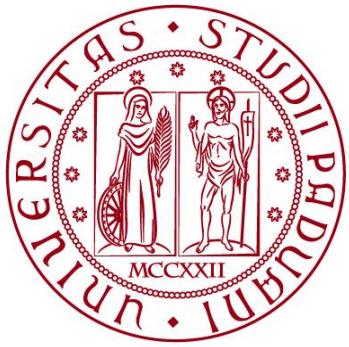
UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea magistrale in Biologia Sanitaria



TESI DI LAUREA

COMPARISON OF PROTOCOLS TO GENERATE INDUCED PLURIPOTENT STEM CELL-DERIVED CARDIOMYOCYTES AND VALIDATION OF A RNA-SEQ DATASET IN *IN VIVO* MODELS OF ARRHYTHMOGENIC CARDIOMYOPATHY.

Relatore: Prof.ssa Alessandra Rampazzo
Dipartimento di Biologia

Correlatore: Dott.ssa Martina Calore
Dipartimento di Biologia

Laureanda: Aurora Badalini

ANNO ACCADEMICO 2022/2023

TABLE OF CONTENTS

ABSTRACT

Arrhythmogenic cardiomyopathy (ACM) is a rare genetic disease, characterized by ventricular arrhythmias and myocardial fibrofatty replacement. Mechanisms underling its pathogenesis are still unknown, however, useful novel insights can arise integrating human *in vitro* and murine *in vivo* models. In this context, reproducibility and comparability of the data are crucial. Human induced pluripotent stem cell (hiPSCs) technology is a valuable source for *in vitro* models preserving the patient-specific background. In the first part of this thesis, two protocols for differentiation into hiPSC-derived cardiomyocytes were compared, providing information regarding their robustness, and guiding the choice of which to use. The results showed validity for both, with higher beating properties for the protocol previously used by our group than for the protocol used in Goettingen Medical Centre.

The second part of this thesis aimed gaining new insights in ACM pathogenesis through an RNA-seq dataset validation from mouse heart samples carrying the p.Q563* DSG2 mutation associated with ACM. Despite only 3 of the 27 analysed genes were validated (*Dsg2, Des, Tmem88*), differences between three- and six-month-old mice were detected in mitochondria, apoptosis and MAPK pathway related genes, suggesting possible changes underlining pathogenesis temporal evolution, that could be further investigated in the future.

INTRODUCTION

1.1 CARDIOMYOPATHIES

Cardiomyopathy is defined as a familiar, non-ischemic, heart muscle disease, typically characterized by altered electrical and muscular function and properties, with common manifestations ranging from microscope changes in cardiac myocytes to rapid heart failure with inadequate tissue perfusion, fluid accumulation, and cardiac rhythm dysfunction [1], [2]. The American Heart Association has proposed a classification system that divides cardiomyopathies into primary (Figure 1), if confined to the heart, and secondary, if cardiac involvement is a feature of a systemic condition. Specifically, several etiologies can be identified for primary cardiomyopathies: genetics, acquired or mixed. In the former, genetic abnormalities affect the heart, in the second, non-genetic causes lead to cardiac complications, and in the third, both genetic and non-genetic defects contribute to the pathogenic phenotype [1]. The classification of primary cardiomyopathies is illustrated in Figure 1. It is important to mention, however, that pathologies with a known and specific cardiovascular cause are not included in the definition of cardiomyopathy; examples are: hypertension, valvular disease, congenital heart disease, and coronary ischemia [1].

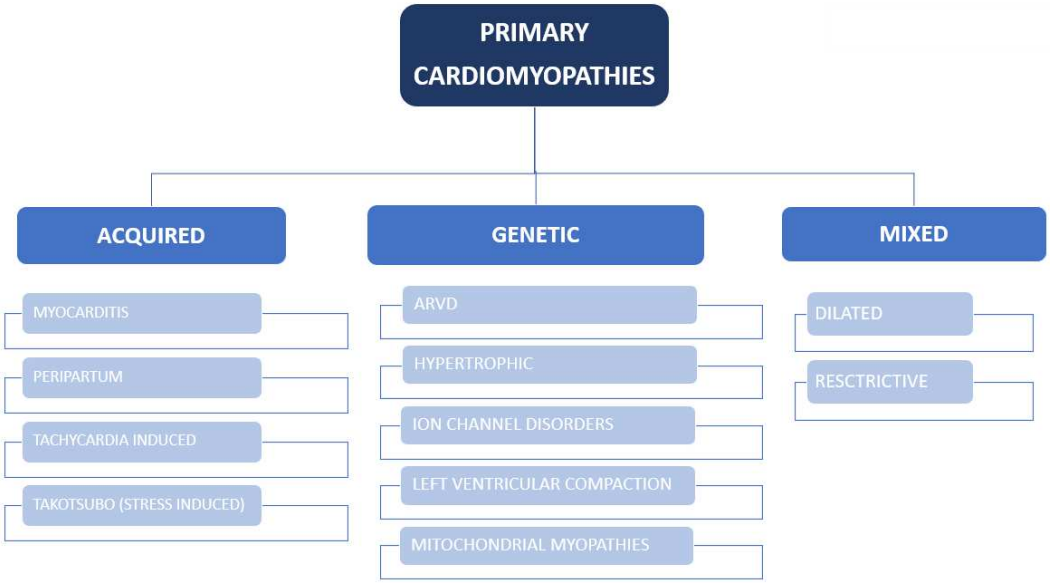


Figure 1 Primary cardiomyopathies classification. ARVD, arrhythmogenic right ventricular dysplasia. Adaptation from [1].

1.2 ARRHYTHMOGENIC CARDIOMYOPATHY (ACM)

Arrhythmogenic cardiomyopathy (ACM), also known as arrhythmogenic right ventricular dysplasia, is a rare genetic heart muscle disease, classified as a primary cardiomyopathy. Its main pathological feature is progressive myocardial dystrophy with fibro-fatty replacement of the ventricular myocardium leading to impairment of ventricular systolic functions, and its main clinical features are prominent

ventricular arrhythmias, which is why it is described as a “structural” [2] and “electrical” [4] cardiomyopathy. It represents one of the leading causes of sudden cardiac death in young athletes. Several studies correlated intense exercise with acceleration and increase of ACM pathogenesis in mutation carriers. The theory behind these studies relies on the fact that intercellular junctions are subjected to high mechanical stress, which sustains their decline [4].

The most frequent cause of ACM is mutation in desmosomal genes, affecting about 50% of cases [5], but non-desmosomal mutation have also been described. Therefore, it is often classified as an intercellular junction pathology [5]. Among the pathogenetic theories under investigation, altered intracellular signalling and consequent myocyte death with fibro-fatty replacement, abnormal cell-cell adhesion, ion channel and gap junction alteration are the main ones. [6]

This disease was originally described as a Venetian disease in 1982 , characterized by the replacement of myocardium in the right ventricle (RV) with fibrofatty tissue. Its classification as cardiomyopathy rather than dysplasia was based on the founding of its genetic basis and not presence at birth. Using the phenotype characterization, it is nowadays clear that the disease spectrum is wider than previously thought, particularly for the left ventricular involvement and prevalence at early stages, which lead to the reclassification of the condition as arrhythmogenic cardiomyopathy (ACM) [6]

1.2.1 EPIDEMIOLOGY OF ACM

The estimated prevalence of arrhythmogenic cardiomyopathy is within the range from 1:2000 to 1:5000. Given that sudden death may represents the only presentation of the disease, this could be an underestimation of the real frequency [6]. The condition exhibits a higher incidence in males, with a ratio of 3 to 1 compared to females. Interestingly, the severity of the disease is also more pronounced in males, even though the carriers of the genetic mutations are evenly distributed between the two genders. The potential protective role of sex hormones like estrogen in females is the most widely accepted of the existing theories for these sex-based differences [6].

ACM global presence has been demonstrated, although historically it has been notable for its high occurrence in the northeastern region of Italy, particularly in Veneto, due to the fact that ACM was discovered in this area, so many studies have been conducted [2]. Groups of individuals with the same genetic mutations have been identified in the Netherlands and other parts of Europe. Cases have also been reported in Asia and the Americas. A syndromic variant known as the Carvajal syndrome, characterized by woolly hair and palmoplantar keratoderma, has been observed in South America, resulting from a homozygous mutation in the *DSP* gene [6]. Another syndromic form, named after the Greek island Naxos, is caused by a homozygous mutation in the *JUP* gene [6]

The disease typically begins to manifest between the second and fourth decades of life, but there are cases in which it may emerge before puberty or even after the age of 70. In the latter case, arrhythmogenic cardiomyopathy (ACM) could be misdiagnosed due to the fibrofatty substitution being potentially mistaken for natural adipose tissue, which is more common in the elderly population. The variability in the presentation of symptoms, linked to the age at which the disease begins, can contribute to a higher percentage of undiagnosed individuals [6].

1.2.2 SYMPTOMATOLOGY AND CLINICAL FEATURES

Typical ACM clinical features are premature ventricular complexes (PVC), or ventricular tachycardia (VT). They exhibit a left bundle branch block (LBBB) pattern, together with T-wave inversion in basal electrocardiogram (ECG), particularly in V1-V3 leads. Left ventricular (LV) or biventricular dilatation can be considered as alternative ACM clinical features, possibly accompanied by heart failure symptoms, which could be misinterpreted as dilated cardiomyopathy manifestation [7]. RV aneurysms are the pathognomonic features of ACM; they can be single or multiple, located in the so-called “triangle of dysplasia” [2]. Clinical features are typically age- and disease stage-related [2].

Four phases characterize ACM pathogenesis. The first, namely the concealed one, is still considered a subclinical condition, with moderate RV structural changes and arrhythmias and cardiac arrest as the only possible manifestations. The “overt electrical disorder” phase represents the second one, when palpitations and syncope occurs, and RV arrhythmias become a common presence, typically effort-stimulated and life-threatening. The progressive loss of RV myocardium leads to the third phase, named the ”RV failure”, with pump impairment. The fourth and last phase, the “biventricular failure”, also engages the LV, when the heart is dilatated due to the extreme thinness of its walls, almost mimicking dilated cardiomyopathy condition [2], [7]. Relevant is to mention though, that arrhythmias and their fatal outcome are present in all phases of the disease [4].

Islands of surviving cardiomyocytes can be observed on histological examination, surrounded by fibrous and fatty tissue. Cardiomyocytes necrosis may be responsible for the inflammatory infiltration reported in up to 75% of samples, composed mainly of T-lymphocytes and macrophages [2], [4]. Important is to mention, however, that apoptotic events have also been proven to occur [2]. Disease progression may manifest as periodic “acute bursts” rather than an ongoing process [2].

1.2.3 DIAGNOSIS AND THERAPY OF ACM

ACM diagnosis remains difficult [4], as no single gold-standard is available [2]. It is based on a set of criteria, established in 1994 and updated in 2010 with the addition of genetic criteria to the clinical ones [8]. This set is organized into several

groups, including global or regional dysfunction and structural changes, tissue characterization of the heart wall, abnormalities in repolarization, depolarization- and conduction-related issues, presence of arrhythmias, and patient's family medical history [9]. Within each of these groups, specific criteria are further categorized as major or minor. To make a diagnosis of ACM, a combination of these criteria is required: either two major criteria, four minor criteria, or one major criterion and two minor criteria from different groups must be present [9].

