Protein thermodynamic destabilization in the assessment of pathogenicity of a variant of uncertain significance in cardiac myosin binding protein C

Maria Rosaria Pricolo[1,2,*], Elías Herrero-Galán[1], Cristina Mazzaccara[2,3], Maria Angela Losi[4], Jorge Alegre-Cebollada[1,*], Giulia Frisso[2,3]

[1] Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain
[2] Dipartimento di Medicina Molecolare e Biotecnologie Mediche. Università Federico II, Naples, Italy.
[3] CEINGE Biotecnologie Avanzate, scarl, Naples, Italy
[4] Dipartimento di Scienze Biomediche Avanzate, Università Federico II, Naples, Italy.

*To whom correspondence should be addressed: mrpricolo@cnic.es, jalegre@cnic.es

address of authors: elias.herrero@cnic.es; cristina.mazzaccara@unina.it; losi@unina.it; gfrisso@unina.it

ABSTRACT

In the era of Next Generation Sequencing (NGS), genetic testing for inherited disorders identifies an ever-increasing number of variants whose pathogenicity remains unclear. These variants of uncertain significance (VUS) limit the reach of genetic testing in clinical practice. The VUS for Hypertrophic Cardiomyopathy (HCM), the most common familial heart disease, constitute over 60% of entries for missense variants shown in ClinVar database. We have studied a novel VUS (c.1809T>G-p.I603M) in the most frequently mutated gene in HCM, *MYBPC3*, which codes for cardiac myosin-binding protein C (cMyBPC). Our determinations of pathogenicity integrate bioinformatics evaluation and functional studies of RNA splicing and protein thermodynamic stability. In silico prediction and mRNA analysis indicated no alteration of RNA splicing induced by the variant. At the protein level, the p.I603M mutation maps to the C4 domain of cMyBPC. Although the mutation does not perturb much the overall structure of the C4 domain, the stability of C4 I603M is severely compromised as detected by circular dichroism and differential scanning calorimetry experiments. Taking into account the highly destabilizing effect of the mutation in the structure of C4, we propose reclassification of variant p.I603M as likely pathogenic. Looking into the future, the workflow described here can be used to refine the assignment of pathogenicity of variants of uncertain significance in *MYBPC3*.

Keywords: VUS, HCM, cMyBC

1. Introduction

Hypertrophic cardiomyopathy (HCM) constitutes the most common inherited disease of the myocardium with an estimated prevalence ranging from 1:500 to 1:200 [1, 2]. Although the clinical manifestations of HCM are highly variable, the disease is characterized by left ventricular hypertrophy (LVH) in the absence of triggers, such as hypertension, and in some cases sudden cardiac death (SCD) may be the first manifestation of the disease [3–5]. Importantly, HCM has a very strong genetic component, and up to 50% of patients who meet the diagnostic criteria have pathogenic mutations in genes coding for sarcomere proteins [2, 6–9]. The genetic testing for HCM is key to identify patients at high risk before the occurrence of clinical manifestations. Thus, the clinical role of genetic testing largely centers on family screening to facilitate presymptomatic diagnosis of family members, clinical surveillance and reproductive advice [10–13]. Due to the increased complexity of analysis and interpretation of genetic tests, general guidelines for the interpretation of variants have been published by the

American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) [14, 15]. By applying the proposed score, variants can be classified into five main groups including pathogenic, likely pathogenic, likely benign, benign, and uncertain significance variants (VUS). VUS should not be used for cascade screening in families and therefore are unable to distinguish between individuals at risk and those free from risk [16–18]. The numbers of mutations classified as VUS has increased in the last years [19], reaching 66.5% of entries in the ClinVar database for HCM missense mutations. However, since the frequency of detection of VUS is higher in HCM patients than in the general population [20], it can be deduced that a fraction of VUS are indeed pathogenic. Hence, to improve care of many HCM patients and their families, there is a pressing need to develop methods that can help reclassification of VUS into actionable categories. According to ACMG, in the absence of enough genetic support, functional studies are the most important criterion to establish pathogenicity of putative disease-causing mutations [15].

Mutations in the *MYBPC3* gene, encoding cardiac myosin-binding protein C (cMyBPC), represent 40–50% of all HCM mutations, making it the most frequently mutated gene in this disease [21–23]. *MYBPC3* is also the HCM gene with the highest number of missense VUS in ClinVar. *MYBPC3* pathogenic missense mutations can lead to stable mutant cMyBPCs that are, at least in part, incorporated into the sarcomere and could act as poison polypeptides by alteration of the structure and/or function of the sarcomere [23]. Missense mutations located in the central domains (C3-C7) of the protein, and clinically causative of HCM, have been identified to interfere with the structure and the stability of the domains [24, 25]. Other pathogenic missense mutations of cMyBPC, are known to cause RNA splicing disruptions [26, 27]. Interestingly, both protein destabilization and alterations of splicing can lead to less cMyBPC levels, a well-known driver of HCM [28–30].

During genetic screening of HCM patients, we found a new putative missense mutation in the *MYBPC3* gene (c.1809T>G-p.I603M), which targets the C4 central domain of cMyBPC. Here, we describe how available population and computational data result in the classification of this variant as VUS according to ACMG criteria. We then demonstrate that this variant induces extensive protein destabilization, which leads to reclassification of c.1809T>G-p.I603M as likely pathogenic.

