Brief Communication
Diagnostic Genetics

Ann Lab Med 2025;45:96-100
https://doi.org/10.3343/alm.2024.0201
ISSN 2234-3806 • eISSN 2234-3814

ANNALS OF LABORATORY MEDICINE

Reanalysis of Next-generation Sequencing Data in Patients With Hypertrophic Cardiomyopathy: Contribution of Spliceogenic *MYBPC3* Variants in an Italian Cohort

Silvia Caroselli, M.Sc.[1,2], Marco Fabiani, Ph.D.[1,3], Caterina Micolonghi, M.D.[1], Camilla Savio, M.Sc.[4], Giacomo Tini, M.D., Ph.D.[5], Beatrice Musumeci, M.D.[5], Erika Pagannone, M.D.[4,5], Aldo Germani, Ph.D.[5], Fabio Libi, M.Sc.[4], Vincenzo Visco, M.D.[4,5], Antonio Pizzuti, M.D.[1,6], Camillo Autore, M.D.[7], Simona Petrucci, M.D., Ph.D.[4,5,6], Speranza Rubattu, M.D.[5,8], and Maria Piane, Ph.D.[4,5]

[1]Department of Experimental Medicine, Sapienza University of Rome, Rome, Italy; [2]Juno Genetics, Reproductive Genetics, Rome, Italy; [3]ALTAMEDICA, Human Genetics, Rome, Italy; [4]Sant'Andrea University Hospital, Rome, Italy; [5]Department of Clinical and Molecular Medicine, Sapienza University of Rome, Rome, Italy; [6]Casa Sollievo della Sofferenza Foundation, San Giovanni Rotondo, Italy; [7]San Raffaele Cassino, Cassino, Italy; [8]IRCCS Neuromed, Pozzilli, Italy

Hypertrophic cardiomyopathy (HCM) is a genetic cardiac muscle disease characterized by clinical and genetic heterogeneity. Genetic testing can reveal the presence of disease-causing variants in genes encoding sarcomere proteins. However, it yields inconclusive or negative results in 40–60% of HCM cases, owing to, among other causes, technical limitations such as the inability to detect pathogenic intronic variants. Therefore, we aimed to increase the diagnostic yield of molecular analysis for HCM by improving the *in-silico* detection of intronic variants in *MYBPC3* that may escape detection by algorithms normally used with tagged diagnostic panels. We included 142 HCM probands with negative results in Illumina TruSight Cardio panel analysis, including exonic regions of 174 cardiomyopathy genes. Raw data were re-analyzed using existing bioinformatics tools. The spliceogenic variant c.1224-80G>A was detected in three patients (2.1%), leading us to reconsider their molecular diagnosis. These patients showed late onset and mild symptoms, although no peculiar phenotypic characteristics were shared. Collectively, rare spliceogenic *MYBPC3* variants may play a role in causing HCM, and their systematic detection should be performed to provide more comprehensive solutions in genetic testing using multigenic panels.

Key Words: High-throughput nucleotide sequencing, Hypertrophic cardiomyopathy, Introns, Loss of function mutation, RNA splicing

Received: April 24, 2024
Revision received: July 23, 2024
Accepted: September 11, 2024
Published online: October 2, 2024

Corresponding author:
Silvia Caroselli, M.Sc.
Department of Experimental Medicine, Sapienza University of Rome, Viale Regina Elena, 324, Rome 00161, Italy
E-mail: silvia.caroselli@uniroma1.it

Hypertrophic cardiomyopathy (HCM), the most common inherited cardiac disease, with a prevalence of one in 200–500 individuals [1], is characterized by left ventricular (LV) hypertrophy unexplained by secondary causes and a broad clinical spectrum. HCM is associated with considerable morbidity and mortality [2]. For personal and family risk management, molecular testing can confirm the genetic origin of HCM. HCM generally follows Mendelian inheritance, predominantly in an autosomal dominant fashion. However, it can exhibit both incomplete penetrance and variable expressivity, resulting in apparently sporadic cases [3].

