Comparative Medicine
Copyright 2020
by the American Association for Laboratory Animal Science

Vol 70, No 5
October 2020
Pages 358–367

Original Research

MYBPC3 Haplotype Linked to Hypertrophic Cardiomyopathy in Rhesus Macaques (*Macaca mulatta*)

Robert F Oldt,[1,2,*] Kimberly J Bussey,[1,3] Matthew L Settles,[4] Joseph N Fass,[4] Jeffrey A Roberts,[5] J Rachel Reader,[5] Srivathsan Komandoor,[6] Victor A Abrich,[6] and Sreetharan Kanthaswamy,[1,2,5]

In humans, abnormal thickening of the left ventricle of the heart clinically defines hypertrophic cardiomyopathy (HCM), a common inherited cardiovascular disorder that can precede a sudden cardiac death event. The wide range of clinical presentations in HCM obscures genetic variants that may influence an individual's susceptibility to sudden cardiac death. Although exon sequencing of major sarcomere genes can be used to detect high-impact causal mutations, this strategy is successful in only half of patient cases. The incidence of left ventricular hypertrophy (LVH) in a managed research colony of rhesus macaques provides an excellent comparative model in which to explore the genomic etiology of severe HCM and sudden cardiac death. Because no rhesus HCM-associated mutations have been reported, we used a next-generation genotyping assay that targets 7 sarcomeric rhesus genes within 63 genomic sites that are orthologous to human genomic regions known to harbor HCM disease variants. Amplicon sequencing was performed on 52 macaques with confirmed LVH and 42 unrelated, unaffected animals representing both the Indian and Chinese rhesus macaque subspecies. Bias-reduced logistic regression uncovered a risk haplotype in the rhesus *MYBPC3* gene, which is frequently disrupted in both human and feline HCM; this haplotype implicates an intronic variant strongly associated with disease in either homozygous or carrier form. Our results highlight that leveraging evolutionary genomic data provides a unique, practical strategy for minimizing population bias in complex disease studies.

Abbreviations: HCM, hypertrophic cardiomyopathy; LVH, left ventricular hypertrophy; SNP, single-nucleotide polymorphism

DOI: 10.30802/AALAS-CM-19-000108

Comparative medicine relies on the evolutionary relationships between humans and a variety of animal species to address health issues that affect all species involved. Such studies are particularly useful for conditions with a genetic risk component because they provide a way to isolate environmental exposures from inherited risk due to the differences in both population structure and exposure between species. Hypertrophic cardiomyopathy (HCM), a common cardiovascular disease with a range of clinical manifestations principally defined by abnormal thickening of the heart's left ventricle (left ventricular hypertrophy [LVH]), affects 0.2% of the human population[34] and is inherited in families typically as an autosomal dominant trait with variable penetrance and expressivity.[85] HCM and the mutations associated with it can be screened by sequencing cardiac sarcomere genes responsible for heart muscle function. Variations in 11 primary sarcomeric genes are associated with HCM, frequently affecting the genes responsible for β-myosin heavy chain (*MYH7*) and myosin-binding (*MYBPC3*) proteins.[62,64] Regions of the genome outside the sarcomere units have limited influence on the HCM phenotype.[101] Despite the disease's prevalence and strong association with sudden cardiac death, a leading cause of natural death in the United States, genetic testing strategies that target these genes detect causal pathogenic variants in only approximately half of patients with HCM.[34,62,93] Establishing a clear relationship between genotype and phenotype remains elusive in HCM, where the large intersection between human life habits and cardiovascular health creates environmental confounding effects unique to each patient.[1] Even identical HCM mutations have been demonstrated to produce differing disease phenotypes.[6,14]

The inability to detect genetic markers linked to human HCM can also be attributed to 2 foundational genomic problems. First, the clinical exome sequencing assays used to find patient mutations focus particularly on variants directly affecting protein structure, that is, the exonic and exon–intron boundaries of the sarcomere genes.[22] Comprehensive sequencing endeavors over the last 4 y have highlighted the role of deep intronic variants in both HCM as well as other complex human diseases.[37,97] Even single-nucleotide polymorphisms (SNP) in conserved intron sequences can result in haploinsufficiency, where expression levels are reduced in patients heterozygous or homozygous for a pathogenic variant.[36] Whole-gene sequencing is necessary to

Received: 07 Nov 2019. Revision requested: 13 Jan 2020. Accepted: 07 Feb 2020.
[1]School of Mathematical and Natural Sciences and [2]Evolutionary Biology Graduate Program, School of Life Sciences, and [3]BEYOND Center for Fundamental Concepts in Science, Arizona State University at the West Campus, Glendale, Arizona; [4]Bioinformatics Core, UC Davis Genome Center, and [5]California National Primate Research Center, University of California, Davis, California; and [6]Division of Cardiovascular Diseases, Mayo Clinic, Scottsdale, Arizona
[*]Corresponding author. Email: roldt@asu.edu

capture mutations affecting protein regulation, such as those in repressor or enhancer regions and at sites associated with transcription factors and long noncoding RNA sequences.[31,48]

The second major compounding factor in human HCM studies is genetic variant misclassification due to population structure. Differences in ancestry between groups of people is known to cause spurious genotype–phenotype correlations, especially when using small sample sizes.[23,92] This phenomenon has significantly inflated the number of purported risk alleles for HCM, evidenced by the high false-positive rate observed in persons of non-European ancestry.[61] The inclusion of even small numbers of control persons of different ancestry would be expected to reduce the amount of misclassified HCM risk alleles.[61] In particular, alleles defined according to exome sequence data are vulnerable to pathogenicity misclassification[4] and can be difficult to distinguish from statistical noise.[42] Whole-gene sequencing methods are demonstrably more capable of detecting biomedically relevant splice variants compared with exome-based approaches.[10]

High-throughput sequencing coupled with enrichment through PCR amplification—i.e., targeted or amplicon sequencing—is a cost-effective method for quickly generating vast quantities of genomic data spanning exonic and noncoding sites. Custom primer pairs can be optimized for uncharacterized genetic variation across diverse populations[15] and can produce quality genotypes from DNA target regions that are 500 base pairs or smaller.[19] Due to the PCR amplification step coupled with sample-specific barcodes, expensive library constructions can be bypassed.[95] Therefore, microfluidic PCR-based sequencing has been established as a robust method for SNP genotyping,[12,16,19,59] and bioinformatic pipelines exist for the rapid analysis of such data.[95]

A comparative medicine approach to LVH offers the ability to mitigate the confounding environmental factors relevant to cardiovascular health in a research colony setting. Comparable manifestations of the disease have been limited to rare observations in nonmodel species, including dogs,[33,56] captive gorillas,[76] and owl monkeys.[46] Cats exhibit HCM naturally, typically due to high-frequency single-amino–acid substitutions inherited in an autosomal dominant fashion.[73,74] In addition, the feline model reflects the myocyte disarray and haploinsufficiency[45,72] observed in the human presentation of the disease.[3,67,70,90,99] Mice genetically engineered to have HCM mutations have been used to model the effects of various sarcomere mutations[100] and haploinsufficiency phenotypes,[99] but rodent models differ significantly from humans in regard to myosin composition and contractile function.[18,60]

During the past 5 y, reviews of records at the California National Primate Research Center have identified numerous rhesus macaques (*Macaca mulatta*) that experienced sudden cardiac death and were found to have HCM, specifically LVH.[41,83] Rhesus macaques serve as valuable biomedical models owing to their close evolutionary relationship with humans[49] and lack of environmental confounding effects, such as differences in diet and activity levels.[51] With a 93% DNA sequence similarity to that of humans, the rhesus macaque genome has been characterized by using whole-genome sequencing[84,104] as well as through copy number variant[78] and SNP mapping techniques[40,79] to define orthologous regions in high resolution. In addition, rhesus macaques share the same genetic susceptibility to complex diseases, such as macular degeneration, as humans.[25] Pedigree analysis has demonstrated that approximately 1% to 2% of the rhesus colony at the California National Primate Research Center is affected by LVH, with the inflated prevalence of the disease in these macaques compared with humans explained by colony founding effects.[41,83] Clinical analysis of the LVH-affected animals in regard to physiologic variance[94] and biomarker recovery[29] is ongoing, but currently, no work has been published regarding the genomic etiology of rhesus HCM.

