



# Pathogenic intronic splice-affecting variants in MYBPC3 in three patients with hypertrophic cardiomyopathy

DOI:

[10.3390/cardiogenetics11020009](https://doi.org/10.3390/cardiogenetics11020009)

## Document Version

Final published version

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Newman, W., Wood, K., Thomas, H., Ellingford, J., O'Keefe, R., Hopton, C., & Eden, J. (2021). Pathogenic intronic splice-affecting variants in MYBPC3 in three patients with hypertrophic cardiomyopathy. *Cardiogenetics*, 11(2), 73-83. <https://doi.org/10.3390/cardiogenetics11020009>

## Published in:

Cardiogenetics

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.





# Pathogenic intronic splice-affecting variants in *MYBPC3* in three patients with hypertrophic cardiomyopathy

Katherine A. Wood<sup>1,2</sup>, Jamie M. Ellingford<sup>1,2</sup>, James Eden<sup>2</sup>, Huw B. Thomas<sup>1</sup>, Raymond T. O'Keefe<sup>1</sup>, Claire Hopton<sup>1,2</sup>, William G. Newman<sup>1,2\*</sup>

**Citation:** Lastname, F.; Lastname, F.; Lastname, F. Title. *Cardiogenetics* 2021, 11, Firstpage–Lastpage. <https://doi.org/10.3390/xxxxx>

Academic Editor: Firstname Lastname

Received: date

Accepted: date

Published: date

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

- <sup>1</sup> Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK; katherine.wood@manchester.ac.uk (KAW); jamie.ellingford@manchester.ac.uk (JME); huw.thomas@manchester.ac.uk (HBT); rokeefe@manchester.ac.uk (RTO); claire.hopton@manchester.ac.uk (CH); william.newman@manchester.ac.uk (WGN)
- <sup>2</sup> Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK; james.eden@mft.nhs.uk (JE)
- \* Correspondence: william.newman@manchester.ac.uk

**Abstract:** Genetic variants in *MYBPC3* are one of the most common causes of hypertrophic cardiomyopathy (HCM). While variants in *MYBPC3* affecting canonical splice site dinucleotides are a well-characterised cause of HCM, only recently has work begun to investigate the pathogenicity of more deeply intronic variants. Here, we present three patients with HCM and intronic splice-affecting *MYBPC3* variants and analyse the impact of variants on splicing using *in vitro* minigene assays. We show that the three variants, a novel c.927-8G>A variant and the previously reported c.1624+4A>T and c.3815-10T>G variants, result in *MYBPC3* splicing errors. Analysis of blood-derived patient RNA for the c.3815-10T>G variant revealed only wild type spliced product, indicating that mis-spliced transcripts from the mutant allele are degraded. These data indicate that the c.927-8G>A variant of uncertain significance and likely benign c.3815-10T>G should be reclassified as likely pathogenic. Furthermore, we find shortcomings in commonly applied bioinformatics strategies to prioritize variants impacting *MYBPC3* splicing and re-emphasize the need for functional assessment of variants of uncertain significance in diagnostic testing.

---

**Keywords:** Hypertrophic cardiomyopathy; MYBPC3; splice variants; minigene assays

---

## 1. Introduction

Hypertrophic cardiomyopathy (HCM) is a relatively common genetic disorder (with a prevalence of 1:500) characterised by left ventricular cardiomyopathy, a non-dilated left ventricle and normal, or increased, ejection fraction, and is associated with myocardial fibre disarray [1–3]. HCM is usually asymmetrical and develops in the absence of an identifiable secondary cause such as hypertension or aortic valvular stenosis. Patients with HCM usually have a relatively benign clinical course, although HCM can cause sudden cardiac death particularly in adolescents and young adults [3]. Risk factors such as severe cardiac hypertrophy, family history of sudden cardiac death, syncope and non-sustained ventricular tachycardia have been associated with sudden cardiac death in HCM patients, and high-risk patients may be offered implantable cardioverter-defibrillators (ICDs) to ameliorate their risk [2–7].

HCM usually shows an autosomal dominant mode of inheritance, although very rare cases of autosomal recessive and X-linked modes of inheritance have been described [3,8–10]. Incomplete penetrance and variable clinical expression are common confounding phenomena. The phenotypic variability is thought to arise, at least in part, from interactions between pathogenic variants and genetic and non-genetic modifiers [3,11–13]. Variants in at least 12 genes encoding sarcomere and sarcomere-associated proteins cause HCM, the two most common of which (accounting for approximately 50% of families with HCM) are  $\beta$ -myosin heavy chain (*MYH7*) and myosin-binding protein C3 (*MYBPC3*) [2,3,14–19]. The  $\beta$ -myosin heavy chain has the ATPase enzymatic activity in the myosin head required for force generation in muscle fibres [20]. The myosin-binding protein C3 interacts with both actin and myosin, as well as titin, and regulates cardiac contraction [21,22]. While the majority of variants causing HCM in *MYH7* and other, less common causative genes are missense variants, pathogenic variants in *MYBPC3* tend to be frameshift, nonsense or splice site variants, which result in premature termination codons (PTCs) and predicted loss of function [3,14,16,23]. PTC-containing *MYBPC3* transcripts may result in truncated MYBPC3 protein upon translation; alternatively, mutant transcripts may be degraded by the nonsense-mediated decay pathway, causing allelic loss-of-function and reduced expression [23,24]. Approximately 15% of *MYBPC3* variants in HCM are non-truncating, including missense and short in-frame insertion/deletion variants. The pathogenic mechanisms of these non-truncating variants are largely unknown, although it has been shown that disease severity and clinical outcome, while highly variable, is largely independent of whether the pathogenic *MYBPC3* variant is truncating or non-truncating [23,25,26].

