Overview of Hypertrophic Cardiomyopathy (HCM) Genomics and Transcriptomics: Molecular Tools in HCM Assessment for Application in Clinical

Susana Rodrigues Santos
Centro Química Estrutural, Instituto Superior Técnico
Universidade de Lisboa, Portugal
Faculdade de Engenharia
Universidade Lusófona de Humanidades e Tecnologias, Portugal

Ana Teresa Freitas
Instituto de Engenharia de Sistemas e Computadores, Instituto Superior Técnico
Universidade de Lisboa, Portugal

Alexandra Fernandes
Departamento Ciências da Vida, Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa, Caparica, Portugal
Centro Química Estrutural
Instituto Superior Técnico, Universidade de Lisboa, Portugal

1 Inherited Cardiomyopathies

Inherited cardiomyopathies are a group of cardiovascular disorders classified based on the morphology and function of the ventricle and include hypertrophic cardiomyopathy (HCM), arrhythmogenic right ventricular cardiomyopathy (ARVC), dilated cardiomyopathy (DCM), left ventricular noncompaction (LVNC), and restrictive cardiomyopathy (RCM) (Teekakirikul *et al.,* 2013).

In this book chapter we will review current status of HCM molecular genetics and the importance of transcriptomics for revealing new diagnostic and therapeutic biomarkers and bioinformatic approaches to improve the translation between the bench and the clinic.

2 Physiopathological, Pathological, and Clinical Characteristics of HCM

HCM is a primary disorder of the myocardium classically characterized by unexplained left ventricular hypertrophy (LVH) (Figure 1) in the absence of an underlying systemic condition or other cardiac disease (such as hypertension or valvular heart disease), and by distinctive histopathologic features of cardiomyocyte hypertrophy, disarray and increased myocardial fibrosis (Figure 2) (Teekakirikul *et al.,* 2013; Spirito *et al.,* 2000, McKenna & Behr, 2002; McKenna *et al.,* 2003, Yaxin *et al.,* 2013). Apart from LVH, due to mitral valve systolic anterior motion and mitral-septal contact, LV outflow tract obstruction occurs in approximately 70% of HCM cases at rest and/or with physiologic provocation (Maron *et al*., 2006). Other HCM common phenotypic features include shortness of breath, chest pain, palpitations, presyncope or syncope and orthostasis (Teekakirikul *et al*., 2013). Despite these more common phenotypic features, HCM clinical manifestations are highly variable ranging from being completely asymptomatic to progressive heart failure and sudden cardiac death (SCD) caused by mechanical or electric defects (Teekakirikul *et al*., 2013; Tian *et al*., 2013, Lopes *et al*., 2013; Spirito *et al*., 2000, McKenna & Behr, 2002; McKenna *et al*., 2003). Indeed, HCM is the leading cause of sudden nontraumatic death in young adults and competitive athletes in the United States (Maron *et al*., 2003). HCM has an estimated prevalence of 1 in 500 in the general population (Gersh *et al*., 2011; Lopes *et al.,* 2013; Maron *et al.,* 1995).

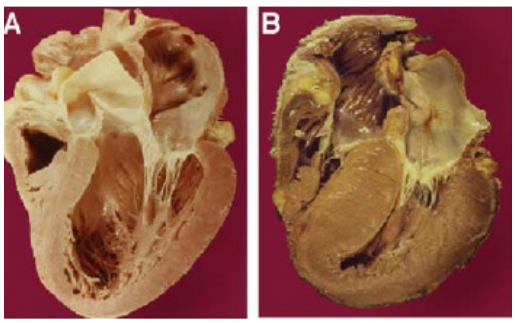
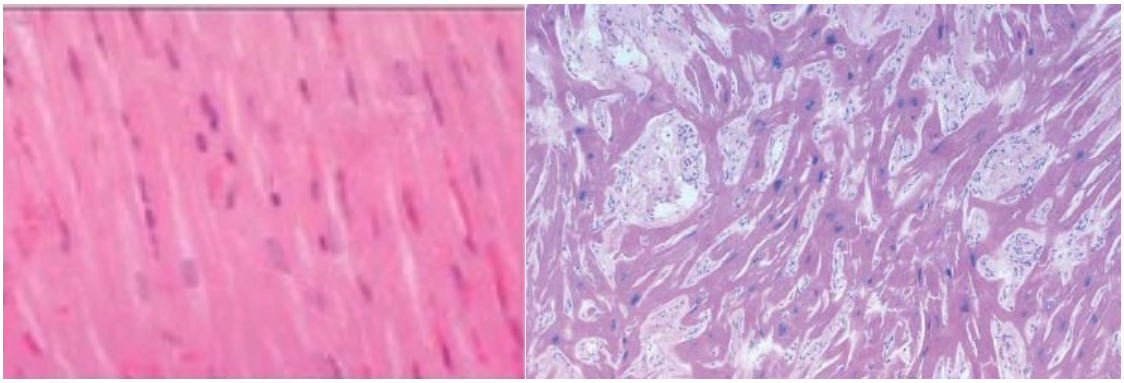


Figure 1: Morphologic features of the normal heart (A), compared to hypertrophic cardiomyopathy (B).We can observe the hypertrophied interventricular septum (arrow) (adapted from Ahmad *et al*., 2005).



(A) **(B)**

Figure 2: Histological images of a heart tissue from A. Normal individual; B. Patient with HCM showing the disarray and increased fibrosis (adapted from Ahmad *et al*., 2005).

HCM is usually diagnosed by a maximal LV wall thickness ≥ 15mm (13-14 mm=borderline, most commonly asymmetric septal and less frequently concentric and apical) through cardiac imaging, commonly done with two-dimensional echocardiography (ECHO) (the "gold standard" technique) but increasingly with cardiac magnetic resonance (CMRI) (Gersh *et al*., 2011). Some characteristic ECHO findings include systolic anterior motion of the mitral valve with associated left ventricular outflow tract obstruction and mitral regurgitation; midventricular obstruction as a result of systolic cavity obliteration; diastolic dysfunction including restrictive physiology (Teekakirikul *et al*., 2013). More recently, contrast-enhanced CMRI with gadolinium (Gd-CMRI) can reliably detect myocardial fibrosis *in vivo* and has been reported to be useful for the differential diagnosis of HCM (Miller *et al*., 2009).

The broad spectrum of HCM clinical manifestations described above and the age-dependent expression of hypertrophy make HCM clinical diagnosis difficult (Tian *et al*., 2013; Lopes *et al*., 2013; Spirito *et al*., 2000, McKenna & Behr, 2002; McKenna *et al*., 2003). A prior history of sudden cardiac arrest, syncope caused by cardiac arrhythmias, repetitive non-sustained or sustained ventricular tachycardia, severe cardiac hypertrophy and a strong family history of SCD are considered important risk factors and can be crucial for an early diagnosis (Elliot *et al.,* 2000, Spirito *et al*., 2000, Marian, 2003, Frenneaux, 2004). Also extremely important is the differential diagnosis of HCM due to HCM phenocopies. HCM phenocopies mimic the phenotypic and clinical features of sarcomeric HCM and include several syndromes that typically manifest with multiorgan involvement but that can also present with isolated or predominant LVH. These syndromes include metabolic cardiomyopathies, such as Danon disease and Wolf-Parkinson-White syndrome (Arad *et al.,* 2005) and the lysosomal storage disorder, Fabry disease (Sachdev *et al*., 2002). LVH in these conditions is not accompanied by myocyte disarray or fibrosis but by a characteristic accumulation of glycogen or glycosphingolipids in cellular vacuoles (Arad *et al*., 2005). LVH is also a part of the phenotypic spectrum of Noonan syndrome (Nishikawa *et al*., 1996) and Friedreich ataxia (Osterziel *et al*., 2002). The incidence of phenocopies in patients with the clinical diagnosis of HCM is unknown but is estimated at 10% (Marian, 2010).

3 HCM as a Genetic Disease: Importance of Genetic Diagnosis

Genetic studies established the paradigm that HCM is a disease of the sarcomere, caused by dominant mutations in genes encoding components of the contractile apparatus (Figure 3) with most of them (80%) present in the *MYH7* and *MYBPC3* genes (Table 1) (Seidman CE & Seidman JG, 2011). Until now more than 900 mutations in sarcomeric and more recently also in nonsarcomeric genes (for example in genes encoding Z-disk proteins and genes encoding proteins located in the sarcoplasmic reticulum and plasma membrane) have been described in HCM patients (Lopes *et al*., 2013; Teekakirikul *et al*., 2013; Santos *et al.,* 2011; Santos *et al*., 2012) (Table 1). However, variants in these nonsarcomeric genes are rare (Table 2), and most studies do not provide complete evidence of their role in HCM. Segregation with disease or *in vivo* functional data are necessary for most of these rare variants (Teekakirikul *et al*., 2013).

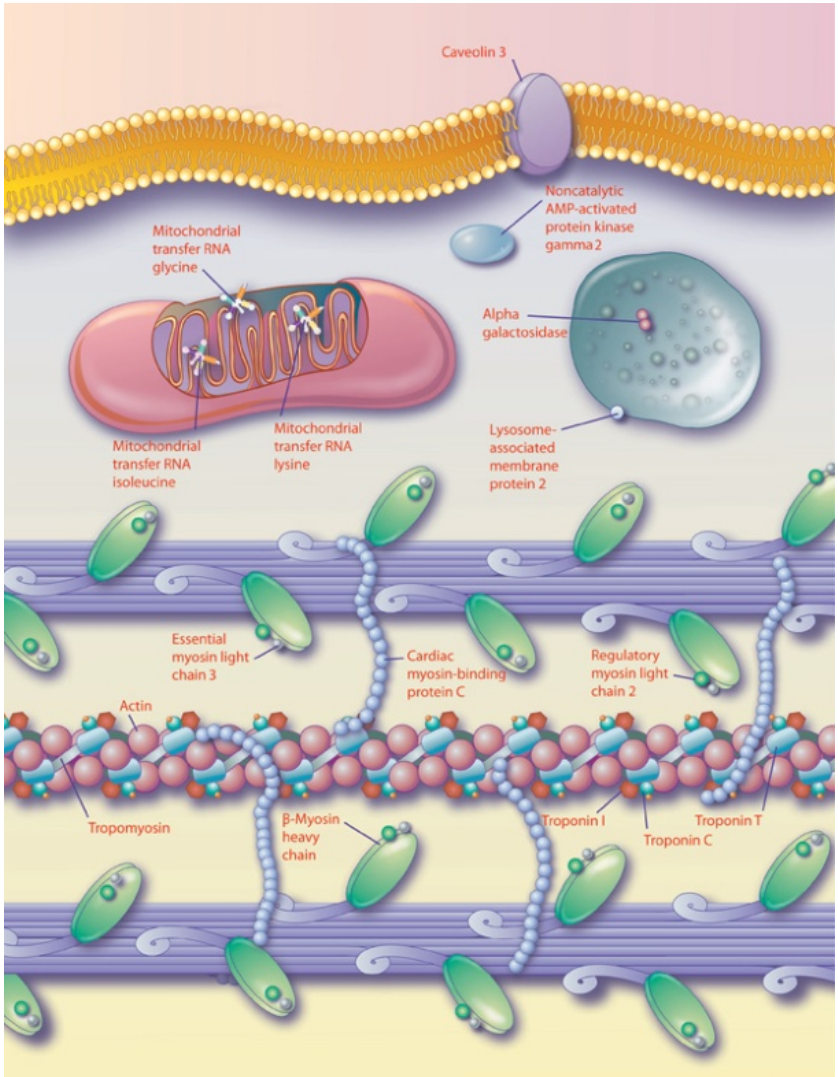


Figure 3: Cardiac sarcomere and proteins involved in HCM (Images courtesy of GeneDX (http://www.genedx.com/)

Most sarcomere mutations are believed to act in a dominant negative manner (i.e affecting the expression of the normal gene product). The exceptions to this dominant negative model for HCM are the loss-of-function variants resulting in insufficient protein for normal function and leading to haploinsufficiency; these variants occur less frequently but are prevalent in the *MYBPC3* gene (Andersen *et al*., 2004).

