with the A31P mutation in cardiac myosin binding protein-C. *Sci. Rep.* **13**(1), 10319. https://doi.org/10.1038/s41598-023-36932-5 (2023).
18. Reader, J. R. et al. Left ventricular hypertrophy in rhesus macaques (Macaca mulatta) at the California National Primate Research Center (1992–2014). *Comp. Med.* **66**(2), 162–169 (2016).
19. Haertel, A. J. et al. Antemortem Screening for Left Ventricular Hypertrophy in Rhesus Macaques (Macaca mulatta). *Comp. Med.* **66**(4), 333–342 (2016).
20. Korcarz, C. E., Padrid, P. A., Shroff, S. G., Weinert, L. & Lang, R. M. Doppler echocardiographic reference values for healthy rhesus monkeys under ketamine hydrochloride sedation. *J. Med. Primatol.* **26**(6), 287–298. https://doi.org/10.1111/j.1600-0684.1997.tb00057.x (1997).
21. Tang, H. L., Wang, L. L., Cheng, G., Wang, L. & Li, S. Evaluation of the cardiovascular function of older adult Rhesus monkeys by ultrasonography. *J. Med. Primatol.* **37**(2), 101–108. https://doi.org/10.1111/j.1600-0684.2007.00249.x (2008).
22. Ueda, Y., Gunther-Harrington, C. T., Cruzen, C. L., Roberts, J. A. & Stern, J. A. Echocardiographic parameters of clinically normal geriatric rhesus macaques (Macacamulatta). *J. Am. Assoc. Lab. Anim. Sci.* **56**(4), 361–368 (2017).
23. Kanthaswamy, S. et al. Large scale pedigree analysis leads to evidence for founder effects of Hypertrophic Cardiomyopathy in Rhesus Macaques (Macaca mulatta). *J. Med. Primatol.* **43**(4), 288–291. https://doi.org/10.1111/jmp.12127 (2014).
24. Ueda, Y., Kovacs, S., Reader, R., Roberts, J. A. & Stern, J. A. Heritability and pedigree analyses of hypertrophic cardiomyopathy in rhesus macaques (*Macaca Mulatta*). *Front Vet. Sci.* **8**, 540493. https://doi.org/10.3389/fvets.2021.540493 (2021).
25. Ueda, Y. et al. Heart rate and heart rate variability of rhesus macaques (*Macaca mulatta*) affected by left ventricular hypertrophy. *Front. Vet. Sci.* **6**, 1. https://doi.org/10.3389/fvets.2019.00001 (2019).
26. Milani-Nejad, N. & Janssen, P. M. Small and large animal models in cardiac contraction research: Advantages and disadvantages. *Pharmacol Ther.* **141**(3), 235–249. https://doi.org/10.1016/j.pharmthera.2013.10.007 (2014).
27. Camacho, P., Fan, H., Liu, Z. & He, J. Q. Small mammalian animal models of heart disease. *Am. J. Cardiovasc. Dis.* **6**(3), 70–80 (2016).
28. Vakrou, S. et al. Differences in molecular phenotype in mouse and human hypertrophic cardiomyopathy. *Sci. Rep.* **11**(1), 13163. https://doi.org/10.1038/s41598-021-89451-6 (2021).
29. Oldt, R. F. et al. MYBPC3 haplotype linked to hypertrophic cardiomyopathy in Rhesus macaques (*Macaca mulatta*). *Comp. Med.* **70**(5), 358–367. https://doi.org/10.30802/AALAS-CM-19-000108 (2020).
30. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Prot. Sci.* **28**(11), 1947–1951. https://doi.org/10.1002/pro.3715 (2019).
31. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**(D1), D587–D592. https://doi.org/10.1093/nar/gkac963 (2023).
32. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).
33. Antonicka, H. et al. Mutations in COX10 result in a defect in mitochondrial heme A biosynthesis and account for multiple, early-onset clinical phenotypes associated with isolated COX deficiency. *Hum. Mol. Genet.* **12**(20), 2693–2702. https://doi.org/10.1093/hmg/ddg284 (2003).
34. Yang, Z. et al. A Novel COX10 deletion polymorphism as a susceptibility factor for sudden cardiac death risk in Chinese populations. *DNA Cell Biol.* **40**(1), 10–17. https://doi.org/10.1089/dna.2020.6086 (2021).
35. Percie du Sert, N. et al. Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0. *PLoS Biol.* **18**(7), e3000411. https://doi.org/10.1371/journal.pbio.3000411 (2020).
36. Animal Welfare Act as Amended. *USC.* 2013;7:S2131-2159.
37. 8 Guide for the Care and Use of Laboratory Animals (The National Academies Press) (2010).
38. Rivas, V. N., Ueda, Y. & Stern, J. A. Sex-specific differences and predictors of echocardiographic measures of diastolic dysfunction in rhesus macaques (*Macaca mulatta*). *J. Med. Primatol.* **52**(6), 374–383. https://doi.org/10.1111/jmp.12662 (2023).
39. Chen, S. et al. Transcriptome analysis of human hypertrophic cardiomyopathy reveals inhibited cardiac development pathways in children. *iScience.* **27**(1), 108642. https://doi.org/10.1016/j.isci.2023.108642 (2024).
40. Pasquali, F. et al. Application of different DNA extraction procedures, library preparation protocols and sequencing platforms: Impact on sequencing results. *Heliyon.* **5**(10), e02745. https://doi.org/10.1016/j.heliyon.2019.e02745 (2019).
41. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics.* **30**(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 (2014).