2. Methods

2.1 Clinical diagnosis of HCM

Patients undergone physical examination, resting EKG, and 2D Doppler echocardiography. Echocardiography measurements were made according to current guidelines [31]. In particular, left ventricular (LV) wall thickness was measured at end-diastole in the parasternal short-axis views in 4 regions at the mitral valve and papillary level and in 2 regions at apical level. Maximal wall thickness (MWT) was defined as the greatest thickness [31]. In patients <18 year old, a diagnosis of HCM was made if MWT was >2 SDs above the body surface area–corrected population mean (z score >2), whereas in patients >18 year if MWT was >12 mm not explained by abnormal loading conditions or in accordance with published criteria for familial disease [31]. LV outflow tract obstruction and diastolic function were investigated by Doppler echocardiography. Twelve-lead EKGs were analyzed checking for the following: QRS axis, abnormal Q waves, Sokolow-Lyon voltage criteria for LV hypertrophy and repolarization abnormalities [32].

2.2 Molecular genetics of HCM patients

Human samples to perform genetic analysis were obtained following informed consent of patients according to the Declaration of Helsinki. Genomic DNA was isolated from peripheral whole blood as previously described [33]. All coding exons, and 5′ and 3′ UTRs of genes involved in HCM were amplified by PCR and analysed by automatic sequencing using previously reported protocols [10].

2.3 RNA Splicing predictions and analysis

Alamut software (Alamut® Visual, Interactive Biosoftware) was used for *in silico* prediction of splice-affecting nucleotide variant [34, 35]. Genomic sequences (WT and mutant) were analysed using five splicing prediction tools (SpliceSiteFinder-like, MaxEntScan, Neural Network Splice, GeneSplicer, and Human Splicing Finder). RNA splicing was experimentally examined using mRNA from peripheral blood. Total RNA was extracted from lymphocyte cells of patients' peripheral blood using Trizol Reagent (Thermo Fischer Scientific, Waltham, MA, USA). RNA retro-transcription was performed by SuperScript VILO (Life Technologies), starting from 1μg of total RNA and using random primers. The cDNA obtained was amplified by PCR using specific consecutive-exon-spanning primers: **MYBPC3_ex15_Fw** 5'-CAAGCGTACCCTGACCATCA-'3 and **MYBPC3_ex20-21_Rv** 5'-GGATCTTGGGAGGTTCCTGC-'3 oligonucleotides which anneal with exon 15 and the region encompassing exons 20 and 21 of the *MYBPC3* mRNA, respectively. The same primers were used for sequencing the PCR fragment. Sequences were analysed with CodonCode Aligner software (CodonCode Corporation, Dedham, MA, USA).

2.4 Bioinformatics predictions of pathogenicity at the protein level

Missense mutation at the protein level was evaluated with three independent bioinformatics tools that predict the potential impact of amino acid substitutions on the structure and function of the protein (PolyPhen-2, SIFT, Provean). **PolyPhen-2** (Polymorphism Phenotyping v2) evaluates the impact of amino acid allelic variants via analysis of multiple sequence alignments and protein 3D-structures [36]. **SIFT** (Sorting Intolerant From Tolerant) presumes that important amino acids will be conserved in the protein family, so changes at well-conserved positions tend to be predicted as deleterious [37]. **PROVEAN** (Protein Variation Effect Analyzer) predicts whether an amino acid substitution has an impact on the biological function of a protein using pairwise sequence alignment scores [38].

2.5 Homology modelling

For modelling of cMyBPC's C4 domain, the protein sequence Q14896 from the UniProt database was used (aa.544-aa.633). The structure was modelled using I-TASSER tool (Iterative Threading ASSEmbly Refinement) [39], indicating immunoglobulin domain four of slow-MyBPC (2YUZ PDB) as the best template. The quantitative measuring of the similarity between two superimposed atomic coordinates is expressed as the RMSD (Root Mean Square Deviation) of the positions of the alpha carbons of equivalent amino acids. PyMol software was used for molecular representation and RMSD calculation[40].

2.6 Protein expression and purification for biophysical characterization

The recombinant WT and mutant I603M C4 domains (Table 1) were engineered and purified

for thermodynamic analysis. With that aim, the cDNA fragment including exons 18 and 19 of MYBPC3, was amplified from cardiac RNA with **cMyBPC_C4_Fw** and **cMyBPC _C4 _Rv** oligonucleotides (Table 2) and cloned in a custom-modified pQE80L (Qiagen) using BamHI and BglII restriction sites. I603M mutant cDNA was generated by PCR site-directed mutagenesis. Final constructs were verified by Sanger sequencing. Proteins were expressed in the BLR(DE3) *E. coli* strain and purified from the soluble fraction using nickel-nitrilotriacetic acid (Ni-NTA) chromatography (Qiagen) with a column volume of 3 ml and pre-equilibrated with phosphate buffer (50 mM sodium phosphate pH 7 and 300 mM NaCl) supplemented with 10 mM DTT. Elution was performed in two steps with increasing imidazole concentration (from 20mM to 250mM). Further purification was achieved by size-exclusion chromatography in an AKTA Pure 25l system using a Superdex 200 Increase 10/300 GL column (GE Healthcare). Proteins were eluted in phosphate buffer buffer (20 mM sodium phosphate pH 6.5, 50 mM NaCl), which was used both in circular dichroism (CD) and differential scanning calorimetry (DSC) experiments. All domains were at least 90–95% pure as estimated by SDS-PAGE analysis.