Gene	Protein	HCM associated Mutations*	Location or Function**
ACTA1	actin, alpha 1	1	Sarcomere, skeletal muscle
ACTC1	actin, alpha, cardiac muscle 1	14	Sarcomere, cardiac muscle
ACTN2	actinin, alpha 2	4	Z-disk
ANKRD1	ankyrin repeat domain 1	3	Z-disk and nucleus (transcription factor)
BRAF	v-raf murine sarcoma viral oncogene homolog B1	1	Cytoplasmic serine/threonine kinase
COA5	cytochrome c oxidase assembly factor 5	1	Mitochondrial
CALM3	calmodulin 3 (phosphorylase kinase, delta)	1	Calcium sensor and signal transducer
CALR3	calreticulin 3	2	Endoplasmic reticulum chaperone
CASQ2	calsequestrin 2	1	Sarcoplasmic reticulum; calcium storage
CAV3	caveolin 3	1	Plasma membrane
COX15	cytochrome c oxidase assembly homolog 15	2	Mitochondrial respiratory chain
CSRP3	cysteine and glycine-rich protein 3	11	Z-disk
DES	Desmin	1	Intermediate filament
FHL1	four and a half LIM domains 1	3	Biomechanical stress sensor
FHOD3	formin homology 2 domain containing 3	1	Actin-organizing protein
FXN	Frataxin	1	Mitochondrial iron transport and respiration
GLA	galactosidase, alpha	1	Lysosome
JPH2	junctophilin 2	4	Junctional membrane complexes; calcium signaling
KLF10	Kruppel-like factor 10	6	Transcriptional repressor; inhibits cell growth
MAP2K1	mitogen-activated protein kinase kinase 1	1	MAP kinase kinase; signal transduction
MAP2K2	mitogen-activated protein kinase kinase 2	1	MAP kinase kinase; signal transduction
MRPL3	mitochondrial ribosomal protein L3	1	Mitochondrial Ribosomal Protein
MTO1	mitochondrial tRNA translation optimization 1	2	Mitochondrial tRNA modification
MYBPC3	myosin binding protein C, cardiac	373	Sarcomere
MYH6	alpha-myosin heavy chain	4	Sarcomere
MYH7	beta-myosin heavy chain	307	Sarcomere

MYL2	ventricular myosin regulatory light chain	14	Sarcomere
MYL3	myosin light chain 3	12	Sarcomere
MYLK2	myosin light chain kinase 2	2	Calcium/calmodulin dependent kinase
MYO6	myosin VI	1	Actin-based reverse-direction motor protein
MYOM1	myomesin 1	1	Sarcomere
MYOZ2	myozenin 2	2	Z-disk
MYPN	Myopalladin	7	Z-disk
NDUFAF1	NADH dehydrogenase (ubiquinone) complex I, assembly factor 1	2	Mitochondrial chaperone
NDUFV2	NADH dehydrogenase (ubiquinone) flavoprotein 2	1	Mitochondrial respiratory chain
NEXN	Nexilin	2	Z-disk
OBSCN	Obscurin	1	Sarcomere
PDLIM3	PDZ and LIM domain 3	1	Z-disk
PRKAG2	5'-AMP-activated protein kinase subunit gamma-2	7	Energy sensor protein kinase
PLN	Phospholamban	7	Sarcoplasmic reticulum; regulates Ca(2+)-ATPase
RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	1	Serine/threonine-protein kinase; signal transduction
SLC25A3	solute carrier family 25, member 3	1	Phosphate carrier protein (cytosol to mitochondria)
SLC25A4	solute carrier family 25, member 4	2	Adenine nucleotide translocator (cytosol/mitochondria)
SOS1	son of sevenless homolog 1	2	Guanine nucleotide exchange factor for RAS proteins; signal transduction
SRI	Sorcin	1	Calcium-binding; modulates excitation-contraction coupling
TCAP	Telethonin	3	Z-disk
TNNC1	troponin C	7	Sarcomere
TNNI3	troponin I	45	Sarcomere
TNNT2	troponin T	50	Sarcomere
TPM1	alpha-tropomyosin	16	Sarcomere
TRIM63	tripartite motif-containing 63	3	Sarcomere; regulates protein degradation
TTN	Titin	3	Sarcomere
VCL	Vinculin	1	Sarcomere

Table 1: HCM associated genes, proteins, mutations and protein localization or function. * - Human Genome Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php); ** - National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/)

These sarcomere protein abnormalities ultimately appear to converge on increased actin activated ATPase activity, disruption of actin-myosin interaction and force generation, and altered intracellular calcium signaling in cardiomyocytes as the major common paths leading to the anatomic (hypertrophy,

myofiber disarray, and fibrosis) and functional features (pathological signaling and diastolic dysfunction) characteristic of HCM (Teekakirikul *et al*., 2013; Fatkin D & Graham RM 2002; Sequeira *et al*., 2013; Witjas-Paalberends *et al*., 2013). More recently scientific data suggests that LVH can be caused by perturbations of transforming growth factor b and CaMKII Mef2 signaling pathways (Teekakirikul *et al.,* 2010).

Sarcomere gene mutations have been identified in up to 60% of familial HCM and in 40% of sporadic HCM (Ho, 2010). Most of HCM associated mutations/variants are private (within a family) and only a small number of variants are present at higher population frequencies, the most common of which, a 25-bp deletion in intron 32 of the *MYBPC3* gene, is present in 4% of the Southeast Asian Indian population and increases the risk for heart failure by >6 fold (Dhandapany *et al*., 2009). In the Netherlands, for instance a single founding mutation also accounts for a substantial percentage of HCM (Alders *et al*., 2003), while in the US, there is marked genetic heterogeneity and most HCM-patients have a unique pathogenic mutation (Alcalai *et al.,* 2008).

The marked genetic heterogeneity, the highly variable intra- and inter- family expressivity and the incomplete penetrance poses some problems in the establishment of complete genotype-phenotype correlations. Despite this, few exceptions has been attempted namely at the gene level. For example, some *TNNT2* mutations have been associated with an appreciable risk of arrhythmia and minor hypertrophy (Watkins *et al*., 1995), *MYH7* variants lead to a significant LVH after the second decade of life and are thought to be associated with an increased risk of heart failure and SCD (Ho *et al*., 2010a), and some pathogenic variants in *MYBPC3* are believed to be associated with a later onset (high prevalence of loss-of-function variants) (Niimura *et al*., 2002) although others have also been identified in a significant proportion of HCM patients with childhood-onset LVH (usually missense variants; possible with more severe functional consequences) (Morita *et al*., 2008). For a more detailed information regarding genes, mutations and clinical phenotype see table 2.

Causal genes for HCM	Encoded Protein	Frequency Familial (%)[¥]	Frequency Sporadic (%)*	Clinical Features, associated phenotype	References
Sarcomere Thick filament					
MYH7	Beta-myosin heavy chain	30-40	25-35*	Moderate to severe hypertrophy, high disease penetrance and variable prognosis. Younger onset is typical. Severe LVH, Heart failure, SCD	Watkins et al 1992; Charron et al 1998; Marian et al 2001; Seidman & Seidman 2001; Fatkin, 2002; Erdmann et al 2003; Richard et al 2003; Ingles et al 2005; Van Driest et al 2005; Morita et al 2008; Ho et al 2010; Keren et al 2008; Lopes et al 2013; 2013b; Seidman & Seidman, 2011; 2013, 2013b; Teekakirikul et al 2013
MYBPC3	Myosin binding protein C, cardiac	30-40	20-30*	Usually milder disease. Although it can be severe; some older onset;	Watkins et al 1995b; Charron et al 1998, Marian et al 2001; Seidman & Seidman 2001; Fatkin 2002; Niimura et al 2002; Erdmann et al 2003; Richard et al 2003; Andersen et al 2004; Ingles et al 2005; Van Driest et al 2005; Keren et al 2008; Morita et al 2008; Ho et al 2010; Seidman & Seidman, 2011; Lopes et al 2013, 2013b; Teekakirikul et al 2013

TNNT2	Troponin T	10-20	3-5*	Mild LVH, SCD more common	Thierfelder et al 1994; Watkins et al 1995; ; Moolman et al 1997; Marian et al 2001; Seidman & Seidman 2001;Varnava et al 2001; Fatkin 2002; Erdmann et al 2003; Richard et al 2003; Ingles et al 2005; Van Driest et al 2005; Keren et al 2008; Morita et al 2008 Ho et al 2010; Lopes et al 2013, 2013b; Teekakirikul et al 2013
MYL2	Ventricular myosin regulatory light chain 2	Rare**	Rare**	Skeletal myopathy Midcavity obstruction Common in HCM families with a malignant prognosis.	Poetter et al 1996; Seidman & Seidman 2001; Fatkin 2002; Richard et al 2003; Ingles et al 2005; Van Driest et al 2005; Morita et al 2008, Lopes et al 2013, 2013b; Teekakirikul et al 2013
MYL3	Myosin light chain 3	Rare**	Rare**	Skeletal myopathy Midcavity obstruction	Poetter et al 1996; Seidman & Seidman 2001; Fatkin 2002; Richard et al 2003; Ingles et al 2005; Van Driest et al 2005; Morita et al 2008, Lopes et al 2013, 2013b; Teekakirikul et al 2013
MYH6	Alpha-myosin heavy chain	Rare**	Rare**	Late onset	Carniel et al 2005; Fatkin 2002; Lopes et al 2013, 2013b; Teekakirikul et al 2013
Sarcomere Thin filament					
TNNI3	Troponin I	2-5	?	Very heterogeneous disease expression. Sudden death with severe disease	Kimura et al 1997; Seidman & Seidman 2001; Fatkin 2002; Niimura et al 2002; Mogensen et al 2004; Van Driest et al 2005; Morita et al 2008; Ho et al 2010;Teekakirikul et al 2013
TPM1	Alpha-tropomyosin	2-5	?	Variable prognosis, sudden death	Thierfelder et al 1994; Watkins et al 1995; Fatkin 2002; Seidman & Seidman 2001; Erdmann et al 2003; Heller et al 2003; Chang et al 2005; Van Driest et al 2005; Morita et al 2008; Ho et al 2010; Karibe et al 2011; Teekakirikul et al 2013
TNNC1	Troponin C	Rare**	Rare**	Typical HCM	Seidman & Seidman 2001; Erdmann et al 2003; Fatkin 2002; Richard et al 2003; Teekakirikul et al 2013
TTN	Titin	Rare**	Rare**	Typical HCM	Satoh et al 1999; Bos et al 2006; Lopes et al 2013b; Teekakirikul et al 2013
ACTC1	Actin, alpha, cardiac muscle	Rare**	Rare**	Apical hypertrophy Midcavity obstruction Some late onset mutations.	Olson et al. 2000; Seidman & Seidman 2001; Wong et al 2001; Fatkin 2002; Niimura et al 2002; Richard et al 2003; Ingles et al 2005; Van Driest et al 2005, Vang et al 2005; Bookwalter et al 2006; Lopes et al 2013, 2013b; Teekakirikul et al, 2013
Z- Disk					
TCAP	Telethonin	Rare**	Rare**	Variable hypertrophy Typical HCM, variable penetrance	Hayashi et al 2004; Bos et al 2006; Teekakirikul et al 2013
CSRP3	Cysteine and glycine-rich	Rare**	Rare**	Variable hypertrophy and skeletal myopathy Late onset, variable	Fatkin 2002; Bos et al 2006; Teekakirikul et al 2013

	protein 3			penetrance	
MYOZ2	Myozenin 2	Rare**	Rare**	Early onset of symptoms, pronounced cardiac hypertrophy, and cardiac arrhythmias	Osio et al 2007; Posch et al 2008; Teekakirikul et al 2013
VCL	Vinculin	Rare**	Rare**	Obstructive midventricular Hypertrophy	Vasile et al 2006: Teekakirikul et al 2013
Ca^{2+}-handling protein					
PLN	Phospholamban	Rare**	?	Typical HCM, variable penetrance	Landstrom et al 2011; 2012; Teekakirikul et al 2013
CASQ2	Calsequestrin 2	Rare**	?	Typical HCM	Keren et al 2008; Landstrom et al 2012
JPH2	Junctophilin 2	Rare**	?	Typical HCM	Keren et al 2008; Landstrom et al 2012

Table 2: Hypertrophic cardiomyopathy: disease-causing genes and associated clinical features and phenotypes; ¥Hershberger et al 2009; *Keren et al 2008; **Rare - Means < 1 %; ? - Means unknown

