UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Cardiac, Thoracic and Vascular Sciences

Ph.D. COURSE IN: Specialistic Medicine "G. B. Morgagni"
CURRICULUM: Cardiothoracic and Vascular Sciences
SERIES: XXX

Application of omic technologies in Arrhythmogenic Cardiomyopathy

Coordinator: Prof. Annalisa Angelini
Supervisor: Prof. Kalliopi Pilichou

Ph.D. student: Marco Cason

Index

Abstract

Background. Arrhythmogenic cardiomyopathy (AC) is an inherited myocardial disease characterized by fibro-fatty replacement of the myocardium and life-threatening arrhythmias. This genetically and phenotypically heterogeneous condition, caused mainly by mutations in desmosomal genes (JUP, DSP, PKP2, DSG2 and DSC2), exhibits reduced penetrance making challenging the diagnosis and the identification of a molecular mechanism underlying disease pathogenesis.

Aims. (1) To identify one or more altered molecular pathways in AC carriers of desmosomal mutations; (2) to evaluate the diagnostic role of JUP immune analysis in AC; (3) to study the frequency of genetic variants in the five major AC-related genes in a healthy population of Veneto region in order to evaluate their role in the pathogenesis of the disease.

Materials and methods. (1) Differential expression analysis was carried out on myocardial tissue of 7 transplanted AC patients harbouring a pathogenic mutation in desmosomal genes and 3 controls. Genetic/epigenetic interference-factors were unbiased analyzing also 3 mouse groups: 8 over-expressing NS-dsg2 mutation [TgNS], 6 over-expressing wild type dsg2 [TgWt] and 2Wt); each group was further subdivided in two age-groups (<2 weeks and >3 weeks) before and after the onset of disease. Data confirmation was obtained by quantitative-PCR.

(2) Heart specimens (HS, either autopsy or transplants) and endomyocardial biopsies (EMB) formalin-fixed from 44 AC unrelated patients and 30 non-AC matched-subjects were evaluated both by conventional immunoperoxidase analysis (IPOX) and immunofluorescence (IF) using two different JUP antibody (Ab) dilutions. Reduced JUP signal level was defined as 50% reduction in distribution and/or intensity of the immunostained areas compared to the control samples. To exclude time-dependent tissue decay, control staining for N-Cadherin was also performed.

(3) 200 unrelated young athletes (mean age 20 yrs, male/female ratio 3:1), eligible at the pre-participation clinical evaluation, underwent conventional genetic screening for major disease causative genes. Variants selection was based on the current ACMG guidelines and the absence or low frequency (minor allele frequency, MAF <0.0002) of the genetic variants in the general population.

Results. (1) 1136 and 822 differentially expressed genes (DEGs) were respectively identified in the right and left human myocardium of AC compared to controls. 204 DEGs were identified comparing TgNS<2 weeks and TgNS>3 weeks gene expression profiling. 82 DEGs were identified comparing human and murine (TgNS>3 weeks) expression-profiling including genes most associated with canonical WNT/β-catenin and TGF-β pathways. On the contrary only 29 DEGs were identified in the comparison between TgNS<2 weeks to age-matched controls (WT and TgWt <2 weeks) mostly associated with inflammatory and pro-apoptotic process, but none with WNT and TGF-β pathways.

(2) Test sensitivity (Se) of 70.6% and specificity (Sp) of 50%, with an Ab dilution 1:50.000 were found among HS, whereas with a 5-fold higher (1:250.000) Ab dilution the test Se was increased to 79.4% and the Sp decreased to 35%. Same analysis was performed on EMB samples showing different results: 40% Se; 80% Sp with 1:50.000 Ab dilution, whereas Se was 50% and Sp 70% with 1:250.000 Ab dilution. IF data were similar both with 1:1000 and 1:50.000 JUP-Ab dilution, indicating a Se of 61.8% and a Sp of 45% Sp in HS and a Se 50% and a Sp 70% in EMB samples.

(3) Genetic screening identified rare genetic desmosomal variants in 20 healthy subjects (10%) reduced to 12 (6%) after appropriate filtering.

Conclusions. Our findings demonstrated the interaction between WNT and TGF pathways at early disease stages, triggering cardiac remodelling. Specifically, we identified probably the ‘culprit molecules’ of disease onset.

Routine IPOX and IF analysis of JUP signal is associated with low Se and limited Sp to be advocated as a diagnostic test. The absolute Se range was much higher in HS than EMB samples. The same for IF, HS specimens showed higher Se and lower Sp than EMB samples. Finally, high Ab dilutions confer higher Se but reduce test Sp.

Comprehensive mutation screening and filtering in a large cohort of unrelated consecutive healthy subjects identified a lower rate of rare variants than the rates (16 and 18%) reported in literature for AC probably due to our selected cohort of healthy subjects.

1 Introduction

1.1 Definition and epidemiology

Arrhythmogenic Cardiomyopathy (AC) is an inherited disorder of the myocardium with a highly heterogeneous clinical presentation causing life-threatening arrhythmias, even though sudden cardiac death (SCD) might be the first manifestation, particularly in the young and athletes [1]. The main hallmark of the disease is the replacement of cardiomyocytes with adipose and fibrous tissues, predominantly in the right ventricle (RV) [2]. The substitution of the myocardium is progressive, starting from the epicardium or midmyocardium and subsequently becomes transmural. This process is associated with structural and functional changes involving predominantly the RV; structural manifestations mainly include ventricular dilatation, hypokinesia, and aneurysms of the ventricular wall.

Myocardial atrophy is the consequence of cell death occurring after birth, usually during childhood and is progressive with time. These changes are frequently associated with inflammatory infiltrates, which probably play a major part in triggering life-threatening arrhythmias [3]. In the early disease, changes are mostly located in the right ventricular inflow, outflow and apical regions, called the "triangle of dysplasia" [4].

Figure 1. A. Heart specimen of an AC patient showing isolated fatty replacement of the RV free wall, and translucent infundibulum. B. Azan-Mallory section showing fibro- fatty substitution of cardiac myocyte.

Clinical manifestations of AC mostly occur between the second and fourth decade of life, the mean age of diagnosis is estimated to be approximately 30 years, and males are more often affected than females, with an estimated gender ratio of 3:1 [5]. In the Veneto region, AC is

the most common cause of SCD in individuals younger than 35 years of age and in young athletes [2, 6]. The prevalence of the disease is estimated to range between 1:2000 and 1:5000 [7], however this value could be higher because of the presence of many undiagnosed and misdiagnosed cases linked to the highly variable expressivity of AC and its progressive nature.

The first historical description of the condition can be found in the book "De Motu Cordis et Aneurysmatibus" published in 1736 by Giovanni Maria Lancisi, the Pope's physician and Professor of Anatomy in Rome. He reported a family with disease recurrence in four generations. Symptoms included palpitations, heart failure, dilatation and aneurysms of the RV, and SCD [1]. However, the first comprehensive clinical description of the disease was reported in 1982, when Marcus and colleagues studied 24 adult subjects with sustained or non-sustained ventricular tachyarrhythmia and left branch block morphology. The authors coined Arrhythmogenic Right Ventricular Dysplasia as the name of the disease [8]. The hereditary nature of the disease and its polymorphic manifestation, was recognized in Padua some years later [9]. Thiene and colleagues in 1988, described the pathological substrate of the disease for the first time. They described the progressive replacement of the myocyte with fibrotic and fatty tissue mostly confined to the free wall of the RV, thus concluding that the term cardiomyopathy (= myocardial disease of unknown origin) was more appropriate to described the disease [2].

1.2 Clinical and pathological manifestations

Progressive structural changes of the working myocardium are preceded or accompanied by arrhythmias and conduction abnormalities. Ventricular tachycardia is a typical characteristic of AC patients that may lead to ventricular fibrillation and SCD [2]. The most common clinical manifestations in AC include T-wave inversion in right precordial leads, epsilon waves, the presence of widening of the QRS complex observed in resting electrocardiograms (ECG), late potentials and ventricular premature complexes which can be registered by signal average ECG and by Holter monitoring, respectively [10]. However often SCD represents the first and only manifestation of AC making its early identification difficult.

The natural history of the disease can be divided into four phases, based on clinical and pathological observations [1, 11, 12]:

- A “concealed” phase typical of very early AC. In this period, subtle structural changes and minor ventricular arrhythmia may be recognized, but the disease is frequently characterized by the absence of clinical symptoms. During this phase, subjects are mainly asymptomatic, and they may be at risk of SCD, especially during physical exercise [6].
- In the “overt electrical disorder” phase, patients manifest symptomatic ventricular arrhythmias, palpitations, syncopes and pre-syncopes. This phase is characterized by the presence of functional and morphological abnormalities of the RV and/or Left ventricle (LV), usually detectable by cardiac imaging techniques.
- The “right ventricular dysfunction” phase is characterized by further extension of disease involving the RV myocardium, leadings to impaired contractility and isolated right heart failure.
- The most advanced phase is characterized by LV involvement leading to biventricular (BV) heart failure, which is difficult to distinguish from dilated cardiomyopathy (DCM). However, LV involvement may be detected also in earlier phases of the disease [13].

1.3 Diagnostic criteria

Diagnosis of AC is based on classification of clinical findings into specific criteria, divided into major and minor according to specific association with the disease, as described in **Table 1**. AC criteria were first proposed by an International Task Force in 1994 [10] to facilitate diagnosis. Recently this criteria have been modified [14], as the original scales were highly specific but showed low sensitivity (Se) especially in evaluating asymptomatic subjects or family members affected by an early AC form [14]. The modified criteria focused on the detection of milder forms of the disease and aimed to facilitate the cascade screening of patients’ relatives based on different clinical features comprising structural, histological and arrhythmic abnormalities along with a detailed familial history of disease and SCD. However, the clinical diagnosis of AC is still complicated, as a single diagnostic test is not enough to ascertain the disease. The definitive diagnosis requires the fulfilment of 2 major, 1 major plus

2 minors, or 4 minor criteria. A borderline diagnosis requires the fulfilment of 1 major plus 1 minor, or 3 minor criteria, and possible diagnosis fulfilment of 1 major or 2 minor criteria.

Table 1. Comparison of original and revised Task Force criteria for diagnosis of AC [14].

Original task force criteria	Revised task force criteria
I. Global or regional dysfunction and structural alterations*	
Major	
	By 2D echo:
• Severe dilatation and reduction of RV ejection fraction with no (or only mild) LV impairment • Localized RV aneurysms (akinetic or dyskinetic areas with diastolic bulging) • Severe segmental dilatation of the RV	• Regional RV akinesia, dyskinesia, or aneurysm • *and* 1 of the following (end diastole): — PLAX RVOT ≥32 mm (corrected for body size [PLAX/BSA] ≥19 mm/m^2) — PSAX RVOT ≥36 mm (corrected for body size [PSAX/BSA] ≥21 mm/m^2) — *or* fractional area change ≤33%
	By MRI:
	• Regional RV akinesia or dyskinesia or dyssynchronous RV contraction • *and* 1 of the following: — Ratio of RV end-diastolic volume to BSA ≥110 mL/m^2 (male) or ≥100 mL/m^2 (female) — *or* RV ejection fraction ≤40%
	By RV angiography:
	• Regional RV akinesia, dyskinesia, or aneurysm
Minor	
	By 2D echo:
• Mild global RV dilatation and/or ejection fraction reduction with normal LV • Mild segmental dilatation of the RV • Regional RV hypokinesia	• Regional RV akinesia or dyskinesia • *and* 1 of the following (end diastole): — PLAX RVOT ≥29 to <32 mm (corrected for body size [PLAX/BSA] ≥16 to <19 mm/m^2) — PSAX RVOT ≥32 to <36 mm (corrected for body size [PSAX/BSA] ≥18 to <21 mm/m^2) — *or* fractional area change >33% to ≤40%
	By MRI:
	• Regional RV akinesia or dyskinesia or dyssynchronous RV contraction • *and* 1 of the following: — Ratio of RV end-diastolic volume to BSA ≥100 to <110 mL/m^2 (male) or ≥90 to <100 mL/m^2 (female) — *or* RV ejection fraction >40% to ≤45%
II. Tissue characterization of wall	
Major	
• Fibrofatty replacement of myocardium on endomyocardial biopsy	• Residual myocytes <60% by morphometric analysis (or <50% if estimated), with fibrous replacement of the RV free wall myocardium in ≥1 sample, with or without fatty replacement of tissue on endomyocardial biopsy
Minor	• Residual myocytes 60% to 75% by morphometric analysis (or 50% to 65% if estimated), with fibrous replacement of the RV free wall myocardium in ≥1 sample, with or without fatty replacement of tissue on endomyocardial biopsy
III. Repolarization abnormalities	
Major	• Inverted T waves in right precordial leads (V_1, V_2, and V_3) or beyond in individuals >14 years of age (in the absence of complete right bundle-branch block QRS ≥120 ms)
Minor	
• Inverted T waves in right precordial leads (V_2 and V_3) (people age >12 years, in absence of right bundle-branch block)	• Inverted T waves in leads V_1 and V_2 in individuals >14 years of age (in the absence of complete right bundle-branch block) or in V_4, V_5, or V_6 • Inverted T waves in leads V_1, V_2, V_3, and V_4 in individuals >14 years of age in the presence of complete right bundle-branch block

Continued

Table I *Continued*

Original task force criteria	Revised task force criteria
IV. Depolarization/conduction abnormalities	
Major	
• Epsilon waves or localized prolongation (>110 ms) of the QRS complex in right precordial leads (V_1 to V_3)	• Epsilon wave (reproducible low-amplitude signals between end of QRS complex to onset of the T wave) in the right precordial leads (V_1 to V_3)
Minor	
• Late potentials (SAECG)	• Late potentials by SAECG in ≥1 of 3 parameters in the absence of a QRS duration of ≥110 ms on the standard ECG • Filtered QRS duration (fQRS) ≥114 ms • Duration of terminal QRS <40 μV (low-amplitude signal duration) ≥38 ms • Root-mean-square voltage of terminal 40 ms ≤20 μV • Terminal activation duration of QRS ≥55 ms measured from the nadir of the S wave to the end of the QRS, including R′, in V_1, V_2, or V_3, in the absence of complete right bundle-branch block
V. Arrhythmias	
Major	
	• Nonsustained or sustained ventricular tachycardia of left bundle-branch morphology with superior axis (negative or indeterminate QRS in leads II, III, and aVF and positive in lead aVL)
Minor	
Minor	
• Left bundle-branch block-type ventricular tachycardia (sustained and nonsustained) (ECG, Holter, exercise) • Frequent ventricular extrasystoles (>1000 per 24 hours) (Holter)	• Nonsustained or sustained ventricular tachycardia of RV outflow configuration, left bundle-branch block morphology with inferior axis (positive QRS in leads II, III, and aVF and negative in lead aVL) or of unknown axis • > 500 ventricular extrasystoles per 24 hours (Holter)
VI. Family history	
Major	
• Familial disease confirmed at necropsy or surgery	• ARVC/D confirmed in a first-degree relative who meets current Task Force criteria • ARVC/D confirmed pathologically at autopsy or surgery in a first-degree relative • Identification of a pathogenic mutation‡ categorized as associated or probably associated with ARVC/D in the patient under evaluation
Minor	
• Family history of premature sudden death (<35 years of age) due to suspected ARVC/D • Familial history (clinical diagnosis based on present criteria)	• History of ARVC/D in a first-degree relative in whom it is not possible or practical to determine whether the family member meets current Task Force criteria • Premature sudden death (<35 years of age) due to suspected ARVC/D in a first-degree relative • ARVC/D confirmed pathologically or by current Task Force Criteria in second-degree relative

1.3.1 Clinical tests enabling AC diagnosis:

Baseline electrocardiography (12-lead ECG) findings associated with AC are abnormalities of depolarization, conduction and repolarization. ECG depolarization manifest as T-wave inversion in anterior precordial leads in the absence of "right bundle branch block" (RBBB). Conduction abnormalities are detected as prolongation of the QRS complex in right precordial leads (>110 ms in leads V1-V3) in the absence of RBBB, reflect slow conduction in the RV

free wall (major criterion). The presence of an epsilon wave, represented by a deflection between the end of the QRS complex and the beginning of the T wave, corresponds to a delayed electric potential that initiates in regions of healthy tissue surrounded by fibrofatty tissue and is indicative of intraventricular impulse conduction delay.

Signal-averaged electrocardiogram (SAECG) allows the registration of low amplitude potentials within the end of the QRS complex (late potential) that are not wide enough to be evident on the 12-lead ECG.

Holter monitoring is used to register the electric activity of the cardiovascular system for a period of normally 24 h, with particular attention to the diurnal rhythm fluctuations.

Echocardiography represents the first-line imaging approach for evaluating patients with suspected AC or for screening of family members, allowing serial examinations with the aim to assess the disease onset and progression.

Magnetic resonance imaging (MRI) is another non-invasive tissue characterization technique that allows the detection of morphological and structural abnormalities of the ventricular wall like micro aneurysms, calculation of RV and LV volumes and ejection fraction [15-17]. Endomyocardial biopsy (EMB) is used to detect myocytes in diverse stages of cell death and of fibrofatty replacement (major criterion). However, a negative biopsy is not enough to exclude AC because of the segmental nature of the disease, especially during early stages [18].

The importance of the genetic screening of the affected probands has been widely recognized and the detection of pathogenic variants now contributes as a major diagnostic criterion for the diagnosis of AC. A comprehensive workup also involves the study of the family history, with particular focus on the co-segregation of the identified genetic variant/variants and cardiac symptoms, the occurrence of SCD and unexplained heart failure [1, 12].

1.4 Molecular genetics of AC

To date, AC is considered an autosomal dominant trait showing reduced penetrance and expression variability [7, 10, 19]. However, compound and digenic heterozygosity are often described in patients with severe forms of the disease [20, 21] and, rare autosomal recessive forms of the myocardial abnormalities in some cases associated to cardio-cutaneous syndromes have been reported [22-28]. More than a decade elapsed between the recognition of familial AC [7] and the identification of the first disease-causing gene mutation [29, 30]. Linkage analysis and candidate gene approaches unmasked the genetic heterogeneity of AC with the identification of 15 genetic loci (**Table 2**). The first molecular genetics studies involved the recessive syndromic form of AC called Naxos disease. Naxos disease was first described in 1986 as a familial, autosomal recessive disease characterized by hair and skin abnormalities (woolly hair and palmoplantar keratoderma) and a form of cardiomyopathy showing AC-like features [31]. The typical woolly hair phenotype presented at birth and the erythema of the palms of the hands appeared within the first year of life, while the cardiac phenotype appeared only during adolescence or early adulthood. The original description comprised nine affected individuals from four families from the Greek island of Naxos. Linkage analysis identified a locus on the long arm of chromosome 17 (17q21.2) [32], and subsequent sequencing of the Plakoglobin (*JUP*) gene in affected subjects identified a deletions of 2 nucleotides leading to a frame shift with introduction of a premature stop codon [29]. Soon after the identification of the *JUP* deletion in Naxos syndrome, another deletion this time in the Desmoplakin (*DSP*) gene, (6p24.3), was described for another syndromic form, so called Carvajal syndrome [33]. Carvajal syndrome, an autosomal recessive disorder first described in 1998 in eighteen subjects belonging to three families from Ecuador, is characterized by epidermolytic palmoplantar keratoderma, woolly hair, and heart anomalies. The homozygous variant linked to this disorder creates a premature stop codon that truncates early DSP protein, living it without the C-domain [34, 35]. Subsequently, a variety of mutations were associated with dominant forms of AC in genes encoding desmosomal components, including: *JUP, DSP,* plakophilin-2 (*PKP2*), desmoglein-2 (*DSG2*), and desmocollin-2 (*DSC2*) [30, 36-39].

Thus, AC is now considered a disorder of the desmosome, as causative variants affecting different components of the cardiac desmosomes, have been reported in approximately 50% of AC subjects [40].

Table 2. Genetic background of AC

MIM entry	Locus	Gene ID	Gene name	Mode of transmission	Reference
#611528	17q21.2	JUP	Plakoglobin	AD	[36]
#601214				AR	[29]
#607450	6p24.3	DSP	Desmoplakin	AD	[30]
#605676				AR	[33]
#609040	12p11.21	PKP2	Plakophilin-2	AD/AR	[37]
#610193	18q12.11	DSG2	Desmoglein-2	AD/AR	[38]
#610476	18q12.1	DSC2	Desmocollin-2	AD/AR	[39]
Non-desmosomal genes					
#600996	1q43	RYR2	Cardiac Ryanodine Receptor 2	AD	[41]
#107970	14q24.3	TGF-β3	Transforming growth factor-beta 3	AD	[42]
#604400	3p25.1	TMEM43	Transmembrane Protein 43	AD	[43]
	2q35	DES	Desmin	AD	[44]
	6q22.31	PLN	Phospholamban	AD	[45]
	2q31.2	TTN	Titin	AD	[46]
	1q22	LMNA	Lamin A/C	AD	[47]
#615616	10q21.3	CTNNA3	alpha-1-catenin	AD	[48]
	7q32.1	FLNC	Filamin C	AD	[49]
	18q12.1	CDH2	Cadherin 2	AD	[50]

Abbreviations AD: autosomal dominant; AR: autosomal recessive.

Nevertheless, about 1% of AC subjects with the disease phenotype or similar, carry a pathogenic variant in genes encoding for extra-desmosomal proteins such as: Ryanodine Receptor 2 (*RYR2*), Transforming Growth Factor, Beta 3 (*TGF-β3*), Transmembrane protein 43 (*TMEM43*), Desmin (*DES*), Titin (*TTN*), Lamin A/C (*LMNA*), Phospholamban (*PLN*), α-catenin 3 (*CTNNA3*), Filamin C (*FLNC)* and Cadherin 2 (*CDH2) [41-50]*

1.4.1 Desmosomes

Desmosome complexes (schematically represented in **Figure 2**) are particularly abundant in tissues subjected to mechanical stress like the epithelium and the myocardium, where they mediate mechanical anchorage of cardiomyocytes by connecting the cytoskeleton to the cell membranes of adjacent cells. In addition to cell adhesion, a possible role in cell-cell communication, tissue differentiation, and apoptosis, has been advanced [51]. These electron-dense symmetrical structures that appear as dense membrane-associated plaques divided by a central mid-line of intercellular space are called desmoglea and intracellular plaque.

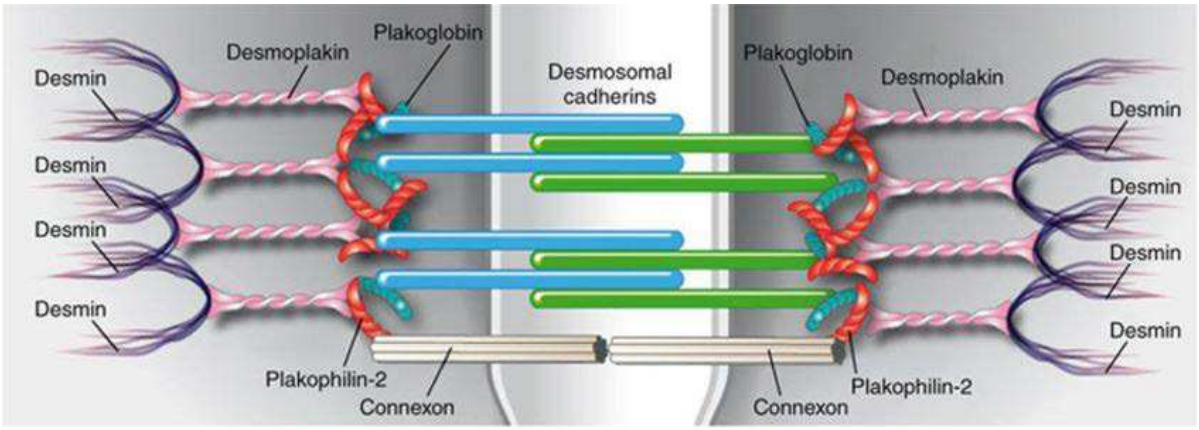

Figure 2. Schematic representation of desmosome structure (modified from Pilichou et al [52]).

The intracellular plaque is commonly described as composed of two areas: the outer dense plaque, and a dense inner plaque [53]. The desmosomal structure comprises transmembrane adhesive glycoproteins (components of the cadherin superfamily) and cytoplasmic proteins (components of the plakin and armadillo families). The outer dense plaque is where the cytoplasmic domains of the cadherins (DSG2, DSC2) attach to plakins (DSP) via armadillo proteins (PKP2 and JUP). The inner dense plaque is where plakins attach to the intermediate filaments of the cell. The connection between desmosomal components and intermediate filaments provides cellular adhesion and structural integrity.

1.4.1.1 Desmosomal genes and proteins

Desmoplakin_DSP (6p24.3) belongs to the plakin protein family which includes a large number of proteins such as plectin, envoplakin and periplakin, all mediating the connection of different junctions with the cytoskeleton. DSP is composed by an N-terminal domain which has a major role in targeting/anchoring the protein to the membrane, a central coiled-coil rod domain involved in the protein dimerization and a C-terminal domain that interacts directly with the intermediate filaments. It is expressed in all tissues containing desmosomes [54] and an alternative splicing of the DSP precursor messenger RNA (mRNA) produces two isoforms, differing in the length of the central -helical domains: DSP I comprise 2871 amino acids whereas DSP II only 2271 amino acids.

The first pathogenic nucleotide variants described in DSP families with autosomal dominant striated palmoplantar keratoderma, were heterozygous nonsense (p.Q331X) and splice site (c.939+1G>A) variants [55]. However, cardiac alterations were not studied. Soon after, with the identification of a missense mutation in exon 7 (p.S299R) involving a highly conserved amino acid and affecting a putative phosphorylation site, *DSP* was linked also to autosomal dominant AC [30]. Since then, more than 100 *DSP* pathogenic variants have been detected in 5 to 16% of AC cases, many of them showing LV involvement [38, 56, 57].

Plakophilin-2_PKP2 (12p11.21) the predominant protein isotype expressed in heart, belongs to the armadillo family of proteins and interacts directly with both, desmosomal cadherins and DSP [58]. Two alternatively spliced mRNA transcripts give origin to two protein isoforms: transcript 2b (881 amino acid) and transcript 2a (837amino acid) [59]. PKP2 comprised of an amino-terminal head domain and nine armadillos repeat motifs and is essential for heart morphogenesis and proper localization of DSP in mice [60]. In 2004, Gerull and colleagues described 26 different *PKP2* variants in 32 out of 120 AC index patients, and suggested that the lack and/or aberrant inclusion of PKP2 into the desmosomes might affect cell-cell contacts [37]. Up-to-date more than 120 pathogenic variants of *PKP2* have been associated with AC accounting for approximately 15 to 50% of reported cases [37, 38, 56, 57]. Most variants show a dominant inheritance with reduced penetrance, but recessive and compound heterozygous variants have also been identified in several patients. Moreover, large deletions involving PKP2 have also been described in subset of affected individuals [57, 61-64].

Desmoglein-2_DSG2 (18q12.1) is a desmosomal cadherin belonging to the cadherin superfamily, which is involved in calcium dependent cell-cell adhesion. In the desmosomes, there are four isoforms of desmogleins (DSG 1 to 4), showing tissue specific expression patterns [65]. DSG2 is expressed in all tissues bearing desmosomes but is the only isoform expressed in cardiac tissue [66, 67]. DSG2 has an intracellular anchoring domain interacting with DSP, a transmembrane domain, four extracellular cadherin domains each containing a calcium binding site that stabilizes the structure and function of cadherins, a small signal and a preprotein domain. *DSG2* was linked for the first time to AC in 2006 with the detection of 9 missense variants in 8 Italian families affected by the disease [38]. These variants which were located on highly conserved amino acids spread throughout the gene and for the first time was described compound heterozygosity in an AC patient [38]. Soon after another study confirmed *DSG2* association with AC by detecting 4 more variants in AC probands [68]. To date, more than 50 *DSG2* variants have been found in 2 to 20% of AC patients [38, 56, 57].

Desmocollin-2_DSC-2 (18q12.1) like DSG", is a glycoprotein, like DSG2, which belongs also to the cadherin superfamily mediating calcium dependent cell-cell adhesion. Of the three desmocollin isoforms (DSC 1, 2 and 3), DSC2 is ubiquitously expressed in tissues bearing desmosomes; however, it is the only isoform present in cardiac tissue [65]. DSC2 is composed of a signal domain, a preprotein domain followed by four highly conserved extracellular subdomains and an extracellular anchor domain at the N-terminus. In 2006 Syrris and colleagues described the presence of heterozygous frameshift variants in *DSC2* in 4 affected AC probands not carrying variants in *DSP*, *JUP*, *PKP2* or *DSG2* [39]. In the same year, a second study involving 88 index patients, identified a single heterozygous variant affecting an acceptor-splicing site in intron 5 (c.613-2A>G) of the *DSC2* in one patients; this variant resulted in skipping of the first 25 base pair (bp) of exon 6 leading to a truncated protein [69]. Less than 50 *DSC2* nucleotide variants have been reported accounting for 1 to 3% of AC cases [39, 56, 57, 69, 70].

Plakoglobin_JUP (γ-catenin), is the major cytoplasmic protein of desmosomes, belonging to the armadillo proteins. It is present both, in adherens junctions and desmosomes, where it binds to the cytoplasmic domain of cadherins acting as a linker molecule between the inner and outer parts of the desmosomal plaque. JUP comprise an N-terminal domain, a central domain containing highly conserved armadillo repeats and a C-terminal domain. JUP has only recently been linked to the dominant AC form, after the description of an in frame insertion of

a serine residue (p.S39_K40insS) in the N-terminus of the protein in a small German family [36]. Less than 20 *JUP* variants have been detected in approximately 1% of AC cases [36, 56] [57].

1.5 Animal models in AC

The genetic basis of AC is largely attributed to the presence of mutations in the desmosomal complex. In AC, mechanical stress in contracting cardiomyocytes probably results in desmosomal instability, ultimately leading to cardiac dysfunction and cell death. It is often difficult to ascribe all phenotypic aspects to defective cell adhesion. Indeed, because desmosomes also participate in intracellular signalling networks, it is frequently not clear whether the symptoms that occur in human diseases targeting desmosomes arise primarily from loss of adhesion, or from the modulation of desmosomal signalling pathways [71]. Desmosomes have also a critical role in embryonic development. Accordingly, it has been observed that homozygous null mutations in the genes encoding for the desmosomal complex are embryonically lethal in mice [60, 72-74], whereas, heterozygous null alleles for *jup* or *pkp2* cause little if any phenotypic changes without further stressors like aging or exercise [73-75].

DSP. It has been described that systemic null mutations in *dsp* in mouse embryos lead to early lethality at embryonic day 6.5 [72]. In heterozygous cardiac restricted transgenic mice expressing a point mutation or a disrupted form of dsp [73, 75], jup is mislocalized to the nucleus, where it suppresses the WNT/β-catenin canonical signalling. Furthermore, in a cardiomyocyte-specific dsp haploinsufficiency mouse model [76], which presents significant alterations in conduction–repolarization kinetics preceding morphological changes detectable on conventional cardiac imaging, reduced expression of dsp has been shown with the consequent mislocalization of the gap junctional protein connexin 43 (cx43). Hearts harbour defects in desmosomal structural integrity also in a cardiomyocyte specific *dsp*-knockout mouse (dsp-cKO) [77], which shows a BV form of human AC. Besides, connexin 40 expression is dramatically under-expressed in cardiomyocytes of the ventricular conduction system in dsp-cKO hearts. In both animal models, the WNT/β-catenin signalling pathway is impaired, but the mechanism by which the adipogenesis switch occurs is different. In the first study authors observed a suppression of the canonical WNT/β-catenin pathway due to a

mislocalization of jup to the nucleus [73, 76] while in the second study authors postulate that a dose-dependent loss – as previously proposed by Pilichou et al 2009 [78] - of dsp would lead to under-expressed of cx43, resulting in conduction abnormalities prior to the molecular dissociation of the mechanical junction complex, preceding the fibro-fatty manifestation observed in AC [77]. These models have assigned a central role to DSP in the desmosomal structure, to anchor desmosomal proteins and to provide structural integrity to the cell membrane [71].

PKP2. Heart rupture during cardiac development results in germline *pkp2* deficiency mice [60]. It has been described that pkp2 deficiency provokes reduced architectural stability of the intercalated discs (IDs) in AC9, consistent with the reduce association of the cardiac junctional plaques with dsp, jup, and desmosomal cadherins. Nevertheless, some elements of the ID remain intact; for example, several armadillo proteins such as β-catenin and jup remain in their correct locations, indicating that their binding to other plaque components is sufficient for their junctional integration. The cross-talk between the desmosome and sodium channel complex was revealed by further studies performed on *pkp2* heterozygous ($pkp2^{+/-}$) mice. These studies revealed myocyte sodium current defects, pointing to the potential contribution of sodium current dysfunction to arrhythmias. In this $pkp2^{+/-}$ murine model, mice do not seem to develop overt structural disease; $pkp2^{+/-}$ hearts show ultrastructural, but not histological or gross anatomical alterations when compared with wild-type (Wt) mice [79]. Although this could be considered a weakness of the model, nonetheless it enables the study of arrhythmogenic substrates in the absence of structural involvement, but also demonstrates that haploinsufficiency is not enough to develop AC phenotypes [71]. Zebrafish models have also been used to study PKP2-related AC. Pkp2 morphant embryos show a reduced number of IDs and increased intracellular spaces, leading to an enlarged atrium and abnormal ventricular looping [80]. As in mice, pkp2 have both structural and signalling roles in the developing zebrafish heart.

DSC2. Regarding dsc2, to date no functional knockout models are available [71]. Nevertheless, a very recent study provided evidence that the microRNA, miR-130a, targets dsc2 [81]. The authors observed over-expression of miR-130a in this model, an 80% reduction in the protein expression of dsc2 in the myocardium and right ventricular dilation. Moreover, surface ECG revealed spontaneous premature ventricular complexes, and histological examination showed fibrosis and lipid accumulation within both ventricles. Thus,

miR-130a overexpression results in a disease phenotype characteristic of AC. To determine the potential contribution of dsc2 to cardiomyopathy and intercellular adhesion pathology, researchers have developed a transgenic mouse model specifically over-expressing dsc2 in cardiac myocytes [82]. Transgenic mice developed a severe cardiac BV dysfunction which was accompanied by necrosis, calcification, and early up-regulation of inflammatory and fibrotic remodelling pathways. However, ventricular arrhythmias or SCD were not observed. From a molecular point of view, the expression of endogenous desmosomal proteins including dsg2, jup and dsp was under-expressed in mutant mice only in areas of fibrosis, indicating a local remodelling process due to the replacement of cardiomyocytes rather than a global under-expressed of endogenous desmosomal genes [82]. DSC2-related AC has also been studied in zebrafish. Morpholino knockdown of dsc2 in zebrafish embryos was shown to recapitulate the effects of a heterozygous mutation (c.631-2A>G) causing AC in humans [69]. Loss of dsc2 resulted in a reduction in the desmosomal plaque area, loss of desmosome extracellular electron-dense midlines, and associated myocardial contractility defects, indicating that DSC2 is necessary for normal myocardial structure and function [69].

JUP. Homozygous jup mutant mice die due to severe heart defects from embryonic day 10.5 onwards [74, 83]; however, some can survive until late gestation and die around birth [83], thanks to a partially compensatory mechanism between jup and β-catenin. To overcome this problem, a cardio-restricted conditional jup knockout model was created [84]. The model developed cardiomyopathy around 6–8 weeks of age, presenting progressive loss of cardiac myocytes, fibrous tissue replacement, extensive inflammatory infiltration, and cardiac dysfunction, mimicking classical AC phenotypes. Nonetheless, no spontaneous or induced cardiac arrhythmias were detected in these mice, and more than 50% of jup cardiac-deficient mice survived 17 months after recombination. Other authors, in order to avoid the embryonic lethality found in the mouse model, have turned to zebrafish models for the study of jup function; also, because zebrafish can survive a few days without a cardiovascular system [69, 85]. Overall, these models indicate that the loss of jup from the desmosome cannot be compensated by β-catenin, which leads to an altered desmosome structure. Several mouse models have successfully reproduced Naxos disease, showing defects in embryonic skin architecture and extreme Se to mechanical stress [83, 86, 87]. Unfortunately, mutant mice die during late embryogenesis or soon after birth, indicating that there might be differences in the mutant JUP expression levels between human patients and mouse models, or divergent requirements for JUP in the two species.

DSG2. Lastly, also the inactivation of the *dsg2* gene by homologous recombination in embryonic stem cells to generate a *dsg2*$^{-/-}$ knockout mouse model results in embryonic lethality [88]. This fact indicates that dsg2 has an essential function in the early embryonic development, but it hinders the characterization of the role of dsg2 in the postnatal heart.