It is relevant to mention that the criteria explained mostly concern the diagnosis of the classic ARVC form of the disease, leading to underestimation of other forms, such as left ventricular or biventricular. To overcome this problem, the "Padua Criteria" have been proposed as new diagnostic method useful to detect all different types of ACM. This version of the criteria is based on the 2010 version, with the addition regarding fibro-fatty replacement in tissue characterization by CE-CMR and new ECG criteria [6]. Relevant to mention is that for risk prediction genetic diagnostic is still important for both patients and relatives [10].

In paediatric age, a few rare cases of ACM have been reported. With limited exception, the diagnosis criteria still appear to be valid, with negative results often present due to the age-related penetrance of ACM [2].

Nowadays, definitive treatment remains absent [4]. The major goals of clinical management of ACM are to reduce of mortality, prevent disease progression, alleviate symptoms and improve quality of life, and increase exercise capacity [4], [7]. Lifestyle modifications should be suggested before starting any curative treatment, and the only effective management that improves survival is represented by the placement of an implantable cardioverter-defibrillator [2], [4]. Antiarrhythmic drugs are still the most common treatment for ACM patients, with sodium blockers, beta-blockers, sotalol, amiodarone, and verapamil the most used among them. Other treatments for the most severe cases include catheter ablation, implantable cardioverter defibrillator and heart transplant [7]. Finally, standard heart failure managements should be applied in late stages of ACM pathogenesis [4].

1.2.4 GENETICS OF ACM

ACM is an autosomal dominant genetic disease, displaying incomplete penetrance (often age-related [2]) and variable expressivity, even among individuals within the same family, with mild phenotypes to SCD range [4]. In addition, element such as de novo mutations, recessive mutations, compound heterozygosity, and digenetic inheritance contribute to make the genetic understanding of this disease extremely difficult [4], [7].

Genetic testing is recommended in order to simplify early diagnosis, based on the knowledge that a genetic basis has been identified in over 60% of cases [8]. Many of the mutated genes identified associated with ACM primarily encode proteins involved in desmosomes and the area composita, which are mechanical junctions

found in intercalated discs [6]. Specifically, causative mutations are most frequently discovered in genes such as *PKP2*, coding for plakophilin-2 [11]; *JUP*, coding for plakoglobin [12]; *DSG2*, which encodes desmoglein-2 [13]; *DSC2*, coding for desmocollin-2 [14]; and *DSP*, which encodes desmoplakin [15] (Table 1). Mutations in these genes most frequently are inherited with a dominant pattern, although also recessive transmission of the disease have been reported. Importantly, the majority of mutations is located within the *PKP2* gene.

GENE	ENCODED PROTEIN	REFERENCE
PKP2	Plakophilin-2	[11]
JUP	Plakoglobin	[12]
DSG2	Desmoglein-2	[13]
DSC2	Desmocollin-2	[14]
DSP	Desmoplakin	[17]

Table 1 Desmosomal genes associated with arrhythmogenic cardiomyopathy.

Non-desmosomal mutations also been reported and associated with ACM, although with a lower frequency and, in many cases, with pathogenic role yet to be clarified. These non-desmosomal mutations are reported in Table 2.

GENE	ENCODED PROTEIN	LOCALIZATION
CDH2	N-cadherin	Area composita
CTNNA3	αT-catenin	Area composita
TJP1	Tight junction protein-1	Tight junctions
DES	Desmin	Intermediate filaments
LDB3	Cypher	Z-bands
ACTN2	A-actinin-2	Z-bands
SCN5A	Cardiac sodium voltage-gated channel subunit α5, NaV1.5	Gap junctions
PLN	Phospholamban	SR membrane
LMNA	Lamin A/C	Nuclear envelope
TMEM43	Transmembrane protein 43	Nuclear envelope
LEMD2	Nuclear transmembrane protein-25	Nuclear envelope
TGFβ3	Transforming growth factor β 3	Secreted
TTN	Titin	Sarcomere
ILK	Integrin linked kinase	Focal adhesion
FLNC	Filamin-C	Z-bands
RYR2	Ryanodine Receptor 2	Sarcoplasmic reticulum

Table 2 Non-desmosomal genes associated with arrhythmogenic cardiomyopathy [4], [17]

1.2.5 MECHANICAL JUNCTIONS

To resist to the high mechanical stress typical of heart tissue, cardiomyocytes are highly interconnected through their intercalated disks (IDs), complex regions composed of both electric and mechanical junctions, along with ion channels. These

IDs should be seen as a single functional unit, essential for cardiomyocytes adhesion and synchrony [17].

Desmosomes and area composita represent the mechanical part of the junctions contained in the IDs, crucial to ensure adhesion and resistance to the mechanical stress, while gap junctions and ion channels are essential for electrical signal transmission, therefor synchrony of contraction [17].

Cardiac desmosomes are multiprotein complexes that connect the intermediate filament networks of adjacent cells through protein-interaction. Desmoglein-2 (DSG2) and desmocollin-2 (DSC2), the desmosomal cadherins, are transmembrane proteins two cadherin proteins and they represent the scaffold for desmosomal plaque assembly, forming homo- and hetero-dimers with the cadherins protruding from the other cell. Mechanical resistance in cardiac tissue is guaranteed by the network composed of these two proteins along with plakoglobin (PG) and plokophilin-2 (PKP2), which interact with the intracellular domains of DSG2 and DSC2. In turn, PG and PKP2 turn in link desmin intermediate filament through desmoplakin (DSP) [17].

The intercalated discs of mammalian post-natal heart are characterized by the presence of both desmosomes and area composita (AC), which is composed of the special mix of two major junctional elements such as desmosomes and adherens junctions (AJ), linked to both actin cytoskeleton and desmin intermediate filaments. The AC represents more than 90% of the ID volume: only few gap junctions, desmosomes and limited junction-free regions stop it. Important is to mention though, that AC is not present in lower vertebrate's heart (fish and amphibian), which remain composed only of separate desmosomes and AJ. AC might therefore have evolved in order to better support the high mechanical stress typical of mammalian hearts, increasing the extension and amount of ID linkage to the cytoskeleton [17].

AJ composition is based on N-cadherin (N-cad) extracellular homodimers, able to interact with β-catenin, a linker protein interacting with F-actin cytoskeleton filament through αE-catenin (Figure 2) Finally, αT-catenin represents the bridge between cytoskeletal actin and PKP2 [17].

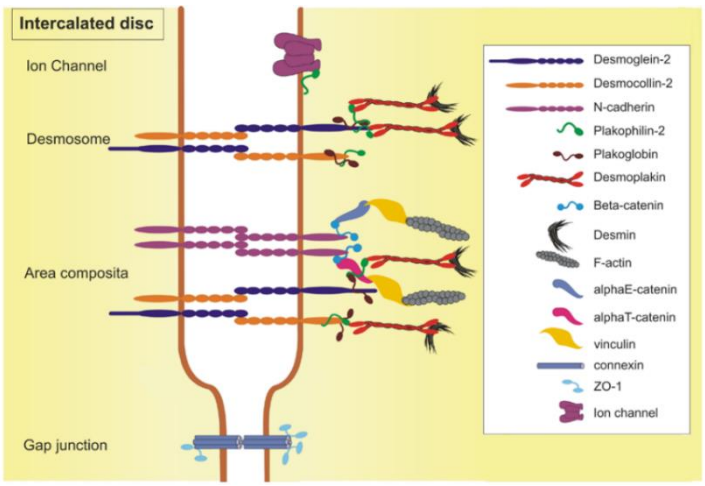


Figure 2 Illustration of the components of intercalated discs (ID). Three junctional complex types plus ion channels compose the IDs: gap junctions, desmosomes and area composita (AC). [17].

The majority of ACM pathogenic mutations involve structural proteins important for ID organization, extending the concept of ACM beyond desmosomes, as originally thought [17].

1.2.6 PATHOGENESIS OF ACM

Loss of myocardium with fibro-fatty replacement is the main feature of ACM, causing electrical instability due to patchy scar formation, resulting in impaired ventricular function and possibly leading to sudden cardiac death [4], [5]. It has been suggested that most of the adipocytes in ACM derive from cardiac progenitor cells of the embryonic second heart field [2], and that pro-fibrotic transforming growth factor beta is overexpressed in ACM samples [4]. Inflammation could be also implicated in the pathogenesis, given the inflammatory infiltration in the form of T-cell abundance found in post-mortem hearts of ACM patients [7], suggesting the role of chronic inflammation in disease progression [4]. However, it is still debated if these cells represent tissue-resident elements, and whether this infiltration represents a secondary event due to cardiomyocyte death or it promotes fibrosis and fibro-fatty replacement is yet to be determinate [4]. Nevertheless, the presence of autoantibodies against DSG2 and a positive correlation between their title and disease severity have been proven in ACM patients, supporting the hypothesis of an important role of inflammation in pathogenesis [4].

Cardiomyocytes are highly connected with other cell types, such as myofibroblasts and immune cells, making hard to understand the cascade of molecular events that lead to ACM cardiac fibrosis [4]. Therefore, ACM pathogenesis remains under investigation, with elements clearly demonstrated to be involved and steps yet to be understood. Among the different existing theories, intercalated discs dysfunction and altered signalling pathways seem to be common elements. Specifically, the first theory, formulated observing ACM patient's hearts samples, highlighted the role of mutations in ID proteins in causing loss of adhesion between adjacent cardiomyocytes, therefore the progressive cardiomyocytes death and the typical loss of ventricular myocardium, with fibro-fatty replacement as the final product. Affected ACM cardiomyocytes appear smaller, with lower desmosome amount and increased intercellular distance [2], [6].

Given the role of desomomes in intra- and intercellular signal transduction, the linkage between ACM pathogenesis and signalling pathways alteration has been proposed [6], [7], with Wnt/β-catenin/Tcf/Lef pathway alteration being the first described in Dsp-deficient mouse models [7].

Wtn pathway represents an essential regulatory element during embryogenesis, maintenance of adult tissue, and remodelling of cardiac adult tissue. During development indeed, it controls cell differentiation, polarization, and migration,

with the ability to reactivate developmental programs during adult age for cardiac remodelling and maintenance of contractile function [5]. It is characterized by timing- and location-specific activation thanks to an overabundance of ligands, receptors and inhibitors [18].

Canonical Wnt pathway suppression has been found to support pro-fibrotic, pro-apoptotic and pro-adipogenic gene expression changes in ACM patients [2]. Its primary signalling molecule is β-catenin, an Armadillo protein able to both bind AJ proteins and enter the cell nucleus to regulate transcriptional gene expression [4].

Degradation of β-catenin is normally prevented by the activation of canonical Wnt pathway, allowing β-catenin to enter in the cell nucleus and playing its role as T-cell factor/Lymphoid-enhancer binding factor (Tcf/Lef), therefor leading to cell growth and differentiation [5]. However, studies conducted on various ACM models have shown that mutations in desmosomal proteins can cause PG to detach from the cell membrane and relocate into the nucleus. Here, it competes with β-catenin, preventing its interaction with Tcf/Lef, ultimately deactivating the canonical Wnt pathway leading to the promotion of adipogenesis. Lipid droplets are indeed typically observed in ACM heart tissue samples [5], [6]. This alteration appears to also impact PPARγ (peroxisome proliferator-activated receptor-γ) expression, a regulator of lipogenesis and adipocyte differentiation, which could be responsible of the fibrofatty tissue development [5]. The role of PPARγ also stays in an interaction with β-catenin supporting its degradation, thus reflecting a direct influence on the canonical Wnt signalling pathway [19]. GSK3β (glycogen synthase kinase 3β) serine/threonine kinase represents another inhibitor of the canonical Wnt signalling pathway through promotion of β-catenin degradation, and it results to be mislocated in the ID in both mouse models and ACM patients, resulting in cardiac function and surviving level improvement [5], [6], [19]. (Figure 3).

