Am J Transl Res 2019;11(3):1724-1735
www.ajtr.org /ISSN:1943-8141/AJTR0078974

Original Article

A study of the pathogenicity of variants in familial heart disease. The value of cosegregation

Esperanza García-Molina[1,2], María Sabater-Molina[1,2], David López-Cuenca[3], María C Olmo[3], Inmaculada Pérez[1], Carmen Muñoz Esparza[2], Juan R Gimeno Blanes[3,4]

[1]Cardiogenetics Laboratory, Inherited Cardiac Disease Unit, IMIB University Hospital Virgen de la Arrixaca, Murcia (Spain); [2]Department of Genetic and Microbiology, Murcia University, Murcia, Spain; [3]Inherited Cardiac Disease Unit, Department of Cardiology, University Hospital Virgen de la Arrixaca, Murcia, Spain; [4]Department of Internal Medicine, Murcia University, Murcia, Spain

Received May 4, 2018; Accepted September 16, 2018; Epub March 15, 2019; Published March 30, 2019

Abstract: With the development of deep sequencing, a significant proportion of mutations already listed in studies have inconclusive pathogenicity. We aim to establish the proportion of cases in which familial studies are possible and cosegregation analysis is informative. We also compare cosegregation analysis with in silico software and a proposed pathogenicity score. 204 consecutive positive tests were reviewed. 4 different in silico software programs were used. Spaendonck-Zwarts' pathogenicity score was also calculated. A total of 73 of the missense variants could be classified by the score as being likely or definitively pathogenic. A high percentage of nonsense variants were found in desmosomal genes and missense variants in sarcomeric genes. 36.3% of the missense variants in our cohort classified as very likely or definitively pathogenic were novel. Cosegregation analysis was positive in 19.5% and could be discarded in 15.6%. There was a significant discrepancy between the in silico tools used in the setting of inherited heart disease. Multiparametric scoring systems which include cosegregation and functional studies seem to perform better than individual prediction software.

Keywords: Cosegregation, genetic variation, in silico tools, cardiomyopathy, channelopathy

Introduction

The development of new generation sequencing technologies and massive sequencing has led to the identification of new genes associated with cardiomyopathy [1]. The number of genetic variants has increased exponentially. In the majority of cases, up to two thirds of rarely (MAF<1%) detected variants are not included in reference databases (Ensembl, NCBI) [2]. Even in the case of already listed mutations, pathogenicity studies turn out to be inconclusive.

A significant proportion of mutations considered to be causally associated with hypertrophic cardiomyopathy (HCM) or dilated cardiomyopathy (DCM) identified in the Sanger era have been identified in exome sequencing projects in control populations of different ethnicities [3, 4].

Establishing pathogenicity in the case of hereditary heart disease has important implications for diagnosis, prognosis and counseling. High risk patients might need the implantation of a defibrillator for the prevention of malignant arrhythmia [5]. In the cases of carriers of mutations in ion-channel genes, even if asymptomatic in a normal ECG, medication is recommended (those with long QT syndrome (LQTS) and catecholaminergic ventricular tachycardia (CPVT) and they are encouraged to adopt lifestyle modifications (avoiding contraindicated medication in LQTS, CPVT, Brugada syndrome (BrS), and disqualification from competitive sports).

Despite genetic studies being recommended in documents published by experts and international clinical guidelines, the yield of the tests has geographical variability [6-8]. The use of deep sequencing or next generation sequencing (NGS) technologies in hereditary heart disease is currently a subject of debate [9].

Once technological issues have been addressed and the cost has been dramatically reduced, the main challenge remaining is to

establish the pathogenicity of new genetic variants [4, 10-16]. Functional analysis and familial studies are crucial to guarantee the correct identification of the pathogenicity of genetic variants.

We aim to establish the proportion of cases in which familial studies are possible and cosegregation analysis is conclusive. We also compare cosegregation analysis with in silico software and a proposed pathogenicity score.

Material and methods

Patient cohort

Patients included in this study were all evaluated at a specialist Inherited Cardiac Disease Unit at the Virgen de la Arrixaca University Hospital (Murcia, Spain) from 2003 to 2015. A total of 260 consecutive apparently unrelated index patients with a positive genetic result were included: 135 (46.4%) HCM, 42 (14.4%) DCM, 24 (8.2%) ARVC, 26 (8.9%) BrS, 20 (6.9%) LQTS, 7 (2.4%) SADS, 5 (1.7%) LVNC, 3 (1.0%) neuromuscular disease with cardiac involvement, and 29 (9.9%) other inherited heart diseases.

Clinical diagnosis was carried out according to published international clinical guidelines [17, 18]. The screening of relatives was carried out when a mutation was identified in an index case. The average number of affected relatives studied when families were available for screening was 3.7±3.8. This study was approved by the local ethics committee at our institution. All patients were informed about the study and provided signed consent for the genetic analysis.

Genetic analysis

Genomic DNA was extracted from blood samples using standard methods (Maxwell 16 blood DNA kit, Promega). Genetic analysis included selected genes associated with each type of heart disease. Variants were grouped into 4 categories, based on their activity and location: ion-channel genes (SCN5A, KCNQ1, KCNH2, KCNE1, KCNE2, RYR2, CACNA1B, CACNA1C, CACNA1D, CACNA1H and CACNAB2), desmosomal genes (PKP2, DSP, DSC, DSG, DES), sarcomeric genes (MYBPC3, MYH7, TNNT2, TNNI3, TPM1, TNNC1), cytoskeletal or Z-bands genes (ZASP, MYOT, TTN, DES, FLNC). Non-missense variations included: ins/del, frameshift, nonsense and intronic variations up to ±6 bp at the end of the exon. Non-missense variants were not included in the study of bioinformatics, cosegregation or function. These variants were considered as radical and causative mutations.

Sanger sequencing exon by exon was the method used in studies from 2003 to 2012. Afterwards, from 2012 to 2015, NGS technologies led to the use of larger panels (45, 29.2% of the tests (per patient), with a mean of 137.7±71.9 genes/NGS study).

