Gene 577 (2016) 227–235

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Molecular analysis of sarcomeric and non-sarcomeric genes in patients with hypertrophic cardiomyopathy

Irene Bottillo [a,*], Daniela D'Angelantonio [a], Viviana Caputo [b], Alessandro Paiardini [c], Martina Lipari [a], Carmelilia De Bernardo [a], Diana Giannarelli [d], Antonio Pizzuti [b], Silvia Majore [a], Marco Castori [a], Elisabetta Zachara [e], Federica Re [e], Paola Grammatico [a]

[a] Medical Genetics, Department of Molecular Medicine, Sapienza University, San Camillo-Forlanini Hospital, Circonvallazione Gianicolense, 87-00152 Rome, Italy
[b] Department of Experimental Medicine, Sapienza University of Rome, Rome, Italy
[c] Department of Biochemical Sciences, Sapienza University of Rome, Rome, Italy
[d] Biostatistic Unit, Regina Elena National Cancer Institute, Rome, Italy
[e] Cardiomyopathies Unit, Division of Cardiology and Cardiac Arrhythmias, San Camillo-Forlanini Hospital, Rome, Italy

ARTICLE INFO

Article history:
Received 29 July 2015
Received in revised form 20 November 2015
Accepted 29 November 2015
Available online 2 December 2015

Keywords:
Hypertrophic cardiomyopathy
HCM
Next generations sequencing
NGS
Genetic testing
Sarcomere

ABSTRACT

Background: Hypertrophic cardiomyopathy (HCM) is a common genetic heart disorder characterized by unexplained left ventricle hypertrophy associated with non-dilated ventricular chambers. Several genes encoding heart sarcomeric proteins have been associated to HCM, but a small proportion of HCM patients harbor alterations in other non-sarcomeric *loci*. The variable expression of HCM seems influenced by genetic modifier factors and new sequencing technologies are redefining the understanding of genotype–phenotype relationships, even if the interpretations of the numerous identified variants pose several challenges.
Methods and results: We investigated 62 sarcomeric and non-sarcomeric genes in 41 HCM cases and in 3 HCM-related disorders patients. We employed an integrated approach that combines multiple tools for the prediction, annotation and visualization of functional variants. Genotype–phenotype correlations were carried out for inspecting the involvement of each gene in age onset and clinical variability of HCM. The 80% of the non-syndromic patients showed at least one rare non-synonymous variant (nsSNV) and among them, 58% carried alterations in sarcomeric *loci*, 14% in desmosomal and 7% in other non-sarcomeric ones without any sarcomere change. Statistical analyses revealed an inverse correlation between the number of nsSNVs and age at onset, and a relationship between the clinical variability and number and type of variants.
Conclusions: Our results extend the mutational spectrum of HCM and contribute in defining the molecular pathogenesis and inheritance pattern(s) of this condition. Besides, we delineate a specific procedure for the identification of the most likely pathogenetic variants for a next generation sequencing approach embodied in a clinical context.

1. Introduction

Hypertrophic cardiomyopathy (HCM) is a genetic heart disorder with a 1/500 prevalence (Maron, 2002). It is a primary genetic cardiomyopathy (Maron et al., 2006) characterized by a hypertrophic and non-dilated left ventricle in the absence of another systemic or cardiac process that could reasonably account for the magnitude of wall thickening. Heart muscle changes affect the electrical stability of the myocardial cells, predisposing to heart failure and/or arrhythmias (Dische, 1972). The HCM phenotype is highly variable, ranging from lifelong absence of symptoms to rapidly progressive heart failure or sudden cardiac death (Watkins et al., 1995). Two thirds of patients display the left ventricular outflow tract obstruction strongly associated with progression to severe symptoms of heart failure and of death (Maron et al., 2003). Since the discovery of the first HCM-causing gene (cardiac muscle β-myosin heavy chain, *MYH7*) (Jarcho et al., 1989), a large number of mutations in other *loci* mostly encoding sarcomeric proteins (i.e. *ACTC1*, *FLNC*, *MYBPC3*, *MYH6*, *MYH7*, *MYL2*, *MYL3*, *MYOM1*, *NEBL*, *TNNC1*, *TNNI3*, *TNNT2*, *TPM1*) have been shown to cause non-syndromic HCM (Marian et al., 2001; Arad et al., 2002; Hershberger et al., 2013; Valdes-Mas et al., 2014). Sarcomeric alterations

Abbreviations: HCM, hypertrophic cardiomyopathy; DCM, dilatative cardiomyopathy; RCM, restrictive cardiomyopathy; ARVC, arrhythmogenic right ventricular cardiomyopathy; NGS, next generation sequencing; NsSNV, non-synonymous single nucleotide variant; ECG, electrocardiography; CMRI, cardiac magnetic resonance; SCD, sudden cardiac death; NSVT, non-sustained ventricular tachycardia; SIFT, sorting intolerant from tolerant; PolyPhen-2, polymorphism phenotyping v2; GERP, genomic evolutionary rate profiling; RMSD, root–mean–square deviation; MWT, mean maximal wall thickness; LVM, left ventricular mass; LVEF, left ventricular ejection fraction; ICD, intracardiac defibrillator; PMK, pacemaker; NMR, nuclear magnetic resonance; VOUS, variants of unknown clinical significance.

* Corresponding author.
E-mail address: i.bottillo@gmail.com (I. Bottillo).

http://dx.doi.org/10.1016/j.gene.2015.11.048
0378-1119/

are inherited as an autosomal dominant trait in about 50–60% of affected adults. Besides, in a minor proportion of patients, mutations have been also identified in several genes encoding Z-disk, cytoskeletal and other non-sarcomeric proteins involved in Ca^{++} homeostasis, including *ANKRD1*, *CALR3*, *CAV3*, *CSRP3*, *DES*, *JPH2*, *MYLK2*, *MYOZ2*, *MYPN*, *NEXN*, *PLN*, *TCAP* and *VCL* (Millat et al., 2010). When occurring in the context of a multisystemic disorder, hypertrophic cardiomyopathy may also represent a relevant feature of non-sarcomeric gene syndromes comprising *GLA* (Fabry's disease, OMIM #301500), *LAMP2* (Danon's disease, OMIM #300257) and *ABCC9* (Cantù syndrome, OMIM #239850).

Both HCM and other cardiomyopathies such as dilatative cardiomyopathy (DCM), restrictive cardiomyopathy (RCM), arrhythmogenic right ventricular cardiomyopathy (ARVC), can be viewed as a continuum of phenotypes sharing typical symptoms of heart failure and pathophysiological features (Frey et al., 2012). Furthermore, the cardiomyopathies constitute an allelic series since they can be caused by mutations at the same genes (Hershberger et al., 2013). Beyond their substantial genetic overlap, only recently some genotype-phenotype associations at a strong level of significance have been described. Particularly for HCM, DNA variants in sarcomeric genes have been associated with asymmetric septal hypertrophy pattern, younger age at presentation, family history of the disease and sudden cardiac death, greater maximum left ventricle wall thickness and an increased incidence of cardiovascular death (Lopes et al., 2014). Although etiology identification may be important for screening of at risk family members, cardiomyopathy molecular diagnosis is still difficult due to the large number of causative genes and the high rate of private mutations. Next generation sequencing (NGS) results in a remarkable increase in speed and efficiency, transforming our insights about rare-variant genomic diseases, and over the past decade, many bioinformatics algorithms have been developed to predict functional consequences of single nucleotide variants in coding regions.

The aim of this study was to explore the role in HCM onset and expression of 62 sarcomeric and non-sarcomeric genes in 41 patients with hypertrophic cardiomyopathy and 3 cases affected by a HCM-related disorder. Each identified rare (frequency ≤0.01) non-synonymous Single Nucleotide Variant (nsSNV) was firstly validated by Sanger sequencing. We also adopted a comprehensive computational schema for variants's annotation and predictions, in order to improve genomic reports generated from the experimental studies, and to better understand the variations implicated in HCM manifestations. Genotype–phenotype correlations were finally carried out for outlining the contribution of each analyzed gene in the onset and clinical variability of HCM.

2. Materials and methods

2.1. Patients

Patients were selected from those attending an outpatient hospital service dedicated to the diagnosis and management of HCM. All patients underwent clinical history registration, physical examination, electrocardiography (ECG), echocardiography, cardiopulmonary exercise test coupled with ambulatory ECG monitoring. Thirty-two out of 44 patients were evaluated by cardiac magnetic resonance (CMRI). Diagnostic criteria for HCM was defined in adults by a maximal left ventricular wall thickness ≥ 13 mm on echocardiography, in the absence of other loading conditions (Klues et al., 1995). Family history of Sudden Cardiac Death (SCD), syncope episodes and the presence of Non-Sustained Ventricular Tachycardia (NSVT) were defined as described by O'Mahony et al. (2014). Electrocardiographic changes that were considered of clinical significance included abnormal Q waves (0.04 s or 25% depth of an R-wave), LVH (voltage criteria), and marked repolarization changes (e.g. T-wave inversion in at least 2 leads). The QT interval corrected for heart rate was calculated using the Bazett's formula (QTc (ms) = QT / √RR, where RR is the RR interval measured in seconds). Familial HCM cases were defined if at least one additional affected family member with HCM, or one case of sudden cardiac death was present in the pedigree. All patients gave informed consent for the DNA analyses, which was approved by local ethic committees in accordance with the principles of the Declaration of Helsinki.