42. Warren, W. C. et al. Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science* https://doi.org/10.1126/science.abc6617 (2020).
43. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* **25**(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 (2009).
44. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* **25**(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 (2009).
45. Garrison EMG. Haplotype-based variant detection from short-read sequencing. *arXiv:12073907v2* 2012
46. Cingolani, P. 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)* **6**(2), 80–92. https://doi.org/10.4161/fly.19695 (2012).
47. Cingolani, P. et al. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front Genet.* **3**, 35. https://doi.org/10.3389/fgene.2012.00035 (2012).
48. Rausch, T. et al. DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics.* **28**(18), i333–i339. https://doi.org/10.1093/bioinformatics/bts378 (2012).
49. Chen, X. et al. Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics.* **32**(8), 1220–1222. https://doi.org/10.1093/bioinformatics/btv710 (2016).
50. Cameron, D. L. et al. GRIDSS2: Comprehensive characterisation of somatic structural variation using single breakend variants and structural variant phasing. *Genome Biol.* **22**(1), 202. https://doi.org/10.1186/s13059-021-02423-x (2021).
51. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: A probabilistic framework for structural variant discovery. *Genome Biol.* **15**(6), R84. https://doi.org/10.1186/gb-2014-15-6-r84 (2014).
52. Jeffares, D. C. et al. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* **8**, 14061. https://doi.org/10.1038/ncomms14061 (2017).
53. Eggertsson, H. P. et al. GraphTyper2 enables population-scale genotyping of structural variation using pangenome graphs. *Nat. Commun.* **10**(1), 5402. https://doi.org/10.1038/s41467-019-13341-9 (2019).
54. Zhou, X., Carbonetto, P. & Stephens, M. Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* **9**(2), e1003264. https://doi.org/10.1371/journal.pgen.1003264 (2013).
55. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**(7), 821–824. https://doi.org/10.1038/ng.2310 (2012).
56. Katsuoka, F. & Yamamoto, M. Small Maf proteins (MafF, MafG, MafK): History, structure and function. *Gene.* **586**(2), 197–205. https://doi.org/10.1016/j.gene.2016.03.058 (2016).
57. von Scheidt, M. et al. Transcription factor MAFF (MAF Basic Leucine Zipper Transcription Factor F) regulates an atherosclerosis relevant network connecting inflammation and cholesterol metabolism. *Circulation.* **143**(18), 1809–1823. https://doi.org/10.1161/CIRCULATIONAHA.120.050186 (2021).
58. Konhilas, J. P., Boucek, D. M., Horn, T. R., Johnson, G. L. & Leinwand, L. A. The role of MEKK1 in hypertrophic cardiomyopathy. *Int. Heart J.* **51**(4), 277–284. https://doi.org/10.1536/ihj.51.277 (2010).
59. Raffaello, A. et al. JunB transcription factor maintains skeletal muscle mass and promotes hypertrophy. *J. Cell Biol.* **191**(1), 101–113. https://doi.org/10.1083/jcb.201001136 (2010).
60. Seidenberg, J. et al. The AP-1 transcription factor Fosl-2 regulates autophagy in cardiac fibroblasts during myocardial fibrogenesis. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms22041861 (2021).
61. Stellato, M. et al. The AP-1 transcription factor Fosl-2 drives cardiac fibrosis and arrhythmias under immunofibrotic conditions. *Commun. Biol.* **6**(1), 161. https://doi.org/10.1038/s42003-023-04534-6 (2023).
62. Liu, S. et al. SOCS3 negatively regulates cardiac hypertrophy via targeting GRP78-Mediated ER stress during pressure overload. *Front Cell Dev. Biol.* **9**, 629932. https://doi.org/10.3389/fcell.2021.629932 (2021).
63. Pedroso, J. A. B. et al. SOCS3 ablation in leptin receptor-expressing cells causes autonomic and cardiac dysfunctions in middle-aged mice despite improving energy and glucose metabolism. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms23126484 (2022).