First section of variant analysis: bioinformatic analysis

To evaluate the pathogenicity of missense variants we used a multiparametric scoring system based on the one published by van Spaendonck-Zwarts and her research group [19]. This scoring system includes 10 items in a first phase (in silico) and 2 items in a second phase (reassessment) which involves functional and cosegregation analysis.

The following in silico tools, more commonly used for missense variant interpretation in clinical laboratories, were applied: Polyphen v.22 [20], SIFT [21] and MutationTaster [22]. Additional tests carried out for bioinformatic analysis included Grantham distance, Align-GVGD and Blosum 62 matrix; and for the evaluation of evolutionary conservation level alignments between different isoforms and different species, the splicing module from Alamut visual v.2.7.0.0.software was used. The allele frequency in the Caucasian population of the control group was available (http://evs.gs.washington.edu/EVS/). A comprehensive literature and database search was preformed (Ensembl, HGMD, the Exome Variant Server and a functional analysis). A detailed explanation of the scoring system used is provided in the supplementary material (Supplementary Table 1).

In the first phase of the analysis, a score was assigned to each one of the parameters, and variants were classified into 4 different categories. The van Spaendonck-Zwarts' score ranged

from 0 to 13.75, and as the score increased, the likelihood of the variant being pathogenic also increased. Variants scoring <25% were considered as benign while those between 25-45% were probably benign, Variant of Uncertain Signficance-1 (VUS1). A score of between 45-70% indicated a variant of Uncertain Clinical Significance (VUS2), and a score of >70% mean it was likely to be pathogenic (VUS3).

A study carried out by Jordan and his research group describes a computational method for the assessment of missense variants in HCM. This predictor uses phylogenetic and structural features specific to genes involved in HCM. Based on its predictions, we have analyzed and compared the final scores of our sarcomeric variants (N=49) with the results using Jordan's scoring system.

We analyzed all 154 missense variants with the in silico software. However in some variants it could not be carried out by some of the software. One hundred and eight variants had a prediction result from all the programs which were used to correlate the quality parameters.

Second section of variant analysis: cosegregation and functional study

The final result of the prediction analysis was calculated including the functional analysis (when it was available in the literature search) and the familial cosegregation analysis (either from the literature or from our own series).

In the second phase of the analysis, and following Van Spaendonck-Zwarts' scoring system, variants were reclassified into 5 categories: definitively benign, VUS1, VUS2, VUS3 and definitively pathogenic.

Cosegregation analysis

In this section we analyzed the cosegregation value in families with informative pedigrees: ones having at least 3 affected relatives. In addition, in families with 2 affected members, with one of them being a non-carrier, the variant was considered as non-disease-causing.

Statistical analysis

Statistical analysis was performed using SPSS (version 15.0) statistical software (SPSS Inc., Chicago, IL, USA). We used the analysis of software described by group of Thusberg [23]. The quality of the predictions is described using 6 parameters: accuracy, precision, sensitivity, specificity, negative predictive value (NPV) and the Matthews correlation coefficient (MCC). The MCC [24] statistic is unaffected by the differing proportion of neutral and pathogenic datasets predicted by the different programs. The MCC is in essence a correlation coefficient between the observed and predicted binary classifications returning a value between -1 and +1. A coefficient of +1 represents a perfect prediction, 0 no better than a random prediction, and -1 indicates total disagreement between the prediction and observation. Because of its insensitivity to differing test set sizes, it gives a more balanced assessment of performance than the other performance measures [25]. Correlations between the program outputs were calculated by counting all of the common cases and those predicted correctly, and using Spearman's rank correlation coefficient.

Results

Genetic study type of variant

Two hundred and sixty consecutive index patients with inherited heart disease who had a positive genetic result were included. A total of 204 variants were detected in 42 different genes, 102 (66.2%) of these were novel variants. The following types of genetic variants were detected: 154 (75.5%) missense variants and 50 (24.5%) radical variations (24 premature stop codons, 12 intronic variants likely to alter splicing, 9 ins/del variants and 5 frameshift mutations). Cascade screening in relatives led to the identification of a total of 554 gene carriers. In 84 missense variants there were 2 or more carriers.

Distribution of missense variants per group of genes was: 50 (32.5%) variants in sarcomeric genes, 18 (11.7%) in desmosomal genes, 45 (29.2%) variants in ion-channels and 15 (9.7%) in Z-band genes. Eleven premature stop codons were identified in desmosomal genes, which accounted for 45.8% of the truncation variants detected (followed by 6 in Z-band, 5 in sarcomeric genes and 3 in ion-channel genes). Two (40.0%) of the 5 frameshift variants were identified in sarcomeric and desmosomal genes respectively while in Z-band and ion-channel

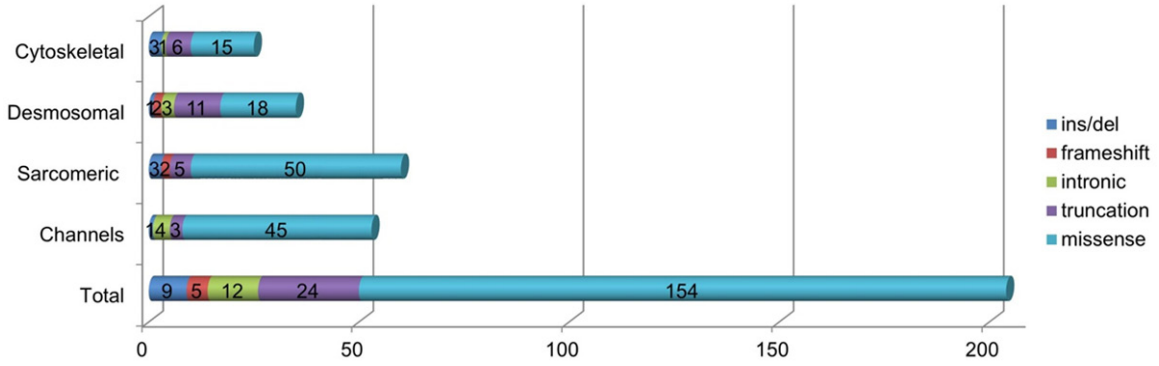

Figure 1. Number of each type of variant (insertion/deletion, frameshift, intronic, truncation and missense) and distributed by group of genes.