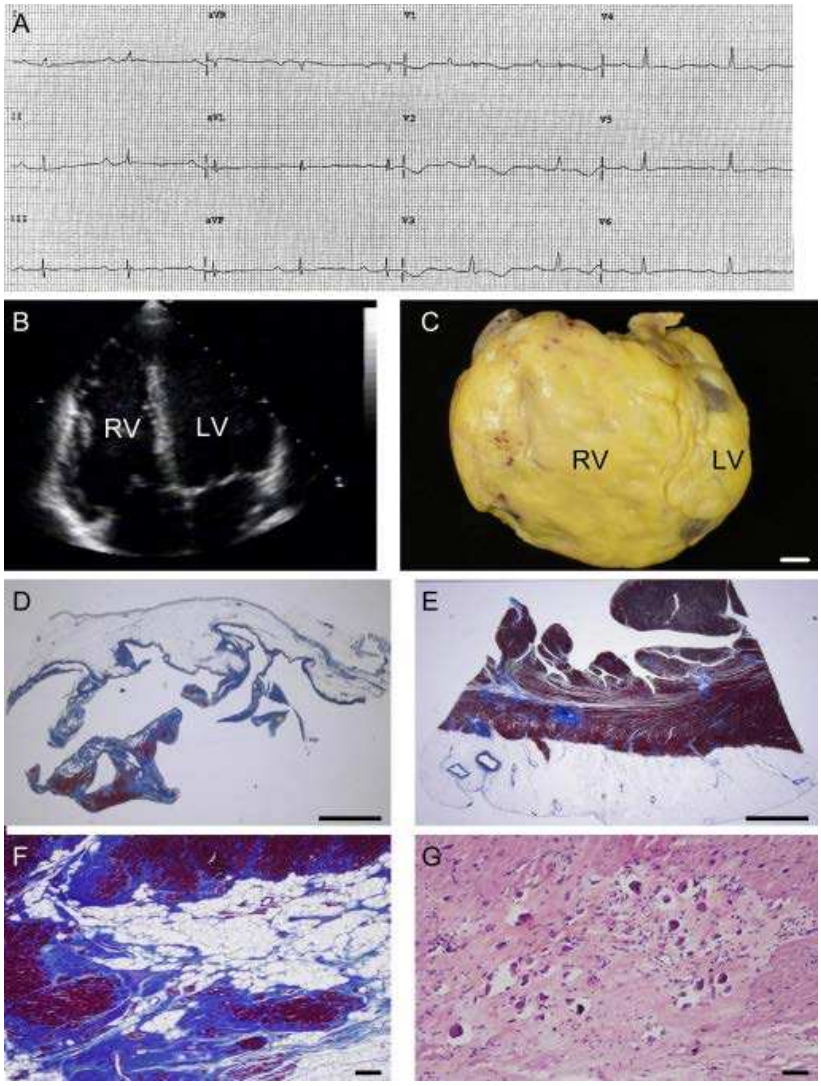

Figure 3. Characteristics of the female patient carrier of the DSG2-N266S mutation who underwent heart transplantation at the age of 65. (A) 12-lead ECG 2 mo before heart transplantation displaying sinus rhythm with PQ and QT prolongation, intraventricular conduction delay, and negative T wave V1–V4. (B) Two-dimensional echocardiography showing BV dilatation. (C) Gross view of the explanted heart specimen showing massive RV dilatation with RV outflow tract aneurysm. Bar, 1 cm. (D and E) RV and LV free wall showing fibro-fatty replacement of the myocardium with almost total disappearance of the myocardium in the RV. Bar, 5 mm. (F) Close-up of E. Bar, 30 μm. (G) At higher magnification, spotty calcification is observed in areas of replacement-type fibrosis. Bar, 50 μm [78].

Furthermore, Pilichou et al. [78] generated a transgenic mouse with cardiac overexpression of the N271S mutation of dsg2 (TgNS), equivalent to N266S found in humans [38]. Transgenic mice recapitulated the clinical features of AC, including SCD at young age, spontaneous

ventricular arrhythmias, cardiac dysfunction, and BV dilatation and aneurysms. In this model, myocyte necrosis was considered the key initiator of myocardial injury, triggering progressive myocardial damage, including an inflammatory response and massive calcification within the myocardium, followed by injury repair with fibrous tissue replacement, and myocardial atrophy. These observations were identical to the findings in the explanted heart from the patient that harbour the N266S mutation on DSG2.

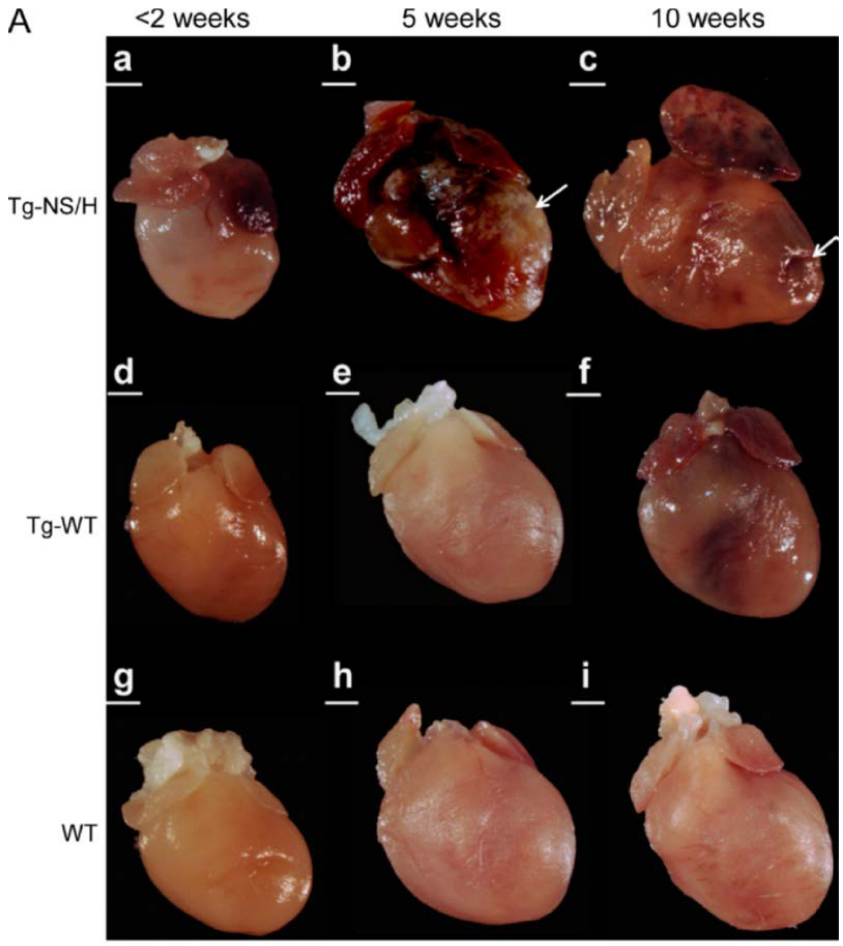

Figure 4. Gross features of explanted mouse hearts. (A) Tg-NS/H, Tg-WT, and WT hearts at age <2 (a, d, and g), 5 (b, e, and h), 10 (c, f, and i) weeks. Note the structurally normal hearts at young age, with later appearance of calcification and dilatation with aneurysms (arrows) in Tg-NS/H hearts. Bars: (a, b, and d–i) 1 mm; (c) 2 mm [78].

The model does not display fatty infiltration, which has traditionally been a characteristic finding in human AC, although no longer a specific AC diagnostic criterion. The fact that the mouse heart does not show evidence of epicardial fat was proposed as explanation for the absence of fat deposition in TgNS mice. This feature is different from the human heart, where

a physiological amount of fatty tissue can always be found in the sub epicardium. Researchers also could observe the age-related progression of the disease. Indeed, young TgNS mice (with less than two weeks of life) did not show the typical alterations of the disease observable in mice with more than 3 weeks of life.

To exclude the hypothesis that the observation made on TgNS were due to the transgene over-expression, mice over-expressing Wt dsg2 (TgWt) were studied as well and no alterations were found [78]. From further study of the N271S transgenic mouse, the same authors found an in vivo interaction between DSG2 and the NaV1.5 sodium channel protein, indicating that conduction disturbances and electrical instability could be early events of AC, evident before onset of structural changes to the myocardium [89]. This analysis also showed intercellular space widening at the level of the IDs, indicating compromised adhesion of cardiomyocytes. These data strongly support an important adhesive role for DSG2.

Several animal models have been developed also for less-common non-desmosomal AC subtypes. Notwithstanding the fact that some of the mutations have not yet been reported in human patients with AC - i.e. LAMR1 [90] and STRN [91], the models recapitulate most of the features of the human disease and could improve our understanding of the mechanisms involved, and ultimately help in the development of new therapies [71].

2 AIM of the study

This study aimed to approach different queries of AC by molecular and omic analysis in humans and animal models.

The first task was the identification of altered molecular pathway in AC subjects, carrier of desmosomal mutation. Given the molecular complexity driven by a different lifestyle, age and sex as well as the variable genetic background in human affected by AC, we carried out expression profiling analysis on a control group of transgenic mice for a specific *dsg2* mutation [78]. To this end, we divided our analysis into three phases. In the first step we studied the gene expression profile in humans comparing transcriptome of AC subjects with that of controls. In the second phase we studied the gene expression profile in the animal model comparing the gene expression profile of mice over-expressing the N271S-*dsg2* mutation (TgNS) with the transcriptome profile of mice harbouring a transgenic over-expression of Wt *dsg2* (TgWt). To subtract the artificial effect of over-expression of the *dsg2* gene, we have also taken into account the transcription profile of Wt mice. Finally, in the third step we compared the transcriptional signature of AC subjects with the profile of the experimental mouse model.

The second task was to evaluate the diagnostic role of JUP immune analysis in AC [92]. Therefore, we have study the variation of the presence of the JUP protein at the IDs among AC subject and non-AC subject. Furthermore, we have assessed the different capabilities of two analytic tools (immunoperoxidase-IPOX and immunofluorescence-IF assay) to discriminate between AC and non-AC subjects.

The third task was to study the genetic variants frequency and pathogenicity in the five major AC-related genes in the Veneto region and to evaluate their role in the pathogenesis of the disease. To this end, 500 young athletes considered healthy by clinical examinations for participation in agonistic sports activity underwent genetic screening for the five major AC associated genes (*JUP, DSP, PKP2, DSG2* and *DSC2*).

3 Materials and methods

3.1 Case Cohort

For the differential gene expression analysis, we enrolled 7 subjects who underwent heart transplantation with an anatomo-clinical diagnosis of AC and were referred to the Referential Clinical Genetic Centre of Arrhythmic Cardiomyopathies in Padua. In addition, we also included 3 patients who died for accidental causes as controls.

The study cohort for the evaluation of the JUP immune analysis as a diagnostic test for AC were chosen retrospectively, from Cardiac, Cerebral and Vascular Centre of the Veneto Region of Italy and from the Clinic and Genetic Centre for Cardiac arrhythmias of Padua. Cases diagnosed as AC were identified based on an electronical database within a 20-year period. In the study, were included 44 unrelated cases of AC (27 male; mean age 35 years ±16) and 30 controls (21 male, mean age 33 years ±12). Both groups were composed of heart specimens (HS; 34 of the AC-cases and 20 of the controls) originally obtained for histological evaluation following autopsy or heart transplantation, and by EMBs (10 of the AC-cases and 10 of the controls). All HS and EMB consisted of formalin-fixed, paraffin-embedded cardiac tissue. Among the AC HS, 27 had BV or RV involvement, while 7 cases had predominant LV involvement. As control materials, we engaged subjects with a diagnosis other than AC; including 10 HS from cardiac death "*sine materia*", 4 HS with valve prolapse, 3 HS with atherosclerosis, 1 HS with myocarditis, 1 HS with anomalous origin of coronary artery and 1 HS from a case of Idiopathic LV Hypertrophy. Among the EMB control cases we included: 6 specimens from normal heart and 4 with a diagnosis of myocarditis.

Finally, in order to study the frequency of genetic variation in the healthy population of the Veneto region, 500 athletes eligible at the pre-participation clinical evaluation were enrolled. Written informed consent was obtained from all study participants before blood sampling and genetic screening.

Experimental Animals cohort. The whole heart of a murine tissue samples from a previously described model [78] was harvested. Specifically, the study was carried out on 16 mice: 8 mice with cardiac-specific over-expression of FLAG-tagged N271S-*dsg2* (TgNS), 6 mice with cardiac-specific over-expression of FLAG-tagged WT *dsg2* (TgWT) and 2Wt mice. After euthanasia the murine heart was stored at -80 °C. For the gene expression study, only

LV myocardium was isolated and used because of its high degree of preservation in young and old mice. In order to study how gene expression varies with age, the three groups of mice (TgNS, TgWt and Wt) were subdivided into 2 age-groups: mice with less than 2 weeks of life (<2wks) and mice with an age of more than 3 weeks (>3wks).

3.2 Nucleic acid and proteins extraction

Genomic DNA (gDNA) was isolated from whole blood and frozen myocardial tissue utilizing specific onboard protocols of MagNA Pure Compact System (Roche Applied Science, Mannheim, Germany), an automated benchtop device that uses the magnetic bead technology for nucleic acids isolation. The MagNA Pure Compact performs all nucleic acids isolation steps preventing cross-contamination between samples. It automatically pipets into the sample a pre-aliquoted volume of lysis buffer containing chaotropic salts and proteinase K, nucleic acids are bound to the surface of the magnetic glass beads and after several washing steps the purified DNA is eluted in DNase-free H_20.

3.2.1 DNA extraction from blood

gDNA extraction was carried out from 200 µl of whole blood using Nucleic acid isolation kit I (Roche Diagnostics GmbH, Mannheim, Germany) on the MagNA Pure Compact System by setting the "DNA_Blood_100_400" protocol on the machine software. The typical DNA yield from 200 µl of human whole blood using the "DNA_Blood_100_400" protocol and an elution volume of 100 µl is 6 µg.

3.2.2 DNA extraction from frozen tissue

gDNA extraction was performed from approximately 10 mg of snap frozen tissue by using Nucleic acid isolation kit I (Roche Applied Science, Mannheim, Germany) with the "DNA culture cells" protocol. An initial pre-treatment step was performed to enhance cell lysis. Briefly 200 µl of MagNA Pure Compact DNA Lysis buffer (Roche Applied Science, Mannheim, Germany) was added to the tissue sample and a 2' disruption on a Tissue Lyzer System (Qiagen, Hilden, Germany) was carried out at 25 Hz. The homogenized sample was then incubated with 20 µl of Proteinase K (20 mg/mL; Roche Diagnostics GmbH, Mannheim, Germany) for 10 min at 56 °C and transferred on the MagNA Pure Compact workstation. The

typical DNA yield using the “DNA culture cells” protocol from 10 mg of snap frozen tissue and an elution volume of 50 µl is 14 µg.

3.2.3 RNA Extraction

RNA extraction from 10 mg of frozen myocardial tissue was obtained by using RNeasy mini kit (Qiagen, Hilden, Germany) according to the manufacturers’ protocol. Briefly, 1 ml of Qiazol Lysis Reagent (Qiagen, Hilden, Germany) was added to the samples and a 2’ disruption on a Tissue Lyzer System (Qiagen, Hilden, Germany) was carried out at 25 Hz. The homogenized sample was then incubated with 20 µl of Proteinase K (20 mg/mL; Roche Diagnostics GmbH, Mannheim, Germany) for 10 min at 56 °C. Subsequently, after the addiction of 200 µl of chloroform, and in order to separate RNA from cell debris, the sample was centrifuged at 12000 rcf for 15 min at 4°C. Nucleic acid was purified using Spin column of the RNeasy mini kit (Qiagen, Hilden, Germany) and DNase I was used to eliminate DNA. The typical RNA yield using RNeasy mini kit protocol from 10 mg of snap frozen tissue and an elution volume of 30 µl is 5-10 µg of RNA.

3.2.4 Proteins extraction

Protein extraction was obtained from 10 mg of frozen myocardial tissue. Briefly, 600 ml of RIPA Buffer (150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) was added to the samples and a 2’ disruption on a Tissue Lyzer System (Qiagen, Hilden, Germany) was carried out at 25 Hz. The homogenized sample was then incubated for 1h at 4°C. Subsequently, the sample was centrifuged at 15000 rcf for 15 min at 4°C, in order to separate proteins from cell debris. The typical protein yield using this protocol from 10 mg of snap frozen tissue is 4-5 µg of protein in ≈ 600 µl of elution volume.

3.3 Quantity and quality controls

Nucleic acid isolation is followed by a quantity and quality (purity, integrity) check before down-stream preparation steps. DNA/RNA quantification allows standardization of the PCR input material and is crucial in NGS RNA library preparation since it influence run quality and efficiency.

3.3.1 Spectrophotometric method

Nucleic acids quantification is performed with the help of spectrophotometers by measuring absorption of ultraviolet light at the wavelength of 260 nm and applying the Lambert-Beer law that correlates absorbance, molar extinction co-efficient and nucleic acids concentration. Absorbance measurements were carried out on a Nanovue spectrophotometer (GE Healthcare Life Sciences, UK) by directly pipetting onto the pedestals 2 µl of the DNA or RNA sample dissolved in DNase or RNase-free H_20, after an initial blank measurement. Nanovue automatically calculates the nucleic acids concentration by applying specific extinction coefficients (50 for dsDNA, 40 for RNA).

The ratios A260/280 and A260/230 are used as indicators of sample purity. The nucleic acid sample ratio A260/280 is generally used as indicator of protein contamination. Indeed the 280 nm is the absorbance wavelength of aromatic amino acid side chains and phenol groups. Pure DNA should present a 260/280 ratio between 1.8 and 2, whereas pure RNA should present a 260/280 ratio between 2 and 2.2. The A260/230 ratio is generally used as indication of organic contaminants. 230 nm is the absorbance wavelength of many organic compounds (i.e. phenol, TRIzol, and chaotropic salts present in the most common lysis buffers). In pure DNA/RNA samples the ratio A260/230 should be between 2 and 2.2.

3.3.2 Qubit Fluorometer

The quantity and quality of DNA, RNA and proteins were further assessed by using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Walthman, MA, USA). According to the manufacturer's instructions analysis were conducted with the dsDNA HS Assay kit (Thermo Fisher Scientific, Walthman, MA, USA) for DNA measurements, the RNA Assay kit (Thermo Fisher Scientific, Walthman, MA, USA) for RNA measurements and the Protein Assay kit (Thermo Fisher Scientific, Walthman, MA, USA) for proteins measurements. Briefly, 1 µl of the sample is mixed with 199 µl of working solution containing specific fluorescent dyes, composed by 199 µl of dsDNA HS buffer, RNA buffer, or Protein Buffer and 1 µl of dsDNA HS reagent, RNA reagent or protein reagent respectively. After an incubation of 2 min (15 min for proteins) the measurement is performed on the Qubit 2.0 by setting the DNA or RNA or Protein protocol. The Qubit 2.0 is a benchtop fluorometer that uses fluorescent dyes that specifically bind to either DNA, RNA or protein, therefore being able to selective quantify them. Moreover, as the DNA dye exclusively binds to double strand

DNA (dsDNA), it further allows the selective quantification of dsDNA that will be exclusively used for downstream applications, minimizing the effect of contaminant RNA and ssDNA.

3.3.3 Agilent Bioanalyzer

The Agilent Bioanalyzer 2100 (Agilent, Santa Clara, Carlsbad, USA) is a microfluidics-based platform that can perform quantification and quality control of nucleic acids by an electrophoretic separation of samples on micro channels containing fluorescent dyes. One µl of RNA sample is sufficient for analysis on 2100 Bioanalyzer with the RNA Nano Chip (Agilent, Santa Clara, Carlsbad, USA). All the assays were performed according to the manufacturer's guidelines. Briefly, RNA chips were prepared by adding 9 µl of gel matrix to the assigned well under pressure, and 9 µl of gel-dye mix to the next 2 wells. After adding 5 µl of marker to each well, 1 µl of RNA ladder and 1 µl of sample were added to the separate wells. Chips were then vortexed for 1 min at 2400 rpm on an IKA vortex mixer and were run on the Agilent 2100 Bioanalyzer. The run data were analysed by Agilent 2100 expert software version B.02.08.SI648 (SR2). The Agilent RNA Nano kit evaluates the ratio between the 18S and the 28S ribosomal subunits, and the presence of degraded small RNA fragments, in order to calculate a RNA integrity value (RIN) that expresses an estimation of the integrity of the RNA sample. RIN values range from 10 (intact) to 1 (totally degraded). An acceptable RNA quality for RNA library preparation is RIN near 8.

3.4 Standard Genetic Screening

The entire cohort of 7 AC probands and 3 controls was screened for pathogenic variants in five desmosomal genes (*DSG2, DSC2, DSP, PKP2* and *JUP*). The same analysis was conducted also on 200 out of 500 healthy athletes.

Each exon and exon-intron boundaries of the 5 genes was amplified by polymerase chain reaction (PCR) and analysed by denaturing high-performance liquid chromatography (DHPLC) and Sanger sequencing.

3.4.1 Polymerase chain reaction

PCR, a method developed by Mullis in the 80s [93], allows the exponential amplification of specific targeted DNA regions. A PCR reaction requires template DNA, deoxynucleotide triphosphates (dNTPs), oligonucleotide primers flanking the target DNA sequence, DNA polymerase enzyme, reaction buffer and magnesium. PCR is composed of different consecutive reactions:

- Template denaturation by heating DNA at 95 °C;
- Annealing temperature (Ta) usually ranges from 55 to 65°C; oligonucleotide primers may align on the target region of the single strand DNA;
- Elongation at 72°C, the DNA polymerase synthesizes new strands of DNA starting from the 3' end of annealed primers. The newly synthesized DNA of the first cycle will be the template of the next and so on, reaching a million-fold increase of the DNA quantity at the end of the reaction.

3.4.1.1 Standard PCR

Amplification reactions were carried out in a final volume of 12.5 µl containing 50 ng of template (final concentration 4 ng/µl), 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM of dNTPs, 0.8 µM of forward and reverse primer and 0.32 U of Taq DNA polymerase (Thermo Fisher Scientific, Walthman, MA, USA). For some amplicons, amplification was enhanced by the use of 1 µl of Dimethyl sulfoxide (DMSO) in the reaction mix. DMSO destabilizes the double helix structure by hydrogen binding to the major and minor grooves of DNA and reduces secondary structure formation in the DNA template, thus facilitating primer annealing. Amplification reactions were performed on Mastercycler Pro (Eppendorf, Hamburg, Germany) at specific Ta for every PCR product, optimized based on the GC content of the sequence and melting temperatures (Tm) of primers. Briefly, each sample was denatured at 95°C for 10' to allow the activation of the hot start DNA polymerase, and exposed to 40 amplification cycles of denaturation for 30'' at 95°C, annealing for 30'' at a range of Ta comprised between 55°C and 65°C and extension at 72°C for 30''; followed by a final extension step of 7 min at 72°C to enhance the amplicon elongation.

The specificity (Sp) and the amplification yield of some PCR products was enhanced by the use of touch-down PCR (TD-PCR). TD-PCR protocols set Ta above the anticipated melting temperature of a primer/template pair and then decrease the Ta of subsequent cycles stepwise.

The initial Ta of the first few cycles is higher than the expected Tm of primers, and then for the subsequent cycles the temperature progressively decreases to lower Ta to allow the correct hybridization of primer to the template. Thermal cycling conditions used for TD-PCR are as follows: an initial incubation at 95 °C for 10 min; followed by 10 cycles at 95 °C for 30'', Ta stepdown every cycle of 1 °C (from 70 °C to 60 °C); extension at 72 °C for 30''. The successive 30 cycles are performed at a Ta of 60 °C with denaturation and extension steps as above; followed by the final extension at 72 °C for 7'.

PCR products (ranging from 150-700 bp) were analysed on a 2% agarose gel.

3.4.1.2 GC rich PCR

GC rich amplicons were amplified by using AmpliTaq Gold 360 Master Mix kit (Thermo Fisher Scientific, Walthman, MA, USA) containing all premixed PCR components according to the manual. Amplification reactions were performed in a final volume of 12.5 µl by adding 7 µl of mix, 2.5 µl of GC solution provided, 2 µl of forward and reverse primer to 50 ng of template DNA.

3.4.2 Agarose gel

For standard 2% agarose gel electrophoresis, 2 g of agarose (Thermo Fisher Scientific, Walthman, MA, USA) were added to 100 ml of 1X Tris-Acetate-EDTA (TAE) buffer. The solution was heated in a microwave to dissolve agarose, gel was added with 5 µl of Nancy-520 DNA Gel Stain (Sigma-Aldrich, Saint Louis, MO, USA), cast in a sealed tray and a proper comb was inserted. Aliquots of 3 µl of PCR products and 3 µl of bromophenol blue loading dye were mixed and loaded into each gel well. In addition, 2 µl of DNA Marker VIII (Roche Applied Science, Manheim, Germany) were loaded in order to determine the fragment sizes. Electrophoretic run was performed at 100 V in 1X TAE buffer. Visualization was achieved on a LAS mini 4000 (Fujifilm, Tokyo, Japan).

3.4.3 Denaturing High Performance Liquid Chromatography (DHPLC)

DHPLC is a pre-analytical technology that can detect the presence of single nucleotide variants (SNV) and small insertions and deletions [94]. It separates heteroduplex molecules from homoduplex molecules by ion-pair reverse-phase liquid chromatography on a column containing nonporous alkylated polystyrene-divinylbenzene particles. Heteroduplexes are made by double stranded amplification products composed of two almost complementary

strands, containing a mismatch at the variant level, derived from two heterozygous alleles. The analysis is performed under partial denaturation temperature; the slow DNA renaturation allowing the formation of homoduplexes and heteroduplexes. The samples are run into the buffer flow that contains triethylammonium acetate (TEAA) and acetonitrile. The positively charged portion of TEAA interacts with the negatively charged phosphate group of the DNA that binds to the hydrophobic chromatographic column. With an increasing linear acetonitrile gradient injected into the column, the DNA binding capacity to TEAA ions decreases and the DNA fragments are released. Heteroduplexes, because of the mismatch, have a lower column binding affinity and thus a reduced retention time with respect to the homoduplexes. Eluted samples pass through the UV detector, which registers absorbance at 260 nm over time. In absence of DNA variants all the homoduplex molecules will show the same retention time and will elute as a single peak; contrarily, if a DNA variant is present, the chromatogram will show two or four peaks.

DHPLC was used to investigate the presence of variants in the 5 target genes, the analysis was performed using an ADS Biotec Limited WAVE System (ADS Biotec Limited, Glasgow, UK) with a DNASep HT Cartridge column.

After PCR reactions, amplified fragments were denatured at 96°C for 5', and then slowly cooled at room temperature to allow the formation of the heteroduplexes. Exons longer than 500 bp were splitted in more amplicons to allow DHPLC analysis. The optimal temperature for fragments analysis was calculated using Navigator Software (ADS Biotec Limited, Glasgow, UK). Runs were carried out at flow rate of 0.9 ml/min with the mutation detection application.

Samples with abnormal DHPLC profiles were purified and sequenced.

3.4.4 Purification of PCR products

Before sequencing, PCR products were purified by a treatment with Exonuclease 1 (EXO) and Shrimp Alkaline Phosphatase (SAP) enzymes, in order to eliminate primers and dNTPs residues from the previous PCR reaction, as they could interfere with the subsequent sequence reaction.

The purification reaction was carried out with the "Illustra ExoProStar 1-Step" kit (GE Healthcare Life Sciences, UK) by mixing 5 µl of PCR product and 2 µl of reaction mix,

containing SAP enzyme that dephosphorylates dNTPs and EXO that hydrolyses residual primers and nonspecific single strand oligonucleotides that could have been amplified during the PCR. The reaction conditions were: an incubation at 37 °C for 20' followed by 80°C for 15' to inactivate the enzymes.

3.4.5 Direct sequencing

The DNA sequencing method developed by Sanger in the 1970s [95] is based on the DNA chain-termination by chemically altered bases called di-deoxy nucleoside triphosphates (ddNTPs) carrying four different fluorophores. The method requires the presence of normal deoxy nucleoside triphosphates (dNTPs), and di-deoxy nucleotide triphosphates lacking a 3'-hydroxy group that, when incorporated into a newly synthesized DNA fragment, terminate the DNA strand elongation at specific bases (A, C, T, or G). This process produces DNA fragments with different sizes that can be separated by capillary electrophoresis and detected with laser-induced fluorescence.

Sequencing reactions were performed with the BigDye terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, Walthman, MA, USA) following the manufacturer's instruction. 1 µl of each purified PCR products was mixed with 2 µl of sequencing buffer at a final concentration 1X, 1 µl of BigDye Terminators solution (containing dNTPs, fluorophore-conjugated ddNTPs, and polymerase), and 10 pmol of forward or reverse primer in a total volume of 10 µl. Sequencing reaction was performed as follows: 96°C for 1', 30 cycles of 96°C for 10'', 55 °C for 5'', 60°C for 2' 30''. Fluorescent-labeled sequences were then cleaned-up to eliminate unincorporated big dye terminators, dNTPs, and salts using ethanol/EDTA precipitation. Briefly, 50 µl of EtOH 96% and 2 µl of EDTA 125 mM were added to each sample, they were incubated at room temperature for 20' and centrifuged at maximum speed for 15' at 20°C. Samples were washed with 150 µl of EtOH 96% and resuspended in 10 µl of H_20. 5 µl of sequence reaction were mixed with 15 µl of HI-Di Formamide (Thermo Fisher Scientific, Walthman, MA, USA), denatured for 5' at 95°C and submitted to sequencing analysis on an AB3500-Dx Genetic Analyzer (Thermo Fisher Scientific, Walthman, MA, USA).

3.4.6 Sequence Analysis

CHROMAS software (Technelysium) was used to view Sanger sequencing electropherograms, while comparison with genomic sequences (GRCh37, hg19) obtained

from UCSC Genome Browser (http://genome.ucsc.edu/) were done by SeqMan II (DNASTAR, Madison, WI, USA). AlaMut software version 2.10-0 (Interactive Biosoftware, Rouen, France) was used to evaluate nucleotide variants by integrating information from different sources (*in-silico* prediction algorithms, conservation, literature informations etc).

Human Genome Variation Society (http://www.hgvs.org/) recommendations (den Dunnen and Antonarakis, 2000) was used for nucleotide variant nomenclature. Whenever a putative pathogenic variant was detected, the presence of the variant was confirmed by performing a new DNA extraction, PCR and sequence reaction.

Sequence nucleotide variants found in the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) were considered known variants and their frequency in the general population was determined by the data available from the Exome Aggregation Consortium (ExAC http://exac.broadinstitute.org//), in the 2188 control chromosomes available from the 1000 Genomes Project and in the Genome aggregation database (http://gnomad.broadinstitute.org/). Intronic variants located less than 50 bp from the exon boundaries were *in-silico* analysed by at least four different algorithms (SpliceSiteFinder, MaxEntScan, NNSPLICE, GeneSplicer), in order to predict their possible effect on the splicing sites of the transcript.

In-silico analysis of missense variants is taking into account the outcome of different prediction tools, listed below:

- Polyphen-2 (http://genetics.bwh.harvard.edu/pph/). Polyphen is a tool that predicts the possible impact of an amino acid substitution on the structure and function of a protein [96]. Three outcomes are possible: probably damaging (the variant is expected with high confidence to affect protein structure), possibly damaging (the variant supposed to affect protein structure), benign (the variant is not expected to have any effect on protein structure);
- Sorting Intolerant From Tolerant (SIFT) (http://sift.jcvi.org/www/SIFT_enst_submit.html). SIFT is a tool that predicts the possible impact of an amino acid substitution on the protein function [97]. It classifies amino acids substitutions as tolerated or deleterious;
- MutationTaster. The software performs a battery of *in-silico* tests to estimate the impact of the variant on the gene product/protein. Tests are made on both, protein and

DNA level. Hence the MutationTaster is not limited to substitutions of single amino acids but can also handle synonymous alterations [98].

3.5 Next Generation Sequencing (NGS)

Microarray gene expression studies pioneered the use of genome-wide techniques in the hunt for sets of genes or gene networks implicated in complex phenotypes [99, 100]. However, microarray technology is limited by its dependence on the use of known probes, requiring a species-specific chip for most accurate results. The advent of next-generation sequencing (NGS) revolutionized gene expression studies by obviating the need for pre-existing transcript probes. RNA sequencing (RNA-seq), uses the high-throughput reads produced by NGS to represent the entire transcriptome: in other words, all transcripts produced in a tissue sample including previously uncharacterized transcribed sequences and novel isoforms. RNA-seq is used for a variety of applications, most commonly to discover lists of genes that are differentially expressed between experimental groups; i.e. samples from different tissues, samples from different treatment groups, or samples from different populations. To identify gene networks associated with inherited diseases or other genetic traits, individuals can be grouped by disease status (affected versus unaffected) or by different haplotypes at the mapped loci. In addition to gene expression differences, some RNA-seq studies may seek differences in isoform expression and allele-specific gene expression [101].

After RNA extraction, mRNA is isolated using either the poly (A) capture or ribosomal RNA (rRNA) depletion protocol. The poly (A) capture protocol results in a bias to the 3' end of transcripts, while the rRNA depletion protocol results in more variation in depth of coverage throughout the length of transcripts. After mRNA isolation, complementary DNA (cDNA) libraries are constructed. Libraries may be nonstrand-specific, or may support strand-specific RNA reads, which allow transcripts to be identified as sense or antisense. Strand-specific reads have been used in transcriptome assembly and may facilitate differentiation of reads from adjacent or overlapping genes transcribed from opposite strands [102]. Additionally, Illumina sequencers support either single- or paired-end read sequencing. Paired-end sequencing may be more expensive, but increases the percentage of reads successfully mapped to the genome. Its use is recommended for the detection of distinct isoforms; however, its increase in the mapping of unique reads may be only marginal, so the use of paired-end reads is not recommended unless maximizing unique read mapping is critical to the project [102].

3.5.1 Sequencing strategy

After the construction of the library, transcripts are typically sequenced on an NGS platform; currently, Illumina sequencers are the most common. Sequencing considerations include determining the appropriate read length and the number of lanes (ie, sequencing depth). Illumina HiSeq 2500 sequencers produce reads of 50–150 bp in length; they employ flow cells with eight lanes, and multiple samples may be run on a single lane. To differentiate reads from different samples after sequencing on the same lane, a unique barcode may be attached to each sample during library preparation. The appropriate number of lanes must be determined by taking into account the necessary depth of sequencing; for example, studies that rely on the detection of rare transcripts or polymorphisms will require greater depth [102]. Artefactual variation per lane may contribute technical variation to a study, but this can be avoided by the use of multiplexing, eg, ensuring that each lane contains a balanced number of samples from each treatment group [102]. Longer reads result in an increased percentage of mapped transcripts and improved handling of splice junctions during alignment. In the past, 454 pyrosequencing (454 Life Sciences) has been used to produce RNA-seq reads between 100 and 500 bp in length. However, this technology has proven prohibitively expensive and is currently not widely available. Emerging platforms such as PacBio (Pacific Biosciences), which provide longer read lengths, may prove popular in the future, perhaps even providing the ability to sequence entire transcripts in a single read.

3.5.2 RNA-seq bioinformatic workflow

A typical bioinformatic workflow using a reference genome and aimed at the identification of differentially expressed genes (DEGs) begins with raw reads, which are aligned to a reference genome. Gene counts are then quantified from the alignment files and used in differential gene expression analysis. This mapping process is complicated by the presence of splice junctions in the reads, originating from post processing of mRNA. Two approaches may be used to reduce the number of reads mismapped as a result of splice junctions. First, a splice-junction-aware aligner should always be used for mapping RNA reads, as alignment of this type of reads with a nonsplice-aware aligners such as Bowtie2 or BWA results in a higher percentage of mapping errors [103]. Top Hat-2 and STAR are two splice-aware aligners that are widely used for RNA read mapping; STAR operates at greater speeds than other aligners but has a correspondingly larger memory footprint [103]. Second, annotation of splice junction locations in the genome should be provided to the aligner when available. Typically,

species annotations are archived at Ensembl (ftp://ftp.ensembl.org/), RefSeq (ftp://ftp.ncbi.nlm.nih.gov), and UCSC (http://genome.ucsc.edu/). Selection of an annotation with broader gene coverage will result in an increased percentage of reads mapped to genes. In humans, the Ensembl annotation provides the broadest gene coverage and, as a result, corresponds to the highest gene mapping rates. The appropriate choice of annotation may vary from species to species, but, in general, the annotation with the broadest coverage should be selected to maximize mapping rates. Alignment of reads to the genome is further complicated by reads that map to complementary sequences at multiple locations in the genome. Ambiguous mapping may be due to conserved domains of paralogous genes, pseudogenes, and repeats. Such reads are particularly problematic in gene differential expression studies, as some gene count quantification tools discard them. A paired-end sequencing approach results in an increased percentage of uniquely mapped reads, though this improvement may be minimal [102]. Visualization of aligned reads offers the opportunity to evaluate the dataset before continuing. Such evaluation can provide opportunities to better understand problems such as coverage bias, intronic or intergenic reads, and overlapping genes. Two such visualizers are GenomeBrowse (Golden Helix) and IGV (Broad Institute). Although exonic reads make up the preponderance of RNA-seq datasets, introns, untranslated regions, and intergenic regions are often retained, albeit at lower depth. These nonexonic regions may not be artefactual but may be a result of pervasive transcription of the genome. Reads that align to intergenic regions may also represent unannotated exons or long noncoding RNA transcripts [104].

3.5.3 Differential gene expression analysis

Gene expression must be quantified in reads before differences in expression can be identified. An assessment of quantification tools shows that while results from different tools are often highly correlated, results from a subset of genes may display differences as great as 10-fold. Identification of the quantification tool with the greatest accuracy is difficult, as accurate counts may not be known for comparison. However, in a comparison of different pipelines composed of a variety of quantification tools and aligners, pipelines including the HTSeq-count quantification tool numbered among those with the best performance [103]. Outlier samples may influence differential expression results, and thus should be identified and removed prior to differential gene analysis. Alternatively, some differential expression analysis tools, such as DESeq2 [105], perform outlier detection and remove them

automatically. Another differential expression tool, edgeR, incorporates outlier detection into its estimate of gene wise dispersion when the robust = TRUE parameter is specified in the estimateDisp() method [106]. Differential gene expression analysis difficulties are: the different depths of sequencing per sample and the gene length. Additionally, these tools must contend with the small replicate numbers that are typical of RNA-seq experiments, often as low as 2–3 replicates. A comparison of 11 different methods for differential expression analysis showed that while some widely used methods have similar accuracy, the sets of DEGs found by different methods vary significantly. The tools DESeq2 and edgeR, have been found to have superior Sp and Se and are widely used. To ensure that the findings from the RNA-Seq analysis are not artefactual, it is recommended that qPCR be used to evaluate a representative set of DEGs [102, 103, 105].