Another element probably linked to the adipogenesis is the Hippo-Yes-associated protein (YAP) signalling pathway, which has been founded dysregulated in both ACM patients and animal models [4]. Particularly, the inactivation of this pathway leads to activation via hypo-phosphorylation of YAP and TAZ through a kinase cascade. YAP and TAZ are transcription factors able to activate the transcriptional enhancer factor domain (TEAD): the interaction of activated TEAD with β-catenin/TCF (T-cell factor) complex, promotes the Wnt canonical pathway and sustaining cell growth resistance to apoptosis [4], [5], [7].

Activation of the Hippo signalling pathways by neurofibromin-2 (NF2), a cell-cell and cell-matrix adhesion protein, on the other hand, leads to MST1 and MST2 (mammalian STE20-like protein kinase 1 and 2) phosphorylation of LATS1 and LATS2 (large tumour suppressor homologs 1 and 2), which phosphorylate YAP and TAZ, avoiding their entrance into the nucleus. They are at this point degraded in the cytosol [5], [7].

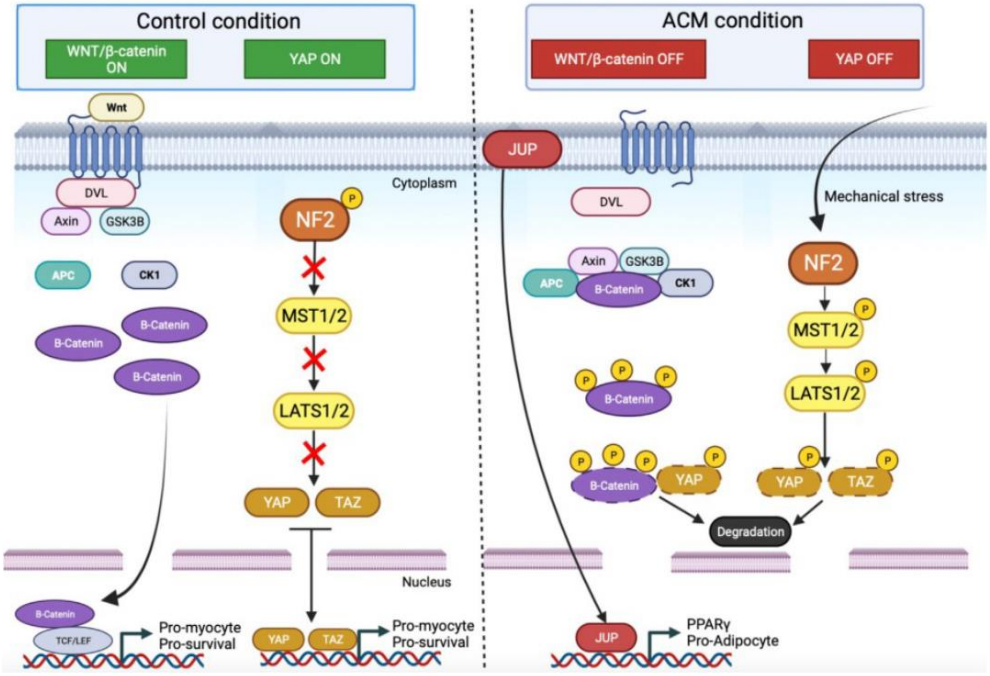


Figure 3 Representation of Wnt and YAP signalling pathways in healthy condition (left) and ACM (right). The left side shows active Wnt and YAP. Particularly, Wnt ligand is able to activate Wnt receptor, stimulating the DVL-Axin-GSK3B complex and, consequently, β-catenin cytosolic accumulation. β-catenin is therefore free to enter the nucleus and bind TCF/LEF transcription factors. The YAP pathway activation starts from NF2 phosphorylation, with consequent suppression of the kinase cascade. As consequence, YAP and TAZ enter the nucleus. The right panel shows the ACM mechanism based on dysregulation of Wnt and YAP pathways. It all starts from desmosomes damage, which leads to PG dissociation from the desmosomal complex and entrance into the nucleus, where it is able to inhibit Wnt/β-catenin signaling and to stimulate PPARγ expression. Hippo pathway activation starts from NF2 and leads to activation of the kinase cascade. YAP phosphorylation blocks its entrance in the nucleus: in the cytoplasm it is able to interact and stimulate β-catenin degradation. In this scenario, GSK3β remains localized at the ID level [5].

Retention of YAP in the cytosol results in inhibition of the Wtn pathway by sequestring β-catenin in the cytoplasm and resulting in the activation of pro-adipogenic genes [4], [7]. The Hippo-Yap pathway remains poorly studied in ACM, further studies are still needed in order to fully understand its exact role in disease pathogenesis [5] (Figure 3).

The last signalling found to be involved in ACM pathogenesis is the TGFβ pathway [4], [19], supporting recent founding that revealed TFGβ overexpression in ACM models due to plakophilin-2 loss [4]. Activation of the canonical TGFβ pathway is based on SMAD2-3 phosphorylation due to TGFβ1 and TGFβ2 binding to receptors. SMAD2-3 represents a transcription factor controlling the expression of extracellular matrix proteins and metalloproteinases. The non-canonical TGFβ pathway is instead able to activate mitogen-activated protein kinase (MAPK) with ROS generation [19]. In animal models activation of MAPK signalling pathway

leads to decreased contractility, increased cardiomyocytes apoptosis and hypertrophy [2].

Alterations in calcium handling support the pro-arrhythmogenic phenotype and impact cell-cell adhesion. Indeed, desmosomes exist in a calcium-independent adhesive conformation and in a less adhesive calcium-dependent one, which allows tissue remodelling. These findings highlight the electrophysiological remodelling occurring together with the morphological and structural remodelling of cardiomyocytes during ACM pathogenesis [5].

Given that the two most important regulatory element of energy production are represented by cytosolic calcium and ADP concentrations, these events support the altered energy metabolism observed in ACM cardiomyocytes [20]. Mitochondrial dysfunction, which has not been demonstrated during pathogenesis but would be coherent with the phenotypic manifestations of the disease, also leads to reactive oxygen species (ROS) production, which may result in oxidative stress with arrhythmogenic effect. Relevant is also to mention that under pathological conditions, cardiac metabolism shifts back from fatty acids oxidation to a greater glucose role as energy production source, increasing its hydrolysis and mirroring the fetal condition. Without changing in glucose oxidation as well, the cell will experience an imbalance in glucose metabolism that, together with the reduction of fatty acids oxidation, leads to a progressive reduction in energy supply. However, the hypothesis of mitochondrial dysfunction resulting in electrical and structural changes in ACM remains to be confirmed. [20]

Desmosomes destabilization based on mutations in genes coding their proteins, also leads to electrical remodelling, one of the hallmarks of ACM. Particularly, alterations in gap junction organization and connexin expression are typically present in the diseased myocardium. PKP2 mutations were shown to lead to a drastic loss of Connexin43 expression in hiPSC-CMs, an essential transmembrane protein for gap junction structure [5]. Inter-cellular coupling alterations are therefore sustaining a strong pro-arrhythmogenic condition. In this scenario, also $Na_v1.5$ voltage-gate sodium channel appears altered: it represents a desmosome/connexin macro-molecular complex protein responsible for sodium current (I_{Na}). Its drastic reduction due to desmosomes alteration could lead to a reduction of I_{Na} density, establishing an arrhythmogenic substrate in early ACM patients. The role of I_{Na} reduction as causal in the disease or resulting from a pathological phenotypic remodelling is yet to be determinate[5].

1.3 ANIMAL MODELS FOR ACM

The first disease-associated gene of ACM was identified more than 20 years ago; since then, different in vivo models have been generated and characterized to understand the pathogenetic process of ACM and to study the role of desmosomal proteins JUP, S2p, Dsg2, and PKP2 in embryonic development. Particularly, in the 1990s knockout mouse models of desmosomal genes revealed for the first time that

cell-cell adhesion defects lead to heart failure and embryonic lethality if in homozygous condition. These findings sustained the hypothesis of disease related to desmosomes. Subsequently, several transgenic, knock-in, and conditional heart-specific mouse models elucidated the pathogenic role of some mutation in desmosomal genes found in affected patients. Thanks to this models, it was demonstrated that even heterozygous knockout of JUP protein in mice leads to right ventricular dilation and ventricular arrhythmia, with exercise that, interestingly, strengths, and accelerates cardiac dysfunction. Additionally, plakoglobin-deficient mice shown severe cardiac fibrosis and dysfunction, together with ventricular arrhythmia. Moreover, cardiac-specific DSP knock-out murine models shown fibro-fatty myocardial replacement and cardiac dysfunction, with deregulation of genes inhibiting Wtn pathway and sustaining inflammation [10].

About 5% of ACM cases are caused by mutation in desmoglein-2 (*DSG2*) gene [10]. In the human genome, there are four DSG isoforms (DSG1 to DSG4), which share a common overall domain structure. Particularly, DSG2 represents the only type expressed in cardiac desmosomes and area composita, thus holding an essential role for cardiac tissue integrity [21]. It shares with the other desmogleins the overall domain organization in five extracellular cadherin (EC) domains, a transmembrane domain passing through the membrane with one single region, and an intracellular domain interacting with the intermediate cytoskeleton filaments through adaptor proteins desmoplakin and plakoglobin [21].

In this thesis, I focused on the murine mutation corresponding to the human c.1672C>T in *DSG2* gene, resulting in the non-sense mutation p.Q558X* (Figure 4). a truncated protein is formed, lacking in the transmembrane and intracellular domains, and preventing *DSG2* plasma membrane localization (figure 4).

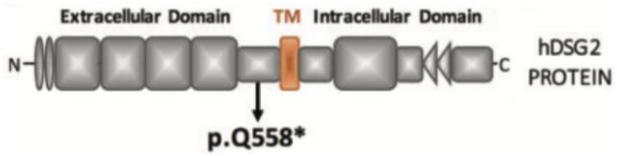


Figure 4 Scheme representing DSG2 protein indicating the Q558 mutation taken into account in the second part of this thesis.*

For Dsg2, two murine models have been generated over the years, harbouring the p-N271S and p.Q554X mutations. The first model shown arrhythmia and alteration in intercalated disk's structures, while the second one developed cardiac fibrosis and low number of cardiac desmosomes. [10]

A new murine model for ACM was recently generated: a transgenic mouse harbouring the Q558* nonsense mutation in DSG2 genes found in an ACM, particularly relevant for the second part of this thesis, given that the same mutation was implicated. From this model, genome-wide RNA-seq highlighted 24 micro-RNAs (miRNAs) deregulated in transgenic mice compared to non-transgenic animals. Particularly, the four most up- or down-regulated are thought to have a regulatory function on the Wtn pathway, which was found to be reduced in transgenic mice. The heart of these animals was characterized by small desmosomes

size and signs of fibrosis [16]. Our group previously generated a knock-in murine model for this mutation, both in homozygous and heterozygous condition. Preliminary data in homozygous mice showed increased left atria dimension and thinning of right atria wall, along with fibrosis signs in both atria and inflammatory infiltrates. Heterozygous mice, in these preliminary data, showed no ACM phenotypic or functional signs.