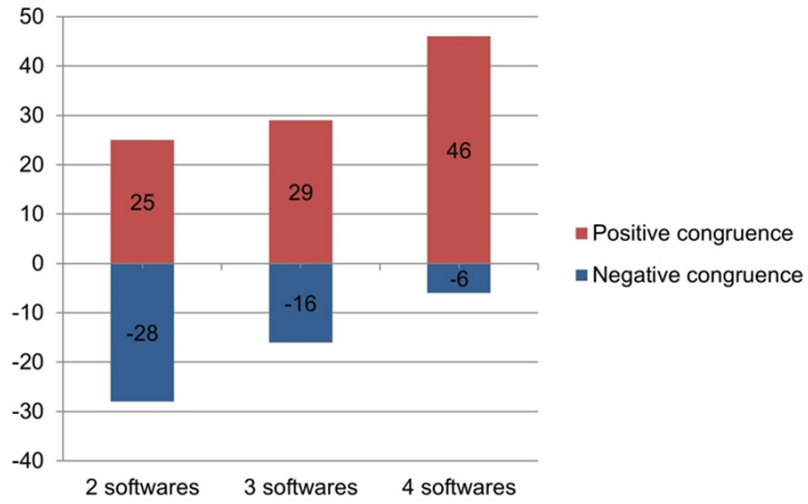

Figure 2. Congruence between 2. 3 or 4 types of software. Congruence is consider positive when different software agree that a variant is pathogenic and negative when the results of the software agree that the variant is non-pathogenic.

genes we failed to find any frameshift variant. We found 12 intronic variants; 4 (33.3%) in ion-channels, 3 (25.0%) in desmosomal genes and 1 (8.3%) in cytoskeletal genes.

When we analyzed each group of genes by type of variant, we found a high percentage of missense variants, particularly in both ion-channel genes (84.9% of which corresponded to missense variants), and sarcomeric genes; missense variants were detected in 83.3%. However, in cytoskeletal and desmosomal genes the percentage was lower, comprising 60.0% and 51.4% respectively. It is noteworthy that truncation variants accounted for 31.4% of those found in desmosomal genes, 24.0% in Z-band genes, 8.3% in sarcomeric genes and 5.6% in ion channel genes (**Figure 1**). Intronic variants accounted for 7.5% of those found in ion-channel genes, 8.6% in desmosomal genes, 4.0% in cytoskeletal genes and 0.0% in sarcomeric genes. Ins/del accounted for 12.0% of those found in cytoskeletal genes, 5.0% in sarcomeric genes and 1.9% in ion-channel genes and 2.8% in desmosomal genes. Frameshift accounted for 5.7% in desmosomal genes, 3.3% of those found in sarcomeric genes, and 0.0% in cytoskeletal genes and ion-channels. Radical variants in desmosomal genes accounted for 48.6% variants, followed by 40.0% in cytoskeletal genes, 16.6% in sarcomeric genes and 15.1% in ion-channel.

Pathogenicity analysis of missense variants

In silico prediction: MutationTaster was the software with the highest proportion of likely pathogenic mutations (120, 78.9%), followed by Pmut (96, 77.4%), Polyphenv.2 (85, 65.9%), and Sift (81, 55.9%). All 4 types of software had a similar positive result in 46 (27.3%) of the variants (positive congruence), and they had a similar negative result in 6 (3.9%) (**Figure 2**). A correlational study was conducted to determine the concordance between the different software used for the prediction of pathogenic-

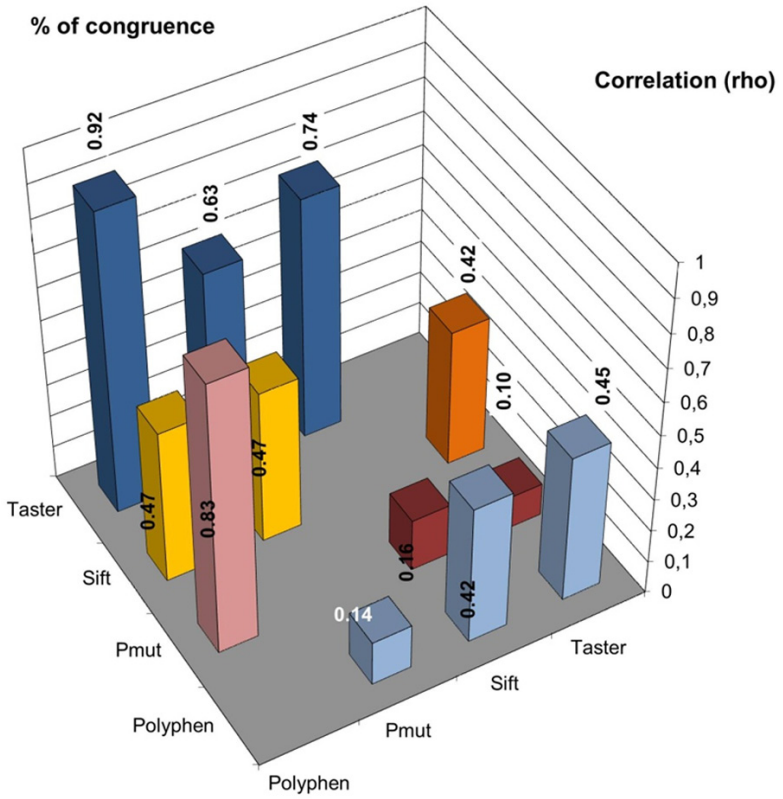

Figure 3. On the left the columns represent the percentage of congruence between the kinds of software. The right columns show the correlation between software calculated as the Spearman correlation coefficient.

ity present in **Figure 3**. MutationTaster software presented a ρ(rho) value of 0.453 using Polyphen v.2 software, interpreted as having moderately positive agreement. Polyphen v.2 with SIFT showed moderate agreement too: 0.424. The prediction program Pmut obtained the least concordance with the rest of the software: with SIFT it was ρ=0.160, with Polyphen v.2, ρ=0.136 and with MutationTaster, ρ=0.101, all interpreted as having a ρ value with poor agreement.