Typical differential expression analyses produce lists of hundreds of DEGs, requiring further analysis to construct a high-level overview of changes between the study groups. A commercial package, Ingenuity Pathway Analysis (QIAGEN), provides a graphical user interface to assist in the discovery of pathways enriched in DEGs, generates publication-quality figures, and offers links to peer-reviewed articles about DEGs and related pathways. Panther classification system database [107] offers a freely available alternative with a web-based interface. Panther associates DEGs with other genes that have similar functions. While Panther does not provide figure generation, visualization may be accomplished by use of the freely available tool STRING (https://string-db.org/). This tool provides a graphical interface to allow the user to specify how to visualize a network of genes.

The use of qPCR, a well-established method for the evaluation of gene expression, provides technical validation of the RNA-seq procedure and data analysis used for the identification of DEGs.

3.5.4 Computing resources

Alignment processes may be computationally expensive and are best performed on a high-performance server or cluster rather than a desktop computer. An eight-core cluster with 32 GB of RAM has been recommended as the minimum hardware requirement for a typical alignment process. However, a desktop computer is generally sufficient for calling DEGs. Additionally, file sizes for reads from an individual sample aligned to a genome may be expected to reach ~10 GB in size, depending on the read depth [102].

3.5.5 The Illumina workflow

Illumina sequencers are based on sequencing-by-synthesis using a fluorescently labelled method. The general workflow includes library preparation, single molecule amplification by bridge PCR and reversible terminator sequencing-by-synthesis. The process starts with DNA library preparation, where DNA is randomly fragmented, and Illumina-specific adaptors are ligated at both ends of the fragments. The fragments are attached on an oligo-derivatized surface of a flow cell, containing oligos that bind with the adaptors. Amplification of the clusters is called bridge amplification: DNA polymerase is used to produce clusters of approximately one million copies of the original fragment. During sequencing the four labeled nucleotides are simultaneously added to the flow cell channels with the DNA polymerase, and are incorporated into the cluster fragments. The four nucleotides are labelled with base-specific fluorescents, the label contains a 3'-OH group that inhibits the fluorescence, the polymerase ligates the fluorescent labelled nucleotides in the clusters and, during this process, the 3'-OH group detaches, and the fluorescence is detected. Illumina sequencer produces reads around 100 bases.

RNA-Seq of 16 mice and 10 genotyped subjects was performed in collaboration with BMR Genomics (Padua, Italy). Mice libraries were sequenced in 3 sequencing runs, while human samples were sequence in 2 runs. Each sequencing runs included 6 samples.

RNA-Seq is composed by three main steps: library preparation, sequencing and bioinformatic data analysis of results including differential gene expression analysis.

3.5.5.1 Library preparation

Libraries were constructed according to the Illumina recommendations, and enriched for the coding regions of 20846 genes using the TruSeq stranded RNA Sample Preparation kit with Ribo Zero gold (Illumina, San Diego, CA, USA) according to the "Low sample protocol", used to process a maximum of 48 samples at one time. The protocol for the library preparation is composed by different steps (**Figure 5**).

Ribo-Zero Depletion and Fragmentation of the RNA. This process depletes rRNA from total RNA. After the rRNA is depleted, the remaining RNA is purified, fragmented, and primed for cDNA synthesis. The protocol provides as starting materials 1 μg of RNA. Subsequently, RNA for each sample is diluted with nuclease-free ultrapure water to a final

volume of 10 µl with the adding of the suitable buffer. After denaturation rRNA is removed with the use of rRNA removal beads.

Synthesize First Strand cDNA. This process reverse transcribes the cleaved RNA fragments that were primed with random hexamers into first strand cDNA using reverse transcriptase and random primers. The addition of Actinomycin D to the First Strand Synthesis Act D mix (FSA) prevents spurious DNA-dependent synthesis, while allowing RNA-dependent synthesis, thus improving strand Sp.

Synthesize Second Strand cDNA. This process removes the RNA template and synthesizes a replacement strand, incorporating dUTP in place of dTTP to generate ds cDNA. The incorporation of dUTP quenches the second strand during amplification, because the polymerase does not incorporate past this nucleotide. AMPure XP beads (Beckman Coulter) are used to separate the ds cDNA from the second strand reaction mix. At the end of this process, the library has blunt-ended cDNA.

Adenylate 3' Ends. A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation.

Ligate Adapters. This process ligates multiple indexing adapters to the ends of the ds-cDNA, preparing them for hybridization on to a flow cell.

Enrich cDNA Fragments. This process uses PCR to selectively enrich those cDNA fragments that have adapter molecules on both ends and to amplify the amount of cDNA in the library. The PCR is performed with a PCR Primer Cocktail that anneals to the ends of the adapters.

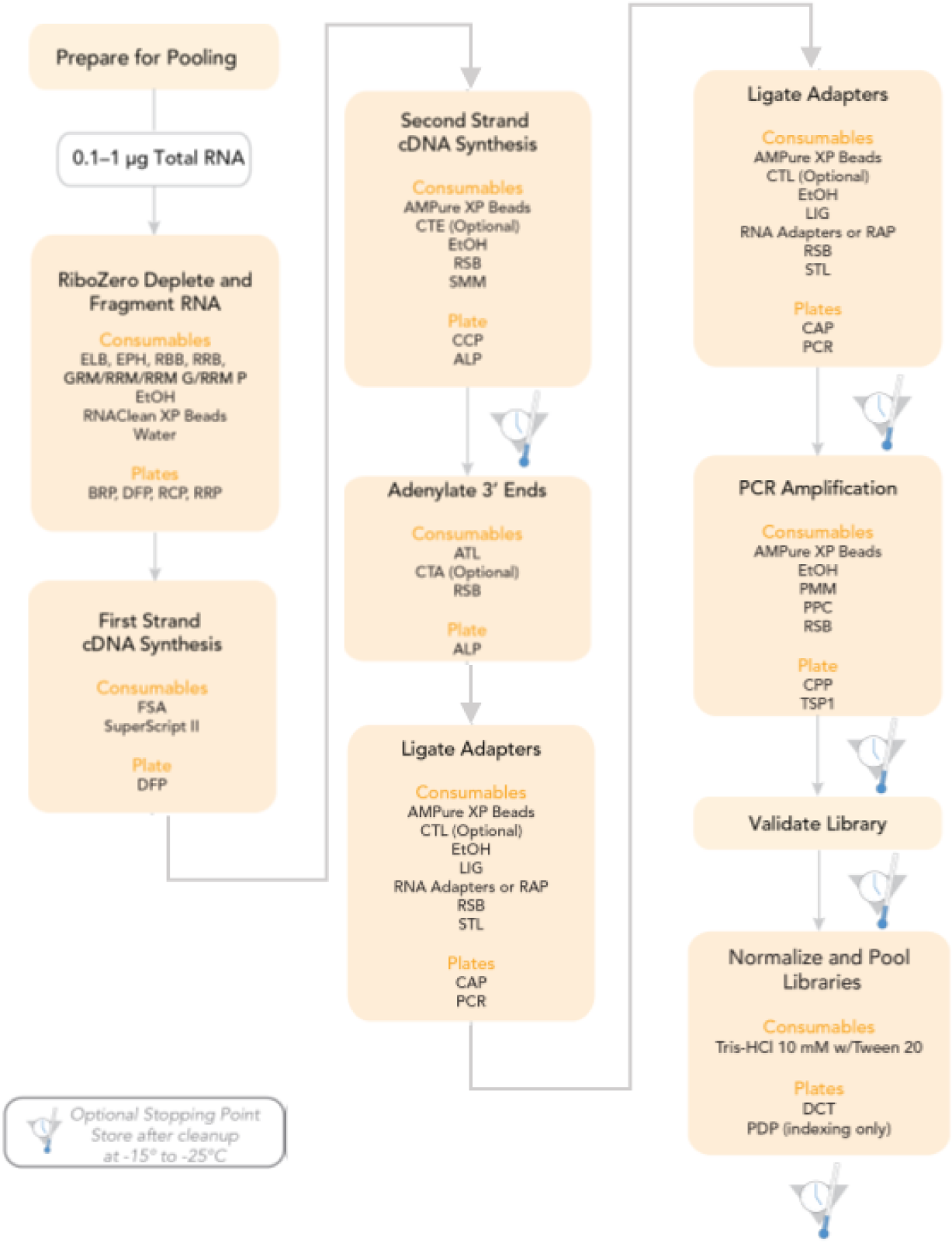

Figure 5. TruSeq RNA Sample Preparation workflow (modified from TruSeq ® RNA Sample Preparation v2 Guide).

Library validation. Quality control analysis of the sample library includes quantification of the cDNA library templates and quality control:

- **Quantification.** An accurate quantification of DNA library is essential to create optimal cluster densities across every lane of the flow cell and obtain high quality sequencing data. Successful enrichment (at least 10 µg) was verified by qPCR according to the Illumina Sequencing Library qPCR Quantification Guide. Briefly, a 100-fold dilution of the 2 mM qPCR control template was prepared. 100 µl of 0.1% Tween 20 were added to 100 µl of the diluted template and a titration curve f six 2x serial dilutions was prepared (20 pM, 16 pM, 8 pM, 6 pM, 4 pM, 2 pM, 1 pM). Then 998 µl of 0,1 % Tween 20 solution were added to 2 µl of the library template to make a 500-fold dilution (4pM), three independent dilutions of the library template were prepared to make triplicate measurements. 18 µl of SYBR reaction mix (10 µl of 2xKAPA SYBR FAST Master Mix Universal, 0.2 µl of 10 µM qPCR Primer 1.1 ad 2.1, 7.6 µl of Nuclease-free Water) were added to each well of the plate, 2 µl of the control template dilutions, the unknown library dilutions, or water were added to each well. The plate was placed on the qPCR machine with the following conditions: 95 °C for 10 min (Hot start), and 40 cycles at 95 °C for 10'', 60 °C for 30''. Finally, qPCR results are analysed, and the initial concentration of the unknown library templates is automatically calculated based on the standard curve.
- **Quality Control**. The size of the enriched fragments was checked by running 1 µl of 1:100 diluted library aliquot of the enriched library on a 2100 Bioanalyzer (Agilent, Santa Clara, Carlsbad, USA) using a DNA 1000 chip as previously described.

Normalize and Pool Libraries. Indexed DNA libraries are normalized to10 nM and then pooled together in equal volumes.

Sequencing. Cluster generation was achieved according to the Illumina Cluster Generation User Guide. In this procedure, the template is attached to the surface of an oligonucleotide-coated flow cell and amplified to produce a cluster bound to the surface of the flow cell. Cluster generation workflow includes cluster amplification, linearization, blocking, and primer hybridization. During cluster amplification the sample is hybridized on the flow cell and amplified, then the amplified sample is prepared for sequencing: one of the two adapters are cleaved off from the surface of the flow cell, 3' OH ends of the linearized dsDNA clusters

is blocked, dsDNA is denatured, and sequencing primers can hybridize. Briefly the template DNA is denatured with 0.1 N NaOH to a DNA concentration of 20 pM, and added with Hybridization Buffer. The DNA library was sequenced on a flow cell with 2x75 (for human samples) and 1x50 bp (for murine samples) run a on a HiSeq2000 (Illumina, San Diego, CA, USA).

Data Analysis. Image acquisition, and image and signal processing were performed during the run using the Illumina software package for Illumina Systems. Sequences were aligned against the reference genome (Mus musculus 10, mm10 and - human genom-19, Hg19 - Ensembl) with "STAR Aligner" software. The "Phred Quality Score" which represents the probability that the sequence contains an error [108] was used to calculate the sequence quality grade. Only sequences with Q≥20 (99% accuracy) were considered suitable for analysis. The gene expression level has been quantified using "HTSeq" software, based on the value "RPKM" (Reads Per Million mapped reads for Kilobase of exon model). This value represents the number of times a given sequence has been aligned to the reference sequence, normalized to the total number of sequences aligned for a given sample. Finally, the difference in gene expression was analysed using the "edge-R" statistical software. To determine if a gene is actually expressed in a differential manner, the statistical analysis was conducted with the "Robinson and Smith" algorithm [109] that used the "False Discovery Rate" (FDR) parameter. DEGs filtering was conducted following criteria: $Log_2FC1 \geq 1$, $Log_2FC2 \leq -1$ and p-value <0.001.

3.6 Reverse transcription

Total RNA was isolated from frozen myocardium of the patient and reverse transcribed. 200 ng of total RNA was mixed with 1 µl of 50 µM random hexamers (Thermo Fisher Scientific, Walthman, MA, USA), 1 µl of 10 mM dNTPs mix (Thermo Fisher Scientific, Walthman, MA, USA) and H_2O to a total volume of 10 µl and incubated on a thermal cycler at 65 °C for 5 min. 2 µl of 5X first strand buffer (Thermo Fisher Scientific, Walthman, MA, USA), 1 µl of 40 U/µL RNaseOUT (Thermo Fisher Scientific, Walthman, MA, USA) and 2 µl of H_2O were added to the mix, mixed by pipetting and heated at 42°C for 2'. 1 µl of SuperScript III Reverse Transcriptase (200 U/µL) (Thermo Fisher Scientific, Walthman, MA, USA) was

added and mixed by pipetting to start the retro-transcription. The reaction mix was incubated at 42°C for 50', followed by inactivation at 72°C for 15'.

3.7 Quantitative real-Time PCR (qPCR)

DEGs detected with RNA-Seq were validated by relative quantitative PCR (qPCR) performed on gDNA on a Light Cycler 480 II (Roche Applied Science, Mannheim, Germany) by SYBR Green-based quantification according to the manufacturer's protocol. The reactions were prepared in a final volume of 20 µl with 1x Master SYBR Green I (Roche Applied Science, Maanheim, Germany), 20 pmol of forward and reverse primers and 5 µl of cDNA. The amplification was carried out under the following conditions: an initial preincubation of 95°C for 10' followed by 45 cycles of denaturation at 95°C for 30'', annealing at 60°C for 20'' and extension at 72°C for 10'', and a final extension at 72°C for 10'. qPCR experiments were carried out in triplicate and included a negative control. The housekeeping gene used as reference for human cDNA samples was Hypoxanthine Phosphoribosyltransferase 1 (HPRT), while for murine cDNA samples Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) was used. Primers were developed by the Primer-Blast software (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and designed to have a Tm around 60°C, to span among two exons in order to non-amplify gDNA, and to produce amplicons approximately 150 bp long.

3.8 Western Blot

Twenty microliters from each sample were separated on NuPAGE 3%–8% Tris Acetate Gel (Thermo Fisher Scientific, Walthman, MA, USA) and blotted onto a polyvinylidene difluoride membrane for 2h at 4°C using X-cell II Blot Module (Thermo Fisher Scientific, Walthman, MA, USA). Primary antibodies were incubated for 3h at the following dilutions: α-TGFβ2 and α-MSTN 1:500, α-cMYC 1:100. α-CDH2 1:100 was used as a loading control. The reactions were visualized by detecting horseradish peroxidase-conjugated secondary antibodies goat anti-mouse 1:2500 (GE Healthcare Life Sciences, UK) or goat anti-rabbit 1:2500 (GE Healthcare Life Sciences, UK) with LAS4000 (Fujifilm/ GE Healthcare Life

Sciences, UK). Relative protein quantification was obtained by using MultiGauge software included in the ScienceLab software provided by Fujifilm.

3.9 Immunohistochemical analysis

3.9.1 Immunoperoxidase

For IPOX experiments using paraffin-embedded tissue samples, deparaffinised thin sections (4–5 mm) were subjected to heat-induced antigen retrieval performed according to standard protocols. A pre-treatment in citrate buffer using a Rapid Microwave Histoprocessor (RHS Milestone Sorisole, Italy), followed by washes in phosphate buffer and 20' incubation in phosphate-buffered saline (PBS) containing 2% milk powder and 0.2% Triton X-100 was carried out.

Antibodies used in the analysis of all patients included: a murine monoclonal Ab directed against an epitope in the aminoterminal part of JUP (Sigma, clone 15F11 ascites, P8087) [110], a murine monoclonal Ab directed against an epitope at the carboxiterminal part of JUP (Progen, clone PG5.1) [111]. For 15F11 Ab we use two different dilutions (1:50.000 and 1:250.000) according to already published data [112]. To determine the optimal dilution for the PG5.1 Ab we performed an initial immunostaining on normal myocardium using a stepwise dilution technique with dilution from 1:100 to 1:1000. The 1:250 dilution was used for PG5.1. A mouse monoclonal α-CDH2 Ab (dilution 1:1000) directed against the N-terminal half of the extracellular domain of CDH2 (clone GC-4, Sigma Aldrich, Dorset, United Kingdom) was used as comparison. Primary Ab binding was revealed by incubation with Envision (Dako) peroxidase-labelled polymer with RealDAB (Dako) as a chromogen substrate. Then, nuclear counterstaining was carried out with Harris's haematoxylin (Sigma). A Zeiss light microscope (Zeiss, West Germany), connected to an Olympus DP70 digital camera, was used for antigen distribution evaluation and image recording.

3.9.2 Immunofluorescence

Formalin-fixed sections were deparaffinized in xylol and rehydrated using four steps of decreasing concentrations of ethanol. Antigen was retrieved by heating in citrate buffer (10 mmol/l, pH 6.0) in a microwave for 11'. Subsequently, sections were permeabilized and blocked by incubating for 40' at RT in PBS containing 1% Triton X-100, 3% normal goat

serum, and 1% Bovine Serum Albumin. Next, sections were double labelled (overnight at 4°C) with primary antibodies against JUP (Sigma, clone 15F11 ascites, P8087) and CDH2 (mouse, GC-4, Sigma, dilution used: 1:1000). We use two different dilutions for 15F11 Ab (1:1.000 and 1:50.000) according to the published data [92]. Secondary labelling was performed with appropriate fluorescein isothiocyanate (FITC)-conjugated (1:250) in presence of 10% NGS. All procedures were performed at room temperature (RT), and between all subsequent steps, sections were washed for 5' with PBS. After immunolabeling, sections were mounted in Vectashield (Vector Laboratories) and examined with a Leica TCS-SC confocal microscope.

3.9.3 Statistical methods

Se, Sp and positive and negative predictive (PPV and NPV) values (all with exact 95% confidence intervals) were calculated from the blinded analysis of patient values to determine the ability of the IHC test to correctly identify AC. Only cases in which the clinical diagnosis was considered to be definite (AC or not AC) and in which an IHC analysis could be performed (normal N-cadherin signal) were included. Receiver Operating Characteristics (ROC) curve analysis was performed using MedCalc Statistical Software version 17.9.4 (MedCalc Software, Ostend, Belgium; http://www.medcalc.org; 2017) to assess the accuracy of IHC as diagnostic test.

4 CHAPTER 1: Differential gene expression analysis

4.1 Gene transcription and regulation

During the transcription process, the information contained in the DNA is transferred to an mRNA molecule. The DNA of a gene serves as a template for complementary base-pairing, and an enzyme called RNA polymerase II catalyses the formation of a pre-mRNA molecule, which is then processed to form mature mRNA. The resulting mRNA is a single-stranded copy of the gene, which next must be translated into a protein molecule [113].

During translation, which is the second major step in gene expression, the mRNA is "read" according to the genetic code, which relates the DNA sequence to the amino acid sequence of the proteins. Each group of three bases in mRNA constitutes a codon, and each codon specifies amino acids. The mRNA sequence is thus used as a template to assemble—in order—the chain of amino acids that form a protein. In eukaryotes, mature mRNA molecules must leave the nucleus and travel to the cytoplasm, where the ribosomes are located. The ribosome is composed of two subunits: the large (50S) subunit and the small (30S) subunit (S, for svedberg unit, is a measure of sedimentation velocity and, therefore, mass). Each subunit exists separately in the cytoplasm, but both join together on the mRNA molecule. The ribosomal subunits contain proteins and specialized RNA molecules called, rRNA. Another type of RNA is transfer RNA (tRNA). tRNA molecules are adaptor molecules—they have one end that can read the triplet code in the mRNA through complementary base-pairing, and another end that attaches to a specific amino acid [114, 115]. The idea that tRNA was an adaptor molecule was first proposed by Francis Crick, co-discoverer of the DNA structure, who did much of the key work in deciphering the genetic code [116]. Within the ribosome, the mRNA and aminoacyl-tRNA complexes are held together closely, which facilitates base-pairing. The rRNA catalyzes the attachment of each new amino acid to the growing chain. Interestingly, not all regions of an mRNA molecule correspond to particular amino acids. In particular, there is an area near the 5' end of the molecule that is known as the untranslated region (UTR) or leader sequence. This portion of mRNA is located between the first nucleotide that is transcribed and the start codon (AUG) of the coding region, and it does not affect the sequence of amino acids in a protein. However, the UTR sequence is important as it contains a ribosome-binding site. In bacteria, this site is known as the Shine-Dalgarno box (AGGAGG), after scientists John Shine and Lynn Dalgarno, who first characterized it. A similar site in vertebrates was characterized by Marilyn Kozak and is thus known as the

Kozak box. In bacterial mRNA, the 5' UTR is normally short; in human mRNA, the median length of the 5' UTR is about 170 nucleotides. If the leader is long, it may contain regulatory sequences, including binding sites for proteins, that can affect the stability of the mRNA or the efficiency of its translation. The translation of mRNA begins with the formation of a complex on the mRNA. First, three initiation factor proteins (known as IF1, IF2, and IF3) bind to the small subunit of the ribosome. This preinitiation complex and a methionine-carrying tRNA then bind to the mRNA, near the AUG start codon, forming the initiation complex. Although methionine (Met) is the first amino acid incorporated into any new protein, it is not always the first amino acid in mature proteins—in many proteins, methionine is removed after translation. In fact, if a large number of proteins are sequenced and compared with their known gene sequences, methionine (or formylmethionine) occurs at the N-terminus of all of them. However, not all amino acids are equally likely to occur second in the chain, and the second amino acid influences whether the initial methionine is enzymatically removed. For example, many proteins begin with methionine followed by alanine. In both prokaryotes and eukaryotes, these proteins have the methionine removed, so that alanine becomes the N-terminal amino acid. However, if the second amino acid is lysine, which is also frequently the case, methionine is not removed (at least in the sample proteins that have been studied thus far). These proteins therefore begin with methionine followed by lysine [117].

Once the initiation complex is formed on the mRNA, the large ribosomal subunit binds to this complex, which causes the release of initiation factors. The large subunit of the ribosome has three sites at which tRNA molecules can bind. The A (amino acid) site is the location at which the aminoacyl-tRNA anticodon base pairs up with the mRNA codon, ensuring that correct amino acid is added to the growing polypeptide chain. The P (polypeptide) site is the location at which the amino acid is transferred from its tRNA to the growing polypeptide chain. Finally, the E (exit) site is the location at which the "empty" tRNA sits before being released back into the cytoplasm to bind another amino acid and repeat the process. The initiator methionine tRNA is the only aminoacyl-tRNA that can bind in the P site of the ribosome, and the A site is aligned with the second mRNA codon. The ribosome is thus ready to bind the second aminoacyl-tRNA at the A site, which will be joined to the initiator methionine by the first peptide bond. The next phase in translation is known as the elongation phase (Figure 6). First, the ribosome moves along the mRNA in the 5'-to-3'direction, which requires the elongation factor G, in a process called translocation. The tRNA that corresponds

to the second codon can then bind to the A site, a step that requires elongation factors (in E. coli, these are called EF-Tu and EF-Ts), as well as guanosine triphosphate (GTP) as an energy source for the process. Upon binding of the tRNA-amino acid complex in the A site, GTP is cleaved to form guanosine diphosphate (GDP), then released along with EF-Tu to be recycled by EF-Ts for the next round. Next, peptide bonds between the now-adjacent first and second amino acids are formed through a peptidyl transferase activity. For many years, it was thought that an enzyme catalysed this step, but recent evidence indicates that the transferase activity is a catalytic function of rRNA (Pierce, 2000). After the peptide bond is formed, the ribosome shifts, or translocates, again, thus causing the tRNA to occupy the E site. The tRNA is then released to the cytoplasm to pick up another amino acid. In addition, the A site is now empty and ready to receive the tRNA for the next codon. This process is repeated until all the codons in the mRNA have been read by tRNA molecules, and the amino acids attached to the tRNAs have been linked together in the growing polypeptide chain in the appropriate order. At this point, translation must be terminated, and the nascent protein must be released from the mRNA and ribosome. The next phase in translation is known as the elongation phase (Figure 6). First, the ribosome moves along the mRNA in the 5'-to-3'direction, which requires the elongation factor G, in a process called translocation. The tRNA that corresponds to the second codon can then bind to the A site, a step that requires elongation factors (in E. coli, these are called EF-Tu and EF-Ts), as well as guanosine triphosphate (GTP) as an energy source for the process. Upon binding of the tRNA-amino acid complex in the A site, GTP is cleaved to form guanosine diphosphate (GDP), then released along with EF-Tu to be recycled by EF-Ts for the next round. Next, peptide bonds between the now-adjacent first and second amino acids are formed through a peptidyl transferase activity. For many years, it was thought that an enzyme catalysed this step, but recent evidence indicates that the transferase activity is a catalytic function of rRNA (Pierce, 2000). After the peptide bond is formed, the ribosome shifts, or translocates, again, thus causing the tRNA to occupy the E site. The tRNA is then released to the cytoplasm to pick up another amino acid. In addition, the A site is now empty and ready to receive the tRNA for the next codon. This process is repeated until all the codons in the mRNA have been read by tRNA molecules, and the amino acids attached to the tRNAs have been linked together in the growing polypeptide chain in the appropriate order. At this point, translation must be terminated, and the nascent protein must be released from the mRNA and ribosome.

4.2 Mechanisms in AC pathogenesis

Several theories have been advocated in the pathogenesis of AC: dysontogenic, degenerative, inflammatory, transdifferentiation; all have been hypothesized as causative or secondary factors for triggering and progression of disease. The dysontogenic theory considered AC as a milder form of Uhl's anomaly, which is a congenital heart defect characterized by hypoplasia of the RV myocardium at birth [118, 119]; however, it has been shown that myocyte loss in AC occurs progressively starting from childhood [120].

The degenerative theory which was postulated in 1996, considered AC as a consequence of myocyte death, either by necrosis or apoptosis, due to inherited ultra-structural defects [121] [122]. Experimental data have shown that the key initiating phenomenon in the cascade of events leading to the fibrofatty replacement of the normal myocardium is myocyte necrosis [78, 89].

In the inflammatory theory, the disease is considered the result of preceding myocarditis, since myocardial inflammation is a common feature in hearts with AC [123, 124]. More recent studies considered viral myocarditis in an already affected heart as accelerator of disease progression, rather than a primary factor in the aetiology of the disease [125].

The transdifferentiation theory suggests that cardiac myocytes undergo a metamorphosis and switch to the fate of adipocytes [126]. Even though this theory seems questionable due to the limited dedifferentiation capabilities of adult cardiomyocytes, recent studies supported the idea that adipocytes in AC derive from progenitor cells of the second heart field, which give rise to the bulbus cordis and the pulmonary infundibulum [127]. According to this hypothesis, the progenitor cells of the second heart field switch into adipocytes because of suppressed WNT/β-catenin signalling as a result of the JUP translocation to the nucleus [73]. This pathway is known to regulate adipogenesis, fibrogenesis and apoptosis, however contradictory results are coming from experimental studies regarding the activation of this pathway [84, 128].

The role of desmosomal gene mutations in the pathophysiology of myocardial injury remains elusive [129]. Pathogenic variants located in the major components of cardiac desmosomes (DSP, DSG2, DSC2, PKP2, JUP) may influence desmosome composition and function, organization of junction assemblies at the IDs level, or alter WNT/β-catenin signalling

pathway. The pathogenic effects of desmosomal genes variants may result in the incorporation of abnormal protein affecting the correct assembly and function of desmosomes (dominant negative effect), or in a quantitatively insufficient incorporation of a normal protein in the desmosomal structure (haploinsufficiency), or in the loss of essential protein-protein interactions. Pathogenic variants in the desmosomal components may cause the disruption of these structures resulting in the reduction of the mechanical contacts between cells. The impairment of the desmosomal organization would lead to loss of myocyte adhesion, and a subsequent cell death, that could be enhanced by physical activity [130]. As the cardiomyocytes have a limited regenerative capacity cell death may activate a repair mechanism implying replacement by fibrous and adipose tissue. It has been demonstrated that myocardial injury due to the death of cardiac myocytes initiates a characteristic inflammatory response. Recent studies have shown that when cells die via necrosis or apoptosis the release of their cytosolic contents into the extracellular space initiates a brisk inflammatory response that phenocopies the inflammatory response triggered by pathogenic bacteria and/or viruses [131, 132]. Mammalian cells possess germline-encoded pattern recognition receptors (PRRs) that are capable of recognizing conserved molecular motifs shared by pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) [133]. The PRRs that have been implicated in sensing both PAMPs and DAMPs include the canonical PRRs such as Toll-like receptors (TLRs), nucleotide-binding oligomerization domain–like receptors and retinoic acid–inducible gene I receptors, or atypical PRRs such as the receptor for advanced glycation end products [134]. It has been demonstrating that DAMP-induced inflammation is essential to restore tissue homeostasis by activating tissue repair mechanisms. In fact, they induce the proliferation of fibroblast and cardiac myogenesis. Although the release of DAMPs from necrotic cells is likely essential for beneficial healing responses in the heart, sustained DAMPs signalling may also lead to excessive inflammation and adverse LV remodelling [135].

Another possible effect of abnormal desmosomal components may result in the structural reorganization of the cardiac junctions known as *area composita*, which comprise both desmosome and gap junction components [136]. The remodelling of gap junction assembly may lead to impaired localization of the gap-junctional protein CX43, resulting in heterogeneous conduction of the electric impulse and an increased propensity for arrhythmias [137-139].

Desmosome components involved in cell-cell adhesion, such as JUP, are also mediators in intercellular signalling. Desmosome dysfunction leads to the translocalization of JUP from the desmosome to the nucleus, where it suppresses the canonical WNT/β-catenin signalling pathway [73]. This would cause increased expression of transcriptional regulators of adipogenesis that have been suggested to mediate the differentiation of second heart field cardiac stem cells into adipocytes [73, 127]. The majority of JUP and β-catenin is engaged at adherens junctions and/or desmosomes, linking adhesion receptors to the cytoskeleton. Although closely related, the armadillo protein β-catenin cannot replace JUP at desmosomal sites [73, 140].

Only JUP links desmosomal cadherins to intermediate filaments of the cytoskeleton via interactions with DSP, whereas JUP and β-catenin link N- or E-cadherins to the actin cytoskeleton through a-catenin [73, 140]. The cytoplasmic pool catenins are targeted by glycogen synthase kinase 3α/β (GSK-3α/β) N-terminal phosphorylation for degradation by the ubiquitin-proteasome system. Stabilization of cytoplasmic β-catenin following WNT ligand stimulation leads to its nuclear accumulation, binding with T cell factor/lymphoid enhancer factor (TCF/LEF) transcription factors, and transactivation of β-catenin/TCF/LEF target genes. In the adult heart, β-catenin/TCF/LEF-dependent gene expression has been shown to regulate both physiologic and pathological growth (i.e., hypertrophy) [141]. In terminally differentiated cardiomyocytes, β-catenin is stabilized by a WNT-independent mechanism involved in the recruitment of activated protein kinase B (PKB/AKT) to the β-catenin degradation complex [142]. In addition, JUP can also bind to TCF/LEF and either activate or suppress gene transcription, depending on the targets and cellular context [143-145]. The mechanism responsible for redistribution of JUP from the junctional pool to the intracellular and nuclear pools in AC has not been fully elucidated. It has been hypothesized that in AC JUP fails to integrate into the desmosome complex because desmosomal assembly is disturbed [73, 140] but exactly how that happens is not understood. This hypothesis raises the possibility that redistribution of JUP from IDs to the myocyte nucleus in AC might interfere with transcriptional activity regulated by β-catenin/TCF/LEF complex [128]. In fact, suppression of nuclear WNT/β-catenin signalling and other features of AC phenotype have been observed after overexpression of Wt JUP in normal cardiac myocytes [128].

Accordingly with the results obtained by Heallen [146], Chen [147] and their collaborators canonical WNT signalling is suppressed also by the activation of Hippo pathway. First

discovered in Drosophila, the Hippo signalling pathway is a conserved regulator of organ size. Central to this pathway is a kinase cascade leading from the tumour suppressor Hippo (Mst1 and Mst2 in mammals) to the oncoprotein Yki (YAP and TAZ in mammals), a transcriptional coactivator of target genes involved in cell proliferation and survival [148]. To determine whether the Hippo signalling pathway functions in the determination of mammalian heart size, investigators inactivated the single mammalian Salv ortholog (a protein forming an active complex with Mst) in the mouse heart by using a Salv conditional null allele. Chromatin immunoprecipitation (ChIP) revealed that Yap and β-catenin were in the same molecular complex and were recruited to Sox2 and Snai2 downstream regulation sites. These results uncover a nuclear mechanism for antagonistic control of cardiomyocyte growth by Hippo and canonical WNT signalling [146]. Subsequently, by studying AC mouse models, investigators found that the activation of the Hippo pathway was correlated to adipogenesis [147]. The pathogenic role of activation of the Hippo pathway in adipogenesis might be related to its suppressive effects on the canonical WNT signalling, infect YAP binds to β-catenin. They suggested that the molecular remodelling of the IDs observed in the human hearts with AC activates NF2 (an IDs component, upstream activator of the Hippo pathway), which might result in membrane localization of YAP. The molecular interactions between YAP and β-catenin could sequester β -catenin and suppress the canonical WNT and Hippo pathways, shifting the cell fate toward an adipogenic lineage [147].

Abnormalities of the Rho/Rho-kinase pathway have been implicated in the pathogenesis of several cardiovascular disorders in human and animal studies. [149]. Rho-kinase regulates a wide range of cellular functions, including actin cytoskeleton assembly, cell contractility, proliferation and differentiation, and gene expression [149]. In addition, the Rho/Rho-kinase system plays an important role in the regulation of adipogenesis. [150]. Indeed, the Rho/Rho-kinase system has been shown to negatively regulate adipogenesis through interaction with WNT signalling [151] and, in part, by controlling the expression of pro- and anti-adipogenic WNT genes. [150]. Ellawindy and colleagues [152] developed a novel mouse model in which Rho-kinase by suppression of the Rho-binding activity, was targeted to the cardiovascular system during development. They showed that these Rho-kinase-deficient mice spontaneously develop unique phenotypes fulfilling criteria of AC in humans associated with altered desmosome structure and aberrant WNT signalling. In the absence of desmosomal mutation in this mouse model, the nuclear translocation of JUP might be explained by the recently uncovered role for RhoA signalling in the assembly of the desmosome protein complex.

Godsel and collaborators [153] have demonstrated that the RhoA-mediated actin reorganization is necessary for directing dsp-enriched desmosomal precursors to sites of cell–cell contact in order to be incorporated into newly forming desmosomes. Considering the early embryonic Rho-kinase inhibition in the Rho-kinase deficient mouse model, the direction of dsp-enriched desmosomal precursors to the cell border and the subsequent incorporation into newly formed desmosomes might have been impaired. JUP, which anchors DSP to the intermediate filament network, will consequently fail to incorporate into the desmosome and is available for translocation to the nucleus. Nuclear JUP competes with β-catenin for its transcriptional factors, thereby inhibiting WNT/β-catenin signalling [73].

4.3 Next Generation Sequencing and RNA-seq

Since the conclusion of the Human Genome Project in 2000, based on Sanger technology, which required about 13 years [154, 155] and 70 million dollars [156], deep sequencing technologies were developed in order to sequence whole genomes and exomes in short time. NGS, also known as “deep sequencing” and “high-throughput sequencing” employ new sequencing apparatuses able to produce millions of small DNA sequences (reads) in a single run. In the last years, the ability of NGS to increase throughput and decrease sequencing costs, has shifted the interest from variant research in specific DNA regions to the identification of variants from genome-wide sequencing data. NGS has the potential to focus on the analysis of entire genomes (Whole Genome Sequencing-WGS), on the sole coding part of the genome (Whole Exome Sequencing, WES) or on specific target genes (Targeted Resequencing). NGS technologies may also be employed for the detection of DNA methylation sites (Methylation Sequencing, Methyl-seq), for DNA-protein interactions studies (Chromatin Immunoprecipitation Sequencing, ChIp-seq), for transcription factor profiling and gene expression quantification studies (RNA-seq) ribosomal Sequencing (Ribo-seq), small RNA profiling including microRNAs and promoter associated RNAs (sRNA-Seq), thus enabling epigenetic and transcriptome analysis [157].

RNA-seq is particularly useful in assessing the current state of a cell or tissue and the possible effects of disease states or treatment conditions on the transcriptome. RNA-seq also provides information on the differences between the transcriptome and the exome that result from RNA editing. Although microarrays have revolutionized the study of transcriptomics and

proved useful in determining gene expression profiles, RNA-seq by comparison is more sensitive, provides absolute quantity levels, is not affected by on-chip sequence biases, and gives additional information on gene expression levels and splice junction variants [158, 159].

In RNA-seq, RNA is commonly first converted to a more stable cDNA through a combination of reverse transcription and the selection process to isolate the RNA from the abundant rRNA. The input RNA quality is very important in RNAseq preparation because RNAse enzymes are ubiquitous and extremely stable and fragmentation can also occur simply when a divalent cation is present. Library preparation and sequencing of cDNA follow the same sequencing procedure as DNA-seq. However, numerous variations of RNA-seq library preparations have been developed, each with its benefits and limitations in terms of relative costs and input requirements. The main differences in these various library preparations are the methods of purifying and isolating RNA of interest (i.e. mRNA). RNA-seq libraries can be made using polyadenylated tail selection, not-so-random primers (for reverse transcription), and ribosomal depletion [160].