Danio rerio (zebrafish) models represented another important tool for understanding molecular mechanisms in ACM pathogenesis and in non-desmosomal genes association with the disease, such as *SCN5A, LMNA, ACNT2,* and *TTN* [10]. Knockdown of *PKP2* in danio rerio models led to cardiac edema and decreased heart rate.

These animal models show at least two of ACM features, including ventricular dysfunctions, electrical abnormalities, gap junction delocalization, cardiomyocytes death, and inflammatory infiltration, but they often fail to display one of the hallmarks of the disease, namely the fibro-fatty replacement of the myocardium [19]. Other important limitations remain in the differences between these animal systems and the human one, such as the bicameral heart of zebrafish. A solution could be the use of larger animals, such as Boxer dogs, who seem to better represent the pathological features of ACM, including fibro-fatty replacement and inflammatory infiltrations [10]. Right-ventricular dilation, thin right-ventricular wall, ventricular tachycardia, and fibro-fatty replacement appear to spontaneously occur in domestic cats, although no genetic mutations associated with ACM are known in cats [10].

1.4 ROLE AND RELEVANCE OF hiPSCs

Cell culture represents a useful tool in the investigation of the molecular mechanisms of ACM pathogenesis, with cardiomyocytes being the most common due to their desmosome-rich adhesion and electrical properties, therefore relevance for the disease [19]. However, adult cardiomyocytes cannot be cultured for a long period of time due to their low viability, limited proliferation, and dedifferentiation in cell culture [19]. In this regard, cardiomyocytes derived from human induced pluripotent stem cells (hiPSC-CMs) based on hiPSCs technology by Yamanaka and colleagues, provide a unique platform to model human disease and, thanks to their self-renewing ability, an unlimited supply of differentiated somatic tissue [22]. In this approach, patient-derived somatic cells are converted, passing through the pluripotency state, into other cell types simply applying on a cocktail of reprogramming factors. Among the advantages of this technology long culture maintenance of hiPSCs, proliferation capacity, and patient's genetic background are highlighted. Cardiomyocytes derived from hiPSCs made it possible to summarize the hallmarks of ACM pathogenesis, namely: desmosomes alterations and destabilization, lipogenesis with lipid droplets accumulation, Wnt signalling pathways abnormality, electrophysiological alterations, and arrhythmias. They also represent a unique platform for drug screening and hence therapy [5].

The first successfully reprogrammed hiPSCs generated from an ACM-symptomatic patient's fibroblasts and differentiated down the cardiac lineage, harboured a PKP2 mutation (c.1841T>C), and showed an under-expression of PKP2 and JUP, along with evident lipid droplets in culture. Exposure to testosterone made apoptosis and lipid accumulation worse, while exposure to estradiol improved the phenotype. These results support the higher AC incidence observed in male compared to female. Successively, hiPSCs lineage harbouring the c.972InsT PKP2 mutation led to altered desmosomal structures and under-expression of Cx43. These early findings were confirmed by several following studies with hiPSCs cell lines carryng other PKP2 mutations founded in ACM patients. The first DGS2-mutated hiPSCs line harboured the p.Gly638Arg mutation, which led to cardiac action potential abnormalities and reduced I_{Na} current. From the c.C355T, p.R119X DSG2 mutation, a second hiPSCs line was generated, which showed reduction in DSG2 and DSC2 abundance in heterozygous cells, and reduced conduction velocity and arrhythmic events in homozygous cells. Finally, hiPSCs carrying the c.2358delA and p.Asp787fs DSG2 mutation were generated and deregulation of major calcium-handling genes *RYR2, CAMK2A, CASQ2*, and *SLN* was observed. hiPSCs lines harbouring DSP mutations have also been generated, specifically for the p.R451G and the p.H1684R ones. The first line showed lower protein levels of Cx43 and high level of phosphorylated Cx43, and the second one reduced I_{Na} current and a phenotype consistent with pro-arrhythmogenic events. Several other studies successively confirmed this initial finding [23].

1.5 PROTOCOLS FOR hiPSCs DIFFERENTIATION TO hiPSC-CMs

hiPSCs differentiation into cardiomyocytes mirrors the subsequent steps of embryonic cardiac development, relying on the activation and deactivation of specific signalling pathways with timing of growth factor application being the most important regulatory element [24]. Nowadays several protocols exist, characterized by peculiar selection, concentration, and time of application of differentiation agents controlling mesodermal induction and cardiac differentiation, such as bone morphogenetic protein 4 (BMP4), activin A, fibroblast growth factors (TGFs), and vascular endothelial growth factor (VEGF), following temporal control of Wnt signalling pathways [24]. The most efficient existing protocols are based on the chemically defined basal medium RPMI supplemented with B27, a mix of 21 components. Metabolic purification represents a key step in differentiation protocol, allowing to eliminate other cell types besides cardiomyocytes due to the replacement of glucose with lactate in the culture medium: cardiomyocytes are the only cell type capable of utilizing the TCA cycle for ATP production [25]. Common steps among the exhisting hiPSC-CMs differentiation protocols are illustrated in Figure 5.

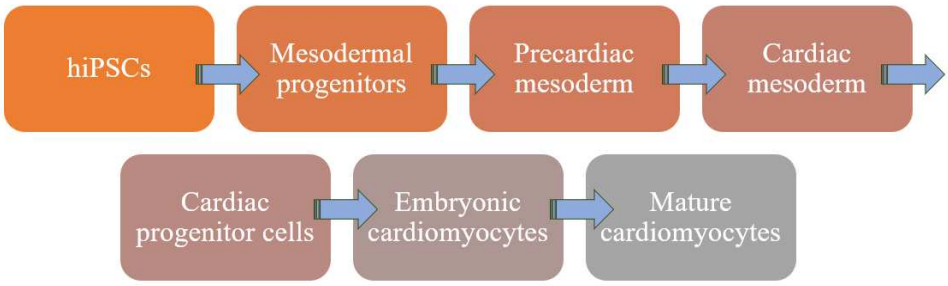


Figure 5 Schematic representation of different steps in hiPSC-CMs differentiation from hiPSCs.

From a general point of view, for hiPSC-CMs differentiation we progressed from complex, uncontrolled systems to relatively robust, specific, and defined protocols, leveraging the knowledge gained on each step from more than two decades of work. However, further improvements are still needed in scalability, time and cost, and increased robustness, with the minimum nutritional requirements still to be found [26]. In addition, cardiomyocytes differentiation protocols still show incoherence and different levels of effectiveness and reproducibility [24]. Comparing available protocols is therefore a useful tool in this regard and to choose the most suitable one for the purpose of the study.

Future prospects for hiPSC-CMs stay in two- and three-dimensional coculture approaches, which reproduce the hole heart environment, to improve and rationalize patient-specific treatment, supporting the personalized-medicine scenario and gene replacing therapy application [5].

2 AIM OF THE THESIS

Arrhythmogenic cardiomyopathy (ACM) is a rare genetic heart muscle disease, characterized by ventricular arrhythmias and fibro-fatty replacement of the myocardium. Several studies support the involvement of the Wnt/b-catenin signaling in the disease pathogenesis; however, the detailed mechanisms underlining this complex disorder haven't been clarified yet. In this context, useful novel insights can arise from the integration of human in vitro and murine in vivo models. Reproducibility and comparability of the data are therefore crucial elements to gain new insights on ACM pathogenesis.

Preserving the patient-specific background, cellular models based on human induced pluripotent stem cells (hiPSCs) provide a unique platform to study this disease. For the differentiation of hiPSCs to cardiomyocytes, several protocols currently exist, exhibiting peculiar efficiency and conferring specific characteristics to the final cells. Therefore, the first aim of this thesis is to compare two protocols to generate hiPSC-derived cardiomyocytes (hiPSC-CMs), in order to offer a useful tool for choosing the most appropriate one for the purpose of the study and to investigate the solidity of the data already achieved.

The second aim consists in the validation of a transcriptomic dataset obtained from murine cardiac samples carrying the desmoglein-2 p.Q563* non-sense mutation, corresponding to the human ACM p.Q558* mutation. Obtained data will be useful to provide new insights on the molecular mechanisms driving the pathogenesis.

3 MATERIALS AND METHODS

3.1 CELL CULTURE

Two protocols for differentiation of hiPSCs into hiPSC-derived cardiomyocytes were performed in order to compare them by analysis of expression of differentiation markers and analysis of beating behaviour. The choice of the two protocols was based on the one most widely used in the laboratory: "A" indicated the protocol adapted from the Heart Research Center Goettingen, University Medical Center Goettingen, and "B" a protocol used in Dr. Calore laboratory at Maastricht University. Wild type blood-derived hiPSCs were used.

3.1.1 COATING PREPARATION

Cells were cultured in 6-well plates and T-25 flasks. For coating preparation, Geltrex (ThermoFisher Scientific, #A1413202) resuspended in DMEM F-12 (ThermoFisher Scientific, #11320033) was used. For 6-wells plates 1ml/well of 0,16 mg/ml Geltrex was used, while 2,5 ml/T25 of it for the flasks. After coating, both wells and flasks were kept at least for 40 minutes at 37°C.

3.1.2 hiPSCs THAWING

hiPSCs generated from epithelial renal cells from healthy patients were used, they were already available in the lab, stored in liquid nitrogen.
After thawing at 37°C in a water bath, cells were mixed with 8 ml of E8++, made of Essential 8 medium (ThermoFisher Scientific, #A1517001), 2% E8 supplement, 0,5% penicillin-streptomycin (P/S), and 0,1% Thiazovivin (TZV) in a 15 ml tube. After centrifugation at 200g for 5 min and resuspension in E8++ with 0,1% TZV, cells were seeded on Geltrex-coated 6-wells plates (1,5 ml/well). The use of TZV helps cells in the adhesion process, but after 24 hours it becomes toxic, so medium change into E8++ without it is needed.

3.1.3 hiPSCs PASSAGING

After thawing, expansion of hiPSCs was performed using SDS-EDTA 0.5 mM: the culture medium was removed, cells were first washed with 1 ml of SDS-EDTA, and then rinsed with 1,5 ml of the same solution to detach the cells. After 4 min of incubation at 37°C, the detachment of the colony edge was checked. Then, the SDS-EDTA solution was removed, 1,5 ml of E8++ was added and the plate was shaken gently to obtain patches of about 15 cells in suspension. Finally, for cell seeding, the medium with the resuspended cells was split into 2 or 3 new wells. The remaining medium was then added to reach 1,5 ml/well. The initial well could be maintained by adding the appropriate volume of E8++ or it could be discarded.

When 80-90% confluence was reached, to prepare the cells for differentiation, they were split in a 1:6 to 1:8 ratio and seeded in new 6-well plates for protocol B and in T25 flasks for protocol A. The culture medium was aspirated and the cells were

washed once with 1 ml of E8. Then, a 5-min incubation at 37°C with 800 µl/well of TrypLE™ Select Enzyme (10X), no phenol red (ThermoFisher Scientific, #A1217703), was performed. TrypLE dilution was made using 1 ml of E8+ medium with 0,5% P/S and 0,1% TZV in each well. The cells were then collected in a 15 ml tube and, after rinsing each well with 1 ml of E8+ with 0,1% TZV to collect the remaining cells, centrifugation was performed for 3 min at 300g. After aspirating the supernatant, the pellet was resuspended in an appropriate volume of E8++ with TZV to seed 1,5 ml in each well of the 6-well plate and 4 ml/T25. After 24 hours, medium was changed with E8++.