64. Cohn, R. et al. A contraction stress model of hypertrophic cardiomyopathy due to sarcomere mutations. *Stem Cell Rep.* **12**(1), 71–83. https://doi.org/10.1016/j.stemcr.2018.11.015 (2019).
65. Harper, A. R. et al. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat. Genet.* **53**(2), 135–142. https://doi.org/10.1038/s41588-020-00764-0 (2021).
66. Tong, Y. F. et al. Cyclin-dependent kinase inhibitor p21WAF1/CIP1 facilitates the development of cardiac hypertrophy. *Cell Physiol. Biochem.* **42**(4), 1645–1656. https://doi.org/10.1159/000479407 (2017).
67. Aung, N. et al. Genome-wide analysis of left ventricular maximum wall thickness in the UK biobank cohort reveals a shared genetic background with hypertrophic cardiomyopathy. *Circ. Genom. Precis. Med.* **16**(1), e003716. https://doi.org/10.1161/CIRCGEN.122.003716 (2023).
68. Hunten, S. et al. p53-regulated networks of protein, mRNA, miRNA, and lncRNA expression revealed by integrated pulsed stable isotope labeling with amino acids in cell culture (pSILAC) and next generation sequencing (NGS) analyses. *Mol. Cell Proteom.* **14**(10), 2609–2629. https://doi.org/10.1074/mcp.M115.050237 (2015).
69. Khan, S. S. et al. Identification of a novel familial fibrotic cardiomyopathy with a loss-of-function mutation in. *J. Card Fail.* **23**(8), S3–S3. https://doi.org/10.1016/j.cardfail.2017.07.004 (2017).
70. Khan, S. S. et al. Identification of cardiac fibrosis in young adults with a homozygous frameshift variant in SERPINE1. *JAMA Cardiol.* **6**(7), 841–846. https://doi.org/10.1001/jamacardio.2020.6909 (2021).
71. Sneath, R. J. & Mangham, D. C. The normal structure and function of CD44 and its role in neoplasia. *Mol. Pathol.* **51**(4), 191–200. https://doi.org/10.1136/mp.51.4.191 (1998).
72. Weng, X. et al. Inhibition of CD44 attenuates pressure overload-induced cardiac and lung inflammation, fibrosis, and heart failure progression. *Eur. Heart J.* **41**, 878–878 (2020).
73. Aydin, S., Ugur, K., Aydin, S., Sahin, I. & Yardim, M. Biomarkers in acute myocardial infarction: Current perspectives. *Vasc. Health Risk Man.* **15**, 1–10. https://doi.org/10.2147/Vhrm.S166157 (2019).
74. Gong, X. J., Song, X. Y., Wei, H., Wang, J. & Niu, M. Serum S100A4 levels as a novel biomarker for detection of acute myocardial infarction. *Eur. Rev. Med. Pharm.* **19**(12), 2221–2225 (2015).
75. Wei, X. et al. Myocardial hypertrophic preconditioning attenuates cardiomyocyte hypertrophy and slows progression to heart failure through upregulation of S100A8/A9. *Circulation.* **131**(17), 1506–1517. https://doi.org/10.1161/Circulationaha.114.013789 (2015).
76. Gao, J., Collyer, J., Wang, M., Sun, F. & Xu, F. Genetic dissection of hypertrophic cardiomyopathy with myocardial RNA-Seq. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms21093040 (2020).
77. Liu, X. et al. Lineage-specific regulatory changes in hypertrophic cardiomyopathy unraveled by single-nucleus RNA-seq and spatial transcriptomics. *Cell Discov.* **9**(1), 6. https://doi.org/10.1038/s41421-022-00490-3 (2023).
78. Lu, J. et al. High-resolution single-cell transcriptomic survey of cardiomyocytes from patients with hypertrophic cardiomyopathy. *Cell Prolif.* **57**(3), e13557. https://doi.org/10.1111/cpr.13557 (2024).
79. Sharpe, A. N. et al. Pharmacokinetics of a single dose of Aficamten (CK-274) on cardiac contractility in a A31P MYBPC3 hypertrophic cardiomyopathy cat model. *J. Vet. Pharmacol. Ther.* **46**(1), 52–61. https://doi.org/10.1111/jvp.13103 (2023).
80. Sharpe, A. N. et al. Effects of Aficamten on cardiac contractility in a feline translational model of hypertrophic cardiomyopathy. *Sci. Rep.* **13**(1), 32. https://doi.org/10.1038/s41598-022-26630-z (2023).