Isolating polyadenylated mRNA and then reverse transcribing is the conventional method of preparing an RNA-seq sample, but it favours the 3′ end of transcripts and does not work well with low-quality or degraded samples, or provide any information about noncoding RNA. Two other general methods have been commercially developed to selectively remove rRNA. For example, a method for selectively amplifying non-rRNA offered in a sequencing preparation kit (NuGEN Ovation – Leek, Netherlands) involves the use of a designed set of reverse transcription primers that contain all variants of random oligonucleotides (random primers), excluding the ones that would amplify rRNA (non-random primers). Ribosomal depletion immobilizes rRNA to remove it before reverse transcription. A commercial kit for this purpose is Ribo-Zero Gold (Illumina, San Diego, CA, USA) The use of one of these methods can recover additional RNA signals that would not be otherwise obtained via a poly-A selection because of the degraded and noncoding RNA. To annotate de novo discovery, strand-specific RNA-seq is used to determine which strand of RNA was the original template in reverse transcription. By pre-processing RNA to select for polyadenylated mRNA, or by selectively removing rRNA, a greater sequencing depth can be achieved. Depending on the experimental design, a greater sequencing depth may be required when complex genomes are being studied or when information on low abundant transcripts or splice variants is required. In general, bioinformatic analysis consists of aligning the sequence reads to a reference

genome, assembling the reads into transcripts, and detecting differences in transcript expression between or among groups. Transcriptome changes assessed by RNA-seq play a valuable role in cardiovascular medicine because transcriptome changes can identify how cardiovascular diseases change with time. Lee et al [161] used RNA-seq to study how murine hearts change during heart failure, whereas Song et al [162] used RNA-seq to decipher the transcriptome differences between physiological hypertrophy and pathological hypertrophy. In addition, Hu et al [163] studied the mRNA and microRNA transcriptome changes that occur during pressure overloading hypertrophy in mice hearts by HTS. By identifying which mRNAs and microRNAs changed during hypertrophy, and which microRNA–mRNA interaction occurred with immunoprecipitations of argonaute 2 RNA-induced silencing complex sequencing, they demonstrated that small changes in microRNA expression could lead to global mRNA changes during heart stress. Other results show that the use of RNA-seq with ChIP-seq information has provided significant advances in the study of how transcription factor binding can influence changes in gene expression. To demonstrate the potential of this approach, RNA-seq was performed on the hearts from Tbx20 knockout mice that had rapidly developed heart failure. By combining the transcriptome changes that occur from the loss of Tbx20 with the putative transcription factor binding sites of Tbx20 previously identified with ChIP-seq [164], a comprehensive analysis of how the loss of Tbx20 leads to heart failure was achieved [165]. In addition, combining ChIP-seq along with RNAseq was also used to successfully identify genes and chromatin marks involved in the progression of cardiomyocyte differentiation from human induced pluripotent stem cells [166].

4.3.1 Gene expression studies in Arrhythmogenic Cardiomyopathy

In recent years, gene expression studies have been designed to elucidate which molecular signalling pathways are altered during the pathogenesis of the disease. The importance of sarcolemmal calcium channel regulation, apoptosis and adipogenesis in the pathogenesis of AC was highlighted by a qPCR based study of candidate molecules [167]. In this study researcher found increased mRNA expression in ventricular myocardial tissue of patients with AC compared to DCM and healthy controls of phospholamban (PLN), a sarcolemmal calcium channel regulating molecule linked to desmosomes, protein 53 apoptosis effector (PERP) and carnitine-palmytoil-transferase -1beta (CPT1B), two molecules involved in apoptosis and adipogenesis [167].

Different expression patterns were identified in AC compared to DCM and non-failing (NF) heart in a microarray-based study [99] conducted on mRNA extracted from myocardial samples of AC, DCM and NF human transplanted heart. No transcriptional regulation of desmosomal protein coding genes was observed in the RV of AC patients. Conversely, in the LV researchers found a dysregulation in PKP2 (over-expressed) and DSC2 (under-expressed) mRNA expression. In the RV of AC samples, investigators also found differential expression of Pleckstrin homology-like domain, family A, 297 member-1 (PHLDA1 associated with significant changes in cell morphology, decreased cell adhesion and promoted detachment-mediated programmed cell-death in endothelial cells) compared to RV samples of NF myocardium. Another gene with a higher expression level in RV of AC patients compared with RV of NF samples was the Secreted frizzled-related protein 1 (SFRP1, it is supposed to inhibit the canonical as well as the non-canonical WNT-signalling) [168]. The Actin-binding Rho-activating protein (ABRA) was found under-expressed in RV samples of AC patients when compared with RV of NF controls. Cardiac overexpression of ABRA was shown to be linked to increased expression of SRF dependent fetal cardiac genes and sensitization of the heart to biomechanical stress [169]. Some genes specifically under-expressed in LV samples of AC subjects were previously shown to be associated with different forms of heart failure [170-173]: inflammation-associated enzyme Phospholipase A2 Type IIA 328 (PLA2G2A), the K^+ channel interacting protein (KCNIP2) with a broad physiological role in cardiac electrophysiology and the Serine protease inhibitor member 3 of clade A (SERPINA3). Intriguingly an over-expressed gene found in LV of AC patients was Spectrin beta non-erythrocytic 1 (SPTBN1) that is involved in the binding activity of Spectrin. Spectrin is a tetrameric cytoskeletal protein, which was claimed to be essential for the determination of the cell shape, and involved in resilience of membranes upon mechanical stress [174]. A recent study based on large-scaled, quantitative proteomics based on TMT-labelled LC−MS/MS and RNA-seq analysed the left and right ventricular myocardium of 4 AC and 4 DCM explanted hearts and compared them with normal hearts [175]. Researchers identified as most significant change in AC hearts compared to DCM heart, the activation of C/EBPα. In addition, they found similar changes in the Peroxisome Proliferator Activated Receptor Gamma (PPARγ) and Forkhead Box A2 (FOXA2) expression. C/EBPα and PPARγ are important regulatory factors in the fate determination of adipocytes and lipid differentiation and form a positive feedback effect to promote lipid droplet deposition [176]. Along with the significant activation of C/EBPα, the proteomics results of the study reveal that many

lipogenesis proteins (e.g., FASN, FABP4, etc.) are overexpressed in the RV of AC hearts. FASN is regulated by the SREBP1c transcription factor, with C/EBP and PPARγ binding sites, and plays an important role in coordinating fatty acid uptake and storage [177].

Our study aimed to depict the three phases of the disease (the concealed, the overt electrical disorders and the RV dysfunction phases) in order to discriminate molecular pathway that are altered in the first phase of the disease from those altered as a consequence of the disease.

4.4 Results

4.4.1 Cohort and experimental design

The study cohort, seven unrelated index cases who previously underwent cardiac transplantation with a clinical diagnosis of AC, was comprised of 2 males and 5 females. The average age at transplantation was 53±10 years (range 41-69) and all were referred for molecular genetic screening to the Cardiovascular Genetics Laboratory (University of Padua) between years 2006 and 2015. Clinical diagnosis of AC was defined according to the 2010 Task Force major and minor criteria [14], as previously described. Diagnostic criteria classified the 7 affected index cases as definite, with BV involvement. Furthermore, three non-diseased donors were included in the study as control cases. All 10 subjects underwent genetic screening for 5 genes encoding desmosomal protein. DNA was extracted from myocardial tissue and the screening was performed by DHPLC, followed by Sanger sequencing of PCR samples showing abnormal elution pattern. Direct sequencing identified putative pathological variants in *DSG2* and *PKP2* in 2 patients respectively and in 1 patient in *DSP*. In 2 patients, genetic screening identified rare variants in both *DSP* and *PKP2* (**Table 3**). No putative pathological variants were found in controls patients.

In order to prevent genetic and epigenetic factors (i.e. lifestyle, age) from interfering with our data, differential expression analysis was also performed on mouse myocardium. Pilichou and colleagues previously described used the murine model in their study from 2009 [78] a transgenic AC murine model overexpressing the N271S mutation in dsg2. This model recapitulated AC major features. Authors also observed for the first time the age-related progression of the disease. Indeed, young mice (with less than 2 weeks of life) did not show cardiac structural alteration whereas the older counterpart showed the typical signs of the

disease such as: BV dilatation, free wall thinning with aneurysm formation, and fibrous substitution of the working myocardium.

Specifically, 16 mouse samples were included in this study cohort: 8 mice over-expressing N271S-dsg2 mutation (TgNS), 6 mice over-expressing Wt dsg2 (TgWt) and 2 Wt mice. TgWt mice were included in the study to exclude alteration of gene expression due to the over-expression of dsg2. In order to identified DEGs altered during the development of the disease the three murine groups (TgNS, TgWt and Wt) were further subdivide in mice with less than 2 weeks of life (TgNS<2wks, TgWt<2wks, Wt<2wks) and mice with more than three weeks of life (TgNS>3wks, TgWt>3wks, Wt>3wks)

Table 3. Genetic characteristics of the 7 AC patients

Nr.	Sex	Age	Ventricular involvement	Gene	Exon	nt change	aa change	MAF	In silico prediction algorithms			
									Clin. Significance	SIFT	MutationTaster	PolyPhen-2
1	F	49	BV	DSP	23	c.4961T>C	L1654P	C=0.00002/2	other	Deleterious (score: 0.02)	Disease causing (p-value: 1)	Probably damaging (score:0.877)
				PKP2	10	c.2119C>T	Q707X	NA	Likely pathogenic	-	-	-
2	F	44	BV	DSP	9	c.1067C>T	T356M	T=0.000008/1	Uncertain	Deleterious (score: 0.02)	Disease causing (p-value: 1)	Probably damaging (score: 1.000)
				PKP2	1	c.39C>G	I13M	-	-	Tolerated (score: 0.29)	Disease causing (p-value: 0.998)	Probably damaging (score: 0.994)
3	M	69	BV	PKP2	7	c.1643delG	G548VfsX15	NA	Pathogenic	-	-	-
4	M	47	BV	PKP2	13	c.2552C>T	T851M	A=0.00006/7	Uncertain	Deleterious (score: 0.02)	Polymorphism (p-value: 0.602)	Probably damaging (score: 0.999)
5	F	41	BV	DSG2	12	c.1672C>T	Q558X	NA	-	-	-	-
6	F	64	BV	DSG2	7	c.797 A>G	N266S	G=0.000008/1	Pathogenic	Deleterious (score: 0)	Disease causing (p-value: 1)	Probably damaging (score: 1.000)
7	F	60	BV	DSP	21	c.2956C>T	Q986X	NA	-	-	-	-

F: female; M: male; BV: biventricular involvement; MAF: Minor Allele Frequency (from ExAC database);

.

4.4.2 RNA isolation and Sequencing

Total RNA extracted from the 20 human specimens (10 from RV and 10 from LV) and the 16 murine samples exhibited good concentrations (250ng/µl - 650 ng/µl), and a good A280/A260 ratio (1.99-2.1). Human and mouse RNA libraries were sequenced as reported and for each sample, 20 to 60 million sequences were obtained (quality threshold used ≥Q20). One murine sample (sample 13 of **Table 5**) was not considered suitable for subsequent analyses due to the low number of reads obtained and was discarded.

Table 4. Alignment metrics of human samples

N°	Sample ID	Read Lenght	Number of reads	% Total Aligned	% Abundant	% Unaligned	%Stranded
1	Sample_1 VDx	75/75	48,766,618	95.48%	5.76%	4.52%	98.72%
2	Sample_1 VSn	75/75	55,880,909	96.46%	4.11%	3.54%	99.41%
3	Sample_2 VDx	75/75	45,187,753	96.29%	2.99%	3.71%	99.45%
4	Sample_2 VSn	75/75	42,058,546	96.11%	4.04%	3.89%	99.37%
5	Sample_3 VDx	76/76	51,473,773	96.20%	6.05%	3.80%	98.74%
6	Sample_3 VSn	76/76	51,324,498	96.06%	5.66%	3.94%	99.29%
7	Sample_4 VDx	75/75	52,429,026	96.29%	8.50%	3.71%	99.06%
8	Sample_4 VSn	75/75	49,110,963	96.28%	3.15%	3.72%	99.35%
9	Sample_5 VDx	75/75	46,110,072	96.27%	6.17%	3.73%	99.32%
10	Sample_5 VSn	75/75	51,237,810	96.32%	5.65%	3.68%	99.37%
11	Sample_6 VDx	75/75	48,623,227	95.74%	7.99%	4.26%	99.18%
12	Sample_6 VSn	75/75	48,996,962	96.48%	6.18%	3.52%	99.40%
13	Sample_7 VDx	76/76	54,927,100	96.21%	4.64%	3.79%	99.25%
14	Sample_7 VSn	76/76	60,916,500	95.91%	7.30%	4.09%	99.22%
15	Cn1Vdx	76/76	55,835,121	96.13%	5.50%	3.87%	99.47%
16	Cn1Vsx	76/76	49,690,231	95.92%	4.50%	4.08%	99.39%
17	Cn2Vdx	76/76	62,634,983	96.26%	3.92%	3.74%	99.40%
18	Cn2Vsx	76/76	53,177,852	95.70%	6.07%	4.30%	99.35%
19	Cn3Vdx	76/76	50,645,418	96.10%	7.48%	3.90%	99.44%
20	Cn3Vsx	76/76	53,670,835	95.84%	7.93%	4.16%	99.35%

Read length: length of reads (bases); Number of reads: total number of reads passing filter for the sample; % Total Aligned: percentage of reads passing filter that aligned to the reference, including abundant reads; % Unaligned: percentage of reads that do not align to the reference; % Abundant: percentage of reads that align to abundant transcripts, such as mitochondrial and ribosomal sequences.

Table 5. Alignment metrics of murine samples

N°	Sample ID	Read Lenght	Number of reads	% Total Aligned	% Abundant	% Unaligned	%Stranded
1	WT<2wks	1x50	59,856,132	98.66%	19.02%	1.34%	99.06%
2	WT>3wks	1x50	55,496,714	98.15%	25.33%	1.85%	99.04%
3	TgWT 2wks-2	1x50	54,736,609	98.36%	15.91%	1.64%	99.03%
4	TgWT 2wks-3	1x50	58,950,942	98.28%	17.62%	1.72%	98.56%
5	TgWT 2 wks-1	1x50	65,049,955	98.37%	15.88%	1.63%	98.95%
6	TgWT>3wks-3	1x50	51,472,250	98.28%	33.36%	1.72%	98.99%
7	TgWT>3wks-2	1x50	53,952,122	98.25%	32.10%	1.75%	98.71%
8	TgWT>3wks-1	1x50	55,497,214	97.98%	34.80%	2.02%	98.98%
9	TgNS<2wks-1	1x50	55,618,833	98.19%	29.39%	1.81%	99.13%
10	TgNS<2wks-2	1x50	61,107,268	98.40%	25.97%	1.60%	99.18%
11	TgNS<2wks-3	1x50	61,642,747	98.28%	19.45%	1.72%	98.93%
12	TgNS<2wks-4	1x50	62,392,164	98.27%	18.18%	1.73%	99.08%
13	TgNS<3wks-1	1x50	21,766,178	98.63%	29.46%	1.37%	99.60%
14	TgNS<3wks-2	1x50	52,366,748	97.74%	17.12%	2.26%	99.12%
15	TgNS<3wks-3	1x50	55,625,383	98.20%	18.09%	1.80%	99.21%
16	TgNS<3wks-4	1x50	60,058,789	98.40%	14.00%	1.60%	99.40%

Read length: length of reads (bases); Number of reads: total number of reads passing filter for the sample; % Total Aligned: percentage of reads passing filter that aligned to the reference, including abundant reads; % Unaligned: percentage of reads that do not align to the reference; % Abundant: percentage of reads that align to abundant transcripts, such as mitochondrial and ribosomal sequences.

4.4.3 Human transcriptome

Transcripts from RV and LV tissues derived from age-matched AC patients and control cases were sequenced in order to identify significantly DEGs. RV AC transcriptomes showed more than 1100 DEGs when compared to RV controls (224 over-expressed genes and 912 under-expressed genes). Prominent increment of profibrotic cytokines expression such as transforming growth factor- β2 (*TGF-β2*) and Myostatin (*MSTN*), cadherin and protocadherin such as Cadherin-8 (*CDH8*), Protocadherin beta-1 (*PCDHB1*) Protocadherin gamma-A1 (*PCDHGA1*) as well as Actin, alpha skeletal muscle (*ACTA1*) was observed. The group of over-expressed genes included those encoding molecules regulated by calcium signalling, including calcium/calmodulin-dependent protein kinase type 1G (*CAMK1G*), calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase 1A (*PDE1A*), neuron-specific calcium-binding protein hippocalcin (*HPCA*), and calcium/calmodulin-dependent protein kinase type 1G (*CAMK1G*). The under-expressed genes included signalling molecules involved in WNT signalling, such as Secreted frizzled-related protein 2 (*SFRP2*) and the Myc proto-oncogene protein (*C-MYC*). Interestingly, genes correlated with oxidative stress such as Matrix metalloproteinase 23 (*MMP23A*) were under-expressed.

On the other hand, LV AC transcriptome showed more than 800 DEGs compared to the control samples (388 over-expressed genes and 434 under-expressed genes). Prominent increases were observed in the expression of profibrotic cytokines, such as *TGF-β2*, *MSTN*, and Growth/differentiation factor 5 (*GDF5*). Cadherin and protocadherin were also over-expressed.

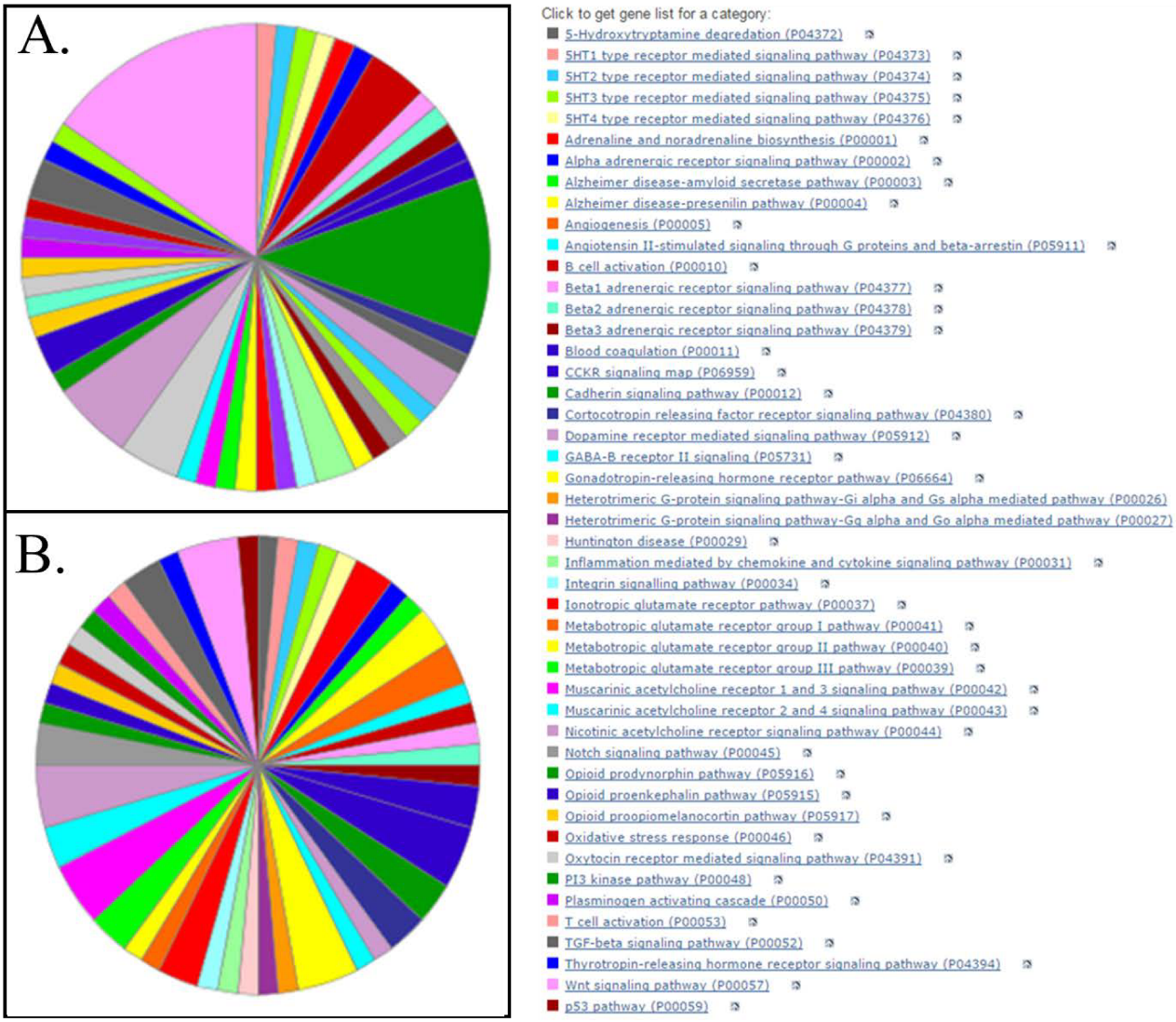

Figure 6. Pathway analysis of the LV; A. over-expressed genes, B. under-expressed genes

Among under-expressed genes in the LV, genes coding for protein involved in the regulation of WNT signalling pathway are listed, such as: *SFRP2, SFRP5, WNT9B* and *C-MYC*. Intriguingly, differently than seen in the RV, genes involved in oxidative stress response were down-expressed: *CHRNA7, NOTCH4, DUPD1* and *GRIN2A*.

We also performed a comparison between AC RV and AC LV transcriptomes. The comparison showed a total of 1342 DEGs (220 over-expressed and 1122 under-expressed genes). Pathway analysis with "Panther classification system" [107] of over-expressed genes in the RV highlights the presence of genes involved in the metabotropic glutamate receptor pathway such as: Glutamate receptor 3 (GRIA3), excitatory amino acid transporter 2 (SLC1A2), Vesicle-associated membrane protein 7 (VAMP7) and metabotropic glutamate receptor 1 and 8 (GRM1 and GRM8). The role of glutamate pathway has been widely described in the brain where it plays a significant protective role in a broad range of diseases in the central nervous system. A recent work of Vincent and collaborators provides the first demonstration that GRM1 activation at the onset of reperfusion is involved in inducing cardio protection and GRM1 activation might represent a putative strategy to prevent ischemia reperfusion injury [178].

Interestingly, WNT and TGF-β pathways genes were under-expressed in the RV in comparison to LV. Indeed, in the RV we found genes such as C-MYC, SFRP2, ACVR1C, TGF-β2 and BMP3 under-expressed, where as in LV were over-expressed.

4.4.4 Murine transcriptome

In order to find DEGs altered due to the expression of mutated dsg2, a comparison between DEGs found in Wt, TgWt and TgNS mice was performed (schematically represented in Figure 7)

Comparison 1. We analysed the differences in gene expression between Wt <2wks and TgWt <2wks. The comparison did not reveal differences in gene expression, indicating that the over-expression of dsg2 Wt does not result in significant alterations in gene expression in young mice. We then analysed the differences in gene expression between TgWt <2wks and TgNS <2wks mice identifying 10 over-expressed and 19 under-expressed genes in TgNS mice. We therefore postulated that some of these differences occur due to the *dsg2* mutation.

Comparison 2. We analysed the differences in gene expression between Wt >3wks and TgWt >3wks. The comparison did not reveal differences in gene expression, indicating that the over-expression of dsg2 Wt does not result in significant alterations in gene expression in older mice. However, the comparison between TgWt> 3wks and TgNS> 3wks mice identifies

653 DEGs (460 over-expressed 193 under-expressed genes). Confirming what previously hypothesized in comparison 1.

Comparison 3. To exclude DEGs caused by aging we compared TgWt <2wks and TgWT> 3wks identifying 77 over-expressed genes and 64 under -expressed genes. These were subtracted from the 104 over-expressed genes and 180 under-expressed genes identified from the comparison of TgNS <2wks and TgNS> 3wks.

Summarizing, from the comparison performed we identified 29 DEGs in TgNS<2wks mice (before disease onset) and 143 DEGs comparing TgNS>3wks mice (after the disease onset) to their young littermates. Further, 653 DEGs were found comparing TgWt>3wks and TgNS>3wks, including 143 DEGs found comparing TgNS<2wks and TgNS>3wks.

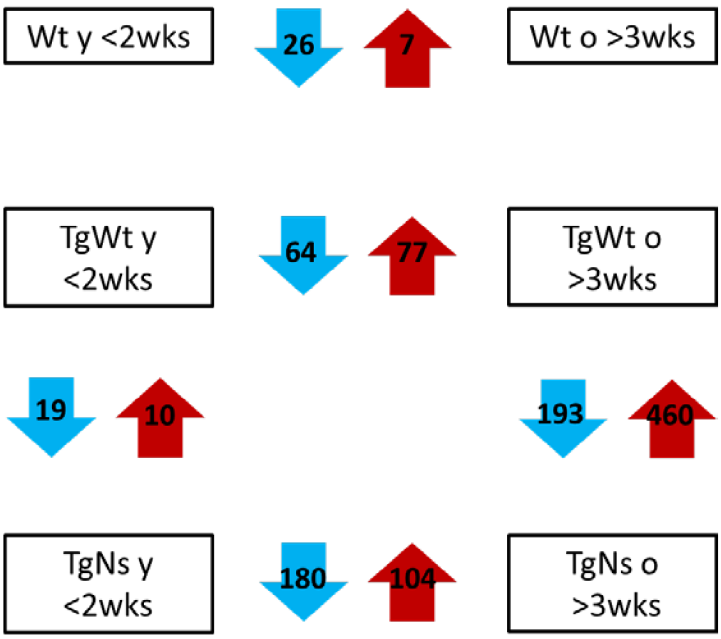

Figure 7 Comparison performed among different mice groups

In TgNS <2wks structural changes in myocardial tissue are not yet visible and the heart appears normal. Differential expression analysis allowed us to identify 10 genes whose expression appears to be increased and 19 genes whose expression level is reduced compared to Wt and TgWt mice (**Table 6**). The molecular mechanisms linked to these genes include

inflammatory processes (Lgals3 and Mapk4), alterations in the expression of proteins involved in the activation of pro-apoptotic pathways (Capn6), genes coding for proteins involved in transport of ion channels to IDs (Unc79). Not much is known about their function and mechanism of action. The over-expression of Lrg1 in TgNs <2wks, which is a modulator of Tgf-β1 activity, was described previously as trigger for the onset of fibrosis in animal models with Dilated Cardiomyopathy [179]. Molecular pathway analysis carried out with "Panther classification system" linked Capn6 and Mapk4 to pathway related to oxidative stress. Furthermore, the Pathway unification database (PathCards) classified Capn6 in the Degradation of the extracellular matrix and Arrhythmogenic Right Ventricular Cardiomyopathy pathways (Relevance score 0,251 and 0.358respectively); on the other hand, Mapk4 was classified as involved in the VEGF Signalling Pathway (Relevance score 0.265).

Table 6. 29 DEGs found in TgNS<2wks in comparison to TgWt<2wks

Over-expressed genes			Under-expressed genes					
Gene ID	**FoldChange**	**pvalue**	**Gene ID**	**FoldChange**	**pvalue**	**Gene ID**	**FoldChange**	**pvalue**
Mgl2	20,38	$3.6x10^{-04}$	Ms4a7	-20,32	$1.7x10^{-04}$	Lgals3	-22,41	$1x10^{-04}$
Lrrn1	20,78	$2x10^{-04}$	Tdrd9	-20,49	$8.x10^{-05}$	Syp	-23,05	$9.3x10^{-06}$
4930481A15Rik	20,95	$6x10^{-04}$	Mapk4	-20,51	$2.7x10^{-06}$	Tgfa	-23,27	$2.1x10^{-05}$
Lrg1	21,15	$4.3x10^{-04}$	Fbxl16	-21,02	$4.3x10^{-05}$	C2	-23,27	$7.4x10^{-05}$
Lrat	21,25	$4.7x10^{-04}$	Scg3	-21,12	$5x10^{-04}$	Cybrd1	-23,44	$6.3x10^{-05}$
Hif3a	21,52	$4.7x10^{-05}$	Crispld1	-21,16	$6.6x10^{-06}$	Doc2b	-23,63	$4.4x10^{-05}$
Slc22a3	21,71	$4.5x10^{-05}$	Erc2	-21,24	$1x10^{-04}$	Ptprv	-24,22	$1x10^{-05}$
Cbr2	21,97	$6.1x10^{-06}$	Unc79	-21,71	$3x10^{-04}$	Spp1	-24,54	$3.6x10^{-05}$
Cd163	24,69	$1.2x10^{-05}$	Pex5l	-22,02	$2x10^{-04}$	Capn6	-24,67	$6.4x10^{-06}$
Kcnv2	25,44	$1.8x10^{-05}$	Tbx15	-22,13	$2.6x10^{-04}$			

To identify the molecular pathways represented by the 143 DEGs found in comparison 3, we performed gene ontological analysis based on "Panther Classification System". Prominent increment in expression of profibrotic cytokines such as Tgf-β2 and Mstn and genes involved in inflammatory processes and T cell activation were observed. Interestingly, over-expression of Mapk4 was found TgNS>3weeks mice. Under-expressed DEGs were annotated mainly as WNT signalling pathway, and genes listed under this Gene Ontological (GO) term are: Sfrp2, Sfrp5 Acvr1c, Wnt9b, Cdh1 and Wnt4. We also observed the under-expression of genes involved in oxidative stress processes such as Actg2, Grin2c, Grik5 and Fos.

4.4.5 Comparison among human and murine transcriptome

Finally, DEGs comparison was performed between murine model and human samples in order to identify common DEGs. Among the 653 DEGs found comparing TgWt>3weeks and TgNS>3weeks we have found 146 DEGs in common with human. When the comparison was performed taking in account only the 143 DEGs found in the comparison among TgNS<2weeks and TgNS>3weeks, the number of DEGs in common with human decrease further to 82 (5 over-expressed genes and 77 under-expressed). Of these 82 DEGs, 45 genes were classified in 7 molecular pathways: Oxidative stress (27%), WNT pathway (16%), TGF-β pathway(13%), inflammation (13%), apoptosis (11%), G-protein signalling pathway (11%) and Metabotropic glutamate receptor signalling pathway (9%).

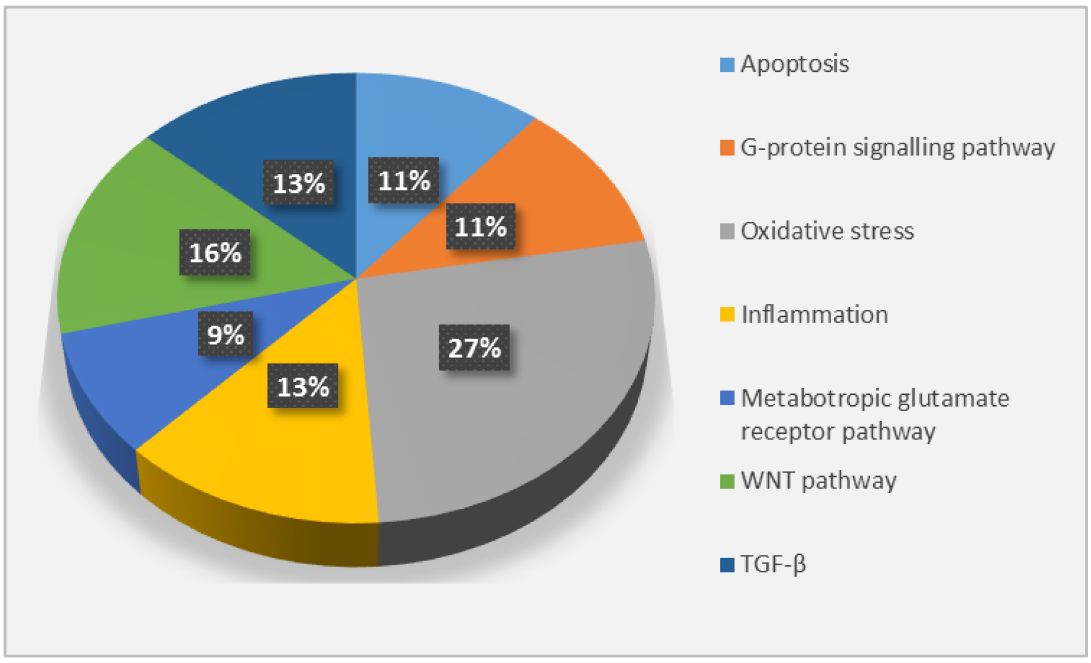

Figure 8. Molecular pathways represented by

Pathway analysis of these DEGs showed that the main represented pathways were WNT and TGF-β signalling pathway. For the WNT signalling pathway we have identified 7 genes, among these: C-MYC which is a target gene of WNT/β-catenin pathway and, among others, it regulates the expression of adipogenic related genes such as C/Ebpα, Pparγ and Fabp4 [180], WNT4 and WNT9B signalling molecules involved in the regulation of adipogenesis and

fibrogenesis [181], SFRP1 and SFRP5 which is supposed to inhibit the canonical as well as the non-canonical WNT signalling [168], SFRP2 which activates fibrogenesis during heart failure [182].

Table 7. AC subject and murine model common genes involved in WNT pathway

	WNT pathway	TgNS<2wks vs TgNS>3wks		TgWt>3wks vs TgNS>3wks		Human VDx		Human VSn	
		FC	pvalue	FC	pvalue	FC	pvalue	FC	pvalue
1	ACTG2	-2,5174	0.000027	-3,0589	0.0007	-5,6451	0.0002	-5,5366	0.000031
2	CDH1	-3,8280	0.000056	-8,7240	0.000013	-7,3411	0.000017	-6,8448	0.000019
3	C-MYC	-2,3232	0,0001	-3,2852	0.000015	-5,2634	0.00002	-3,8280	0.000056
4	SFRP2	-2,1350	0,0033	-3,1624	0.00005	-2,5174	-2,5174437	-3,1102	0.000081
5	SFRP5	-2,1320	0,0063	-3,1932	0.00005	-2,5520	0.000011	-3,9231	0.000017
6	WNT4	-2,0870	0,0008	-2,7683	0.00004	-3,2593	0.00004	-2,5544	0.000064
7	WNT9B	-3,2104	0,0001	-2,7992	0.00003	-3,4174	0.000085	-3,9231	0.00039

Finally, 6 genes were found also targeted for TGF-β pathway, among these: TGF-β2, BMP7, TLL2 which are extracellular ligands of TGFβR1 and TGFβR2, involved in fibrogenesis and inflammatory processes, MSTN which its prolonged over-expression stimulates fibroblast activation and fibrosis via P38/TAK1 signalling pathway [183].

Table 8. AC subject and murine model common genes involved in TGF-β pathway

	TGF-β pathway	TgNS<2wks vs TgNS>3wks		TgWt>3wks vs TgNS>3wks		Human VDx		Human VSn	
		FC	pvalue	FC	pvalue	FC	pvalue	FC	pvalue
1	MSTN	8,3799	0,0001	8,4661	0.0000259	7,1503	0.000081	6,4576	0.000064
2	LEFTY2	4,5884	0.000091	5,9094	0.0000009	7,3411	0.00002	6,8448	0.00039
3	TGFB2	4,3875	0,0000	4,4015	0,00002	8,0222	0.000017	7,0421	0.000031
4	TLL2	-2,8120	0,0001	-2,8540	0.00019	-3,8462	0.0000086	-3,5366	0.000038
5	BMP7	-2,2824	0,0010	-2,8959	0.000124	-3,8280	0.000015	-3,1089	0.000067
6	ACVR1C	-2,8120	0,0001	-3,0441	0.00022	-3,2204	0.000056	-3,6807	0.00013

4.4.6 Data confirmation

Gene expression data obtained with RNA-Seq for all genes involved in the TGF-β and WNT pathways found to be differentially expressed both, in humans and the mouse model was subsequently confirmed by qPCR. The primers used for data confirmation were designed as described in the chapter 3.7.

Subsequently, differential expression was confirmed also at the protein level for MSTN, C-MYC and TGF-β2 in human sample by using western blot analysis.

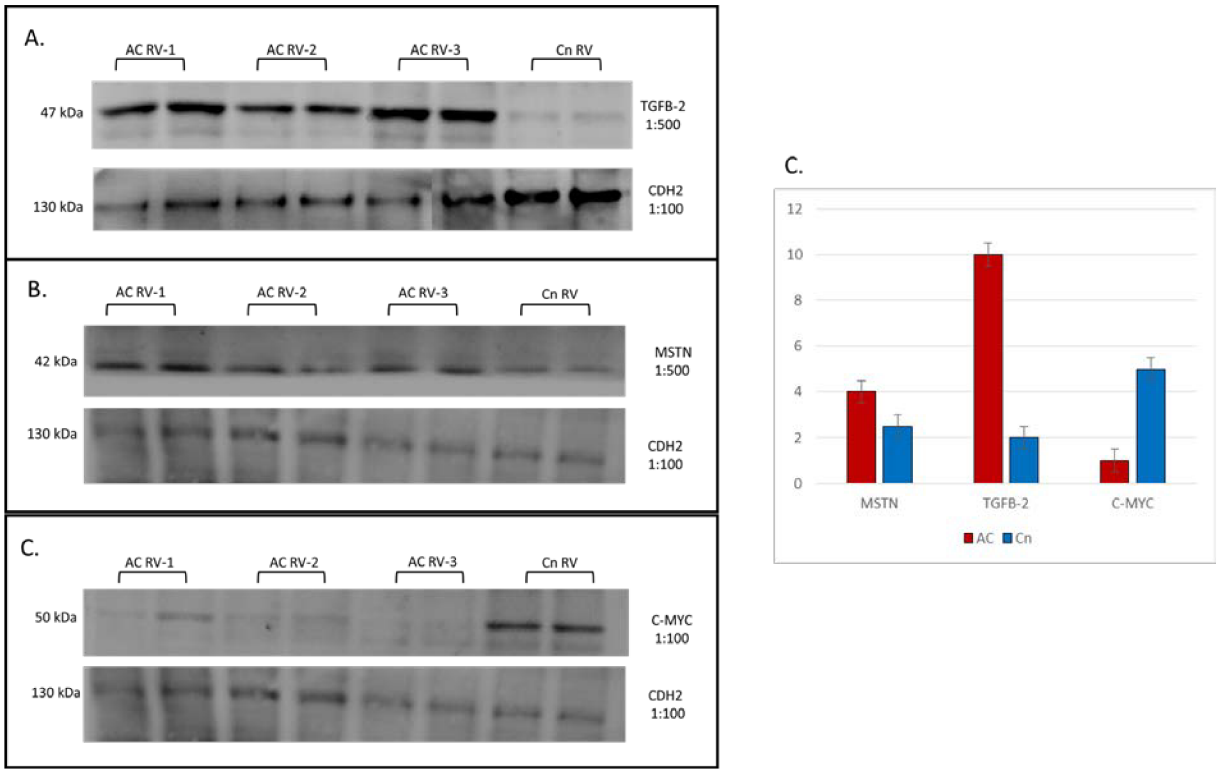

Figure 9. WB results confirmed RNA-seq findings. Here are shown results obtained from human RV.