While genetic testing is now common in patients with HCM, the diagnostic yield is only 40–50% and the causative genotype is unknown for at least 40% of patients, even for those with family histories of the disease [27,28]. There are many factors which may explain this low genetic testing yield. For example, phenocopying occurs in other syndromic conditions such as Noonan syndrome and storage disorders, including Anderson-Fabry disease, while in other cases complex and as-yet unknown genetic mechanisms may cause HCM [27,29–32]. However, conventional genetic testing methods may also fail to identify pathogenic non-coding variants, in particular cryptic splice variants (i.e. intronic variants out-with the canonical splice acceptor and splice donor dinucleotides which alter pre-mRNA splicing) in known HCM genes, which may be a major contributor to the missing heritability in HCM. For example, several recent publications have identified novel intronic cryptic splice variants within *MYBPC3* in cohorts of HCM patients, which would not have been detected as pathogenic variants by standard genetic testing [27,33–37]. These intronic variants were shown to alter *MYBPC3* splicing and result in allelic loss-of-function and haploinsufficiency.

Here, we present three patients with HCM presenting with intronic *MYBPC3* splice variants out-with the invariant splice donor/acceptor dinucleotides. We use *in vitro* minigene splicing assays in human cell lines, combined with direct analysis of patient RNA in one case, to confirm the aberrant effect of these variants on *MYBPC3* splicing. This study supports the importance of sequencing intronic regions in *MYBPC3* to increase the detection of pathogenic variants causing HCM

## 2. Materials and Methods

### 2.1. Patient recruitment and genetic testing

The three probands were recruited by the Manchester Centre for Genomic Medicine. For proband 1 clinical DNA sequencing by NGS was undertaken of the entire coding sequence of four genes – *MYH7*, *MYBPC3*, *TNNT2* and *TNNI3*. For probands 2 and 3 a panel of >20 genes associated with HCM was screened by NGS as previously described [38].

### 2.2. Bioinformatics analysis

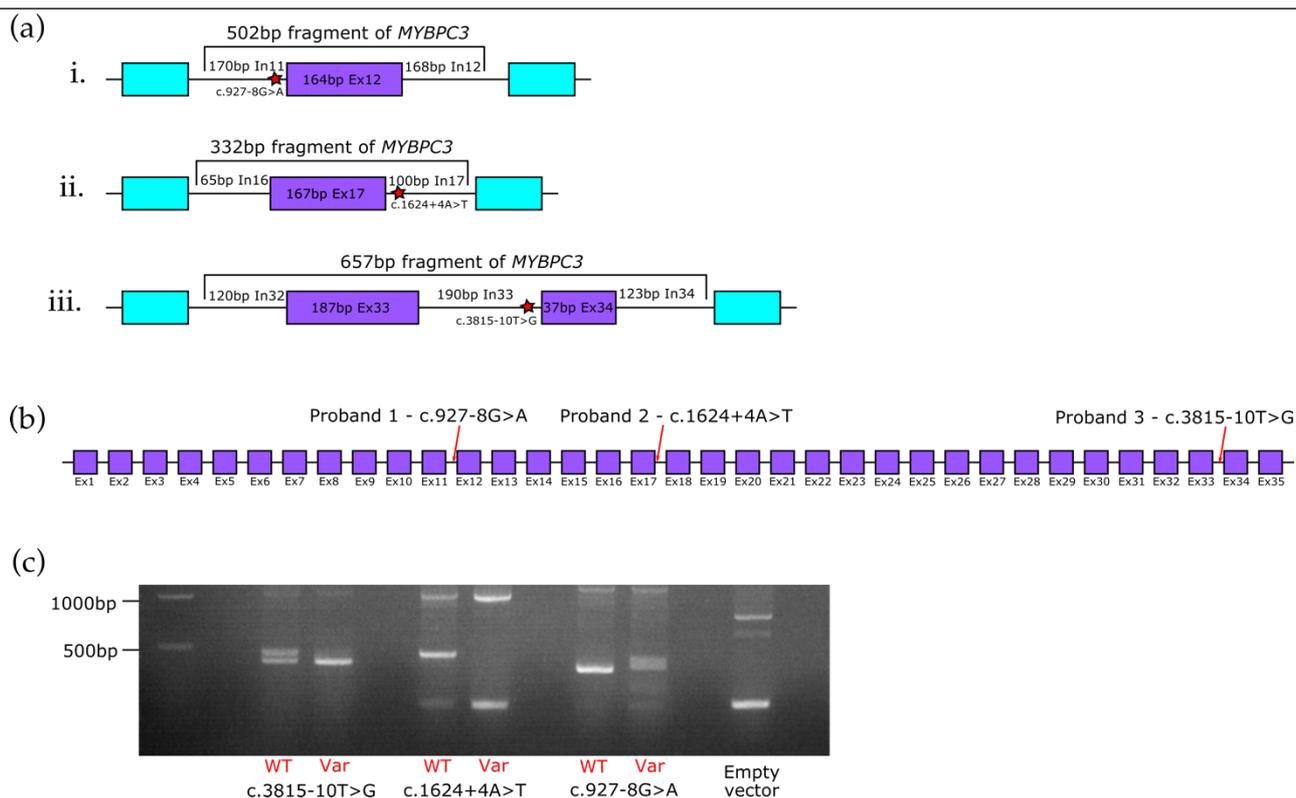
*In silico* pathogenicity prediction with SpliceAI v1.3.1 and Human Splicing Finder v3.1 were used to score cryptic splice-affecting variants in *MYBPC3* [39,40]. Population frequency of *MYBPC3* variants was analysed using gnomAD v.3.1 (<https://gnomad.broadinstitute.org/>) [41].

### 2.3. Human K562 cell culture

Human K562 cells were cultured under standard tissue culture conditions in RPMI-1640 medium (Sigma) supplemented with 10% Fetal Bovine Serum (FBS) (Sigma) and incubated in 5% CO<sub>2</sub> at 37°C. Original K562 cell stocks were supplied at passage 10.