2.7 Circular Dichroism (CD)

CD spectra of domain C4 were collected on a Jasco-810 spectropolarimeter. Temperature was controlled using a Peltier thermoelectric system. Purified proteins in phosphate were loaded into a 0.1cm or 1cm path-length quartz cuvette for data collection in the far-UV (195-250 nm) or the near-UV (250-350nm), respectively. Protein concentration was 0.3 mg/ml (far-UV) and 1 mg/ml (near-UV). Spectra were collected for protein samples and bufferat 25°C and 85°C with four accumulations. The buffer baseline spectrum was subtracted from each protein spectrum at each corresponding temperature to correct for the background signal. Thermal denaturation analyses were carried out collecting the variations of ellipticity at 230 nm as a function of temperature (25°C-85°C, at a rate of 30ºC/h). Protein refolding was analysed with a temperature ramp-down (85°C-25°C) at the same speed. We plotted the ellipticity versus temperature and data points were fit to a Boltzmann sigmoidal curve using the Igor Pro software to obtain midpoint unfolding temperatures (T_m). The changes in CD as a function of temperature were used to determine the van't Hoff enthalpy (ΔH_v) of unfolding considering $\Delta Cp = 0$ [41].

2.8 Differential scanning calorimetry (DSC)

Calorimetric measurements were performed using a Microcal VP-DSC differential scanning calorimeter with 0.5 ml cells. Experiments were done by increasing temperature from 25°C to 85°C at a rate of 30ºC/h, using 0.085 mM (1 mg/ml) protein concentration in phosphate buffer. The reversibility of the thermal transitions was assessed by reheating the sample immediately after cooling from the previous scan, using the same rate of temperature change. Calorimetric traces were subtracted the baseline of a scan performed with buffer in both cells and corrected for the instrumental background. The temperature dependence of the excess heat capacity was analysed using the Origin software (MicroCal, Northampton, MA). The thermal stability of the proteins was described by its T_m, and the calorimetric enthalpy (ΔH_{cal}), which was calculated as the area under the excess heat capacity function. The van't Hoff enthalpy (ΔH_v) associated with DSC thermograms was determined by a two-state fit of the thermograms [42, 43]. The Gibbs free energy change (ΔG) was then calculated at any temperature using three experimental parameters: T_m, enthalpy change at T_M ($\Delta H(T_m)$) and heat capacity change at T_m (ΔC_p). ΔG was derived for C4 WT and C4 I603M considering $\Delta C_p = 0$.

3. Results

3.1 c.1809T>G (p.I603M) is classified as VUS according to available population and bioinformatics predictions

The genetic screening of an HCM patient highlighted a missense variant in MYBPC3 (c.1809T>G-p.I603M) that was not reported before as associated with HCM. The variant occurs in exon 19 of the gene, whereas at the protein level Ile603 maps to the central domain C4 of cMyBPC. The p.I603M substitution was identified in a proband with family history of HCM (Table 3). Although segregation analysis was possible, the family tree was not informative (*Figure 1, a*). The genetic screening of other family members evidenced the presence of the c.1809T>G-p.I603M variant in two affected subjects in heterozygosis with another MYBPC3 mutation, c.98_99del CA-p.T33RfsX15. One subject (III.5 in *Figure 1, a*) is clinically affected but did not carry the p.I603M mutation, although hypertrophy in this individual can be secondary to coarctation of the aorta. Two young subjects carry the c.1809T>G-p.I603M variant and show so far no clinical evidence of HCM (IV.1 and IV.2 in *Figure 1, a;* Table 3).

We applied the ACMG classification criteria to assess this variant, looking at population and computational evidence of pathogenicity. We first evaluated Minor Allele Frequency (MAF) in the ExAC and GnomAD databases [44, 45]. The MAF of allele c.1809G is 0.00004267 and 0.0000142 in ExAC and GnomAD, respectively. These MAF values indicate that c.1809T>G is a rare variant in the population, compatible with HCM [20]. Evaluation using Polyphen, SIFT and PROVEAN tools showed strong evidence for the deleterious effect of the p.I603M variant. According to the low MAF, the results from the bioinformatics predictors and the co-segregation of this variant with disease in multiple affected family members (PP3, PM2 and PP1 of ACMG criteria), the variant c.1809T>G (p.I603M) is classified as VUS.

Hence, we experimentally determined if the mutation can affect the bioavailability of cMyBP-C at the level of alterations of RNA processing or protein stability, as reported for other mutants [23–27, 46, 47].

3.2 c.1809T>G (p.I603M) does not induce alterations in RNA splicing

The removal of introns from pre-mRNA and the joining of exons are critical aspects of gene expression. Introns are removed from primary transcripts by the process of RNA splicing, which links together the flanking exons to generate mRNA [48]. We investigated potential RNA splicing alterations induced by variant c.1809T>G-p.I603M. Bioinformatics predictions did not suggest loss of natural splicing sites. Indeed, all five algorithms did not show any probability of canonical splicing site loss compared to a well-established pathogenic mutation c.1624G>C (p.E542Q) of *MYBPC3* [29, 49] (*Figure 1, b*). Similar results were obtained for the prediction of activation of cryptic splicing sites (data not shown). To verify predictions, we amplified the fragment spanning from exon 15 to exon 21 using mRNA isolated from peripheral blood of the index case III-3. Gel electrophoresis showed an identical migration pattern for both mutant carrier and a healthy control (*Figure 1, c right*). Sanger sequencing confirmed the correct splicing in both cases (*Figure 1, c left*). Furthermore, both c.98_99del CA and c.1809T>G mRNAs are detected in the sequencing electropherogram (Figure 1,c).fdr0ki