HCM is primarily caused by pathogenic variants (PVs) in sarcomeric core genes, followed by variants in non-sarcomeric genes with varying levels of supportive evidence [4]. HCM-causing variants are typically identified using large gene panels based on targeted next-generation sequencing (tNGS), which facilitates differential diagnosis with other heart diseases. The overall detection rate varies depending on patient baseline characteristics (i.e., demographic, family, and clinical history), ranging from 20–30% in sporadic cases to 40–60% in familial HCM. This implies that genetic analysis yields inconclusive or negative results in 40% to >60% of HCM cases [5]. Several clinical (e.g., phenocopies/misdiagnoses) and technical (e.g., unidentified genes and/or variants) factors may contribute to this relatively low diagnostic yield. tNGS panels target exonic regions and are unable to effectively detect intronic positions. Notably, *MYBPC3* accounts for most HCM cases, and PVs in this gene are generally loss-of-function (LoF) variants, leading to nonsense-mediated decay (NMD) and subsequent haploinsufficiency. In this scenario, variants that alter RNA splicing (i.e., spliceogenic variants) can produce aberrant transcripts that also result in NMD. They are increasingly recognized as causative of HCM in several cohorts [6-9], potentially representing >3% of HCM probands [7]. However, these studies screened for previously characterized spliceogenic *MYBPC3* variants using NGS covering all *MYBPC3* introns (6, 7) or using whole-exome sequencing (WES) including 100-bp intronic regions beyond intron-exon boundaries [8, 9]. Torrado, *et al*. [6] focused on the discovery and characterization of rare unknown variants using several deep learning-based splicing algorithms and functional assays at the mRNA and protein levels.

Here, we re-evaluated raw data obtained with a cardiac tNGS panel to improve the molecular HCM diagnostic rate. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Sant'Andrea Hospital (Rome, Italy; approval No.: 42 of September 28, 2007).

Between 2021 and 2023, 142 consecutive unrelated patients diagnosed with HCM had negative genetic testing results. This cohort of HCM patients was characterized by the absence of a family history and a broad spectrum of clinical features. The conventional molecular analysis covered exonic regions of 174 genes associated with hereditary heart diseases, including HCM, and consisted of (i) genomic DNA extraction from peripheral blood (DNeasy Blood & Tissue Kit and QIAamp DNA Blood Mini Kit, Qiagen, Hilden, Germany), (ii) library preparation (TruSight Cardio panel, Illumina, San Diego, CA, USA), (iii) sequencing on a MiniSeq system (Illumina), and (iv) standard bioinformatics analysis using BaseSpace platforms (i.e., BaseSpace Sequence Hub and BaseSpace Variant Interpreter, Illumina). As a default filter setting, the padding size was 20 bp on both sides of exonic regions.

To include the annotation of deeper intronic variants as well, sequencing data were re-analyzed based on an *ad-hoc* parallel workflow (Fig. 1). Through a first approach for the screening of known disease-causing variants, a list of intronic variants previously described and interpreted as pathogenic in the literature [6] was compiled, considering up to 100 bp of flanking sequenced regions (Supplemental Data Table S1). A corresponding *in-silico* hotspot for these PVs was built by annotating their genomic coordinates using Integrative Genomics Viewer v2.16.2 [10] and was later applied to raw sequencing data (i.e., BAM files) obtained from the BaseSpace Sequence Hub. PVs thus identified and visualized were validated using Sanger sequencing. In a second approach used for identifying unknown variants, all *MYBPC3* intronic variants detected by tNGS were considered. Common intronic variants were filtered out, and rare variants were analyzed using the SpliceAI software [11] (a web-based interface) to evaluate their probability of being splice-altering. In the latter case, a score of >0.8 indicated an increased damaging impact with high precision [12]. Subsequently, mRNA-level evaluation was performed to assess the impact of emerging variants (i.e., SpliceAI score >0.5) on RNA splicing. Upon identifying a potential splice-altering variant, it was classified according to the American College of Medical Genetics and Genomics (ACMG) criteria [13] and validated using Sanger sequencing. By applying the filtering approach, 23 additional intronic variants were identified in our cohort. Considering a frequency threshold of 1% in the general population, 13 were common, and 10 were rare variants. All rare variants had a SpliceAI score <0.5, a more permissive cutoff to discriminate potential splice-altering variants from non-altering variants [12] (Supplemental Data Table S2). Therefore, none of these variants required further investigation.