81. Kaplan, J. L., Rivas, V. N. & Connolly, D. J. Advancing treatments for feline hypertrophic cardiomyopathy: The role of animal models and targeted therapeutics. *Vet. Clin. N. Am. Small Anim. Pract.* **53**(6), 1293–1308. https://doi.org/10.1016/j.cvsm.2023.05.011 (2023).
82. Kawas, R. F. et al. A small-molecule modulator of cardiac myosin acts on multiple stages of the myosin chemomechanical cycle. *J. Biol. Chem.* **292**(40), 16571–16577. https://doi.org/10.1074/jbc.M117.776815 (2017).
83. Woodland, M. & Al-Horani, R. A. New Era: Mavacamten for obstructive hypertrophic cardiomyopathy. *Cardiovasc. Hematol. Agents Med. Chem.* **21**(2), 78–83. https://doi.org/10.2174/1871525721666221019095218 (2023).
84. Olivotto, I. et al. Mavacamten for treatment of symptomatic obstructive hypertrophic cardiomyopathy (EXPLORER-HCM): A randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet.* **396**(10253), 759–769. https://doi.org/10.1016/S0140-6736(20)31792-X (2020).
85. Tuohy, C. V., Kaul, S., Song, H. K., Nazer, B. & Heitner, S. B. Hypertrophic cardiomyopathy: The future of treatment. *Eur. J. Heart Fail.* **22**(2), 228–240. https://doi.org/10.1002/ejhf.1715 (2020).
86. Hegde, S. M. et al. Effect of mavacamten on echocardiographic features in symptomatic patients with obstructive hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.* **78**(25), 2518–2532. https://doi.org/10.1016/j.jacc.2021.09.1381 (2021).
87. Stephens, S. B. & Wallen, K. Environmental and social influences on neuroendocrine puberty and behavior in macaques and other nonhuman primates. *Horm. Behav.* **64**(2), 226–239. https://doi.org/10.1016/j.yhbeh.2013.05.003 (2013).
88. Dixson, A. F. & Nevison, C. M. The socioendocrinology of adolescent development in male rhesus monkeys (*Macaca mulatta*). *Horm. Behav.* **31**(2), 126–135. https://doi.org/10.1006/hbeh.1997.1374 (1997).

Author contributions

J.R.R., J.A.R., and J.A.S. conceived the experiment(s). V.N.R., M.W.V., Y.U., J.L.K., J.R.R., and J.A.S., conducted the experiment(s). V.N.R., M.W.V., and J.A.S. analyzed the results. V.N.R., W.W.V., and J.A.S. wrote the original manuscript. All authors reviewed and revised the final manuscript.

Funding

This work was supported by the California National Primate Research Center at the University of California-Davis (Davis, CA) under base grant #P51OD011107 and MyoKardia Inc. Translational Research Grant-In-Aid (#A19-4267); industry sponsors did not have any influence into the results or conclusions of this study. Support for V.N.R. was provided by NIH T32 HL086350 and TL1 TR001861. Support for J.L.K. was provided by NIH T32 OD011147 and Morris Animal Foundation D22FE-027.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-82770-4.

Correspondence and requests for materials should be addressed to J.A.S.