3.1.4 hiPSCs CRYOPRESERVATION

For long-term cell preservation, cryopreservation and storage in liquid nitrogen was performed. The procedure used was the same as previously described for cell passaging. The only difference is in the final resuspension after centrifugation: in this case, the cells were resuspended in 0,4 ml of E8++ combined with 0,4 ml of cold cryopreservation medium, consisting of E8++, 0,25% DMSO and 0,0025% TZV. After transfer into cryovials (0,8 ml each), the cells were stored for 48 hours at -80°C inside the Mr. Frosty freezing container. Finally, the vials were placed in liquid nitrogen.

3.1.5 hiPSCs CULTURE

hiPSCs were cultured in 6-welL plates, using 1,5 ml E8++/well. Medium change was performed daily until 80-90% confluence was reached. At this point, cells were split at 1:6 to 1:8 ratio and transferred to T-25 flasks for differentiation protocol A and to new 6-well plates for differentiation protocol B.

3.1.6 hiPSCs DIFFERENTIATION to hiPSC-CMs

3.1.6.1 PROTOCOL A

For differentiation protocol A, the following culture media were prepared:

- Basal Serum Free Medium minus insulin (BSFM(-)), consisting of RPMI 1640 with Glutamax (ThermoFisher Scientific, #61870-010), 1% of 100X sodium pyruvate (ThermoFisher Scientific, #11360) , 0,4% of B27 supplement minus insulin (ThermoFisher Scientific, #A1895601, 200 µl in 50 ml final volume), 200 µM ASC (Sigma-Aldrich, #A8960-5G).
- Basal Serum Free Medium with insulin (BSFM(+)), consisting of RPMI 1640 with Glutamax (ThermoFisher Scientific, #61870-010), 1% of 100X sodium pyruvate (ThermoFisher Scientific, #11360), 0,4% of B27 supplement (ThermoFisher Scientific, #17504-044, 200µl in 50 ml final volume), 200 µM ASC (Sigma-Aldrich, #A8960-5G).
- Selection Medium (SM), consisting of 49,15 ml RPMI without glucose (ThermoFisher, #61870-010), 0,25 ml sodium lactate (final concentration

2,2 mM), 0,1 ml of 2-mercaptoethanol (final concentration of 0,1 mM), in 50 ml final volume.

Differentiation protocol A (Figure 6) began at day 0, when 80-90% confluent cells cultured in T25 flasks were rinsed twice with RPMI medium before adding 7 ml/flask of mesodermal induction medium (MIM) to induce mesodermal differentiation, consisting of BSFM(-), 1µM CHIR (1µl of 10mM stock in 10 ml of medium), 5 ng/ml BMP4 (5 µL of 10 µg/ml), 9 ng/ml Activin A (9 µl of 10 µg/ml stock in 10 ml of medium) and 5 ng/ml bFGF (5 µL of 10 µg/ml stock in 10 ml of medium). MIM should be prepared fresh every day using prewarmed BSFM(-). MIM was changed on day 1 and day 2.

On day 3, flasks were rinsed twice with 5 ml RPMI, then 12 ml of Cardiac Differentiation Medium (CDM) was added in each one to induce cardiac differentiation, consisting of BSFM(+) and 5µM IWP4. CDM should be prepared fresh every day using prewarmed BSFM(+). On day 5, 6 and 8, cells were rinsed once with 5 ml RPMI, then 7 ml CDM was added. On day 10, CDM was changed without rinsing.

On day 12, CDM was replaced with BSFM (+) for maintenance. On day 14 medium change was performed without rinsing.

On day 15, metabolic selection started in order to increase the percentage of cardiomyocytes in culture by reducing the amount of other cell types. After washing once with 1X DPBS, differentiated cardiac cultures were treated with 7 ml/flask of SM for 5 days (day 15 to day 19).

After selection, purified cardiomyocytes were cultured in BSFM(+) medium for 5 days (day 20 to day 24) for recovery.

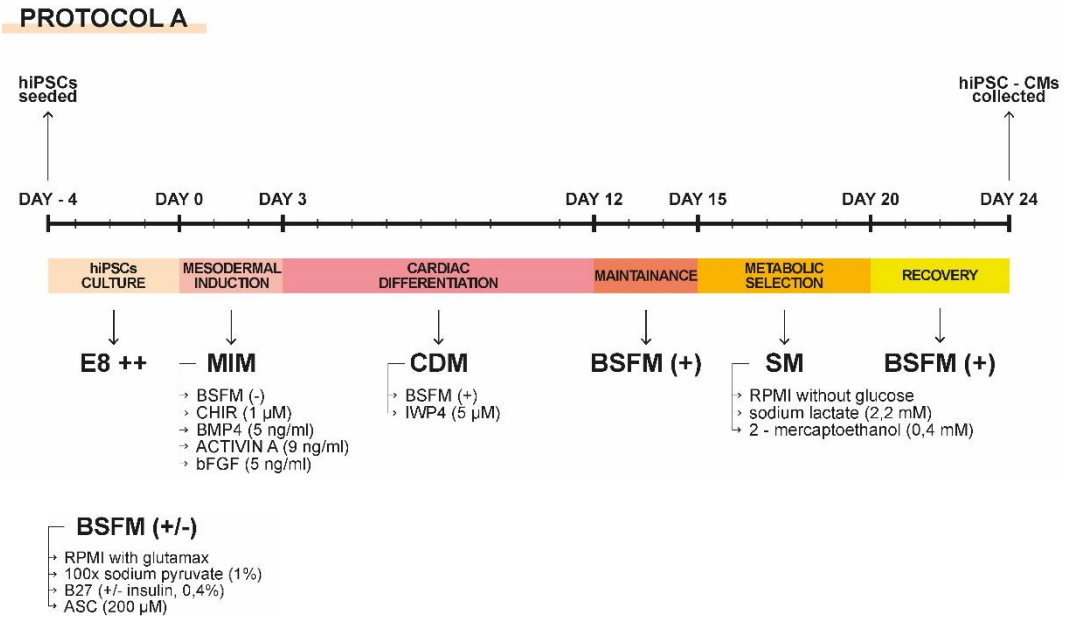


Figure 6 Schematic representation of protocol A. MIM, Mesodermal Induction Medium. BSFM, Basal Serum Free Medium. CDM, Cardiac Differentiation Medium. SM, Selection Medium.

3.1.6.2 PROTOCOL B

For differentiation protocol B, the following culture media were prepared:

- Cardiac differentiation medium (CDM), consisting of RPMI 1640 with HEPES and GlutaMAX™ (ThermoFisher Scientific, #72400021), 2% B-27™ supplement minus insulin (ThermoFisher Scientific, #A1895601), 0,2 mg/ml L-Ascorbic acid 2-phosphate (Sigma-Aldrich, #A8960), 0,5 mg/ml human Albumin (Sigma-Aldrich, #A9731) and 0,5% P/S.
- Cardio culture medium (CCM), consisting of RPMI 1640 medium with HEPES and GlutaMAX™ (ThermoFisher Scientific, #72400021), 2% serum-free B-27™ supplement (ThermoFisher Scientific, #17504044) and 0,5% P/S.
- Cardiac selection medium (CSM), consisting of RPMI 1640 glucose-free medium (ThermoFisher Scientific, #11879020), 0,2% 2-mercaptoethanol (50 mM, ThermoFisher Scientific, #31350010), 1% P/S and 0,44% sodium lactate/HEPES solution (1 M). 3 ml of sodium DL-lactate solution (60% [w/w] in H_2O) (Chemcruz, SC-301818A) was diluted in 18 ml of 1M HEPES solution (Sigma-Aldrich, #H0887) to prepare sodium lactate/HEPES solution (1 M).

Differentiation protocol B (Figure 7) was performed in 6-well plates and started at day 0, when cells with 75-85% confluence were selected, washed once with 1X DPBS and treated with 2 ml/well of CDM supplemented with 0,84 µl/ml (10 µM) CHIR99021 12 mM to induce mesodermal differentiation. On day 1 (24 h after CHIR administration), the medium was changed to 2 ml/well of CDM.

On day 3, cells were washed once with 1X DBPS and 2 ml/well of CDM supplemented with 1 ml/ml (5 µM) of IWP2 was added to induce cardiac differentiation. On day 5, CDM was changed after washing once with 1X DPBS, while on day 7, CDM was changed without washing.

On days 9 and 11, after washing with 1X DPBS, cells were cultured with 2 ml/well of CCM for maintenance. After day 11, the medium was changed twice a week.

Cardiac metabolic selection was started on day 14 to increase the percentage of cardiomyocytes in culture by reducing the amount of other cell types. After rinsing with CSM, 2ml/well was added. Then, the medium was changed daily for 5 days (day 14 to day 18), followed by CCM with weekly change until day 24 for recovery.

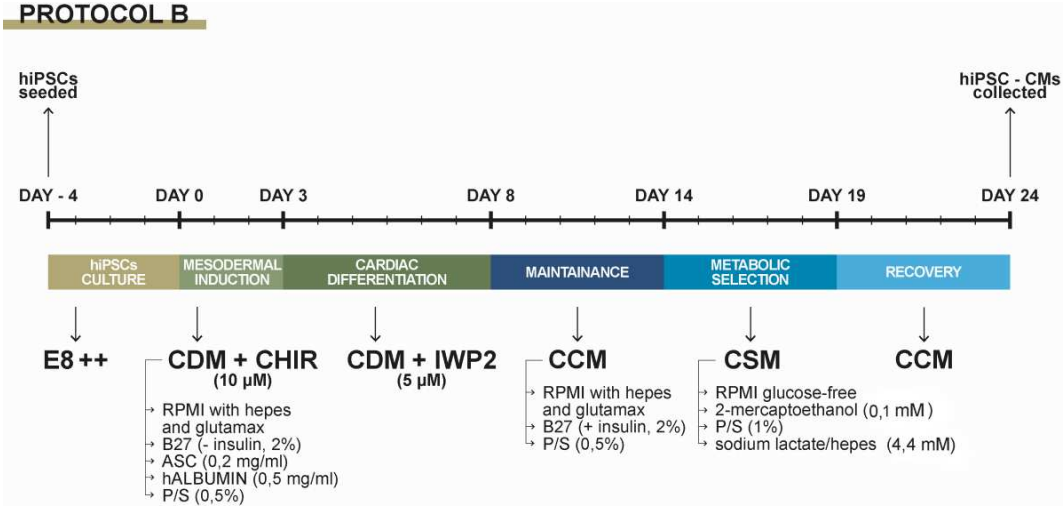


Figure 7 Schematic representation of protocol B. CDM, Cardiac differentiation medium. CCM, Cardio culture medium. CSM, Cardiac selection medium. P/S, pen/strep.