The performance of sensitivity and specificity of the software programs were compared with the considered "gold standard", cosegregation, in variants with an available prediction for the 4 kinds of software (**Table 1**). MutationTaster had the highest sensitivity (82.0%) but the worst negative predictive value (NPV) (22.0%). Polyphen v.2 showed a sensitivity of 72.0% and a 41% NPV, Pmut had 74% sensitivity and a 41% NPV and SIFT showed a sensitivity of 53.0% and an NPV of 25%.

We compared the concordance between genes with a structural (N=62) and ionic transporter function (N=38) which needed to have predictions from all 4 types of software. All 4 kinds of prediction software matched in 55.2% of the ionic transporter variants and 3 matched in 42.1% (97.3%, 3 or 4). The congruence of the different kinds of software was lower for variants in structural genes, and 4 types of software gave similar results in 33.9% and 3 matched in 32.3% (in 66.2% 3 or 4 had similar results).

Phase I Spaendonck-Zwarts' scoring system

All missense variants were analyzed and scored using this algorithm. Fifty-eight (37.7%) variants were classified as very likely to be pathogenic (VUS3) in the first phase of in silico prediction. Sixty (38.9%) were classified as VUS2 and 35 (22.7%) as VUS1. Only 1 (0.6%) variant could be classified as clearly benign.

The Jordan prediction system for sarcomeric variants

We applied the Jordan prediction system to sarcomeric variants; 49 (31.8%) of the 154 variants could be predicted with this score. Fifteen (9.7%) were not predicted, with readings such as "no call", 10 (6.5%) variants were predicted to be benign, and 11 (7.1%) as pathogenic.

Phase II Spaendonck-Zwarts' scoring

In the second phase we introduced cosegregation and functional study information to carry out the final prediction of pathogenicity.

Cosegregation obtained a score of 2 (possible cosegregation) for 11 variants, 3 (probable cosegregation) for 3 variants and 4 (very likely pathogenic) for 10 variants. Only 24 (15.6%)

Table 1. Performance of prediction methods

Performance	*Taster*	*Polyphen*	*Pmut*	*Sift*
tp	31	28	28	20
tn	2	7	7	6
fp	14	11	11	12
fn	7	11	10	18
Cases +	84	70	82	64
Cases -	24	38	26	44
Accuracy = TP+TN/TP+TN+FP+FN	0.61	0.61	0.63	0.46
Precision = TP/TP+FP	0.69	0.72	0.72	0.63
Specificity = TN/TN+FP	0.13	0.39	0.39	0.33
Sensitivity = TP/TP+FN	0.82	0.72	0.74	0.53
NPV = TN/TN+FN	0.22	0.39	0.41	0.25
MCC = (TPxTN)-(FNxFP)/$\sqrt{}$(TP+FN)(TP+FP)(TN+FP)	-0.07	0.11	0.13	-0.13

TN: true negative; TP: true positive; FN: false negative; FP: false positive; NPV: Negative predicted value; MCC: Matthews correlation coefficient.

could be classified using this cosegregation score.

A functional analysis was available in the literature in only 12 (7.8%) variants, being convincingly deleterious in 6 (50.0%) cases, likely to be aberrant in 4 (33.3%), and inconclusive or not functionally aberrant in 2 (16.6%).

In the final prediction 16 variants (10.4%) were classified as definitively causative, 8 (57.1%) in sarcomeric genes, 3 (21.4%) in the ion-channels and 1 in cytoskeletal genes. Fifty-seven (37.0%) variants were classified as very likely to be causative (VUS3); 23 (40.3%) in sarcomeric genes, 16 (28.1%) in ion-channel genes, 6 (10.5%) in desmosomal genes and 4 (7.0%) in cytoskeletal genes. Fifty-one variants were classified as VUS2, 31 as VUS1, and only 1 as clearly not pathogenic. Eleven of the definitively causative variants were described in the literature but only 3 (21.4%) appeared in the ClinVar public archive (http://www.ncbi.nlm.nih.gov/clinvar/) classified as pathogenic. Application of the score with the data presented in this study led to the reassessment of the pathogenic variant of 4 variants with conflicting interpretations of pathogenicity.

When we analyzed the yield of genetic testing in the different groups of genes, shown in **Figure 4**, in the sarcomeric group (N=50), 23 (46.0%) of the variants were classified as VUS3 and 8 (16.0%) as definitively pathogenic. In the ion-channel group (N=45) 16 (35.5%) variants were shown to be VUS3 and 3 (6.6%) were definitively pathogenic. In the desmosomal group (N=18), 6 (33.3%) variants were classified as VUS3 and none as definitively pathogenic. Finally, in the cytoskeletal group (N=15), 4 (26.6%) variants were classified as VUS3 and 1 (6.6%) as being definitively pathogenic.

An analysis of phases I and II of the Spaendonck-Zwarts' scoring system is shown in **Figure 4**.

Cosegregation analysis

Cosegregation analysis was possible in 44 (28.6%) variants with informative pedigrees (with at least 3 affected relatives with available DNA samples). Cosegregation was probable (positive) in 31 (19.5%) and could be ruled out in 14 (9.1%) of the variants in which a wildtype affected relative was found. Thirty-three variants were present in families with 2 affected relatives. In these cases the cosegregation analysis was inconclusive for 3 variant while 10 variants were classified as definitively negative and 20 as possible. In total, positive or negative cosegregation was possible in 54 (35.1%) variants (**Figure 5**).