3.3 C4 WT and C4 I603M share a similar β-sheet rich structure

p.I603M is a missense variant that affects a highly conserved residue in the central domain C4 of cMyBPC (*Figure 2, a,b)*. Since the structure of domain C4 is not known, to predict how the protein structure might be affected by the p.I603M mutation, the homology modelling of the domain was carried out using I-TASSER. The resulting model showed that the C4 domain adopts an Ig-like folding (*Figure 2, b*) [50]. The predicted structure of C4 I603M is very similar to WT. The RMSD (Root Mean Square Deviation) of both predictions with respect to the template is 1.9±1.6Å for WT and 2.0±1.6Å for p.I603M. The RMSD of WT and p.I603M homology models is 0.2634Å.
Both isoleucine and methionine share similar physicochemical properties, suggesting no effect on protein structure by mutation p.I603M. Still, it is interesting to observe that Ile603 is buried in the core of the WT domain (*Figure 2, b*), so a mutation affecting this residue may not be easily accommodated, potentially leading to destabilization. To study this scenario, recombinant WT and I603M C4 domains were produced (*Figure 2, c*) and analysed by Circular Dichroism (CD) in the far-UV (200-250 nm, informative of secondary structure) and near-UV (250-350 nm, resulting from tertiary structure) at 25°C (*Figure 3, a,b*). In agreement with the homology modelling prediction, the far-UV spectrum of C4 WT showed a minimum of ellipticity at 215 nm, characteristic of β-sheet structure [51, 52]. C4 WT and I603M spectra showed similar CD spectra, suggesting that the mutation does not have a major impact in the fold of the protein at 25ºC (*Figure 3, a,b*).

3.4 Mutation p.I603M destabilizes the C4 domain

Thermodynamic protein stability was investigated by monitoring the CD signal at 230nm while increasing temperature from 25°C to 85°C. The unfolding profiles showed a T_m of 58°C for C4 WT and 45°C for the I603M mutant (*Figure 3, c; Supplementary Figure 1;* Table 4). A 13ºC decrease in T_m highly suggests that the I603M mutation induces considerable destabilization of the C4 domain.

For further thermodynamic characterization, we undertook refolding CD experiments by ramping the temperature down after the initial heating ramp. The refolding ability of the domains was determined by collecting a final far-UV CD spectrum at 25°C. Most protein refolded into original secondary structure for both WT and I603M (*Supplementary Figure 2*), suggesting that thermal denaturation of both proteins is a reversible process. We used the Gibbs-Helmholtz equation to fit the change of ellipticity at 230 nm as a function of temperature, which enabled estimation of the van't Hoff enthalpies (ΔH_v). We obtained that ΔH_v of C4 WT is 220 kJ/mol, whereas the value for I603M is 130 kJ/mol for C4 I603M (Table 4). Such difference in enthalpies shows considerable thermodynamic destabilization induced by the I603M mutation.

When the unfolding of a protein is reversible and two-state, the thermodynamic parameters evaluated by CD are almost identical to the ones estimated by calorimetric methods [41]. However, to confirm thermodynamic destabilization, differential scanning calorimetry (DSC) experiments were carried out in the same conditions as CD experiments [43]. The thermal peaks of both proteins indicate an exothermic transition with heat release, typical of protein unfolding transitions (*Figure 3, d*). The T_m values calculated in DSC are 58°C and 45°C for WT and I603M respectively, matching the values obtained in CD. The calorimetric enthalpy (ΔH_{cal}) was calculated by integration of the area under the thermogram peak and resulted in 364 kJ/mol for C4 WT and 196 kJ/mol for C4 I603M. ΔH_v was determined by a fit of the thermogram to a two-state model of unfolding. Again, the ΔH_v (WT) is higher than ΔH_v (I603M), 330 kJ/mol and 222 kJ/mol, respectively (Table 4). The $\Delta H_{cal}/\Delta H_v$ ratio is 1.1 for WT and 0.89 for mutant, suggesting that the two-state model of unfolding is a reasonable approximation to the unfolding process of both proteins. The Gibbs free energy change (ΔG)

of unfolding was calculated at 25°C considering $\Delta C_p=0$. The $\Delta\Delta G$ associated with the I603M mutation is 4.5 kcal/mol, confirming its highly destabilizing properties.

4. DISCUSSION

In the era of Next Generation Sequencing, an ever-increasing number of variants whose pathogenicity remains unclear is detected during genetic testing. In particular, many VUS that have no clinical value are identified [17, 19] . Indeed, the majority of genetic variants across all actionable genes are currently classified as VUS. This issue has become particularly challenging in the genetic diagnosis of HCM, one of the most common cardiac inherited diseases [4, 12]. Pathogenic mutations in HCM-associated genes are found in more than 60% of HCM patients, allowing cascade predictive genetic testing to early diagnosed relatives who might be at risk of disease-related complications, such as sudden cardiac death. A different scenario follows identification of a VUS in the proband. Due to the absence of pathogenicity assignment of VUS, relatives cannot be reassured about their potential to develop HCM. Indeed, the Association for Clinical Genetic Science does not recommend predictive testing on family members following the finding of a VUS in an index case [53], which limits the utility of genetic testing in clinical practice. Functional assessment of VUS can help define their pathogenic nature. According to the ACMG, in the absence of enough genetic and population data, the most important criterion to establish causality of putative disease-causing mutations is a well-established functional study to estimate the impact of mutations on a gene or protein function.