Materials and Methods

Candidate sequencing targets. We aggregated 428 human variants associated with LVH by using the NCBI ClinVar database[50] for 8 sarcomere genes: *MYBPC3*, *MYH7*, *TNNT2*, *TPM1*, *TNNI3*, *ACTC1*, *MYL3*, and *MYL2*.[35,63,69,71] Variant flanking sites in each gene were mapped to the human reference genome hg19 featured on the UCSC Genome Browser[43] by using the Burrow–Wheeler aligner (bwa) v.0.7.13[53] to ensure unique mapping of a single allele. We found 11 variants in gene *MYH7* to have mapping qualities greater than 1; these were treated as paralogous sites and subsequently removed from further analysis to prevent downstream inflation of association significance. Of the remaining 417 human sarcomere sites, 63 mapped uniquely to the rheMac3 reference genome, which also was accessed by using the UCSC Genome Browser. These 63 orthologous gene regions were treated as the core candidate sequencing sites to screen for pathogenic variation (Figure 1).

Samples. A total of 52 Indian-derived rhesus macaques from the CNPRC confirmed to have LVH[41,83] were treated as a cohort of HCM-affected animals. These affected animals were derived from 4 different pedigrees representing 9 generations and, as such, these macaques reflected various degrees of relatedness to each other. We assembled a reference nonpathogenic sample set from unaffected rhesus macaques of Chinese (n = 13) or Indian (n = 29) ancestry that had no kinship to any of the affected study animals[39,105] as biologic replicates. DNA extracts comprising this sample set (n = 94) originally purified from blood samples of colony animals were quantified by using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA) and diluted to 25 to 50 ng/μL prior to sequencing.

This research followed the American Society of Primatologists' principles for the ethical treatment of primates. The animals studied have been managed in compliance with IACUC regulations, NIH guidelines, or the US Department of Agriculture regulations prescribing the humane care and use of laboratory animals.

Amplicon sequencing and haplotype calling. Deep targeted amplicon sequencing on a MiSeq 600 instrument was facilitated by using the Access Array System (Fluidigm, San Francisco, CA) at the UC Davis Genome Center DNA Technologies Core by using two 48×48 integrated fluidic circuits. Illumina assays designed to capture the 63 conserved candidate sites were augmented for the rhesus macaque genome through the Fluidigm workflow and condensed into 35 primer pairs targeting 150- to 500-bp sequence windows. These target-specific primer pairs were optimized and tagged for downstream pooling according to the Fluidigm Access Array System user manual. Reads were demultiplexed by sample by using barcodes and by primer pair by using adaptor sequences as described previously.[17] Raw FASTQ files were processed into consensus sequences by using the dbcAmplicons R script (https://github.com/msettles/dbcAmplicons) incorporating polymorphisms present in over 5% of reads to reduce the inclusion of common sequencing errors.[95] After manually removing poorly performing assay targets according to multiple sequence alignments of each amplicon, we used the reduce-amplicons option available within the dbcAmplicons repository (https://github.com/msettles/dbcAmplicons/blob/master/scripts/R/reduce_amplicons.R) to establish

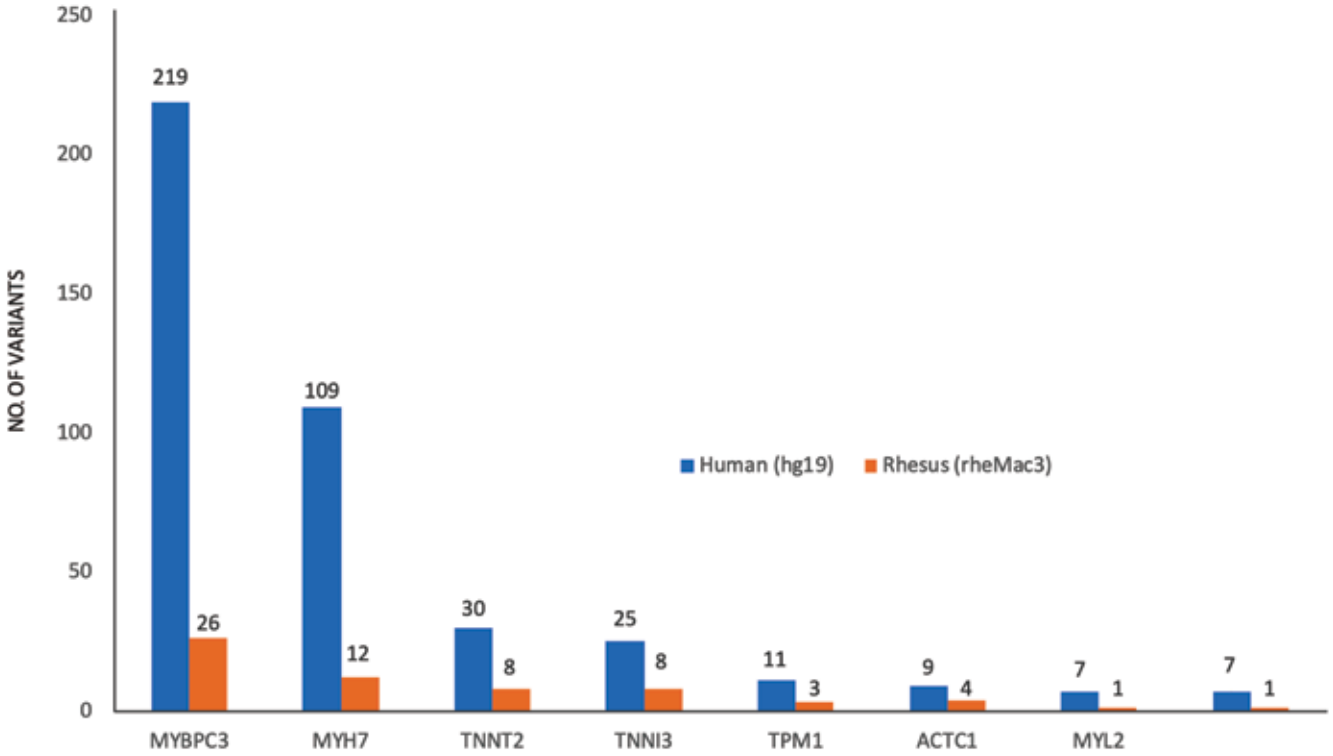

Figure 1. Bar plot displaying uniquely mapped human variants associated with LVH in the major sarcomere genes. The first column denotes 417 variants in complete alignment to human genome hg19; the second column denotes a subset of 63 variants in complete alignment to rhesus genome rheMac3. The overlapping orthologous variant sites were treated as candidate pathogenic gene regions for sequencing.

representative sequences. Paired-end reads were joined into a single contiguous sequence by using FLASH2 (https://github.com/dstreett/FLASH2) and standardized to the most frequent amplicon length per sample and primer pair. We ensured amplicon quality by removing assay regions amplified in fewer than 80 animals and by removing animals not represented by at least 25 amplicon sequences. Sequences represented by less than 8% of a target's reads were eliminated also. Due to the abundance of sequence reads per target region, homozygosity was confirmed when an allele was present at a frequency of 28% within each amplicon and when no alternative alleles were observed to be within 8% of the wildtype allele frequency. Sequences failing this test were considered heterozygous.

Results

Amplicon sequencing. A total of 23.3 million Illumina reads were generated and subsequently assessed for quality by using a LabChip GX. Approximately 80.9% of reads surpassed the Phred Q30 quality filter; 8.95% of reads represented PhiX library quality-control sequences. On average, more than 242,000 reads were produced for each rhesus macaque, representing an approximate coverage of 7000 reads per target region. Given that the average consensus sequence of each amplicon was 425 base pairs, this figure indicates more than 16-fold coverage across target gene regions. Individual amplicon performance per sample is represented by the heatmap in Figure 2.

Risk haplotype determination. After postsequence reduction and consensus sequence generation, we tested representative haplotypes denoting polymorphic sites for association with LVH by using Firth bias-reduced logistic regression.[24] Because colony macaques are expected to share genotypes as a consequence of close shared ancestry, typical logistic regression approaches may yield skewed odds ratio estimates. Firth regression, however, minimizes the effect of alleles shared through identity-by-descent in LVH-affected animals. In our analysis, we incorporated the Indian or Chinese ancestry status of each rhesus macaque as covariates to correct for population structure. Because the risk haplotype indeed may be segregating in the Chinese population, we repeated this experiment without any ancestry information and separately with only Indian-origin animals, that is, by excluding all Chinese animals from the analysis.

Only haplotypes represented in at least 3 animals, from target regions with at least 2 haplotypes, were used for statistical correlation to ensure high variant quality; no haplotypes from the *MYL3* gene met this criterion. After Bonferroni correction for multiple comparisons,[32] which represents a stricter threshold for statistical significance compared with other procedures that mitigate false-positive results,[11] we found a haplotype within the *MYBPC3* gene that was highly associated with the disease in both homozygous (adjusted P value = 0.004) and carrier (adjusted P value = 0.005) animals. Significant association between the *MYBPC3* haplotype and LVH was observed also, both among homozygous (adjusted P value = 0.004) and heterozygous (adjusted P value = 0.005) animals, when neither ancestry information nor Chinese animals were considered. Two additional potential risk haplotypes were identified in *MYBPC3* as well as in intronic regions of genes *TPM1*, *TNNT2*, and *MYH7*, but these results were not statistically significant after Bonferroni correction (Figure 3).