The variants with at least 2 affected relatives had an average of 4.3±4.4 affected wildtype subjects and 3.7±3.8 affected carriers per variant.

Recurrent variants

Fifteen of the 154 variants were present in at least 2 families. These variants were described

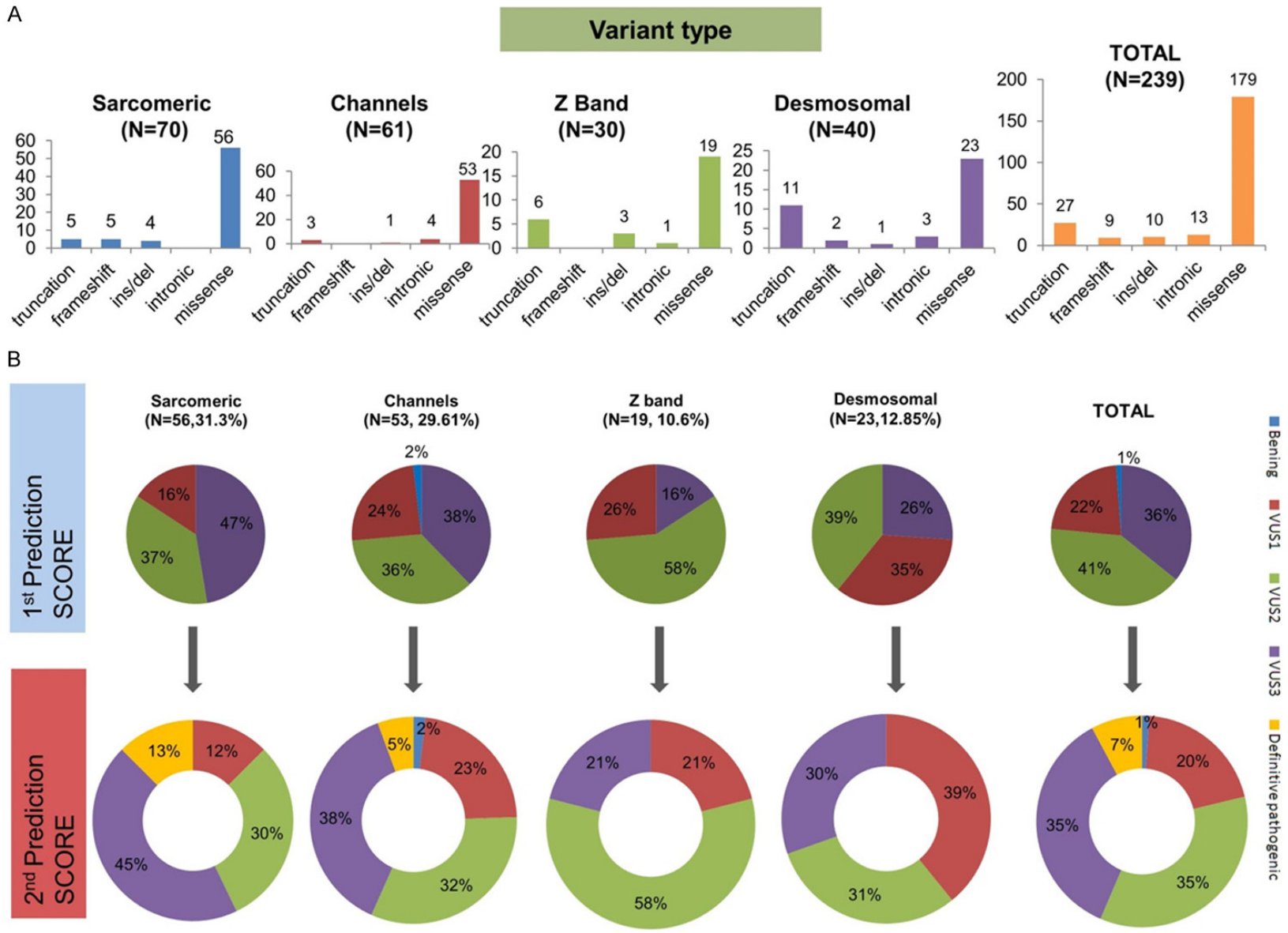

Figure 4. A. Number of variants detected (N) according to the type of variant grouped by cluster of genes. B. Study of prediction of pathogenicity in 2 phases. The first score, it is not considered the cosegregation study. The second score, the cosegregation study was included, and a reclassification of the variants studied is showed. The variants are classified as benign, VUS1, VUS2 and VUS3. VUS (variant of uncertain significance).

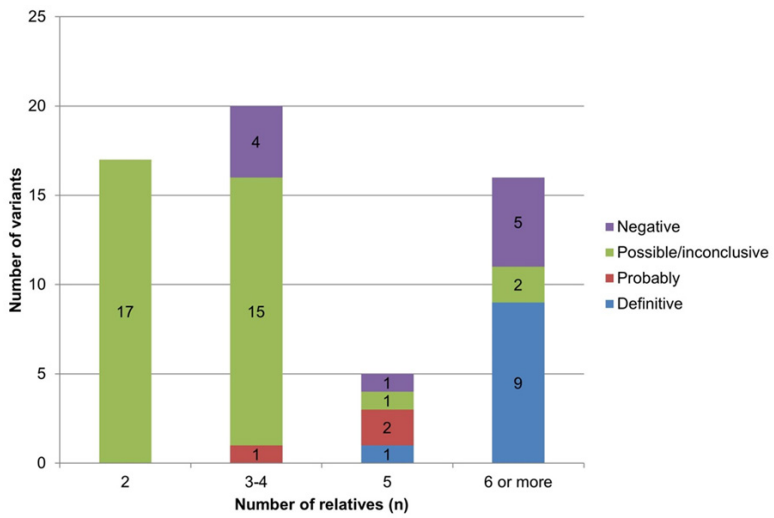

Figure 5. Cosegregation value according to the number of affected relatives per variant. In the X axis the number of relatives carrying a variant is showed.

in the literature but only 11 were registered in ClinVar: 9 had conflicting interpretations of pathogenicity, 1 had associated clinical significance which was benign and the other was pathogenic. Twelve variants were in sarcomeric genes (5 MYBPC3, 1 MYH6, 4 MYH7, 1 TNNT2), 2 in SCN5A, 1 in a DSP gene and 1 in a ZASP. Five variants were upgraded one level after cosegregation and a functional analysis, and 3 variants were upgraded two levels. In total, 6 variants were reassessed as definitively causing variations. Described recurrent variations are listed in **Table 2**.