The c.1809T>G-p.I603M variant identified during genetic testing of an HCM patient was classified as VUS following ACMG criteria PP1, PM2, PP3. Hence, the variant was selected for functional study to determine pathogenicity. We first evaluated if the c.1809T>G (p.I603M) variant induces alterations in RNA splicing. During splicing, introns are removed from primary transcripts, in a process that links together the flanking exons to generate mRNA [48]. During the editing of pre-mRNA transcripts, the splicing machinery recognizes consensus sequences that also include the exons, particularly in the regions close to the intron-exon boundaries [54]. Results show that c.1809T>G-p.I603M does not induce alterations in RNA splicing (*Figure 1, b,c*).
Regarding protein stability, the variant does not perturb much the structure of the domain at 25ºC, according to the far-UV and near-UV CD spectra (*Figure 3, a,b*). However, the thermodynamic stability of the C4 I603M mutant is severely compromised, as shown by lower T_m in thermal denaturation experiments by CD and DSC (*Figure 3, c,d*). Therefore, the p.I603M variant alters the stability of the central C4 domain of MYBPC3, which can lead to more frequent protein unfolding that can result in degradation of cMyBPC containing the mutation I603M [55]. This scenario may induce cMyBPC haploinsufficiency, which is believed to be a major pathogenic mechanism in truncating variants [56]. If we consider this damaging effect on protein stability as causative of HCM (ACMG PS3 criterion), the pathogenicity assessment would change to likely pathogenic. Our results exemplify that protein destabilization could be useful to guide clinical assessment of genetic variants of cMyBPC. To achieve this goal, exhaustive screening of pathogenic and non-pathogenic variants will be needed.

Remarkably, variant p.I603M was identified in clinically affected subjects in combination with the p.T33RfsX15 truncating variant of cMyBPC (*Figure 1, a*). The two young subjects that have only the p.I603M do not show the phenotype although, according to our results, we do not exclude that they can develop the disease in the future. There is only one subject of the

family that has cardiac hypertrophy in the absence of the mutation p.I603M, although the clinical manifestations probably are secondary to the coarctation of the aorta caused by an accident. In addition, there is one patient that carries only the truncated mutation and does not show HCM phenotype. This family reflects the difficulties associated with genetic testing in a clinical setting. In principle, frameshift mutations in MYBPC3 are thought to be causative of HCM [29]. Hence, if p.I603M is also pathogenic, two of the members of the family would be compound heterozygotes for pathogenic mutations. This situation usually results in pediatric-onset HCM, which we did not observe in our case although our limited pedigree data is in agreement with a more severe phenotype in compound heterozygous. This observation suggests that at least one of the variants is not highly damaging. Regarding p.T33RfsX15, there has been no study on whether N-terminal truncations in cMyBPC are as pathogenic as central ones. The existence of re-start of translation downstream to the mutation-driven stop codon has been observed [57, 58], which may lead to less damaging molecular phenotypes in these N-terminal truncations. In support of this scenario, we have detected mRNA coming from both alleles of the index case subject (III-3), suggesting that the p.T33RfsX15 is not targeted by non-sense mediated decay in peripheral blood. Regarding p.I603M, it remains to be determined how the clinical severity of a mutation scales with the degree of destabilization, for which many pathogenic mutations need to be analysed.

In summary, our study shows the advantages offered by functional assessment of mutations in the assignment of pathogenicity in the context of HCM. To fulfil its potential, we identify the need of extensive profiling of pathogenic mutations to find molecular phenotypes, such as protein destabilization [47, 59], that can be used to guide functional assessment.

Acknowledgements

J.A.C. acknowledges funding from the Ministerio de Ciencia, Innovación y Universidades (MCNU) through grants BIO2017-83640-P (AEI/FEDER, UE) and RYC-2014-16604, the European Research Area Network on Cardiovascular Diseases (ERA-CVD/ISCIII, MINOTAUR, AC16/00045), the *Comunidad de Madrid* (P2018/NMT-4443), and the CNIC-Severo Ochoa intramural grant program (03-2016 IGP). The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), MCNU and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).
G.F. acknowledges funding from the Ministero dell'Istruzione, dell'Università e della Ricerca-RomePS35-126/IND.
We acknowledge the BioBank of CEINGE SCaRL Advanced Biotechnology, Naples, Italy, which provided the biological samples.

Table 1: Properties of recombinant C4 WT and I603M. Sequences corresponding to C4 are indicated in bold whereas extra amino acids resulting from cloning are shown in regular type. Predicted extinction coefficients were obtained from ProtParam tool[60].

Protein	Sequence	Number of amino acids	Extinction coefficient (cm^2 x mg^{-1})
C4 WT	MRGSHHHHHHGS**KLEVYQSIADLMVG AKDQAVFKCEVSDENVRGVWLKNGKE LVPDSRIKVSHIGRVHKLTIDDVTPADE ADYSFVPEGFACNLSAKLHFME**RS	104	0.736

C4 I603M	MRGSHHHHHHGS**KLEVYQSIADLMVG AKDQAVFKCEVSDENVRGVWLKNGKE LVPDSRIKVSHIGRVHKLTMDDVTPAD EADYSFVPEGFACNLSAKLHFME**RS	104	0.724

Table 2: Primers used to clone C4 (5'-3')

Primer name	Adapter sequence	Primer Sequence
cMyBPC_C4_Fw	CGCGGATCC	AAGCTGGAGGTGTACCAGAGCAT
cMyBPC_C4_Rv	ATAGGTACCTTAGCA ACAAGATCT	CTCCATGAAGTGGAGCTTGGCTG

Table 3: Clinical data

PATIENT	AGE AT ONSET	SYNCOPE	CONDUCTION DEFECTS	ARRHYTHMIAS	LVEDD (mm)	LVESD (mm)	EF (%)	MWT (mm)
III.1	39	-	LAH	-	40	23	70	26
III.2	asymptomatic	-	-	-	45	29	56	10
III.3*	17	-	RBBB	NSVT	48	27	58	19
III.5	21	-	-	-	NR	NR	65	19
IV.1	asymptomatic	-	-	-	42	21	57	9
IV.2	asymptomatic	-	-	-	44	23	65	7