Functional inference. We inspected the significant *MYBPC3* risk haplotype, which spanned exon–intron flanking regions (chr14: 18690553 to 18690755), in the Ensembl genome browser by using the available rhesus macaque genome build rheMac8. *MYBPC3* is a gene known to undergo heavy alternative splicing,[88] in which introns near actively spliced exons contain high numbers of pathogenic HCM mutations.[26] The core C-to-T transition

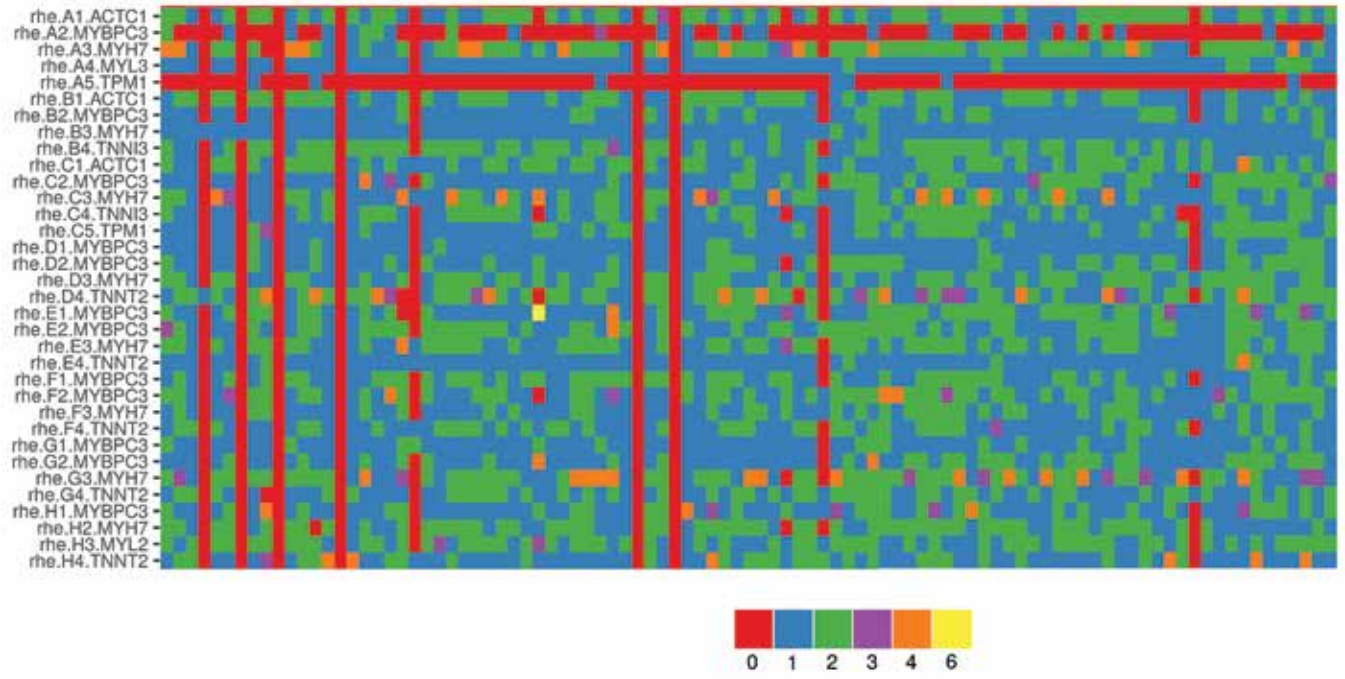

Figure 2. Heat map designating performance of each target sequence site by sample (*y*-axis) and primer pair (*x*-axis). Colors indicate the number of unique amplicon sequences that differ by one or more ambiguous positions. 'Failure to produce amplicons' (in red) represents sample-specific behavior, rather than stochastic or assay-specific nonamplification (as observed in target sites *rhe.A2.MYBPC3* and *rhe.A5.TPM1*).

variant within the haplotype is an intronic allele immediately downstream of an alternatively spliced exon designated as exon 24, 25, and 22 in the *MYBPC3*-201(ENSMMUT00000021887.3), 202 (ENSMMUT00000059815.1), and 203 (ENST00000544791.1) isoforms, respectively. Given the acute effect of mutations on RNA splicing capabilities[55,98] and lack of publicly available tools to specifically analyze the functional effects of rhesus variation, we assessed the pathogenic polymorphism observed in the primary *MYBPC3* risk haplotype for transcriptional disruption potential by using the free-use program Spliceman (http://fairbrother.biomed.brown.edu/spliceman/). Spliceman uses positional distribution analysis based on hexamers proximate to splice sites and generates L1 distances based on hexamer frequency that are converted to splicing mutation likelihoods from 1 to 100.[54] The program assimilated both the wild-type and risk variant *MYBPC3* sequences to evaluate the effect of the hexamer modification. In rheMac2, the rhesus reference genome available in Spliceman, we found the *MYBPC3* variant to have a 65% probability of causing splicing disruption.

Discussion

Among a subset of animals from the rhesus macaque colony at the California National Primate Research Center that exhibited LVH, we detected a single, high-quality risk allele significantly associated with the disease. Our approach leveraged 1) extant orthologous genomic variation across major subpopulations of the animal species under study, 2) comparative human genetic data to narrow down the number of pathogenic candidate sites, 3) deep sequencing technology inclusive of exons and introns, capable of producing high-quality and high-coverage results across a diversity of species,[15] and 4) a nonbiased correlation analysis preventing false-positive discovery due to genotypes shared by ancestry. This approach is essential in light of the fact that rhesus macaques exhibit greater than 2-fold increased nucleotide diversity compared with humans, including in functional regions,[104] thus impeding confidence in variant calling using singular reference genomes.[13]

We found that the *MYBPC3* gene, which is the major source of HCM mutations in humans,[7,35,71] was strongly associated with LVH in rhesus macaques. More than 147 mutations in this gene have been reported in humans alone, accounting for between 15%[91] to more than one third[57] of patient cases. Furthermore, human HCM mutations present in populations at high frequencies due to founder effects are primarily localized in the *MYBPC3* gene.[21,75] Although prior studies have neither identified a clear mode of inheritance nor the genetic variants responsible for rhesus LVH, founder effects best explain the prevalence of LVH in the rhesus colony at the California National Primate Research Center.[41]

Mutations in *MYBPC3* also are associated with increased variability in human subjects' age when the diagnostic test was performed,[80] correlating with the wide age span of rhesus animals diagnosed with LVH.[83] Because 78% of the databased disease-associated variants in *MYBPC3* represent intronic, synonymous, or rare missense mutations,[37] successful sequencing strategies for HCM across species must prioritize nonexonic elements as well as exons and exon–intron boundary sites.

We detected a candidate sequencing region targeting a splicing-relevant exon in *MYBPC3* that was highly correlated to disease status (Figure 3). This result is consistent with previous findings of an elevated concentration of HCM mutations located in intron regions flanking the *MYBPC3* exons involved in alternative splicing.[26] The target region amplicon encompassed both the upstream and downstream intron boundaries of the exon, exhibiting a deep intron variant highly correlated with LVH in both homozygous and carrier macaques. Intron mutations such as these are known to disrupt the regulation of expression by modifying repressor and depressor elements or introducing cryptic splice sites.[97] Genetic changes that might alter RNA splicing have been characterized extensively in *MYBPC3*.[27,37,58]

