Matters of the heart: genetic and molecular characterisation of cardiomyopathies

Posafalvi, Anna

Matters of the heart: genetic and molecular characterisation of cardiomyopathies

Pósafalvi Anna

The work described in this thesis was supported by the University Medical Center Groningen, the Jan Kornelis de Cock Foundation, the NutsOhra Foundation and the Netherlands Heart Foundation.

Printing of this thesis was supported by the Graduate School of Medical Sciences and the University Library, University of Groningen, Groningen, the Netherlands.

Cover photograpy: ©Anna Posafalvi

Design: dreamed of by Anna Posafalvi, dreams made come true by Joanna Smolonska

Layout and printing:  Lovebird Design & Printing Solutions

ISBN:
978-90-367-7767-4 (printed)
978-90-367-7766-7 (electronic)

university of groningen

Matters of the heart: genetic and molecular characterisation of cardiomyopathies

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 20 April 2015 at 16:15 hours

by

Anna Posafalvi

born on 17 June 1986
in Debrecen, Hungary

Supervisor
Prof. RJ Sinke

Co-supervisor
Dr. JDH Jongbloed

Assessment committee
Prof. MH Breuning
Prof. RA de Boer
Prof. RMW Hofstra

Drága Nagyapámnak...

To my dearest grandfather...

"Wheresoever you go, go with all your heart."

(Confucius)

Paranymphs

Ena Sokol
Eva Teuling

TABLE OF CONTENTS

…about cardiomyopathies in a nutshell

The disease

Cardiomyopathy is an insidious disease of the heart muscle (myocardium) leading to decreased pumping capacity, and resulting in a wide range of symptoms. These range from mild (dizziness, fatigue, chest pain or oedema) to severe (heart failure, arrhythmia, embolism, or even sudden death).