2.2. Next generation sequencing (NGS)

Genomic DNA from peripheral blood was tested by NGS with a custom design for the cardiomyopathy panel, based on AmpliSeq strategy (ThermoFisher, Carlsbad, CA, USA). The panel was designed to analyze coding, intronic junctions and UTR sequences of 62 genes. Those comprised 12 sarcomeric *loci* (*ACTC1*, *MYBPC3*, *MYH6*, *MYH7*, *MYL2*, *MYL3*, *MYOM1*, *NEBL*, *TNNC1*, *TNNI3*, *TNNT2*, *TPM1*), 9 cytoskeletal (*CRYAB*, *DES*, *DMD*, *DTNA*, *EMD*, *FXN*, *LAMA4*, *PDLIM3*, *SGCD*), 9 Z-disk (*ACTN2*, *ANKRD1*, *CSRP3*, *LDB3*, *MYOZ2*, *MYPN*, *NEXN*, *TCAP*, *VCL*), 6 desmosomal (*DSC2*, *DSG2*, *DSP*, *FHL2*, *JUP*, *PKP2*), 5 intracellular Ca^{++} homeostasis (*CALR3*, *CASQ2*, *JPH2*, *PLN*, *RYR2*), 3 genes encoding for K^+ and Na^+ channels and interacting proteins (*ABCC9*, *CAV3*, *SCN5A*) and 18 other non-sarcomeric loci (*BAG3*, *CTF1*, *EYA4*, *GATAD1*, *GLA*, *ILK*, *LAMP2*, *LMNA*, *MYLK2*, *PRKAG2*, *PTPN11*, *RAF1*, *RBM20*, *TAZ*, *TMEM43*, *TMPO*, *TTR*, *TXNRD2*).

In summary, the panel included 36 genes known to be associated with HCM and dilated cardiomyopathy (*ACTC1*, *ACTN2*, *ANKRD1*, *BAG3*, *CALR3*, *CAV3*, *CRYAB*, *CSRP3*, *DES*, *EYA4*, *GATAD1*, *ILK*, *JPH2*, *LAMA4*, *LDB3*, *LMNA*, *MYBPC3*, *MYH6*, *MYH7*, *MYL2*, *MYL3*, *MYLK2*, *MYOZ2*, *MYPN*, *NEXN*, *PDLIM3*, *PLN*, *RBM20*, *SGCD*, *TCAP*, *TMPO*, *TNNC1*, *TNNI3*, *TNNT2*, *TPM1* and *VCL*), 7 genes related with arrhythmogenic right ventricular cardiomyopathy and left ventricular non-compaction (*DSC2*, *DSG2*, *DSP*, *DTNA*, *JUP*, *PKP2* and *TMEM43*), 7 genes associated to malformation syndromes and storage disorders (*ABCC9*, *GLA*, *LAMP2*, *PTPN11*, *PRKAG2*, *RAF1* and *TTR*), 4 genes related to myopathies and neuromuscular disorders (*DMD*, *EMD*, *FXN* and *TAZ*), 3 genes implicated in arrhythmia syndromes/ion-channel disease (*CASQ2*, *RYR2* and *SCN5A*) and other 5 cardiovascular candidate genes (*CTF1*, *FHL2*, *MYOM1*, *NEBL* and *TXNRD2*).

Supplementary Table 1 summarizes the genes included in the study as well as their function, associated disorder(s), chromosomal position and sequencing details. The design allowed the targeted resequencing of 2890 amplicons (global size: 345,39 kb/patient) by Ion Torrent PGM instrument (ThermoFisher, Carlsbad, CA, USA). In order to validate the applied NGS protocol the entire coding sequences of *MYH7*, *MYBPC3*, *TNNT2* and *TNNI3* genes of 23/44 patients were at first analyzed by Sanger sequencing.

2.3. Analysis of raw NGS data

The NGS read depth of each analysed gene is shown in Fig. S1. The genes included in the NGS panel, as well as their function within the cell, associated disorder(s), chromosomal location and sequencing details are shown in Supplementary Table 1. The total number and type of DNA point variants identified for each patient is given in Supplementary Table 2. For each patient, the summary of clinical and molecular data is tabulated in Supplementary Table 3. The synonymous variants identified for each patient are given in Supplementary Table 4.

2.4. Variant calling, annotation and prioritization

Variant calls were annotated by wANNOVAR Web Server (wannovar2.usc.edu/). The nucleotide variants with a Minor Allele Frequency (MAF) ≤0.01 both in the 1000 Genome Project (global and European) and in the NHLBI-ESP 6500 exome project (global) were filtered and designed as “rare variants”. Between the selected variants, only the exonic and ±10 bp intronic ones were prioritized

and validated by Sanger sequencing. The confirmed DNA changes were subjected to different in silico predictions by the following methods: SIFT, PolyPhen-2_HDIV, PolyPhen-2_HVAR, Provean, LRT, Mutation Taster, Mutation Assessor, FATHMM, RadialSVM, LRT, CADD v1.3 and molecular modeling, for the analysis of missense variants; Human Splicing Finder (HSF) 3.0, for the analysis of intronic changes; GERP++, PhyloP placental, PhyloP vertebrate and SiPhy, for exploring nucleotide-specific estimates of evolutionary constraint (Bottillo et al., 2015).

2.5. Statistical assessment of genotype–phenotype correlations

All statistical analyses were performed using SPSS (statistical software package version 20.0). For each patient, the clinical data and molecular data (i.e. total number nsSNVs, number of any class of nsSNVs) were tabulated (Supplementary Table 3).

Phenotype data were presented as continuous variables obtained from clinical data and instrumental measurements. Categorical variables were shown as the presence/absence or grade, of the clinical feature. Continuous variables were summarized using means and standard deviations, while categorical variables were summarized using frequencies and percentages.

The molecular data were presented as number of DNA rare (frequency ≤0.01) nsSNVs variants identified in each functional class. The genes carrying rare variants were indeed grouped into sarcomeric, desmosomal, K^+ and Na^+ channels and interacting proteins, *loci* for mRNA splicing and cellular enzymes, cytoskeletal, Z-disk and Ca^{++} homeostasis.

Statistical correlations for non-syndromic patients were evaluated between the presence of rare DNA variants (or the family history) and the prevalence of phenotypic traits with the following methods: (i) Chi-square test when referred to categorical variables; (ii) unpaired two-tailed Student's t-test when related to continuous variable. Due to the relatively low number of patients no adjustments was planned for multiple testing; the analysis is therefore exploratory and results to be considered as hypotheses-generating.

3. Results

3.1. Phenotype analysis

Forty-four unrelated Caucasian/Italian patients (23 males and 21 females) were included in this study. Three patients displayed HCM associated with a syndromic phenotype: case 10, 24 and 27 were respectively affected by a mild form of Fabry's disease, by Cantù's syndrome and by Danon's disease. Supplementary Table 3 summarizes the clinical features of the patients at evaluation.

3.2. Genotyping results

Among the 41 non-syndromic HCM cases, 33 (80%) were found to carry at least one rare nsSNV. In total, 95 non-synonymous sequence changes, of which 87 different, were identified. These included 73 missense, 13 intronic, 3 frameshift, 4 nonsense, one stop-loss and one in frame deletion. Among the 87 different nsSNVs, 21 were already reported in literature (Table 1).

The proportion of nsSNVs identified for each analyzed gene is shown in Fig. 1a.

About 33% of the changes mapped in sarcomeric *loci*, 19% in desmosomal *loci*, 15% in genes coding for K^+ and Na^+ channels and for channels-interacting proteins, 12% in genes coding for mRNA splicing and other cellular enzymes, 9% in cytoskeletal *loci*, 6% in genes related to intracellular Ca^{++} homeostasis, 5% in Z-disk *loci* and 5% in other non-sarcomeric *loci* (Fig. 1b).

The 27% of the patients harbored two heterozygous DNA changes in different genes, 15% a single heterozygous change and 39% three or more variants (Fig. 1c).