The hotspot approach helped detect the c.1224-80G>A variant in *MYBC3* in three patients (2.1%; 95%CI: 0.7–6.0%). No additional intronic variants were identified. According to the ACMG criteria (PM2 and PP5), the c.1224-80G>A variant is classified as likely pathogenic. This variant has been documented previously in four HCM cohorts: N=1/93 [7], N=1/720 [8], N=3/1644 [9], and N=6/3437 [6]. The frequency observed in our cohort corresponds with those reported in the literature, albeit influenced by the population considered.

From a functional perspective, RNA splicing requires the rec-

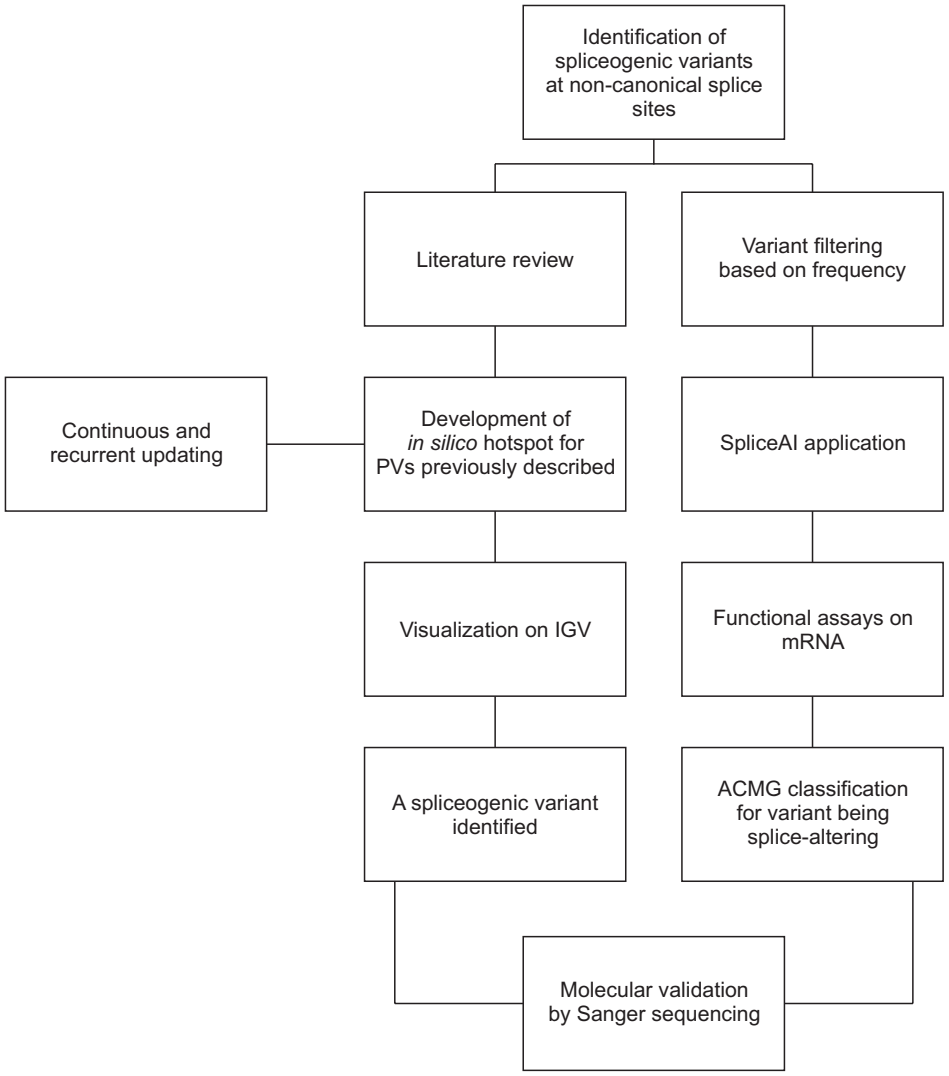

Fig. 1. Parallel multistep workflow for diagnostic setting. Left side (first approach): development and application of an *in-silico* hotspot for previously reported PVs; right side (second approach): *in-silico* filtering and *in-vitro* analysis of unknown variants.
Abbreviations: PVs, pathogenic variants; IGV, Integrative Genomics Viewer; ACMG, American College of Medical Genetics and Genomics.

ognition of splice donor and acceptor sites located at the boundaries of introns, generally represented by the canonical dinucleotides "GT" and "AG," respectively. Briefly, the c.1224-80G>A variant potentially generates a novel splice acceptor site ("AG") in intron 13 of *MYBPC3*, leading to the retention of 78 nucleotides upstream of exon 14 in the transcript (Fig. 2). The splice-altering property of this variant, predicted *in silico* and confirmed in functional studies, includes a SpliceAI Δscore of 0.94 [6] and results in an *in-frame* insertion of 26 codons, diminishing *MYBPC3* transcripts by 50% [7]. Additionally, the splicing of micro-exons, protein-coding regions spanning 3 to 30–50 bp, warrants further attention [14]. Particularly, *MYBPC3* contains three micro-exons, including the conserved exon 14 [15, 16]. Micro-exons, with their weak splicing signals, rely on robust signals from flanking intron regions [17, 18]. Therefore, these regions prove critical as several damaging variants have been identified [19]. The c.1224-80G>A variant establishes a new, stronger acceptor splice site in intron 13, potentially favored over the canonical site [7], thus altering the splicing of micro-exon 14 and resulting in an unstable transcript.

Clinical data, such as comprehensive cardiological assessments and familial/personal histories, were collected from the three individuals carrying the c.1224-80G>A variant (Table 1). Their ages ranged from 56–74 years, with two reporting a familial history of HCM. Maximal LV wall thickness varied from 13–27 mm; patient #1 (P1) exhibited systolic anterior motion of the mitral valve (SAM), and P2 experienced left ventricular outflow tract (LVOT) obstruction with late onset. These patients did not share notable phenotypic characteristics. Overall, this variant might be associated with mild hypertrophy, as supported by prior