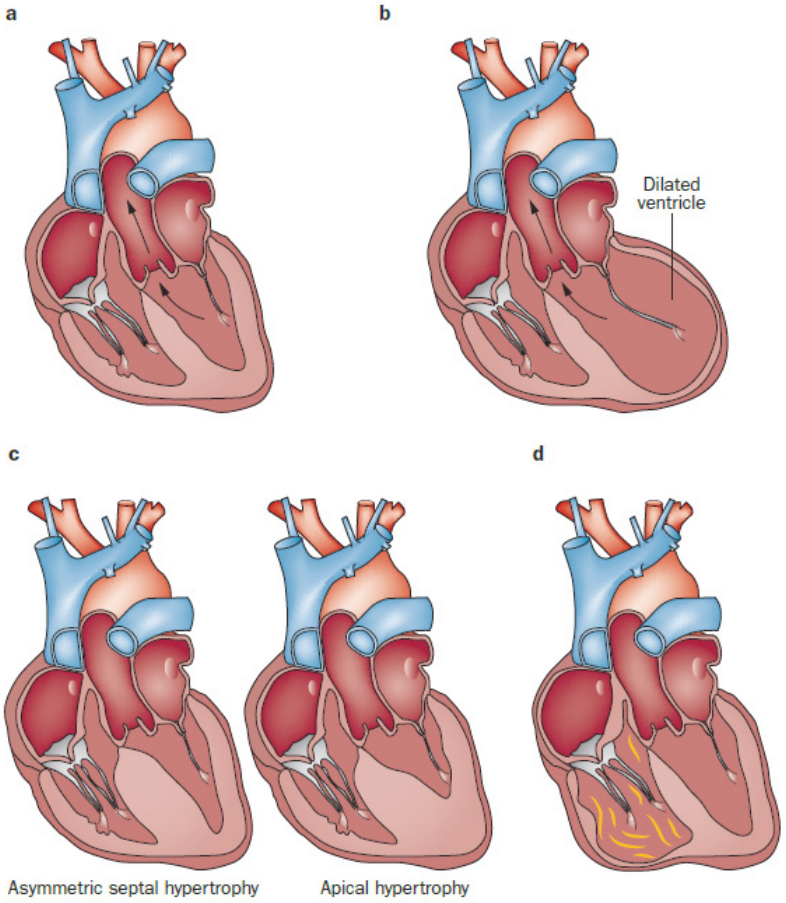

Figure 1. Schematic cross-section of a healthy heart (a) and hearts with DCM (b), HCM (c) and ARVC (d)
In dilated cardiomyopathy (DCM), the left ventricle becomes enlarged with a thin, weakened muscle wall, and is unable to generate enough pumping force during contractions; the myocardium is thickened in hypertrophic cardiomyopathy (HCM); in arrhythmogenic right ventricular cardiomyopathy (ARVC) fibrofatty infiltration of the myocardium leads to arrhythmia.
Figure published by Wilde & Behr, Nature Reviews Cardiology, 2013; used with permission.
For more information on cardiomyopathy types, see box 1.

Clinical diagnosis and treatment guidelines

When the symptoms of cardiomyopathy appear, the **diagnosis of the disease** is most frequently made by electrocardiogram (ECG), and non-invasive imaging techniques such as an X-ray of the chest, echocardiography (imaging of the heart with ultrasound), or MRI. In addition, patients receive a general medical examination combined with a simple blood test (measuring, for instance, molecular markers of heart failure or kidney function). Less regularly, cardiac catheterization or coronary angiography is used. Both these methods are "minimally invasive", only a thin tube is inserted in one of the biggest veins of the body and threaded to the heart, instead of an open surgery. These help physicians acquire either a myocardial biopsy for further experimental analysis or enough information to exclude potential blockage (stenosis) of the heart and the coronary blood vessels.

The **therapy** for cardiomyopathy largely depends on the disease type and the severity of the symptoms. Therapy aims at slowing down the progression of the disease or at disease prevention in susceptible individuals through life style changes and medical treatment using different antihypertensive, antiarrhythmic, diuretic or anticoagulant drugs (e.g., ACE inhibitors or calcium antagonists). In serious cases of arrhythmia, the implementation of an ICD (a small, implantable defibrillator) or a pacemaker may be the solution. Heart transplantation is only considered as a last resort in patients with end-stage heart failure.

The genetic causes

Even though there are several environmental factors that may trigger the onset of cardiomyopathy (viral infections, the use of certain drugs, alcoholism, and other cardiovascular conditions, as well as certain systemic disorders), we often see the disease running in families (30-50% of ARVC and DCM cases; see box 1 for definitions). Most of these familial cardiomyopathy cases can be explained by an **autosomal dominant** (AD) inheritance pattern. To date, about 76 genes are known to be involved in different types of cardiomyopathy, which often show considerable genetic overlap (figure 2), and the majority of these 76 genes show AD inheritance. Additionally, a few genes, such as *DMD*, *EMD*, *GLA*, *LAMP2*, or *TAZ*, are involved in the **X-linked** form of the disease. Exceptionally, **autosomal recessive** inheritance is also observed. These patients usually exhibit more severe symptoms, and the disease generally begins in infancy or early childhood (paediatric cardiomyopathies; the genes involved include *ANO5, MYL2, PKP2, TNNI3*).

Although it is also known that abnormalities in **mitochondrial DNA** can contribute to the pathogenesis of different cardiomyopathies (e.g., mutations of *MTTL1*), this has not yet been extensively studied. The possible complex, **oligogenic** or multifactorial causes for cardiomyopathies have also not been investigated in detail, nor have the potential roles of **risk alleles** of lower effect size, **copy number variations** (such as those including the *BAG3* or *PRDM16* genes), or **microRNAs**. To date, a significant proportion of familial cardiomyopathies (about 30-40% of HCM, 40-50% of ARVC, and around 50% of DCM cases; see box 1 for definitions) remain genetically unexplained.