For the last decades the management strategy for families living with HCM has been described in the literature and followed routinely by cardiologists over the world (Ho, 2010, O'Mahony *et al*., 2013). Serial screening in HCM families includes an echocardiogram, electrocardiogram and evaluation by a cardiologist (Maron *et al*., 2003, O'Mahony *et al*., 2013). This strategy is considered time consuming and an expensive process for families and medical community. In the last decade, major progresses have been achieved in the understanding of the genetic basis underlying HCM and genetic testing has been offer as a service (Ackerman *et al.,* 2011; Wang *et al*., 2010, Ho *et al.,* 2011). Clinical guidelines for HCM recommend comprehensive testing for at least five HCM genes (*MYBPC3*, *MYH7*, *TNNI3*, *TNNT2*, and *TPM1*) (Ackerman *et al.,* 2011, Ho *et al*., 2011). Genetic testing for HCM is primarily used to identify families with a detectable genetic cause of disease and to screen at-risk family members. Testing can also help rule out nongenetic conditions, such as athlete's heart, although only when a pathogenic variant is identified (Ho, 2010; Wheeler *et al*., 2009). Since the late 90s traditional genotyping strategies such as SSCP (Single Strand Conformation Polymorphism), DHPLC (Denaturing High Pressure Liquid Chromatography) and dideoxy sequencing of amplified protein encoding exons have been employed (Larsen *et al.,* 1999, Mogensen *et al*., 2003). Nevertheless, these technologies are considered low throughput, with high cost and has a low sensitivity. In 2003, genetic testing entered the mainstream of the healthcare system with automated DNA sequencing providing rapid, reliable, and comprehensive molecular diagnosis on a fee-for-service basis (Maron *et al*., 2012). Although current clinical guidelines recommend routine genetic testing in patients with HCM (Anderson *et al.,* 2004; Hershberger *et al*., 2009; Ho, 2010) its use in everyday clinical practice has been limited by the cost and complexity of conventional sequencing technologies. Advances in high throughput technologies have the potential to solve this problem by analyzing substantially larger genomic regions at a lower cost than conventional capillary Sanger sequencing (Geier *et al*., 2008). In the last years, high throughput technologies such as High Resolution Melting (Millat *et al*., 2010; Santos *et al*., 2012), resequencing DNA arrays (Fokstuen *et al.,* 2008), Next Generation Sequencing (NGS) (Lopes *et al*., 2012; Lopes *et al.*, 2013) and iPlex Mass Array (Santos *et al*., 2011) have emerged to detect HCM pathological mutations enabling the test of a large number of genes simul-

taneously with lower costs and more rapidly. Nevertheless, some of these new technologies also pose new challenges, in particular, NGS and HRM given their potential to identify a large number of variants of unclear clinical significance (VUS) (Lopes *et al*., 2013; Santos *et al*., 2012). Indeed, distinguishing pathogenic mutations from VUS or rare nonpathogenic variants regarding genetic testing interpretation is considered a new challenge. Notably, generally accepted guidelines for interpreting VUS are currently lacking, with estimated frequencies varying dramatically from 5% to 50% (Maron *et al*., 2012), and being so, when a VUS is identified, its utility in confirming the diagnosis in a proband with suspected disease is limited and it cannot be used for predictive gene testing in at-risk relatives. Additionally, laboratory commercial testing strategies vary widely, employing from 3 to 7 descriptive pathogenicity classes to construct formal reports, thereby creating the distinct possibility that different interpretations of pathogenicity may emanate from different laboratories for the same mutation (Maron *et al*., 2012). In this regard, it has become critical to carefully review the scientific evidence underlying a proclaimed disease gene. Genes with low detection rates (Table 2) usually have a limited number of studies and many studies infer disease association based on the absence from a small number of healthy control individuals and evolutionary conservation of the affected amino acid. Indeed, screening 300 chromosomes identifies only 80% of all rare benign variation and that more than 6000 chromosomes are necessary to identify the full spectrum (Ionita-Laza *et al*., 2009). Additional genetic evidence, such as segregation with disease and functional data that would bring a stronger support, is rarely available (Teekakirikul *et al*., 2013).

Taking all this into consideration the process of genetic testing involves more and more a close collaboration between molecular geneticists and genetic counselors at the testing laboratory and the referring health care provider and the family. Genetic test results often initiate a cascade of events, including familial segregation studies and iterative reevaluation of clinical and molecular evidence, to determine the clinical significance of novel variants. Although clinical sequencing of emerging disease genes is increasingly performed, it is important to realize that it can take years until the spectrum of pathogenic and benign variation has been characterized. Until then, testing often yields many VUSs, and these inconclusive results can be problematic for the patient (Aatre *et al*., 2011). The rapidly increasing number of sequenced genomes is beginning to remedy this problem, as the spectrum of rare benign variation is better defined (Teekakirikul *et al*., 2013). As this is a rapidly expanding field, recommendations for HCM genetic testing will be reinforced. Nevertheless, it should be noted that there is a positive impact of identifying a mutation in an individual since it often provides an explanation for why the disease has occurred and also allows for cascade testing of other affected and unaffected family members. Testing of affected family members is performed as a confirmation of their disease status and to exclude the possibility of a "phenocopy". This enables an accurate risk assessment to be given for their offspring. Asymptomatic family members can be offered a predictive genetic test to clarify whether they are at risk of developing clinical disease and to determine the inheritance risk to their children. As the decision of having a predictive genetic test is complex, it should be made in the context of a clinical genetics service offering pre- and post-test genetic counselling. Additionally, it should be noted that, in most cases, the presence of a mutation on a genetic test will not alter the clinical management of an individual who is already known to be affected. However, in cases where an unaffected individual receives a positive predictive gene test result, an ongoing surveillance programme should be recommended. Due to variable age of onset, surveillance should be started early, and may be determined by the natural history in the family. Genotype-phenotype correlations, if known, could be an important consideration in determining screening for asymptomatic mutation positive individuals (Teekakirikul *et al*., 2013; Maron *et al.,* 2010). Indeed, in the clinical practice genotyping HCM-patients has proven to be useful for risk stratification, with particular

regard to the SCD event (Elliot *et al*., 2000, Spirito *et al.,* 2000, Marian, 2003, Frenneaux, 2004; Hershberger *et al*., 2009). An accurate diagnostic molecular approach to establish meaningful relationships among individual gene mutations, phenotype, and outcome in HCM cohorts is therefore considered of utmost importance (Arad 2002, Frenneaux , 2004, Bos *et al.,* 2006, Santos *et al.,* 2012, Seidman & Seidman, 2011). The management of the clinical aspects of HCM, integrating the clinical, diagnostic, and therapeutic recommendations based on a synthesis of all data is particularly relevant with HCM because of the complexity of decision analysis for clinical interventions (eg, the assessment of outflow tract obstruction, and if present, selection of a treatment plan that may involve surgical or catheter-based interventions or a positive family history of SCD, may increase the threshold for preventive implantation of an automated implantable cardioverter defibrillator) (Teekakirikul *et al*., 2013; Maron *et al*., 2010; Hershberger *et al*., 2009).This also poses several implications for the mutation-positive individuals since they should receive appropriate lifestyle modification advices as part of result follow-up.

Through large-scale gene sequencing, new data have emerged in HCM suggesting that double (or triple) or compound pathogenic mutations can be associated with more severe disease expression and adverse prognosis (e.g., advanced heart failure or SCD, even in the absence of conventional risk markers (Maron *et al*., 2012). While it is possible that multiple mutations will prove to be prognostic markers or arbitrators of ambiguous risk profiles, current evidence is preliminary and prospective long-term studies in large populations are required (Arad 2002, Frenneaux , 2004, Bos *et al*., 2006, Maron *et al*., 2013, Santos *et al.,* 2012, Seidman & Seidman, 2011).

The benefits for establishing the molecular basis for HCM are not limited to researchers. Indeed, for clinicians, knowledge about the genetic causes of human HCM is of immediate and practical value. The practical information about genetic causes has considerable implications provides insights into prognosis, accurate diagnosis and screening strategies for at-risk familial members (Hershberger *et al*., 2009).

Another important aspect is the use of genetic test results in guiding clinical management such as the case of enzyme replacement therapy for storage disorders that can present as isolated LVH (Rozenfeld & Neumann, 2011). In this regard, an emerging area is the use of genotype data to guide therapeutic decisions in preclinical individuals. Studies using animal model suggest that calcium channel blockers, such as diltiazem, may delay the clinical progression of disease (Semsarian *et al*., 2002). Additionally, animal studies have also shown evidence of a connection between HCM and increased transforming growth factor b signaling. A combined therapy using an anti-transforming growth factor b antibody and losartan (an angiotensin II receptor type 1 antagonist) was shown to prevent cardiac fibrosis and hypertrophy in a sarcomere mutation positive mice (Teekakirikul *et al*., 2010). All these studies may open additional therapeutic avenue for HCM.

4 HCM New Insights: Translating Transcriptomics into Diagnostic Markers

One of the major concerns regarding HCM is the phenotypic variability even in related patients carrying the same disease-causing mutation (Keren *et al*., 2008). This has drawn several concerns in the establishment of complete genotype-phenotype correlations based only on their genetic analysis results and raises the fact that HCM phenotypic and genetic variability cannot be explained exclusively on the basis of the type of genetic defect or the gene that is altered per se (Fatkin *et al*., 2002). Indeed, disease manifestation are likely to be influenced not only by genetic factors such as mutations, but also by modifier

genes, and environmental factors such as lifestyle, degree of physical exercise and blood pressure (Fatkin *et al*., 2002).

Despite significant advances in the relations between structure and function of the components of the contractile apparatus, there is limited knowledge of how a particular mutation can induce clinical HCM. Cardiac muscle development is characterized by the activation of contractile protein genes and subsequent modulation of expression by physiological and pathological stimuli (Clerk *et al.,* 2007; Kehat *et al.,* 2010). Structural remodeling, meaning changes in the size, shape and function of the heart, is a key determinant of whether cardiac hypertrophy is classified as pathological or physiological. Pathological hypertrophy is associated with structural rearrangement of components of the normal chamber wall that involves cardiac myocyte hypertrophy, cardiac fibroblast proliferation, fibrosis, and cell death (apoptotic and necrotic) (Manabe *et al*., 2002). During the HCM development, cardiomyocytes undergo a remodeling process that, while initially compensatory, ultimately accelerates functional deterioration namely caused by cardiomyocyte disarray and tissue fibrosis responsible for the onset of cardiac failure (Maron *et al*., 1987; Klues *et al*., 1995; Nimura *et al*., 1998; Marian *et al*., 2000; Olivotto *et al*., 2001; Varnava *et al*., 2001; Maron *et al*., 2004). In this regard, the understanding of the molecular changes within the myocardial tissue that for instances occur in response to HCM causal mutations may advance mechanistic insights that account for disease. This important issue could be tackled by characterizing the key players in cardiac tissues during HCM induced sarcomere remodeling. In this context the transcriptional profile involved in the contractile function and HCM progression, may represent valuable candidates for recognition of HCM (Watkins *et al*., 1996; Cooper *et al.,* 2005; Lawler *et al*., 2007; Theis *et al*., 2009). Gene expression during HCM remodelling in myocardial tissue – ventricle and auricle was performed by our group (our results not published; Santos *et al*., 2009 and Santos *et al*., 2010 both oral presentation in international meetings; Santos *et al*., 2012; Fernandes *et al*., 2013 both poster presentation in international meetings). In an expression study in human HCM tissue an increased myofilamet proteins level in patients with either MYBPC3- or MYH7- mediated HCM suggest a dominant negative mechanism (Theis *et al*., 2009). Our results point out to the hypothesis that transcriptional variants may correlated with the clinical and genetic profile (Santos *et al*., 2009 and Santos *et al*., 2010 both oral presentation in congress; Santos *et al*., 2012; Fernandes *et al*., 2013 both poster presentation in congress). The extraordinary degree of analytical power of the current high-throughput technologies provides opportunities to gain novel insights into cell biology and associated pathophysiological processes. Cardiac molecular changes that prompt cardiac remodeling could be evaluated by comprehensive transcriptional analyses that quantify key molecules as myocardial mRNAs and microRNAs (miRNAs) within HCM heart tissues. High throughput transcriptomic analyses of mRNAs enable scientists to gain a comprehensive view of transcriptional changes within tissue. Heart tissue mRNAs could be systematically studied using different and complementary methods, namely RNA sequencing (RNA-Seq), Quantitative Real Time gene expression (qRT-PCR), Deep Serial Analysis Gene Expression (DSAGE). These approaches allow quantifying mRNA and defining gene splice variation. Also, custom or commercially available cDNA microarrays can simultaneously examine the expression of thousands of transcripts from several biological specimens. Most important, the correlation of genomic and transcriptomic strategies allow characterizing transcriptional responses related to HCM mutations, and as so to gain fundamental mechanistic insights that could explain the range of clinical manifestations and symptoms associated with HCM (Watkins *et al.,* 1996; Cooper *et al*., 2005; Margulies & Matiwala, chapter 36; Lawler *et al*., 2007; Theis *et al.,* 2009).