*proband; LVEDD: left ventricular end diastolic diameter; LVESD: left ventricular end systolic diameter; EF: ejection fraction; MWT: maximal wall thickness; RBBB: right bundle branch block; LAH: left anterior hemiblocks; NSVT: nonsustained ventricular tachycardia; NR: not reported

Table 4: thermal stability parameters calculated by CD and DSC

	T_m CD	ΔH_v CD_{230}	T_m DSC	ΔH_{cal} DSC	ΔH_v DSC
C4 WT	58 °C	220 kJ/mol	58°C	364 kJ/mol	330 kJ/mol
C4 I603M	45 °C	130 kJ/mol	45°C	196 kJ/mol	222 kJ/mol

Conflict of Interest
The authors declare that they have no conflict of interest.

Ethical approval
Human samples to perform genetic analysis were obtained following informed consent of patients according to the Declaration of Helsinki and its later amendments or comparable ethical standards. Research involving humans was authorized by the Ethics Committee of the Naples University Federico II "Carlo Romano" (Protocol number 157/13).
This article does not contain any studies with animals performed by any of the authors.

References

1. Maron, B. J., Gardin, J. M., Flack, J. M., Gidding, S. S., Kurosaki, T. T., & Bild, D. E. (1995). Prevalence of Hypertrophic Cardiomyopathy in a General Population of Young Adults. *Circulation*, *92*(4), 785–789. doi:10.1161/01.CIR.92.4.785
2. Semsarian, C., Ingles, J., Maron, M. S., & Maron, B. J. (2015). New perspectives on the prevalence of hypertrophic cardiomyopathy. *Journal of the American College of Cardiology*, *65*(12), 1249–1254. doi:10.1016/j.jacc.2015.01.019
3. Soler, R., Méndez, C., Rodríguez, E., Barriales, R., Ochoa, J. P., & Monserrat, L. (2018). Phenotypes of hypertrophic cardiomyopathy. An illustrative review of MRI findings. *Insights into Imaging*. doi:10.1007/s13244-018-0656-8
4. Wijnker, P. J. M., Sequeira, V., Kuster, D. W. D., & Velden, J. van der. (2019). Hypertrophic Cardiomyopathy: A Vicious Cycle Triggered by Sarcomere Mutations and Secondary Disease Hits. *Antioxidants & redox signaling*, *31*(4), 318–358. doi:10.1089/ars.2017.7236
5. Geske, J. B., Ommen, S. R., & Gersh, B. J. (2018). Hypertrophic Cardiomyopathy: Clinical Update. *JACC: Heart Failure*, *6*(5), 364–375. doi:10.1016/j.jchf.2018.02.010
6. Watkins, H., Ashrafian, H., & Redwood, C. (2011). Inherited Cardiomyopathies. *New England Journal of Medicine*, *364*(17), 1643–1656. doi:10.1056/NEJMra0902923
7. Salman, O. F., El-Rayess, H. M., Abi Khalil, C., Nemer, G., & Refaat, M. M. (2018). Inherited Cardiomyopathies and the Role of Mutations in Non-coding Regions of the Genome. *Frontiers in Cardiovascular Medicine*, *5*, 77. doi:10.3389/fcvm.2018.00077
8. Ingles, J., Doolan, A., Chiu, C., Seidman, J., Seidman, C., & Semsarian, C. (2005). Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. *Journal of medical genetics*, *42*(10), e59. doi:10.1136/jmg.2005.033886
9. Alfares, A. A., Kelly, M. A., McDermott, G., Funke, B. H., Lebo, M. S., Baxter, S. B., … Rehm, H. L. (2015). Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: Expanded panels offer limited additional sensitivity. *Genetics in Medicine*. doi:10.1038/gim.2014.205
10. Girolami, F., Frisso, G., Benelli, M., Crotti, L., Iascone, M., Mango, R., … Olivotto, I. (2018). Contemporary genetic testing in inherited cardiac disease. *Journal of Cardiovascular Medicine*, *19*(1), 1–11. doi:10.2459/JCM.0000000000000589
11. Limongelli, G., Monda, E., Tramonte, S., Gragnano, F., Masarone, D., Frisso, G., … Pacileo, G. (2019). Prevalence and clinical significance of red flags in patients with hypertrophic cardiomyopathy. *International Journal of Cardiology*.