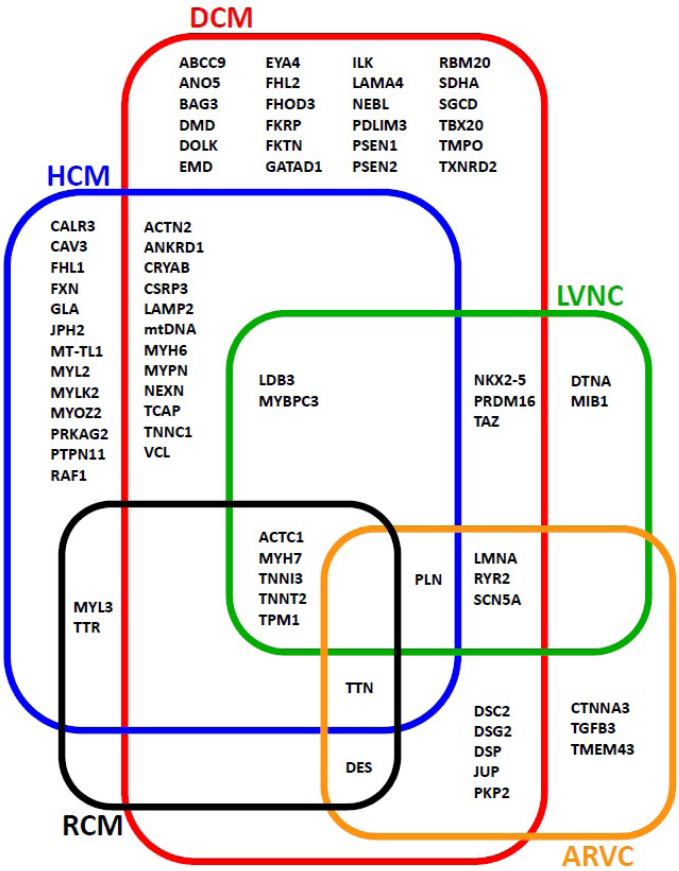

Figure 2. Cardiomyopathy disease genes and the genetic overlap between subtypes of the disease (updated from Jongbloed at el, EOMD 2011)

Not only is there considerable phenotypic overlap between the subtypes of cardiomyopathy, many genes are also involved in multiple forms of the disease. The official full names of the abbreviated genes, according to OMIM, are listed in appendix 1.

Types of cardiomyopathy

There are various forms of cardiomyopathy, each with different underlying causes for the insufficient circulation. The cardiomyopathies investigated in this thesis include:

1. **dilated cardiomyopathy (DCM):** one or both of the ventricles (in most cases only the left one) become enlarged with a thin, weakened muscle wall unable to generate enough pumping force during contractions (figure 1)
2. **arrhythmogenic right ventricular cardiomyopathy (ARVC):** the replacement of the degenerating myocardium with scar (fibrofatty) tissue results in disturbed electrical signals and conduction in the heart (arrhythmia)
3. **hypertrophic cardiomyopathy (HCM):** a thickened myocardium due to abnormal growth and arrangement (hypertrophy and disarray) of muscle fibres results in smaller chamber volume and sometimes blocks the blood flow (obstruction)
4. **restrictive cardiomyopathy (RCM):** due to their stiffness, the ventricles do not get refilled with enough blood during relaxation, hence the heart cannot supply the organs with sufficient circulation during contraction
5. **left-ventricular non-compaction cardiomyopathy (LVNC):** the wall of the left ventricle is spongiform, characterized by a meshwork of muscle fibres
6. **peripartum cardiomyopathy (PPCM):** a special form of dilated cardiomyopathy that becomes manifest towards the end of pregnancy or within a few months following delivery
7. **paediatric cardiomyopathy:** this type of cardiomyopathy becomes manifest in infancy or early childhood, and is usually characterized by more severe symptoms and worse outcomes than when the disease manifests in adulthood (from a structural-functional point of view, most frequently it is DCM>HCM>RCM>ARVC)

Our methods

Candidate gene screening

- **Sanger-sequencing:** This method of DNA-sequencing allows us to detect single nucleotide changes and small indels of DNA fragments with an average size of 400-500 base pairs. It can be used for screening candidate genes in a large cohort of patients, as well as for segregation analysis of a variant within a family, or for confirmation of DNA-variations detected by high-throughput sequencing.