Twenty-four/41 of the cases harbored at least one sarcomeric nsSNV, and three patients carried only desmosomal changes. In total, 58% carried alterations in sarcomeric *loci*, 14% in desmosomal and 7% in other non-sarcomeric ones without any sarcomere change. Fig. 1d illustrates the fraction of mutated cases respect to the different genes' categories.

Among the syndromic patients, case 10 was found to carry the p.N215S mutation in the *GLA* gene, previously associated with a mild form of Fabry disease (Davies et al., 1993; Eng et al., 1993), case 24 harbored the ABCC9 p.R1154W mutation causing Cantù's syndrome (Harakalova et al., 2012; van Bon et al., 2012) and case 27 carried the novel *LAMP2* p.F151fs mutation. In addition, they were found to carry other rare nsSNVs (Table 1).

The molecular analyses also identified 49 heterozygous synonymous variants, of which 36 different, that are reported in Supplementary Table 4.

3.3. In silico predictions

Twelve methods were employed for assessing the possible pathogenicity of the identified missense changes (see Table 1 in (Bottillo et al., 2015)). The variants with at least 6/12 deleterious predictions have been considered possibly pathogenetic, while a "no predicted deleterious effect" was accounted for those scored with two or less deleterious predictions (Table 1). The in silico scores were completely concordant for 20 variants (14 benign and 6 deleterious), but ambiguous for 19 changes which resulted in 3 to 5 deleterious predictions among 12 ones

Nine among the 14 different intronic variants were predicted to alter a splicing site.

Among the 85 different nsSNVs identified in this study, 24 resulted in the alteration of an allele that is highly conserved across evolutionarily distant species (Table 1). For those DNA changes, the four computational methods employed for the nucleotides' conservation analysis gave concordant scores (see Table 1 in Bottillo et al. (2015)).

Inferring from the type of the variant, querying of frequency and mutational databases, as well as from the output of in silico predictions, all but three (patient 3, 37 and 52) mutated cases harbored at least one non-synonymous DNA change predicted to be pathogenetic, or already reported as disease-modifying factor (Table 1).

3.4. Genotype–phenotype correlations

We observed an inverse correlation between the age at diagnosis and the total number of nsSNVs: the mean age at diagnosis was indeed 60 years in the group without variants and 43.8 years in the group with variants ($P = 0.01$).

The presence of nsSNVS in the genes involved in intracellular Ca^{++} homeostasis was significantly correlated to an earlier HCM onset: the mean age at diagnosis was in fact 35.8 years between the cases carrying those variants, versus 48.2 years in the cases without them ($P = 0.03$).

The mean value of maximum wall thickness was lower both in the group carrying nsSNVs in genes for K^+ and Na^+ channels ($P = 0.03$), and in the group not harboring variants for Ca^{++} homeostasis ($P = 0.05$). Moreover we observed a correlation between maximum wall thickness and the family history, as the mean of the MWT value of was lower in the sporadic group (18.055 mm) than in the familial group (22.3 mm) ($P = 0.05$).

The occurrence of non-synonymous changes in cytoskeletal *loci* was correlated with the presence of non-sustained ventricular tachycardia: NSTV was indeed present in 40% of the patients

Table 1
Rare nsSNVs identified in 36/44 HCM patients. A comment inferred from review of literature, in silico predictions and evolution conservation analysis, is given for each DNA change in the last column. Based on those considerations and on the type of variant, comments with a gray background highlight DNA changes predicted to be possibly pathogenetic. Comments in bold stand for DNA variants that have been previously reported as phenotype modifying factors.

Pt	Gene	Mutation	Status	1000G ALL	1000G EUR	ESP6500si ALL	Type	ClinVar SIG	Reference	Comment
1	*MYBPC3*	c.T1664C,p.M555T	Htz	–	–	–	missense	unk	Girolami et al. 2006	No predicted deleterious effect
	RyR2	c.A3380G,p.E1127G	Htz	0.0006	–	0.0006	missense	unk	–	8/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
3	*TMPO*	c.A1037G,p.H346R	Htz	0.0002	–	–	missense	–	–	No predicted deleterious effect
5	*MYH6*	c.643–5C>T	Htz	–	–	0.0005	intronic	unk|PnPath	–	Branch–point splice site broken
8	*RBM20*	c.G3373A,p.E1125K	Htz	0.002	0.007	0.0037	missense	PnPath	–	Ambiguous *in silico* predictions
	MYH6	c.2928+5G>A	Htz	0.0002	0.001	0.0018	intronic	PnPath	–	Donor splice site broken
	CSRP3	c.T10C,p.W4R	Htz	0.0018	0.005	0.0037	missense	unk|PnPath	Knoll et al. 2002 Geier et al. 2008	6/12 deleterious predictions Previously reported alone in CMD1M patients and in conjunction with a sarcomeric mutation in HCM patients The reference allele is conserved across evolutionarily distant species
	LAMA4	c.A4937G,p.E1646G	Htz	–	–	0.0001	missense	–	–	11/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	LAMA4	c.G1565C,p.R522T	Htz	–	–	–	missense	–	–	No predicted deleterious effect
9	*MYBPC3*	c.A3825C,p.*1275Cysext*33	Htz	–	–	–	stop loss	unk	–	Stop loss
	NEBL	c.G604A,p.G202R	Htz	0.0014	0.003	0.0018	missense	–	Purevjav et al. 2010	Previously identified in one DCM individual. The Gly202Arg mouse model exhibits DCM features
	RyR2	c.9450–9C>T	Htz	–	–	–	intronic	–	–	No predicted deleterious effect
10*	*MYBPC3*	c.C1112A,p.P371Q	Htz	–	–	–	missense	Path	–	8/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	GLA	c.A644G,p.N215S	Htz	–	–	–	missense	–	Eng et al. 1993 Davies et al. 1993	9/12 deleterious pr edictions Previously associated with a mild forms of Fabry disease The reference allele is conserved across evolutionarily distant species
	DSP	c.G137A,p.G46D	Htz	–	–	0.0001	missense	–	–	No predicted deleterious effect
	TXNRD2	c.G1150A,p.G384S	Htz	0.0018	0.006	0.0041	missense	–	Sibbing et al. 2011	Ambiguous *in silico* predictions
11	*JPH2*	c.G1536C,p.W512C	Htz	–	–	–	missense	–	–	Ambiguous *in silico* predictions
	SCN5A	c.A5605T,p.I1869F	Htz	–	–	–	missense	PnPath	–	12/12 deleterious predictions
	FHL2	c.G109T,p.A37S	Htz	0.0016	–	0.0023	missense	unk	–	Ambiguous *in silico* predictions
	DSC2	c.C1787T,p.A596V	Htz	0.0022	0.002	0.0009	missense	PnPath	den Haan et al. 2009	Ambiguous *in silico* predictions Previously reported in one individual with ARVC The reference allele is conserved across evolutionarily distant species
14	*PKP2*	c.G76A,p.D26N	Htz	0.003	0.011	0.0049	missense	–	Christensen et al. 2010	6/12 deleterious predictions Previously reported as modifier of disease, over–represented in ARVC cases
15	*MYOM1*	c.C2131T,p.R711C	Htz	–	–	0.0001	missense	–	–	11/12 deleterious predictions
	RAF1	c.[124G>A;125C>T] p.A42I	Htz	–	–	–	missense	–	–	Ambiguous *in silico* predictions
17	*TNNT2*	c.A252T,p.R84S	Htz	–	–	–	missense	unk	–	10/12 deleterious predictions Previously identified in one Asian HCM individual
	MYOM1	c.4485–6T>C	Htz				intronic	–	–	Branch–point splice site broken
	RBM20	c.A59G,p.D20G	Htz	–	–	–	missense	–	–	Ambiguous *in silico* predictions
	CAV3	c.C233T,p.T78M	Htz	0.002	0.005	0.0043	missense	untested	Fulizio et al. 2005 Vatta et al 2006 Cronk et al 2007	8/12 deleterious predictions Previously found in 3 unrelated individuals with long QT syndrome. The mutant causes a 5–fold increase in late sodium current compared to wild–type
18	*DSG2*	c.A208G,p.I70V	Htz	–	–	–	missense	–	–	No predicted deleterious effect
	DSP	c.G5218A,p.E1740K	Htz	0.0004	0.002	0.0009	missense	unk	Cox et al. 2011	Ambiguous *in silico* predictions Previously reported in two ARVC individuals; both individuals carried a second likely pathogenic variant
19	*DSP*	c.4441_4443del,p.K1481_1481del	Htz				in frame del	–	–	In frame deletion
	MYOM1	c.G2132A,p.R711H	Htz	0.0002	–	–	missense	Ppath	–	11/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	SCN5A	c.C5837T,p.S1946F	Htz	–	–	–	missense	–	Hermida JS 2010	7/12 deleterious predictions Reported as "probably pathogenetic" in ClinVar database
	DSC2	c.C2328G,p.I776M	Htz	–	–	0.0001	missense	–	–	No predicted deleterious effect
	LAMA4	c.T3482C,p.M1161T	Htz	–	–	–	missense	–	–	No predicted deleterious effect
	LAMP2	c.C418G,p.L140V	Hem	–	–	–	missense	–	–	No predicted deleterious effect
20	*DSG2*	c.G2147A,p.G716E	Htz	–	–	–	missense	nPath	–	No predicted deleterious effect
	LAMA4	c.4665+8G>T	Htz	0.0004	–	0.0016	intronic	other|Path	–	IIE site splice broken
21	*MYBPC3*	c.2309–2A>G	Htz	–	–	–	intronic	–	Van Driest 2004 Roncarati 2011	Acceptor splice site broken Reported as "pathogenetic" in ClinVar database
	RyR2	c.G1454A,p.R485Q	Htz	–	–	–	missense	PPath|nPath	–	10/12 deleteri ous predictions The reference allele is conserved across evolutionarily distant species
	ABCC9	c.G2200A,p.V734I	Hmz	0.004	0.014	0.0092	missense	–	Minoretti et al. 2006	Ambiguous *in silico* predictions Previously reported as influencing susceptibility to precocious myocardial infarction
	CALR3	c.A1036G,p.I346V	Htz	–	–	–	missense	PnPath|other	–	No predicted deleterious effect
	TNNT2	c.264+7G>A	Htz	0.01	–	0.011	intronic	–	Kassem et al 2013	Donor splice site broken. It is a common variant in the Black population