### 2.4. *In vitro* splicing minigene assay

Fragments of *MYBPC3* (reference transcript NM\_000256) were amplified by polymerase chain reaction (PCR) from commercially available reference genomic DNA (Promega) using either wildtype or mutagenic primers (**Table S1**) using Phusion High Fidelity DNA Polymerase (NEB) according to manufacturer's recommendations. For the c.927-8G>A variant, the total *MYBPC3* fragment length was 502bp, containing the 164bp exon 12 plus 170bp of intron 11 and 168bp of intron 12 of *MYBPC3* (**FIGURE 1a(i)**). For the c.1624+4A>T variant, the total *MYBPC3* fragment length was 332bp, containing the 167bp exon 17 plus 65bp of intron 16 and 100bp of intron 17 (**FIGURE 1a(ii)**). For the c.3815-10T>G variant, the *MYBPC3* fragment was 657bp and contained the 187bp exon 33 and 37bp exon 34, plus 120bp of intron 32, the 190bp intron 33, and 123bp of intron 34 (**FIGURE 1a(iii)**). The wildtype and mutagenic *MYBPC3* fragments were cloned into the SK3 plasmid (a derivative of the pSpliceExpress minigene splice reporter vector, gifted from Stefan Stamm, Addgene #32485) using the Gibson method [42,43]. Constructs were transformed into competent bacteria and candidate colonies were cultured and vector DNA isolated using the GenElute Plasmid Miniprep kit (Sigma). Sequences of the minigene vector constructs were verified by Sanger sequencing (performed by Eurofins genomics).



**Figure 1. *MYBPC3* intronic splice variants in three probands with hypertrophic cardiomyopathy.** a) Schematic of the minigene constructs used to test pathogenicity of *MYBPC3* splice variants i. c.927-8G>A, ii. c.1624+4A>T and iii. c.3815-10T>G. *MYBPC3* exons are shown in purple. Endogenous minigene exons are shown in light blue. The location of the splice site variants are indicated by red stars. Ex = exon, In = intron. b) Location of *MYBPC3* splice variants analysed in this paper and the probands in which the variants were identified, using the reference transcript NM\_000256. Purple boxes refer to exons (Ex), black lines refer to introns. c) Representative RT-PCR results showing the outcome of the minigene assays. Identities of RT-PCR products were confirmed by Sanger sequencing. High molecular weight bands above 500bp correspond to unspliced vector DNA contamination. WT = wild type, Var = variant. n = 2.

Human K562 cells were plated in 2ml RPMI-1640 medium (Sigma) supplemented with 10% Fetal Bovine Serum (FBS) (Sigma) at approximately 80% confluency in 6-well tissue culture plates (Corning) on the day of transfection. Cells were transiently transfected with 2.5µg plasmid DNA using Lipofectamine LTX (Invitrogen) following the manufacturer's protocol and incubated for 18 – 24 hours at 37°C with 5% CO<sub>2</sub>. The splicing minigene assays were performed twice independently.

### 2.5. Reverse-transcription PCR (RT-PCR)

For the splicing minigene assay, total RNA was extracted from K562 cells using TRIzol reagent (Invitrogen) according to manufacturer's instructions. The resulting RNA was cleaned up using the RNeasy Mini RNA Isolation Kit (Qiagen), including an on-column DNA digestion step. An equal amount of RNA for each sample was converted to cDNA using Superscript IV (Invitrogen) with random hexamers (Invitrogen) according to the supplier's recommendations. cDNA was amplified using Phusion polymerase and primers listed in **Table S2**. PCR products were separated by agarose gel electrophoresis supplemented with SafeView nucleic acid stain (NBS Biologicals) and visualised under a blue-light transilluminator. Products were extracted using QIAquick gel extraction kit (Qiagen) and sequences by Sanger sequencing to confirm their identity.

For proband 3 with the *MYBPC3* c.3815-10T>G variant, an RNA sample (Paxgene) was derived from blood following manufacturer's instructions. A cDNA synthesis was performed as above using patient RNA as a template and the cDNA amplified using

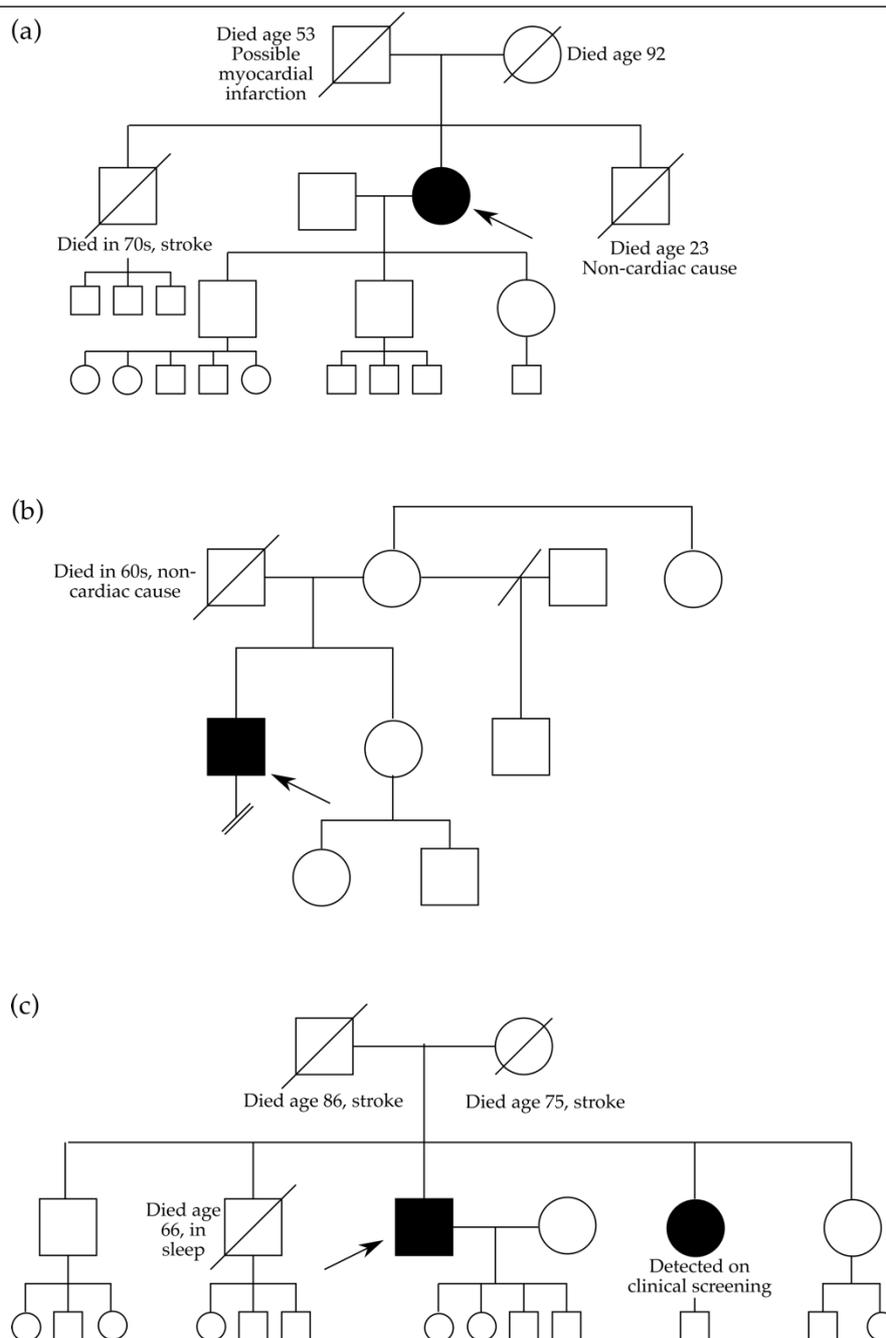
Phusion polymerase and primers listed in **Table S2**. PCR products were separated, gel extracted and sequenced as described above.