Disease gene mapping

- **haplotype sharing test (HST):** An ideal, SNP-genotyping-based method for small cardiomyopathy families, who are usually not suitable for classical linkage analysis. With this method, we aim to identify chromosomal regions shared among affected family members, hypothesizing that the highest chance of finding the mutation is in the largest shared region of the family. We use this method as a filtering step in exome sequencing data analysis – if a variant is located in the 2nd largest shared haplotype of 10 cM, it is more likely to be causative than a variant located in the 57th largest shared haplotype of only 0.1 cM.

High-throughput sequencing

- **exome sequencing:** Sequencing all coding parts (exons) of all genes (about 1% of the genome). Though costly and requiring intensive data analysis, this method is suitable for identifying private coding mutations of novel disease genes in families with an unknown genetic cause of cardiomyopathy.
- **gene-panel based (targeted) sequencing:** High-throughput sequencing of a DNA sample previously enriched for the small set of genes we are interested in. Since this method results in very high coverage across the regions of interest and high data quality, it has recently been implemented in routine diagnostics.

The challenges we face

Identifying a novel disease gene carrying the **heterozygous** causal variant (heterozygous because of the dominant inheritance) is usually more challenging than working on a recessive disease, but there are also other complications to be considered in our research.

Cardiomyopathy is, in general, a **late onset disease**. For example, DCM usually begins between 20 and 50 years of age, while most ARVC patients are diagnosed before 40 years of age. Thus, low penetrance of the disease at young age makes it difficult to make the genetic diagnosis in a family as the disease status of young relatives is uncertain (partly due to the variety in the nature and severity of the symptoms). Furthermore, **phenocopies** also occur, with family members having comparable symptoms due to an independent cause (e.g. developing disease on the basis of another, often non-genetic, cardiovascular event: coronary artery disease). In consequence, the medical diagnosis of cardiomyopathy is based on exclusion criteria and performing segregation analysis for a putative pathogenic variant in families without being absolutely sure of the healthy/affected status of the screened individuals can be complicated.

Since cardiomyopathy can be so difficult to diagnose, and because the chances of a successful treatment rapidly decline with time, our aims are (1) to obtain an **early (molecular) diagnosis** of the inherited form of the disease before severe symptoms become manifest, and (2) to enable **preventive treatment** (including life-style changes as well as medical treatment if necessary) of the endangered individuals, combined with regular, thorough cardiological check-ups.

Recommended literature

website of the National Heart, Lung, and Blood Institute, health topic on cardiomyopathies: http://www.nhlbi.nih.gov/health/health-topics/topics/cm/

website of the Children's Cardiomyopathy Foundation: http://www.childrenscardiomyopathy.org/site/main_brochure.htm

Herschberger RE, Lindenfeld J, Mestroni L et al: Genetic evaluation of cardiomyopathy – a Heart Failure Society of America guideline. J Cardiac Fail 2009;15:83-97

Wilde AA & Behr ER: Genetic testing for inherited cardiac disease. Nat Rev Cardiol 2013;10:571-83

Teekakirikul P, Kelly MA, Rehm HL et al: Inherited cardiomyopathies – Molecular genetics and clinical genetic testing in the postgenomic era. J Mol Diagn 2013;15:158-70

Jongbloed JD, Pósafalvi A, Kerstjens-Frederikse WS et al: New clinical molecular diagnostic methods for congenital and inherited heart disease. Expert Opin Med Diagn 2011;5:9-24

Posafalvi A, Herkert JC, Sinke RJ et al: Clinical utility gene card for: dilated cardiomyopathy (CMD). Eur J Hum Genet 2013;21 doi: 10.1038/ejhg.2012.276