https://doi.org/10.3343/alm.2024.0201

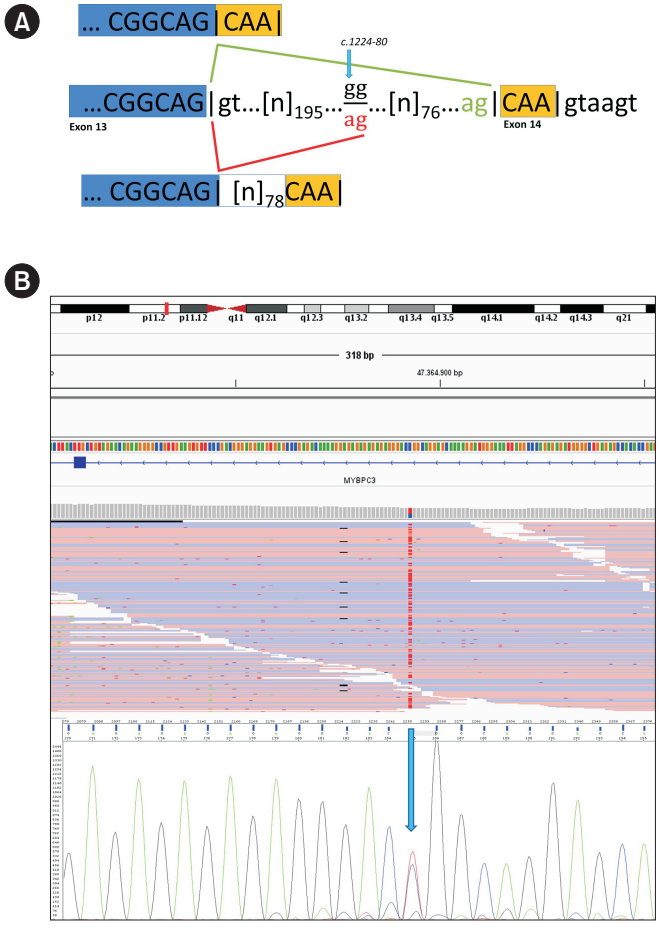

Fig. 2. *MYBPC3* exon 14 splicing. (A) Schematic representation of the nucleotide sequences of exon 13, intron 13, and exon 14. Canonical splicing of intron 13 with the juxtaposition of the 3′-exon 13/5′-exon 14 junction is shown above (green lines). The variant c.1224-80G > A (in red) creates a cryptic splice acceptor site in intron 13 (dinucleotide "AG"), which leads to the juxtaposition of 3′-exon 13 with a locus upstream of 5′-exon 14 (bottom, red lines), inserting 78 intronic nucleotides into the transcript. (B) Representation of aligned NGS reads (top) and Sanger electropherogram (bottom) corresponding to the heterozygous variant c.1224-80G > A (arrow).
Abbreviation: NGS, next-generation sequencing.

reports. Janin, *et al*. [7] described a 34-year-old man with sporadic asymptomatic HCM and an LV wall thickness of 17 mm, whereas Singer, *et al*. [8] reported an asymptomatic 28-year-old man with a family history of HCM.

To the best of our knowledge, we report the first-known application of a strategy using a multigene panel for detecting pathogenic intronic *MYBPC3* variants, previously identified only through intron sequencing or WES methodologies. Specifically, the filtering approach based on both frequency and *in-silico* predictions helped exclude new potential intronic variants. The design and application of a hotspot for previously described intronic PVs yielded additional evidence for their prevalence, which is crucial for pathogenicity interpretation. Notably, this facilitated accurate genetic diagnoses in some patients with HCM who had shown gene elusiveness in standard analyses, increasing the detection rate of the cardio panel from 20.2% (N = 36/178) to 21.9% (N = 39/178) in our study population. The clinical utility of our findings would have been enhanced by complete intronic sequencing of *MYBPC3* and other HCM-associated genes, and clinical validity could be supported further by mRNA sequencing to assess the functional impact of identified intronic variants. However, these analyses were impractical owing to practical limitations, such as sample availability, quality, cost, and resource constraints.

Collectively, our findings are consistent with the evidence in the literature and support broader *MYBPC3* analysis to provide health professionals and manufacturers with better information for the systematic recognition of additional splice-disrupting variants at non-canonical splice sites.

Table 1. Clinical characteristics of patients with HCM carrying the spliceogenic c.1224-80G > A variant in *MYBPC3*

Characteristics	Patient 1	Patient 2	Patient 3
Age at first observation (yrs)	56	74	60
Sex	Female	Male	Male
Familial history	No	Yes	Yes
Symptoms			
NYHA class	-	II	-
Imaging			
LV thickness (mm)	13	27	14
Obstruction (type)	No (SAM)	Yes (LVOT gradient = 50 mmHg)	No

Abbreviations: HCM, hypertrophic cardiomyopathy; NYHA, New York Heart Association; LV thickness, left ventricular thickness; SAM, systolic anterior motion of the mitral valve; LVOT gradient, left ventricular outflow tract gradient.

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via https://doi.org/10.3343/alm.2024.0201

ACKNOWLEDGEMENTS

None.

AUTHOR CONTRIBUTIONS

Conceptualization: Rubattu S, Piane M, Petrucci S. Methodology: Caroselli S, Fabiani M, and Savio C. Investigation: Micolonghi C, Tini G, Musumeci B, Pagannone E, Germani A, and Libi F. Data curation: Caroselli S. Visualization: Caroselli S. Project administration: Rubattu S and Piane M. Supervision: Visco V, Pizzuti A, and Autore C. Writing – original draft: Caroselli S. Writing – review & editing: Rubattu S, Piane M, and Petrucci S.