### 3. Results

Three patients with a clinical diagnosis of HCM, but no known genetic cause underwent genetic and molecular screening by next-generation sequencing (NGS) to establish the genetic basis for their specific cardiac phenotype.

#### 3.1. Clinical characteristics associated with three intronic MYBPC3 variants

The proband in family 1 (**FIGURE 2a**) was a 77 year old woman with a history of persistent atrial fibrillation. An echocardiogram revealed severe asymmetrical left ventricular hypertrophy with a septum thickness of 2cm. The proband's father was reported to have died suddenly at the age of 53, however there was no other family history consistent with a diagnosis of cardiomyopathy. Genetic testing revealed a heterozygous c.927-8G>A variant in intron 11 of *MYBPC3* (**FIGURE 1b**). The affected nucleotide is 8 base pairs (bp) upstream of the canonical 3' splice site in intron 11, and potentially creates a new cryptic 3' splice site AG. This c.927-8G>A variant has not been described previously in HCM patients in the literature, is not present in the gnomAD population genomic sequence database and is classified as a variant of uncertain significance. However, this variant has been identified in seven additional unrelated individuals with HCM in diagnostic laboratories performing cardiomyopathy testing in the United Kingdom. Additionally, a variant in the adjacent residue (c.927-9G>A) has been reported as likely pathogenic, although analysis of patient RNA derived from blood for this adjacent variant did not confirm any splicing errors caused by the variant [37,44]. The c.927-8G>A variant was not predicted to have any significant impact on splicing signals by Human Splicing Finder. In agreement with Human Splicing Finder, the SpliceAI score for this variant was 0.11 (acceptor loss). Commonly applied SpliceAI scores of 0.2 as potentially splice-altering and 0.5 as likely splice-altering for variant prioritisation would (falsely) exclude this variant from prioritisation [39,45].



**Figure 2. Family pedigrees for the three probands with HCM presented in this study.** Pedigrees for a) family 1, b) family 2, and c) family 3. Arrow indicates the probands subject to genetic testing in this study. Filled shapes indicate clinically affected individuals.

The proband in family 2 (FIGURE 2b) was a 46 year old man who presented with palpitations and shortness of breath on exertion. A cardiac MRI scan showed moderate to severe concentric left ventricular hypertrophy with septal predominance (basal antero-septum measured 2.4cm). There was also evidence of mid layer fibrosis involving the hypertrophied regions. The proband had type 1 diabetes mellitus but no history of hypertension or other significant medical issues. There was no known family history of cardiomyopathy or sudden cardiac death. The proband was identified as having a heterozygous c.1624+4A>T variant in intron 17 of *MYBPC3* (FIGURE 1b). The affected nucleotide is 4bp downstream of the 5' splice site, but still within the splice donor consensus

sequence. This c.1624+4A>T variant is present in three of 224880 alleles in gnomAD and is predicted by Human Splicing Finder to affect splicing by abrogating the splice donor site at the 5' end of intron 17, although the SpliceAI score for the variant is only 0.16 (donor loss) again suggesting no impact on splicing. The c.1624+4A>T variant is classified as pathogenic, having been reported in a number of HCM patients, both within and outside of the UK, and previous blood-derived RNA analysis from these patients revealed the variant causes exon 17 skipping, which would result in a PTC and loss of the C-terminal region of the protein upon translation removing the major myosin and titin binding sites [24,37,46–49].

The proband in family 3 (FIGURE 2c) was diagnosed with HCM after presenting with chest pain whilst walking in his fifth decade. He had a history of paroxysmal atrial fibrillation and had a stroke at the age of 58 years. A cardiac MRI scan performed shortly after his stroke revealed a maximal wall thickness of 2.6cm and left ventricular outflow tract obstruction, consistent with HCM. On gadolinium imaging, there was evidence of very extensive enhancement involving most of the left ventricle. One of the proband's brothers was reported to have died in his sleep at the age of 66 years and one of the proband's sisters was identified as having mild HCM on screening shortly after his diagnosis. Genetic testing in the proband revealed a heterozygous c.3815-10T>G variant in intron 33 of *MYBPC3* (FIGURE 1b). The variant is 10 nucleotides upstream of the 3' splice site. This c.3815-10T>G variant has been reported once in 242,620 alleles in gnomAD and has been reported previously in one HCM patient in ClinVar, although is currently classified as a likely benign variant [33]. *In silico* pathogenicity prediction by Human Splicing Finder suggested no impact of this variant on splicing signals, while the variant had a SpliceAI score of 0.37 (acceptor loss) indicating possible impact on splicing. Genetic testing also revealed a heterozygous *MYL2* c.37G>A, p.Ala13Thr variant in this proband.

### 3.2. *In vitro* minigene assays to assess the impact of intronic *MYBPC3* splice-affecting variants

To test the effects of the identified *MYBPC3* intronic variants on *MYBPC3* pre-mRNA splicing, we conducted *in vitro* minigene splicing assays in human cell lines. The assays compared spliced RNA transcripts extracted from human K562 cells following transfection with pairs of wild type and variant-containing minigene constructs containing fragments of *MYBPC3* (FIGURE 1a).