Discussion

Geneticists and cardiologists working in inherited heart disease units worldwide would agree that the challenge of the integration of genetic information into medical practice is necessary for the correct evaluation of the pathogenicity of detected variants.

It is notable that half of the missense variants detected were already reported in the literature (53.1%). The proportion of novel missense variants in our cohort classified with a score indicating a very likely or definitively pathogenic variant was 34.2%. This highlights the importance of causality studies in order to appropriately classify novel variants and to reassess the pathogenicity of old variants.

In keeping with other series, there was a high percentage of non-sense mutations in desmosomal genes and missense mutations in sarcomeric genes [26, 27]. The percentage of missense mutations in MYBPC3 (66.6%), the genes most frequently associated with HCM, was higher in our series than in the one reported by Jordan (35%) [11].

With the help of the proposed Spaendonck-Zwarts score, a total of 73 (47.4%) of the missense variants identified in our inherited heart disease unit could be classified as very likely or definitively causative. This percentage was higher in sarcomeric genes (62.0%) and lower in cytoskeletal genes (33.2%).

The value of cosegregation based on the study of informative pedigrees was measured to be 35.1% in our cohort. Cosegregation analysis was possible in a significant proportion of families. This is the first time cosegregation analysis has been evaluated in this context of the study of inherited heart diseases.

Another novelty of the present study is the analysis of the value of different types of commonly-used in silico software and the correlation between them. In silico software only gave consistent results in a minority of missense variants evaluated by all 4 software programs (31.2%).

The majority of prediction software available have a low level of accuracy of between 65-80% [23]. Most tools tend to have a low specificity, resulting in the over-prediction of missense changes as being deleterious, and they are not as reliable at predicting missense variants with a mild effect [28]. Unfortunately, current in silico classification schemes for predicting the pathogenicity of missense variants have a low predictive power for classifying cardiomyopathy variants [26]. In our study the software with the best prediction results for assessing the clinical significance of variants was MutationTaster.

Due to the fact that genetic information can influence clinical decision making, the implementation of objective multiparametric methods (integrating in silico software, functional analysis and cosegregation analysis) for the

Table 2. Described recurrent variants found in at least 2 different families

Gene	Variants	Nº families	Affected	Affected WT	Affected carriers	Associated pathology	Affected age	Sub-classification pathogenicity	ClinVar Interpretation	Reassesment causality
MYBPC3										
	p.E1179K	3	4	0	4	HCM	35.3	VUS2	CI	VUS3
	p.E441K	2	5	3	1	HCM/BSr	46.4	VUS3	CI	VUS3
	p.R326Q	3	5	2	3	DCM/HCM/MCE	62.5	VUS2	BPB	VUS2
	p.E258K	5	13	0	12	HCM	40.9	VUS3	Pathogenic	DC
MYH6										
	p.A1004S	2	5	0	5	HCM	63.0	VUS2	CI	DC
MYH7										
	p.R1382Q	3	4	0	3	HCM	33.5	VUS3	----	DC
	p.D928N	3	7	0	7	HCM	34.5	VUS3	----	DC
	p.A355S	4	7	1	6	HCM	53.3	VUS2	----	VUS2
	p.T1377M	5	17	1	15	HCM	39.8	VUS3	CI	VUS3
TNNT2										
	p.K253R	2	3	0	0	HCM	40.0	VUS1	Benign	VUS1
	p.R278C	2	2	0	1	HCM	44.0	VUS3	CI	VUS3
SCN5A										
	p.G1743R	4	9	0	8	BSr	30.8	VUS3	----	DC
	p.D772N	2	2	0	2	BSr	49.0	VUS2	CI	VUS2
DSP										
	p.D2070N	2	2	1	1	BSr/QTL	41.0	VUS2	CI	VUS2
ZASP										
	*p.D117N	2	27	6	19	DCM	35.3	VUS2	CI	DC
RBM20										
	*p.R636H	1	24	3	21	DCM		VUS2	Pathogenic	DC

*These variants were detected in a big family and we confirmed their pathogenicity. Conflicting interpretations: CI. Benign/probably benign: BPB. Definitively causing: DC.

appropriate classification of new variants is of paramount importance. No single in silico software program seems to perform well enough to be used in isolation.

We have demonstrated significant discrepancies between the in silico tools used in the setting of inherited heart disease (45-65%, **Figure 3**). MutationTaster with Polyphen v.2 showed a high level of congruence (92%). Polyphen v.2 turned out to be the one with the best agreement among the software used, and is the most consistent with the results obtained using cosegregation. The software with most false negative results was SIFT. MutationTaster was the software with the highest sensitivity. This information is important for the interpretation of reports from genetic laboratories. Furthermore, algorithms can have vastly different predictive capabilities for different genes [29]. In this regard, we have demonstrated significantly higher percentages of agreement in ion-channel genes compared to structural genes (97% vs. 65%, for 3 or 4 congruent results).