For the transcriptional profiling of HCM in human cardiac tissue a range of logistical and technical factors tend to undermine the settlements that could be acquired by this approach. Foremost, all

studies obtain failing human ventricular myocardium at the time of myectomy or cardiac transplantation (Margulies & Matiwala, chapter 36). This sampling approach excludes some of the information that could be obtained by a time sampling strategy to enquire about disease progression. If possible, gene expression analysis should be conducted at multiple time points of sampling that has proved to be informative in animal models studies. Also, myocardial profiling from failing hearts is focused in end state of the disease and probably the inherent transcriptional changes correspond to an adaptive cellular and molecular mechanisms rather than the cause of the pathology. Moreover, the time of onset, etiology and therapeutics associated to each individual can confound the true significance of the obtained data. Another logistical issue concerns with the complicated issue to obtain an appropriate healthy human myocardial control tissue. Because of rapid mRNA degradation, samples obtained at necropsy are not suitable.

Despite the above considerations, the use of gene expression analysis does provide an opportunity to take aware of the transcriptional portrait of cardiomyopathies by simultaneous analysis of several genes from diverse pathways. Studies regarding the profiling of transcription patterns in myocardial tissues have already produced novel evidences that contribute to the understanding of cardiomyopathies cardiac remodeling (Lee *et al*., 2000; Lim *et al.,* 2001; Hawng *et al*., 2002). An example is related with the molecular distinction between HCM and DCM being these data obtained from transcriptional profiling data (Hwang *et al*., 2002). In most cases, familial HCM is caused by mutation in one of sarcomeric proteins, while cytoskeletal mutations can lead to a DCM (Fatkin *et al*., 2002). Hwang *et al.* demonstrated sets of genes that were differentially expressed among the patients with DCM and HCM, despite the common endpoint of severe heart failure (Hwang *et al*., 2002). Among upregulated genes, the ones related to immune responses were more mostly identified in DCM, while the genes related to protein synthesis were more distinguished in HCM. On the other hand, the genes related to metabolism were more downregulated in DCM than in HCM. Genes related to cell signaling and cell structure tended to be further reduced in HCM compare to DCM (Hwang *et al*., 2002). In this example the obtained data of transcriptional profiling supports the concept that distinct etiologies of cardiomyopathies, progress through different patterns of cardiac remodeling, involving distinct molecular dynamics. Another example, is regarded the end-stage DCM in which Barrans and collaborators (Barrans *et al*., 2002) identified a significant transcriptional convergence. These author compared DCM and healthy cardiac tissues and observed the deregulation of more than 100 genes in the failing hearts. In addition it was identified an upregulation of classic markers of hypertrophy, such as sarcomeric and cytoskeletal proteins, transcriptional factors and genes involved with energy metabolism (Barrans *et al*., 2002).

It is well documented that the cellular responses characteristic of cardiac hypertrophy include an increase in cell size due to accelerated synthesis of sarcomeric and structural proteins, and reprogramming of the fetal cardiac genes (Komuro *et al*., 1993; Sadoshima *et al*., 1997). Some of the contractile proteins, ion channels and metabolic enzymes have both fetal and adult isoforms, which have similar, but not identical, functions. Studies regarding gene expression reported that the transcriptional control during cardiac hypertrophy involves switching of gene expression from normally expressed adult isoforms to fetal isoforms (Komuro *et al*., 1993; Sadoshima *et al.,* 1997; Friddle *et al*., 2000).

Tan and collaborators (Tan *et al*., 2002) reported an accentuated transcriptional convergence among the different diseased hearts. Interestingly, many of the deregulated genes in end-stage failing human hearts include the genes also expressed during fetal development. This data supports the evidence that many of the transcriptional changes observed in severely failing human hearts obtained at the time of transplantation are responses to, rather than causes of, sustained myocardial stress and overall the heart failure phenotype (Vikstrom *et al*., 1998).

Marston *et al* (2009) studied the cellular mechanisms responsible for sarcomere stoichiometric alteration in HCM using human myectomy samples, and were able to demonstrate that mutations in cardiac myosin-binding protein-C (MyBP-C), cause HCM through a mechanism of haploinsufficiency. Most of sarcomere gene mutations that cause HCM are missense alleles that encode dominant negative proteins. MyBP-C mutations are considered an exception as frequently encode truncated proteins. The authors compared ventricular muscle from patients undergoing surgical myectomy with samples from donor hearts (Marston *et al*., 2009). MyBP-C protein and mRNA levels were quantified from cardiac tissues using immunoblotting and RT-PCR. The absence of any detectable truncated MyBP-C argues against its incorporation in the myofiber and any dominant negative effect. In contrast, the lowered relative level of full length protein in both truncation and missense *MYBPC3* mutations argues strongly that haploinsufficiency is sufficient to cause the disease (Marston *et al*., 2009).

In addition, recent data have demonstrated that using PMAGE (for "polony multiplex analysis of gene expression") is possible to identify early transcriptional changes that preceded pathological manifestations of HCM in mice carrying a disease-causing mutation (Kim *et al*., 2007). The authors describe a sensitive mRNA profiling technology that provided a comprehensive profile of cardiac mRNAs, including low-abundance mRNAs encoding signaling molecules and transcription factors that are likely to participate in disease pathogenesis.

In recent years, it has become increasingly apparent that therapeutic interventions have effects on transcriptional profiles in failing hearts prompting the ability to induce regression of the pathological phenotype (Margulies & Matiwala, chapter 36). In both animal models and clinical settings, this phenomenon of so-called 'reverse remodeling' has been observed via both medical interventions and surgical interventions. The cellular and organ level mechanisms that drive the process of myocardial recovery and reverse remodeling are still under enquire in a way similar to the processes that drive the progression of cardiomyopathy in diseased hearts. As so the use of transcriptional profiling may provide new insights into the molecular biology of myocardial reverse remodeling. Interestingly, in what concerns DCM therapeutics and the intrinsic targets, Yasumura and collaborators (2003) identified an increase in SERCA and phospholamban abundance and a decrease in β myosin heavy chain (β -MHC) and sodium-calcium exchanger. Gene expression of right ventricular endomyocardium was assessed by qRT-PCR (Yasamura *et al*., 2003). These results provide further evidence that β-blocker treatment affects expression of sarcomeric proteins and calcium regulatory proteins (Yasamura *et al.,* 2003). The authors reported that ventricular functional recovery by β-blocker therapy is attributed to time-dependent biologic effects on cardiomyocyte (Yasamura *et al*., 2003). The obtained data indicate that transcriptional profiling may provide clues of the mechanisms of myocardial adaptations observed during pharmacological therapies. Commonly prescribed therapeutic agents for cardiovascular disease exert pleiotropic effects on cardiomyocytes and cardiac fibroblasts having beneficial outcomes on the remodeling heart. These include drugs for reducing hypertension, as ACE inhibitors, angiotensin receptor blockers, beta-blockers, for cholesterol levels as statins, fibrates and for insulin resistance as thiazolidinediones (Porter *et al.,* 2009).

The above transcriptional profiling studies using human myocardial specimens with cardiomyopathies, and HCM could reveal the mechanisms how gene mutations trigger a variety of transcriptional adaptations. These studies further demonstrate that transcriptional regulation of severe hypertrophy or dysfunction, are more likely to represent model-specific changes rather than non-specific responses to the development of heart failure and myocardial stress or injury (Margulies & Matiwala, chapter 36). Nevertheless, a more accurate evaluation of transcriptional profiling must be complemented with assessments of protein abundance and post-translational modifications.

In parallel with mRNA evaluation, microRNAs (miRNAs), non-protein-coding small RNAs of 20-23 nucleotides, have emerged as one of the central players of gene expression regulation (Thum, 2011). miRNAs act as negative regulators of gene expression by inhibiting the translation or promoting the degradation of target mRNAs (Thum, 2011). Recent studies have uncovered important roles for miRNAs, in the control of diverse aspects of cardiac function (including cardiomyocyte growth, integrity of the ventricular wall and contractility) and of the pathological heart (Olson & Rooj, 2007; Cheng *et al*., 2007; da Costa Martins *et al*., 2008; Thum *et al*., 2008; Wang & Yang, 2012). Specific miRNAs are misexpressed in diseased heart, and gain- and loss-of-function experiments in mice have shown the importance of these miRNAs as necessary and sufficient to evoke cardiac hypertrophy and heart failure (Olson & Rooj, 2007; Divakaran *et al*., 2008). Microarray and qRT-PCR analyses also have demonstrated a collection of miRNAs that are up/downregulated during pathological cardiac remodeling in hypertrophic and failing hearts (Olson & Rooj 2007; Cheng *et al*., 2007; da Costa Martins *et al*., 2008; Thum *et al*., 2008; Wang & Yang, 2012). *In vitro* experiments using either overexpression or knockdown of miRNAs in cultured cardiomyocytes indicate that a subset of miRNAs are indeed actively involved in cardiomyocyte hypertrophy. However, a major challenge remains in to identify the mRNA targets of the miRNAs that participate in cardiac remodeling and to understand the functions of their target mRNAs. Indeed each miRNA could repress up to hundreds of transcripts, and it is thus hypothesized that miRNAs form large-scale regulatory networks across the transcriptome through miRNA response elements (MREs) (Cheng *et al.,* 2007).

In our research we have been tackling some of the pathways involved in HCM remodeling through an integrated evaluation of the transcriptional profile in human HCM cardiac tissue not only at the level of sarcomere genes expression levels but also miRNA profiling (our unpublished results; Fernandes *et al*., 2013; Santos *et al*., 2013 - poster presentations in international meetings). These miRNA expression profiling studies are important for revealing novel miRNA-based pathways underlying the cardiac remodeling and therefore to inquire the potential role for miRNAs in regulating the changes in gene expression that occur during cardiac hypertrophy. As so, the use of specific miRNA for targeting deregulated mRNA could be an opportunity for potential therapeutic of heart diseases, namely in controlling cardiac remodeling typical of HCM, and so specific miRNAs may themselves become therapeutic targets in HCM.

To date, several tools are available to selectively target miRNA pathways. Chemically engineered oligonucleotides, termed “mimic” and “antagomirs”, and adenovirus that expresses specific sense or antisense miRNAs, have been developed and evaluated for their therapeutic effect on cardiac diseases (Wang & Yang, 2012). It is possible to use chemically modified oligonucleotide to target specific miRNAs and/or to disrupt the binding between a specific miRNA and a specific mRNA target. Dysregulation of cardiomyocyte miRNAs disturbs cardiac homeostasis by disrupting the cellular responses of cardiomyocytes to various signaling pathways. Recent studies have demonstrated that re-expression of downregulated anti-hypertrophic miRNAs or knockdown of upregulated pro-hypertrophic miRNAs is able to modulate cardiac remodeling, and serve as a promising therapeutic approach (Wang & Yang, 2012).

As tools for diagnosis and clinical evaluation of HCM improve, the lack of rationale therapies for this condition remains a major obstacle in care and management of HCM patients and families. However, genomics and transcriptomics inherent knowledge allow the development of new therapies to slow or prevent disease development. The identification of disease genes in several inherited cardiac diseases has raised expectations for new forms of treatment for some inherited cardiomyopathies there are realistic prospects that molecular insights will soon lead to novel treatments (Watkins *et al*., 2011). Despite these promising advances, further research continues to determine how the binding of multiple miRNAs affects

the expression of individual targeted mRNAs. Also, it is important to understand how multiple miRNA targets are interlinked to affect the various pathways and cardiac remodeling. A new concept- "competing endogenous RNA (ceRNA)" has emerged trying to explain how different types of RNAs interact to each other using MREs (Cheng *et al*., 2007).

The high frequency of novel genetic alterations, for instance in new genes, identified in genome of HCM patients trough the new sequencing technologies illustrate the difficulties in translating genetics complexity into the clinical framework. Also gene expression evaluation using transcriptomic arrays of mRNA and miRNA still fail to be informative at a diagnostic level. Genetic and transcriptional datasets correlated with clinical datasets will be important to understand how different HCM causal mutations produce HCM clinical manifestations (variable hypertrophy, fibrosis, diastolic dysfunction, arrhythmias and heart failure). Importantly, the knowledge obtained from the data integration of these datasets will be useful in the assessment of genotype–phenotype correlations by identifying low penetrance cases with important clinical implications on HCM therapeutic and prevention. Informatics tools are currently being used to store, integrate and analyse biological, genetic and clinical data.