For the c.927-8G>A variant, we found that while the wild type sequence led to canonical splicing of *MYBPC3* only, the variant resulted in loss of the wild type splice acceptor and the use of an upstream cryptic splice acceptor site instead, resulting in the inclusion of an additional 19bp at the 5' end of exon 12 (FIGURE 1c). This 19bp exon extension would result in a frameshift and the creation of a PTC in exon 13.

For the previously characterised c.1624+4A>T variant, the wild type sequence resulted in a primary product containing canonically spliced exon 17, and a small fraction of product with exon 17 skipping. However, the mutant sequence resulted in complete exon 17 skipping (due to mutation of the wild type splice donor consensus sequence) and no canonically spliced product, in agreement with previous reports (FIGURE 1c, FIGURE S1) [33,37].

Finally, for the c.3815-10T>G variant, the wild type sequence resulted in two spliced products, one containing exon 34 and one with exon 34 skipped. However, for the mutant construct, there was complete skipping of exon 34 and no exon 34 retention, suggesting the variant led to complete loss of splice acceptor recognition at the 3' end of intron 33 (FIGURE 1c). Interestingly, exon 34 (of 35 total exons for the NM\_000256 reference transcript used in this study) is the final coding exon of *MYBPC3*, and so the skipping event observed here would result in loss of the normal stop codon. RNA from blood was available for this patient, and direct analysis of the spliced products in blood-

derived patient RNA revealed only a wild type product, suggesting that mutant transcripts produced from the variant-containing allele were degraded (data not shown). This finding suggests the c.3815-10T>G variant results in decreased *MYBPC3* expression, in agreement to the known disease mechanism of haploinsufficiency.

#### 4. Discussion

In this study, we present three patients with *MYBPC3* intronic variants and assess the effects of these variants on *MYBPC3* splicing in minigene assays *in vitro*. Confirming the pathogenicity of these intronic splice variants makes it possible to utilise the findings clinically and offer targeted genetic testing to at risk family members. For the c.927-8G>A and c.3815-10T>G variants, we can now apply PS3 from the American College of Medical Genetics (ACMG) guidelines for assigning pathogenicity to sequence variants as we have shown, using a well-established *in vitro* functional study, the damaging effect on splicing of these two variants [50].

While the *MYBPC3* c.1624+4A>T variant has been well-characterised and previously classified as pathogenic, we included it in our minigene assays as a positive control to validate our methodology and confirm the sensitivity of our *in vitro* assay for detecting splice-altering variant effects [24,37,46–49]. In our system, this variant resulted in complete *MYBPC3* exon 17 skipping, in agreement with these previous analyses.

The novel c.927-8G>A variant, described for the first time in proband 1 of our study, had a clear impact on *MYBPC3* splicing, activates a cryptic splice site upstream of the canonical 3' splice site at intron 11, resulting in the extension of exon 12 by 19bp at the 5' end, leading to a frameshift and a PTC. It is interesting that the report of a variant in the adjacent nucleotide (c.927-9G>A) could not confirm aberrant splicing in patient RNA samples by RT-PCR; the authors only observed the wild type spliced product [37]. We postulate that the adjacent variant may have the same effect on *MYBPC3* splicing as our c.927-8G>A variant as both variants affect the same canonical splice site in intron 11 and likely activate the same cryptic upstream 3' splice site. It is likely that transcripts from the mutant allele are degraded by nonsense mediated decay (NMD), such that they are not observed by RT-PCR of patient RNA. Quantitative analysis of the *MYBPC3* expression by qPCR and/or heterozygous variant ratios to assess whether transcripts from the mutant allele are present in patient RNA samples could be used to confirm this mechanism, as well as assessing the variant in a minigene assay.

Finally, the previously described c.3815-10T>G variant in proband 3 of our study had a clear effect on splicing in our minigene assay, resulting in total loss of the splice donor recognition at the 3' end of intron 33 leading to complete exon 34 skipping and loss of the normal stop codon. The nucleotide sequence of exon 35 codes for 46 amino acids in-frame with exon 33 with an in-frame stop codon and so the mis-splicing event associated with the variant could result in a change to the 3' end of the encoded protein. This event would not be detected from our minigene assay because the construct only contained *MYBPC3* exons 33 and 34, and not exon 35. However, RT-PCR analysis of patient RNA indicated that this event causes NMD of transcripts from the mutant allele, as opposed to production of a protein with an altered C-terminal. Two points of interest are worth noting here. Firstly, even for our wild type minigene construct, there is a relatively strong RT-PCR product (approximately the same intensity of band on the gel as the canonically spliced RT-PCR product) corresponding to the exon 34-skipped splice product and so the variant appears to shift the ratio of splice isoforms towards an exon 34-skipped product which is already normally observed. In contrast, while a very faint band corresponding to the exon 17-skipped splice product is also present for the wild type construct for the c.1624+4A>T variant, this band is much weaker than the canonically spliced product. *MYBPC3* exon 34 is a very short exon (37bp), and previous studies have indicated shorter exon length is associated with more frequently skipped exons, which may explain the relatively high prevalence of the shorter splice isoform even for the wild type minigene constructs [51,52]. The second point of interest with the

c.3815-10T>G variant is that this individual also carries a heterozygous *MYL2* c.37G>A, p.Ala13Thr variant. This amino acid substitution has been shown to have an important effect on calcium binding and subsequent myosin binding of the encoded protein, and has been previously reported in patients with HCM, particularly associated with left ventricular obstruction. This observation is consistent with the phenotype of left ventricular outflow tract obstruction in our patient, but the high minor allele frequency means that the clinical significance of this variant is unclear [53–57]. It may be that because the *MYBPC3* c.3815-10T>G splice variant enhances an exon skipping event, which is normally observed at a relatively high frequency, this variant must be combined with a second variant in another sarcomere-associated gene and/or must occur in the presence of a background of particular genetic modifiers to be fully penetrant. Another non-mutually exclusive possibility is that the common *MYL2* variant in this patient is only pathogenic in the presence of a second variant in a sarcomere-associated gene (in this case the *MYBPC3* c.3815-10T>G variant), explaining the link between the variant and specific cardiac phenotypes even though its population frequency is relatively high. It is not clear whether the previously reported individuals with the *MYBPC3* c.3815-10T>G variant also carried the *MYL2* c.37G>A variant [33]. One of this proband's sisters was identified as having a mild HCM phenotype. Genetic testing will now be extended to the affected sibling.