3.2 RNA EXTRACTION

For isolation of RNA from hiPSC-CMs, the Direct-Zol RNA MiniPrep Kit (Zymo Research, #R2051) was used following the protocol provided by the manufacturer. First, the culture medium was aspirated from the wells/flasks, which were then washed once with 1X DPBS. Cell membrane lysis was then performed by adding 950 µl/well and 2,5 ml/flask of TRIzol® and incubating at room temperature for 3 minutes. To stop the reaction, an equal volume of absolute ethanol was added. The wells/flasks were then scraped with a cell scraper and the volume was transferred to a 2 ml tube. RNA purification was then performed in a Zymo-Spin™ IICR column in a collection tube. After centrifugation at 15000 g for 30 seconds, the flow was discarded, collection tubes were replaced and columns were washed with 400 µl of RNA-wash buffer. After another centrifugation at 15000 g for 30 seconds, 5 µl of Dnase I (6 U/ul) and 75 µl of DNA digestion buffer were added to the columns, followed by 15 min of incubation at room temperature. Two washes were then performed with 400 µl of Direct-zol™ RNA PreWash, centrifuging each time at 15000g for 30 seconds and discarding the stream. Finally, 700 µl of RNA Wash Buffer was added and centrifugation at 15000g for 1 min was performed to remove all the wash buffer. To elute the RNA, the columns were transferred to new RNase-free tubes, then 25 µl of DNase/RNase-free water was added and centrifugation at 15000g for 30 seconds allowed the RNA to be collected in the tubes.

3.3 RNA QUANTIFICATION (NANODROP)

Quantification of the eluted RNA was performed with the Thermo™ Scientific NanoDrop™ 2000 Spectrophotometer, using as blank Milli-Q water and 2 µl as samples volume.

3.4 REVERSE TRANSCRIPTION

For reverse transcription, reverse transcriptase was used, a retrovirus-encoded enzyme that starts from the 3' end of the primers to synthesis the cDNA, using the RNA sequencing as template. Reagents used for the reaction are shown Table 3.

REAGENT	VOLUME PER SAMPLE (µl)
5x buffer M-MLV (Promega, #M531A)	8
M-MLV RT (Promega, #M369B)	2
Oligo dT (IDT, #51011505)	2
Hexamer (IDT, #0000434088)	2
RNAsin (Promega, #N261A)	2
dNTPs (10 mM) (Promega, #U1330)	2
RNA (1 µg, 45,45 ng/µL)	22
TOTAL VOLUME	**40**

Table 3 Reverse transcription reagents and dosages.

3.5 PRIMER DESIGN

For primer pair design, the online Primer3 tool was used https://primer3.ut.ee/. Using the mRNA sequences of the gene of interest as input, setting the maximum difference in melting temperature between the two primers as 1°C and the length range of the amplicon between 100 and 210 nucleotides, the software allows the user to choose among different primers pairs. The choice was made based on the prediction of the presence of nonspecific annealing regions made by BLAT (https://genome.ucsc.edu/cgi-bin/hgBlat?command=start), a UCSC Genome Browser tool, using GRCm39/mm39 as reference for murine genome.

Alternatively, an online primer bank was used: PrimerBank (Harvard University, https://pga.mgh.harvard.edu/primerbank/). In this case, specificity is favored over chemical properties and the primers pairs are tested experimentally by DNA amplification and amplicons analysis by gel electrophoresis and NCBI Blast analysis.[27]

3.6 POLYMERASE CHAIN REACTION (PCR)

3.6.1 GRADIENT PCR

The choice of the most suitable temperature for each newly designed primer pair was based on the results of gradient PCR, testing different temperatures spanning from 58°C to 68°C. Each primers pair, in fact, has its own suitable annealing temperature because of its characteristics such as length and GC percentage. Primer pair sequences and chosen temperatures for real-time PCR are shown in Table 4.

SPECIES	TARGET GENE	FORWARD SEQ. (5'-3')	REVERSE SEQ. (5'-3')	TEMP (°C)
mmu	*l7 (hk)*	GAAGCTCATCTATGAGAAGGC	AAGACGAAGGAGCTGCAGAAC	58-68
mmu	*ppp1r13l*	GAAAAGTGCGACCCTTACCG	CCACCACCAATCAGTCTCCT	60
mmu	*enc1*	CTGTTACCTCCCGGAACTGT	CCACGCCGTCATTCTGTAAG	60
mmu	*txn2*	TCAGTCATCTCCAGGAAGCC	TCTGGTGGTGTGTACTGTCC	60
mmu	*cfl1*	ATGACATGAAGGTTCGCAAGT	GACAAAAGTGGTGTAGGGGTC	60
mmu	*col4a2*	ACAAGCTCTGGAGTGGGTAC	CAGTAGGACTTGTCGTTGCG	68
mmu	*gpx1*	ATAGAAACCCTGCTGTCCCA	CAGGTGTTTCTCCGTGCAAA	60
mmu	*arhgap5*	AGTTGGACTCTCTGGGACTG	CTCGGCCTCCAAAGTCAATG	60
mmu	*sod1*	AACCATCCACTTCGAGCAGA	AGCCTTGTGTATTGTCCCCA	60
mmu	*hrc*	AAGGTGAGAAGGACAGAGGC	CGTGGAGCTGTCTAACATGC	62
mmu	*noc2l*	GACAGCAGAGCGTAAACAGG	ACAAAGTGGAGAGGAGAGGC	65
mmu	*sap18*	CTGTGGCTGCTAAAGGTGTG	CTGGACCTACGGAAGAGCAA	62
mmu	*mbd3*	CCCCAGCGGGAAGAAGTTC	GACCTGGTTGGAAGAATCATAGC	60
mmu	*ndufb10*	CTCCCTAACCCCATCACCTAC	TGGCACTCGACGGTACTGT	60
mmu	*rb1cc1*	GACACTGAGCTAACTGTGCAA	GCGCTGTAAGTACACACTCTTC	60
mmu	*etfb*	CTGTCAAGAGGGTCATCGACT	CACAGAAGGGGTTCATGGAGT	60
mmu	*litaf*	CCACAGGGCTCATTACAGGC	GCTGCACATAAACGGTCTGC	65
mmu	*loxl1*	TGCCCGACAACTGGAGAGA	TGCGGATAGGGGAACTGCT	65
mmu	*ctla2a*	CTCCACCCCCTGATCCAAGT	ACACGAGTCTTCTGTGTCTTTCT	65
mmu	*dsg2*	AAGACGAAGGAGCTGCAGAAC	CCAGGAACAGGTCATCGAGTT	60
mmu	*myl2*	TTGAGCAGACCCAGATCCAG	GGACCTGGAGCCTCTTTGAT	60
mmu	*isg20*	TTCTGAAAGGCAAGCTGGTG	ATGTTCTTGTGTAGCAGGCG	60
mmu	*tfgb1*	CTTCAATACGTCAGACATTCGGG	GTAACGCCAGGAATTGTTGCTA	60
mmu	*lars2*	TGTGGATTCTGCGTGGTACT	AGCATGTTCCTTTCCTCCGA	58
mmu	*slc2a3*	ATGAAGTCCCTGGTCCCTTG	AGAGCCTTCCAACCTCAGAC	62
mmu	*sik1*	CCACTTTGCCGCCATTTACT	AGGGTCACACGGGAGAATTT	60
mmu	*cfap276*	TGTATAAGAACCCAGCTCACCTT	GTTATACAAGGTTGCTAAGCGGA	58
mmu	*tmem88*	GGACCTTGTGCCTTCCTTTG	ATGTGCCACCCCTGTATTCA	60
mmu	*des*	TATCAGGACAACATTGCGCG	CCTCCAGTAGCTTCCGGTAG	60
mmu	*nol3*	AGACACCCAAAGCCATCTCA	TGTCCTTGCTACCCTTCCTG	60

Table 4 Primers and temperatures for real-time PCR. Mmu, mus musculus. HK, housekeeping.

A set of primer pairs used was already available and tested in the laboratory by gradient PCR to identify the most suitable annealing temperature for qPCR. The sequences and temperatures chosen from this set are shown in Table 5.

SPECIES	TARGET GENE	FORWARD SEQ. (5'-3')	REVERSE SEQ. (3'-5')	TEMP. (°C)
hsa	*L7 (hk)*	AATTCGAATGGCGAGGATGGCAAG	TGCAACACCTTTCGAACCTTTGGG	59-64
hsa	*OCT4*	TTCAGCCAAACGACCATCTG	CACGAGGGTTTCTGCTTTGC	63,9
hsa	*SOX2*	CCCACCTACAGCATGTCCTACTC	TGGAGTGGGAGGAAGAGGTAAC	63,9
hsa	*NANOG*	TTCCCTCCTCCATGGATCTG	TGTTTCTTGACTGGGACCTTGTC	63,9
hsa	*BRA*	TATGAGCCTCGAATCCACATAGT	CCTCGTTCTGATAAGCAGTCAC	59
hsa	*TBX5*	CTGTGGCTAAAATTCCACGAAGT	GTGATCGTCGGCAGGTACAAT	59
hsa	*EOMES2*	AAATGGGTGACCTGTGGCAAAGC	CTCCTGTCTCATCCAGTGGGAA	59
hsa	*TNNT2*	AAGAGGATGCTGAAGCAGAG	TGGTTTGGACTCCTCCATTG	63,9
hsa	*GATA4*	CGACACCCCAATCTCGATATG	GTTGCACAGATAGTGACCCGT	59
hsa	*NKX2-5*	AGTTTGTGGCGGCGATTAT	AGCTCAGTCCCAGTTCCA	61
hsa	*IRX4*	TTGGACTCCTGGGAACATGGACAA	ATGCTTCAGGGTATCTGGCCTCTT	63,9
hsa	*MYL2*	TGTCCCTACCTTGTCTGTTAGCCA	ATTGGAACATGGCCTCTGGATGGA	63,9
hsa	*NR2F2*	TGATGTAGCCCATGTGGAAAG	GCTGCCGGACAGTAACATATC	63,9
hsa	*KCNJ3*	TCATCAAGATGTCCCAGCCCAAGA	CACCCGGAACATAAGCGTGAGTTT	63,9

Table 5 Primers and temperatures for real-time PCR already available in the lab. Hsa, homo sapiens. HK, housekeeping

Table 6 shows reagents and quantities used for gradient PCR protocol.

REAGENT	µl / WELL
F primer (10 µM)	0,5
R primer (10 µM)	0,5
LongAmp® Taq 2X Master Mix	6,25
cDNA (100 ng)	*x*
H_2O Milli-Q RNase/DNase free	12,5-0,5-0,5-6,25-*x*
TOTAL VOLUME	**12,5**

Table 6 Reagents and quantities used for gradient PCR.

LongAmp® Taq 2X Master Mix (New England Biolabs, #M0287S) consists of Recombinant Taq DNA polymerase and Deep Vent® DNA polymerase, which enables increased amplification fidelity due to its 3'→5' exonuclease activity. Gradient PCR protocol is shown in Table 7.

PHASE	TEMPERATURE	DURATION	REPETITION
Taq Activation	95°C	30 sec	1 cycle
Denaturation	95°C	30 sec	

Annealing	58-68°C	30 sec	35 cycles
Elongation	68°C	30 sec	
Final extension	68°C	10 min	1 cycle
Conservation	4°C	As needed	

Table 7 Gradient PCR protocol

Amplicon analysis was performed by 2% agarose gel electrophoresis using 5 µl/100ml of ethidium bromide solution 10mg/ml for DNA visualization. Agarose UltraPure™ in 1X TAE buffer was used for the run. For the run, 2,5 µl Orange G (6X) in 12,5 µl samples and 7,5 µl GeneRuler DNA Ladder Mix (ThermoFischer, #SM0334) were used. BioRad Molecular Imager® Gel Doc™ XR+ with Image Lab™ Software allowed amplicons visualization after the run.