5 Translational Bioinformatics Approaches in HCM Research

HCM is, as previously demonstrated, a complex disease whose study requires the effective integration and analysis of a number of heterogeneous features that originate from genotypic, phenotypic, and environmental sources. Computational approaches that take into account heterogeneous features, to improve clinical guidelines on treatments and disease prevention, are being developed in a very promising area of research named translational bioinformatics (Drolet & Lourenzi. 2011; Roque *et al*., 2011). This emerging research field is a discipline that was built on the successes of bioinformatics and health informatics for the study of complex diseases. The American Medical Informatics Association (AMIA), which considers the new field as a third major domain of informatics, defined translational bioinformatics as "... the development of storage, analytic, and interpretive methods to optimize the transformation of increasingly voluminous biomedical data into proactive, predictive, preventative, and participatory health." (Butte, 2009).

In this section we describe how computational approaches, that take advantages of Semantic Web technologies (Berners-Lee *et al*., 2001), are being developed to model, integrate and analyse large volumes of heterogeneous data, in order to provide new insights for the diagnosis and prognosis of complex diseases like HCM.

Semantic Web and Linked Open Data technologies are being adopted in science and business to overcome the limitations of the conventional data integration approaches and make data as open as possible (see, http://www.w3.org/). During the last decade, several approaches have been put into practice to integrate heterogeneous data sources for the domains of biomedicine, medicine and bioinformatics (Chen *et al*., 2013; Goble & Stevens, 2008; Antezana *et al*., 2009a; Agorastos *et al*., 2009). These new technologies open a new dimension to data integration, a big current challenge in biological and biomedical knowledge management (Attwood *et al*., 2009; Antezana *et al*., 2009b). The first approaches that have been developed, using Semantic Web technologies, aimed only to browse, visualize and search RDF data (RDF Primer http://www.w3.org/TR/2004/REC-rdf-primer-20040210/). Now, new developments already include more powerful tools such as the ones required to perform hypothesis testing.

Decision support based on translational bioinformatics means better information and workflow management, efficient literature and resource retrieval, and communication improvement. As new high-throughput typing and sequencing technologies gain popularity and provide unprecedented opportunities to characterize individual genomic landscapes and identify mutations relevant for diagnosis and therapy, data analysis poses significant challenges that led to the development of a large number of tools supporting specific parts of the analysis workflow or providing a complete solution. A typical computational workflow in this context encompasses four stages: (1) Data capture; (2) Data integration; (3) Data Analysis; and (4) Data visualization and reporting.

Acquisition and dissemination of genotypic, phenotypic and clinical data requires the availability and cross compatibility of simple and cost-effective information management systems, to enable more rapid adoption of data collection standards and broader use of comprehensive data and metadata tracking. In a recently proposed translational bioinformatics approach for HCM, the authors focused on the identification of associations between clinical and genotypic data with the objective of helping physicians predict the likely outcome of the disease, for every individual patient, with respect to the occurrence of a sudden cardiac death event (Machado *et al*., 2012). The proposed framework integrates clinical and genetic data mediated by a semantic data model representing the disease and explores data mining models depicting the clinical-genetic associations.

Still in the context of cardiovascular diseases, another work reasoned that an integrative genomics-phenomics approach could expedite disease candidate gene identification and prioritization. To approach the problem of inferring likely causality roles, the authors generated Semantic Web methods-based network data structures and performed centrality analyses to rank genes according to model-driven semantic relationships (Gudivada *et al.,* 2008). In a different direction, methods have also been proposed to exploit bio-ontologies for guiding data selection within the preparation step of the Knowledge Discovery in Databases (KDD) process. For familial hypercholesterolemia dataset, three scenarios have been proposed in which domain knowledge and ontology elements such as properties and class descriptions have been taken into account for data selection, before the data mining step (Coulet *et al*., 2008). In the context of cerebrovascular diseases, the Neuroweb European Project (Colombo *et al.,* 2010) was created to support genetic association studies, through the integration of clinical and genetic databases from four clinical institutions. The data is maintained in relational format and at its original location, thus in accordance with a database federation approach. All these systems envisage supporting the development of an open network of data, improving interoperability.

Regarding data integration, it does not only provide better data access, but also improves knowledge representation. Ontologies and controlled vocabularies play an important role at this stage since they provide a standard way of representing knowledge. Usually, these vocabularies are references accepted by a specific user community, such as the Gene Ontology (Ashburner *et al*., 2000) and the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) (http://www.ihtsdo.org/snomed-ct/). Strategies based on multiple vocabularies have also been developed, namely in pharmacogenomics considering the Human Disease Ontology and the Pharmacogenomics Knowledge Base (Hoehndorf *et al*., 2012). New efforts like the Cardiovascular Gene Ontology Annotation Initiative encourage the creation of an information-rich resource for the cardiovascular-research community, enabling researchers to rapidly evaluate and interpret existing data (http://www.ucl.ac.uk/cardiovasculargeneontology).

Most of the times the data integration stage starts with the creation of a semantic model for the domain under study, however one of the important features of using semantic web technologies is that

they enable linking the data in a way that the structure that represents the knowledge emerges from the data. This means that the semantic model can emerge from the links between the data.

Usually, building a semantic model for a given disease includes collecting its domain knowledge, e.g. clinical and genetic, in the form of keywords and scientific publications directly from the biomedical experts. This domain knowledge is used to identify relevant biomedical ontologies, to which modularization techniques can be applied in order to extract the modules that contain the concepts of interest. Previous work on semantic modelling on HCM (Machado *et al*., 2012), has shown that the semantic data model provides a useful framework for the integration of data not only from two heterogeneous domains of knowledge, clinics and genetics, but also from different medical institutions and research groups. In this work the authors propose a semantic model with three modules: 1) *Genotype Analysis* (with 19 concepts and approximately 39 properties), containing concepts associated with the genetic testing of biological samples; 2) *Medical Classifications* (with two high-level concepts: Angina Classification and Heart Failure Classification), containing medical standards used in the characterization of clinical elements such as patient symptoms; and 3) *Clinical Evaluation* (with a total of 63 concepts and approximately 60 object and data properties), that is the main module and that imports the other two. This last module additionally contains administrative concepts and clinical data elements that play an important role in the diagnosis and the prognosis of HCM patients. All the concepts have also been mapped to external controlled vocabularies, such as ontologies, to facilitate the interaction with other systems. Controlled vocabularies like SNOMED CT (version 2010 01 31), the National Cancer Institute Thesaurus (NCIt) (version 10.03), the Gene Regulation Ontology (version 0.5, released on 04 20 2010) and to the Sequence Ontology (released on 11 22 2011) have been considered. The concepts modelled in this work have also been identified and defined with the help of physicians, geneticists and molecular biologists based on the data elements collected during their activities. The model was populated with data from four medical institutions and two research centers. In opposition to ontology modularization state-of-the-art approaches, the work presented by Machado et al. (Machado *et al*., 2012) did not rely on the exploitation of a single biomedical ontology (Wennerberg *et al*., 2011), since clinical and genetic conceptualizations have been collected from multiple ontologies. In these scenarios, methods like local evidence content (Couto *et al*., 2005), ontology matching (Pesquita *et al*., 2011) and disjunctive common ancestors identification (Couto *et al*., 2011), can be used to retrieve concepts and keywords from different places. Although these techniques can help in building a semantic model from scratch, it is very important and useful if semantic models can be reused from other diseases or if manually developed models, validated by experts, are made available (Machado *et al*., 2010). For data storage and access purposes, this framework makes use of the sdlink system (http:/kdbio.inesc-id.pt/sdlink/), which allows the input of the data according to a semantic model or ontology. This system was developed based on Semantic Web technologies (Francisco *et al*., 2012), and its utilization as a clinical decision support system is under evaluation by physicians and biomedical experts.

Many workflows do not consider data integration and analysis as being part of two different stages, since data-mining techniques can be used both for knowledge extraction to support data representation and to obtain correlations between heterogeneous data sources that will help on the clinical diagnosis. In this context we are considering that upon completion of the data integration stage, the data is analysed by using data mining techniques in order to infer genotype-phenotype correlations, or, more specifically, to develop models for the association between the presence of certain mutations and the resulting physical traits.

In the case of HCM data analysis, the data elements that usually need to be integrated correspond to the presence/absence of each mutation in the genome of the patients (genotypic data) and to the clinical elements upon which the clinicians rely to provide a diagnose (phenotypic and clinical data). The latter normally include the results from physical examinations (e.g. electrocardiogram, echocardiogram), as well as the clinical history of the individual (e.g. age at diagnosis, sudden deaths in the family). Since there is a dynamic aspect associated with this type of disease, an object *date* was included in the top concepts (Person, Procedure, Clinical Finding, Health Care Site and Observable Entity) of the *Clinical Evaluation* module. By detailing a date to every Clinical Finding, that describes a specific stage of the disease, and to the Procedure or Observable entity, which can be a physical examination or a cardiovascular measurement, respectively, it is possible to analyse the evolution of the clinical scenario for each patient.

Although a large number of studies have been conducted on the linkage between specific mutations and the risk of specific illnesses (Roque *et al*., 2011; Aslam *et al*., 2011), models for the more general case of genotype to phenotype association in the presence of high disease complexity, both genetic and clinical, remain largely unexplored. Supervised machine learning techniques, such as decision trees and support vector machines, offer the potential to identify more complex relationships than those identified using simple correlation analysis, the standard practice in genotype-phenotype association models. Standard statistical analysis may identify correlations between one or a small set of specific mutations, but in more complex cases these correlations will not be significant enough to lead to concrete diagnosis methods. The models obtained using data mining techniques are expected to be of great interest both in terms of their predictive ability and their practical usability for physicians.

The exponential growth of genomic and also transcriptomic data, along with parallel achievements in acquiring and analyzing clinical data position the biomedical research enterprise to deliver on the promise of personalize medicine. Presently, data generated by new tests overwhelms current information technology systems and human interpretation capabilities. The need for sophisticated data analysis tools is leading the genetics industry to a fundamental shift from a clinical science focus to a translational bioinformatics focus.

6 Key Points and Concluding Remarks

HCM pathogenic mutations remain elusive in 30% to 40% of investigated HCM patients. This may be explained by the limitations of genotyping strategies with respect to yet-to-be discovered genes. Although cardiac hypertrophy is initially compensatory and beneficial, prolongation of this process leads to deleterious outcomes such as heart failure, arrhythmia, and in last instance SCD. Cellular and molecular studies to enquire about cardiac process of adaptation suggest that the genetic mutations cause functional defects that activate signaling molecules that ultimately prompts for cardiac remodeling adaptation. One fundamental issue in HCM has been whether cardiac remodeling encompassing cardiac hypertrophy and fibrosis, once established, can be reversed or prevented. Accordingly, the elucidation of molecular mechanisms underlying key processes generating cardiac hypertrophy is an important subject of intense research from a clinical point of view. The identification of key molecules that contribute to explain HCM phenotypic heterogeneity may provide important clues about the mechanisms by which HCM causes heart cellular remodeling. The knowledge about specific mRNA, miRNAs and their regulated networks will allow using these molecules as key molecular biomarkers related with their function in HCM and the molecular mechanism underlying cardiac hypertrophy and heart failure. Biological heterogeneity has

considerable impact on the translation of basic discoveries into clinical ascertainment of genetic cause of human cardiomyopathies. Computational approaches that take into account heterogeneous features for better providing clinical guidelines on treatments and disease prevention are being developed in a very promising area of research named translational bioinformatics. By using new semantic web based knowledge systems it is possible to have an integrated environment for querying, retrieving and analysing linked data and integrate heterogeneous data resources. In the case of HCM data analysis, the data elements that usually need to be integrated correspond to the presence/absence of each mutation in the genome of the patients (genotypic data), transcriptomic data (at the level of mRNA and miRNA) and to the clinical elements upon which the clinicians rely to provide a diagnose (phenotypic and clinical data). Several factors can impact on the availability and utility of genetic testing, such as access to testing, cost of testing and the mutation detection rate. These factors vary greatly depending on the disease and the technology used, nevertheless, genetic testing for HCM is available now at a much lower cost and timeframe and has a relatively high probability (up to 75%) of finding a mutation in a proband (Teekakirikul *et al*., 2013). Incumbent on the professional ordering genetic testing for HCM is the need to be skilled in interpreting the genetic test results and the consequent counseling based on the integration of the results (positive or negative), the family history, the clinical data of the patient, and any other known affected or unaffected family members. Ideally, the practitioner will also be skilled in the management of the clinical aspects of HCM, integrating the clinical, diagnostic, and therapeutic recommendations based on a synthesis of all data.