A further point of interest is that none of the intronic splice-affecting variants presented in this study had particularly high SpliceAI scores and only the c.1624+4A>T variant was predicted by Human Splicing Finder to affect *MYBPC3* splicing. Indeed, for this variant, the SpliceAI score was only 0.16, and yet the variant resulted in total loss of the intron 17 splice donor and complete exon 17 skipping, and is already classified as pathogenic on the basis of previous RNA analysis. Similarly, the c.927-8G>A variant had a SpliceAI score of just 0.11, and yet had a very clear impact on the *MYBPC3* exon 12 splice acceptor site, while the c.3815-10T>G had a SpliceAI score of 0.37 but had a major impact in shifting *MYBPC3* splicing towards exon 34 skipping. In general, our laboratory policy is to consider variants with SpliceAI scores greater >0.2 as potentially splice-altering, while scores >0.5 are considered as very likely to be splice-altering, and is similar to previously described SpliceAI cut-offs used for variant prioritisation [39,45]. However, some reports investigating *MYBPC3* intronic splice variants have used very stringent SpliceAI cut-offs of  $\geq 0.9$  to initially prioritise variants [27]. Nonetheless, when these studies expanded their inclusion criteria to include variants with lower SpliceAI scores, an *MYBPC3* c.1898-23A>G variant with a SpliceAI score of just 0.04 was identified; assessment of patient RNA using RT-PCR showed inclusion of intron 19, leading to a premature stop codon and presumed NMD [27]. In our study we further demonstrate that intronic variants beneath commonly applied SpliceAI filters do have a significant impact on splicing, and these data reemphasise that functional investigation alongside *in silico* assessments remains an important consideration for diagnostic testing in the context of splicing.

## 5. Conclusions

Intronic *MYBPC3* cryptic splice variants are an important genetic cause of HCM and may not be accurately predicted by *in silico* bioinformatic methods. Analysis of intronic *MYBPC3* splice variants increases the diagnostic yield in HCM patients.

**Supplementary Materials:** The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1). Table S1: Primers used for generating *MYBPC3* minigene constructs. Table S2: Primers used for reverse-transcription polymerase chain reaction (RT-PCR) experiments. Figure S1: Sequence traces showing the 19bp insertion at the 5' end of *MYBPC3* exon 12 for the c.927-8G>A variant in the minigene assay.

**Author Contributions:** Conceptualization, WGN; Methodology, HBT, KAW; Software, JME; Validation, KAW, HBT, JME; Formal Analysis, KAW, HBT, JME; Investigation, KAW, JME; Resources, WGN, RTO; Data Curation, KAW, JME, CH; Writing – Original Draft Preparation, KAW, CH; Writing – Review & Editing, HBT, JME, JE, RTO, WGN; Visualization, KAW; Supervision, WGN, RTO; Project Administration, WGN; Funding Acquisition, WGN, RTO, JME.

**Funding:** KAW is funded by a Medical Research Council PhD studentship (1916606). WGN is supported by the Manchester NIHR BRC (IS-BRC-1215-20007). CH is a Clinical Lecturer supported by the NIHR. RTO and HBT are supported by the BBSRC (BB/N000258/1). JME is supported by the Health Education England Genomics Education Programme.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the South Manchester Research Ethics Committee (11/H10003/3).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in this study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to sensitivity of patient information.

**Acknowledgments:** The authors acknowledge the NW Genomic Laboratory Hub for genetic screening by NGS of patient samples.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Towbin JA. Hypertrophic cardiomyopathy. *PACE - Pacing and Clinical Electrophysiology*. 2009. doi:10.1111/j.1540-8159.2009.02381.x
2. Hensley N, Dietrich J, Nyhan D, Mitter N, Yee MS, Brady MB. Hypertrophic cardiomyopathy: A review. *Anesthesia and Analgesia*. 2015. doi:10.1213/ANE.0000000000000538
3. Marian AJ, Braunwald E. Hypertrophic cardiomyopathy: Genetics, pathogenesis, clinical manifestations, diagnosis, and therapy. *Circ Res*. 2017; doi:10.1161/CIRCRESAHA.117.311059
4. Elliott PM, Poloniecki J, Dickie S, Sharma S, Monserrat L, Varnava A, et al. Sudden death in hypertrophic cardiomyopathy: Identification of high risk patients. *J Am Coll Cardiol*. 2000; doi:10.1016/S0735-1097(00)01003-2
5. Frenneaux MP, Counihan PJ, Caforio ALP, Chikamori T, McKenna WJ. Abnormal blood pressure response during exercise in hypertrophic cardiomyopathy. *Circulation*. 1990; doi:10.1161/01.CIR.82.6.1995
6. Spirito P, Bellone P, Harris KM, Bernabò P, Bruzzi P, Maron BJ. Magnitude of Left Ventricular Hypertrophy and Risk of Sudden Death in Hypertrophic Cardiomyopathy. *N Engl J Med*. 2000; doi:10.1056/nejm200006153422403
7. McKenna WJ, Franklin RCG, Nihoyannopoulos P, Robinson KC, Deanfield JE, Dickie S, et al. Arrhythmia and prognosis in infants, children and adolescents with hypertrophic cardiomyopathy. *J Am Coll Cardiol*. 1988; doi:10.1016/0735-1097(88)90181-7
8. Greaves SC, Roche AHG, Neutze JM, Whitlock RML, Veale AMO. Inheritance of hypertrophic cardiomyopathy: A cross sectional and M mode echocardiographic study of 50 families. *Heart*. 1987; doi:10.1136/hrt.58.3.259
9. Branzi A, Romeo G, Specchia S, Lolli C, Binetti G, Devoto M, et al. Genetic heterogeneity of hypertrophic cardiomyopathy. *Int J Cardiol*. 1985; doi:10.1016/0167-5273(85)90352-3
10. Hartmannova H, Kubanek M, Sramko M, Piherova L, Noskova L, Hodanova K, et al. Isolated X-linked hypertrophic cardiomyopathy caused by a novel mutation of the four-and-a-half LIM domain 1 gene. *Circ Cardiovasc Genet*. 2013; doi:10.1161/CIRCGENETICS.113.000245
11. Marian AJ. Modifier genes for hypertrophic cardiomyopathy. *Current Opinion in Cardiology*. 2002. doi:10.1097/00001573-200205000-00006
12. Zhang X, Xie J, Zhu S, Chen Y, Wang L, Xu B. Next-generation sequencing identifies pathogenic and modifier mutations in a consanguineous Chinese family with hypertrophic cardiomyopathy. *Med (United States)*. 2017; doi:10.1097/MD.00000000000007010
13. Chen Y, Xu F, Munkhsaikhan U, Boyle C, Borcky T, Zhao W, et al. Identifying modifier genes for hypertrophic cardiomyopathy. *J Mol Cell Cardiol*. 2020;144: 119–126. doi:10.1016/j.yjmcc.2020.05.006
14. Millat G, Bouvagnet P, Chevalier P, Dauphin C, Simon Jouk P, Da Costa A, et al. Prevalence and spectrum of mutations in a cohort of 192 unrelated patients with hypertrophic cardiomyopathy. *Eur J Med Genet*. 2010; doi:10.1016/j.ejmg.2010.07.007
15. Kaski JP, Syrris P, Esteban MTT, Jenkins S, Pantazis A, Deanfield JE, et al. Prevalence of sarcomere protein gene mutations in preadolescent children with hypertrophic cardiomyopathy. *Circ Cardiovasc Genet*. 2009; doi:10.1161/CIRCGENETICS.108.821314
16. Erdmann J, Daehmlow S, Wischke S, Senyuva M, Werner U, Raible J, et al. Mutation spectrum in a large cohort of unrelated consecutive patients with hypertrophic cardiomyopathy. *Clin Genet*. 2003; doi:10.1034/j.1399-0004.2003.00151.x
17. Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, et al. Hypertrophic cardiomyopathy: Distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation*. 2003; doi:10.1161/01.CIR.0000066323.15244.54
18. Wigle ED, Rakowski H, Kimball BP, Williams WG. Hypertrophic cardiomyopathy: Clinical spectrum and treatment. *Circulation*. 1995. doi:10.1161/01.CIR.92.7.1680