3.6.2 REAL TIME PCR

Real-time PCR was performed in 96-well plates to quantify gene expression. Thanks to the intercalation of SYBR Green into double stranded DNA (dsDNA), it is indeed possible to follow the amplification of template cDNA. From the threshold cycle (Ct value) obtained at the end of the procedure, which indicates the first cycle in which a signal above the background noise was detected, it is possible to calculate the level of gene expression relative to the house keeping one. This procedure also allows to assess the presence of aspecific PCR products from the quality of the melting curve obtained at the end of the reaction.

Reagents and amounts used for real-time PCR primer mix are shown in Table 8.

PRIMER MIX	µl / WELL
iQ SYBR Green Supermix (Bio-Rad, #1708885)	12
Primer Forward 10µM	1
Primer Reverse 10µM	1
H_2O	10
TOTAL VOLUME	24

Table 8 Reagents for real-time PRIMER mix

24 µl of Primer Mix and 1 µl of 1 ng/µl cDNA were added to each well to start the reaction with 1 ng of cDNA. A negative control consisting of Primer Mix and water was used for each Primer Mix. The plate was then analyzed in the BioRad CFX96 Touch Real-Time PCR detection system, with the protocol given in the table below. Each gene was analyzed using three biological replicates and three technical replicates, normalizing its Ct values with those of housekeeping (L7). Real time PCR protocol phases and durations are shown in Table 9.

PHASE	TEMPERATURE	DURATION	REPETITIONS
Taq Activation	95°C	10 min	1 cycle
Denaturation	94°C	15 sec	4 cycles
Annealing	58-68°C	30 sec	
Elongation	72°C	1 min	
Termination	95°C	1 min	1 cycle
Preparation	55°C	1 min	1 cycle
Melting curve	55 to 94,5°C + 0,5°C		1cycle

Table 9 Real-time PCR protocol

3.6.3 REAL TIME PCR DATA ANALYSIS

Analysis of qPCR results was performed in Excel using the ΔΔ Ct method followed by Student's t-test. In this method, the average Ct values of the gene of interest is normalized using the housekeeping gene, obtaining the Δ value. Next, to obtain the ΔΔvalue, the difference between the Δ of the control and the Δ of the experimental condition was calculated. To determine the Log2FoldChange, log2 must be calculated. To determine the power, 2 was taken to the power of the negative ΔΔCt values.

Statistical analysis was performed using Student's t-test for pairwise comparisons and ANOVA for multiple comparisons. The value 0,05 was used as the p-value threshold below which the data were considered significant.

Finally, to obtain the expression, the power value of the condition was divided by the power value of the control. Finally, Prism v.10.0.0 (GraphPad Software) was used to graphically display the data.

3.7 EVALUATION OF hiPSC-CMs BEATING PROPERTIES

Videos of beating hiPSC-CM were taken using a tablet with an ocular attached to a phase-contrast inverted microscope, at 10-fold magnification, 1280x960 resolution, and 30 fps framerate. Videos were then converted from PM4 format to AVI format with Fiji ImageJ image processing software (https://fiji.sc). Analysis was performed with the MYOCYTER macro for ImageJ, manually selecting regions of interest (ROIs). Among the different parameters generated by the software running the macro, beat time, frequency, and peak time with thresholds of 10%, 50%, and 90% were selected and exported to Excel for statistical analysis (one-way ANOVA test), considering 0,05 as the p-value threshold below which the data were considered significant. [32]

3.8 RNA-SEQUENCING VALIDATION

The selection of genes for validation started from a transcriptomic dataset already available in the lab, obtained from hearts isolated from three-month old knock-in (KI) mice harboring the homozygous or heterozygous DSG2 p.Q563* mutation, and related wild type controls. First, from a GeneOntology (GO) term dataset, produced through an R-enrichment analysis already available in the lab and showing all genes related to each term, a selection of interesting GO terms was performed based on significance, disease literature, and prior knowledge. Secondly, comparison of the selected genes with a differentially expressed gene dataset was performed to enrich the selection. The selected genes were then ranked according to several parameters, such as lowest false discovery rate (FDR; cutoff <0,05), lowest adjusted p-value, and Log2FoldChange. Finally, literature was considered an important parameter to identify genes involved in pathways involved in or potentially relevant for the disease.

In contrast, two genes (COL4A2 and CFL1) were selected from an allele-specific expression analysis based on RNA sequencing data to determine the relative expression of individual alleles due to the cis-acting regulatory system causing allelic imbalance. [29]

After validation, the same genes were also analysed in hearts harvested from 6-month-old KI mice harboring the same mutation, or controls, to study disease progression compared with 3-month-old samples.

4. RESULTS

In the first part of this thesis, two protocols for differentiation of hiPSCs into hiPSC-CMs were applied and compared in order to determine the robustness and reproducibility of obtained data.

The second part of this thesis consists of the validation of a transcriptomic dataset already available in the laboratory, obtained from hearts isolated from three- and six-months-old knock-in (KI) mice carrying the homozygous (only in 3-months-old animals) or heterozygous *Dsg2* p.Q563* mutation (corresponding to the p.Q558* mutation in human) and related wild type controls.

4.1 COMPARISON OF PROTOCOLS FOR DIFFERENTIATION OF hiPSCs INTO hiPSC-CMs

To characterize and compared the two protocols, the expression of differentiation markers, cell morphology and organization, and beating behaviour were analysed. The protocol adapted from the Goettingen Heart Research Center, Goettingen University Medical Center was denoted "A," while the protocol previously used in our lab was denoted "B."

4.1.1 DIFFERENCES IN REAGENTS USED IN THE TWO PROTOCOLS

The main differences observed between protocol A and protocol B are shown in Table 10.

DIFFERENCE	PHASE	PROTOCOL A	PROTOCOL B
B27 supplement concentration	All phases except selection	0,4 %	2 %
CHIR concentration	Mesodermal	1 µM	10 µM
Mesodermal induction molecules	Mesodermal	BMP4, ActivinA, bFGF	Human albumin
Cardiac phase duration	Cardiac	9 days (from day 3 to day 12)	5 days (from day 3 to day 8)
IWP molecule	Cardiac	IWP4	IWP2
2-mercaptoethanol concentration	Metabolic selection	0,4 mM	0,1 mM
Sodium lactate concentration	Metabolic selection	2,2 mM	4,4 mM

Table 10 Main differences observed between protocol A and protocol B.

4.1.2 DIFFERENTIATION MARKERS ANALYSIS

The differentiation of hiPSCs into hiPSC-CMs was followed and characterized by analysis of the differentiation markers shown in Table 11 and examination of cell morphology and organization. Three biological replicates and three technical replicates were performed for each marker analysis.

TIMEPOINT	DIFFERENTIATION PHASE	MARKERS
day0	pluripotency	*SOX2, OCT4, NANOG*
day3	mesodermal	*BRA, TBX5, EOMES*
day10	cardiac	*TNNT2, GATA4, NKX2.5*
day24	cardiomyocytes	*NR2F2, KCNJ3, IRX4, MYL2*

Table 11 Markers analysed with respective timing and differentiation phase.

4.1.2.1 DAY0: PLURYPOTENCY STATE

Both protocols started with hiPSCs seeding at day -4 and culturing them to day 0 with E8++ medium, when 80-90% confluence was reached for protocol A and 75-85% confluence was reached for protocol B. Figure 8 shows cells at day 0 for both protocols.

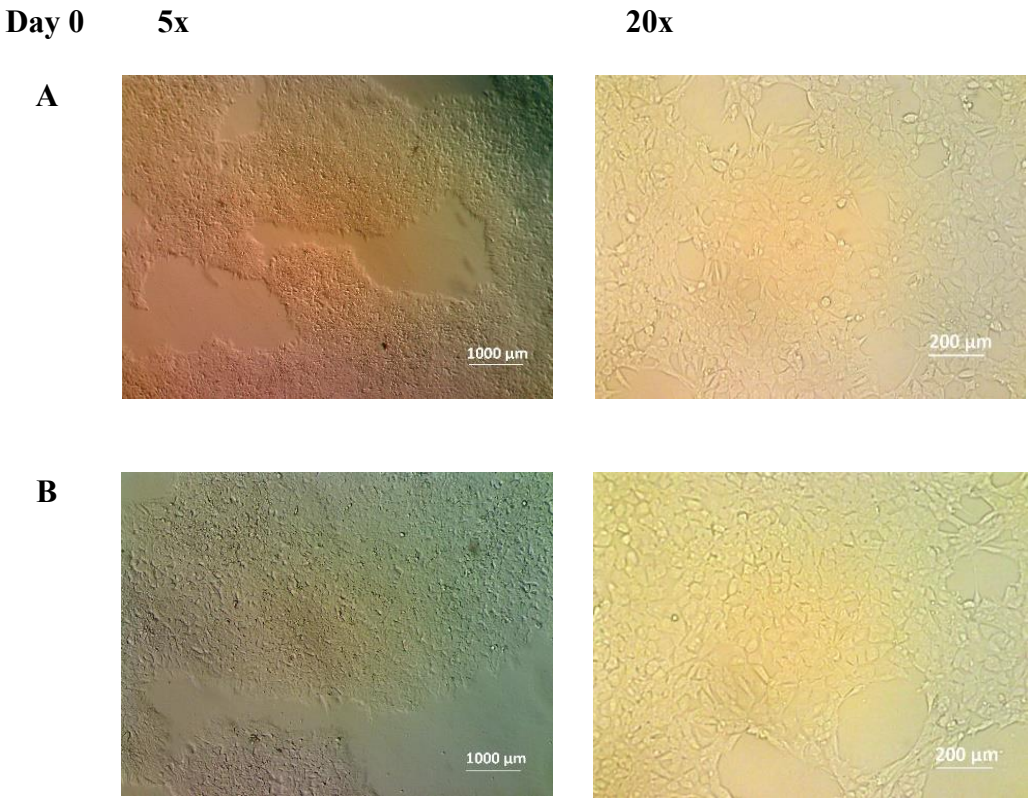


Figure 8 Light microscope images illustrating cells at day 0 of protocol A (top) and protocol B (bottom), at 5X and 20X magnification.

In both protocols, hiPSCs appeared small and mostly round at day 0, reflecting the typical morphology of pluripotent cells. At this point, the expression levels of *SOX2, OCT4* and *NANOG* were analysed as pluripotency markers. For both protocols, cells at day 10 were used to compare the expression levels of these markers as control samples in which the expression of pluripotency markers is expected to be very low because already differentiated cells. Since expression level of the markers analysed at day 10 of protocol A was not significantly different from that of protocol B, day 10A was arbitrarily chosen as the reference for the analysis.