Te Rijdt WP, Jongbloed JD, de Boer RA et al: Clinical utility gene card for: arrhythmogenic right ventricular cardiomyopathy (ARVC). Eur J Hum Genet 2014;22. doi: 10.1038/ejhg.2013.124

Udeoji DU, Philip KJ, Morrissey RP et al: Left ventricular noncompaction cardiomyopathy: updated review. Ther Adv Cardiovasc Dis 2013;7:260-73

Caleshu C, Sakhuja R, Nussbaum RL et al: Furthering the link between the sarcomere and primary cardiomyopathies: restrictive cardiomyopathy associated with multiple mutations in genes previously associated with hypertrophic or dilated cardiomyopathy. Am J Med Genet A 2011;155A:2229-35

Peled Y, Gramlich M, Yoskovitz G et al: Titin mutation in familial restrictive cardiomyopathy. Int J Cardiol 2014;171:24-30

Wooten EC, Hebl VB, Wolf MJ et al: Formin homology 2 domain containing 3 variants associated with hypertrophic cardiomyopathy. Circ Cardiovasc Genet 2013;6:10-8

Chang B, Nishizawa T, Furutani M et al: Identification of a novel TPM1 mutation in a family with left ventricular noncompaction and sudden death. Mol Genet Metab 2011;102:200-6

Luxán G, Casanova JC, Martínez-Poveda B et al: Mutations in the NOTCH pathway regulator MIB1 cause left ventricular noncompaction cardiomyopathy. Nat Med 2013;19:193-201

Purevjav E, Varela J, Morgado M et al: Nebulette mutations are associated with dilated cardiomyopathy and endocardial fibroelastosis. J Am Coll Cardiol 2010;56:1493-502

Arndt AK, Schafer S, Drenckhahn JD et al: Fine mapping of the 1p36 deletion syndrome identifies mutation of PRDM16 as a cause of cardiomyopathy. Am J Hum Genet 2013;93:67-77

Ohno S, Omura M, Kawamura M et al: Exon 3 deletion of RYR2 encoding cardiac ryanodine receptor is associated with left ventricular non-compaction. Europace 2014;16:1646-54

Pinto JR, Yang SW, Hitz MP et al: Fetal cardiac troponin isoforms rescue the increased Ca2+ sensitivity produced by a novel double deletion in cardiac troponin T linked to restrictive cardiomyopathy: a clinical, genetic, and functional approach. J Biol Chem 2011;286:20901-12

van Hengel J, Calore M, Bauce B et al: Mutations in the area composita protein αT-catenin are associated with arrhythmogenic right ventricular cardiomyopathy. Eur Heart J 2013;34:201-10

Pruszczyk P, Kostera-Pruszczyk A, Shatunov A et al: Restrictive cardiomyopathy with atrioventricular conduction block resulting from a desmin mutation. Int J Cardiol 2007;117:244-53

OUTLINE OF THIS THESIS

The aims of this thesis are (1) to provide a better understanding of the genetic background and the molecular pathomechanism of familial cardiomyopathies, (2) to identify novel disease genes in unsolved families, and (3) to improve the existing methods of molecular diagnostic testing.

Chapter 1 is a detailed introduction to the field of cardiogenetics. This chapter **reviews** congenital and late onset heart diseases (the latter referring to cardiomyopathies and arrhythmia syndromes), categorizes the genes involved in the different types of heritable heart diseases, and thoroughly describes the research methods with special attention paid to their potential future diagnostic applications in cardiovascular diseases.

The subsequent chapters contain experimental data and are subdivided based on the research methods used.

In ***chapter 2***, we applied the classical **candidate gene screening** approach Sanger sequencing. We were interested if (and to what extent) the known DCM gene *RBM20* contributes to the genetic background of the disease in Dutch patients (*2.1*). In addition, we hypothesized that the desmosomal *PLEC* gene may play a role in the development of ARVC. In an attempt to prove this, we studied the clustering of sequence variations in patients compared to that in a healthy control population (*2.2*).