Table 1 (*continued*)

22	*MYH7*	c.4402G>A,p.E1468K	Htz	–	–	–	missense	–	–	9/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	ILK	c.G1086C,p.Q362H	Htz	–	–	0.0001	missense	unk	–	11/12 deleterious predictions
23	*DSC2*	c.C1787Tp.A596V	Htz	0.0022	0.002	0.0009	missense	–	den Haan et al. 2009	Ambiguous *in silico* predictions Previously reported in one individual with ARVC The reference allele is conserved across evolutionarily distant species
	LAMA4	c.G1967A,p.S656N	Htz	–	–	–	missense	–	–	No predicted deleterious effect
	MYH6	c.G68A,p.R23H	Htz	–	–	–	missense	–	–	11/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	TPM1	c.C134T,p.A45V	Htz	–	–	–	missense	unk|unk	–	9/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
24*	*PKP2*	c.C2299A,p.R767S	Htz	0.0006	0.002	0.0004	missense	Path|other	Fressart 2010 Klauke 2010 Tan 2010	6/12 deleterious predictions Previously identified in 4 ARVC, one HCM, and one VT patients
	DSP	c.G88A,p.V30M	Htz	0.0022	0.001	0.0012	missense	–	Yang et al. 2006	Previously found in one ARVC patient. The mutant allele fails to localize to the cell membrane in a desmosome–forming cell line and fails to bind to and coimmunoprecipitate junction plakoglobin
	RyR2	c.A4273G,p.T1425A	Htz	–	–	–	missense	–	–	Ambiguous *in silico* predictions The reference allele is conserved across evolutionarily distant species
	MYBPC3	c.821+3G>A	Htz				intronic	Path	–	IIE splice site broken
	ABCC9	c.C3460T,p.R1154W	Htz	–	–	–	missense	–	van Bon 2012 Harakalova 2012	11/12 deleterious predictions Reported as "pathogenetic" in ClinVar database Causes Cantù syndrome
25	*MYBPC3*	c.2846dupT,p.M949fs	Htz	–	–	–	frameshift	unk|PnPath	–	Truncating mutation
	TXNRD2	c.375-8C>T	Htz	0.0004	–	0.0007	intronic	PnPath	–	No predicted deleterious effect
	CSRP3	c.T10C,p.W4R	Htz	0.0018	0.005	0.0037	missense	unk	Knoll et al. 2002 Geier et al. 2008	Previously reported alone in CMD1M patients and in conjunction with a sarcomeric mutation in HCM patients
	RBM20	c.G3373A,p.E1125K	Htz	0.002	0.007	0.0037	missense	unk	–	Ambiguous *in silico* predictions
26	*DSP*	c.C3956T,p.T1319I	Htz	–	–	0.0001	missense	untested	–	8/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	ABCC9	c.2238-1G>A	Htz	0.0018	0.001	0.0008	intronic	other|Path	–	Acceptor splice site broken The reference allele is conserved across evolutionarily distant species
	CAV3	c.C233T,p.T78M	Htz	0.002	0.005	0.0043	missense	–	Fulizio et al. 2005 Vatta et al 2006 Cronk et al 2007	8/12 deleterious predictions Previously found in 3 unrelated individuals with long QT syndrome. The mutant causes a 5–fold increase in late sodium current compared to wild–type
	MYBPC3	c.2309–2A>G	Htz	–	–	–	intronic	–	–	Acceptor splice site broken Reported as "pathogenetic" in ClinVar database
	DTNA	c.C2095T,p.R699C	Htz				missense	PnPath	–	No predicted deleterious effect
27*	*LMNA*	c.1363–7T>C	Htz				intronic	unk|unk	–	No predicted deleterious eff ect
	LAMP2	c.453delT,p.F151fs	Hem				frameshift	Path|nPath	–	Truncating mutation
29	*ABCC9*	c.816+11G>A	Htz	0.0004	0.001	0.0002	intronic	unk|PnPath|nPath	–	Donor splice site broken
31	*PKP2*	c.C2299A,p.R767S	Htz	0.0006	0.002	0.0004	missense	–	Fressart 2010 Klauke 2010 Tan 2010	6/12 deleterious predictions Previously identified in 4 ARVC, one HCM, and one VT patients
	JPH2	c.G1513A,p.G505S	Htz	0.015	0.011	0.0071	missense	–	Matsushita 2007 Bean 2013	Previously identified in 4 Japanese HCM probands
	MYBPC3	c.3192dupC,p.K1065fs	Htz	–	–	–	frameshift	–	Girolami 2006 Girolami 2010 Olivotto 2011	Truncating mutation
32	*JUP*	c.909+6C>T	Htz	–	–	0.0014	intronic	–	–	No predicted deleterious effect
	RAF1	c.G1858A,p.A620T	Htz	–	–	–	missense	PnPath	–	6/12 deleterious pre dictions The reference allele is conserved across evolutionarily distant species
	SCN5A	c.C1810G,p.L604V	Htz	–	–	–	missense	unk|PPath	–	Ambiguous *in silico* predictions
33	*PDLIM3*	c.G163T,p.E55X	Htz	–	–	–	nonsense	–	–	Truncating mutation
	MYBPC3	c.C3775Tp.Q1259X	Htz	–	–	–	nonsense	PnPath	–	Truncating mutation
34	*ABCC9*	c.2867–5T>C	Htz	–	–	0.0001	intronic	untested|Path|PnPath	–	No predicted deleterious effect
	MYOM1	c.T199G,p.S67A	Htz	–	–	0.0001	missense	–	–	No predicted deleterious effect
	PKP2	c.G76A,pD26N	Htz	0.003	0.011	0.0049	missense	–	Christensen et al. 2010	6/12 deleterious predictions Previously reported as modifier of disease, over–represented inARVC cases
	TNNT2	c.A755T,p.K252I	Htz	–	–	–	missense	–	–	11/12 deleterious predictions Reported as "probably pathogenetic" in ClinVar database
	SCN5A	c.G2614A,p.D872N	Htz	–	–	–	missense	–	–	6/12 deleterious predictions
	BAG3	c.G463A,p.A155T	Htz	0.0018	0.008	–	missense	untested	–	No predicted deleterious effect
35	*MYPN*	c.C3335T,p.P1112L	Htz	0.0026	0.005	0.0021	missense	PPath|Path	Duboscq–Bidot 2007 Adzhubei 2010 Purevjav 2012	7/12 deleterious predictions Previously reported in HCM and DCM patients The reference allele is conserved across evolutionarily distant species
	CAV3	c.G221A,p.R74H	Htz	0.0002	–	–	missense	PnPath	El Huneidi 2014	11/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
36	*MYH7*	c.G676A,p.A226T	Htz	–	–	–	missense	–	–	9/12 deleterious predictions
37	*DSP*	c.G6188A,p.R2063Q	Htz	–	–	0.0001	missense	unk	–	Ambiguous *in silico* predictions
	DSC2	c.C595T,p.R199C	Htz	–	–	–	missense	PnPath	–	Ambiguous *in silico* predictions

(*continued on next page*)

Table 1 (*continued*)