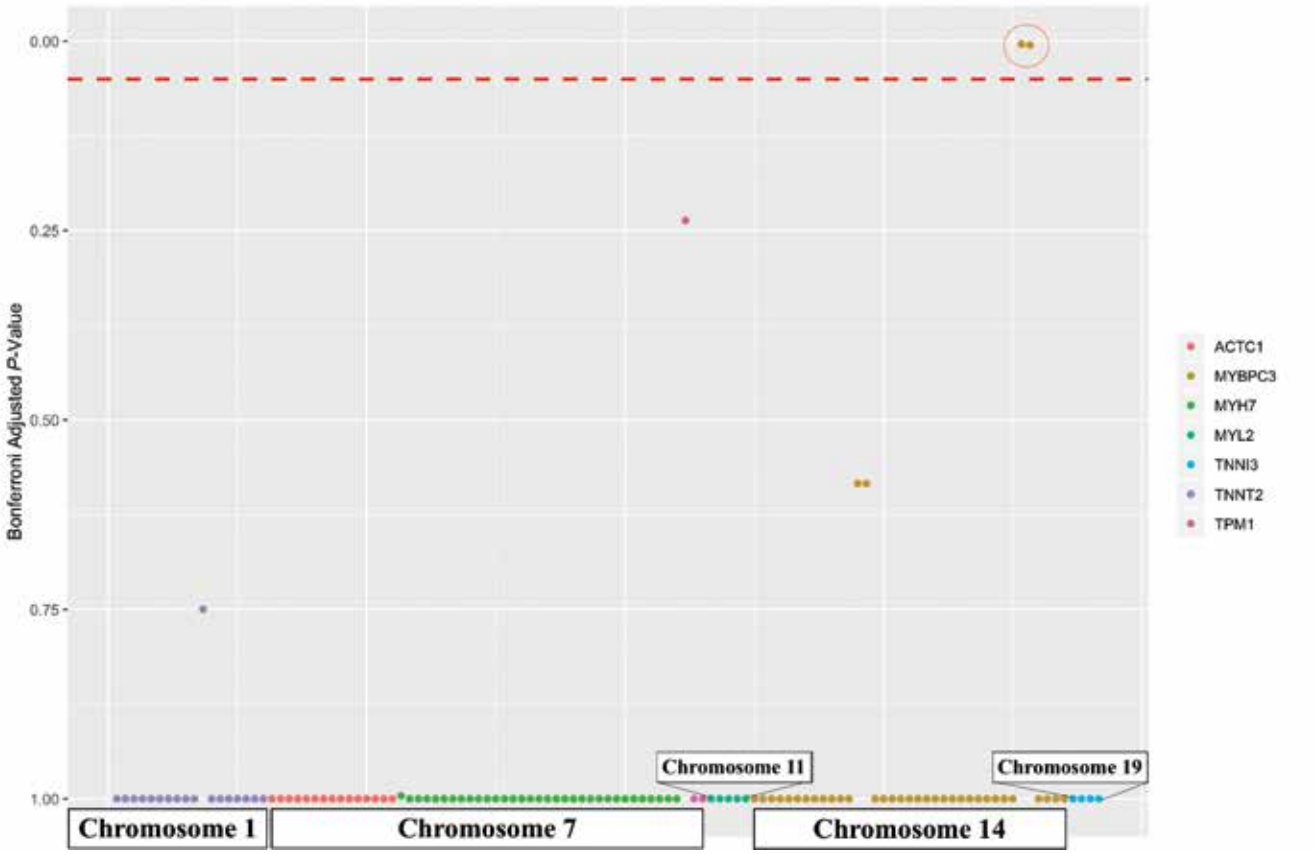

Figure 3. Amplicon-disease status correlation graph comprising all haplotypes tested for association with rhesus LVH incidence by using non-biased regression. Risk haplotypes statistically associated with animals exhibiting LVH are shown above the red line, which designates the significance threshold of $P = 0.05$ after Bonferroni adjustment for multiple corrections. Colors indicate individual sarcomere genes.

The C6 domain of *MYBPC3*[20] contains the highest proportion of clinical variants typed for HCM[30,77] in the gene; this domain is where the rhesus risk haplotype was localized. The intronic space within the exon 24–25 junction of *MYBPC3*, the exact same location of the high-association mutation in rhesus macaques, has specifically been linked to haploinsufficiency and abnormal splicing in HCM.[96]

The intronic risk allele was implicated with aberrant splicing behavior, because it overlapped at least 3 different isoforms of the *MYBPC3* transcript and exhibited a significant likelihood of disrupting regulatory activity. An algorithm capable of analyzing rhesus macaque intron data to predict cryptic splicing effects returned a pathogenic probability score (65%) comparable to splice-site mutations linked to other complex diseases, such campomelic dysplasia (67%),[68] Müllerian aplasia (69%),[87] and nephropathic cystinosis (55% to 85%).[89] In a genome-wide association study of patients with chronic otitis media, noncoding variants were highly associated with the disease and confirmed to be pathogenic by using the same predictive splicing algorithm.[2] Deep intron mutations, as revealed by large-scale sequencing, have important functional significance in complex disease[5] and in human HCM specifically.[71]

Whereas sarcomere genes acquire missense mutations that lead to altered protein structure or 'poison peptides,'[63] *MYBPC3* mutations are commonly associated with haploinsufficiency as the primary mechanism for human HCM[65-67] and some cases of feline[72] and mouse HCM.[9] Allelic imbalance, a phenomenon in which alleles exhibit differential transcription levels due to polymorphisms, has been linked to human HCM and the *MYBPC3* gene as a cause of functional haploinsufficiency.[28] Human *MYBPC3* mutations result in heterogeneous clinical outcomes[80] and diminished or variable protein expression levels in human patients.[9,69,96] This clinical variability due to alterations in expression levels is consistent with recent observations in rhesus macaques at the California National Primate Research Center, in which neither major cardiac biomarkers[29] nor physiologic metrics[94] are predictive of LVH.

Although direct evidence of alternative *MYBPC3* splicing or reduced mRNA levels would have supported our haploinsufficiency hypothesis for HCM in Indian rhesus macaques, these approaches were not feasible because the current retrospective study relied on samples that had been stored for decades. Cardiac tissue samples from recently diagnosed animals would be ideal for confirming that transcript levels are significantly reduced in affected compared with unaffected macaques. However, it is a noteworthy caution that the heterogeneity of disease presentation observed in the affected study cohort may further obstruct functional studies.[45,54,55]

A previous pedigree analysis involving LVH-affected animals at the California National Primate Research Center was unable to confirm an autosomal dominant inheritance pattern due to genetic drift creating a higher prevalence of the disease.[41] However our current results affirm that LVH in rhesus macaques is an autosomal dominant trait due to the strong correlation between a deep intron variant, in either homozygous or heterozygous form, and the disease. This mode of inheritance, coupled with the cross-species relevance of the *MYBPC3* gene in HCM, establishes that the rhesus macaques at the California National

Primate Research Center are a suitable and uniquely useful model for human HCM.

This comprehensive analysis of sarcomere variants in rhesus macaques with LVH underscores the critical need for full gene sequencing and population genomic variation in complex disease. Exonic and exon–intron flanking mutations are known to have devastating consequences in the HCM and sudden cardiac death phenotypes, and next-generation sequencing has confirmed a strong intronic component to human HCM.[37,71,86] In general, access to high-throughput sequencing technologies has led to the discovery of deep intron variants highly correlated with a disease,[103] and in genes that undergo extensive splicing, such as those that encode sarcomeres, these variants may activate dormant splice sites as well as alter enhancer or depressor regions.[8,52,97,102] This notion is consistent with the variety of ages among the colony rhesus macaques that experienced sudden cardiac death.[83]

Recent advancements highlighting the role of intronic and noncoding-region mutations in HCM[44] warrant further deep sequencing of subjects—both humans and other animals—that exhibit clear clinical signatures such as LVH. In addition, haplotype-based methods can be used to efficiently investigate loci for disease association, because genetic studies of HCM require that the recovery of combinations of alleles is precise.[69] The discovery of more disease variants and the expansion of reference genomic information for biomedically relevant model species is crucial for accurate and replicative sequencing results. Not only do therapeutic strategies and in vivo knockout models using *MYBPC3*[38] need to account for confounding intronic effects on phenotype, but the efficacy of resources for determining functional consequences of mutations is dependent on robust reference genomic data. Our amplicon sequencing strategy has been shown to be a powerful alternative to whole-exome approaches, which ignore most intronic sequences and have a substantial fail rate for causal mutation recovery in human HCM patients.

Although the LVH-affected rhesus macaques we studied may carry exonic mutations in untested sarcomere regions or non-sarcomere genes, no exonic variant interrogated in our study reached statistical significance when factoring population structure and unbiased genotype correlation. Even if it is not the primary causal HCM mutation, the intronic risk allele identified in *MYBPC3* may be linked to another locus involved with LVH susceptibility, as similarly postulated in other macaque complex disease studies.[25] Alternatively, this disease-associated allele could represent an ancestry-biased variant indicative of the Indian macaque founding population responsible for the transmittance of LVH in the rhesus colony at the California National Primate Research Center. These suppositions can be explored comprehensively in future research by combining high-resolution linkage analysis methods with multiplex LVH pedigree reconstructions (based on parentage, clinical, and pathologic data[44,45]) and colony-wide genotype information from a rapid detection SNP technique, such as that which exists for ABO blood group determination in both rhesus and cynomolgus macaques (*M. fascicularis*).[81,82]

The recent description of a cynomolgus (*M. fascicularis*) macaque in China with HCM and LVH[47] accentuates the applicability of, and need for, genotyping tools capable of capturing inter- and intraspecies genomic variation. Cynomolgus macaques housed at Chinese breeding centers exhibit substantial rhesus ancestry due to their origin in the hybridization zone between the 2 species located in Laos, Cambodia, and Vietnam.[105] If pathogenic variants for HCM have been shared across different biomedical model macaque species, sequence-based approaches are critical to assessing disease risk, in large part owing to the heterogeneous clinical manifestation of the disease. Our discovery of a deep intron variant heavily associated with LVH and HCM in rhesus macaques represents an overall success in using primate evolutionary genomic data to focus on high-probability candidate regions for next-generation sequencing.