19. Maron BJ, Maron MS. Hypertrophic cardiomyopathy. *The Lancet*. 2013. doi:10.1016/S0140-6736(12)60397-3
20. Colegrave M, Peckham M. Structural implications of  $\beta$ -cardiac myosin heavy chain mutations in human disease. *Anat Rec*. 2014; doi:10.1002/ar.22973
21. Previs MJ, Previs SB, Gulick J, Robbins J, Warshaw DM. Molecular Mechanics of Cardiac Myosin-Binding Protein C in Native Thick Filaments. *Science* (80- ). 2012;337: 1215–1218. doi:10.1126/science.1223602
22. Previs MJ, Prosser BL, Mun JY, Previs SB, Gulick J, Lee K, et al. Myosin-binding protein C corrects an intrinsic inhomogeneity in cardiac excitation-contraction coupling. *Sci Adv*. 2015; doi:10.1126/sciadv.1400205
23. Helms AS, Thompson AD, Glazier AA, Hafeez N, Kabani S, Rodriguez J, et al. Spatial and functional distribution of MYBPC3 pathogenic variants and clinical outcomes in patients with hypertrophic cardiomyopathy. *Circ Genomic Precis Med*. 2020; doi:10.1161/CIRCGEN.120.002929
24. Marston S, Copeland O, Jacques A, Livesey K, Tsang V, McKenna WJ, et al. Evidence From Human Myectomy Samples That MYBPC3 Mutations Cause Hypertrophic Cardiomyopathy Through Haploinsufficiency. *Circ Res*. 2009;105: 219–222. doi:10.1161/CIRCRESAHA.109.202440
25. Erdmann J, Raible J, Maki-Abadi J, Hummel M, Hammann J, Wollnik B, et al. Spectrum of clinical phenotypes and gene variants in cardiac myosin-binding protein C mutation carriers with hypertrophic cardiomyopathy. *J Am Coll Cardiol*. 2001; doi:10.1016/S0735-1097(01)01387-0
26. Page SP, Kounas S, Syrris P, Christiansen M, Frank-Hansen R, Andersen PS, et al. 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*. 2012; doi:10.1161/CIRCGENETICS.111.960831
27. Lopes LR, Barbosa P, Torrado M, Quinn E, Merino A, Ochoa JP, et al. Cryptic Splice-Altering Variants in MYBPC3 Are a Prevalent Cause of Hypertrophic Cardiomyopathy. *Circulation: Genomic and Precision Medicine*. 2020. doi:10.1161/CIRCGEN.120.002905
28. Ko C, Arscott P, Concannon M, Saberi S, Day SM, Yashar BM, et al. Genetic testing impacts the utility of prospective familial screening in hypertrophic cardiomyopathy through identification of a nonfamilial subgroup. *Genet Med*. 2018;20: 69–75. doi:10.1038/gim.2017.79
29. Marian AJ. Challenges in the Diagnosis of Anderson-Fabry Disease: A Deceptively Simple and Yet Complicated Genetic Disease\*. *Journal of the American College of Cardiology*. 2016. doi:10.1016/j.jacc.2016.06.026
30. Gelb BD, Roberts AE, Tartaglia M. Cardiomyopathies in Noonan syndrome and the other RASopathies. *Progress in Pediatric Cardiology*. 2015. doi:10.1016/j.ppedcard.2015.01.002
31. Elliott P, Baker R, Pasquale F, Quarta G, Ebrahim H, Mehta AB, et al. Prevalence of Anderson-Fabry disease in patients with hypertrophic cardiomyopathy: The European Anderson-Fabry Disease Survey. *Heart*. 2011; doi:10.1136/heartjnl-2011-300364
32. Sankaranarayanan R, Fleming EJ, Garratt CJ. Mimics of Hypertrophic Cardiomyopathy – Diagnostic Clues to Aid Early Identification of Phenocopies. *Arrhythmia Electrophysiol Rev*. 2013;2: 36. doi:10.15420/aer.2013.2.1.36
33. Ito K, Patel PN, Gorham JM, McDonough B, DePalma SR, Adler EE, et al. Identification of pathogenic gene mutations in LMNA and MYBPC3 that alter RNA splicing. *Proc Natl Acad Sci U S A*. 2017; doi:10.1073/pnas.1707741114
34. Bagnall RD, Ingles J, Dinger ME, Cowley MJ, Ross SB, Minoche AE, et al. Whole Genome Sequencing Improves Outcomes of Genetic Testing in Patients With Hypertrophic Cardiomyopathy. *J Am Coll Cardiol*. 2018;72: 419–429. doi:10.1016/j.jacc.2018.04.078
35. Janin A, Chanavat V, Rollat-Farnier PA, Bardel C, Nguyen K, Chevalier P, et al. Whole MYBPC3 NGS sequencing as a molecular strategy to improve the efficiency of molecular diagnosis of patients with hypertrophic cardiomyopathy. *Hum Mutat*. 2020; doi:10.1002/humu.23944
36. Frank-Hansen R, Page SP, Syrris P, McKenna WJ, Christiansen M, Andersen PS. 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*. 2008;16: 1062–1069. doi:10.1038/ejhg.2008.52
37. Singer ES, Ingles J, Semsarian C, Bagnall RD. Key Value of RNA Analysis of MYBPC3 Splice-Site Variants in Hypertrophic Cardiomyopathy. *Circ Genomic Precis Med*. 2019; doi:10.1161/CIRCGEN.118.002368
38. Xu M, Orsborne C, Eden J, Wallace A, Church HJ, Tylee K, et al. Mosaic Fabry Disease in a Male Presenting as Hypertrophic Cardiomyopathy. *Cardiogenetics*. 2020; doi:10.3390/cardiogenetics11010001
39. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019; doi:10.1016/j.cell.2018.12.015
40. Desmet F-O, Hamroun D, Lalande M, Collod-B eroud G, Claustres M, B eroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37: e67–e67. doi:10.1093/nar/gkp215
41. Koch L. Exploring human genomic diversity with gnomAD. *Nat Rev Genet*. 2020;21: 448–448. doi:10.1038/s41576-020-0255-7
42. Kishore S, Khanna A, Stamm S. Rapid generation of splicing reporters with pSpliceExpress. *Gene*. 2008;427: 104–110. doi:10.1016/j.gene.2008.09.021
43. Thomas HB, Wood KA, Buczek WA, Gordon CT, Pingault V, Atti -Bitach T, et al. EFTUD2 missense variants disrupt protein function and splicing in mandibulofacial dysostosis Guion-Almeida type. *Hum Mutat*. 2020;41: 1372–1382. doi:10.1002/