38	*CAV3*	c.C233T,p.T78M	Htz	0.002	0.005	0.0043	missense	other|Path	Fulizio et al. 2005 Vatta et al 2006 Cronk et al 2007	8/12 deleterious predictions Previously found in 3 unrelated individuals with long QT syndrome. The mutant causes a 5–fold increase in late sodium current compared to wild–type
	MYH7	c.G3346A,p.E1116K	Htz	–	–	–	missense	PnPath|unk	–	10/12 deleterious predict ions Reported as "pathogenetic" in ClinVar database The reference allele is conserved across evolutionarily distant species
43	*MYH6*	c.G622A,p.D208N	Htz	0.002	0.005	0.0049	missense	other|Path	Granados-Riveron et al. 2010	No predicted deleterious effect
	MYBPC3	c.C1302A,p.Y434X	Htz	–	–	–	nonsense	PnPath	–	Truncating mutation
44	*MYH6*	c.G4193A,p.R1398Q	Htz	–	–	0.0001	missense	–	–	7/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
	DSG2	c.G2033C,p.G678A	Htz	–	–	0.0002	missense	nPath	–	No predicted deleterious effect
	PKP2	c.C1445T,p.T482M	Htz	0.0038	0.005	0.0037	missense	unk	–	No predicted deleterious effect
	MYH7	c.G428A,p.R143Q	Htz	–	–	–	missense	unk	Song 2005 Van Driest 2004 Wang 2007 Kimura 2010 Gruner 2011 Fokstuen 2011	9/12 deleterious predictions Previously reported in 6 HCM individuals The reference allele is conserved across evolutionarily distant species
45	*MYBPC3*	c.A649G,p.S217G	Htz	0.0018	0.001	0.001	missense	unk	Lakdawala et al. 2012	Ambiguous *in silico* predictions Likely benign when present in isolation
	MYH7	c.C2167T,p.R723C	Htz	–	–	–	missense	unk|PnPath	Watkins 1992 Tesson 1998 Richard 2003 Ingles 2005 Girolami 2010	10/12 deleterious predictions Previously reported in 5 HCM families Residue 723 is conserved among all known cardiac MHCs and all vertebrate striated muscle MHCs except the human perinatal and rabbit skeletal isoforms Mutation to a Cys changes Thr net charge
46	*MYBPC3*	c.3192dupC,p.K1065fs	Htz	–	–	–	frameshift	–	Girolami 2006 Girolami 2010 Olivotto 2011	Truncating mutation
49	*RBM20*	c.G3373A,p.E1125K	Htz	0.002	0.007	0.0037	missense	–	–	Ambiguous *in silico* predictions
	DSP	c.C5851T,p.R1951X	Htz	–	–	–	nonsense	ClinVar_SIG	–	Truncating mutation
	TMPO	c.G1696A,p.D566N	Htz	–	–	–	missense	unk	–	Ambiguous *in silico* predictions
	MYL2	c.G34T,p.G12C	Htz	–	–	–	missense	unk	–	7/12 deleterious predictions The reference allele is conserved across evolutionarily distant species
51	*VCL*	c.A1907G,p.H636R	Htz	0.0004	0.001	0.0009	missense	–	Zimmerman 2010	9/12 deleterious predictions Detected in 1/>250 Caucasian DCM individuals and in another DCM proband of unknown ethnicity The reference allele is conserved across evolutionarily distant species
	MYBPC3	c.T1664C,p.M555T	Htz	–	–	–	missense	unk|PnPath	Girolami et al. 2006	No predicted deleterious effect
52	*ACTN2*	c.C1484T,p.T495M	Htz	–	–	0.0001	missense	PnPath		Ambiguous *in silico* predictions The reference allele is conserved across evolutionarily distant species
	SCN5A	c.C2074A,p.Q692K	Htz	0.0002	0.001	0.0002	missense	PnPath	Van Langen 2003 Ackerman 2004	Ambiguous *in silico* predictions
	DMD	c.G1646A,p.R549Q	Hem	–	–	–	missense	unk|PnPath	–	Ambiguous *in silico* predictions The reference allele is conserved across evolutionarily distant species
	DMD	c.C3164T,p.T1055I	Hem	–	–	–	missense	–	–	No predicted deleterious effect

Based on Table 1 in Bottillo et al. (2015), missense changes with at least 6/12 deleterious in silico predictions have been considered "risk variants" possibly pathogenetic (gray background). Nucleotides highly conserved across evolutionarily distant species are those for which all the conservation analyses were concordant.
Pt: patient; **1000G ALL**: MAF in 1000 Genomes Project global; **1000G EUR**: MAF in 1000 Genomes Project European; **ESP6500si ALL**: MAF in exome sequencing project global; **Htz**: heterozygosity; **Hem**: hemizygosity; **Hmz**: homozygosity; **ClinVar SIG**: significance in ClinVar database, including unknown (**unk**), non-pathogenic (**nPath**), probable-non-pathogenic (**PnPath**), probable-pathogenic (**PPath**), pathogenic (**Path**); **HCM**: hypertrophic cardiomyopathy; **DCM**: dilated cardiomyopathy; **CMD1M**: cardiomyopathy, dilated, 1 M; **ARVC**: arrhythmogenic right ventricular cardiomyopathy; **IIE**: intron identity element; **VT**: ventricular tachycardia; **MHC**: myosin heavy chain.
* Syndromic patients.

carrying cytoskeletal nsSNVs, versus 11% of patients not carrying them ($P = 0.05$). NSTV was also correlated with familial cases, respect to sporadic ones ($P = 0.02$).

Cytoskeletal variants were correlated with a minor grade (grade 0–1) of diastolic dysfunction ($P = 0.004$) and, regarding New York Heart Association (NYHA) functional classification, sarcomeric variants were present in all grade I (no limitation of physical activity), in 56% of grade II (slight limitation of physical activity), and in 33% of grade III (marked limitation of physical activity) patients ($P = 0.05$).

The prevalence of male sex was higher, respect to female in the group carrying at least one nsSNV compared with the HCM patients with no rare nsSNV ($P = 0.01$).

Finally we identified a statistically significant correlation between the presence of an implantable cardioverter-defibrillator and family history ($P = 0.03$): the implantation of ICD resulted to be linked to a familial framework.

4. Discussion

Here we present a NGS analysis of the prevalence of sarcomeric and non-sarcomeric gene variants in 41 patients affected by primary hypertrophic cardiomyopathy, and in 3 cases affected by syndromic HCM. This study was aimed the discovery of disease causing genetic variation for an appropriate genetic counseling, as well as for the enlargement of the HCM mutational spectrum thorough a broad genetic test. In the context of cardiomyopathies, the characterization of the family mutation helps indeed in planning surveillance and early detecting possible complications in close relatives. This also implies