Acknowledgments

We thank the staff of the Evolutionary and Forensic Genetics Lab at Arizona State University–West for their aid with sample preparation and the Genome Center Biotechnology and Bioinformatics core labs at University of California–Davis and Fluidigm Corporation for their assistance in optimizing our analysis and bioinformatics pipeline. This study was supported by NIH CNPRC base grant OD 5P51OD011107 and by the ASU Athletics Research Grant program graduate student award to RFO.

References

1. **Alcalai R, Seidman JG, Seidman CE.** 2008. Genetic basis of hypertrophic cardiomyopathy: from bench to the clinics. J Cardiovasc Electrophysiol **19:**104–110.
2. **Allen EK, Chen WM, Weeks DE, Chen F, Hou X, Mattos JL, Mychaleckyj JC, Segade F, Casselbrant ML, Mandel EM, Ferrell RE, Rich SS, Daly KA, Sale MM.** 2013. A genome-wide association study of chronic otitis media with effusion and recurrent otitis media identifies a novel susceptibility locus on chromosome 2. J Assoc Res Otolaryngol **14:**791–800. https://doi.org/10.1007/s10162-013-0411-2.
3. **Andersen PS, Havndrup O, Bundgaard H, Larsen LA, Vuust J, Pedersen AK, Kjeldsen K, Christiansen M.** 2004. Genetic and phenotypic characterization of mutations in myosin-binding protein C (MYBPC3) in 81 families with familial hypertrophic cardiomyopathy: total or partial haploinsufficiency. Eur J Hum Genet **12:**673–677. https://doi.org/10.1038/sj.ejhg.5201190.
4. **Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, Sajadieh A, Haunso S, Svendsen JH, Olesen MS.** 2013. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. Eur J Hum Genet **21:**918–928. https://doi.org/10.1038/ejhg.2012.283.
5. **Anna A, Monika G.** 2018. Splicing mutations in human genetic disorders: examples, detection, and confirmation. J Appl Genet **59:**253–268. https://doi.org/10.1007/s13353-018-0444-7. Correction: J Appl Genet 2019.**60:**231.
6. **Arad M, Seidman JG, Seidman CE.** 2002. Phenotypic diversity in hypertrophic cardiomyopathy. Hum Mol Genet **11:**2499–2506. https://doi.org/10.1093/hmg/11.20.2499.
7. **Authors/Task Force members, Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, Hagege AA, Lafont A, Limongelli G, Mahrholdt H, McKenna WJ, Mogensen J, Nihoyannopoulos P, Nistri S, Pieper PG, Pieske B, Rapezzi C, Rutten FH, Tillmanns C, Watkins H.** 2014. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). Eur Heart J 35:2733–2779. https://doi.org/10.1093/eurheartj/ehu284.
8. **Baralle D, Baralle M.** 2005. Splicing in action: assessing disease causing sequence changes. J Med Genet **42:**737–748. https://doi.org/10.1136/jmg.2004.029538.
9. **Barefield D, Kumar M, Gorham J, Seidman JG, Seidman CE, de Tombe PP, Sadayappan S.** 2015. Haploinsufficiency of MYBPC3 exacerbates the development of hypertrophic cardiomyopathy in heterozygous mice. J Mol Cell Cardiol **79:**234–243. https://doi.org/10.1016/j.yjmcc.2014.11.018.
10. **Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, Shang L, Boisson B, Casanova JL, Abel L.** 2015. Whole-genome sequencing is more powerful than whole-exome sequencing for

detecting exome variants. Proc Natl Acad Sci USA **112:**5473–5478. https://doi.org/10.1073/pnas.1418631112.

11. **Benjamini Y, Hochberg Y.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol **57:**289–300.
12. **Bhat S, Polanowski AM, Double MC, Jarman SN, Emslie KR.** 2012. The effect of input DNA copy number on genotype call and characterising SNP markers in the humpback whale genome using a nanofluidic array. PLoS One **7:**1–9. https://doi.org/10.1371/journal.pone.0039181.
13. **Bimber BN, Ramakrishnan R, Cervera-Juanes R, Madhira R, Peterson SM, Norgren RB Jr, Ferguson B.** 2017. Whole genome sequencing predicts novel human disease models in rhesus macaques. Genomics **109:**214–220. https://doi.org/10.1016/j.ygeno.2017.04.001.
14. **Brito D, Richard P, Isnard R, Pipa J, Komajda M, Madeira H.** 2003. Familial hypertrophic cardiomyopathy: the same mutation, different prognosis. Comparison of two families with a long follow-up. Rev Port Cardiol **22:**1445–1461. [Article in English, Portuguese].
15. **Bybee SM, Bracken-Grissom H, Haynes BD, Hermansen RA, Byers RL, Clement MJ, Udall JA, Wilcox ER, Crandall KA.** 2011. Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. Genome Biol Evol **3:**1312–1323. https://doi.org/10.1093/gbe/evr106.
16. **Byers RL, Harker DB, Yourstone SM, Maughan PJ, Udall JA.** 2012. Development and mapping of SNP assays in allotetraploid cotton. Theor Appl Genet **124:**1201–1214. https://doi.org/10.1007/s00122-011-1780-8.
17. **Byrne AQ, Rothstein AP, Poorten TJ, Erens J, Settles ML, Rosenblum EB.** 2017. Unlocking the story in the swab: A new genotyping assay for the amphibian chytrid fungus *Batrachochytrium dendrobatidis*. Mol Ecol Resour **17:**1283–1292. https://doi.org/10.1111/1755-0998.12675.
18. **Camacho P, Fan H, Liu Z, He JQ.** 2016. Small mammalian animal models of heart disease. Am J Cardiovasc Dis **6:**70–80.
19. **Campbell NR, Harmon SA, Narum SR.** 2014. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Mol Ecol Resour **15:**855–867. https://doi.org/10.1111/1755-0998.12357.
20. **Carrier L, Bonne G, Bährend E, Yu B, Richard P, Niel F, Hainque B, Cruaud C, Gary F, Labeit S, Bouhour JB, Dubourg O, Desnos M, Hagège AA, Trent RJ, Komajda M, Fiszman M, Schwartz K.** 1997. Organization and sequence of human cardiac myosin binding protein C gene (MYBPC3) and identification of mutations predicted to produce truncated proteins in familial hypertrophic cardiomyopathy. Circ Res **80:**427–434. https://doi.org/10.1161/01.res.0000435859.24609.b3.
21. **Christiaans I, Nannenberg EA, Dooijes D, Jongbloed RJ, Michels M, Postema PG, Majoor-Krakauer D, van den Wijngaard A, Mannens MM, van Tintelen JP, van Langen IM, Wilde AA.** 2010. Founder mutations in hypertrophic cardiomyopathy patients in the Netherlands. Neth Heart J **18:**248–254. https://doi.org/10.1007/BF03091771.
22. **Coto E, Reguero JR, Palacin M, Gomez J, Alonso B, Iglesias S, Martin M, Tavira B, Diaz-Molina B, Morales C, Moris C, Rodriguez-Lambert JL, Corao AI, Diaz M, Alvarez V.** 2012. Resequencing the whole MYH7 gene (including the intronic, promoter, and 3′ UTR sequences) in hypertrophic cardiomyopathy. J Mol Diagn **14:**518–524. https://doi.org/10.1016/j.jmoldx.2012.04.001.
23. **Finno CJ, Aleman M, Higgins RJ, Madigan JE, Bannasch DL.** 2014. Risk of false positive genetic associations in complex traits with underlying population structure: a case study. Vet J **202:**543–549. https://doi.org/10.1016/j.tvjl.2014.09.013.
24. **Firth D.** 1993. Bias reduction of maximum likelihood estimates. *Biometrika* **80:**27–38. https://doi.org/10.1093/biomet/80.1.27.
25. **Francis PJ, Appukuttan B, Simmons E, Landauer N, Stoddard J, Hamon S, Ott J, Ferguson B, Klein M, Stout JT, Neuringer M.** 2008. Rhesus monkeys and humans share common susceptibility genes for age-related macular disease. Hum Mol Genet **17:**2673–2680. https://doi.org/10.1093/hmg/ddn167.
26. **Frank-Hansen R, Page SP, Syrris P, McKenna WJ, Christiansen M, Andersen PS.** 2008. Micro-exons of the cardiac myosin binding protein C gene: flanking introns contain a disproportionately large number of hypertrophic cardiomyopathy mutations. Eur J Hum Genet **16:**1062–1069. https://doi.org/10.1038/ejhg.2008.52.
27. **Frisso G, Detta N, Coppola P, Mazzaccara C, Pricolo MR, D'Onofrio A, Limongelli G, Calabro R, Salvatore F.** 2016. Functional studies and in silico analyses to evaluate non-coding variants in inherited cardiomyopathies. Int J Mol Sci **17:**1–13.
28. **Glazier AA, Thompson A, Day SM.** 2018. Allelic imbalance and haploinsufficiency in MYBPC3-linked hypertrophic cardiomyopathy. Pflugers Arch **471:**781–793. https://doi.org/10.1007/s00424-018-2226-9.
29. **Haertel AJ, Stern JA, Reader JR, Spinner A, Roberts JA, Christe KL.** 2016. Antemortem screening for left ventricular hypertrophy in rhesus macaques (*Macaca mulatta*). Comp Med **66:**333–342.
30. **Harris SP, Lyons RG, Bezold KL.** 2011. In the thick of it: HCM-causing mutations in myosin binding proteins of the thick filament. Circ Res **108:**751–764. https://doi.org/10.1161/CIRCRESAHA.110.231670.
31. **Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R, Hubbard TJ.** 2012. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res **22:**1760–1774. https://doi.org/10.1101/gr.135350.111.
32. **Haynes W.** 2013.Bonferroni correction, p 154–154. In: Dubitzky W, Wolkenhauer O, Yokota H, Cho K-H, editors. Encyclopedia of systems biology. New York (NY): Springer New York.
33. **Hittinger L, Ghaleb B, Chen J, Edwards JG, Kudej RK, Iwase M, Kim SJ, Vatner SF, Vatner DE.** 1999. Reduced subendocardial ryanodine receptors and consequent effects on cardiac function in conscious dogs with left ventricular hypertrophy. Circ Res **84:**999–1006. https://doi.org/10.1161/01.RES.84.9.999.
34. **Ho CY.** 2012. Hypertrophic cardiomyopathy in 2012. Circulation **125:**1432–1438. https://doi.org/10.1161/CIRCULATIONAHA.110.017277.
35. **Ho CY, Charron P, Richard P, Girolami F, Van Spaendonck-Zwarts KY, Pinto Y.** 2015. Genetic advances in sarcomeric cardiomyopathies: state of the art. Cardiovasc Res **105:**397–408. https://doi.org/10.1093/cvr/cvv025.
36. **Hsu AP, Johnson KD, Falcone EL, Sanalkumar R, Sanchez L, Hickstein DD, Cuellar-Rodriguez J, Lemieux JE, Zerbe CS, Bresnick EH, Holland SM**. 2013. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood **121:**3830–3837, S1–S7. https://doi.org/10.1182/blood-2012-08-452763
37. **Ito K, Patel PN, Gorham JM, McDonough B, DePalma SR, Adler EE, Lam L, MacRae CA, Mohiuddin SM, Fatkin D, Seidman CE, Seidman JG.** 2017. Identification of pathogenic gene mutations in *LMNA* and *MYBPC*$_3$ that alter RNA splicing. Proc Natl Acad Sci USA **114:**7689–7694. https://doi.org/10.1073/pnas.1707741114.
38. **Judge DP, Neamatalla H, Norris RA, Levine RA, Butcher JT, Vignier N, Kang KH, Nguyen Q, Bruneval P, Perier MC, Messas E, Jeunemaitre X, de Vlaming A, Markwald R, Carrier L, Hagege AA.** 2015. Targeted *Mybpc3* knock-out mice with cardiac hypertrophy exhibit structural mitral valve abnormalities. J Cardiovasc Dev Dis **2:**48–65. https://doi.org/10.3390/jcdd2020048.
39. **Kanthaswamy S, Johnson Z, Trask JS, Smith DG, Ramakrishnan R, Bahk J, Ng J, Wiseman R, Kubisch HM, Vallender EJ, Rogers J, Ferguson B.** 2014. Development and validation of a SNP-based assay for inferring the genetic ancestry of rhesus macaques (*Macaca mulatta*). Am J Primatol **76:**1105–1113. https://doi.org/10.1002/ajp.22290.
40. **Kanthaswamy S, Ng J, Ross CT, Trask JS, Smith DG, Buffalo VS, Fass JN, Lin D.** 2013. Identifying human-rhesus macaque gene orthologs using heterospecific SNP probes. Genomics **101:**30–37. https://doi.org/10.1016/j.ygeno.2012.09.001.