5 CHAPTER 2: Plakoglobin immune analysis as a diagnostic test for arrhythmogenic cardiomyopathy.

AC is a disease characterized by structural alterations of desmosomes and progressive replacement of ventricular myocardium by fibrofatty tissue. The ventricular arrhythmias appear to occur out of proportion to structural changes. As already mentioned there are four accepted stages of the disease (see chapter 1.2). The variability in clinical presentation end genetic penetrance makes diagnosis challenging, especially during the early stages of disease. A set of clinical criteria has been developed by the international Task Force to aid the diagnosis of this disease, but no gold standard is currently available [14]. EMB is a helpful method to establish the diagnosis of AC [184]. The histological finding of fibrofatty replacement of the myocardium constitutes a major criterion for diagnosis. AC, however, typically spares the cardiac septum, the usual site of EMB, and obtaining biopsy samples from the more commonly involved RV free wall may be associated with significant complications. Additionally, even in patients with septal involvement, the patchy nature of AC may lead to false-negative biopsy results. Thus, there is a need for different, more diffuse histological alterations that can be assessed in any area of heart or even in other more accessible tissues. This need is particularly great in patients with early or mild forms of the disease, in whom the diagnosis based on other clinical characteristics is often even more difficult [92].

A reduction of the immunoreactive signal level for the desmosomal protein JUP has been observed at IDs from patients with Naxos disease and Carvajal syndrome, two rare recessive forms of AC [35, 185]. Furthermore, Asimaki and colleagues have shown that reduced JUP immunoreactive signal level at IDs is a consistent feature in patients with AC and is not seen in other forms of heart-muscle disease. This finding induced the authors to put forward this method as a diagnostic test for AC [92].

The possibility of conducting analysis on EMB makes the test interesting for an early diagnosis of the disease. However, other authors have questioned the potency of the test in AC diagnosis (**Table 9**). In 2012, Munkholm and collaborators, performed immunohistochemistry on 36 AC EMBs and 14 EMBs control samples. They observed a reduction of the JUP immunoreactive signal level in AC samples, but even if the Se of the test is high (85%), the Sp is only 57% [112]. Thus, they concluded advocating caution in basing clinical decision making on this test. Successively the authors published data obtained from

another set of samples with an autopsy diagnosis of AC (40 cases) and 15 cases without heart disease as controls. In this study, no control samples showed reduced JUP staining, while 68% of AC samples displayed an overall reduction [186]. Again, the reduction of JUP IF signal level was observed in AC samples, but not in the controls, in a further study conducted, on both EMBs and autopsy specimens [187]. Subsequently, the ability of the test to distinguish between AC and other cardiomyopathies was proved [188]. This study, like the aforementioned, has as among its limitation the low number of cases (9 AC samples and 7 controls) which makes robust statistical analysis difficult. Noteworthy, each of the studies listed above observed that the dilution of the primary Ab is a factor highly impacting on the test results: the greater the dilution, the greater the test capacity to distinguish between AC and the controls. For example: with a dilution of 1:50.000, there were no differences between the two categories, but when increasing the dilution to 1:250:000 the Se of the test turned 85% and the Sp 57% [112]. Therefore, it is not surprising if some authors described no differences between AC and controls with a Se and Sp nearby 0% with an Ab dilution of 1:100 [189]. The perfect dilution for the best test results is still a matter of debate [1:50.000 [92], 1:75.000 [188], and 1:250.000 [112].

Finally, it has been suggested that the alteration of the desmosomal structure in AC determines the masking of the protein's Ab recognized epitope, instead of a wrong localization of JUP. This idea came from observations on an animal model expressing a DSG2 lacking two exons, recapitulating the AC phenotype [190] and on human samples. The reduction of JUP staining is observed only with JUP ab clone 15F11 that recognizes an epitope at the N-term domain of the protein (the same used in previous papers), but is not observed with other antibodies such as Ab PG5.1 (Progen), which recognizes an epitope at the C-term of the protein. Consequently, the reduction of JUP immunoreactive signal level at the IDs would not be specific to the disease but only to the test used [191].

Table 9. Schematic representation of the systematic literature review

References	Juornal	Sample Source	AC samples	not AC samples*	Control samples**	Antibody	IF		IHC		Sensitivity	Specificity
							Dil. Tested	Dil. Used	Dil. Tested	Dil. Used		
Kaplan et al, 2004	Heart Rhythm	Autopsy	4 (JUP mut)		NA	P8087 - 15F11		NA		NA	100%	100%
		Transplantation										
		EMB										
Kaplan et al, 2004	Cardiovasc Pathology	Autopsy	1 (DSP mut)		NA	P8087 - 15F11		NA		NA	100%	100%
		Transplantation										
		EMB										
Asimaki et al, 2009	N Engl J Med.	Autopsy	9		10	P8087 - 15F11	NA	1:50.000	NA	1:50.000	95%	90%
		Transplantation		15								
		EMB	18	14								
Munkholm et al, 2012	Am J Cardiol	Autopsy				P8087 - 15F11			1:4.000 1:300.000	1:50.000 1:250.000	85% (1:250.000)	57% (1:250.000)
		Transplantation										
		EMB	36	14								
Munkholm et al, 2015	Forensic Sci Med Pathol	Autopsy	38		15	P8087 - 15F11			1:2.000 1:64.000	1:4000	68%	100%
		Transplantation										
		EMB										
Noorman et al, 2013	Cardiovasc Pathology	Autopsy	5		5	P8087 - 15F11	1:100 1:150.000	1:100.000			74%	100%
		Transplantation										
		RVSB	15									
Tavora et al, 2013	Open Cardiovasc Med J	Autopsy	23	6	15	ab12083 - 15F11			NA	1:100	0%	0%
		Transplantation										
		EMB										
Ermakov et al, 2014	PACE	Autopsy				P8087 - 15F11			1:50.000 1:150.000	1:50.000 1:75.000	13% 60%	100%
		Transplantation										
		EMB	9	7								
Kant et al, 2016***	Cardiovasc Res	Autopsy				P8087 - 15F11	NA	15F11 1:30.000 (H); 1:20.000 (m) PG5.1 1:450 (H); 1:30 (m)			15F11: Significant reduced JUP staining at intercalated disk in mice and human; PG5.1: no differences in intercalated disks localization of JUP in mice and human	
		Transplantation										
		EMB	10	14	10	ABIN112236 - PG5.1						
		murine heart	6		6							
Kwon et al, 2013 [*]	Int J Clin Pathol	Autopsy			20	Cell signaling, policlonale C-term			NA	1:100	76%**** (77%)	84%**** (100%)
		Transplantation										
		EMB	22									
Yoshida et al, 2015	Int. Heart J.	Autopsy				Sigma monoclonal			NA	1:100	90%	54%
		Transplantation										
		EMB	10	13								
Padova Ab 15F11		Autopsy	30	24				1:1.000 - 1:50.000		1:50.000 - 1:250.000		
		Transplantation	16									
		EMB	10	14								
Padova Ab PG5.1		Autopsy	15	16					1:100-1:1000	1:250		
		Transplantation	12									
		EMB										

*with other cardiomyopathies; **no clinical or pathological evidence of heart disease; *** Cryosections rather than formalin fixed-paraffin embedded materials; ****Calculated with AC borderline among CN; [*] polyclonal antibody (different from the other studies); EMB: endomyocardial biopsy; RVSB: Right ventricular septum biopsy.

5.1 Results

5.1.1 Cohort

We performed IHC analysis on cardiac tissue samples (biopsies or autopsy samples), obtained from 44 unrelated cases of AC (27 male; mean age 35 years ±16). Genetic analysis was previously conducted on 33 of them. Twenty-nine (79%) of these subjects had a specific mutation in a desmosomal protein gene. The study also included 30 controls (21 male, mean age 33 years ±12). Both, AC and control group are composed of HS (34 of the AC-cases and 20 of the controls) originally obtained for histological examination (autopsy and heart transplantation), and by EMB (10 of the AC-cases and 10 of the controls). All HS and EMB consisted of formalin-fixed, paraffin-embedded cardiac tissue. Among AC HS 27 had BV or RV involvement, while 7 cases had predominantly LV involvement. As control materials, we engaged subjects with a diagnosis other than AC; including 10 HS from cardiac death "*sine materia*", 4 HS with valve prolapse, 3 HS with atherosclerosis, 1 HS with myocarditis, 1 HS with anomalous origin of coronary artery and 1 HS from a case of Idiopathic LV Hypertrophy. Among the EMB control cases we included: 6 specimens from normal heart and 4 with a diagnosis of myocarditis.

5.1.2 IHC analysis

We first executed an IPOX analysis on areas of structurally normal myocardium, investigating the capability of two JUP Ab dilutions (1:50.000 and 1:250.000) to discriminate between AC and non-AC cases. Among HS, while using the lower dilution, we found that 24 out of the 34 AC-patients had a reduced JUP signal level; thus, the test Se was 70.59%. However, among the non-AC HS we found that the JUP immune-signal level was reduced in 10 of the 20 cases, revealing a test Sp of only 50% (**Table 10**).

Table 10. IPOX analysis: results obtained with 1:50000 JUP antibody dilution.

Samples	N° Samples	Jup reduction	Se	Sp	PPV	NPV
AC HS	34	24	70,5	50,0	70,6	50,0
Cn HS	20	10				
AC EMB	10	4	40,0	80,0	66,7	57,1
Cn EMB	10	2				

Se: test sensitivity; Sp: test specificity; PPV: test positive predictive value; NPV: test Negative Predictive Value.

The same analysis was performed on EMBs, with the following results. Reduction of the JUP signal level was observed in 4 out of the 10 AC-cases, and in 2 of the 10 non-AC cases (1:50.000 Ab dilution). Accordingly, when using the same Ab dilution, the test Se (40%) is lower in EMBs than in HS, while the Sp is higher (80%).

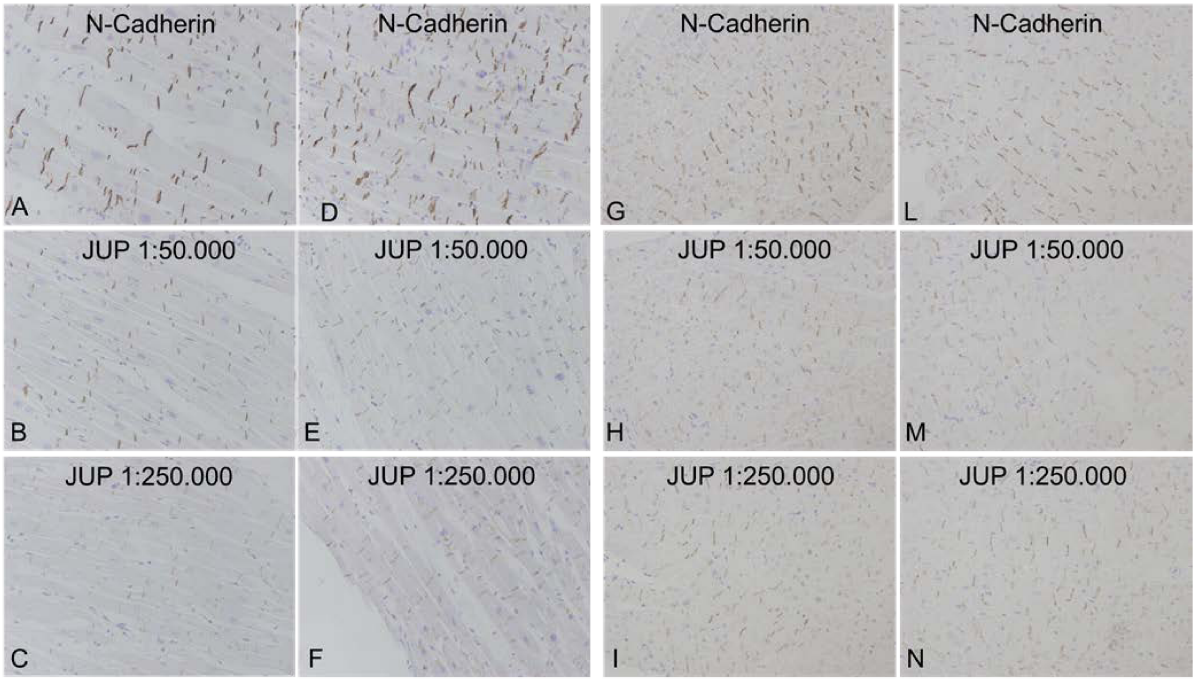

Figure 10. A, B, C images depict IPOX performed on HS of a 'not AC case', whereas images D, E, F are coming from HS of an AC case. G, H, I images depict IPOX performed on EMB of a 'not AC case', whereas the images L, M, N are coming from an AC case. Note that no JUP signal reduction was observed on AC specimens for both Ab dilutions.

With an Ab dilution of 1:250.000 we found a reduced JUP signal level in 27 out of the 34 AC cases and in 13 out of the 20 non-AC cases, thus increasing test Se up to 79.41% and reducing Sp to 35%. Lastly, the analysis on EMBs with an Ab dilution 1:250.000 revealed a decrease in the JUP immune-signal in 5 out of the 10 AC, and in 3 out of the 10 non-AC cases, consequently showing a Se of 50% and a Sp of 70% (**Table 11**).

Table 11. IPOX analysis: results obtained with 1:250000 JUP antibody dilution.

Samples	N° Samples	Jup reduction	Se	Sp	PPV	NPV
AC HS	34	27	79,4	35,0	67,5	50,0
Cn HS	20	13				
AC EMB	10	5	50,0	70,0	62,5	58,3
Cn EMB	10	3				

Se: test sensitivity; Sp: test specificity; PPV: test positive predictive value; NPV: test Negative Predictive Value.

The same cases were analysed also by IF using two different JUP Ab dilutions, 1:1000 and 1:50.000. Among HS a decreased JUP signal level was observed in 21 out of 34 AC cases and in 11 out of 20 non-AC cases (Se: 61,7%, Sp: 45%). Results regarding EMB samples showed a Se dropped to 50% and a Sp of 70%, in fact decreased JUP signal levels were identified in on 5 out of 10 AC cases and on 3 out of 10 non-AC cases (**Table 12**).

Table 12. IF analysis: results obtained with 1:1000 JUP antibody dilution.

Samples	N° Samples	Jup reduction	Se	Sp	PPV	NPV
AC HS	34	21	61,8	45,0	65,6	40,9
Cn HS	20	11				
AC EMB	10	5	50,0	70,0	62,5	58,3
Cn EMB	10	3				

Se: test sensitivity; Sp: test specificity; PPV: test positive predictive value; NPV: test Negative Predictive Value.

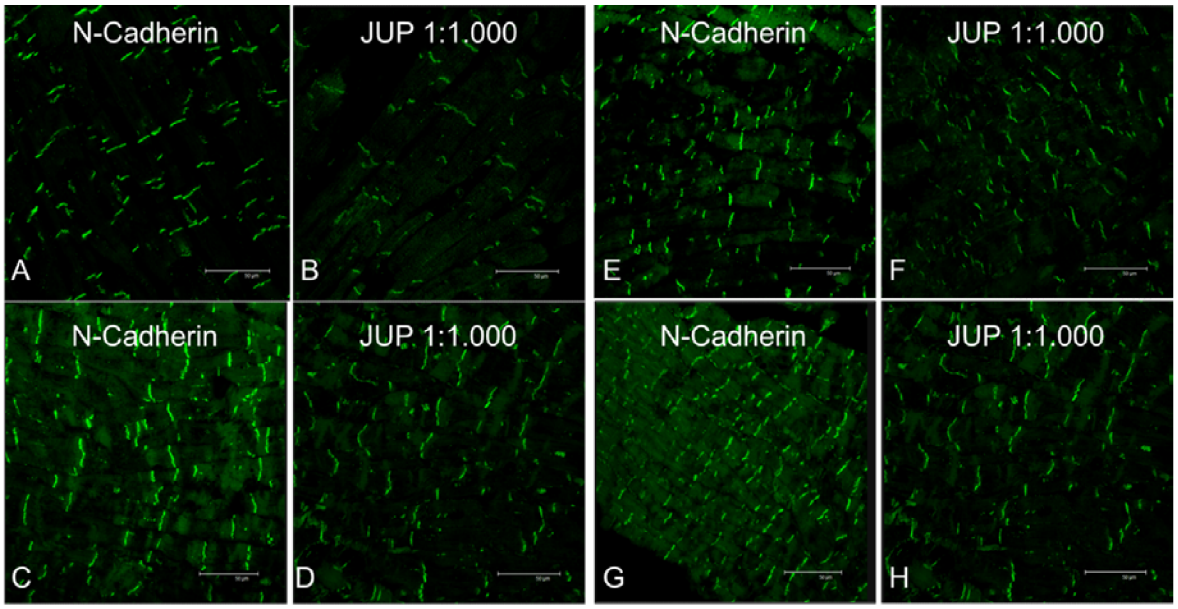

Figure 11. Upper images (A, B) depict IF performed on HS of a 'not AC case', whereas lower images (C, D) are coming from an AC case. Upper images (E, F) depict IF performed on EMB of a 'not AC case', whereas lower images (G, H) are coming from EMB of an AC case. Note that no JUP signal reduction was observed on AC specimens for both Ab dilutions.

Similar data were obtained with a 1:50.000 dilution of the JUP Ab (for HS Se: 79.41%, Sp: 35%; EMB Se: 60%, Sp: 70%) as reported in **Table 13**.

Table 13. IF analysis: results obtained with 1:50000 JUP antibody dilution.

Samples	N° Samples	Jup reduction	Se	Sp	PPV	NPV
AC HS	34	27	79,4	35,0	67,5	50,0
Cn HS	20	13				
AC EMB	10	6	60,0	70,0	66,7	63,6
Cn EMB	10	3				

Se: test sensitivity; Sp: test specificity; PPV: test positive predictive value; NPV: test Negative Predictive Value.

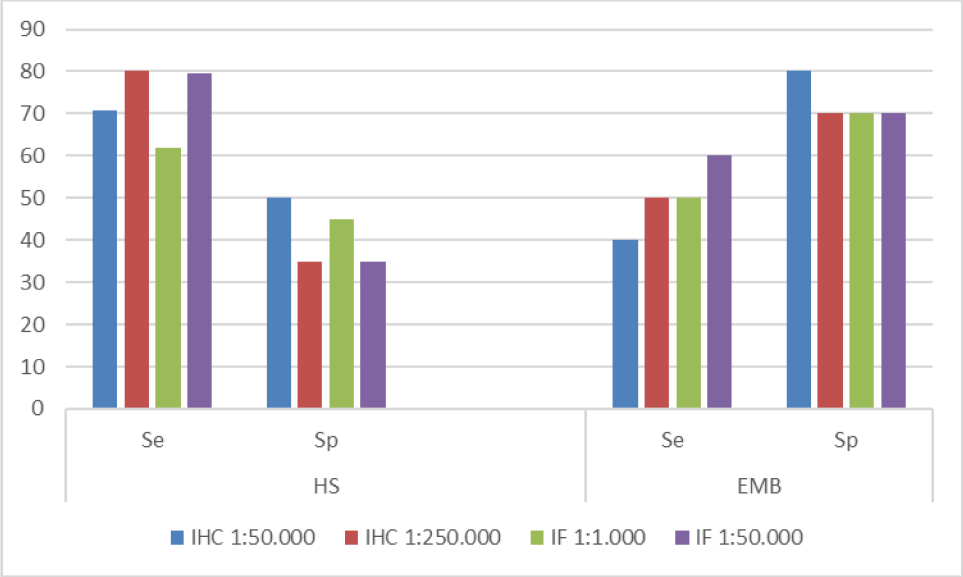

Figure 12. Test Se and Sp: On the left, the graph is showing the Se and Sp of the test when performed on HS, while on the right the graph shows the Se and Sp of the test when performed on the EMB.

For each test, ROC analysis was conducted using MedCalc software. As shown in **Figure 13** and **Figure 14** in both IPOX and IF the area under the curve (AUC) is lower than 80% (the threshold value over which a diagnostic test is considered appropriate) [192].

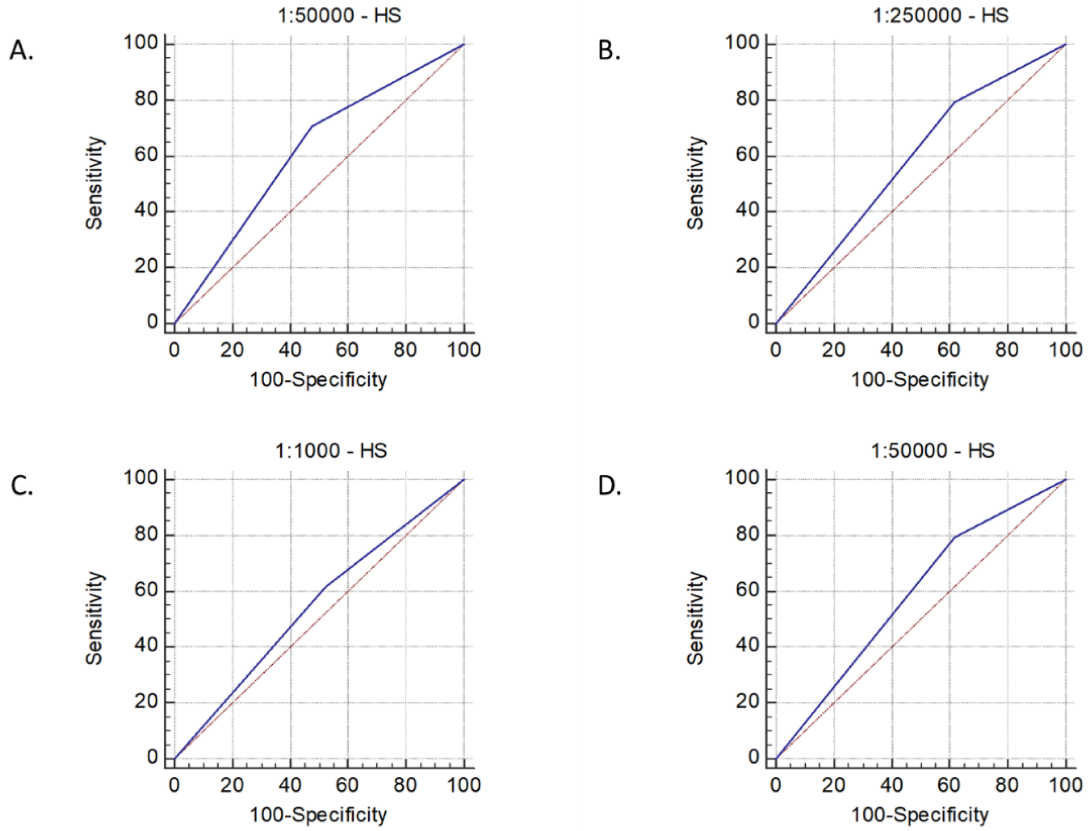

Figure 13. ROC curve analysis of test performed on HS. A. IPOX analysis performed with antibody dilution 1:50000; B. IPOX analysis performed with antibody dilution 1:250000; C. IF analysis performed with antibody dilution 1:1000; D. IF analysis performed with antibody dilution of 1:50000.

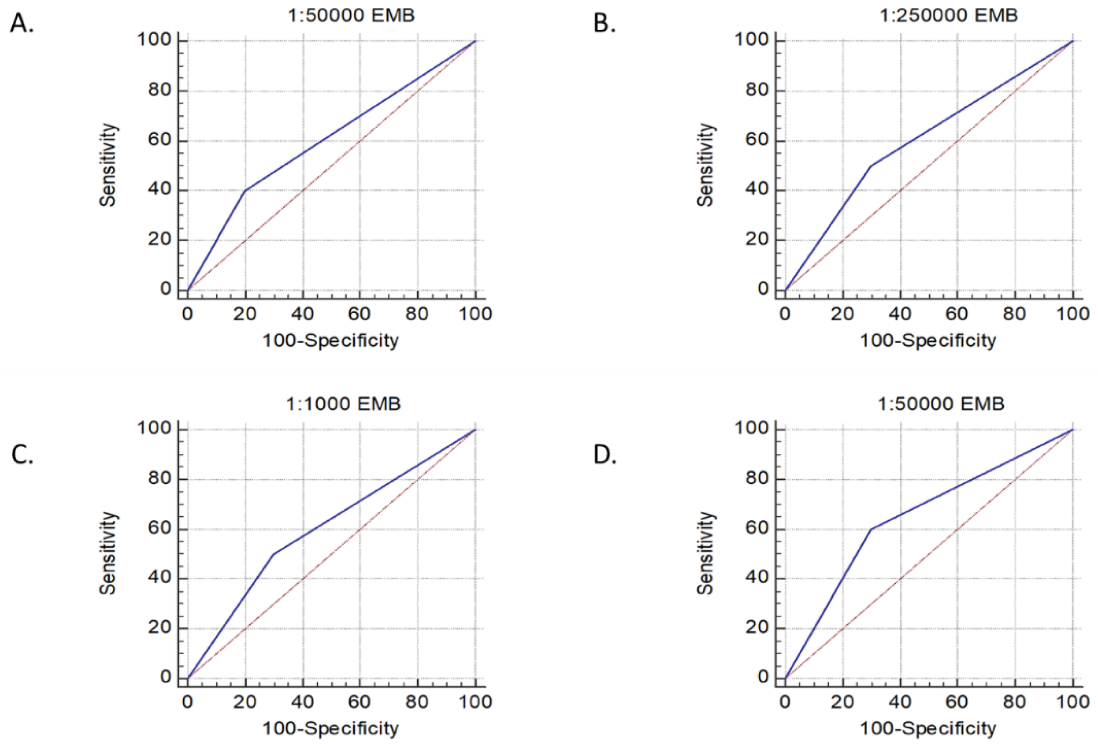

Figure 14. ROC curve analysis of test performed on EMB. A. IPOX analysis performed with antibody dilution 1:50000; B. IPOX analysis performed with antibody dilution 1:250000; C. IF analysis performed with antibody dilution 1:1000; D. IF analysis performed with antibody dilution of 1:50000

All data described above were obtained by using the 15F11 Ab clone that recognizes an epitope at the N-terminal part of the protein. To confirm results, proposed by Kant [191] analysis was conducted also with a second Ab that recognizes another protein epitope. The Ab used in the second part of this analysis, PG5.1 (Progen) recognize an epitope located in the C-terminal part of the protein. IPOX analysis of HS cases has allowed us to detect JUP signal reduction in 11 of the 25 AC cases and in 7 of the 15 control cases. The test in this case shows a Se of 33% and a Sp of 53% as shown in Table 14.

Table 14. IPOX analysis: results obtained with 1:250 of PG5.1 antibody

Samples	N° Samples	Jup reduction	Se	Sp	PPV	NPV
AC HS	15	5	33,0	53,3	41,6	44,4
Cn HS	15	7				

Se: test sensitivity; Sp: test specificity; PPV: test positive predictive value; NPV: test Negative Predictive Value.

6 CHAPTER 3: Study of the genetic variants frequency in the population of the Veneto region

Roughly one-half of all AC cases is recognized as familial, with an autosomal dominant inheritance as well as incomplete penetrance and variable expressivity. As previously reported, mutations in several AC susceptibility genes that encode essential desmosomal proteins have been identified (J*UP, DSP, PKP2, DSG2, DSC2*) [52]. The detection of novel genetic variants is rapidly increased with the introduction of high throughput technologies into the clinical diagnostics. This make the interpretation of variant's clinical significance a challenge. Even with the lack of a comprehensive and collectively-accepted interpretation protocol, variants classification is aided by available interpretation guidelines [193], which recently have been improved and updated [194]. The American College of Medical Genetics (ACMG) recommends a correct annotation and a standardized terminology in the reporting of nucleotide variants [195]. Furthermore, they recommend a careful assessment of pathogenicity by evaluating a string of parameters such as evolutionary conservation, population studies, functional in vitro studies, clinical presentation, family history and co-segregation. The availability of reliable databases collecting genetic variants and information on genotype-phenotype associations helps to discern between benign variants and disease-causing variants, for example by reporting for instance whether a specific variant had been already described in another patient. A number of established public databases such as Human Gene Mutation Database (HGMD) [196], ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) [197] focusing on inherited diseases and reporting information on genotype-phenotype associations are available. Several web-based, gene-centred, locus-specific databases have also been developed, such as the Leiden Open Variation Database (LOVD) [198], that provides a flexible, freely available tool for gene-centred collection and display of DNA variants. In particular, comprehensive information about all known AC-associated variants are collected and shared in the ARVD/C database since 2009 [199]. Further, the 1000 Genomes Project and Exome Variant Server (EVS) report millions of SNV obtained from deep sequencing of large population cohorts (1000 Genomes Project Consortium, 2010; Exome Variant Server, 2011), and report information about the variant frequencies in the population.

As with other clinical tests, a true understanding of the extent and spectrum of genetic variation found in a healthy population is necessary for correct interpretation. Although significant research has been dedicated to the identification of pathogenic mutations, little is

known about the background genetic variation in genes constituting the clinically available AC genetic tests. Understanding this signal-to-noise ratio is critical to determining whether an identified variant of undetermined or uncertain significance might be the biomarker responsible for disease in the proband or whether it is a rare genetic variant with no relevance to the disease state. Specifically, missense mutations were hosted by 16% of controls versus a similar 21% of AC cases [200]. This poses several challenges to the interpretation of AC genetic tests, as 1 in 6 healthy individuals would meet current criteria for a so-called positive AC genetic test result, even with proper qualification of these rare "mutations" with the clinically ambiguous designation as a "variant of uncertain significance" Another study provided confirmatory data by reporting that 18% (38 genetic variants) of previously reported AC pathogenic genetic variants were identified in the Exome Sequencing Project (ESP) population (1 nonsense and 37 missense) [201]. It is noteworthy that convincing family co-segregation was reported in only 1 of these 38 pathogenic genetic variants and 3 variants had also functional characterization showing significant differences between mutant and Wt transfected cells. Thus far, 190 pathogenic missense variants have been associated with AC, all reported in Human Gene Mutation Database (HGMD): these variants are distributed as follows: 56 in DSP, 49 in DSG2, 48 in PKP2, and 24 in DSC2. From all reported pathogenic missense pathogenic genetic variants associated with AC, only 62 % predicted to be deleterious by in silico platforms, using PROVEAN (Protein Variation Effect Analyzer) (http://provean.jcvi.org/ index.php). Moreover, only 57 % are predicted to be deleterious using Condel (CONsensus DELeteriousness score of non- synonymous SNV) (http://bg.upf.edu/ condel/analysis), which integrates five different platforms (Polyphen-2, SIFT, MAPP, LogR Pfam E -value, and Mutation Assessor). On the other hand, in the Kapplinger study, radical mutations (insertions or deletions, splice junction mutations, and nonsense mutations), were significantly more prevalent in AC cases compared to controls (50 % vs 0.5 % respectively, p=9.8x10-44), indicating that this type of genetic variant has a high likelihood of being associated with AC pathogenicity [200]. Notably, radical mutations constituted the majority (75 of 102, 73.5%) of genetic alterations identified in mutation-positive AC cases. Thus far, 155 radical mutations have been associated with AC, all reported in HGMD, and distributed as following: 75 in PKP2, 52 in DSP, 16 in DSG2, and 12 in DSC2. Despite of being reported in the HGMD database, not all radical genetic variants denote a pathogenic role in clinical practice, especially if no family segregation has been reported. For this purpose, guidelines in genetics [193, 194], highly recommend performing

accurate in silico analysis, in vitro assays, and, as key point, co-segregation studies in families to assess the pathogenicity on AC.

6.1 Results

6.1.1 Genetic screening of 5 desmosomal genes

200 out of 500 subjects underwent genetic screening for 5 genes encoding for desmosomal protein (see chapter 1.4).

We identified a total of 131 different variants. Forty-three of the 131 (32.8%) variants were intronic, 1 (0.76%) resided in a splice site and 2 (1.5%) were located in the untranslated regions (UTR) of the gene. The remaining 86 (65.6%) variants were located within the coding regions; 56 (61.6%) of these were missense variants, 30 (34.9%) were synonymous, and 1 (1.2%) was a frame shift variant, no nonsense mutations were identified. A summary of the detected nucleotide variants is reported in ***Figure 15***.

Variants in JUP. The 14 exons of the JUP gene were analysed for the presence of variants in all 200 index subjects. A total of 13 (9.9%) variants were identified; 5 of which were intronic, 1 was located in the UTR, and the remaining 7 were exonic. 4 were missense variants and 3 were synonymous.

Variants in DSP. Genetic screening of all 24 DSP gene exons identified 47 (35.9%) variants in the 200 index subjects 14 of the identified variants were intronic, 1 was a splice site variant, 1 was located in the UTR, 16 were missense variants, and 14 were synonymous No frameshift or nonsense variants were identified.

Variants in PKP2. The 14 exons of the PKP2 gene were examined for variants in all 200 index subjects. A total of 22 (16.8%) variants were identified, of which 6 were intronic, 15 were missense variants and 1 was a synonymous change.

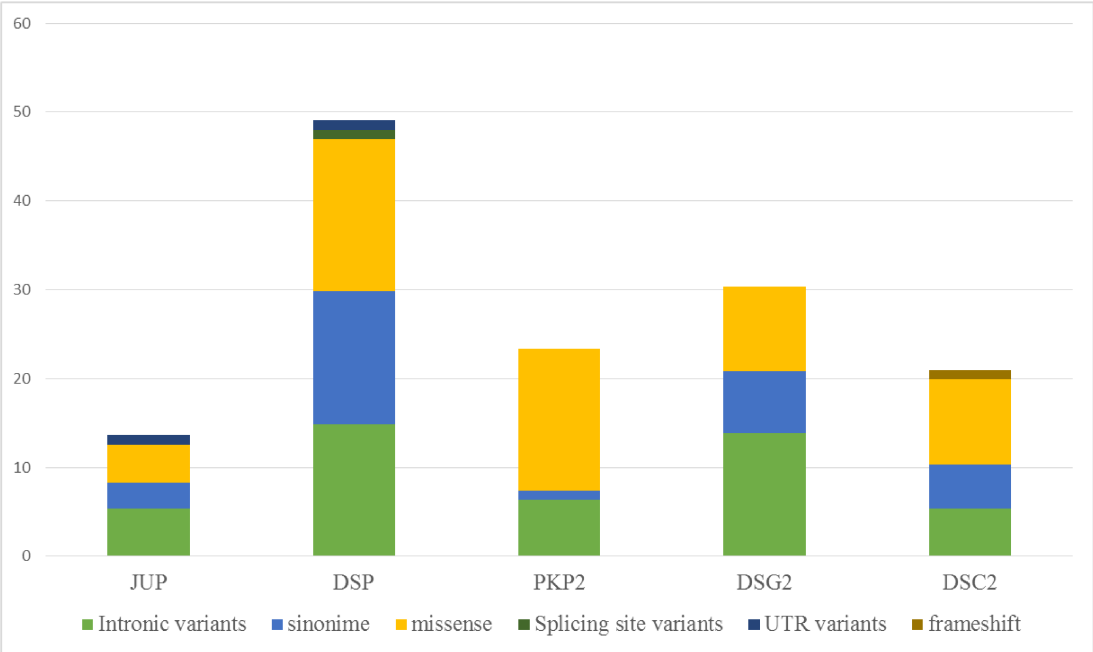

Figure 15. The 131 variants detected in 5 genes, sorted per gene and type (intronic, missense, synonymous, frameshift, nonsense, splicing, UTR).

Variants in DSG2. The 15 exons of the *DSG2* gene were examined for variants in all 200 index subjects. A total of 29 (22.1%) variants were identified in *DSG2*, 13 of them were intronic, while 16 were located within the coding region. Nine of them were missense and 7 were synonymous.

Variants in DSC2. The 14 exons of DSC2 gene were examined for variants in all 200 index subjects. A total of 20 (15.3%) variants were identified; 5 of them were intronic, and the remaining 15 were located in the coding region (9 missense variants, 1 frameshift, 5 synonymous, 0 nonsense variants).

6.1.2 Variants filtering

A first filtering step based on the position and function of nucleotide variants, excluded 44 of the 131 unique variants due to intronic localization without predicted functional effects on the normal splicing of the transcript, 2 variants were excluded due to UTRs localization and 30 synonymous variants were excluded as not changing the functional properties of the protein. The remining 55 variants are listed in **Table 15**.

Table 15. Missense, frameshift and splicing site variants found in healthy athletes.