High-throughput sequencing is a recent technological development that is revolutionizing the science of genetics. We applied two different experimental designs of this method to elucidate genetic causes for cardiomyopathies. In ***chapter 3***, we have described families where mutations in known cardiomyopathy genes had been excluded, and we successfully applied **exome sequencing** to identify novel disease genes in both autosomal dominant (*3.1*) and recessive (*3.2*) cardiomyopathies, while *3.3* is an interesting case report on a family suffering from both forms of the disease.

In ***chapter 4***, we applied targeted enrichment of DNA samples to a set of well-defined candidate disease genes. We address the applicability and the quantitative advantages of **targeted sequencing** in routine diagnostics for a cohort of 252 unselected cardiomyopathy patients in *4.1*, while report our findings on targeted sequencing of PPCM/DCM families in *4.2*.

The work described in this thesis is then discussed in a broader context, and future perspectives for the use of high-throughput sequencing in research and diagnostic settings, as well as potential research directions in the field of cardiogenetics, are presented in ***chapter 5.***

APPENDIX 1

List of cardiomyopathy genes:

(official abbreviations and names of genes included in figure 2 of the preface)

ABCC9	ATP-binding cassette, subfamily C (CFTR/MRP), member 9
ACTC1	actin, alpha, cardiac muscle 1
ACTN2	actinin, alpha 2
ANKRD1	ankyrin repeat domain 1 (cardiac muscle)
ANO5	anoctamin 5
BAG3	BCL2-associated athanogene 3
CALR3	calreticulin 3
CAV3	caveolin 3
CRYAB	crystalline, alpha B
CSRP3	cysteine and glycine-rich protein 3 (cardiac LIM protein)
CTNNA3	catenin (cadherin-associated protein), alpha 3
DES	desmin
DMD	dystrophin
DOLK	dolichol kinase
DSC2	desmocollin 2
DSG2	desmoglein 2
DSP	desmoplakin
DTNA	dystrobrevin, alpha
EMD	emerin
EYA4	EYA transcriptional coactivator and phosphatase 4
FHL1	four and a half LIM domains 1
FHL2	four and a half LIM domains 2
FHOD3	formin homology 2 domain containing 3
FKRP	fukutin related protein
FKTN	fukutin
FXN	frataxin
GATAD1	GATA zinc finger domain containing protein 1
GLA	galactosidase, alpha
ILK	integrin-linked kinase
JPH2	junctophilin 2
JUP	junction plakoglobin
LAMA4	laminin, alpha 4
LAMP2	lysosomal-associated membrane protein 2
LDB3	LIM domain binding 3
LMNA	lamin A/C
MIB1	mindbomb E3 ubiquitin protein ligase 1
MT-TL1	mitochondrially encoded tRNA leucine 1 (UUA/G)

MYBPC3	myosin-binding protein C, cardiac
MYH6	myosin, heavy chain 6, cardiac muscle, alpha
MYH7	myosin, heavy chain 7, cardiac muscle, beta
MYL2	myosin, light chain 2, regulatory, cardiac, slow
MYL3	myosin, light chain 3, alkali; ventricular, skeletal, slow
MYLK2	myosin light chain kinase 2
MYOZ2	myozenin 2
MYPN	myopalladin
NEBL	nebulette
NEXN	nexilin (F actin binding protein)
NKX2-5	NK2 homeobox 5
mtDNA	mitochondrial DNA
PDLIM3	PDZ and LIM domain 3
PKP2	plakophilin 2
PLN	phospholamban
PRDM16	PR domain containing 16
PRKAG2	protein kinase, AMP-activated, gamma 2 noncatalytic subunit
PSEN1	presenilin 1
PSEN2	presenilin 2
PTPN11	protein tyrosine phosphatase, non-receptor type 11
RAF1	Raf-1 proto-oncogene, serine/threonine kinase
RBM20	RNA binding motif protein 20
RYR2	ryanodine receptor 2 (cardiac)
SCN5A	sodium channel, voltage-gated, type V, alpha subunit
SDHA	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
SGCD	sarcoglycan, delta (35kDa dystrophin-associated glycoprotein)
TAZ	tafazzin
TBX20	T-box 20
TCAP	titin-cap
TGFB3	transforming growth factor, beta 3
TMEM43	transmembrane protein 43
TMPO	thymopoietin
TNNC1	troponin C type 1 (slow)
TNNI3	troponin I type 3 (cardiac)
TNNT2	troponin T type 2 (cardiac)
TPM1	tropomyosin 1 (alpha)
TTN	titin
TTR	transthyretin
TXNRD2	thioredoxin reductase 2
VCL	vinculin