Genetic testing for HCM has seen major advancements in recent years with the introduction of new technologies for mutation detection, such as genome-wide sequencing which promises to rapidly increase the rate of variant detection and should enable more families to acquire genotype results. However, this has brought also new challenges for sequence variant interpretation and many variants of uncertain significance are likely to be found. Nevertheles, the positive identification of a mutation or a negative result in a family member can bring great benefits to management of HCM due to a better surveillance programme in the first case and less frequent intervals for clinical screening due to the reduced evidence of genetic risk.The understanding of HCM clinical and genetic heterogeneity and the development of new targeted therapies and accurate genetic diagnosis depends of a multidisciplinary approach that integrates experts in the fields of basic research, pathology, cardiology, medical geneticists and bioinformaticians working closely to bring cutting-edge research advances to manage and improve HCM patient care.

Acknowledgement

Isabel Carreira, Manuel Antunes, Carolino Monteiro, Nuno Cardim, Isabel Gaspar, Cátia Machado and Francisco Couto for their contribution to our work.

References

Aatre, R.D., Day, S.M. (2011). Psychological issues in genetic testing for inherited cardiovascular diseases. Circ Cardiovasc Genet (4):81-90.

Ackerman, M.J., Priori, S.G., Willems, S., Berul, C., Brugada, R., Calkins, H., Camm, A.J., Ellinor, P.T., Gollob, M., Hamilton, R., Hershberger, R.E., Judge, D.P., Le Marec, H., McKenna, W.J., Schulze-Bahr, E., Semsarian, C., Towbin, J.A., Watkins, H., Wilde, A., Wolpert, C., Zipes, D.P. (2011). HRS/EHRA expert consensus statement on the state of

genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). Heart Rhythm (8):1308-1339.

Agorastos, T., Koutkias, V., Falelakis, M., Lekka, I., Mikos, T., Delopoulos, A., Mitkas, P.A., Tantsis, A., Weyers, S., Coorevits, P., Kaufmann, A.M., Kurzeja, R. & Maglaveras, N. (2009). Semantic Integration of Cervical Cancer Data Repositories to Facilitate Multicenter Association Studies: the ASSIST Approach. Cancer Informatics. (8): 31-44.

Alcalai, R., Seidman, J.G., Seidman, C.E. (2008). Genetic basis of hypertrophic cardiomyopathy: From bench to the clinics. J Cardiovasc Electrophysiol (19):104-110.

Alders, M., Jongbloed, R., Deelen, W., van den Wijngaard, A., Doevendans, P., Ten Cate, F., Regitz-Zagrosek, V., Vosberg, H.P., van Langen, I., Wilde, A., et al (2003). The 2373insG mutation in the MYBPC3 gene is a founder mutation, which accounts for nearly one-fourth of the HCM cases in the Netherlands. Eur Heart J (24):1848-1853.

Ahmad F., Seidman J.G., Seidman C.E. (2005). The genetic basis for cardiac remodeling. Annu Rev Genomics Hum Genet (6):185-216.

Andersen, P.S., Havndrup, O., Bundgaard, H., Larsen, L.A., Vuust, J., Pedersen, A.K., Kjeldsenm K., Christiansen, M. (2004). Genetic and phenotypic characterization of mutations in myosin-binding protein C (MYBPC3) in 81 families with familial hypertrophic cardiomyopathy: total or partial haploinsufficiency. Eur J Hum Genet (12):673-677.

Antezana, E., Blond, W., Egaa, M., Rutherford, A., Stevens, R., De Baets, B., Mironov, V., & Kuiper, M. (2009a). BioGateway: a semantic systems biology tool for the life sciences. BMC Bioinformatics, (10):S11

Antezana, E., Kuiper, M., & Mironov, V. (2009b). Biological knowledge management: the emerging role of the Semantic Web technologies Briefings on Bioinformatics (10): 392-407

Arad, M., Maron, B.J., Gorham, J.M., Johnson, W.H. Jr, Saul, J.P., Perez-Atayde, A.R., Spirito, P., Wright, G.B., Kanter, R.J., Seidman, C.E., Seidman, J.G. (2005). Glycogen storage diseases presenting as hypertrophic cardiomyopathy. N Engl J Med (352):362-37214.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel- Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. & Sherlock, G. (2000). Gene Ontology: Tool for the Unification of Biology. Nature Genetics (25): 25-29.

Attwood, T.K., Kell, D.B., McDermott, P., Marsh, J., Pettifer, S.R., & Thorne, D. (2009). Calling International Rescue: knowledge lost in literature and data landslide! Biochemical Journal (424): 317-333.

Barrans, J.D., Allen, P.D, Stamatiou, D., Dzau, V.J. Liew C.C. (2002). Global Gene Expression Profiling of End-Stage Dilated Cardiomyopathy Using a Human Cardiovascular-Based cDNA Microarray. Am J Pathol. (160):2035-4.

Basso, C., Thiene, G., Corrado, D., Buja, G., Melacini, P., Nava A. (2000). Hypertrophic cardiomyopathy and sudden death in the young: pathologic evidence of myocardial ischemia. Human pathology.(31):988–998.

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The SemanticWeb. Scientific American, 29-37.

Biagini, E., Coccolo, F., Ferlito, M., Perugini, E., Rocchi, G., Bacchi-Reggiani, L., Lofiego, C., Boriani, G., Prandstraller, D., Picchio, F.M., Branzi, A.,Rapezzi, C. (2005) Dilated-hypokinetic evolution of hypertrophic cardiomyopathy: prevalence, incidence, risk factors, and prognostic implications in pediatric and adult patients. J Am Coll Cardiol (46): 1543-1550

Butte, AJ. (2009). Translational bioinformatics applications in genome medicine. Genome Medicine (1):64

Bookwalter, C.S., Trybus, K.M. (2006). Functional consequences of a mutation in an expressed human a-cardiac actin at a site implicated in familial hypertrophic cardiomyopathy. J Biol Chem (281):16777–16784.

Bos, J.M., Poley, R.N., Ny, M., Tester, D.J., Xu, X., Vatta, M., Towbin, J.A., Gersh, B.J, Ommen, S.R, Ackerman, M.J. (2006). Genotype phenotype relationships involving hypertrophic cardiomyopathyassociated mutations in titin, muscle LIM protein, and telethonin. Mol Genet Metab (88):78–85.

Carniel, E., Taylor, M.R., Sinagra, G., Di Lenarda, A., Ku, L., Fain, P.R., Boucek, M.M., Cavanaugh, J., Miocic, S., Slavov, D., Graw, S.L., Feiger, J., Zhu, X.Z., Dao, D., Ferguson, D.A., Bristow M,R., Mestroni, L. (2005). Alpha-

myosin heavy chain: a sarcomeric gene associated with dilated and hypertrophic phenotypes of cardiomyopathy. Circulation. (5);112(1):54-9.

Chang, A.N., Harada, K., Ackerman, M.J., Potter, J.D. (2005). Functional consequences of hypertrophic and dilated cardiomyopathy-causing mutations in alfa-tropomyosin. J Biol Chem (280):34343–34349.

Charron, P. , Dubourg, O., Desnos, M., Isnard, R., Hagege, A., Bonne, G., Carrier, L., Tesson, F., Bouhour, J.B., Buzzi, J.C., Feingold, J., Schwartz, K., Komajda, M. (1998) Genotype–phenotype correlations in familial hypertrophic cardiomyopathy: a comparison between mutations in the cardiac protein-C and the β-myosin heavy chain genes. Eur Heart J (19): 139–145

Chen, H., Yu, T., & Chen, J.Y. (2013). Semantic Web meets Integrative Biology: a survey Briefings on Bioinformatics (14): 109-125.

Cheng, Y., Ji, R., Yue, J., Yang, J., Liu, X., Chen, H., Dean, D.B., Zhang C. (2007). MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy? Am. J. Pathol. (170):1831-40.

Cirino, A.L., & Ho, C.Y. (2011). Familial Hypertrophic Cardiomyopathy Overview. GeneReviews™ Pagon RA, Adam MP, Bird TD, editors. Seattle, University of Washington.

Clerk, A., Cullingford, T.E., Fuller, S.J., Giraldo, A., Markou, T., Pikkarainen, S., Sugden, P.H.. (2007). Signaling pathways mediating cardiac myocyte gene expression in physiological and stress responses. J Cell Physiol. (212):311-22.

Colombo, G., Merico, D., Boncoraglio, G., Paoli, F.D., Ellul, J., Frisoni, G., Nagy, Z., van der Lugt, A., Vassanyi, I. & Antoniotti, M. (2010). An Ontological Modeling Approach to Cerebrovascular Disease Studies: the NEUROWEB Case. Journal Biomedical Informatics (43): 469-484

Cooper, T.A. (2005). Alternative Splicing Regulation Impacts Heart Development. Cell. (120):1-2.

Coulet, A., Smail-Tabbone, M., Benlian, P., Napoli, A., & Devignes, M., (2008). Ontology-guided data preparation for discovering genotype-phenotype relationships. BMC Bioinformatics (9): S3

Couto, F.M., Silva, M.J. & Coutinho, P.M. (2005). Finding genomic ontology terms in text using evidence content. BMC Bioinformatics (6): S21

Dhandapany, P.S., Sadayappan, S., Xue, Y., Powell, G.T., Rani, D.S., Nallari, P., Rai, T.S., Khullar, M., Soares, P., Bahl, A., et al (2009). A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. Nat Genet (41):187-191.

Divakaran, V. & Mann, D.L. (2008).The Emerging Role of MicroRNAs in Cardiac Remodeling and Heart Failure. Circulation Research.(103): 1072-1083.

Drolet, BC. & Lorenzi, NM. (2011). Translational research: understanding the continuum from bench to bedside, Translational Research (157): 1-5.

Erdmann, J., Daehmlow, S., Wischke, S., Senyuva, Mm, Werner, U., Raible, J., Tanis, N., Dyachenko, S., Hummel, M., Hetzer, R., Regitz-Zagrosek, V. (2003). Mutation spectrum in a large cohort of unrelated consecutive patients with hypertrophic cardiomyopathy. Clin Genet. (64): 339-349.

Fatkin, D, Graham, R.M. (2002). Molecular Mechanisms of Inherited Cardiomyopathies. Physiol Rev (82): 945–980.

Fernandes, A.R., Marques, V., Nunes, A.C., Freitas, A.T., Gouveia, M.R., Antunes, M., Carreira, I.M., Gaspar, I.M., Monteiro, C., Santos S. (2013). microRNA Transcriptional Evaluation in Obstructive Hypertrophic Cardiomyopathy: a preliminary study using human myectomy samples. Cardiac Remodeling, Signaling Matrix and Heart function. Flash Presentation. Keystone Symposia 2013.

Francisco A., Reis P.M., Abdulrehman D., Vaz C., Santos M., & Freitas A.T. (2012). sdlink: An Integrated System for Linking Biological and Biomedical Semantic Data. Conference on Semantics in Healthcare and Life Sciences (CSHALS).

Friddle, C.J., Koga, T., Rubin, E.M., Bristow, J. (2000). Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. Proc Natl Acad Sci U S A. (97):6745–50.

Geier, C., Gehmlich, K., Ehler, E., Hassfeld, S., Perrot, A., Hayess, K., Cardim, N., Wenzel, K., Erdmann, B., Krackhardt, F., Posch, M.G., Osterziel, K.J., Bublak, A., Nagele, H., Scheffold, T., Dietz, R., Chien, K.R., Spuler, S., Furst, D.O., Nurnberg, P., Ozcelik, C. (2008). Beyond the sarcomere: cSRP3 mutations cause hypertrophic cardiomyopathy. Hum Mol Genet (17):2753-2765.

Gersh, B.J., Maron, B.J., Bonow, R.O., Dearani, J.A., Fifer, M.A., Link, M.S., Naidu, S.S., Nishimura, R.A., Ommen, S.R., Rakowski, H., Seidman, C.E., Towbin, J.A., Udelson, J.E., Yancy, C.W. (2011). ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: executive summary: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation (124): 2761–2796.

Girolami, F. Olivotto, I. Passerini,. I. , Zachara, E., Nistri, S., Re, F., Fantini, S., Baldini, K., Torricelli, F., Cecchi, F.(2006). A molecular screening strategy based on beta-myosin heavy chain, cardiac myosin binding protein C and troponin T genes in Italian patients with hypertrophic cardiomyopathy. J Cardiovasc Med (Hagerstown). (7:)601-607

Goble, C. & Stevens R. (2008). State of the nation in data integration for bioinformatics. Journal of Biomedical Informatics (41): 687-693.