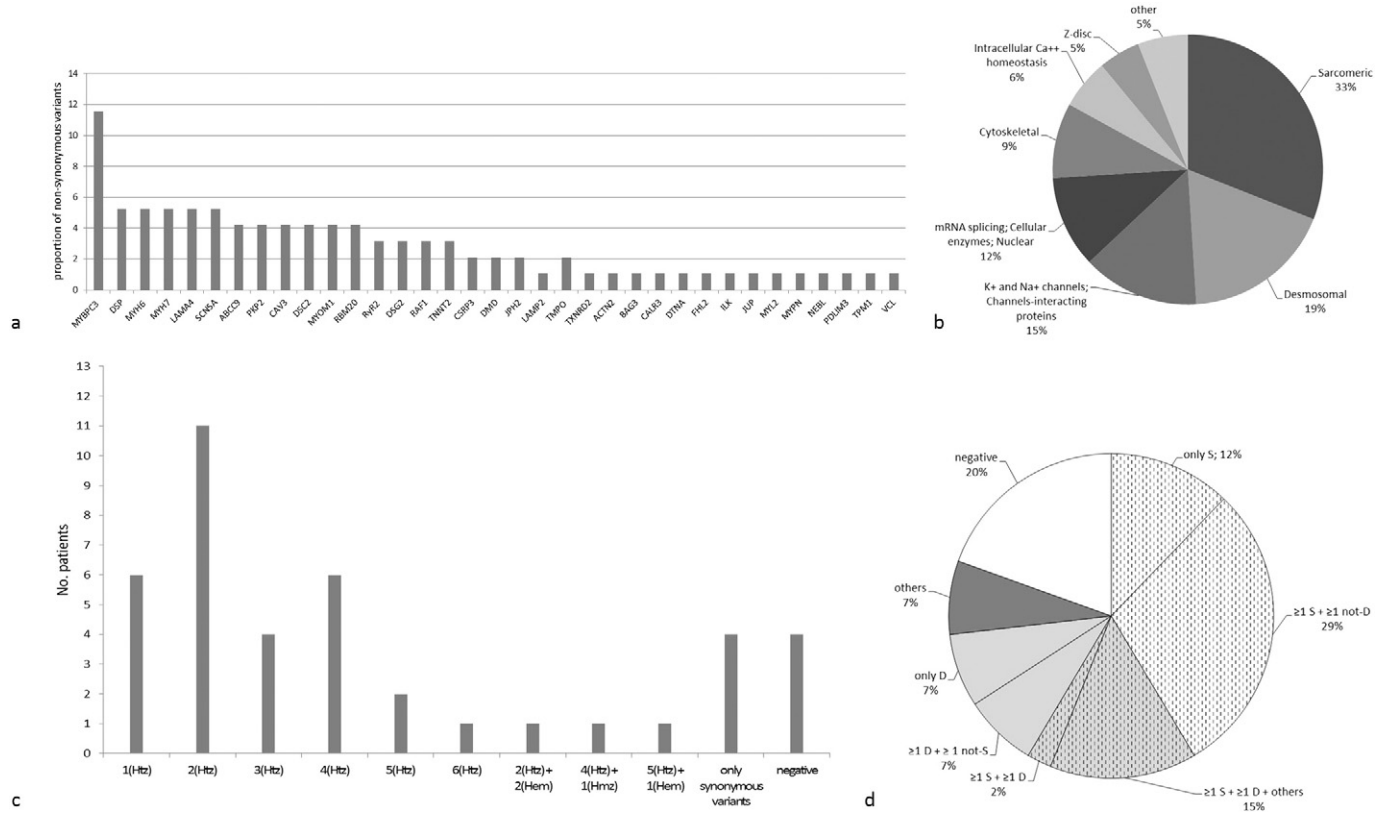

Fig. 1. Distribution of identified nsSNVs. (a–b) Proportion of nsSNVs identified for (a) each analyzed gene, and for (b) each functional genic category. "Other" includes the *loci* coding for lysosomal, thioredoxin reductase, anti-apoptotic and membrane proteins. (c) Number and zygosity of DNA changes identified over 37 HCM patients. Six cases were found to carry a single heterozygous nsSNV, 11 cases two heterozygous nsSNVs and 16 cases three or more nsSNVs. Four patients carried only synonymous changes, and four resulted negative to the NGS screening. (d) Proportion of mutated patients respect to the genes' category. Dashed slices (both white and gray) stand for cases harboring at least one sarcomeric non-synonymous variant. Light gray slices (both dashed and not dashed) represent patients not carrying any sarcomeric nsSNV, but at least one desmosomal one.

that relatives discovered without the familial mutation can avoid unnecessary follow-up.

The employed high throughput sequencing technology enabled the generation of large amounts of sequence data, and determined the need of an accurate assessment of the identified genetic variations. The experimental characterization of all the observed DNA changes would have support the functional and/or regulatory impact of the various mutations, but it requires a long experimental time that was impractical in our clinical setting. Likewise, co-segregation analysis of the DNA changes within families would have been appropriate, but it is uninformative in small pedigrees and it is hindered by HCM incomplete penetrance and variable expressivity.

A high number of bioinformatics solutions for the annotation, scoring and classification of variants are currently available to address the challenge of predicting the functional consequences of a mutation. Our results show the relevance of incorporating integrated computational workflows to predict the biomedical impact of the various DNA variants resulting from a NGS approach, and to identify functionally significant or clinically actionable variants.

After the initial quality assessment of the sequencing reads, their alignment with the reference genome, and the subsequent variant calling, the resulting DNA changes were annotated for facilitating the filtering and prioritization steps. Some online free sequence databases (i.e. 1000 Genome Project, NHLBI-ESP 6500 exome projects and dbSNP138), were helpful in the annotation process. Since all the patients included in this study were Italian, we checked the frequency of each variant both in 1000 Genomes Project Global and Exome Sequencing Project Global, but also in 1000 Genomes Project European. Based on those datasets we considered only variants with an allele frequency ≤0.01.

Sanger sequencing was employed to rule out false positive calls from all the filtered and prioritized DNA changes. These analyses generated a subset of variants presenting several data interpretation challenges, but also some interesting genotype–phenotype cues. For further delineating the likelihood to be disease-relevant nsSNVs, we hence established a bioinformatics framework for the assessment of each variant functional role, based on previously reported data and in silico predictions. We evaluated the performance of several different independently published methods that aim to predict the functional consequences of alleles that result in amino acid substitutions. Moreover, the impact of a sequence variant with respect to the evolutionary conservation was derived from genomic evolutionary rate profiling (GERP) score, phylogenetic P-value (PhyloP) score and PhastCons score. In our hands, those predictions aimed to hypothesize the probable impact of a particular genetic variant on function or regulation and suggested that most of the identified nsSNVs have the potential of being functionally pathogenetic.

Several free online tools (i.e. SIFT, Polyphen 2_HDIV, Polyphen2_HVAR, Provean, LRT, Mutation Taster, Mutation Assessor, FATHMM, RadialSVM, LR, CADD v1.3) were employed to assess sequence- and structure-based features. Provean gave back a prediction score for every input variant. Between the other tools, LRT failed more frequently to give a result (16 not predicted variants out of 85 different ones). Regarding our set of DNA changes, Mutation Taster and Mutation Assessor were the methods resulting respectively in the greater and in the fewer number of deleterious predictions. We therefore observed that combining multiple prediction tools provides a more even balance between sensitivity and specificity than most of the individual methods.

Besides the score-based predictions, performing molecular modeling gave us the opportunity of visually and directly testing the impact of amino acid substitutions on the proteins' specific

tertiary or quaternary structure. The location of a coding SNV with respect to the surface—interior or interface of the protein structure—could indeed influence disease manifestation.

Our NGS approach identified likely pathogenic sarcomeric variants in 58% non-syndromic patients, consistent with previous studies that have used both conventional genetic sequencing techniques(Marian et al., 2001; Arad et al., 2002; Hershberger et al., 2013), and a NGS approach in larger cohorts. In particular, in 2013 Lopes and coauthors analyzed 223 unrelated cases for 41 sarcomeric and non-sarcomeric cardiovascular genes and found that 121 patients (54%) carried alterations in 9 sarcomeric genes (including *ACTC1*, *MYPBC3*, *MYH6*, *MYH7*, *MYL2*, *MYL3*, *TNNI3*, *TNNT2* and *TPM1*) (Lopes et al., 2013). In our study we found alterations in two additional sarcomeric loci (i.e. *MYOM1* and *NEBL*) not analyzed by Lopes et al. (2013), maybe explaining our slight increased rate of cases with sarcomere gene mutations.

Inclusion of not only sarcomeric *loci*, but also of many different genes implicated in cardiomyopathies, resulted in the identification of several rare (frequency ≤0.01) DNA changes. We found that 82% of the patients harbored at least one rare nsSNV. Our results confirm the well-established role of *MYBPC3* a major gene in the HCM pathogenesis(Millat et al., 2010; Lopes et al., 2013), but also show that the HCM mutational spectrum consists mainly in missense mutations. In fact, only ~7% of the identified variants are truncating (frameshift or nonsense SNVs), and they map preferentially in *MYBPC3*. This finding suggests that *MYBPC3* molecular alterations can be both amino acid substitutions but also loss of function mutations, in contrast to the other analyzed *loci* that resulted mostly affected by missense change. We did not observe any strong mutation hot-spot, and only 8/106 nsSNVs recurred in more than one patient.

In our cohort, sarcomeric *loci* resulted as the most affected ones and, among them *MYBPC3*, *MYH6* and *MYH7* showed the main proportion of nsSNVs. These data confirm the major role of cardiac myosin binding protein C and of myosin heavy chain 7 in hypertrophic cardiomyopathy pathogenesis (Millat et al., 2010; Lopes et al., 2013), and following the results by Lopes et al. (2013) clench myosin heavy chain 6 as one of the most recently established sarcomeric HCM genes.

Eleven percent of the patients showed only sarcomere nsSNVs, while 20% of cases harbored at least one sarcomeric nsSNV with at least a desmosomal one. Moreover we found that 14% of the HCM patients displayed at least one desmosomal nsSNV but no other sarcomere change. Such a direct association of desmosomal alterations with HCM has not been described to date. In 2013, Lopes et al. (2013) analyzed by NGS a large cohort of HCM cases and found that 10% of them carried nsSNVs in titin gene (*TTN*) only in association with desmosomal or ion channel variants, but not other sarcomere ones. Furthermore, mutations in the filamin C gene (*FLNC*) have very recently been associated with HCM (Valdes-Mas et al., 2014). A proportion of our not-sarcomeric cases could then harbor additional alterations of *TTN* or *FLNC*, that were not included in the present NGS panel. Moreover, some of those cases might also have alterations of genomic region that were not covered by our NGS panel (i.e. 3'UTR, 5'UTR, ncRNAs). Future studies both of *TTN*, *FLNC* and of sequences regulating the cardiomyopathies genes' expression will shed some light on rising role of desmosomes in HCM as well as in DCM (Haas et al., 2014).