Nr.		Gene	cDNA	gDNA	aa	Exon/ Intron	dbSNP	MAF (ExAC)	dbSNP Clinical significance	Subjects
1	1	JUP	c.425 G>A	g.39925713	R142H	exon 3	rs41283425	0.028/61	other	31
2	2	JUP	c.427G>A	g.39925711	A143T	exon 3	rs375788626	0.0001/11	uncertain	1
3	3	JUP	c.1942G>A	g.39913771	V648I	exon 12	rs143043662	0.003/13	benign	1
4	4	JUP	c.2089A>T	g.39912145	M697L	exon 14	rs1126821	0.429/936	other	173
5	1	DSP	c.88G>A	g.7542236	V30M	exon 1	rs121912998	0.002/4	Pathogenic	2
6	2	DSP	c.273+5G>C	g.7556058		intron 2	rs.200473206	0.0004/39		1
7	3	DSP	c.535G>A	g.7559571	G179S	exon 4	rs774829433	0.00002/2	NA	1
8	4	DSP	c.913A>T	g.7565727	I305F	exon 7	rs17604693	0.017/1	uncertain	12
9	5	DSP	c.1481A>T	g.7569480	Y494F	exon12	rs28763961	0.018/89	other	1
10	6	DSP	c.3075A>C	g.7578786	E1025D	exon 22	NA	NA	NA	1
11	7	DSP	c.3388G>A	g.7579811	D1130N	exon 23	rs765769174	0.00002/3	NA	1
12	8	DSP	c.3616T>A	g.7580039	L1206I	exon 23	rs151115778	0.00009/11	uncertain	1
13	9	DSP	c.4372C>G	g.7580795	R1458G	exon 23	rs28763965	0.0018/9	other	2
14	10	DSP	c.4455G>T	g.7580878	R1485S	exon 23	rs113902911	0.0018/211	Likely benign	1
15	11	DSP	c.4471G>A	g.7580894	D1491N	exon 23	rs768402359	0.00002/3	NA	1
16	12	DSP	c.4535A>G	g.7580958	Y1512C	exon 23	rs2076299	0.183/399	benign	12
17	13	DSP	c.5005C>A	g.7581428	L1669I	exon 23	NA	NA	NA	1
18	14	DSP	c.5178 C>A	g.7581601	N1726K	exon 23	rs147415451	0.0009/104	other	1
19	15	DSP	c.5213 G>A	g.7581636	R1738Q	exon 23	rs6929069	0.222/485	benign	26
20	16	DSP	c.5498A>T	g.7582993	E1833V	exon 24	rs78652302	0.008/18	other	4
21	17	DSP	c.6208G>A	g.7583703	D2070K	exon 24	rs41302885	0.0039/471	other	2
22	1	PKP2	c.76G>A	g.33049590	D26N	exon 1	rs143004808	0.003/6	other	2
23	2	PKP2	c.89T >C	g.33049577	L30P	exon 1	NA	NA	NA	1
24	3	PKP2	c.174 G>T	g.33049492	E58D	exon 1	rs146708884	0.003	other	3
25	4	PKP2	c.209G>T	g.33049457	S70I	exon 1	rs75909145	0.009/20	Likely benign	12
26	5	PKP2	c.419C>T	g.33031395	S140F	exon 3	rs150821281	0.001/3	other	1
27	6	PKP2	c.505A>G	g.33031309	S169G	exon 3	rs139139859	0.0002/1	other	1
28	7	PKP2	c.605T>C	g.33031209	V202A	exon 3	rs751341106	NA	NA	1
29	8	PKP2	c.989 G>T	g.33030825	G330V	exon 3	rs373041797	NA	NA	1
30	9	PKP2	c.1012A>G	g.33030802	T338A	exon 3	rs139851304	0.001/2	other	1
31	10	PKP2	c1093 A>G	g.33021938	M365V	exon 4	rs143900944	0.0006/3	uncertain	1
32	11	PKP2	c.1097T>C	g.33021934	L366P	exon 4	rs1046116	0.143/312	Likely benign	70
33	12	PKP2	c.1592 T>G	g.32994057	I531S	exon 7	rs148240502	0.003/6	other	2
34	13	PKP2	C.1759 G<A	g.32977026	V587I	exon 8	rs146102241	0.001	other	3
35	14	PKP2	c.2083 C>T	g.32974352	R695C	exon 10	rs199583774	0.0001/18	uncertain	1
36	15	PKP2	c.2431C>A	g.32949101	R811S	exon 12	rs139734328	0.001/3	uncertain	1
37	1	DSG2	c.166G>A	g. 29099850	V56M	exon 3	rs121913013	0.001/1	other	1
38	2	DSG2	c.473 T>G	g.29101156	V158G	exon 5	rs191143292	0.001/7	other	2
39	3	DSG2	c.877A>G	g.29104714	I293V	exon 8	rs2230234	0.039/84	Likely benign	23
40	4	DSG2	c.1303 G>A	g.29115255	D435N	exon 10	rs370509593	0,0006/3	other	1
41	5	DSG2	c.1912 G>A	g.29121188	G638K	exon 13	rs201564919	0.0001/15	uncertain	1
42	6	DSG2	c.2137G>A	g.29122618	E713K	exon 14	rs79241126	0.0777/82	Likely benign	29
43	7	DSG2	c.2318G>A	g.29122799	R773K	exon 14	rs2278792	0.272/594	Likely benign	67
44	8	DSG2	c.2759T>G	g.29126108	V920G	exon 15	rs142841727	0.004/8	other	4
45	9	DSG2	c.2953 G>A	g.29126302	V985I	exon 15	rs749540432	0.00002/2	uncertain	1
46	1	DSC2	c.304 G>A	g.28672114	E102K	exon 3	rs144799937	0,0007	other	2
47	2	DSC2	c.1484 C>T	g.28660098	A495V	exon 10	NA	NA	NA	1
48	3	DSC2	c.1609 G>T	g.28659867	A537S	exon 11	NA	NA	NA	1
49	4	DSC2	c.1787 C>T	g.28654750	A596V	exon 12	rs148185335	0.001/3	uncertain	1
50	5	DSC2	c.1879 G>A	g.28654658	A627T	exon 12	rs762223215	0.00002/2	NA	1
51	6	DSC2	c.1914 G>C	g.28651782	Q638H	exon 13	rs147742157	0,0004/50	other	1
52	7	DSC2	c.2326 A>G	g.28649042	I776V	exon 15	rs1893963	0.19/415	Likely benign	31
53	8	DSC2	c.2393 G>A	g.28648975	R798Q	exon 15	rs61731921	0.189/412	uncertain	17
54	9	DSC2	c.2657A>C	g.28648030	N886T	exon 16	NA	NA	NA	1
55	10	DSC2	c.2686_2687dupGA	g.28648001	A897Kfs*4	exon 16	rs200056085	0.006/12	other	5

MAF: Minor allele Frequency (from ExAC database); NA: data not available.

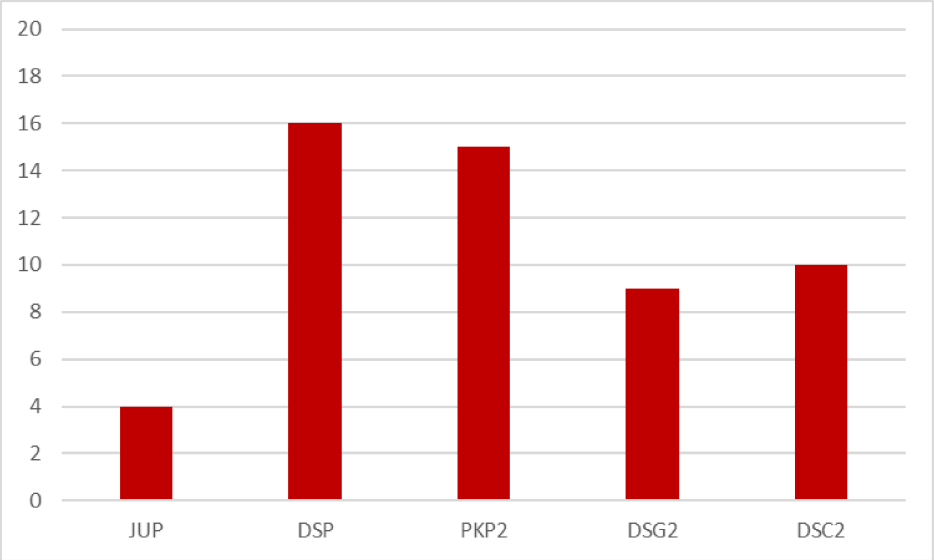

Figure 16. Variants distribution among the 5 desmosomal genes after the first filtering step.

6.1.2.1 Allelic frequency variant filtering

Considering that the prevalence of the disease in the general population is estimated between 1:2000-1:5000 [7], we used two cut-offs for the selection of variants. Those with Minor allele frequency (MAF) comprised among 0.0005 and 0.0002 and those with MAF <0.0002. 18 of the identified variants presented a MAF <0.0002 or are not described in the databases, were present in a total of 18 subjects (9%), while in the MAF range 0.0002 to 0.0005 only 2 variants were selected (1 missense variant and 1 variant that could alter splicing during maturation of the mRNA) as summarized in Table 16. All identified variants were present in the heterozygous state. It is worth mentioning that the duplication c.2686_2687dup in the DSC2 gene leading to a frame-shift (p.A897KfsX4) was also excluded due to high MAF (0.006/12). Indeed, this variant, which was recently identified in 1.5% of the Italian healthy control subjects, was demonstrated to affect only the DSC2 isoform a, while the isoform b (more expressed in the heart) was not altered [202]. Intriguingly, one excluded variant (p.V30M in DSP) harboured by 2 subjects in our study group is currently labelled as pathogenic by the dbSNP database. An in vitro study reported that this variant determined a wrong localization of DSP due to the alteration of the JUP binding site of the protein [75]. This variant was excluded not only because of its high frequency not only in the freely available database, but also in our study population.

After this filtering process, the distribution of the spectrum of heterozygous variants as reported in Figure 17 showing a high prevalence of variants in DSP (7, 35%), followed by PKP2 and DSC2 (5, 25%), DSG2 (2, 10%) and JUP (1, 5%).

Table 16. Variants with MAF <0.0002; in grey are highlighted variant with 0.0005>MAF>0.0002.

Nr.		Gene	cDNA	gDNA	aa	Exon/ Intron	dbSNP	MAF (ExAC)	Clinical significance	Subjects	aa conservation
1	1	JUP	c.427G>A	g.39925711	A143T	exon 3	rs375788626	0.0001/11	uncertain (dbSNP)	1	Moderately
2	1	DSP	c.273+5G>C	g.7556058		intron 2	rs.200473206	0.0004/39	pathogenic (HGMD)	1	-
3	2	DSP	c.535G>A	g.7559571	G179S	exon 4	rs774829433	0.00002/2	NA	1	Conserved
4	3	DSP	c.3075A>C	g.7578786	E1025D	exon 22	NA	NA	NA	1	Conserved
5	4	DSP	c.3388G>A	g.7579811	D1130N	exon 23	rs765769174	0.00002/3	NA	1	Moderately
6	5	DSP	c.3616T>A	g.7580039	L1206I	exon 23	rs151115778	0.00009/11	uncertain	1	Moderately
7	6	DSP	c.4471G>A	g.7580894	D1491N	exon 23	rs768402359	0.00002/3	NA	1	Weakly
8	7	DSP	c.5005C>A	g.7581428	L1669I	exon 23	NA	NA	NA	1	Weakly
9	1	PKP2	c.89T >C	g.33049577	L30P	exon 1	NA	NA	NA	1	Conserved
10	2	PKP2	c.505A>G	g.33031309	S169G	exon 3	rs139139859	0.0002/1	other	1	Weakly
11	3	PKP2	c.605T>C	g.33031209	V202A	exon 3	rs751341106	NA	NA	1	Weakly
12	4	PKP2	c.989 G>T	g.33030825	G330V	exon 3	rs373041797	NA	NA	1	Moderately
13	5	PKP2	c.2083 C>T	g.32974352	R695C	exon 10	rs199583774	0.0001/18	uncertain	1	Conserved
14	1	DSG2	c.1912 G>A	g.29121188	G638K	exon 13	rs.201564919	0.0001/15	uncertain	1	Conserved
15	2	DSG2	c.2953 G>A	g.29126302	V985I	exon 15	rs749540432	0.00002/2	uncertain	1	Weakly
16	1	DSC2	c.1484 C>T	g.28660098	A495V	exon 10	NA	NA	NA	1	Conserved
17	2	DSC2	c.1609 G>T	g.28659867	A537S	exon 11	NA	NA	NA	1	Conserved
18	3	DSC2	c.1879 G>A	g.28654658	A627T	exon 12	rs762223215	0.00002/2	NA	1	Weakly
19	4	DSC2	c.1914 G>C	g.28651782	Q638H	exon 13	rs147742157	0,0004/50	other	1	Weakly
20	5	DSC2	c.2657A>C	g.28648030	N886T	exon 16	NA	NA	NA	1	Weakly

MAF: Minor Allele Frequency (from ExAC database); NA: data not available.

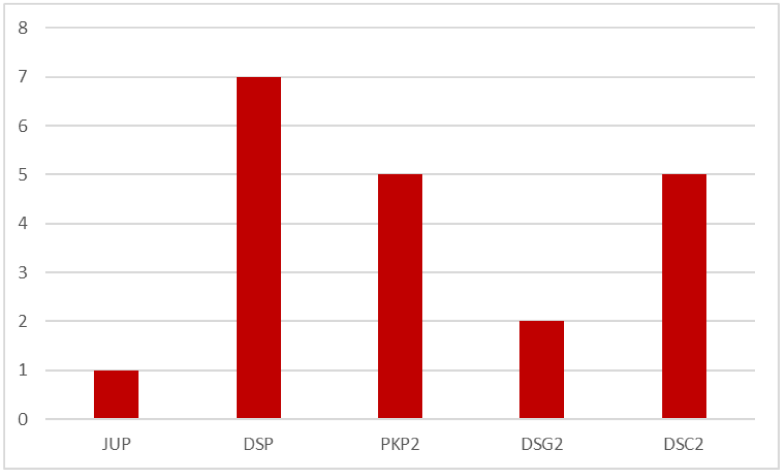

Figure 17. Spectrum of heterozygous variants distribution after the variants allelic frequency filtering step.

Considering the approach proposed by Whiffin and colleagues [203], we have set the MAF cut off at 0.000018. Indeed, using their algorithm:

$$\text{MAF cut-off} = 1/5000 \times ½ \times 0.092 \times 1/0.5 = 0.000018$$

1/5000 x 1/2: dividing disease prevalence per individual by the number of chromosomes per individual.

0.092: frequency of the commonest causative variant [203]

0.5: disease penetrance

Using this MAF cut-off the unique variants remaining after this filtering step were those with not available MAF.

6.1.2.2 Evolutionary conservation based filtering

Amino acid conservation analysis excluded other 8 missense variants, as shown in **Table 17** and subsequently reduced the number of variants to 12.

Table 17. Variants with MAF<0.0005 with weakly amino acidic conservation.

Nr.		Gene	cDNA	gDNA	aa	Exon/ Intron	dbSNP	MAF (ExAC)	Clinical significance	Subjects	aa conservatio
1	1	DSP	c.4471G>A	g.7580894	D1491N	exon 23	rs768402359	0.00002/3	NA	1	Weakly
2	2	DSP	c.5005C>A	g.7581428	L1669I	exon 23	NA	NA	NA	1	Weakly
3	1	PKP2	c.505A>G	g.33031309	S169G	exon 3	rs139139859	0.0002/1	other	1	Weakly
4	2	PKP2	c.605T>C	g.33031209	V202A	exon 3	rs751341106	NA	NA	1	Weakly
5	1	DSG2	c.2953 G>A	g.29126302	V985I	exon 15	rs749540432	0.00002/2	uncertain	1	Weakly
6	1	DSC2	c.1879 G>A	g.28654658	A627T	exon 12	rs762223215	0.00002/2	NA	1	Weakly
7	2	DSC2	c.1914 G>C	g.28651782	Q638H	exon 13	rs147742157	0,0004/50	other	1	Weakly
8	3	DSC2	c.2657A>C	g.28648030	N886T	exon 16	NA	NA	NA	1	Weakly

MAF: Minor Allele Frequency (from ExAC database); NA: data not available.

Among the obtained 12-different putative pathogenic variants (see Table 18), 5 occurred in DSP, 3 in PKP2, 2 in DSC2 and 1 in DSG2 and JUP respectively. Eleven of them were missense variants and 1 occurred in a putative donor acceptor splicing site located within the exon-intron boundaries (±bp).

Table 18. Variants with moderately/highly conserved amino acids.

Nr.		Gene	cDNA	gDNA	aa	Exon/ Intron	dbSNP	MAF (ExAC)	Clinical significance	aa conservation	Subjects
1	1	JUP	c.427G>A	g.39925711	A143T	exon 3	rs375788626	0.0001/11	uncertain (dbSNP)	Moderately	1
2	1	DSP	c.273+5G>C	g.7556058		intron 2	rs.200473206	0.0004/39	pathogenic (HGMD)	-	1
3	2	DSP	c.535G>A	g.7559571	G179S	exon 4	rs774829433	0.00002/2	NA	Conserved	1
4	3	DSP	c.3075A>C	g.7578786	E1025D	exon 22	NA	NA	NA	Conserved	1
5	4	DSP	c.3388G>A	g.7579811	D1130N	exon 23	rs765769174	0.00002/3	NA	Moderately	1
6	5	DSP	c.3616T>A	g.7580039	L1206I	exon 23	rs151115778	0.00009/11	uncertain	Moderately	1
7	1	PKP2	c.89T >C	g.33049577	L30P	exon 1	NA	NA	NA	Conserved	1
8	2	PKP2	c.989 G>T	g.33030825	G330V	exon 3	rs373041797	NA	NA	Moderately	1
9	3	PKP2	c.2083 C>T	g.32974352	R695C	exon 10	rs199583774	0.0001/18	uncertain	Conserved	1
10	1	DSG2	c.1912 G>A	g.29121188	G638K	exon 13	rs.201564919	0.0001/15	uncertain	Conserved	1
11	1	DSC2	c.1484 C>T	g.28660098	A495V	exon 10	NA	NA	NA	Conserved	1
12	2	DSC2	c.1609 G>T	g.28659867	A537S	exon 11	NA	NA	NA	Conserved	1

MAF: Minor Allele Frequency (from ExAC database); NA: data not available.

Thus, after this filtering step 12 variants were present in 12 (6%) healthy subjects.

6.1.2.3 In-silico pathogenicity prediction analysis

After these three filtering steps (position and functional consideration and allelic frequency and aminoacidic conservation analysis) the *in-silico* pathogenicity prediction for all the remaining 12 variants were taken in account as it is shown in Table 19. The three "*in-silico* pathogenicity predictors" used for the analysis were SIFT (Craig Venter Institute), MutationTaster (Berlin Medical University) and Polyphen-2 (Harvard University). 5 out of 12 filtered variants were predicted to be deleterious by SIFT, while MutationTaster predicted 10 out of 12 variants as disease causing. Polyphen-2 predicted probably pathogenic for 8/12 variants. Only 1 variants (DSC2 c.1609C>T) was recognized as non-pathogenic by each one of the three-different algorithm. *In-silico* predictors responses were not available due to its intronic localization for one variant (DSP c.273+5G>C). Splicing finder algorithms predict to change the donor site in a splicing site (MaxEnt: -100.0%; NNSPLICE: -97.7%; HSF: -14.6%).

Table 19. In-silico predictors' results for the filtered variants.

Nr.		Gene	cDNA	gDNA	aa	Exon/ Intron	dbSNP	MAF (ExAC)	Clinical significance	aa conservation	Subjects	In silico predictor		
												SIFT	Mutation Taster	Polyphen-2
1	1	JUP	c.427G>A	g.39925711	A143T	exon 3	rs375788626	0.0001/11	uncertain (dbSNP)	Moderately	1	Deleterious (score: 0)	Disease causing (p-value: 1)	Probably damaging (1.000)
2	1	DSP	c.273+5G>C	g.7556058		intron 2	rs.200473206	0.0004/39	pathogenic (HGMD)	-	1			
3	2	DSP	c.535G>A	g.7559571	G179S	exon 4	rs774829433	0.00002/2	NA	Conserved	1	Tolerated (score: 0.81)	Disease causing (p-value: 1)	Probably damaging (0.997)
4	3	DSP	c.3075A>C	g.7578786	E1025D	exon 22	NA	NA	NA	Conserved	1	Tolerated (score: 0.3)	Disease causing (p-value: 0.967)	Benign (0.003)
5	4	DSP	c.3388G>A	g.7579811	D1130N	exon 23	rs765769174	0.00002/3	NA	Moderately	1	Tolerated (score: 0.11)	Disease causing (p-value: 1)	Probably damaging (0.997)
6	5	DSP	c.3616T>A	g.7580039	L1206I	exon 23	rs151115778	0.00009/11	uncertain	Moderately	1	Tolerated (score: 0.32)	Disease causing (p-value: 0.943)	Benign (0.001)
7	1	PKP2	c.89T >C	g.33049577	L30P	exon 1	NA	NA	NA	Conserved	1	Deleterious (score: 0)	Disease causing (p-value: 1)	Probably damaging (1.000)
8	2	PKP2	c.989 G>T	g.33030825	G330V	exon 3	rs373041797	NA	NA	Moderately	1	Tolerated (score: 0.12)	Disease causing (p-value: 1)	Probably damaging (0.994)
9	3	PKP2	c.2083 C>T	g.32974352	R695C	exon 10	rs199583774	0.0001/18	uncertain	Conserved	1	Deleterious (score: 0.01)	Disease causing (p-value: 0.97)	Probably damaging (1.000)
10	1	DSG2	c.1912 G>A	g.29121188	G638K	exon 13	rs.201564919	0.0001/15	uncertain	Conserved	1	Deleterious (score: 0)	Disease causing (p-value: 1)	Probably damaging (1.000)
11	1	DSC2	c.1484 C>T	g.28660098	A495V	exon 10	NA	NA	NA	Conserved	1	Deleterious (score: 0)	Disease causing (p-value: 1)	Probably damaging (0.999)
12	2	DSC2	c.1609 G>T	g.28659867	A537S	exon 11	NA	NA	NA	Conserved	1	Tolerated (score: 0.81)	Polymorphism (p-value: 1)	Benign (0.035)

MAF: Minor Allele Frequency (from ExAC database); NA: data not available.

7 Discussion

Arrhythmogenic Cardiomyopathy is an inherited heart muscle disease that may result in arrhythmias and SCD, especially in the young and athletes [2, 6]. The main pathological feature of AC is the progressive loss of myocardial cells, which are replaced by fibrous and adipose tissue [2, 7]. Major and minor criteria were established and revised by an international Task Force, to be specific for AC diagnosis [14]. AC criteria consider cardiac morphology and function, tissue characterization, electric rhythm conduction and the presence of arrhythmic events as well as family history, including the identification of pathogenic mutations as a major diagnostic criterion. Although the first genetic findings were accomplished on the recessive form known as Naxos disease, AC is nowadays most commonly considered as an autosomal dominant trait with variable expressivity and age-dependent penetrance [7, 10, 204, 205].

7.1 Differential gene expression analysis in Arrhythmogenic Cardiomyopathy

In recent years, gene expression studies have been designed to elucidate which molecular signalling pathways are altered during the pathogenesis of the disease. Some authors highlighted the importance of sarcolemmal calcium channel regulation, apoptosis and adipogenesis in the AC pathogenesis [167]. Different expression patterns were identified in AC failing hearts (explanted) when compared with dilated cardiomyopathy (DCM) and non-failing hearts [99]. Specifically, no transcriptional regulation of desmosomal protein encoding genes was observed in the RV of AC patients but, in the LV PKP2 over-expression and DSC2 under-expression was observed, as well as altered transcription of genes involved in cell morphology definition, adhesion and WNT signalling pathway. Our data did not show differences in the desmosome genes expression both in RV and LV as well as in the animal models.

Another study based on large-scaled, quantitative proteomics based on proteomics and RNA-seq analysis of RV and LV myocardium of 4 AC and 4 DCM explanted hearts and compared them with normal hearts [175]. These authors identified as most significant change, in AC hearts, the activation of CCAAT-enhancer-binding proteins (C/EBPα) and similar changes in

the Peroxisome Proliferator Activated Receptor Gamma (PPARγ) and Fork-head Box A2 (FOXA2) expression. C/EBPα and PPARγ are important regulatory factors in the adipocytes fate determination and lipid differentiation, forming a positive feedback effect to promote lipid droplet deposition [176]. Along with the significant activation of C/EBPα, their proteomics findings revealed the overexpression in the RV of AC hearts of many lipogenesis proteins e.g., FASN, FABP4, etc. Specifically, FASN is regulated by the transcription factor of Sterol regulatory element-binding proteins and plays an important role in coordinating fatty acid uptake and storage [177].

However, both studies carried out expression analysis on failing hearts tissue from AC patients, which is an uncommon finding of the disease and is creating a bias regarding their findings. Thus, the only molecular pathways linked to the disease until now, both in humans and experimental models, remain: WNT/β-catenin pathway [73, 127], Hippo (YAP/TAZ) pathway [146, 147], and Rho/Rho-kinase pathway [152].

Our study aimed to depict three disease phases (i) the concealed phase prior myocyte structural changes, represented by TgNS<2wks mice, (ii) the overt electrical disorder phase represented by mice TgNS>3wks and (iii) the RV dysfunction phase represented by humans carriers of pathogenic mutations on desmosomal genes who previously underwent heart transplantation.

First, we focused our study on the transcriptional profile of AC patients which was compared to the counterpart of TgNS mice [78] in order to identify shared molecular pathways among humans and Tg animals carrying pathogenic desmosomal mutations; in this way we unbias our findings from environmental and epigenetic factors such as age and lifestyle, stage of the disease and genetic background. Indeed, the number of DEGs found in humans (1136 in the RV and 822 in the LV) was drastically reduced to 82 when we searched for similarities of the transcriptional profile of TgNS mice.

Further *in-silico* analysis of these 82 DEGs allowed to target 45 genes in different molecular signalling pathways among which WNT and TGF-β, previously associated with the disease. Specifically, for the WNT signalling pathway we found 7 genes under-expressed (16% of the 45 DEGS found in common between humans and mice) among which: *C-MYC, WNT4, WNT9B, SFRP2* and *SFRP5* whereas at least 6 of the 45 DEGs were found linked to TGF-β pathway respectively (13%) among these: *TGF-β2, BMP7, MSTN* and *TLL2*.

C-MYC. The protein is a basic helix loop helix transcription factor that at basal levels interacts with Max (another transcription factor) on actively transcribed genes via E-box sequences [206]. Its role in cellular reprogramming was initially attributed to its ability to promote cell proliferation and to activate a set of pluripotency genes and microRNAs through enhancement of proliferation-associated genes [207]. Recently, the role of C-MYC during transcription has been revisited, and it has been demonstrated that C-MYC does not regulate a unique set of target genes but rather acts as a general amplifier of gene expression, increasing the transcription at all active promoters [208]. C-MYC binding to promoter regions is associated with open chromatin marks including H3K4me3 and H3K27ac and is correlated with the amount of RNA polymerase recruited at those promoters [208]. It is a target gene of WNT/β-catenin pathway and regulates the expression of adipogenic related genes such as *C/Ebpα, Pparγ* and *Fabp4* [180, 209].

SFRP1, SFRP2 and SFRP5. These proteins are related to Frizzled receptor, although lack the transmembrane domain may compete with WNTs for the binding site on the Frizzled receptors inhibiting the canonical as well as non-canonical WNT pathways. Their role in cardiac fibrosis after myocardial infarction has been extensively demonstrated. As a matter of fact, SFRPs interfere with canonical WNT/β-catenin signalling and contribute to the prevention of aggravation of cardiac function [210].

WNT4, WNT9b. WNT gene family consists of 19 structurally related genes which encode secreted signalling proteins. These proteins have been implicated in oncogenesis and in several developmental processes, including regulation of cell fate and patterning during embryogenesis. Some members of the WNT genes (such as *WNT10b, WNT1* and *WNT5b*) have been reported to be able to inhibit or exacerbate adipogenesis through activation of the β-catenin pathway [211]. However, the precise mechanism to promote adipogenesis remains unknown. Interestingly, WNT4 negative regulates the expression of *MSTN* gene [212]. Our findings demonstrates that *WNT4* under-expression results in *MSTN* over-expression.

MSTN. Myostatin is also referred to as growth differentiation factor-8 (GDF-8) and belongs to the TGF-β superfamily. Differently to TGF-β, MSTN binds to Activin Type II receptor (ActRIIB), which initiates an intercellular signalling cascade leading to the phosphorylation and activation of Smad 2 and 3. MSTN is an important regulator of skeletal muscle development and adult homeostasis. Interaction of WNT signalling with MSTN during the

transition from cell proliferation to myogenic differentiation was found in mouse myoblast-derived cells (Terada K int J Cell boil 2013;2013:616294, TEE JM Plos One 2009;4:e5880) Naturally occurring mutations, as well as experimental knockout of the *MSTN* gene, lead to hyper muscular phenotype [213]. Later studies established that MSTN regulates the size and the number of muscle fibers by inhibiting myoblast proliferation and differentiation [214]. Finally, MSTN enhances nuclear translocation of β-catenin inhibiting adipogensis [215].

TGF-β2, BMP7. Further evidence of a cross-link between WNT and TGF-β signalling pathways is provided by the down regulation of TGF-β2 and BMP7 proteins, which are extracellular ligands of TGF-βR1 and TGF-βR2 receptors, involved in fibrogenesis and inflammatory processes. The binding of TGF-β to the receptors initiates specific intracellular signals mediated by Smad proteins, a family of 8 proteins that mediate signalling for BMPs proteins [216]. The TGF-β superfamily of growth factors consists of multiple proteins that govern a wide range of physiological processes, such as stem cell pluripotency, cell fate determination, proliferation, and differentiation. In humans, over 30 members of the TGF-β superfamily have been documented, including TGF-β and BMPs [216]. Selective activation of TGF-β family members was previously observed to distinguish fibrotic pathways in HCM and DCM mice. Mice with either DCM or HCM have increased expression of *Tgf-β2* and *Tgf-β3*. However, HCM mice also have a significant, 4-fold increase in *Tgf-β1* [217]. Our findings highlight the over-expression only of *TGF-β2* both in the mouse model and in the human samples.

Finally we performed a longitudinal transcriptome analysis in TgNS mice, to mark temporal changes in gene expression, In order to assess whether the inactivation of the WNT and the activation of TGF-β pathways in older mice and in human hearts are ongoing cellular processes of the disease such as fibrosis or are the trigger of this last.

In this setting, comparing the transcriptional profile of TgNS<2wk (concealed phase) with no observed changes at gross and histological level, to the age matched mice (both TgWt and Wt<2wks), we found 29 DEGs (19 under-expressed and 10 over-expressed genes).

Among these DEGS our study in TgNS <2wks mice highlighted the overexpression of an angiogenetic factor, Leucine Rich Alpha-2-Glycoprotein 1 (Lrg1), which modulates Tgf-β1 activity. Convincing evidence showed that a reduced expression of Lrg1 is associated with an increased Tgf-β1 activity during the cardiac remodelling processes in response to injury with

increased fibrosis and aberrant vascular properties. Indeed, Lrg1 can compete with Tgf-β1 for binding to the Tgf-βR2 receptor complex [179]. However, it is not clear whether TFG exerts its function by inhibiting the expression of LRG1 and what happens when LRG1 is overexpressed.

On the other hand, Lgals3 gene was found under-expressed in TgNS <2wks mice. This ubiquitary gene is the single gene coding for Galectin-3 protein in the human genome, a galectin family member of mammalian lectins. It is characterized by a carbohydrate recognition domain (CRD) and its affinity for β-galactosides. Galectin-3 mediates cell–cell and cell–matrix interactions by binding to lactosamine-containing cell surface glycoconjugates via its CRD. Higher concentration of Galectin-3, a marker of cardiac fibrosis, is associated with increased risk for incident HF and mortality [218, 219]. Some authors, have demonstrate that Galectin-3 binds to β-catenin and activates WNT signalling target genes such as cyclin D1 and c-Myc in breast cancer cells [220, 221]. Subsequently, other authors suggest that galectin-3 mediates WNT signalling, at least in part, by regulating GSK-3β phosphorylation and activity and, thus, the degradation of β-catenin in colon cancer cells [222]. Down-regulation of galectin-3 resulted in AKT and glycogen synthase kinase-3β (GSK-3β) de-phosphorylation and increased GSK-3β activity, increasing β-catenin phosphorylation and degradation and suppressing WNT/β-catenin signalling pathway causing the reduction of WNT target genes expression [222]. Recent studies have revealed the important role of GSK-3β in the regulation of cardiac function and remodelling via genetically modified animal models or pharmacological inhibition [223, 224]. They showed how the inhibition of GSK-3β has cardio-protective effect with a reduction of myocardial apoptosis, fibrosis, and LV dysfunction [223, 224]. The central role of GSK-3β in the complex phenotype of AC was demonstrated also by an experimental study in which researchers identified a small molecule (SB2) that rescues the disease phenotype inhibiting the activity of GSK-3 β [225, 226].

Mitogen activated protein kinase 4, *Mapk4* gene is under-expressed in TgNS <2wks mice. It encodes for a 65 KDa protein named Extracellular Signal-Regulated Kinase (ERK4) which along with ERK3 is part of an atypical subfamily of MAPKs. ERK4 is mainly localized in the cytoplasm. ERK4 interact with MAPK-Activated Protein Kinase 5 (MK5) and is responsible for its redistribution (from nucleus to cytoplasm) and activation [227]. MK5 is a regulator of cardiac fibroblast function but its exact role is not clear. The down-stream effector and

upstream activators of these signalling pathway remain to be identified [228]. Interestingly, in TgNS>3wks mice as well as in human samples, we found the under-expression of Dual Specificity Phosphatase 2 (*DUSP2*). This gene encodes for a protein that modulate the activity of ERK4 and its capability to bind to MK5 [228].

Further studies are needed to understand the role of these proteins in AC pathogenesis, but our findings demonstrated the interaction between WNT and TGF-β pathways at early disease stages, triggering cardiac remodelling. Specifically, we identified probably the 'culprit molecules' of disease onset.

7.2 Plakoglobin immune analysis as diagnostic test for AC

Since there is no "gold standard" to reach the diagnosis of AC, multiple categories of diagnostic information have been combined and the criteria recently updated, to improve diagnostic Se while maintaining Sp. Fibrofatty replacement of the myocardium constitutes a major diagnostic criterion detected by histochemistry on EMB [18]. However, morpho-functional ventricular wall abnormalities can be absent in the early disease stages or in the cardiac septum which is the usual site of EMB leading to false-negative results. In the past, Asimaki and colleagues have shown that JUP immunoreactive signal level reduction at IDs, independent of EMB site of cleavage, is a consistent feature in patients with AC, proposing this method as a highly sensitive diagnostic tool for disease diagnosis (Asimaki et al. 2009). Many other studies have been carried out to confirm these findings [112, 186-189, 191] nevertheless these studies are not homogeneous in terms of starting material and used techniques. These differences seem to highly impact the results. Indeed, Asimaki found a Se = 95% and a Sp = 90% with an Ab dilution of 1:50000 [92], while Munkholm showed lower data both for Se and Sp (85% and 57% respectively) [112]. Asimaki and Munkholm used two different diagnostic techniques: IPOX and IF respectively. Successively the authors published data obtained from another set of samples with an autopsy diagnosis of AC (40 cases) and 15 cases without heart disease as controls. In this case, no control samples showed reduced JUP staining, while the 68% of AC samples displayed the reduction [186]. The reduction of JUP IF signal level was observed on AC samples, but not in the controls, in a study conducted on both EMBs and autopsy specimens [187]. Subsequently, the ability of the test to distinguish among AC and other cardiomyopathies was proved [188]. This study, like the aforementioned, has as its limit the low number of analysed cases (9 AC samples and 7 controls) which makes it difficult to make robust statistical analysis. Noteworthy, each study listed above observed that the dilution of the primary Ab is a factor with a high impact on the test results: the greater the dilution the greater is the test capacity to distinguish between AC and the controls. As an example, with a dilution of 1:50.000, there were no differences between the two categories, but increasing the dilution to 1:250:000 the Se of the test was 85% and the Sp 57% [112]. Therefore, it is not surprising if some authors described no differences between AC and controls with a Se and Sp nearby 0% with an Ab dilution of 1:100 [189]. The dilution that returns the best test results is questioned: 1:50.000 [92], 1:75.000 [188], and 1:250.000 [112].

In this study we have used different dilutions and tests. We first executed an IPOX analysis on areas of structurally normal myocardium testing the capability of two JUP Ab reduction (1:50.000 and 1:250.000) on HS and EMB. We have demonstrated that with lower dilution the test Se increase in both HS and EMB. However, at the same time, test Sp dropped down. We found same results analysing HS and EMB with IF technique. We performed ROC analysis to define the usefulness of this test in diagnostic purpose. The AUC of a ROC analysis is a measure of diagnostic accuracy of the test. The percentage of AUC for all our test is comprise between $0.5<AUC\leq0.7$ defining a not accurate test [229].

7.3 Study of the genetic variants frequency in the Veneto region population

Desmosomal point mutations are detected in nearly half of the AC probands [40]. However, recent studies highlighted the presence of desmosomal gene variants in the general population, intensifying the debate regarding the pathogenic effects of missense and "radical" mutations. At present "radical" mutations are considered highly pathogenic due to abnormal protein length advocating a haploinsufficiency mechanism, whereas missense mutations are cautiously interpreted using complex bioinformatics algorithms. Moreover, the identification of compound and digenic heterozygous carriers, associated with a more severe form of the disease, increased the genetic complexity of the disease suggesting a recessive inheritance pattern [20, 21].