Gudivada, R.C., Qu, X.A., Chen, J., Jegga, A.G., Neumann, E.K. & Aronow, B.J. (2008). Identifying Disease-Causal Genes Using Semantic Web-based Representation of Integrated Genomic and Phenomic Knowledge. Journal Biomedical Informatics (41): 717-729.

Harris, K.M., Spirito, P., Maron, M.S., Zenovich, A.G., Formisano, F., Lesser, J.R., Mackey-Bojack, S., Manning, W.J., Udelson, J.E., Maron, B.J. (2006). Prevalence, clinical profile, and significance of left ventricular remodeling in the end-stage phase of hypertrophic cardiomyopathy. Circulation (114): 216-225.

Harvey, P.A. & Leinwand, L.A. (2011). Cellular mechanisms of cardiomyopathy. J. Cell Biol. (194):355–365.

Hayashi, T., Arimura, T., Itoh-Satoh, M., Ueda, K., Hohda, S., Inagaki, N., Takahashi, M., Hori, H., Yasunami, M., Nishi, H., Koga, Y., Nakamura, H., Matsuzaki, M., Choi, B.Y., Bae, S.W., You, C.W., Han, K.H., Park, J.E., Knöll, R., Hoshijima, M., Chien, K.R., Kimura, A. (2004). Tcap gene mutations in hypertrophic cardiomyopathy and dilated cardiomyopathy. J Am Coll Cardiol (44):2192–2201.

Heller, M.J., Nili, M., Homsher, E., Tobacman, L.S. (2003). Cardiomyopathic tropomyosin mutations that increase thin filament Ca2þ sensitivity and tropomyosin N-domain flexibility. J Biol Chem (278):41742–41748.

Hershberger, R.E., Lindenfeld, J., Mestroni, L.,Seidman, C.E. Taylor, M.R.G., Towbin, J.A. (2009). Genetic Evaluation of Cardiomyopathy - A Heart Failure Society of America Practice Guideline. Journal of Cardiac Failure (15): 83-97.

Ho, C.Y. (2010). Genetics and clinical destiny: improving care in hypertrophic cardiomyopathy. Circulation (122):2430-2440.

Ho, C.Y. (2010a). Hypertrophic cardiomyopathy. Heart Fail Clin (6):141-159.

Hoehndorf, R., Dumontier, M. & Gkoutos, G. (2012). Identifying aberrant pathways through integrated analysis of knowledge in pharmacogenomics. Bioinformatics, (28) :2169-2175.

Hwang, J.J, Allen, P.D., Tseng, G.C, Lam, C.W, Fananapazir, L., Dzaul, V.J., Liew, C.C. (2002). Microarray gene expression profiles in dilated and hypertrophic cardiomyopathic end-stage heart failure. Physiol Genomics. (10): 31–44.

Ingles, J., Doolan, A., Chiu, C., Seidman, J., Seidman, C., Semsarian, C. (2005). Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. J Med Genet. (42):10,e59.

Ionita-Laza, I., Lange, C., Laird, M.N. (2009). Estimating the number of unseen variants in the human genome. Proc Natl Acad Sci USA (106):5008-5013.

Karibe, A., Tobacman, L.S., Strand, J., Butters, C., Back, N., Bachinski, L.L., Arai, A.E., Ortiz, A., Roberts, R., Homsher, E., Fananapazir, L. (2001). Hypertrophic cardiomyopathy caused by a novel alpha-tropomyosin mutation (V95A) is associated with mild cardiac phenotype, abnormal calcium binding to troponin, abnormal myosin cycling, and poor prognosis.Circulation (103):65–71.

Keren, A., Syrris, P, McKenna, W.J. (2008). Hypertrophic cardiomyopathy: the genetic determinants of clinical disease expression. Nat Clin Pract Cardiovasc Med. (3):158-68.

Kehat, I., Molkentin, J.D. (2010). Molecular Pathways Underlying Cardiac Remodeling During Pathophysiological Stimulation. Circulation (122): 2727-2735.

Kim, J.B., Porreca, G.J., Song, L., Greenway, S.C., Gorham, J.M., Church G.M., Seidman, C.E., Seidman, J.G. (2007). Polony multiplex analysis of gene expression (PMAGE) in mouse hypertrophic cardiomyopathy. Science (316):1481–1484

Kimura A, Harada H, Park JE, Nishi H, Satoh M, Takahashi M et al. Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. Nat Genet 1997;16:379–382.

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.

Komuro, I., Yazak,i Y. (1993). Control of cardiac gene expression by mechanical stress. Annu Rev Physiol. (55): 55–75.

Landstrom, A.P., Adekola, B.A., Bos. J.M., Ommen, S.R, Ackerman, M.. (2011). PLN-encoded phospholamban mutation in a large cohort of hypertrophic cardiomyopathy cases: summary of the literature and implications for genetic testing. J Am Heart J.161(1):165-71

Landstrom, A.P., Ackerman, M.J. (2012). Beyond the cardiac myofilament: hypertrophic cardiomyopathy- associated mutations in genes that encode calcium-handling proteins. Curr Mol Med.12(5):507-18.

Lawler, P.R., Yongzhong, W., Steinbrüchel D., Blagoja, D., Paulsson-Berne, G., Kastrup, J., Hansson, G.K. (2007). Gene expression signals involved in ischemic injury, extracellular matrix composition and fibrosis defined by global mRNA profiling of the human left ventricular myocardium. Journal of Molecular and Cellular Cardiology. (42): 870–883

Lee, M.L., Kuo, F.C., Whitmore, G.A., Sklar, J. (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations.Proc Natl Acad Sci USA (97): 9834–9839.

Lim, D.S., Roberts, R., Marian, A.J. (2001). Expression profiling of cardiac genes in human hypertrophic cardiomyopathy: insight into the pathogenesis of phenotypes. J Am Coll Cardiol (38):1175–1180.

Lopes, L.R., Rahman, M.S., Elliott, P.M. (2013). A systematic review and meta-analysis of genotype–phenotype associations in patients with hypertrophic cardiomyopathy caused by sarcomeric protein mutations. Heart doi:10.1136/heartjnl-2013-303939.

Lopes, L.R., Zekavati, A., Syrris, P., Hubank, M., Giambartolomei, C., Dalageorgou, C., Jenkins, S., McKenna, W., Elliott, P.M.. (2013b). Genetic complexity in hypertrophic cardiomyopathy revealed by high-throughput sequencing. J Med Genet.;50(4):228-39.

Manabe, I., Shindo, T., Nagai, R. (2002). Gene expression in fibroblasts and fibrosis: involvement in cardiac hypertrophy. Circ Res. (91):1103–13

Machado, C., Couto, F.M., Fernandes, A.R., Santos, S., & Freitas, A.T. (2012). Toward a Translational Medicine Approach for Hypertrophic Cardiomyopathy. 3rd International Conference on Information Technology in Bio- and Medical Informatics, LNCS (7451): 151-165.

Machado. C., Couto, F.M., Fernandes, A.R., Santos, S., Cardim, N., & Freitas, A.T. (2010). Semantic characterization of hypertrophic cardiomyopathy disease. First Workshop on Knowledge Engineering, Discovery and Dissemination in Health (KEDDH).

Margulies, K.B & Matiwala, S. (2005). Molecular Mechanisms Of Cardiac Hypertrophy And Failure. Taylor and Francis. Edited by Richard A Walsh. Chapter 36: 797-816.

Marian, A.J. (2000). Pathogenesis of diverse clinical and pathological phenotypes in hypertrophic cardiomyopathy. Lancet (355):58-60.

Marian, A.J., Roberts, R. (2001). The molecular genetic basis for hypertrophic cardiomyopathy. J Mol Cell Cardiol (33):655.

Maron, B.J., Bonow, R.O., Cannon, R.O., Leon, M.B., Epstein, S.E. (1987). Hypertrophic cardiomyopathy: interrelations of clinical manifestations, pathophysiology, and therapy. N Engl J Med. (316):780-9.

Maron, B.J., Casey, S.A., Hauser, R.G., Aeppli, D.M. (2003). Clinical course of hypertrophic cardiomyopathy with survival to advanced age. J Am. Coll Cardiol (42):882-888.

Maron, B.J., Gardin, J.M., Flack, J.M., Gidding, S.S., Kurosaki, T.T., Bild, D.E. (1995). Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. Circulation. (92):785–789.

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.

Maron, M.S., Olivotto, I., Zenovich, A.G., Link, M.S., Pandian, N.G., Kuvin, J.T., Nistri, S., Cecchi, F., Udelson, J.E., Maron, B.J. (2006). Hypertrophic cardiomyopathy is predominantly a disease of left ventricular outflow tract obstruction. Circulation (114): 2232–2239.

Maron, B.J., Seidman, J.G., Seidman, C.E. (2004). Proposal for contemporary screening strategies in families with hypertrophic cardiomyopathy. J Am Coll Cardiol. (44): 2125–2132.

Maron BJ, Semsarian C. (2010). Emergence of gene mutation carriers and the expanding disease spectrum of hypertrophic cardiomyopathy. Eur Heart J. (31): 1551-3

Marston, S., Copeland, O., Jacques, A., Livesey, K., Tsang, V., McKenna, W.J., Jalilzadeh, S., Carballo, S., Redwood,, C., Watkins, H. (2009). Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency. Circ Res. (105):219-22.

Millat, G., Chanavat, V., Crehalet, H., Rousson, R. (2010). Development of a high resolution melting method for the detection of genetic variations in hypertrophic cardiomyopathy. Clin Chim Acta Int J Clin Chem (411):1983-1991.

Miller, S.W. (2009). Cardiac Imaging: The Requisites. By Stephen W. Miller, Suhny Abbara, Lawrence Boxt, Mosby Elsevier, Philadelphia.

Mogensen, J., Murphy, R.T., Kubo, T., Bahl, A., Moon, J.C., Klausen, I.C., Elliott, P.M., McKenna, W.J. (2004). Frequency and clinical expression of cardiac troponin I mutations in 748 consecutive families with hypertrophic cardiomyopathy. J Am Coll Cardiol (44):2315–2531.

Moolman, J.C., Corfield, V.A., Posen, B., Ngumbela, K., Seidman, C., Brink, P.A., Watkins, H. (1997). Sudden death due to troponin T mutations. J Am Coll Cardiol (29): 549–555 32.

Moon, J.C., McKenna, W.J., McCrohon, J.A., Elliott, P.M., Smith, G.C., Pennell, D.J. (2003). Toward clinical risk assessment in hypertrophic cardiomyopathy with gadolinium cardiovascular magnetic resonance. J Am Coll Cardiol. (41):1561-7.

Morita, H., Rehm, H.L., Menesses, A., McDonough, B., Roberts, A.E., Kucherlapati, R., Towbin, J.A., Seidman, J.G., Seidman, C.E. (2008). Shared genetic causes of cardiac hypertrophy in children and adults. N Engl J Med (358):1899-1908.

Niimura, H., Bachinski, L.L., Sangwatanaroj, S., Watkins, H., Chudley, A.E., McKenna, W., Kristinsson, A., Roberts, R., Sole, M., Maron, B.J.,Seidman, J.G., Seidman, C.E. (1998). Mutations in the gene for cardiac myosin-binding protein C and late- onset familial hypertrophic cardiomyopathy. N Engl J Med. 338:1248–1257.

Niimura, H., Patton, K.K., McKenna, W.J., Soults, J., Maron, B.J., Seidman, J.G., Seidman, C.E. (2002). Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. Circulation (105):446-451.

Nishikawa, T., Ishiyama, S., Shimojo, T., Takeda, K., Kasajima, T., Momma, K. (1996). Hypertrophic cardiomyopathy in Noonan syndrome. Acta Paediatr Jpn (38):91-98.

Olivotto, I., Cecchi, F., Casey, S.A., Dolara, A., Traverse, J.H, Maron, B.J. (2001). Impact of atrial fibrillation on the clinical course of Hypertrophic Cardiomyopathy. Circulation. (104):2517-24.

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):1687–94.

Osio, A., Tan, L., Chen, S.N., Lombardi, R., Nagueh, S.F., Shete, S., Roberts, R., Willerson, J.T., Marian, A.J. (2007). Myozenin 2 is a novel gene for human hypertrophic cardiomyopathy. Circ Res (100):766–768.