41. **Kanthaswamy S, Reader R, Tarara R, Oslund K, Allen M, Ng J, Grinberg C, Hyde D, Glenn DG, Lerche N.** 2014. Large scale pedigree analysis leads to evidence for founder effects of hypertrophic cardiomyopathy in rhesus macaques (*Macaca mulatta*). J Med Primatol **43:**288–291. https://doi.org/10.1111/jmp.12127.
42. **Kapplinger JD, Landstrom AP, Bos JM, Salisbury BA, Callis TE, Ackerman MJ.** 2014. Distinguishing hypertrophic cardiomyopathy-associated mutations from background genetic noise. J Cardiovasc Transl Res **7:**347–361. https://doi.org/10.1007/s12265-014-9542-z.
43. **Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D.** 2002. The human genome browser at UCSC. Genome Res **12:**996–1006. https://doi.org/10.1101/gr.229102.
44. **Kittleson MD, Meurs KM, Harris SP.** 2015. The genetic basis of hypertrophic cardiomyopathy in cats and humans. J Vet Cardiol **17 Suppl 1:**S53–S73. https://doi.org/10.1016/j.jvc.2015.03.001.
45. **Kittleson MD, Meurs KM, Munro MJ, Kittleson JA, Liu SK, Pion PD, Towbin JA.** 1999. Familial hypertrophic cardiomyopathy in Maine coon cats: an animal model of human disease. Circulation **99:**3172–3180. https://doi.org/10.1161/01.CIR.99.24.3172.
46. **Knowlen GG, Weller RE, Perry RL, Baer JF, Gozalo AS.** 2013. Hypertrophic cardiomyopathy in owl monkeys (*Aotus* spp.). Comp Med **63:**279–287.
47. **Konishi S, Kotera T, Koga M, Ueda M.** 2018. Spontaneous hypertrophic cardiomyopathy in a cynomolgus macaque (*Macaca fascicularis*). J Toxicol Pathol **31:**49–54. https://doi.org/10.1293/tox.2017-0027.
48. **Krupp DR, Soldano KL, Garrett ME, Cope H, Ashley-Koch AE, Gregory SG.** 2014. Missing genetic risk in neural tube defects: can exome sequencing yield an insight? Birth Defects Res A Clin Mol Teratol **100:**642–646. https://doi.org/10.1002/bdra.23276.
49. **Kumar S, Hedges SB.** 1998. A molecular timescale for vertebrate evolution. Nature **392:**917–920. https://doi.org/10.1038/31927.
50. **Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR.** 2015. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res **44 D1:**D862–D868. https://doi.org/10.1093/nar/gkv1222.
51. **Lankau EW, Turner PV, Mullan RJ, Galland GG.** 2014. Use of nonhuman primates in research in North America. J Am Assoc Lab Anim Sci **53:**278–282.
52. **Lee M, Roos P, Sharma N, Atalar M, Evans TA, Pellicore MJ, Davis E, Lam AN, Stanley SE, Khalil SE, Solomon GM, Walker D, Raraigh KS, Vecchio-Pagan B, Armanios M, Cutting GR.** 2017. Systematic computational identification of variants that activate exonic and intronic cryptic splice sites. Am J Hum Genet **100:**751–765. https://doi.org/10.1016/j.ajhg.2017.04.001.
53. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25:**1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
54. **Lim KH, Fairbrother WG.** 2012. Spliceman—a computational web server that predicts sequence variations in pre-mRNA splicing. Bioinformatics **28:**1031–1032. https://doi.org/10.1093/bioinformatics/bts074.
55. **Lim KH, Ferraris L, Filloux ME, Raphael BJ, Fairbrother WG.** 2011. Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. Proc Natl Acad Sci USA **108:**11093–11098. https://doi.org/10.1073/pnas.1101135108.
56. **Liu SK, Maron BJ, Tilley LP.** 1979. Hypertrophic cardiomyopathy in the dog. Am J Pathol **94:**497–508.
57. **Liu X, Jiang T, Piao C, Li X, Guo J, Zheng S, Zhang X, Cai T, Du J.** 2015. Screening mutations of *MYBPC*$_3$ in 114 unrelated patients with hypertrophic cardiomyopathy by targeted capture and next-generation sequencing. Sci Rep **5:**1–8. https://doi.org/10.1038/srep11411.
58. **Lopes LR, Zekavati A, Syrris P, Hubank M, Giambartolomei C, Dalageorgou C, Jenkins S, McKenna W, Uk10k Consortium; Plagnol V, Elliott PM.** 2013. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. J Med Genet **50:**228–239. https://doi.org/10.1136/jmedgenet-2012-101270.
59. **Lu R, Gao X, Chen Y, Ni J, Yu Y, Li S, Guo L.** 2012. Association of an *NFKB1* intron SNP (rs4648068) with gastric cancer patients in the Han Chinese population. BMC Gastroenterol **12:**1–7. https://doi.org/10.1186/1471-230X-12-87.
60. **Maass A, Leinwand LA.** 2000. Animal models of hypertrophic cardiomyopathy. Curr Opin Cardiol **15:**189–196. https://doi.org/10.1097/00001573-200005000-00012.
61. **Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, Margulies DM, Loscalzo J, Kohane IS.** 2016. Genetic misdiagnoses and the potential for health disparities. N Engl J Med **375:**655–665. https://doi.org/10.1056/NEJMsa1507092.
62. **Marian AJ, Braunwald E.** 2017. Hypertrophic cardiomyopathy: genetics, pathogenesis, clinical manifestations, diagnosis, and therapy. Circ Res **121:**749–770. https://doi.org/10.1161/CIRCRESAHA.117.311059.
63. **Marian AJ, Roberts R.** 2001. The molecular genetic basis for hypertrophic cardiomyopathy. J Mol Cell Cardiol **33:**655–670. https://doi.org/10.1006/jmcc.2001.1340.
64. **Maron BJ, Maron MS, Semsarian C.** 2012. Genetics of hypertrophic cardiomyopathy after 20 years: clinical perspectives. J Am Coll Cardiol **60:**705–715. https://doi.org/10.1016/j.jacc.2012.02.068.
65. **Marsiglia JD, Pereira AC.** 2014. Hypertrophic cardiomyopathy: how do mutations lead to disease? Arq Bras Cardiol **102:**295–304. [Article in English, Portuguese].
66. **Marston S, Copeland O, Gehmlich K, Schlossarek S, Carrier L.** 2011. How do *MYBPC3* mutations cause hypertrophic cardiomyopathy? J Muscle Res Cell Motil **33:**75–80. https://doi.org/10.1007/s10974-011-9268-3.
67. **Marston S, Copeland O, Jacques A, Livesey K, Tsang V, McKenna WJ, Jalilzadeh S, Carballo S, Redwood C, Watkins H.** 2009. Evidence from human myectomy samples that *MYBPC3* mutations cause hypertrophic cardiomyopathy through haploinsufficiency. Circ Res **105:**219–222. https://doi.org/10.1161/CIRCRESAHA.109.202440.
68. **Mattos EP, Sanseverino MT, Magalhaes JA, Leite JC, Felix TM, Todeschini LA, Cavalcanti DP, Schuler-Faccini L.** 2015. Clinical and molecular characterization of a Brazilian cohort of campomelic dysplasia patients, and identification of seven new SOX9 mutations. Genet Mol Biol **38:**14–20. https://doi.org/10.1590/S1415-475738120140147.
69. **McNally EM, Barefield DY, Puckelwartz MJ.** 2015. The genetic landscape of cardiomyopathy and its role in heart failure. Cell Metab **21:**174–182. https://doi.org/10.1016/j.cmet.2015.01.013.
70. **Mearini G, Stimpel D, Geertz B, Weinberger F, Kramer E, Schlossarek S, Mourot-Filiatre J, Stoehr A, Dutsch A, Wijnker PJ, Braren I, Katus HA, Muller OJ, Voit T, Eschenhagen T, Carrier L.** 2014. *Mybpc3* gene therapy for neonatal cardiomyopathy enables long-term disease prevention in mice. Nat Commun **5:**1–10. https://doi.org/10.1038/ncomms6515.
71. **Mendes de Almeida R, Tavares J, Martins S, Carvalho T, Enguita FJ, Brito D, Carmo-Fonseca M, Lopes LR.** 2017. Whole gene sequencing identifies deep-intronic variants with potential functional impact in patients with hypertrophic cardiomyopathy. PLoS One **12:**1–19. https://doi.org/10.1371/journal.pone.0182946.
72. **Messer AE, Chan J, Daley A, Copeland O, Marston SB, Connolly DJ.** 2017. Investigations into the sarcomeric protein and Ca^{2+}-regulation abnormalities underlying hypertrophic cardiomyopathy in cats (*Felix catus*). Front Physiol **8:**1–18. https://doi.org/10.3389/fphys.2017.00348.
73. **Meurs KM, Norgard MM, Ederer MM, Hendrix KP, Kittleson MD.** 2007. A substitution mutation in the myosin binding protein C gene in ragdoll hypertrophic cardiomyopathy. Genomics **90:**261–264. https://doi.org/10.1016/j.ygeno.2007.04.007.
74. **Meurs KM, Sanchez X, David RM, Bowles NE, Towbin JA, Reiser PJ, Kittleson JA, Munro MJ, Dryburgh K, Macdonald KA, Kittleson MD.** 2005. A cardiac myosin binding protein C mutation in the Maine Coon cat with familial hypertrophic cardiomyopathy. Hum Mol Genet **14:**3587–3593. https://doi.org/10.1093/hmg/ddi386.