APPENDIX 1

Frequently used abbreviations:

ACE	angiotensin convertase enzyme
AD	autosomal dominant inheritance pattern
AGVGD	align Grantham variation Grantham distance (pathogenicity prediction software for missense variants)
AR	autosomal recessive
ARVC	arrhythmogenic right ventricular cardiomyopathy
bp	base pair
CGH	comparative genomic hybridization
CHD	congenital heart disorders
cM	centimorgan
CNV	copy number variation
DCM	dilated cardiomyopathy
dbSNP	NCBI's SNP database
DMEM	Dulbecco's Modified Eagle Medium
DNA	deoxyribonucleic acid
EBS	epidermolysis bullosa simplex
ECG	electrocardiogram
ES	exome sequencing
ESP	exome sequencing project (variant database of the NHLBI)
E. coli	Escherichia coli
FBS	fetal bovine serum
GERP	genomic evolutionary rate profiling (a score indicating the evolutionary conservation of a nucleotide)
GoNL	Genome of the Netherlands (database of the genomes of 500 individuals, used as a frequency database of "the Dutch wild type")
GWAS	genome-wide association study
HCM	hypertrophic cardiomyopathy
HEK	human embryonic kidney 293T cells
HF	heart failure
HiSeq	Illumina's Next Generation Sequencer system
HLA	major histocompatibility complex genes
HST	haplotype-sharing test
H_2O	hydrogen oxide (water)
H_2O_2	hydrogen peroxide
ICD	implantable cardioverter-defibrillator
LDB3	LIM domain binding 3 gene
LSH	longest shared haplotype

LVNC	left ventricular non-compaction cardiomyopathy
MD	muscular dystrophy
MiSeq	Illumina's "personal sequencer", the "little sister" of the HiSeq system in benchtop size, with faster workflow, allowing the assembly of small genomes or target regions
MRI	magnetic resonance imaging
mRNA	messenger RNA
NCBI	National Center for Biotechnology Information
NGS	next generation sequencing
NHLBI	National Heart Lung and Blood Institue, a division of National Institutes of Health in the USA
OMIM	"Online Mendelian Inheritance in Man" – a comprehensive database of human genes and genetic phenotypes authored and edited by the Johns Hopkins University
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PLEC	plectin
PolyPhen	Polymorphism phenotype (pathogenicity prediction software for missense variants)
PPCM	peripartum cardiomyopathy
RBM20	RNA binding motif protein 20
RCM	restrictive cardiomyopathy
ROS	reactive oxygen species
RNA	ribonucleic acid
RT	reverse transcription
SCD	sudden cardiac death
SIFT	sorting intolerant from tolerant (pathogenicity prediction software for missense variants)
SNP	single nucleotide polymorphism
SOD2	superoxide dismutase 2
TFC	task force criteria (diagnostic criteria of ARVC)
tRNA	transfer RNA
TTN	titin, the longest gene of the human genome
VOUS	variant of unknown significance
VUS	variant of unknown significance
1000G	1000 Genomes catalog of human genetic variation