- humu.24027
44. Rodríguez-García MI, Monserrat L, Ortiz M, Fernández X, Cazón L, Núñez L, et al. Screening mutations in myosin binding protein C3 gene in a cohort of patients with Hypertrophic Cardiomyopathy. *BMC Med Genet.* 2010; doi:10.1186/1471-2350-11-67
  45. Ellingford JM, Thomas HB, Rowlands C, Arno G, Beaman G, Gomes-Silva B, et al. Functional and in-silico interrogation of rare genomic variants impacting RNA splicing for the diagnosis of genomic disorders. *bioRxiv.* 2019. doi:10.1101/781088
  46. Ingles J, Doolan A, Chiu C, Seidman J, Seidman C, Semsarian C. Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. *Journal of medical genetics.* 2005. doi:10.1136/jmg.2005.033886
  47. Helms AS, Davis FM, Coleman D, Bartolone SN, Glazier AA, Pagani F, et al. Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy. *Circ Cardiovasc Genet.* 2014; doi:10.1161/CIRCGENETICS.113.000448
  48. Carrier L, Bonne G, Bährend E, Yu B, Richard P, Niel F, et al. 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.* 1997; doi:10.1161/01.res.0000435859.24609.b3
  49. Flavigny J, Souchet M, Sébillon P, Berrebi-Bertrand I, Hainque B, Mallet A, et al. COOH-terminal truncated cardiac myosin-binding protein C mutants resulting from familial hypertrophic cardiomyopathy mutations exhibit altered expression and/or incorporation in fetal rat cardiomyocytes. *J Mol Biol.* 1999; doi:10.1006/jmbi.1999.3276
  50. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17: 405–423. doi:10.1038/gim.2015.30
  51. Zheng CL, Xiang-Dong FU, Gribkov M. Characteristics and regulatory elements defining constitutive splicing and different modes of alternative splicing in human and mouse. *RNA.* 2005; doi:10.1261/rna.2660805
  52. Wang Y, Liu J, Huang B, Xu Y-M, Li J, Huang L-F, et al. Mechanism of alternative splicing and its regulation. *Biomed Reports.* 2015;3: 152–158. doi:10.3892/br.2014.407
  53. Walsh R, Thomson KL, Ware JS, Funke BH, Woodley J, McGuire KJ, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet Med.* 2017; doi:10.1038/gim.2016.90
  54. Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, et al. Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet.* 1996; doi:10.1038/ng0596-63
  55. Szczesna D, Ghosh D, Li Q, Gomes A V., Guzman G, Arana C, et al. Familial Hypertrophic Cardiomyopathy Mutations in the Regulatory Light Chains of Myosin Affect their Structure, Ca<sup>2+</sup> Binding, and Phosphorylation. *J Biol Chem.* 2001; doi:10.1074/jbc.M009823200
  56. Kazmierczak K, Muthu P, Huang W, Jones M, Wang Y, Szczesna-Cordary D. Myosin regulatory light chain mutation found in hypertrophic cardiomyopathy patients increases isometric force production in transgenic mice. *Biochem J.* 2012;442: 95–103. doi:10.1042/BJ20111145
  57. Farman GP, Muthu P, Kazmierczak K, Szczesna-Cordary D, Moore JR. Impact of familial hypertrophic cardiomyopathy-linked mutations in the NH2 terminus of the RLC on  $\beta$ -myosin cross-bridge mechanics. *J Appl Physiol.* 2014; doi:10.1152/jappphysiol.00798.2014