75. **Moolman-Smook JC, De Lange WJ, Bruwer EC, Brink PA, Corfield VA.** 1999. The origins of hypertrophic cardiomyopathy-causing mutations in two South African subpopulations: a unique profile of both independent and founder events. Am J Hum Genet **65:**1308–1320. https://doi.org/10.1086/302623.
76. **Murphy HW, Dennis P, Devlin W, Meehan T, Kutinsky I.** 2011. Echocardiographic parameters of captive western lowland gorillas (Gorilla gorilla gorilla). J Zoo Wildl Med **42:**572–579. https://doi.org/10.1638/2010-0139.1.
77. **Nadvi NA, Michie KA, Kwan AH, Guss JM, Trewhella J.** 2016. Clinically linked mutations in the central domains of cardiac myosin-binding protein c with distinct phenotypes show differential structural effects. Structure **24:**105–115. https://doi.org/10.1016/j.str.2015.11.001.
78. **Ng J, Fass JN, Durbin-Johnson B, Smith DG, Kanthaswamy S.** 2015. Identifying rhesus macaque gene orthologs using heterospecific human CNV probes. Genom Data **6:**202–207. https://doi.org/10.1016/j.gdata.2015.09.016.
79. **Ng J, Trask JS, Smith DG, Kanthaswamy S.** 2015. Heterospecific SNP diversity in humans and rhesus macaque (*Macaca mulatta*). J Med Primatol **44:**194–201. https://doi.org/10.1111/jmp.12174.
80. **Page SP, Kounas S, Syrris P, Christiansen M, Frank-Hansen R, Andersen PS, Elliott PM, McKenna WJ.** 2012. Cardiac myosin binding protein-C mutations in families with hypertrophic cardiomyopathy: disease expression in relation to age, gender, and long term outcome. Circ Cardiovasc Genet **5:**156–166. https://doi.org/10.1161/CIRCGENETICS.111.960831.
81. **Premasuthan A, Kanthaswamy S, Satkoski J, Smith DG.** 2011. A simple multiplex polymerase chain reaction to determine ABO blood types of rhesus macaques (*Macaca mulatta*). Tissue Antigens **77:**584–588. https://doi.org/10.1111/j.1399-0039.2010.01602.x.
82. **Premasuthan A, Ng J, Kanthaswamy S, Trask JS, Houghton P, Farkas T, Sestak K, Smith DG.** 2012. Molecular ABO phenotyping in cynomolgus macaques using real-time quantitative PCR. Tissue Antigens **80:**363–367. https://doi.org/10.1111/j.1399-0039.2012.01935.x.
83. **Reader JR, Canfield DR, Lane JF, Kanthaswamy S, Ardeshir A, Allen AM, Tarara RP.** 2016. Left ventricular hypertrophy in rhesus macaques (*Macaca mulatta*) at the California national primate research center (1992–2014). Comp Med **66:**162–169.
84. **Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, Siepel A, Sikela JM, Attaway T, Bell S, Bernard KE, Buhay CJ, Chandrabose MN, Dao M, Davis C, Delehaunty KD, Ding Y, Dinh HH, Dugan-Rocha S, Fulton LA, Gabisi RA, Garner TT, Godfrey J, Hawes AC, Hernandez J, Hines S, Holder M, Hume J, Jhangiani SN, Joshi V, Khan ZM, Kirkness EF, Cree A, Fowler RG, Lee S, Lewis LR, Li Z, Liu YS, Moore SM, Muzny D, Nazareth LV, Ngo DN, Okwuonu GO, Pai G, Parker D, Paul HA, Pfannkoch C, Pohl CS, Rogers YH, Ruiz SJ, Sabo A, Santibanez J, Schneider BW, Smith SM, Sodergren E, Svatek AF, Utterback TR, Vattathil S, Warren W, White CS, Chinwalla AT, Feng Y, Halpern AL, Hillier LW, Huang X, Minx P, Nelson JO, Pepin KH, Qin X, Sutton GG, Venter E, Walenz BP, Wallis JW, Worley KC, Yang SP, Jones SM, Marra MA, Rocchi M, Schein JE, Baertsch R, Clarke L, Csuros M, Glasscock J, Harris RA, Havlak P, Jackson AR, Jiang H, Liu Y, Messina DN, Shen Y, Song HX, Wylie T, Zhang L, Birney E, Han K, Konkel MK, Lee J, Smit AF, Ullmer B, Wang H, Xing J, Burhans R, Cheng Z, Karro JE, Ma J, Raney B, She X, Cox MJ, Demuth JP, Dumas LJ, Han SG, Hopkins J, Karimpour-Fard A, Kim YH, Pollack JR, Vinar T, Addo-Quaye C, Degenhardt J, Denby A, Hubisz MJ, Indap A, Kosiol C, Lahn BT, Lawson HA, Marklein A, Nielsen R, Vallender EJ, Clark AG, Ferguson B, Hernandez RD, Hirani K, Kehrer-Sawatzki H, Kolb J, Patil S, Pu LL, Ren Y, Smith DG, Wheeler DA, Schenck I, Ball EV, Chen R, Cooper DN, Giardine B, Hsu F, Kent WJ, Lesk A, Nelson DL, O'Brien W E, Prufer K, Stenson PD, Wallace JC, Ke H, Liu XM, Wang P, Xiang AP, Yang F, Barber GP, Haussler D, Karolchik D, Kern AD, Kuhn RM, Smith KE, Zwieg AS.** 2007. Evolutionary and biomedical insights from the rhesus macaque genome. Science **316:**222–234.
85. **Roma-Rodrigues C, Fernandes AR.** 2014. Genetics of hypertrophic cardiomyopathy: advances and pitfalls in molecular diagnosis and therapy. Appl Clin Genet **7:**195–208.
86. **Salman OF, El-Rayess HM, Abi Khalil C, Nemer G, Refaat MM.** 2018. Inherited cardiomyopathies and the role of mutations in non-coding regions of the genome. Front Cardiovasc Med **5:**1–12. https://doi.org/10.3389/fcvm.2018.00077.
87. **Sandbacka M, Laivuori H, Freitas E, Halttunen M, Jokimaa V, Morin-Papunen L, Rosenberg C, Aittomaki K.** 2013. *TBX6, LHX1* and copy number variations in the complex genetics of Müllerian aplasia. Orphanet J Rare Dis **8:**125. https://doi.org/10.1186/1750-1172-8-125.
88. **Singer ES, Ingles J, Semsarian C, Bagnall RD.** 2019. Key value of RNA analysis of *MYBPC3* splice-site variants in hypertrophic cardiomyopathy. Circ Genom Precis Med **12:**e002368. https://doi.org/10.1161/CIRCGEN.118.002368.
89. **Soliman NA, Elmonem MA, van den Heuvel L, Abdel Hamid RH, Gamal M, Bongaers I, Marie S, Levtchenko E.** 2014. Mutational spectrum of the CTNS gene in Egyptian patients with nephropathic cystinosis. JIMD Rep **14:**87–97. https://doi.org/10.1007/8904_2013_288.
90. **Strande JL.** 2015. Haploinsufficiency *MYBPC3* mutations: another stress induced cardiomyopathy? Let's take a look! J Mol Cell Cardiol **79:**284–286. https://doi.org/10.1016/j.yjmcc.2014.12.008.
91. **Tanjore RR, Rangaraju A, Kerkar PG, Calambur N, Nallari P.** 2008. *MYBPC3* gene variations in hypertrophic cardiomyopathy patients in India. Can J Cardiol **24:**127–130. https://doi.org/10.1016/S0828-282X(08)70568-3.
92. **Tian C, Gregersen PK, Seldin MF.** 2008. Accounting for ancestry: population substructure and genome-wide association studies. Hum Mol Genet **17 R2:**R143–R150. https://doi.org/10.1093/hmg/ddn268.
93. **Tower-Rader A, Desai MY.** 2017. Phenotype-genotype correlation in hypertrophic cardiomyopathy: less signal, more noise? Circ Cardiovasc Imaging **10:**e006066. https://doi.org/10.1161/CIRCIMAGING.117.006066.
94. **Ueda Y, Slabaugh TL, Walker AL, Ontiveros ES, Sosa PM, Reader R, Roberts JA, Stern JA.** 2019. Heart rate and heart rate variability of rhesus macaques (*Macaca mulatta*) affected by left ventricular hypertrophy. Front Vet Sci **6:**1–11. https://doi.org/10.3389/fvets.2019.00001.
95. **Uribe-Convers S, Settles ML, Tank DC.** 2016. A phylogenomic approach based on pcr target enrichment and high throughput sequencing: resolving the diversity within the South American species of *Bartsia* L. (Orobanchaceae). PLoS One **11:**1–28. https://doi.org/10.1371/journal.pone.0148203.
96. **van Dijk SJ, Dooijes D, dos Remedios C, Michels M, Lamers JM, Winegrad S, Schlossarek S, Carrier L, ten Cate FJ, Stienen GJ, van der Velden J.** 2009. Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. Circulation **119:**1473–1483. https://doi.org/10.1161/CIRCULATIONAHA.108.838672.
97. **Vaz-Drago R, Custodio N, Carmo-Fonseca M.** 2017. Deep intronic mutations and human disease. Hum Genet **136:**1093–1111. https://doi.org/10.1007/s00439-017-1809-4.
98. **Vaz-Drago R, Pinheiro MT, Martins S, Enguita FJ, Carmo-Fonseca M, Custodio N.** 2015. Transcription-coupled RNA surveillance in human genetic diseases caused by splice site mutations. Hum Mol Genet **24:**2784–2795. https://doi.org/10.1093/hmg/ddv039.
99. **Vignier N, Schlossarek S, Fraysse B, Mearini G, Kramer E, Pointu H, Mougenot N, Guiard J, Reimer R, Hohenberg H, Schwartz K, Vernet M, Eschenhagen T, Carrier L.** 2009. Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice. Circ Res **105:**239–248. https://doi.org/10.1161/CIRCRESAHA.109.201251.
100. **Vikstrom KL, Factor SM, Leinwand LA.** 1996. Mice expressing mutant myosin heavy chains are a model for familial hypertrophic