In our study genetic screening was conducted by standard sequencing with the use of DHPLC technique followed by Sanger sequencing. Disadvantages of DHPLC are the requirement and maintenance of a specialized and expensive equipment and the optimization of each reaction required to achieve the highest sensitivity for mutation detection, moreover it can also be difficult to identify homozygotes unless the sample is spiked with a known control [230].

For many years the use of the DHPLC technique together with Sanger sequencing was an absolute necessity, due to the lack of affordable and high-throughput methods. Today, NGS sequencing strategy made this technique obsolete, although this study began before the introduction of the NGS technique in our laboratory.

Summarizing, our analysis allowed us to detect 11 (10 missense and 1 that could alter splicing site) rare variants in desmosomal genes in 20 healthy subjects (10%) Intriguingly, in our study "no radical" (frameshift, nonsense) variant were found, even if a probably splicing variant was found in DSP. Furthermore, 2 healthy athletes harbour a missense variant on DSP classified as pathogenic variant due to in vitro function evidences proving its association with the disease in an AC cohort [75]. Our data are lower than values (16 and 18%) reported by others [200, 201] probably due to our selected cohort of healthy subjects. Indeed, our cohort was previously subjected to the pre-participation clinical evaluation.

Moreover, in our analysis we have used as variants filtering cut off a MAF of 0.0002 corresponding to a diseases frequency of 1:5000. Recently, Whiffin and colleagues [203] have

proposed a new approach to calculate the MAF cut off that considers not only the disease frequency but also the frequency of the most common causative variant and the disease penetrance. Using this approach we have lowered the MAF filtering cut off from 0.0002 to 0.000018. After this filtering step the only remaining variants were those with not available MAF (5; 2.5%).

Second level cardiac screening of mutation carriers and cascade family screening along with phenotype-genotype correlation are needed to determine the pathogenicity of these variants.

8 References.

1. Basso, C., Corrado, D., Marcus, F.I., Nava, A., and Thiene, G., *Arrhythmogenic right ventricular cardiomyopathy.* Lancet, 2009. **373**(9671): p. 1289-1300.
2. Thiene, G., Nava, A., Corrado, D., Rossi, L., and Pennelli, N., *Right ventricular cardiomyopathy and sudden death in young people.* N Engl J Med, 1988. **318**(3): p. 129-133.
3. Basso, C., Thiene, G., Corrado, D., Angelini, A., Nava, A., and Valente, M., *Arrhythmogenic right ventricular cardiomyopathy. Dysplasia, dystrophy, or myocarditis?* Circulation, 1996. **94**(5): p. 983-991.
4. Fontaine, G., Frank, R., Guiraudon, G., Pavie, A., Tereau, Y., Chomette, G., and Grosgogeat, Y., *[Significance of intraventricular conduction disorders observed in arrhythmogenic right ventricular dysplasia].* Arch Mal Coeur Vaiss, 1984. **77**(8): p. 872-879.
5. Nava, A., Bauce, B., Basso, C., Muriago, M., Rampazzo, A., Villanova, C., Daliento, L., Buja, G., Corrado, D., Danieli, G.A., and Thiene, G., *Clinical profile and long-term follow-up of 37 families with arrhythmogenic right ventricular cardiomyopathy.* J Am Coll Cardiol, 2000. **36**(7): p. 2226-2233.
6. Corrado, D., Thiene, G., Nava, A., Rossi, L., and Pennelli, N., *Sudden death in young competitive athletes: clinicopathologic correlations in 22 cases.* Am J Med, 1990. **89**(5): p. 588-596.
7. Nava, A., Thiene, G., Canciani, B., Scognamiglio, R., Daliento, L., Buja, G., Martini, B., Stritoni, P., and Fasoli, G., *Familial occurrence of right ventricular dysplasia: a study involving nine families.* J Am Coll Cardiol, 1988. **12**(5): p. 1222-1228.
8. Marcus, F.I., Fontaine, G.H., Guiraudon, G., Frank, R., Laurenceau, J.L., Malergue, C., and Grosgogeat, Y., *Right ventricular dysplasia: a report of 24 adult cases.* Circulation, 1982. **65**(2): p. 384-398.
9. Nava, A., Scognamiglio, R., Thiene, G., Canciani, B., Daliento, L., Buja, G., Stritoni, P., Fasoli, G., and Dalla Volta, S., *A polymorphic form of familial arrhythmogenic right ventricular dysplasia.* Am J Cardiol, 1987. **59**(15): p. 1405-1409.
10. McKenna, W.J., Thiene, G., Nava, A., Fontaliran, F., Blomstrom-Lundqvist, C., Fontaine, G., and Camerini, F., *Diagnosis of arrhythmogenic right ventricular dysplasia/cardiomyopathy. Task Force of the Working Group Myocardial and Pericardial Disease of the European Society of Cardiology and of the Scientific Council on Cardiomyopathies of the International Society and Federation of Cardiology.* Br Heart J, 1994. **71**(3): p. 215-218.
11. Corrado, D., Fontaine, G., Marcus, F.I., McKenna, W.J., Nava, A., Thiene, G., and Wichter, T., *Arrhythmogenic right ventricular dysplasia/cardiomyopathy: need for an international registry. European Society of Cardiology and the Scientific Council on Cardiomyopathies of the World Heart Federation.* J Cardiovasc Electrophysiol, 2000. **11**(7): p. 827-832.
12. Thiene, G., Corrado, D., and Basso, C., *Arrhythmogenic right ventricular cardiomyopathy/dysplasia.* Orphanet J Rare Dis, 2007. **2**: p. 45.
13. Sen-Chowdhry, S., Syrris, P., Prasad, S.K., Hughes, S.E., Merrifield, R., Ward, D., Pennell, D.J., and McKenna, W.J., *Left-dominant arrhythmogenic cardiomyopathy: an under-recognized clinical entity.* J Am Coll Cardiol, 2008. **52**(25): p. 2175-2187.
14. Marcus, F.I., McKenna, W.J., Sherrill, D., Basso, C., Bauce, B., Bluemke, D.A., Calkins, H., Corrado, D., Cox, M.G., Daubert, J.P., Fontaine, G., Gear, K., Hauer, R., Nava, A., Picard, M.H., Protonotarios, N., Saffitz, J.E., Sanborn, D.M., Steinberg, J.S., Tandri, H., Thiene, G., Towbin,

J.A., Tsatsopoulou, A., Wichter, T., and Zareba, W., *Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria.* Eur Heart J, 2010. **31**(7): p. 806-814.

15. Tandri, H., Saranathan, M., Rodriguez, E.R., Martinez, C., Bomma, C., Nasir, K., Rosen, B., Lima, J.A., Calkins, H., and Bluemke, D.A., *Noninvasive detection of myocardial fibrosis in arrhythmogenic right ventricular cardiomyopathy using delayed-enhancement magnetic resonance imaging.* J Am Coll Cardiol, 2005. **45**(1): p. 98-103.

16. Sen-Chowdhry, S., Prasad, S.K., Syrris, P., Wage, R., Ward, D., Merrifield, R., Smith, G.C., Firmin, D.N., Pennell, D.J., and McKenna, W.J., *Cardiovascular magnetic resonance in arrhythmogenic right ventricular cardiomyopathy revisited: comparison with task force criteria and genotype.* J Am Coll Cardiol, 2006. **48**(10): p. 2132-2140.

17. Perazzolo Marra, M., Thiene, G., Rizzo, S., De Lazzari, M., Carturan, E., Tona, F., Caforio, A.L., Cacciavillani, L., Marcolongo, R., Tarantini, G., Corbetti, F., Iliceto, S., and Basso, C., *Cardiac magnetic resonance features of biopsy-proven endomyocardial diseases.* JACC Cardiovasc Imaging, 2014. **7**(3): p. 309-312.

18. Basso, C., Ronco, F., Marcus, F., Abudureheman, A., Rizzo, S., Frigo, A.C., Bauce, B., Maddalena, F., Nava, A., Corrado, D., Grigoletto, F., and Thiene, G., *Quantitative assessment of endomyocardial biopsy in arrhythmogenic right ventricular cardiomyopathy/dysplasia: an in vitro validation of diagnostic criteria.* Eur Heart J, 2008. **29**(22): p. 2760-2771.

19. Thiene, G., *The research venture in arrhythmogenic right ventricular cardiomyopathy: a paradigm of translational medicine.* Eur Heart J, 2015. **36**(14): p. 837-846.

20. Bauce, B., Nava, A., Beffagna, G., Basso, C., Lorenzon, A., Smaniotto, G., De Bortoli, M., Rigato, I., Mazzotti, E., Steriotis, A., Marra, M.P., Towbin, J.A., Thiene, G., Danieli, G.A., and Rampazzo, A., *Multiple mutations in desmosomal proteins encoding genes in arrhythmogenic right ventricular cardiomyopathy/dysplasia.* Heart Rhythm, 2010. **7**(1): p. 22-29.

21. Rigato, I., Bauce, B., Rampazzo, A., Zorzi, A., Pilichou, K., Mazzotti, E., Migliore, F., Marra, M.P., Lorenzon, A., De Bortoli, M., Calore, M., Nava, A., Daliento, L., Gregori, D., Iliceto, S., Thiene, G., Basso, C., and Corrado, D., *Compound and digenic heterozygosity predicts lifetime arrhythmic outcome and sudden cardiac death in desmosomal gene-related arrhythmogenic right ventricular cardiomyopathy.* Circ Cardiovasc Genet, 2013. **6**(6): p. 533-542.

22. Alcalai, R., Metzger, S., Rosenheck, S., Meiner, V., and Chajek-Shaul, T., *A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder, and woolly hair.* J Am Coll Cardiol, 2003. **42**(2): p. 319-327.

23. Simpson, M.A., Mansour, S., Ahnood, D., Kalidas, K., Patton, M.A., McKenna, W.J., Behr, E.R., and Crosby, A.H., *Homozygous mutation of desmocollin-2 in arrhythmogenic right ventricular cardiomyopathy with mild palmoplantar keratoderma and woolly hair.* Cardiology, 2009. **113**(1): p. 28-34.

24. Cabral, R.M., Liu, L., Hogan, C., Dopping-Hepenstal, P.J., Winik, B.C., Asial, R.A., Dobson, R., Mein, C.A., Baselaga, P.A., Mellerio, J.E., Nanda, A., Boente Mdel, C., Kelsell, D.P., McGrath, J.A., and South, A.P., *Homozygous mutations in the 5' region of the JUP gene result in cutaneous disease but normal heart development in children.* J Invest Dermatol, 2010. **130**(6): p. 1543-1550.

25. Gerull, B., Kirchner, F., Chong, J.X., Tagoe, J., Chandrasekharan, K., Strohm, O., Waggoner, D., Ober, C., and Duff, H.J., *Homozygous founder mutation in desmocollin-2 (DSC2) causes arrhythmogenic cardiomyopathy in the Hutterite population.* Circ Cardiovasc Genet, 2013. **6**(4): p. 327-336.

26. Al-Sabeq, B., Krahn, A.D., Conacher, S., Klein, G.J., and Laksman, Z., *Arrhythmogenic right ventricular cardiomyopathy with recessive inheritance related to a new homozygous desmocollin-2 mutation.* Can J Cardiol, 2014. **30**(6): p. 696 e691-693.

27. Lorenzon, A., Pilichou, K., Rigato, I., Vazza, G., De Bortoli, M., Calore, M., Occhi, G., Carturan, E., Lazzarini, E., Cason, M., Mazzotti, E., Poloni, G., Mostacciuolo, M.L., Daliento, L., Thiene, G., Corrado, D., Basso, C., Bauce, B., and Rampazzo, A., *Homozygous Desmocollin-2 Mutations and Arrhythmogenic Cardiomyopathy.* Am J Cardiol, 2015. **116**(8): p. 1245-1251.

28. Qadri, S., Anttonen, O., Viikila, J., Seppala, E.H., Myllykangas, S., Alastalo, T.P., Holmstrom, M., Helio, T., and Koskenvuo, J.W., *Case reports of two pedigrees with recessive arrhythmogenic right ventricular cardiomyopathy associated with homozygous Thr335Ala variant in DSG2.* BMC Med Genet, 2017. **18**(1): p. 86.

29. McKoy, G., Protonotarios, N., Crosby, A., Tsatsopoulou, A., Anastasakis, A., Coonar, A., Norman, M., Baboonian, C., Jeffery, S., and McKenna, W.J., *Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease).* Lancet, 2000. **355**(9221): p. 2119-2124.

30. Rampazzo, A., Nava, A., Malacrida, S., Beffagna, G., Bauce, B., Rossi, V., Zimbello, R., Simionati, B., Basso, C., Thiene, G., Towbin, J.A., and Danieli, G.A., *Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy.* Am J Hum Genet, 2002. **71**(5): p. 1200-1206.

31. Protonotarios, N., Tsatsopoulou, A., Patsourakos, P., Alexopoulos, D., Gezerlis, P., Simitsis, S., and Scampardonis, G., *Cardiac abnormalities in familial palmoplantar keratosis.* Br Heart J, 1986. **56**(4): p. 321-326.

32. Coonar, A.S., Protonotarios, N., Tsatsopoulou, A., Needham, E.W., Houlston, R.S., Cliff, S., Otter, M.I., Murday, V.A., Mattu, R.K., and McKenna, W.J., *Gene for arrhythmogenic right ventricular cardiomyopathy with diffuse nonepidermolytic palmoplantar keratoderma and woolly hair (Naxos disease) maps to 17q21.* Circulation, 1998. **97**(20): p. 2049-2058.

33. Carvajal-Huerta, L., *Epidermolytic palmoplantar keratoderma with woolly hair and dilated cardiomyopathy.* J Am Acad Dermatol, 1998. **39**(3): p. 418-421.

34. Norgett, E.E., Hatsell, S.J., Carvajal-Huerta, L., Cabezas, J.C., Common, J., Purkis, P.E., Whittock, N., Leigh, I.M., Stevens, H.P., and Kelsell, D.P., *Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma.* Hum Mol Genet, 2000. **9**(18): p. 2761-2766.

35. Kaplan, S.R., Gard, J.J., Carvajal-Huerta, L., Ruiz-Cabezas, J.C., Thiene, G., and Saffitz, J.E., *Structural and molecular pathology of the heart in Carvajal syndrome.* Cardiovasc Pathol, 2004. **13**(1): p. 26-32.

36. Asimaki, A., Syrris, P., Wichter, T., Matthias, P., Saffitz, J.E., and McKenna, W.J., *A novel dominant mutation in plakoglobin causes arrhythmogenic right ventricular cardiomyopathy.* Am J Hum Genet, 2007. **81**(5): p. 964-973.

37. Gerull, B., Heuser, A., Wichter, T., Paul, M., Basson, C.T., McDermott, D.A., Lerman, B.B., Markowitz, S.M., Ellinor, P.T., MacRae, C.A., Peters, S., Grossmann, K.S., Drenckhahn, J., Michely, B., Sasse-Klaassen, S., Birchmeier, W., Dietz, R., Breithardt, G., Schulze-Bahr, E., and Thierfelder, L., *Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy.* Nat Genet, 2004. **36**(11): p. 1162-1164.

38. Pilichou, K., Nava, A., Basso, C., Beffagna, G., Bauce, B., Lorenzon, A., Frigo, G., Vettori, A., Valente, M., Towbin, J., Thiene, G., Danieli, G.A., and Rampazzo, A., *Mutations in desmoglein-*

2 gene are associated with arrhythmogenic right ventricular cardiomyopathy. Circulation, 2006. **113**(9): p. 1171-1179.

39. Syrris, P., Ward, D., Evans, A., Asimaki, A., Gandjbakhch, E., Sen-Chowdhry, S., and McKenna, W.J., *Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2.* Am J Hum Genet, 2006. **79**(5): p. 978-984.

40. Basso, C., Corrado, D., Bauce, B., and Thiene, G., *Arrhythmogenic right ventricular cardiomyopathy.* Circ Arrhythm Electrophysiol, 2012. **5**(6): p. 1233-1246.

41. Tiso, N., Stephan, D.A., Nava, A., Bagattin, A., Devaney, J.M., Stanchi, F., Larderet, G., Brahmbhatt, B., Brown, K., Bauce, B., Muriago, M., Basso, C., Thiene, G., Danieli, G.A., and Rampazzo, A., *Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2).* Hum Mol Genet, 2001. **10**(3): p. 189-194.

42. Beffagna, G., Occhi, G., Nava, A., Vitiello, L., Ditadi, A., Basso, C., Bauce, B., Carraro, G., Thiene, G., Towbin, J.A., Danieli, G.A., and Rampazzo, A., *Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1.* Cardiovasc Res, 2005. **65**(2): p. 366-373.

43. Merner, N.D., Hodgkinson, K.A., Haywood, A.F., Connors, S., French, V.M., Drenckhahn, J.D., Kupprion, C., Ramadanova, K., Thierfelder, L., McKenna, W., Gallagher, B., Morris-Larkin, L., Bassett, A.S., Parfrey, P.S., and Young, T.L., *Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene.* Am J Hum Genet, 2008. **82**(4): p. 809-821.

44. van Tintelen, J.P., Van Gelder, I.C., Asimaki, A., Suurmeijer, A.J., Wiesfeld, A.C., Jongbloed, J.D., van den Wijngaard, A., Kuks, J.B., van Spaendonck-Zwarts, K.Y., Notermans, N., Boven, L., van den Heuvel, F., Veenstra-Knol, H.E., Saffitz, J.E., Hofstra, R.M., and van den Berg, M.P., *Severe cardiac phenotype with right ventricular predominance in a large cohort of patients with a single missense mutation in the DES gene.* Heart Rhythm, 2009. **6**(11): p. 1574-1583.

45. van der Zwaag, P.A., van Rijsingen, I.A., Asimaki, A., Jongbloed, J.D., van Veldhuisen, D.J., Wiesfeld, A.C., Cox, M.G., van Lochem, L.T., de Boer, R.A., Hofstra, R.M., Christiaans, I., van Spaendonck-Zwarts, K.Y., Lekanne dit Deprez, R.H., Judge, D.P., Calkins, H., Suurmeijer, A.J., Hauer, R.N., Saffitz, J.E., Wilde, A.A., van den Berg, M.P., and van Tintelen, J.P., *Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy.* Eur J Heart Fail, 2012. **14**(11): p. 1199-1207.

46. Taylor, M., Graw, S., Sinagra, G., Barnes, C., Slavov, D., Brun, F., Pinamonti, B., Salcedo, E.E., Sauer, W., Pyxaras, S., Anderson, B., Simon, B., Bogomolovas, J., Labeit, S., Granzier, H., and Mestroni, L., *Genetic variation in titin in arrhythmogenic right ventricular cardiomyopathy-overlap syndromes.* Circulation, 2011. **124**(8): p. 876-885.

47. Quarta, G., Syrris, P., Ashworth, M., Jenkins, S., Zuborne Alapi, K., Morgan, J., Muir, A., Pantazis, A., McKenna, W.J., and Elliott, P.M., *Mutations in the Lamin A/C gene mimic arrhythmogenic right ventricular cardiomyopathy.* Eur Heart J, 2012. **33**(9): p. 1128-1136.

48. van Hengel, J., Calore, M., Bauce, B., Dazzo, E., Mazzotti, E., De Bortoli, M., Lorenzon, A., Li Mura, I.E., Beffagna, G., Rigato, I., Vleeschouwers, M., Tyberghein, K., Hulpiau, P., van Hamme, E., Zaglia, T., Corrado, D., Basso, C., Thiene, G., Daliento, L., Nava, A., van Roy, F., and Rampazzo, A., *Mutations in the area composita protein alphaT-catenin are associated with arrhythmogenic right ventricular cardiomyopathy.* Eur Heart J, 2013. **34**(3): p. 201-210.

49. Ortiz-Genga, M.F., Cuenca, S., Dal Ferro, M., Zorio, E., Salgado-Aranda, R., Climent, V., Padron-Barthe, L., Duro-Aguado, I., Jimenez-Jaimez, J., Hidalgo-Olivares, V.M., Garcia-Campo, E., Lanzillo, C., Suarez-Mier, M.P., Yonath, H., Marcos-Alonso, S., Ochoa, J.P., Santome, J.L., Garcia-Giustiniani, D., Rodriguez-Garrido, J.L., Dominguez, F., Merlo, M., Palomino, J., Pena, M.L., Trujillo, J.P., Martin-Vila, A., Stolfo, D., Molina, P., Lara-Pezzi, E., Calvo-Iglesias, F.E., Nof, E., Calo, L., Barriales-Villa, R., Gimeno-Blanes, J.R., Arad, M., Garcia-Pavia, P., and Monserrat, L., *Truncating FLNC Mutations Are Associated With High-Risk Dilated and Arrhythmogenic Cardiomyopathies.* J Am Coll Cardiol, 2016. **68**(22): p. 2440-2451.

50. Mayosi, B.M., Fish, M., Shaboodien, G., Mastantuono, E., Kraus, S., Wieland, T., Kotta, M.C., Chin, A., Laing, N., Ntusi, N.B., Chong, M., Horsfall, C., Pimstone, S.N., Gentilini, D., Parati, G., Strom, T.M., Meitinger, T., Pare, G., Schwartz, P.J., and Crotti, L., *Identification of Cadherin 2 (CDH2) Mutations in Arrhythmogenic Right Ventricular Cardiomyopathy.* Circ Cardiovasc Genet, 2017. **10**(2).

51. Merritt, A.J., Berika, M.Y., Zhai, W., Kirk, S.E., Ji, B., Hardman, M.J., and Garrod, D.R., *Suprabasal desmoglein 3 expression in the epidermis of transgenic mice results in hyperproliferation and abnormal differentiation.* Mol Cell Biol, 2002. **22**(16): p. 5846-5858.

52. Pilichou, K., Thiene, G., Bauce, B., Rigato, I., Lazzarini, E., Migliore, F., Perazzolo Marra, M., Rizzo, S., Zorzi, A., Daliento, L., Corrado, D., and Basso, C., *Arrhythmogenic cardiomyopathy.* Orphanet J Rare Dis, 2016. **11**: p. 33.

53. North, A.J., Bardsley, W.G., Hyam, J., Bornslaeger, E.A., Cordingley, H.C., Trinnaman, B., Hatzfeld, M., Green, K.J., Magee, A.I., and Garrod, D.R., *Molecular map of the desmosomal plaque.* J Cell Sci, 1999. **112 (Pt 23)**: p. 4325-4336.

54. Leung, C.L., Green, K.J., and Liem, R.K., *Plakins: a family of versatile cytolinker proteins.* Trends Cell Biol, 2002. **12**(1): p. 37-45.

55. Armstrong, D.K., McKenna, K.E., Purkis, P.E., Green, K.J., Eady, R.A., Leigh, I.M., and Hughes, A.E., *Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma.* Hum Mol Genet, 1999. **8**(1): p. 143-148.

56. Fressart, V., Duthoit, G., Donal, E., Probst, V., Deharo, J.C., Chevalier, P., Klug, D., Dubourg, O., Delacretaz, E., Cosnay, P., Scanu, P., Extramiana, F., Keller, D., Hidden-Lucet, F., Simon, F., Bessirard, V., Roux-Buisson, N., Hebert, J.L., Azarine, A., Casset-Senon, D., Rouzet, F., Lecarpentier, Y., Fontaine, G., Coirault, C., Frank, R., Hainque, B., and Charron, P., *Desmosomal gene analysis in arrhythmogenic right ventricular dysplasia/cardiomyopathy: spectrum of mutations and clinical impact in practice.* Europace, 2010. **12**(6): p. 861-868.

57. Cox, M.G., van der Zwaag, P.A., van der Werf, C., van der Smagt, J.J., Noorman, M., Bhuiyan, Z.A., Wiesfeld, A.C., Volders, P.G., van Langen, I.M., Atsma, D.E., Dooijes, D., van den Wijngaard, A., Houweling, A.C., Jongbloed, J.D., Jordaens, L., Cramer, M.J., Doevendans, P.A., de Bakker, J.M., Wilde, A.A., van Tintelen, J.P., and Hauer, R.N., *Arrhythmogenic right ventricular dysplasia/cardiomyopathy: pathogenic desmosome mutations in index-patients predict outcome of family screening: Dutch arrhythmogenic right ventricular dysplasia/cardiomyopathy genotype-phenotype follow-up study.* Circulation, 2011. **123**(23): p. 2690-2700.

58. Chen, X., Bonne, S., Hatzfeld, M., van Roy, F., and Green, K.J., *Protein binding and functional characterization of plakophilin 2. Evidence for its diverse roles in desmosomes and beta -catenin signaling.* J Biol Chem, 2002. **277**(12): p. 10512-10522.

59. Gandjbakhch, E., Charron, P., Fressart, V., Lorin de la Grandmaison, G., Simon, F., Gary, F., Vite, A., Hainque, B., Hidden-Lucet, F., Komajda, M., and Villard, E., *Plakophilin 2A is the*

dominant isoform in human heart tissue: consequences for the genetic screening of arrhythmogenic right ventricular cardiomyopathy. Heart, 2011. **97**(10): p. 844-849.

60. Grossmann, K.S., Grund, C., Huelsken, J., Behrend, M., Erdmann, B., Franke, W.W., and Birchmeier, W., *Requirement of plakophilin 2 for heart morphogenesis and cardiac junction formation.* J Cell Biol, 2004. **167**(1): p. 149-160.

61. Roberts, J.D., Herkert, J.C., Rutberg, J., Nikkel, S.M., Wiesfeld, A.C., Dooijes, D., Gow, R.M., van Tintelen, J.P., and Gollob, M.H., *Detection of genomic deletions of PKP2 in arrhythmogenic right ventricular cardiomyopathy.* Clin Genet, 2013. **83**(5): p. 452-456.

62. Li Mura, I.E., Bauce, B., Nava, A., Fanciulli, M., Vazza, G., Mazzotti, E., Rigato, I., De Bortoli, M., Beffagna, G., Lorenzon, A., Calore, M., Dazzo, E., Nobile, C., Mostacciuolo, M.L., Corrado, D., Basso, C., Daliento, L., Thiene, G., and Rampazzo, A., *Identification of a PKP2 gene deletion in a family with arrhythmogenic right ventricular cardiomyopathy.* Eur J Hum Genet, 2013. **21**(11): p. 1226-1231.

63. Sonoda, K., Ohno, S., Otuki, S., Kato, K., Yagihara, N., Watanabe, H., Makiyama, T., Minamino, T., and Horie, M., *Quantitative analysis of PKP2 and neighbouring genes in a patient with arrhythmogenic right ventricular cardiomyopathy caused by heterozygous PKP2 deletion.* Europace, 2017. **19**(4): p. 644-650.

64. Pilichou, K., Lazzarini, E., Rigato, I., Celeghin, R., De Bortoli, M., Perazzolo Marra, M., Cason, M., Jongbloed, J., Calore, M., Rizzo, S., Regazzo, D., Poloni, G., Iliceto, S., Daliento, L., Delise, P., Corrado, D., Van Tintelen, J.P., Thiene, G., Rampazzo, A., Basso, C., Bauce, B., Lorenzon, A., and Occhi, G., *Large Genomic Rearrangements of Desmosomal Genes in Italian Arrhythmogenic Cardiomyopathy Patients.* Circ Arrhythm Electrophysiol, 2017. **10**(10).

65. Green, K.J. and Simpson, C.L., *Desmosomes: new perspectives on a classic.* J Invest Dermatol, 2007. **127**(11): p. 2499-2515.

66. Schafer, S., Koch, P.J., and Franke, W.W., *Identification of the ubiquitous human desmoglein, Dsg2, and the expression catalogue of the desmoglein subfamily of desmosomal cadherins.* Exp Cell Res, 1994. **211**(2): p. 391-399.

67. Nuber, U.A., Schafer, S., Schmidt, A., Koch, P.J., and Franke, W.W., *The widespread human desmocollin Dsc2 and tissue-specific patterns of synthesis of various desmocollin subtypes.* Eur J Cell Biol, 1995. **66**(1): p. 69-74.

68. Awad, M.M., Dalal, D., Tichnell, C., James, C., Tucker, A., Abraham, T., Spevak, P.J., Calkins, H., and Judge, D.P., *Recessive arrhythmogenic right ventricular dysplasia due to novel cryptic splice mutation in PKP2.* Hum Mutat, 2006. **27**(11): p. 1157.

69. Heuser, A., Plovie, E.R., Ellinor, P.T., Grossmann, K.S., Shin, J.T., Wichter, T., Basson, C.T., Lerman, B.B., Sasse-Klaassen, S., Thierfelder, L., MacRae, C.A., and Gerull, B., *Mutant desmocollin-2 causes arrhythmogenic right ventricular cardiomyopathy.* Am J Hum Genet, 2006. **79**(6): p. 1081-1088.

70. Beffagna, G., De Bortoli, M., Nava, A., Salamon, M., Lorenzon, A., Zaccolo, M., Mancuso, L., Sigalotti, L., Bauce, B., Occhi, G., Basso, C., Lanfranchi, G., Towbin, J.A., Thiene, G., Danieli, G.A., and Rampazzo, A., *Missense mutations in desmocollin-2 N-terminus, associated with arrhythmogenic right ventricular cardiomyopathy, affect intracellular localization of desmocollin-2 in vitro.* BMC Med Genet, 2007. **8**: p. 65.

71. Padron-Barthe, L., Dominguez, F., Garcia-Pavia, P., and Lara-Pezzi, E., *Animal models of arrhythmogenic right ventricular cardiomyopathy: what have we learned and where do we go? Insight for therapeutics.* Basic Res Cardiol, 2017. **112**(5): p. 50.

72. Gallicano, G.I., Kouklis, P., Bauer, C., Yin, M., Vasioukhin, V., Degenstein, L., and Fuchs, E., *Desmoplakin is required early in development for assembly of desmosomes and cytoskeletal linkage.* J Cell Biol, 1998. **143**(7): p. 2009-2022.

73. Garcia-Gras, E., Lombardi, R., Giocondo, M.J., Willerson, J.T., Schneider, M.D., Khoury, D.S., and Marian, A.J., *Suppression of canonical Wnt/beta-catenin signaling by nuclear plakoglobin recapitulates phenotype of arrhythmogenic right ventricular cardiomyopathy.* J Clin Invest, 2006. **116**(7): p. 2012-2021.

74. Ruiz, P., Brinkmann, V., Ledermann, B., Behrend, M., Grund, C., Thalhammer, C., Vogel, F., Birchmeier, C., Gunthert, U., Franke, W.W., and Birchmeier, W., *Targeted mutation of plakoglobin in mice reveals essential functions of desmosomes in the embryonic heart.* J Cell Biol, 1996. **135**(1): p. 215-225.

75. Yang, Z., Bowles, N.E., Scherer, S.E., Taylor, M.D., Kearney, D.L., Ge, S., Nadvoretskiy, V.V., DeFreitas, G., Carabello, B., Brandon, L.I., Godsel, L.M., Green, K.J., Saffitz, J.E., Li, H., Danieli, G.A., Calkins, H., Marcus, F., and Towbin, J.A., *Desmosomal dysfunction due to mutations in desmoplakin causes arrhythmogenic right ventricular dysplasia/cardiomyopathy.* Circ Res, 2006. **99**(6): p. 646-655.

76. Gomes, J., Finlay, M., Ahmed, A.K., Ciaccio, E.J., Asimaki, A., Saffitz, J.E., Quarta, G., Nobles, M., Syrris, P., Chaubey, S., McKenna, W.J., Tinker, A., and Lambiase, P.D., *Electrophysiological abnormalities precede overt structural changes in arrhythmogenic right ventricular cardiomyopathy due to mutations in desmoplakin-A combined murine and human study.* Eur Heart J, 2012. **33**(15): p. 1942-1953.

77. Lyon, R.C., Mezzano, V., Wright, A.T., Pfeiffer, E., Chuang, J., Banares, K., Castaneda, A., Ouyang, K., Cui, L., Contu, R., Gu, Y., Evans, S.M., Omens, J.H., Peterson, K.L., McCulloch, A.D., and Sheikh, F., *Connexin defects underlie arrhythmogenic right ventricular cardiomyopathy in a novel mouse model.* Hum Mol Genet, 2014. **23**(5): p. 1134-1150.

78. Pilichou, K., Remme, C.A., Basso, C., Campian, M.E., Rizzo, S., Barnett, P., Scicluna, B.P., Bauce, B., van den Hoff, M.J., de Bakker, J.M., Tan, H.L., Valente, M., Nava, A., Wilde, A.A., Moorman, A.F., Thiene, G., and Bezzina, C.R., *Myocyte necrosis underlies progressive myocardial dystrophy in mouse dsg2-related arrhythmogenic right ventricular cardiomyopathy.* J Exp Med, 2009. **206**(8): p. 1787-1802.

79. Cerrone, M., Noorman, M., Lin, X., Chkourko, H., Liang, F.X., van der Nagel, R., Hund, T., Birchmeier, W., Mohler, P., van Veen, T.A., van Rijen, H.V., and Delmar, M., *Sodium current deficit and arrhythmogenesis in a murine model of plakophilin-2 haploinsufficiency.* Cardiovasc Res, 2012. **95**(4): p. 460-468.

80. Moriarty, M.A., Ryan, R., Lalor, P., Dockery, P., Byrnes, L., and Grealy, M., *Loss of plakophilin 2 disrupts heart development in zebrafish.* Int J Dev Biol, 2012. **56**(9): p. 711-718.

81. Mazurek, S.R., Calway, T., Harmon, C., Farrell, P., and Kim, G.H., *MicroRNA-130a Regulation of Desmocollin 2 in a Novel Model of Arrhythmogenic Cardiomyopathy.* Microrna, 2016.

82. Brodehl, A., Belke, D.D., Garnett, L., Martens, K., Abdelfatah, N., Rodriguez, M., Diao, C., Chen, Y.X., Gordon, P.M., Nygren, A., and Gerull, B., *Transgenic mice overexpressing desmocollin-2 (DSC2) develop cardiomyopathy associated with myocardial inflammation and fibrotic remodeling.* PLoS One, 2017. **12**(3): p. e0174019.

83. Bierkamp, C., McLaughlin, K.J., Schwarz, H., Huber, O., and Kemler, R., *Embryonic heart and skin defects in mice lacking plakoglobin.* Dev Biol, 1996. **180**(2): p. 780-785.

84. Li, J., Swope, D., Raess, N., Cheng, L., Muller, E.J., and Radice, G.L., *Cardiac tissue-restricted deletion of plakoglobin results in progressive cardiomyopathy and activation of {beta}-catenin signaling.* Mol Cell Biol, 2011. **31**(6): p. 1134-1144.

85. Martin, E.D., Moriarty, M.A., Byrnes, L., and Grealy, M., *Plakoglobin has both structural and signalling roles in zebrafish development.* Dev Biol, 2009. **327**(1): p. 83-96.

86. Bierkamp, C., Schwarz, H., Huber, O., and Kemler, R., *Desmosomal localization of beta-catenin in the skin of plakoglobin null-mutant mice.* Development, 1999. **126**(2): p. 371-381.

87. Zhang, Z., Stroud, M.J., Zhang, J., Fang, X., Ouyang, K., Kimura, K., Mu, Y., Dalton, N.D., Gu, Y., Bradford, W.H., Peterson, K.L., Cheng, H., Zhou, X., and Chen, J., *Normalization of Naxos plakoglobin levels restores cardiac function in mice.* J Clin Invest, 2015. **125**(4): p. 1708-1712.

88. Eshkind, L., Tian, Q., Schmidt, A., Franke, W.W., Windoffer, R., and Leube, R.E., *Loss of desmoglein 2 suggests essential functions for early embryonic development and proliferation of embryonal stem cells.* Eur J Cell Biol, 2002. **81**(11): p. 592-598.

89. Rizzo, S., Lodder, E.M., Verkerk, A.O., Wolswinkel, R., Beekman, L., Pilichou, K., Basso, C., Remme, C.A., Thiene, G., and Bezzina, C.R., *Intercalated disc abnormalities, reduced Na(+) current density, and conduction slowing in desmoglein-2 mutant mice prior to cardiomyopathic changes.* Cardiovasc Res, 2012. **95**(4): p. 409-418.

90. Asano, Y., Takashima, S., Asakura, M., Shintani, Y., Liao, Y., Minamino, T., Asanuma, H., Sanada, S., Kim, J., Ogai, A., Fukushima, T., Oikawa, Y., Okazaki, Y., Kaneda, Y., Sato, M., Miyazaki, J., Kitamura, S., Tomoike, H., Kitakaze, M., and Hori, M., *Lamr1 functional retroposon causes right ventricular dysplasia in mice.* Nat Genet, 2004. **36**(2): p. 123-130.

91. Meurs, K.M., Mauceli, E., Lahmers, S., Acland, G.M., White, S.N., and Lindblad-Toh, K., *Genome-wide association identifies a deletion in the 3' untranslated region of striatin in a canine model of arrhythmogenic right ventricular cardiomyopathy.* Hum Genet, 2010. **128**(3): p. 315-324.

92. Asimaki, A., Tandri, H., Huang, H., Halushka, M.K., Gautam, S., Basso, C., Thiene, G., Tsatsopoulou, A., Protonotarios, N., McKenna, W.J., Calkins, H., and Saffitz, J.E., *A new diagnostic test for arrhythmogenic right ventricular cardiomyopathy.* N Engl J Med, 2009. **360**(11): p. 1075-1084.

93. Mullis, K.B., *The unusual origin of the polymerase chain reaction.* Sci Am, 1990. **262**(4): p. 56-61, 64-55.

94. Xiao, W. and Oefner, P.J., *Denaturing high-performance liquid chromatography: A review.* Hum Mutat, 2001. **17**(6): p. 439-474.

95. Sanger, F., Nicklen, S., and Coulson, A.R., *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-5467.

96. Adzhubei, I., Jordan, D.M., and Sunyaev, S.R., *Predicting functional effect of human missense mutations using PolyPhen-2.* Curr Protoc Hum Genet, 2013. **Chapter 7**: p. Unit7 20.