Osterziel, K.J., Bit-Avragim, N., Bunse, M. (2002). Cardiac hypertrophy in Friedreich's ataxia. Cardiovasc Res (54):694.

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.

Porter, K.E., Turner, N.A. (2009). Cardiac fibroblasts: at the heart of myocardial remodeling . Pharmacol Ther. (123):255-78.

Posch,, M.G., Thiemann, L., Tomasov, P., Veselka,J., Cardim, N., Garcia-Castro, M., Coto, E., Perrot, A., Geier, C., Dietz, R., Haverkamp, W., Ozcelik, C. (2008). Sequence analysis of myozenin 2 in 438 european patients with familial hypertrophic cardiomyopathy. Med Sci Monit. (14): CR372-CR374.

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., Benaiche, A., Isnard, R., Dubourg, O., Burban, M., Gueffet, J.P., Millaire, A., Desnos, M., Schwartz, K., Hainque, B., Komajda, M.; EUROGENE Heart Failure Project. (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation (107):2227-2232.

Rooij, E. & Olson, E.N. (2007). MicroRNAs: powerful new regulators of heart disease and provocative therapeutic targets. J Clin Invest (117):2369–2376.

Roque, F.S., Jensen, P.B., Schmock, H., Dalgaard. M., Andreatta. M., et al. (2011) Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts. PLoS Computational Biology (7): e1002141.

Rozenfeld, P., Neumann, P.M. (2011). Treatment of fabry disease: current and emerging strategies. Curr Pharm Biotechnol (12):916-922.

Sachdev, B., Takenaka, T., Teraguchi, H., Tei, C., Lee, P., McKenna, W.J., Elliott, P.M. (2002). Prevalence of Anderson-Fabry disease in male patients with late onset hypertrophic cardiomyopathy. Circulation (105): 1407-1411.

Sadoshima, J., Izumo, S. (1997). The cellular and molecular response of cardiac myocytes to mechanical stress. Annu Rev Physiol. (59):551–71.

Santos, S., Cavaco, D., Adragão, P., Sá, I.; Carreira, I.; Antunes, M., Cardim, N., Monteiro, C. (2009). Familial mutation screening and gene expression evaluation in hypertrophic cardiomyopathy profiling: implications for a molecular diagnosis strategy. Oral presentation. 28th Annual Scientific Meeting of the Belgian Society of Cardiology.

Santos, S., Fernandes, A.R., Freitas, A.T., Machado, C.M, Branco, P., Silveira, L., Carreira, I., Antunes, M., Monteiro, C. (2010). HCM sarcomere gene expression analysis: a machine learning approach". Key Note Speaker. Advances in qPCR European Conference, Select Biosciences. Dublin, Ireland.

Santos, S, Lança, V., Oliveira, H., Branco, P., Silveira, L., Marques, V., Brito, D., Madeira, H., Bicho, M., Fernandes, A.R. (2011). "Genetic diagnostic of hypertrophic cardiomyopathy using Mass Spectrometry and High Resolution Melting". Rev Port Cardiol. (30):7-18.

Santos, S., Marques, V., Pires, M., Nunes, A.C., Freitas, A.T., Gouveia, R., Antunes, M., Carreira, I., Gaspar, I.M., Monteiro, C., Fernandes, A.R. (2013). Novel insights in Hypertrophic Cardiomyopathy (HCM) evaluation: miRs as biomarkers of HCM cardiac remodeling. Flash Presentation. Heart Failure Congress. European Society of Cardiology.

Santos, S., Marques, V., Pires, M., Silveira, L., Oliveira, H., Lança, V., Brito, D., Madeira, H., Carreira, I.M., Gaspar, I.M., Monteiro, C., Fernandes, A.R. (2012). "High resolution melting: improvements in the genetic diagnosis of Hypertrophic Cardiomyopathy in a Portuguese cohort". BMC Medical Genetics (13):13-17.

Satoh, M., Takahashi, M., Sakamoto, T., Hiroe, M., Marumo, F., Kimura, A. (1999). Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. Biochem Biophys Res Commun (262):411–417.

Seidman, J.G. & Seidman, C. (2001). The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. Cell. Feb 23;104(4):557-67

Seidman, C.E. & Seidman, J.G. (2011). Identifying sarcomere gene mutations in hypertrophic cardiomyopathy: a personal history. Circ Res (108):743-750.

Sequeira, V., Wijnker, P.J.M., Nijenkamp, L.L.A.M., Kuster, D.W.D., Najafi, A., Witjas-Paalberends, E.R., Regan, J.A., Boontje, N., ten Cate, F.J., Germans, T., Carrier, L., Sadayappan, S., van Slegtenhorst, M.A., Zaremba, R., Foster, D.B., Murphy, A.M., Poggesi, C., dos Remedios, C., Stienen, G.J.M., Ho, C.Y., Michels, M. & van der Velden, J. (2013). Perturbed Length-Dependent Activation in Human Hypertrophic Cardiomyopathy With Missense Sarcomeric Gene Mutations Novelty and Significance. Circ Res. (112):1491-1505.

Semsarian, C., Ahmad, I., Giewat, M., Georgakopoulos, D., Schmitt, J.P., McConnell, B.K., Reiken, S., Mende, U., Marks, A.R., Kass, D.A., Seidman, C.E., Seidman, J.G. (2002). The L-type calcium channel inhibitor diltiazem prevents cardiomyopathy in a mouse model. J Clin Invest (109):1013-1020.

Tan, F.L., Moravec, C.S., Li, J., Apperson-Hansen, C, McCarthy, P.M., Young J.B.,, Bond, M. (2002). The gene expression fingerprint of human heart failure. Proc Natl Acad Sci USA. (99):11387–92.

Teekakirikul, P., Eminaga, S., Toka, O., Alcalai, R., Wang, L., Wakimoto, H., Nayor, M., Konno, T., Gorham, J.M., Wolf, C.M., Kim, J.B., Schmitt. J.P., Molkentin, J.D., Norris, R.A., Tager, A.M., Hoffman, S.R., Markwald, R.R., Seidman, C.E., Seidman, J.G. (2010). Cardiac fibrosis in mice with hypertrophic cardiomyopathy is mediated by non-myocyte proliferation and requires Tgf-b. J Clin Invest (120):3520-3529.

Teekakirikul, P, Kelly, M.A., Rehm, H.L., Lakdawala, N.K., & Funke, B.H. (2013).Inherited Cardiomyopathies Molecular Genetics and Clinical Genetic Testing in the Postgenomic Era. The Journal of Molecular Diagnostics (15):158-170.

Tian T., Liu Y., Zhou X. & Song L. (2013). Progress in the Molecular Genetics of Hypertrophic Cardiomyopathy: A Mini-Review. Gerontology (59):199–205.

Theis, J.L., Bos, J.M., Theis, J.D., Miller, D.V., Dearani, J.A., Schaff, H.V., Gersh, B.J., Ommen, S.R., Moss, R.L., Ackerman, M.J. (2009). Expression Patterns of Cardiac Myofilament Proteins - Genomic and Protein Analysis of Surgical Myectomy Tissue from Patients with Obstructive Hypertrophic Cardiomyopathy. Circ Heart Fail. (2):325-33.

Thierfelder, L., Watkins, H., MacRae, C., Lamas, R., McKenna, W., Vosberg, H.P., Seidman, J.G., Seidman, C. (1994). Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. Cell (77):701–12.

Thum, T. (2011). MicroRNA therapeuthics in cardiovascular Medicine. EMBO Mol Med (4): 3-14.

Thum, T, Catalucci, D., Bauersachs, J. (2008). MicroRNAs: novel regulators in cardiac development and disease. Cardiovascular Research. (79), 562–570.

Van Driest ,S.L., Ommen, S.R., Tajik, A.J., Gersh, B.J., Ackerman, M.J. (2005). Sarcomeric genotyping in hypertrophic cardiomyopathy. Mayo Clin Proc. (80):463-469.

Varnava, A.M., Elliott, P.M., Baboonian, C., Davison, F., Davies, M.J., McKenna, W.J. (2001). Hypertrophic cardiomyopathy: histopathological features of sudden death in cardiac troponin T disease. Circulation (104):1380–1384

Varnava, A.M., Elliott, P.M., Mahon, N., Davies, M.J., McKenna, W.J. (2001). Relation between myocyte disarray and outcome in hypertrophic cardiomyopathy. The American journal of cardiology. (88):275–279

Vang, S., Corydon, T.J., Borglum, A.D., Scott, M.D., Frydman, J., Mogensen, J., Gregersen, N., Bross, P. (2005). Actin mutations in hypertrophic and dilated cardiomyopathy cause inefficient protein folding and perturbed filament formation. Febs J (272):2037–2049.′

Vasile, V.C., Will, M.L., Ommen, S.R., Edwards, W.D., Olson, T.M., Ackerman, M.J. (2006). Identification of a metavinculin missense mutation, R975W, associated with both hypertrophic and dilated cardiomyopathy. Mol Genet Metab. (87):169-174.

Vikstrom, K.L., Bohlmeyer, T., Factor, S.M., Leinwand, L.A. (1998). Hypertrophy, pathology, and molecular markers of cardiac pathogenesis. Circ Res. (82):773–8.

Villard E, Perret C, Gary F, Proust C, Dilanian G, Hengstenberg C, Ruppert V, Arbustini E, Wichter T, Germain M, Dubourg O, Tavazzi L, Aumont MC, DeGroote P, Fauchier L, Trochu JN, Gibelin P, Aupetit JF, Stark K, Erdmann J, Hetzer R, Roberts AM, Barton PJ, Regitz-Zagrosek V; Cardiogenics Consortium, Aslam U, Duboscq-Bidot L, Meyborg M, Maisch B, Madeira H, Waldenström A, Galve E, Cleland JG, Dorent R, Roizes G, Zeller T, Blankenberg S, Goodall AH, Cook S, Tregouet DA, Tiret L, Isnard R, Komajda M, Charron P, Cambien F. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. (2011). European Heart Journal, (32): 1065-1076.

Wang, L., Seidman, J.G., Seidman, C.E. (2010). Narrative review: harnessing molecular genetics for the diagnosis and management of hypertrophic cardiomyopathy. Ann. Intern. Med. (152):513–520: W181.

Wang, J. & Yang, X. (2012). The function of miRNA in cardiac hypertrophy. Cell Mol Life Sci.(69): 3561–3570.

Watkins, H., Rosenzweig, A., Hwang, D.S., Levi, T., McKenna, W., Seidman, C.E., Seidman, J.G. (1992). Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy. N Engl J Med (326):1108–14.

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., Seidman, C.E. (1995). Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy. N Engl J Med (332):1058-1064.

Watkins, H., Conner, D., Thierfelder, L., Jarcho, J.A., MacRae, C., McKenna, W.J., Maron, B.J., Seidman, J.G., Seidman, C.E. (1995b). Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. Nat Genet (11):434–7.

Watkins, H., Ashrafian, H., Redwood, C. (2011). Inherited Cardiomyopathies. N Engl J Med. (364):1643-56

Wennerberg, P., Schulz, K., & Buitelaar, P. (2011). Ontology modularization to improve semantic medical image annotation. Journal Biomedical Informatics (44): 155-162.

Wheeler, M., Pavlovic, A., DeGoma, E., Salisbury, H., Brown, C., Ashley, E.A. (2009). A new era in clinical genetic testing for hypertrophic cardiomyopathy. J Cardiovasc Transl Res (2):381-391.

Witjas-Paalberends, E.R. Piroddi, N., Stam, K., van Dijk, S.J., Sequeira, V., Ferrara, C., Scellini, B., Hazebroek, M., ten Cate, F.J., van Slegtenhorst, M., dos Remedios, C., Niessen, H.W.M., Tesi, C., Stienen, G.J.M., Heymans, S., Michels, M., Poggesi, C., & van der Velden, J. (2013). Mutations in MYH7 reduce the force generating capacity of sarcomeres in human familial hypertrophic cardiomyopathy. Cardiovascular Research doi: 10.1093/cvr/cvt119.

Wong, W.W., Doyle, T.C., Cheung, P., Olson, T.M., Reisler, E. (2001). Functional studies of yeast actin mutants corresponding to human cardiomyopathy mutations. J Muscle Res Cell Motil (22):665–674

Yasumura, Y, Takemura, K, Sakamoto, A, Kitakaze, M, Miyatake, K. (2003). ‚Journal of Cardiac Failure. Changes in myocardial gene expression associated with β-blocker therapy in patients with chronic heart failure. J Card Fail. (9):469-74.