cardiomyopathy. Mol Med **2:**556–567. https://doi.org/10.1007/BF03401640.

101. **Walsh R, Buchan R, Wilk A, John S, Felkin LE, Thomson KL, Chiaw TH, Loong CCW, Pua CJ, Raphael C, Prasad S, Barton PJ, Funke B, Watkins H, Ware JS, Cook SA.** 2017. Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes. Eur Heart J **38:**3461–3468.
102. **Ward AJ, Cooper TA.** 2010. The pathobiology of splicing. J Pathol **220:**152–163.
103. **Xiong HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, Yuen RK, Hua Y, Gueroussov S, Najafabadi HS, Hughes TR, Morris Q, Barash Y, Krainer AR, Jojic N, Scherer SW, Blencowe BJ, Frey BJ.** 2015. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. Science **347:**1254806. https://doi.org/10.1126/science.1254806.
104. **Xue C, Raveendran M, Harris RA, Fawcett GL, Liu X, White S, Dahdouli M, Rio Deiros D, Below JE, Salerno W, Cox L, Fan G, Ferguson B, Horvath J, Johnson Z, Kanthaswamy S, Kubisch HM, Liu D, Platt M, Smith DG, Sun B, Vallender EJ, Wang F, Wiseman RW, Chen R, Muzny DM, Gibbs RA, Yu F, Rogers J.** 2016. The population genomics of rhesus macaques (*Macaca mulatta*) based on whole-genome sequences. Genome Res **26:**1651–1662. https://doi.org/10.1101/gr.204255.116.
105. **Zhang X, Meng Y, Houghton P, Liu M, Kanthaswamy S, Oldt R, Ng J, Trask JS, Huang R, Singh B, Du H, Smith DG.** 2017. Ancestry, Plasmodium Cynomolgi prevalence and rhesus macaque admixture in cynomolgus macaques (*Macaca fascicularis*) bred for export in Chinese breeding farms. J Med Primatol **46:**31–41. https://doi.org/10.1111/jmp.12256.