CONFLICTS OF INTEREST

None declared.

RESEARCH FUNDING

None declared.

REFERENCES

1. Maron BJ and Maron MS. Hypertrophic cardiomyopathy. Lancet 2013;381:242-55.
2. Zampieri M, Salvi S, Fumagalli C, Argirò A, Zocchi C, Del Franco A, et al. Clinical scenarios of hypertrophic cardiomyopathy-related mortality: relevance of age and stage of disease at presentation. Int J Cardiol 2023;374:65-72.
3. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, et al. 2023 ESC guidelines for the management of cardiomyopathies. Eur Heart J 2023;44:3503-626.
4. Mazzarotto F, Olivotto I, Boschi B, Girolami F, Poggesi C, Barton PJR, et al. Contemporary insights into the genetics of hypertrophic cardiomyopathy: toward a new era in clinical testing? J Am Heart Assoc 2020;9:e015473.
5. Walsh R, Buchan R, Wilk A, John S, Felkin LE, Thomson KL, et al. Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes. Eur Heart J 2017;38:3461-8.
6. Torrado M, Maneiro E, Lamounier Junior A, Fernández-Burriel M, Sánchez Giralt S, Martínez-Carapeto A, et al. Identification of an elusive spliceogenic *MYBPC3* variant in an otherwise genotype-negative hypertrophic cardiomyopathy pedigree. Sci Rep 2022;12:7284.
7. Janin A, Chanavat V, Rollat-Farnier PA, Bardel C, Nguyen K, Chevalier P, et al. Whole MYBPC3 NGS sequencing as a molecular strategy to improve the efficiency of molecular diagnosis of patients with hypertrophic cardiomyopathy. Hum Mutat 2020;41:465-75.
8. Singer ES, Crowe J, Holliday M, Isbister JC, Lal S, Nowak N, et al. The burden of splice-disrupting variants in inherited heart disease and unexplained sudden cardiac death. NPJ Genom Med 2023;8:29.
9. Lopes LR, Barbosa P, Torrado M, Quinn E, Merino A, Ochoa JP, et al. Cryptic splice-altering variants in *MYBPC3* are a prevalent cause of hypertrophic cardiomyopathy. Circ Genom Precis Med 2020;13:e002905.
10. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24-6.
11. SpliceAI Lookup. Broad Institute, https://spliceailookup.broadinstitute.org/#. Updated on March 29, 2024.
12. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell 2019;176:535-48.e24.
13. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.
14. Ustianenko D, Weyn-Vanhentenryck SM, Zhang C. Microexons: discovery, regulation, and function. Wiley Interdiscip Rev RNA 2017;8: 0.1002/wrna.1418.
15. Carrier L, Bonne G, Bährend E, Yu B, Richard P, Niel F, et al. Organization and sequence of human cardiac myosin binding protein C gene (MYBPC3) and identification of mutations predicted to produce truncated proteins in familial hypertrophic cardiomyopathy. Circ Res 1997;80:427-34.
16. Oakley CE, Hambly BD, Curmi PMG, Brown LJ. Myosin binding protein C: structural abnormalities in familial hypertrophic cardiomyopathy. Cell Res 2004;14:95-110.
17. Sterner DA and Berget SM. In vivo recognition of a vertebrate mini-exon as an exon-intron-exon unit. Mol Cell Biol 1993;13:2677-87.
18. Carlo T, Sterner DA, Berget SM. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. RNA 1996;2:342-53.
19. Frank-Hansen R, Page SP, Syrris P, McKenna WJ, Christiansen M, Andersen PS. Micro-exons of the cardiac myosin binding protein C gene: flanking introns contain a disproportionately large number of hypertrophic cardiomyopathy mutations. Eur J Hum Genet 2008;16:1062-9.

https://doi.org/10.3343/alm.2024.0201