The yield of genetic testing per condition published in other series differs from our results. In the cohort of Hofman, the scoring system proposed by Spaendock-Zwarts et al. was used to detect a possible disease-causing variant (pathogenic mutation, VUS3 or VUS2) in 702 families (31%), a percentage that is lower than in ours. We found 73 (47.4%) variants which could possibly be disease-causing: an average of 62% in HCM, 42% in several channelopathies (LQTS, BrS and CPVT) and 22% in dilated cardiomyopathy [30]. In the Spaendonck-Zwarts series, they found a mutation in 82 out of 418 index patients with idiopathic dilated cardiomyopathy (20%) [19]. Regarding the Jordan method for sarcomeric variants [11], 81.8% matched with our score compared to 40% of the benign variants. The rest of the analyzed variants (N=28) were “uncovered”

(variants not predicted) by the Jordan scoring system, compared with our prediction; 4 variants could be classified as definitively pathogenic, 12 as VUS3, 11 as VUS2 and 1 as VUS1. Therefore, the yield of our prediction method is better in terms of sarcomeric variants: 32.6% (N=16) of the variants (48 families), would have not received the familial study and closer monitoring if the cosegregation analysis would not have been used in the present study.

Acknowledgements

This study has been partly funded by the fundacio La Marató de TV3 (project: 218/C/2015), a national grant provided by the FIS (PI11/02459) and the Cardiovascular Research Network (RIC; RD12/0042/0049) from the Carlos III Health Institute-Unión Europea, Fondo Europeo de Desarrollo Regional, <<Una manera de hacer Europa>>. We want to particularly acknowledge the collaboration of our patients and the BioBank "Biobanco en Red de la Región de Murcia" (PT13/0010/0018) which is part of the Spanish National Biobanks Network (B.000-859). Institutional affiliations: the researchers are part of a cardiovascular research network (RIC; RD12/0042/0049,29) and the Murcian Institute of biosanitary research (IMIB) both of them from the Carlos III Health Institute. M. Sabater-Molina, E. García-Molina and J. Gimeno also work at University of Murcia. J. Gimeno and M. Sabater-Molina are part of the CIBERCV (CB16/11/00385).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. María Sabater-Molina, Cardiogenetics Laboratory, Inherited Cardiac Disease Unit, IMIB-Hospital Clínico Universitario Virgen de la Arrixaca, LAIB Campus de Ciencias de la Salud, Avda, Buenavista s/n, El Palmar, 30120 Murcia, Spain. Tel: +34-868885084; E-mail: mariasm79es@hotmail.com; mariasm@um.es

References

[1] Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, Camm AJ, Ellinor PT, Gollob M, Hamilton R, Hershberger RE, Judge DP, Le Marec H, McKenna WJ, Schulze-Bahr E, Semsarian C, Towbin JA, Watkins H, Wilde A, Wolpert C, Zipes DP. Heart Rhythm Society (HRS); European Heart Rhythm Association (EHRA). HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies. Europace 2011; 13: 1077-1109.

[2] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010; 7: 248-249.

[3] Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, Sajadieh A, Haunsø S, Svendsen JH, Olesen MS. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. Eur J Hum Genet 2013; 21: 918-928.

[4] Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 2000; 16: 412-424.

[5] Cerrone M, Priori SC. Genetics of sudden death: focus on inherited channelopathies. Eur Heart J 2011; 32: 2109-2120.

[6] Charron P, Arad M, Arbustini E, Basso C, Bilinska Z, Elliott P, Helio T, Keren A, McKenna WJ, Monserrat L, Pankuweit S, Perrot A, Rapezzi C, Ristic A, Seggewiss H, van Langen I, Tavazzi L; European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Genetic counselling and testing in cardiomyopathies: a position statement of the European society of cardiology working group on myocardial and pericardial diseases. Eur Heart J 2010; 31: 2715-2728.

[7] Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One 2012; 7: e46688.

[8] Ingles J, Bagnall RD, Semsarian C. Determining pathogenicity of genetic variants in hypertrophic cardiomyopathy: importance of periodic reassessment. Genet Med 2014; 16: 286-293.

[9] 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 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 2014; 35: 2733-2779.

[10] Hofman N, Tan HL, Alders M, Kolder I, de Haij S, Mannens MM, Lombardi MP, Dit Deprez RH, van Langen I, Wilde AA. Yield of molecular and