As sarcomere and desmosomes nsSNVs were the major changes associated with HCM, and we speculate that the other non-sarcomeric *loci* might have a modifying effect on HCM phenotype. As demonstrated in a recent study, non-sarcomeric variants may indeed influence the disease expression, outlining the complexity of HCM genetic basis (Lopes et al., 2014). Statistical analyses performed in our cohort, revealed that not only a higher number of nsSNVs seems to correlate with an earlier disease onset, but also that those alterations in genes for Ca^{++} homeostasis, for K^+ and Na^+ channels, and for cytoskeletal proteins can modulate HCM expression. Ion (Ca^{++}, Na^+ and K^+) channel variants correlate to an earlier disease onset and to a lower maximal wall thickness, while changes in cytoskeletal loci correlate with non-sustained ventricular tachycardia onset and low-grade diastolic dysfunction. Moreover, sarcomeric nsSNVs were correlated to a lower NYHA class. A recent study on a large HCM population, reported that the prevalence of male sex was lower in sarcomere-positive individuals (Lopes et al., 2014). We did not replicate this finding, and we found a higher proportion of males respect to women in the group of positive nsSNVs cases. The comparison of data from familial and sporadic cases generated 3 significant correlations: (i) the mean of the maximum wall thickness value of was lower in the sporadic group; (ii) non-sustained ventricular tachycardia was correlated with familial cases; (iii) the implantation of ICD resulted to be linked to a familial framework.

We also identified 49 synonymous rare variants that might act as phenotype modifiers. Additional studies about their function on mRNA transcription, splicing, transport, translation or modification are required for determining their possible non-silent role.

Among the 97 identified nsSNVs, a single variant was identified in homozygosity (i.e. *ABCC9* p.V734I in patient 21). Such a variant was already associated to a higher risk of developing precocious myocardial infarction (Minoretti et al., 2006), but patient 21 did not show that occurrence up to now.

Including both the non-syndromic and the syndromic cases, 25 out of 62 genes resulted negative to the mutation screening. Between them, there were not desmosomal *loci*, but four sarcomeric ones: *ACTC1*, *MYL3*, *TNNC1* and *TNNI3*. In line with the present results, previous studies reported a rate of heterozygous mutated patients less than 1% for *ACTC1* (Olson et al., 2000), *MYL3* (Poetter et al., 1996) and *TNNC1* (Landstrom et al., 2008). *TNNI3* alterations are expected in about 5% of HCM cases (Hershberger et al., 2013), but we did not found any mutated patient for this gene. If a clear *TNNI3* genotype–phenotype association exists, this discordance might be due to a bias in patients' selection, but of course additional studies on larger cohorts are needed to investigate this observation.

Between the 25 negative genes, there were 16 *loci* already associated with hypertrophic or dilated cardiomyopathy (*ACTC1*, *ANKRD1*, *CRYAB*, *DES*, *EYA4*, *GATAD1*, *LDB3*, *MYL3*, *MYLK2*, *MYOZ2*, *NEXN*, *PLN*, *SGCD*, *TCAP*, *TNNC1* and *TNNI3*), 3 with myopathies and neuromuscular disorders (*EMD*, *FXN* and *TAZ*), 2 with storage disorders (*PRKAG2* and *TTR*), one with ventricular tachycardia (*CASQ2*), one with arrhythmogenic right ventricular cardiomyopathy (*TMEM43*), one with Noonan syndrome (*PTPN11*) and one cardiovascular candidate gene (*CTF1*). Six of the negative loci (*ANKRD1*, *CRYAB*, *CTF1*, *EMD*, *EYA4* and *FXN*) were not included in a recent study about NGS analysis of a large cohort of HCM patients (Lopes et al., 2014).

Three patients showed syndromic HCM and the NGS analysis allowed the confirmation of their clinical diagnosis. Patient 10 was a 60-year-old woman with apparently isolated non-obstructive HCM featuring a maximum left ventricular wall thickness of 18 mm. Subsequently, she resulted heterozygous carrier of the p.N215S mutation in *GLA* gene, responsible of the X-linked recessive Fabry disease. This mutation was previously associated to a cardiac variant of the disease by Eng et al. (1993) and by Davies et al. (1993), and our data confirm this genotype–phenotype association. Patient 24 was a 6-year-old girl originally ascertained for the clinical suspect of Cantù syndrome, a syndromic form of early-onset HCM also featuring coarse face, hirsutism, persistence of fetal circulation, overgrowth of prenatal onset and mild bone dysplasia, which was recently associated with specific *ABCC9* mutations (van Bon et al., 2012). Accordingly, we found the recurrent heterozygous *ABCC9* p.R1154W mutation, which subsequently resulted de novo (Harakalova et al., 2012; van Bon et al., 2012). Finally, patient 27 showing some intellectual impairment and severe cardiac disease requiring heart transplant and limb weakness, harbored the novel p.F151fs mutation in *LAMP2*, the gene causing Danon's disease, an X-linked dominant disorder predominantly affecting cardiac muscle (Nishino et al., 2000).

5. Conclusion

By the use of NGS technology and in view of the genetic overlap between different types of cardiomyopathies, we developed a molecular test suitable for a broad series of both non-syndromic and syndromic affected patients. Moreover, despite the large amount of data coming out from an NGS protocol, we delineated a prompt informatic pipeline for the prioritization of the most likely pathogenetic variants in a clinical context.

Thanks to the possibility of analyzing many sarcomeric and non-sarcomeric *loci*, the conducted genotype–phenotype correlations represent a starting point for expanding the present results to larger cohorts and for delineating the contribution of each analyzed gene in the onset and clinical variability of HCM. In the future indeed, a broad range of molecular causes (i.e. desmosomal or other non-sarcomeric alterations) and environmental factors will need to be investigated in wider populations of sporadic and familial cases.

In conclusion, our results enlarge the mutational spectrum of hypertrophic cardiomyopathy patients with the intent of contributing to the definition of a molecular paradigm for explaining the clinical HCM diversity.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2015.11.048.

Conflict of Interest

The authors disclose any conflicts of interest that may affect the objectivity of the present work, including any financial, personal or other influences.

Acknowledgments

This work was funded by the Department of Molecular Medicine, Sapienza University of Rome (research grant AII/Dip7_2014, MED03). This work was also partially funded by the Department of Biochemical Sciences Sapienza University of Rome (prot. C26A149EC4).

References

Arad, M., Seidman, J.G., Seidman, C.E., 2002. Phenotypic diversity in hypertrophic cardiomyopathy. Hum. Mol. Genet. 11, 2499–2506.

Bottillo, I., D'Angelantonio, D., Caputo, V., Paiardini, A., Lipari, M., De Bernardo, C., Majore, S., Castori, M., Zachara, E., Re, F., Grammatico, P., 2015. Prediction and Visualization Data for the Interpretation of Sarcomeric and non-Sarcomeric DNA Variants Found in Patients with Hypertrophic Cardiomyopathy (Data in Brief submitted).

Davies, J.P., Winchester, B.G., Malcolm, S., 1993. Mutation analysis in patients with the typical form of Anderson–Fabry disease. Hum. Mol. Genet. 2, 1051–1053.

Dische, M.R., 1972. Observations on the morphological changes of the developing heart. Cardiovasc. Clin. 4, 175–191.

Eng, C.M., Resnick-Silverman, L.A., Niehaus, D.J., Astrin, K.H., Desnick, R.J., 1993. Nature and frequency of mutations in the alpha-galactosidase A gene that cause Fabry disease. Am. J. Hum. Genet. 53, 1186–1197.

Frey, N., Luedde, M., Katus, H.A., 2012. Mechanisms of disease: hypertrophic cardiomyopathy. Nat. Rev. Cardiol. 9, 91–100. http://dx.doi.org/10.1038/nrcardio.2011.159.

Haas, J., Frese, K.S., Peil, B., Kloos, W., Keller, A., Nietsch, R., Feng, Z., Muller, S., Kayvanpour, E., Vogel, B., Sedaghat-Hamedani, F., Lim, W.K., Zhao, X., Fradkin, D., Kohler, D., Fischer, S., Franke, J., Marquart, S., Barb, I., Li, D.T., Amr, A., Ehlermann, P., Mereles, D., Weis, T., Hassel, S., Kremer, A., King, V., Wirsz, E., Isnard, R., Komajda, M., Serio, A., Grasso, M., Syrris, P., Wicks, E., Plagnol, V., Lopes, L., Gadgaard, T., Eiskjaer, H., Jorgensen, M., Garcia-Giustiniani, D., Ortiz-Genga, M., Crespo-Leiro, M.G., Deprez, R.H., Christiaans, I., van Rijsingen, I.A., Wilde, A.A., Waldenstrom, A., Bolognesi, M., Bellazzi, R., Morner, S., Bermejo, J.L., Monserrat, L., Villard, E., Mogensen, J., Pinto, Y.M., Charron, P., Elliott, P., Arbustini, E., Katus, H.A., Meder, B., 2014. Atlas of the clinical genetics of human dilated cardiomyopathy. Eur. Heart J. http://dx.doi.org/10.1093/eurheartj/ehu301.