97. Kumar, P., Henikoff, S., and Ng, P.C., *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protoc, 2009. **4**(7): p. 1073-1081.

98. Schwarz, J.M., Cooper, D.N., Schuelke, M., and Seelow, D., *MutationTaster2: mutation prediction for the deep-sequencing age.* Nat Methods, 2014. **11**(4): p. 361-362.

99. Gaertner, A., Schwientek, P., Ellinghaus, P., Summer, H., Golz, S., Kassner, A., Schulz, U., Gummert, J., and Milting, H., *Myocardial transcriptome analysis of human arrhythmogenic right ventricular cardiomyopathy.* Physiol Genomics, 2012. **44**(1): p. 99-109.

100. Rajan, S., Pena, J.R., Jegga, A.G., Aronow, B.J., Wolska, B.M., and Wieczorek, D.F., *Microarray analysis of active cardiac remodeling genes in a familial hypertrophic cardiomyopathy mouse model rescued by a phospholamban knockout.* Physiol Genomics, 2013. **45**(17): p. 764-773.

101. Wood, D.L., Nones, K., Steptoe, A., Christ, A., Harliwong, I., Newell, F., Bruxner, T.J., Miller, D., Cloonan, N., and Grimmond, S.M., *Recommendations for Accurate Resolution of Gene and Isoform Allele-Specific Expression in RNA-Seq Data.* PLoS One, 2015. **10**(5): p. e0126911.

102. Wolf, J.B., *Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial.* Mol Ecol Resour, 2013. **13**(4): p. 559-572.

103. Fonseca, N.A., Marioni, J., and Brazma, A., *RNA-Seq gene profiling--a systematic empirical comparison.* PLoS One, 2014. **9**(9): p. e107026.

104. Xiao, X.G., Touma, M., and Wang, Y., *Decoding the noncoding transcripts in human heart failure.* Circulation, 2014. **129**(9): p. 958-960.

105. Love, M.I., Huber, W., and Anders, S., *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

106. Robinson, M.D., McCarthy, D.J., and Smyth, G.K., *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-140.

107. Mi, H.Y., Muruganujan, A., Casagrande, J.T., and Thomas, P.D., *Large-scale gene function analysis with the PANTHER classification system.* Nature Protocols, 2013. **8**(8): p. 1551-1566.

108. Richterich, P., *Estimation of errors in "raw" DNA sequences: A validation study.* Genome Research, 1998. **8**(3): p. 251-259.

109. Robinson, M.D. and Smyth, G.K., *Small-sample estimation of negative binomial dispersion, with applications to SAGE data.* Biostatistics, 2008. **9**(2): p. 321-332.

110. Sacco, P.A., McGranahan, T.M., Wheelock, M.J., and Johnson, K.R., *Identification of plakoglobin domains required for association with N-cadherin and alpha-catenin.* J Biol Chem, 1995. **270**(34): p. 20201-20206.

111. Wahl, J.K., Sacco, P.A., McGranahan-Sadler, T.M., Sauppe, L.M., Wheelock, M.J., and Johnson, K.R., *Plakoglobin domains that define its association with the desmosomal cadherins and the classical cadherins: identification of unique and shared domains.* J Cell Sci, 1996. **109 (Pt 5)**: p. 1143-1154.

112. Munkholm, J., Christensen, A.H., Svendsen, J.H., and Andersen, C.B., *Usefulness of immunostaining for plakoglobin as a diagnostic marker of arrhythmogenic right ventricular cardiomyopathy.* Am J Cardiol, 2012. **109**(2): p. 272-275.

113. Clancy, S. and Brown, W., *Translation: DNA to mRNA to Protein.* Nature Education, 2008. **1**(1): p. 7.

114. Chapeville, F., Lipmann, F., Von Ehrenstein, G., Weisblum, B., Ray, W.J., Jr., and Benzer, S., *On the role of soluble ribonucleic acid in coding for amino acids.* Proc Natl Acad Sci U S A, 1962. **48**: p. 1086-1092.

115. Grunberger, D., Weinstein, I.B., and Jacobson, K.B., *Codon recognition by enzymatically mischarged valine transfer ribonucleic acid.* Science, 1969. **166**(3913): p. 1635-1637.

116. Crick, F.H., *The genetic code--yesterday, today, and tomorrow.* Cold Spring Harb Symp Quant Biol, 1966. **31**: p. 1-9.

117. Flinta, C., Persson, B., Jornvall, H., and von Heijne, G., *Sequence determinants of cytosolic N-terminal protein processing.* Eur J Biochem, 1986. **154**(1): p. 193-196.

118. Uhl, H.S., *A previously undescribed congenital malformation of the heart: almost total absence of the myocardium of the right ventricle.* Bull Johns Hopkins Hosp, 1952. **91**(3): p. 197-209.

119. Dokuparti, M.V., Pamuru, P.R., Thakkar, B., Tanjore, R.R., and Nallari, P., *Etiopathogenesis of arrhythmogenic right ventricular cardiomyopathy.* J Hum Genet, 2005. **50**(8): p. 375-381.

120. Daliento, L., Turrini, P., Nava, A., Rizzoli, G., Angelini, A., Buja, G., Scognamiglio, R., and Thiene, G., *Arrhythmogenic right ventricular cardiomyopathy in young versus adult patients: similarities and differences.* J Am Coll Cardiol, 1995. **25**(3): p. 655-664.

121. Mallat, Z., Tedgui, A., Fontaliran, F., Frank, R., Durigon, M., and Fontaine, G., *Evidence of apoptosis in arrhythmogenic right ventricular dysplasia.* N Engl J Med, 1996. **335**(16): p. 1190-1196.

122. Valente, M., Calabrese, F., Thiene, G., Angelini, A., Basso, C., Nava, A., and Rossi, L., *In vivo evidence of apoptosis in arrhythmogenic right ventricular cardiomyopathy.* Am J Pathol, 1998. **152**(2): p. 479-484.

123. Bowles, N.E., Ni, J., Marcus, F., and Towbin, J.A., *The detection of cardiotropic viruses in the myocardium of patients with arrhythmogenic right ventricular dysplasia/cardiomyopathy.* J Am Coll Cardiol, 2002. **39**(5): p. 892-895.

124. Thiene, G., Corrado, D., Nava, A., Rossi, L., Poletti, A., Boffa, G.M., Daliento, L., and Pennelli, N., *Right ventricular cardiomyopathy: is there evidence of an inflammatory aetiology?* Eur Heart J, 1991. **12 Suppl D**: p. 22-25.

125. Calabrese, F., Basso, C., Carturan, E., Valente, M., and Thiene, G., *Arrhythmogenic right ventricular cardiomyopathy/dysplasia: is there a role for viruses?* Cardiovasc Pathol, 2006. **15**(1): p. 11-17.

126. d'Amati, G., di Gioia, C.R., Giordano, C., and Gallo, P., *Myocyte transdifferentiation: a possible pathogenetic mechanism for arrhythmogenic right ventricular cardiomyopathy.* Arch Pathol Lab Med, 2000. **124**(2): p. 287-290.

127. Lombardi, R., Dong, J., Rodriguez, G., Bell, A., Leung, T.K., Schwartz, R.J., Willerson, J.T., Brugada, R., and Marian, A.J., *Genetic fate mapping identifies second heart field progenitor cells as a source of adipocytes in arrhythmogenic right ventricular cardiomyopathy.* Circ Res, 2009. **104**(9): p. 1076-1084.

128. Lombardi, R., da Graca Cabreira-Hansen, M., Bell, A., Fromm, R.R., Willerson, J.T., and Marian, A.J., *Nuclear plakoglobin is essential for differentiation of cardiac progenitor cells to adipocytes in arrhythmogenic right ventricular cardiomyopathy.* Circ Res, 2011. **109**(12): p. 1342-1353.

129. Basso, C., Bauce, B., Corrado, D., and Thiene, G., *Pathophysiology of arrhythmogenic cardiomyopathy.* Nat Rev Cardiol, 2011. **9**(4): p. 223-233.

130. Delmar, M. and McKenna, W.J., *The cardiac desmosome and arrhythmogenic cardiomyopathies: from gene to disease.* Circ Res, 2010. **107**(6): p. 700-714.

131. Kono, H. and Rock, K.L., *How dying cells alert the immune system to danger.* Nat Rev Immunol, 2008. **8**(4): p. 279-289.

132. Rock, K.L., Latz, E., Ontiveros, F., and Kono, H., *The sterile inflammatory response.* Annu Rev Immunol, 2010. **28**: p. 321-342.

133. Janeway, C.A., Jr. and Medzhitov, R., *Introduction: the role of innate immunity in the adaptive immune response.* Semin Immunol, 1998. **10**(5): p. 349-350.

134. Takeuchi, O. and Akira, S., *Pattern recognition receptors and inflammation.* Cell, 2010. **140**(6): p. 805-820.

135. Zhang, W., Lavine, K.J., Epelman, S., Evans, S.A., Weinheimer, C.J., Barger, P.M., and Mann, D.L., *Necrotic myocardial cells release damage-associated molecular patterns that provoke fibroblast activation in vitro and trigger myocardial inflammation and fibrosis in vivo.* J Am Heart Assoc, 2015. **4**(6): p. e001993.

136. Franke, W.W., Borrmann, C.M., Grund, C., and Pieperhoff, S., *The area composita of adhering junctions connecting heart muscle cells of vertebrates. I. Molecular definition in intercalated disks of cardiomyocytes by immunoelectron microscopy of desmosomal proteins.* Eur J Cell Biol, 2006. **85**(2): p. 69-82.

137. Patel, D., Gemel, J., Xu, Q., Simon, A.R., Lin, X., Matiukas, A., Beyer, E.C., and Veenstra, R.D., *Atrial fibrillation-associated connexin40 mutants make hemichannels and synergistically form gap junction channels with novel properties.* FEBS Lett, 2014. **588**(8): p. 1458-1464.

138. Kleber, A.G. and Saffitz, J.E., *Role of the intercalated disc in cardiac propagation and arrhythmogenesis.* Front Physiol, 2014. **5**: p. 404.

139. Asimaki, A., Kleber, A.G., and Saffitz, J.E., *Pathogenesis of Arrhythmogenic Cardiomyopathy.* Can J Cardiol, 2015. **31**(11): p. 1313-1324.

140. Swope, D., Li, J., and Radice, G.L., *Beyond cell adhesion: the role of armadillo proteins in the heart.* Cell Signal, 2013. **25**(1): p. 93-100.

141. Chen, X., Shevtsov, S.P., Hsich, E., Cui, L., Haq, S., Aronovitz, M., Kerkela, R., Molkentin, J.D., Liao, R., Salomon, R.N., Patten, R., and Force, T., *The beta-catenin/T-cell factor/lymphocyte enhancer factor signaling pathway is required for normal and stress-induced cardiac hypertrophy.* Mol Cell Biol, 2006. **26**(12): p. 4462-4473.

142. Haq, S., Michael, A., Andreucci, M., Bhattacharya, K., Dotto, P., Walters, B., Woodgett, J., Kilter, H., and Force, T., *Stabilization of beta-catenin by a Wnt-independent mechanism regulates cardiomyocyte growth.* Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4610-4615.

143. Maeda, O., Usami, N., Kondo, M., Takahashi, M., Goto, H., Shimokata, K., Kusugami, K., and Sekido, Y., *Plakoglobin (gamma-catenin) has TCF/LEF family-dependent transcriptional activity in beta-catenin-deficient cell line.* Oncogene, 2004. **23**(4): p. 964-972.

144. Miravet, S., Piedra, J., Miro, F., Itarte, E., Garcia de Herreros, A., and Dunach, M., *The transcriptional factor Tcf-4 contains different binding sites for beta-catenin and plakoglobin.* J Biol Chem, 2002. **277**(3): p. 1884-1891.

145. Zhurinsky, J., Shtutman, M., and Ben-Ze'ev, A., *Plakoglobin and beta-catenin: protein interactions, regulation and biological roles.* J Cell Sci, 2000. **113 (Pt 18)**: p. 3127-3139.

146. Heallen, T., Zhang, M., Wang, J., Bonilla-Claudio, M., Klysik, E., Johnson, R.L., and Martin, J.F., *Hippo pathway inhibits Wnt signaling to restrain cardiomyocyte proliferation and heart size.* Science, 2011. **332**(6028): p. 458-461.

147. Chen, S.N., Gurha, P., Lombardi, R., Ruggiero, A., Willerson, J.T., and Marian, A.J., *The hippo pathway is activated and is a causal mechanism for adipogenesis in arrhythmogenic cardiomyopathy.* Circ Res, 2014. **114**(3): p. 454-468.

148. Pan, D., *The hippo signaling pathway in development and cancer.* Dev Cell, 2010. **19**(4): p. 491-505.

149. Satoh, K., Fukumoto, Y., and Shimokawa, H., *Rho-kinase: important new therapeutic target in cardiovascular diseases.* Am J Physiol Heart Circ Physiol, 2011. **301**(2): p. H287-296.

150. Cristancho, A.G. and Lazar, M.A., *Forming functional fat: a growing understanding of adipocyte differentiation.* Nat Rev Mol Cell Biol, 2011. **12**(11): p. 722-734.

151. Schlessinger, K., Hall, A., and Tolwinski, N., *Wnt signaling pathways meet Rho GTPases.* Genes Dev, 2009. **23**(3): p. 265-277.

152. Ellawindy, A., Satoh, K., Sunamura, S., Kikuchi, N., Suzuki, K., Minami, T., Ikeda, S., Tanaka, S., Shimizu, T., Enkhjargal, B., Miyata, S., Taguchi, Y., Handoh, T., Kobayashi, K., Kobayashi, K., Nakayama, K., Miura, M., and Shimokawa, H., *Rho-Kinase Inhibition During Early Cardiac Development Causes Arrhythmogenic Right Ventricular Cardiomyopathy in Mice.* Arterioscler Thromb Vasc Biol, 2015. **35**(10): p. 2172-2184.

153. Godsel, L.M., Dubash, A.D., Bass-Zubek, A.E., Amargo, E.V., Klessner, J.L., Hobbs, R.P., Chen, X., and Green, K.J., *Plakophilin 2 couples actomyosin remodeling to desmosomal plaque assembly via RhoA.* Mol Biol Cell, 2010. **21**(16): p. 2844-2859.

154. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J. and International Human Genome Sequencing, C., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

155. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A. and Zhu, X., *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-1351.

156. Metzker, M.L., *Sequencing technologies - the next generation.* Nat Rev Genet, 2010. **11**(1): p. 31-46.

157. Churko, J.M., Mantalas, G.L., Snyder, M.P., and Wu, J.C., *Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases.* Circ Res, 2013. **112**(12): p. 1613-1623.

158. Matkovich, S.J., Zhang, Y., Van Booven, D.J., and Dorn, G.W., 2nd, *Deep mRNA sequencing for in vivo functional analysis of cardiac transcriptional regulators: application to Galphaq.* Circ Res, 2010. **106**(9): p. 1459-1467.

159. Labaj, P.P., Leparc, G.G., Linggi, B.E., Markillie, L.M., Wiley, H.S., and Kreil, D.P., *Characterization and improvement of RNA-Seq precision in quantitative transcript expression profiling.* Bioinformatics, 2011. **27**(13): p. i383-391.

160. Hoeijmakers, W.A., Bartfai, R., and Stunnenberg, H.G., *Transcriptome analysis using RNA-Seq.* Methods Mol Biol, 2013. **923**: p. 221-239.

161. Lee, J.H., Gao, C., Peng, G., Greer, C., Ren, S., Wang, Y., and Xiao, X., *Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts.* Circ Res, 2011. **109**(12): p. 1332-1341.

162. Song, H.K., Hong, S.E., Kim, T., and Kim, D.H., *Deep RNA sequencing reveals novel cardiac transcriptomic signatures for physiological and pathological hypertrophy.* PLoS One, 2012. **7**(4): p. e35552.

163. Hu, Y., Matkovich, S.J., Hecker, P.A., Zhang, Y., Edwards, J.R., and Dorn, G.W., 2nd, *Epitranscriptional orchestration of genetic reprogramming is an emergent property of stress-regulated cardiac microRNAs.* Proc Natl Acad Sci U S A, 2012. **109**(48): p. 19864-19869.

164. Shen, T., Aneas, I., Sakabe, N., Dirschinger, R.J., Wang, G., Smemo, S., Westlund, J.M., Cheng, H., Dalton, N., Gu, Y., Boogerd, C.J., Cai, C.L., Peterson, K., Chen, J., Nobrega, M.A., and Evans, S.M., *Tbx20 regulates a genetic program essential to adult mouse cardiomyocyte function.* J Clin Invest, 2011. **121**(12): p. 4640-4654.

165. Sakabe, N.J., Aneas, I., Shen, T., Shokri, L., Park, S.Y., Bulyk, M.L., Evans, S.M., and Nobrega, M.A., *Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function.* Hum Mol Genet, 2012. **21**(10): p. 2194-2204.

166. Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., Erwin, G., Kattman, S.J., Keller, G.M., Srivastava, D., Levine, S.S., Pollard, K.S., Holloway, A.K., Boyer, L.A., and Bruneau, B.G., *Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage.* Cell, 2012. **151**(1): p. 206-220.

167. Akdis, D., Medeiros-Domingo, A., Gaertner-Rommel, A., Kast, J.I., Enseleit, F., Bode, P., Klingel, K., Kandolf, R., Renois, F., Andreoletti, L., Akdis, C.A., Milting, H., Luscher, T.F., Brunckhorst, C., Saguner, A.M., and Duru, F., *Myocardial expression profiles of candidate molecules in patients with arrhythmogenic right ventricular cardiomyopathy/dysplasia compared to those with dilated cardiomyopathy and healthy controls.* Heart Rhythm, 2016. **13**(3): p. 731-741.

168. Barandon, L., Dufourcq, P., Costet, P., Moreau, C., Allieres, C., Daret, D., Dos Santos, P., Daniel Lamaziere, J.M., Couffinhal, T., and Duplaa, C., *Involvement of FrzA/sFRP-1 and the Wnt/frizzled pathway in ischemic preconditioning.* Circ Res, 2005. **96**(12): p. 1299-1306.

169. Kuwahara, K., Teg Pipes, G.C., McAnally, J., Richardson, J.A., Hill, J.A., Bassel-Duby, R., and Olson, E.N., *Modulation of adverse cardiac remodeling by STARS, a mediator of MEF2 signaling and SRF activity.* J Clin Invest, 2007. **117**(5): p. 1324-1334.

170. Radicke, S., Cotella, D., Graf, E.M., Banse, U., Jost, N., Varro, A., Tseng, G.N., Ravens, U., and Wettwer, E., *Functional modulation of the transient outward current Ito by KCNE beta-subunits and regional distribution in human non-failing and failing hearts.* Cardiovasc Res, 2006. **71**(4): p. 695-703.

171. Saegusa, J., Akakura, N., Wu, C.Y., Hoogland, C., Ma, Z., Lam, K.S., Liu, F.T., Takada, Y.K., and Takada, Y., *Pro-inflammatory secretory phospholipase A2 type IIA binds to integrins alphavbeta3 and alpha4beta1 and induces proliferation of monocytic cells in an integrin-dependent manner.* J Biol Chem, 2008. **283**(38): p. 26107-26115.

172. Thomsen, M.B., Wang, C., Ozgen, N., Wang, H.G., Rosen, M.R., and Pitt, G.S., *Accessory subunit KChIP2 modulates the cardiac L-type calcium current.* Circ Res, 2009. **104**(12): p. 1382-1389.

173. Yang, J., Moravec, C.S., Sussman, M.A., DiPaola, N.R., Fu, D., Hawthorn, L., Mitchell, C.A., Young, J.B., Francis, G.S., McCarthy, P.M., and Bond, M., *Decreased SLIM1 expression and*

increased gelsolin expression in failing human hearts measured by high-density oligonucleotide arrays. Circulation, 2000. **102**(25): p. 3046-3052.

174. Hayes, N.V., Scott, C., Heerkens, E., Ohanian, V., Maggs, A.M., Pinder, J.C., Kordeli, E., and Baines, A.J., *Identification of a novel C-terminal variant of beta II spectrin: two isoforms of beta II spectrin have distinct intracellular locations and activities.* J Cell Sci, 2000. **113 (Pt 11)**: p. 2023-2034.

175. Chen, L., Yang, F., Chen, X., Rao, M., Zhang, N.N., Chen, K., Deng, H., Song, J.P., and Hu, S.S., *Comprehensive Myocardial Proteogenomics Profiling Reveals C/EBPalpha as the Key Factor in the Lipid Storage of ARVC.* J Proteome Res, 2017. **16**(8): p. 2863-2876.

176. Ali, A.T., Hochfeld, W.E., Myburgh, R., and Pepper, M.S., *Adipocyte and adipogenesis.* Eur J Cell Biol, 2013. **92**(6-7): p. 229-236.

177. Schoonjans, K., Martin, G., Staels, B., and Auwerx, J., *Peroxisome proliferator-activated receptors, orphans with ligands and functions.* Curr Opin Lipidol, 1997. **8**(3): p. 159-166.

178. Vincent, A., Sportouch, C., Covinhes, A., Barrere, C., Gallot, L., Delgado-Betancourt, V., Lattuca, B., Solecki, K., Boisguerin, P., Piot, C., Nargeot, J., and Barrere-Lemaire, S., *Cardiac mGluR1 metabotropic receptors in cardioprotection.* Cardiovasc Res, 2017. **113**(6): p. 644-655.

179. Song, W. and Wang, X., *The role of TGFbeta1 and LRG1 in cardiac remodelling and heart failure.* Biophys Rev, 2015. **7**(1): p. 91-104.

180. Buganim, Y., Faddah, D.A., and Jaenisch, R., *Mechanisms and models of somatic cell reprogramming.* Nat Rev Genet, 2013. **14**(6): p. 427-439.

181. Kanazawa, A., Tsukada, S., Sekine, A., Tsunoda, T., Takahashi, A., Kashiwagi, A., Tanaka, Y., Babazono, T., Matsuda, M., Kaku, K., Iwamoto, Y., Kawamori, R., Kikkawa, R., Nakamura, Y., and Maeda, S., *Association of the gene encoding wingless-type mammary tumor virus integration-site family member 5B (WNT5B) with type 2 diabetes.* Am J Hum Genet, 2004. **75**(5): p. 832-843.

182. Lin, H., Angeli, M., Chung, K.J., Ejimadu, C., Rosa, A.R., and Lee, T., *sFRP2 activates Wnt/beta-catenin signaling in cardiac fibroblasts: differential roles in cell growth, energy metabolism, and extracellular matrix remodeling.* Am J Physiol Cell Physiol, 2016. **311**(5): p. C710-C719.

183. Biesemann, N., Mendler, L., Kostin, S., Wietelmann, A., Borchardt, T., and Braun, T., *Myostatin induces interstitial fibrosis in the heart via TAK1 and p38.* Cell Tissue Res, 2015. **361**(3): p. 779-787.

184. Corrado, D., Basso, C., Pilichou, K., and Thiene, G., *Molecular biology and clinical management of arrhythmogenic right ventricular cardiomyopathy/dysplasia.* Heart, 2011. **97**(7): p. 530-539.

185. Kaplan, S.R., Gard, J.J., Protonotarios, N., Tsatsopoulou, A., Spiliopoulou, C., Anastasakis, A., Squarcioni, C.P., McKenna, W.J., Thiene, G., Basso, C., Brousse, N., Fontaine, G., and Saffitz, J.E., *Remodeling of myocyte gap junctions in arrhythmogenic right ventricular cardiomyopathy due to a deletion in plakoglobin (Naxos disease).* Heart Rhythm, 2004. **1**(1): p. 3-11.

186. Munkholm, J., Andersen, C.B., and Ottesen, G.L., *Plakoglobin: a diagnostic marker of arrhythmogenic right ventricular cardiomyopathy in forensic pathology?* Forensic Sci Med Pathol, 2015. **11**(1): p. 47-52.

187. Noorman, M., Hakim, S., Asimaki, A., Vreeker, A., van Rijen, H.V., van der Heyden, M.A., de Jonge, N., de Weger, R.A., Hauer, R.N., Saffitz, J.E., and van Veen, T.A., *Reduced plakoglobin*

immunoreactivity in arrhythmogenic cardiomyopathy: methodological considerations. Cardiovasc Pathol, 2013. **22**(5): p. 314-318.

188. Ermakov, S., Ursell, P.C., Johnson, C.J., Meadows, A., Zhao, S., Marcus, G.M., and Scheinman, M., *Plakoglobin immunolocalization as a diagnostic test for arrhythmogenic right ventricular cardiomyopathy.* Pacing Clin Electrophysiol, 2014. **37**(12): p. 1708-1716.

189. Tavora, F., Zhang, M., Cresswell, N., Li, L., Fowler, D., Franco, M., and Burke, A., *Quantitative Immunohistochemistry of Desmosomal Proteins (Plakoglobin, Desmoplakin and Plakophilin), Connexin-43, and N-cadherin in Arrhythmogenic Cardiomyopathy: An Autopsy Study.* Open Cardiovasc Med J, 2013. **7**: p. 28-35.

190. Krusche, C.A., Holthofer, B., Hofe, V., van de Sandt, A.M., Eshkind, L., Bockamp, E., Merx, M.W., Kant, S., Windoffer, R., and Leube, R.E., *Desmoglein 2 mutant mice develop cardiac fibrosis and dilation.* Basic Res Cardiol, 2011. **106**(4): p. 617-633.

191. Kant, S., Krusche, C.A., Gaertner, A., Milting, H., and Leube, R.E., *Loss of plakoglobin immunoreactivity in intercalated discs in arrhythmogenic right ventricular cardiomyopathy: protein mislocalization versus epitope masking.* Cardiovasc Res, 2016. **109**(2): p. 260-271.

192. Swets, J.A., *Measuring the Accuracy of Diagnostic Systems.* Science, 1988. **240**(4857): p. 1285-1293.

193. Richards, C.S., Bale, S., Bellissimo, D.B., Das, S., Grody, W.W., Hegde, M.R., Lyon, E., Ward, B.E., and Molecular Subcommittee of the, A.L.Q.A.C., *ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007.* Genet Med, 2008. **10**(4): p. 294-300.

194. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L., and Committee, A.L.Q.A., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-424.

195. den Dunnen, J.T., Dalgleish, R., Maglott, D.R., Hart, R.K., Greenblatt, M.S., McGowan-Jordan, J., Roux, A.F., Smith, T., Antonarakis, S.E., and Taschner, P.E., *HGVS Recommendations for the Description of Sequence Variants: 2016 Update.* Hum Mutat, 2016. **37**(6): p. 564-569.

196. Stenson, P.D., Mort, M., Ball, E.V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A.D., and Cooper, D.N., *The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies.* Hum Genet, 2017. **136**(6): p. 665-677.

197. Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-Salomon, R., Rubinstein, W., and Maglott, D.R., *ClinVar: public archive of interpretations of clinically relevant variants.* Nucleic Acids Res, 2016. **44**(D1): p. D862-868.

198. Fokkema, I.F., Taschner, P.E., Schaafsma, G.C., Celli, J., Laros, J.F., and den Dunnen, J.T., *LOVD v.2.0: the next generation in gene variant databases.* Hum Mutat, 2011. **32**(5): p. 557-563.

199. Lazzarini, E., Jongbloed, J.D., Pilichou, K., Thiene, G., Basso, C., Bikker, H., Charbon, B., Swertz, M., van Tintelen, J.P., and van der Zwaag, P.A., *The ARVD/C genetic variants database: 2014 update.* Hum Mutat, 2015. **36**(4): p. 403-410.

200. Kapplinger, J.D., Landstrom, A.P., Salisbury, B.A., Callis, T.E., Pollevick, G.D., Tester, D.J., Cox, M.G., Bhuiyan, Z., Bikker, H., Wiesfeld, A.C., Hauer, R.N., van Tintelen, J.P., Jongbloed, J.D., Calkins, H., Judge, D.P., Wilde, A.A., and Ackerman, M.J., *Distinguishing arrhythmogenic right*

ventricular cardiomyopathy/dysplasia-associated mutations from background genetic noise. J Am Coll Cardiol, 2011. **57**(23): p. 2317-2327.

201. Andreasen, C., Nielsen, J.B., Refsgaard, L., Holst, A.G., Christensen, A.H., Andreasen, L., Sajadieh, A., Haunso, S., Svendsen, J.H., and Olesen, M.S., *New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants.* Eur J Hum Genet, 2013. **21**(9): p. 918-928.

202. De Bortoli, M., Beffagna, G., Bauce, B., Lorenzon, A., Smaniotto, G., Rigato, I., Calore, M., Li Mura, I.E., Basso, C., Thiene, G., Lanfranchi, G., Danieli, G.A., Nava, A., and Rampazzo, A., *The p.A897KfsX4 frameshift variation in desmocollin-2 is not a causative mutation in arrhythmogenic right ventricular cardiomyopathy.* Eur J Hum Genet, 2010. **18**(7): p. 776-782.

203. Whiffin, N., Minikel, E., Walsh, R., O'Donnell-Luria, A.H., Karczewski, K., Ing, A.Y., Barton, P.J.R., Funke, B., Cook, S.A., MacArthur, D., and Ware, J.S., *Using high-resolution variant frequencies to empower clinical genome interpretation.* Genet Med, 2017. **19**(10): p. 1151-1158.

204. Sen-Chowdhry, S., Lowe, M.D., Sporton, S.C., and McKenna, W.J., *Arrhythmogenic right ventricular cardiomyopathy: clinical presentation, diagnosis, and management.* Am J Med, 2004. **117**(9): p. 685-695.

205. van Tintelen, J.P., Entius, M.M., Bhuiyan, Z.A., Jongbloed, R., Wiesfeld, A.C., Wilde, A.A., van der Smagt, J., Boven, L.G., Mannens, M.M., van Langen, I.M., Hofstra, R.M., Otterspoor, L.C., Doevendans, P.A., Rodriguez, L.M., van Gelder, I.C., and Hauer, R.N., *Plakophilin-2 mutations are the major determinant of familial arrhythmogenic right ventricular dysplasia/cardiomyopathy.* Circulation, 2006. **113**(13): p. 1650-1658.

206. Blackwood, E.M. and Eisenman, R.N., *Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc.* Science, 1991. **251**(4998): p. 1211-1217.

207. Meyer, N. and Penn, L.Z., *Reflecting on 25 years with MYC.* Nat Rev Cancer, 2008. **8**(12): p. 976-990.

208. Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., Zhao, K., and Levens, D., *c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells.* Cell, 2012. **151**(1): p. 68-79.

209. van de Schans, V.A., Smits, J.F., and Blankesteijn, W.M., *The Wnt/frizzled pathway in cardiovascular development and disease: friend or foe?* Eur J Pharmacol, 2008. **585**(2-3): p. 338-345.

210. Tao, H., Yang, J.J., Shi, K.H., and Li, J., *Wnt signaling pathway in cardiac fibrosis: New insights and directions.* Metabolism, 2016. **65**(2): p. 30-40.

211. Ross, S.E., Hemati, N., Longo, K.A., Bennett, C.N., Lucas, P.C., Erickson, R.L., and MacDougald, O.A., *Inhibition of adipogenesis by Wnt signaling.* Science, 2000. **289**(5481): p. 950-953.

212. Bernardi, H., Gay, S., Fedon, Y., Vernus, B., Bonnieu, A., and Bacou, F., *Wnt4 activates the canonical beta-catenin pathway and regulates negatively myostatin: functional implication in myogenesis.* Am J Physiol Cell Physiol, 2011. **300**(5): p. C1122-1138.

213. Williams, M.S., *Myostatin mutation associated with gross muscle hypertrophy in a child.* N Engl J Med, 2004. **351**(10): p. 1030-1031; author reply 1030-1031.

214. Trendelenburg, A.U., Meyer, A., Rohner, D., Boyle, J., Hatakeyama, S., and Glass, D.J., *Myostatin reduces Akt/TORC1/p70S6K signaling, inhibiting myoblast differentiation and myotube size.* Am J Physiol Cell Physiol, 2009. **296**(6): p. C1258-1270.

215. Guo, W., Flanagan, J., Jasuja, R., Kirkland, J., Jiang, L., and Bhasin, S., *The effects of myostatin on adipogenic differentiation of human bone marrow-derived mesenchymal stem cells are mediated through cross-communication between Smad3 and Wnt/beta-catenin signaling pathways.* J Biol Chem, 2008. **283**(14): p. 9136-9145.

216. Piersma, B., Bank, R.A., and Boersema, M., *Signaling in Fibrosis: TGF-beta, WNT, and YAP/TAZ Converge.* Front Med (Lausanne), 2015. **2**: p. 59.

217. Burke, M.A., Chang, S., Wakimoto, H., Gorham, J.M., Conner, D.A., Christodoulou, D.C., Parfenov, M.G., DePalma, S.R., Eminaga, S., Konno, T., Seidman, J.G., and Seidman, C.E., *Molecular profiling of dilated cardiomyopathy that progresses to heart failure.* JCI Insight, 2016. **1**(6).

218. Shah, R.V., Chen-Tournoux, A.A., Picard, M.H., van Kimmenade, R.R., and Januzzi, J.L., *Galectin-3, cardiac structure and function, and long-term mortality in patients with acutely decompensated heart failure.* Eur J Heart Fail, 2010. **12**(8): p. 826-832.

219. Ho, J.E., Liu, C., Lyass, A., Courchesne, P., Pencina, M.J., Vasan, R.S., Larson, M.G., and Levy, D., *Galectin-3, a marker of cardiac fibrosis, predicts incident heart failure in the community.* J Am Coll Cardiol, 2012. **60**(14): p. 1249-1256.

220. Shimura, T., Takenaka, Y., Tsutsumi, S., Hogan, V., Kikuchi, A., and Raz, A., *Galectin-3, a novel binding partner of beta-catenin.* Cancer Res, 2004. **64**(18): p. 6363-6367.

221. Shimura, T., Takenaka, Y., Fukumori, T., Tsutsumi, S., Okada, K., Hogan, V., Kikuchi, A., Kuwano, H., and Raz, A., *Implication of galectin-3 in Wnt signaling.* Cancer Res, 2005. **65**(9): p. 3535-3537.

222. Song, S., Mazurek, N., Liu, C., Sun, Y., Ding, Q.Q., Liu, K., Hung, M.C., and Bresalier, R.S., *Galectin-3 mediates nuclear beta-catenin accumulation and Wnt signaling in human colon cancer cells by regulation of glycogen synthase kinase-3beta activity.* Cancer Res, 2009. **69**(4): p. 1343-1349.

223. Lal, H., Ahmad, F., Woodgett, J., and Force, T., *The GSK-3 family as therapeutic target for myocardial diseases.* Circ Res, 2015. **116**(1): p. 138-149.

224. Hirotani, S., Zhai, P., Tomita, H., Galeotti, J., Marquez, J.P., Gao, S., Hong, C., Yatani, A., Avila, J., and Sadoshima, J., *Inhibition of glycogen synthase kinase 3beta during heart failure is protective.* Circ Res, 2007. **101**(11): p. 1164-1174.

225. Asimaki, A., Kapoor, S., Plovie, E., Karin Arndt, A., Adams, E., Liu, Z., James, C.A., Judge, D.P., Calkins, H., Churko, J., Wu, J.C., MacRae, C.A., Kleber, A.G., and Saffitz, J.E., *Identification of a new modulator of the intercalated disc in a zebrafish model of arrhythmogenic cardiomyopathy.* Sci Transl Med, 2014. **6**(240): p. 240ra274.

226. Chelko, S.P., Asimaki, A., Andersen, P., Bedja, D., Amat-Alarcon, N., DeMazumder, D., Jasti, R., MacRae, C.A., Leber, R., Kleber, A.G., Saffitz, J.E., and Judge, D.P., *Central role for GSK3beta in the pathogenesis of arrhythmogenic cardiomyopathy.* JCI Insight, 2016. **1**(5).

227. Aberg, E., Perander, M., Johansen, B., Julien, C., Meloche, S., Keyse, S.M., and Seternes, O.M., *Regulation of MAPK-activated protein kinase 5 activity and subcellular localization by the atypical MAPK ERK4/MAPK4.* J Biol Chem, 2006. **281**(46): p. 35499-35510.

228. Sahadevan, P. and Allen, B.G., *MK5: A novel regulator of cardiac fibroblast function?* IUBMB Life, 2017. **69**(10): p. 785-794.

229. Swets, J.A., *Measuring the accuracy of diagnostic systems.* Science, 1988. **240**(4857): p. 1285-1293.

230. Oefner, P.J. and Underhill, P.A., *Comparative DNA-Sequencing by Denaturing High-Performance Liquid-Chromatography (Dhplc).* American Journal of Human Genetics, 1995. **57**(4): p. 1547-1547.

9 Appendice A: abbreviations

Ab: antibody;

AC: Arrhythmogenic Cardiomyopathy;

Bp: base pair;

cDNA: complementary DNA;

dsDNA: double strand DNA;

DEGs: Differencially Expressed Genes;

EMB: endomyocardial biopsy;

HS: heart Specimens;

ICD: implantable cardioverter-defibrillator;

IDs: intercalated disks;

IF: Immunofluorescence;

IHC: Immunohistochemistry;

IPOX: immunoperoxidase;

LV: left ventricle/left ventricular;

NGS: Next Generation Sequencing;

NPV: Negative Predictive Value;

PCR: Polymerase Chain Reaction;

PPV: Positive Predictive Value;

qPCR: real time quantitative PCR;

RV: right ventricle/right ventricular;

RNA-Seq: RNA sequencing with NGS;

SCD: Sudden Cardiac Death;

Se: Sensitivity;

SNP: Single Nucleotide Variant;

Sp: Specificity;

ssDNA: single strand DNA;

Ta: annealing temperature;

TgNS: murine model over-expressing N271S-dsg2;

Tm: melting temperature;

TgWt: murine model over-expressing Wt-dsg2;

Wt: wild type.