doi:10.1016/j.ijcard.2019.06.073
12. Ingles, J., Burns, C., Barratt, A., & Semsarian, C. (2015). Application of Genetic Testing in Hypertrophic Cardiomyopathy for Preclinical Disease Detection. *Circulation: Cardiovascular Genetics*, *8*(6), 852–859. doi:10.1161/CIRCGENETICS.115.001093
13. Alejandra Restrepo-Cordoba, M., Campuzano, O., Ripoll-Vera, T., Cobo-Marcos, M., Mademont-Soler, I., Gámez, J. M., … Garcia-Pavia, P. (2017). Usefulness of Genetic Testing in Hypertrophic Cardiomyopathy: an Analysis Using Real-World Data. *Journal of Cardiovascular Translational Research*, *10*(1), 35–46. doi:10.1007/s12265-017-9730-8
14. Hershberger, R. E., Givertz, M. M., Ho, C. Y., Judge, D. P., Kantor, P. F., McBride, K. L., … Ware, S. M. (2018). Genetic Evaluation of Cardiomyopathy—A Heart Failure Society of America Practice Guideline. *Journal of Cardiac Failure*, *24*(5), 281–302. doi:10.1016/j.cardfail.2018.03.004
15. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Rehm, H. L. (2015). 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. *Genetics in Medicine*, *17*(5), 405–423. doi:10.1038/gim.2015.30
16. Das K, J., Ingles, J., Bagnall, R. D., & Semsarian, C. (2014). Determining pathogenicity of genetic variants in hypertrophic cardiomyopathy: importance of periodic reassessment. *Genetics in Medicine*, *16*(4), 286–293. doi:10.1038/gim.2013.138
17. Hoffman-Andrews, L. (2017). The known unknown: the challenges of genetic variants of uncertain significance in clinical practice. *Journal of Law and the Biosciences*, *4*(3), 648–657. doi:10.1093/jlb/lsx038
18. Burns, C., Bagnall, R. D., Lam, L., Semsarian, C., & Ingles, J. (2017). Multiple Gene Variants in Hypertrophic Cardiomyopathy in the Era of Next-Generation Sequencing. *Circulation: Cardiovascular Genetics*, *10*(4). doi:10.1161/CIRCGENETICS.116.001666
19. Weile, J., & Roth, F. P. (2018). Multiplexed assays of variant effects contribute to a growing genotype–phenotype atlas. *Human Genetics*, *137*(9), 665–678. doi:10.1007/s00439-018-1916-x
20. Walsh, R., Thomson, K. L., Ware, J. S., Funke, B. H., Woodley, J., McGuire, K. J., … Watkins, H. (2017). Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genetics in Medicine*, *19*(2), 192–203. doi:10.1038/gim.2016.90
21. Marston, S., Copeland, O., Jacques, A., Livesey, K., Tsang, V., McKenna, W. J., … Watkins, H. (2009). Evidence From Human Myectomy Samples That MYBPC3 Mutations Cause Hypertrophic Cardiomyopathy Through Haploinsufficiency. *Circulation Research*, *105*(3), 219–222. doi:10.1161/CIRCRESAHA.109.202440
22. Da'as, S. I., Fakhro, K., Thanassoulas, A., Krishnamoorthy, N., Saleh, A., Calver, B. L., … Nomikos, M. (2018). Hypertrophic cardiomyopathy-linked variants of cardiac myosin-binding protein C3 display altered molecular properties and actin interaction. *Biochemical Journal*, *475*(24), 3933–3948. doi:10.1042/BCJ20180685
23. Carrier, L., Mearini, G., Stathopoulou, K., & Cuello, F. (2015). Cardiac myosin-binding protein C (MYBPC3) in cardiac pathophysiology. *Gene*, *573*(2), 188–197. doi:10.1016/j.gene.2015.09.008
24. Nadvi, N. A., Michie, K. A., Kwan, A. H., Guss, J. M., & 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*(1), 105–115. doi:10.1016/j.str.2015.11.001

25. Pfuhl, M., & Gautel, M. (2012). Structure, interactions and function of the N-terminus of cardiac myosin binding protein C (MyBP-C): who does what, with what, and to whom? *Journal of Muscle Research and Cell Motility*, *33*(1), 83–94. doi:10.1007/s10974-012-9291-z
26. Carrier, L., Bonne, G., Bährend, E., Yu, B., Richard, P., Niel, F., … 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. *Circulation research*, *80*(3), 427–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9048664
27. Helms, A. S., Davis, F. M., Coleman, D., Bartolone, S. N., Glazier, A. A., Pagani, F., … Day, S. M. (2014). Sarcomere Mutation-Specific Expression Patterns in Human Hypertrophic Cardiomyopathy. *Circulation: Cardiovascular Genetics*, *7*(4), 434–443. doi:10.1161/CIRCGENETICS.113.000448
28. van Dijk, S. J., Dooijes, D., dos Remedios, C., Michels, M., Lamers, J. M. J., Winegrad, S., … van der Velden, J. (2009). Cardiac Myosin-Binding Protein C Mutations and Hypertrophic Cardiomyopathy. *Circulation*, *119*(11), 1473–1483. doi:10.1161/CIRCULATIONAHA.108.838672
29. Marston, S., Copeland, O., Gehmlich, K., Schlossarek, S., & Carrrier, L. (2012). How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *Journal of Muscle Research and Cell Motility*, *33*(1), 75–80. doi:10.1007/s10974-011-9268-3
30. Toepfer, C. N., Wakimoto, H., Garfinkel, A. C., McDonough, B., Liao, D., Jiang, J., … Seidman, C. E. (2019). Hypertrophic cardiomyopathy mutations in MYBPC3 dysregulate myosin. *Science Translational Medicine*, *11*(476), eaat1199. doi:10.1126/scitranslmed.aat1199
31. Authors/Task Force members, Elliott, P. M., Anastasakis, A., Borger, M. A., Borggrefe, M., Cecchi, F., … 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). *European heart journal*, *35*(39), 2733–79. doi:10.1093/eurheartj/ehu284
32. Norrish, G., Jager, J., Field, E., Quinn, E., Fell, H., Lord, E., … Kaski, J. P. (2019). Yield of Clinical Screening for Hypertrophic Cardiomyopathy in Child First-Degree Relatives. *Circulation*. doi:10.1161/CIRCULATIONAHA.118.038846
33. Carsana, A., Frisso, G., Tremolaterra, M. R., Ricci, E., De Rasmo, D., & Salvatore, F. (2007). A Larger Spectrum of Intragenic Short Tandem Repeats Improves Linkage Analysis and Localization of Intragenic Recombination Detection in the Dystrophin Gene. *The Journal of Molecular Diagnostics*, *9*(1), 64–69. doi:10.2353/jmoldx.2007.060056
34. Houdayer, C. (2011). In silico prediction of splice-affecting nucleotide variants. *Methods in molecular biology (Clifton, N.J.)*, *760*, 269–81. doi:10.1007/978-1-61779-176-5_17
35. Frisso, G., Detta, N., Coppola, P., Mazzaccara, C., Pricolo, M., D'Onofrio, A., … Salvatore, F. (2016). Functional Studies and In Silico Analyses to Evaluate Non-Coding Variants in Inherited Cardiomyopathies. *International Journal of Molecular Sciences*, *17*(11), 1883. doi:10.3390/ijms17111883
36. Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, *Chapter 7*, Unit7.20. doi:10.1002/0471142905.hg0720s76
37. Ng, P. C. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, *31*(13), 3812–3814. doi:10.1093/nar/gkg509
38. Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747.