Harakalova, M., van Harssel, J.J., Terhal, P.A., van Lieshout, S., Duran, K., Renkens, I., Amor, D.J., Wilson, L.C., Kirk, E.P., Turner, C.L., Shears, D., Garcia-Minaur, S., Lees, M.M., Ross, A., Venselaar, H., Vriend, G., Takanari, H., Rook, M.B., van der Heyden, M.A., Asselbergs, F.W., Breur, H.M., Swinkels, M.E., Scurr, I.J., Smithson, S.F., Knoers, N.V., van der Smagt, J.J., Nijman, I.J., Kloosterman, W.P., van Haelst, M.M., van Haaften, G., Cuppen, E., 2012. Dominant missense mutations in ABCC9 cause Cantu syndrome. Nat. Genet. 44, 793–796. http://dx.doi.org/10.1038/ng.2324.

Hershberger, R.E., Hedges, D.J., Morales, A., 2013. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. Nat. Rev. Cardiol. 10, 531–547. http://dx.doi.org/10.1038/nrcardio.2013.105.

Jarcho, J.A., McKenna, W., Pare, J.A., Solomon, S.D., Holcombe, R.F., Dickie, S., Levi, T., Donis-Keller, H., Seidman, J.G., Seidman, C.E., 1989. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. N. Engl. J. Med. 321, 1372–1378. http://dx.doi.org/10.1056/NEJM198911163212005.

Klues, H.G., Schiffers, A., Maron, B.J., 1995. Phenotypic spectrum and patterns of left ventricular hypertrophy in hypertrophic cardiomyopathy: morphologic observations and significance as assessed by two-dimensional echocardiography in 600 patients. J. Am. Coll. Cardiol. 26, 1699–1708. http://dx.doi.org/10.1016/0735-1097(95)00390-8.

Landstrom, A.P., Parvatiyar, M.S., Pinto, J.R., Marquardt, M.L., Bos, J.M., Tester, D.J., Ommen, S.R., Potter, J.D., Ackerman, M.J., 2008. Molecular and functional characterization of novel hypertrophic cardiomyopathy susceptibility mutations in TNNC1-encoded troponin C. J. Mol. Cell. Cardiol. 45, 281–288. http://dx.doi.org/10.1016/j.yjmcc.2008.05.003.

Lopes, L.R., Zekavati, A., Syrris, P., Hubank, M., Giambartolomei, C., Dalageorgou, C., Jenkins, S., McKenna, W., Uk10k, C., Plagnol, V., Elliott, P.M., 2013. Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. J. Med. Genet. 50, 228–239. http://dx.doi.org/10.1136/jmedgenet-2012-101270.

Lopes, L.R., Syrris, P., Guttmann, O.P., O'Mahony, C., Tang, H.C., Dalageorgou, C., Jenkins, S., Hubank, M., Monserrat, L., McKenna, W.J., Plagnol, V., Elliott, P.M., 2014. Novel genotype-phenotype associations demonstrated by high-throughput sequencing in patients with hypertrophic cardiomyopathy. Heart http://dx.doi.org/10.1136/heartjnl-2014-306387.

Marian, A.J., Salek, L., Lututcuta, S., 2001. Molecular genetics and pathogenesis of hypertrophic cardiomyopathy. Minerva Med. 92, 435–451.

Maron, B.J., 2002. Hypertrophic cardiomyopathy: a systematic review. JAMA 287, 1308–1320.

Maron, M.S., Olivotto, I., Betocchi, S., Casey, S.A., Lesser, J.R., Losi, M.A., Cecchi, F., Maron, B.J., 2003. Effect of left ventricular outflow tract obstruction on clinical outcome in hypertrophic cardiomyopathy. N. Engl. J. Med. 348, 295–303. http://dx.doi.org/10.1056/NEJMoa021332.

Maron, B.J., Towbin, J.A., Thiene, G., Antzelevitch, C., Corrado, D., Arnett, D., Moss, A.J., Seidman, C.E., Young, J.B., American Heart, A., Council on Clinical Cardiology, H.F., Transplantation, C., Quality of, C., Outcomes, R., Functional, G., Translational Biology Interdisciplinary Working, G., Council on, E. and Prevention, 2006. Contemporary definitions and classification of the cardiomyopathies: an American Heart Association scientific statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. Circulation 113, 1807–1816. http://dx.doi.org/10.1161/CIRCULATIONAHA.106.174287.

Millat, G., Bouvagnet, P., Chevalier, P., Dauphin, C., Jouk, P.S., Da Costa, A., Prieur, F., Bresson, J.L., Faivre, L., Eicher, J.C., Chassaing, N., Crehalet, H., Porcher, R., Rodriguez-Lafrasse, C., Rousson, R., 2010. Prevalence and spectrum of mutations in a cohort of 192 unrelated patients with hypertrophic cardiomyopathy. Eur. J. Med. Genet. 53, 261–267. http://dx.doi.org/10.1016/j.ejmg.2010.07.007.

Minoretti, P., Falcone, C., Aldeghi, A., Olivieri, V., Mori, F., Emanuele, E., Calcagnino, M., Geroldi, D., 2006. A novel Val734Ile variant in the ABCC9 gene associated with myocardial infarction. Clin. Chim. Acta 370, 124–128. http://dx.doi.org/10.1016/j.cca.2006.02.007.

Nishino, I., Fu, J., Tanji, K., Yamada, T., Shimojo, S., Koori, T., Mora, M., Riggs, J.E., Oh, S.J., Koga, Y., Sue, C.M., Yamamoto, A., Murakami, N., Shanske, S., Byrne, E., Bonilla, E., Nonaka, I., DiMauro, S., Hirano, M., 2000. Primary LAMP-2 deficiency causes X-linked vacuolar cardiomyopathy and myopathy (Danon disease). Nature 406, 906–910. http://dx.doi.org/10.1038/35022604.

Olson, T.M., Doan, T.P., Kishimoto, N.Y., Whitby, F.G., Ackerman, M.J., Fananapazir, L., 2000. Inherited and de novo mutations in the cardiac actin gene cause hypertrophic cardiomyopathy. J. Mol. Cell. Cardiol. 32, 1687–1694. http://dx.doi.org/10.1006/jmcc.2000.1204.

O'Mahony, C., Jichi, F., Pavlou, M., Monserrat, L., Anastasakis, A., Rapezzi, C., Biagini, E., Gimeno, J.R., Limongelli, G., McKenna, W.J., Omar, R.Z., Elliott, P.M., Hypertrophic Cardiomyopathy Outcomes, I., 2014. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM risk-SCD). Eur. Heart J. 35, 2010–2020. http://dx.doi.org/10.1093/eurheartj/eht439.

Poetter, K., Jiang, H., Hassanzadeh, S., Master, S.R., Chang, A., Dalakas, M.C., Rayment, I., Sellers, J.R., Fananapazir, L., Epstein, N.D., 1996. 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. 13, 63–69. http://dx.doi.org/10.1038/ng0596-63.

Valdes-Mas, R., Gutierrez-Fernandez, A., Gomez, J., Coto, E., Astudillo, A., Puente, D.A., Reguero, J.R., Alvarez, V., Moris, C., Leon, D., Martin, M., Puente, X.S., Lopez-Otin, C., 2014. Mutations in filamin C cause a new form of familial hypertrophic cardiomyopathy. Nat. Commun. 5, 5326. http://dx.doi.org/10.1038/ncomms6326.

van Bon, B.W., Gilissen, C., Grange, D.K., Hennekam, R.C., Kayserili, H., Engels, H., Reutter, H., Ostergaard, J.R., Morava, E., Tsiakas, K., Isidor, B., Le Merrer, M., Eser, M., Wieskamp, N., de Vries, P., Steehouwer, M., Veltman, J.A., Robertson, S.P., Brunner, H.G., de Vries, B.B., Hoischen, A., 2012. Cantu syndrome is caused by mutations in ABCC9. Am. J. Hum. Genet. 90, 1094–1101. http://dx.doi.org/10.1016/j.ajhg.2012.04.014.

Watkins, H., McKenna, W.J., Thierfelder, L., Suk, H.J., Anan, R., O'Donoghue, A., Spirito, P., Matsumori, A., Moravec, C.S., Seidman, J.G., et al., 1995. Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy. N. Engl. J. Med. 332, 1058–1064. http://dx.doi.org/10.1056/NEJM199504203321603.