clinical testing for arrhythmia syndromes: report of 15 years' experience. Circulation 2013; 128: 1513-1521.
[11] Jordan DM, Kiezun A, Baxter SM, Agarwala V, Green RC, Murray MF, Pugh T, Lebo MS, Rehm HL, Funke BH, Sunyaev SR. Development and validation of a computational method for assessment of missense variants in hypertrophic cardiomyopathy. Am J Hum Genet 2011; 88: 183-192.
[12] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009; 4: 1073-1081.
[13] Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. Mutation taster evaluates disease-causing potential of sequence alterations. Nat Methods 2010; 7: 575-576.
[14] Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim Biophys Acta 1975; 405: 442-451.
[15] Li Mura IE, Bauce B, Nava A, Fanciulli M, Vazza G, Mazzotti E, Rigato I, De Bortoli M, Beffagna G, Lorenzon A, Calore M, Dazzo E, Nobile C, Mostacciuolo ML, Corrado D, Basso C, Daliento L, Thiene G, Rampazzo A. Identification of a PKP2 gene deletion in a family with arrhythmogenic right ventricular cardiomyopathy. Eur J Hum Genet 2013; 21: 1226-1231.
[16] Lindor NM, Guidugli L, Wang X, Vallée MP, Monteiro AN, Tavtigian S, Goldgar DE, Couch FJ. A review of a multifactorial probability based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). Hum Mutat 2012; 33: 8-21.
[17] Lopes LR, Zekavati A, Syrris P, Hubank M, Giambartolomei C, Dalageorgou C, Jenkins S, McKenna W; Uk10k Consortium, Plagnol V, Elliott PM. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. J Med Genet 2013; 50: 228-239.
[18] Mogensen J, van Tintelen JP, Fokstuen S, Elliott P, van Langen IM, Meder B, Richard P, Syrris P, Caforio AL, Adler Y, Anastasakis A, Gimeno JR, Klingel K, Linhart A, Imazio M, Pinto Y, Newbery R, Schmidtke J, Charron P. The current role of next-generation DNA sequencing in routine care of patients with hereditary cardiovascular conditions: a viewpoint paper of the European society of cardiology working group on myocardial and pericardial diseases and members of the European society of human genetics. Eur Heart J 2015; 36: 1367-1370.
[19] Norton N, Robertson PD, Rieder MJ, Züchner S, Rampersaud E, Martin E, Li D, Nickerson DA, Hershberger RE; National Heart, Lung and Blood Institute GO Exome Sequencing Project. Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. Circ Cardiovasc Genet 2012; 5: 167-174.
[20] Pan S, Caleshu CA, Dunn KE, Foti MJ, Moran MK, Soyinka O, Ashley EA. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. Circ Cardiovasc Genet 2012; 5: 602-610.
[21] Phelan D, Wilson GR, James PA, Lockhart PJ. The genetics of cardiomyopathy, new technologies and the path to personalised medicine. OA Genetics 2013; 1: 9-14.
[22] Priori SG, Wilde AA, Horie M, Cho Y, Behr ER, Berul C, Blom N, Brugada J, Chiang CE, Huikuri H, Kannankeril P, Krahn A, Leenhardt A, Moss A, Schwartz PJ, Shimizu W, Tomaselli G, Tracy C. HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of patients with inherited primary arrhythmia syndromes: document endorsed by HRS, EHRA, and APHRS in May 2013 and by ACCF, AHA, PACES, and AEPC in June 2013. Heart Rhythm 2013; 10: 1932-1963.
[23] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. 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: 405-424.
[24] Sabater-Molina M, García-Molina E, Tovar I, Ruiz-Espejo F, Gimeno JR, Valdés M. Cost-effectiveness of genetic studies in inherited heart diseases. Cardiogenetics 2013; 3: 28-30.
[25] Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB; IARC Unclassified Genetic Variants Working Group. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 2008; 29: 1327-1336.
[26] Teekakirikul P, Kelly MA, Rehm HL, Lakdawala NK, Funke BH. Inherited cardiomyopathies. Molecular genetics and clinical genetic testing in the postgenomic era. J Mol Diagn 2013; 15: 158-170.
[27] Thompson BA, Greenblatt MS, Vallee MP, Herkert JC, Tessereau C, Young EL, Adzhubey IA, Li B, Bell R, Feng B, Mooney SD, Radivojac P, Sunyaev SR, Frebourg T, Hofstra RM, Sijmons RH, Boucher K, Thomas A, Goldgar DE, Spurdle AB, Tavtigian SV. Calibration of multiple in silico tools for predicting pathogenicity of mismatch

repair gene missense substitutions. Hum Mutat 2013; 34: 255-265.

[28] Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. Hum Mutat 2011; 32: 358-368.

[29] van Spaendonck-Zwarts KY, van Rijsingen IA, van den Berg MP, Lekanne Deprez RH, Post JG, van Mil AM, Asselbergs FW, Christiaans I, van Langen IM, Wilde AA, de Boer RA, Jongbloed JD, Pinto YM, van Tintelen JP. Genetic analysis in 418 index patients with idiopathic dilated cardiomyopathy: overview of 10 years' experience. Eur J Heart Fail 2013; 15: 628-636.

[30] Wallis Y, Payne S, McAnulty C, Bodmer D, Sistermans E, Robertson K, Moore D, Abbs S, Deans Z, Devereau A. Practice guidelines for the evaluation of pathogenicity and the reporting of sequence variants in clinical molecular medicine. Birmingham, UK: Association for Clinical Genetic Science; 2013.

Supplementary Table 1. Pathogenicity score

1. Prediction pathogenicity softwares	2. Evolutive conservation and amino acids differences	3. Conservation	4. Data bases and Literature (HGMD, TGP, Ensembl, EVS, ClinVar)	5. In silico analysis of splicing
1.1. SIFT.....score Intolerant.....1 Tolerant.....0	2.1. Grantham distance.....score Large distances (>140).....2 Moderate distance (≤140).....1 Benign (≤70).....0	3.1. Conservation between isoforms.....score 75%-100%.....0.5 35%-74%.....0.25 0%-34%.....0	4.1. Minor allele Frecuency (%).....score No present.....2 0< frec ≤0.02.....1.5 0.02< frec ≤0.05.....1 0.05< frec ≤0.01.....0.5 >0.1.....0 >40 control allele .....no pathogenic	Analysis of splicing.....score Alteration of splicing..... 2 Very likely.....1.5 Probably.....1 Possibly.....0.5 Not likely.....0
1.2. Polyphen v.2.....score Disease causing.....1 Polymorphism.....0	2.2. Aling-GVGD.....score Clase 65.....1.25 Clase 55.....1.00 Clase 45.....0.75 Clase35 .....0.50 Clase 15/25.....0.25 Clase C0.....0.00 2.3. Blosum 62.....score <-2.....1 -1.....0.5 ≥0.....0	3.2. Conservation between species.....score All mammals and almost all lower animals.....1 All mammals and a few lower animals.....0.75 Almost all mammals and no lower animals.....0.5 Other.....0	4.2. Own laboratory and other sources...score Single proband.....0 >1 (absent in local controls).....1 Reported literature in affected.....2	

SCORE sub-classification.....(0-13.75)

% SCORE
<25%.....not pathogenic
25%≤ score <45%.....VUS 1
45%≤ score <70%.....VUS 2
>70%.....VUS 3

6. Label	Premise	Score	7. Functional analysis	Score
Very likely pathogenic	De novo mutation or ≥6 affected family members with the mutation and no affected without the mutation	4	Functionally aberrant	3
Probably cosegregation	5 affected family members with the mutation and no affected without the mutation	3	Possibly functionally aberrant	1
Possible cosegregation	3-4 affected family members with the mutation and no affected without the mutation	2	Unclear/not functionally aberrant	0
Cosegregation unclear	2 affected family members with the mutation and no affected without the mutation	1		
Only index/no cosegregation	affected family members without the mutation	0		

Study in silico.