doi:10.1093/bioinformatics/btv195
39. Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*(1), 40. doi:10.1186/1471-2105-9-40
40. Mier, P. D., & van den Hurk, J. J. (1975). Lysosomal hydrolases of the epidermis. 2. Ester hydrolases. *The British journal of dermatology*, *93*(4), 391–8. doi:10.1111/j.1365-2133.1975.tb06512.x
41. Greenfield, N. J. (2006). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nature Protocols*, *1*(6), 2527–2535. doi:10.1038/nprot.2006.204
42. Spink, C. H. (2008). Differential scanning calorimetry. *Methods in cell biology*, *84*, 115–41. doi:10.1016/S0091-679X(07)84005-2
43. Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics*, *531*(1–2), 100–109. doi:10.1016/j.abb.2012.09.008
44. Lek, M., Karczewski, K. J., Minikel, E. V, Samocha, K. E., Banks, E., Fennell, T., … MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291. doi:10.1038/nature19057
45. Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., … MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*. doi:10.1101/531210
46. Yue, P., Li, Z., & Moult, J. (2005). Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease. *Journal of Molecular Biology*, *353*(2), 459–473. doi:10.1016/j.jmb.2005.08.020
47. Smelter, D. F., de Lange, W. J., Cai, W., Ge, Y., & Ralphe, J. C. (2018). The HCM-linked W792R mutation in cardiac myosin-binding protein C reduces C6 FnIII domain stability. *American Journal of Physiology-Heart and Circulatory Physiology*, *314*(6), H1179–H1191. doi:10.1152/ajpheart.00686.2017
48. Papasaikas, P., & Valcárcel, J. (2016). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in Biochemical Sciences*, *41*(1), 33–45. doi:10.1016/j.tibs.2015.11.003
49. Rafael, J. F., Cruz Filho, F. E. dos S., Carvalho, A. C. C. de, Gottlieb, I., Cazelli, J. G., Siciliano, A. P., & Dias, G. M. (2017). Myosin-binding Protein C Compound Heterozygous Variant Effect on the Phenotypic Expression of Hypertrophic Cardiomyopathy. *Arquivos Brasileiros de Cardiologia*. doi:10.5935/abc.20170045
50. Kufareva, I., & Abagyan, R. (2012). Methods of protein structure comparison. *Methods in molecular biology (Clifton, N.J.)*, *857*, 231–57. doi:10.1007/978-1-61779-588-6_10
51. Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, *1*(6), 2876–2890. doi:10.1038/nprot.2006.202
52. Shen, C., Du, Y., Qiao, F., Kong, T., Yuan, L., Zhang, D., … Wu, Y.-D. (2018). Biophysical and structural characterization of the thermostable WD40 domain of a prokaryotic protein, Thermomonospora curvata PkwA. *Scientific Reports*, *8*(1), 12965. doi:10.1038/s41598-018-31140-y
53. Ellard, S., Baple, E. L., Owens, M., Eccles, D. M., Abbs, S., & Zandra, C. (2017). ACGS Best Practice Guidelines for Variant Classification 2017. *Association for Clinical Genetic Science*, 1–12.
54. Lee, Y., & Rio, D. C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual Review of Biochemistry*, *84*(1), 291–323. doi:10.1146/annurev-biochem-060614-034316
55. Hohfeld, J., Cyr, D. M., & Patterson, C. (2001). From the cradle to the grave: molecular chaperones that may choose between folding and degradation. *EMBO*

reports, *2*(10), 885–890. doi:10.1093/embo-reports/kve206
56. Harris, S. P., Lyons, R. G., & Bezold, K. L. (2011). In the thick of it: HCM-causing mutations in myosin binding proteins of the thick filament. *Circulation research*, *108*(6), 751–64. doi:10.1161/CIRCRESAHA.110.231670
57. Cohen, S., Kramarski, L., Levi, S., Deshe, N., Ben David, O., & Arbely, E. (2019). Nonsense mutation-dependent reinitiation of translation in mammalian cells. *Nucleic Acids Research*, *47*(12), 6330–6338. doi:10.1093/nar/gkz319
58. Zhang, J. (1997). Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. *The EMBO Journal*, *16*(4), 826–833. doi:10.1093/emboj/16.4.826
59. Ababou, A., Rostkova, E., Mistry, S., Masurier, C. Le, Gautel, M., & Pfuhl, M. (2008). Myosin Binding Protein C Positioned to Play a Key Role in Regulation of Muscle Contraction: Structure and Interactions of Domain C1. *Journal of Molecular Biology*, *384*(3), 615–630. doi:10.1016/j.jmb.2008.09.065
60. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)*, *112*, 531–52. doi:10.1385/1-59259-584-7:531