As expected, all three pluripotency markers showed significantly higher expression level at day 0 compared to day 10, both for protocol A and protocol B. No significant differences can be appreciated in the expression level of *SOX2* and *OCT4* markers between A and B, while higher level of *NANOG* is founded in A (Figure 9)

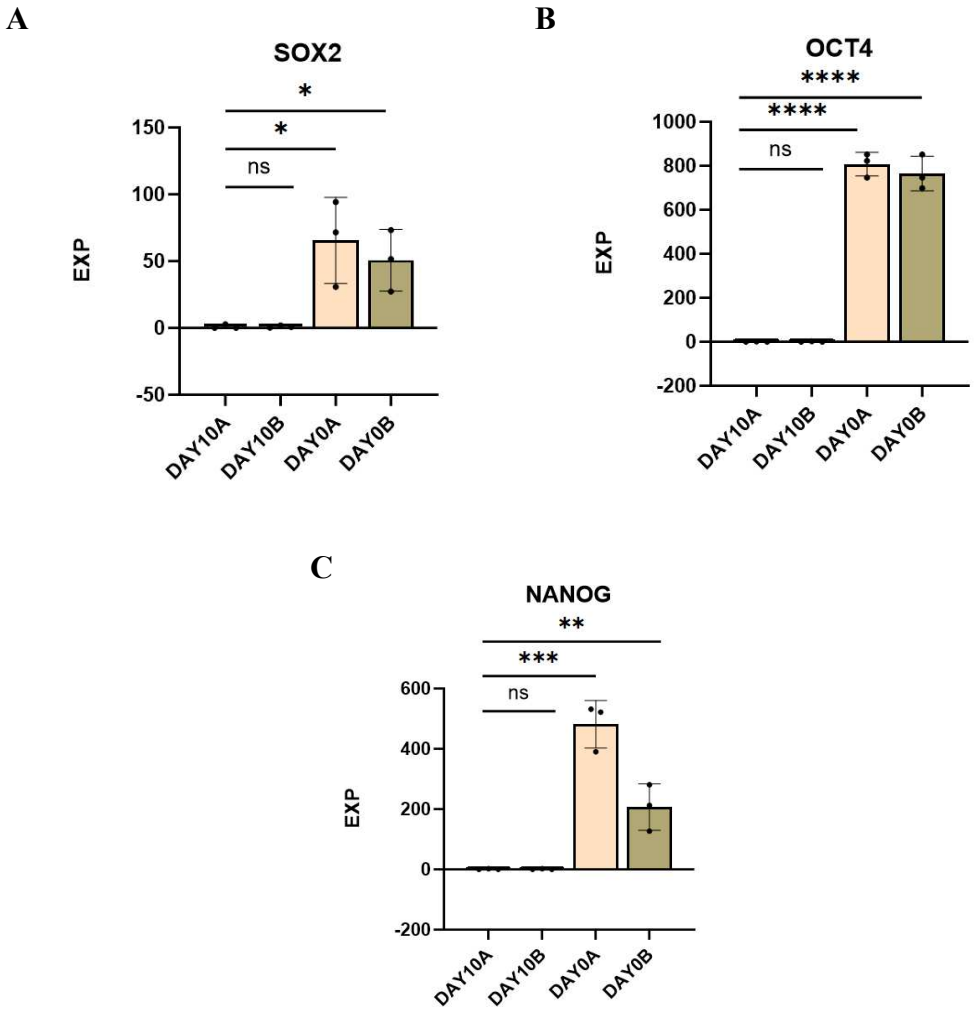


Figure 9 Pluripotency marker expression level analysis at day 0 in cells obtained using protocol A or B. Samples from day 10A were used as control. P-value of 0,05 was used as the threshold below which the data were considered statistically

*significant, with * indicating p-value<0,05, ** for p-value<0,01, *** for p-value<0,001 and **** for p-value< of 0,0001. **A** SOX1. **B** OCT4. **C** NANOG.*

4.1.2.2 DAY3: MESODERMAL STATE

For mesodermal induction, mesodermal induction medium (MIM) was used for protocol A, consisting of Basal Serum Free Medium minus insulin (BSFM(-)), 1µM CHIR (1µl of 10mM stock in 10 ml of medium), 5 ng/ml BMP4 (5 µL of 10 µg/ml), 9 ng/ml Activin A (9 µl of 10 µg/ml stock in 10 ml of medium) and 5 ng/ml bFGF (5 µL of 10 µg/ml stock in 10 ml of medium). BSFM(-) consisted of RPMI 1640 with Glutamax, 1% of 100X sodium pyruvate, 0,4% of B27 supplement minus insulin, 200 µM ASC.

For protocol B, CDM supplemented with 0,84 µl/ml (10 µM) CHIR99021 12 mM was used at day 0 to induce mesodermal differentiation. CDM consisted of RPMI 1640 with HEPES and GlutaMAX™, 2% B-27™ supplement minus insulin, 0,2 mg/ml L-Ascorbic acid 2-phosphate, 0,5 mg/ml human Albumin and 0,5% P/S. After 24 h, medium was change to CDM alone.

At Day 3, in both protocols, the cells began to become organized, creating inter-cellular structures. Cellular shape started to appear more elongated. The morphology and organization of the cells at day 3 are shown in Figure 10.

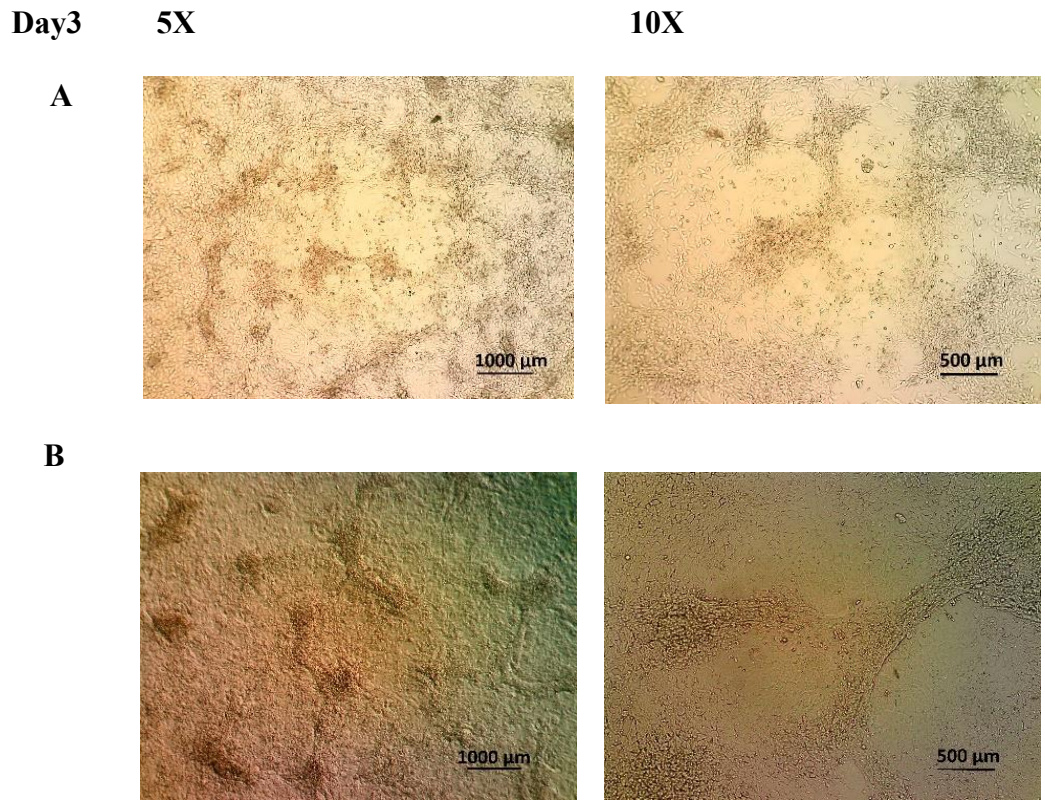


Figure 10 Light microscope images illustrating cells at day 3 of protocol A (top) and protocol B (bottom), at 5X and 10X magnification.

At day3, *BRACHYURY, TBX5* and *EOMES* expression level were checked as mesodermal markers. Day 0 was used as control for the analysis, expecting low expression level of mesodermal markers because of the pluripotency state of the cells. Since expression level of the markers analysed at day 0 of protocol A was not significantly different from that of protocol B, day 0A was arbitrarily chosen as the reference for the analysis (Figure 11).

As expected, all the three analysed markers showed significantly increased expression level at day 3 compared with day 0 in both protocols. *BRACHYURY* expression level resulted to be not significantly different between A and B, *TBX5* expression level resulted higher in A, while *EOMES* one higher in B.

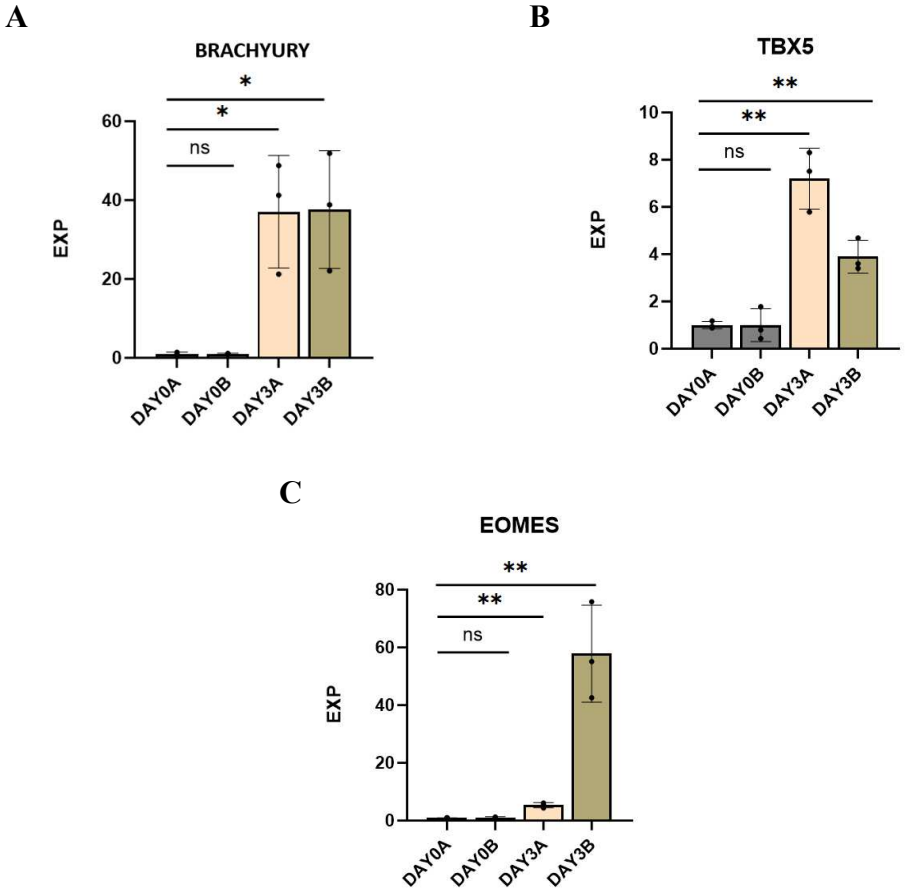


*Figure 11 Mesodermal markers expression level analysis in cells at day 3 obtained with protocol A or B, using day 0A as control. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, ** for p-value<0,01. **A** BRA. **B** OCT4. **C** EOMES.*

4.1.2.3 DAY10: CARDIAC STATE

For protocol A, to induce cardiac differentiation, from day 3 to day 12, the Cardiac Differentiation Medium (CDM) was used, consisting of Basal Serum Free Medium plus insulin (BSFM(+)) and 5µM IWP4. BSFM(+) was made of RPMI 1640 with Glutamax, 1% of 100X sodium pyruvate, 0,4% of B27 supplement, 200 µM ASC.

For protocol B, cardiac differentiation was induced from day 3 until day 8 using CDM supplemented with 1 ml/ml (5 µM) of IWP2. Then, cells were cultured from day 9 until day 13 in Cardio Culture Medium (CCM), consisting of RPMI 1640 medium with HEPES and GlutaMAX™, 2% serum-free B-27™ supplement and 0,5% P/S, for maintenance.

By day 10, the cells of both protocols showed greater inter-cellular organization. Particularly, in protocol A spherical structures were the most present, both rich in cells and cave, while in protocol B more elongated and filamentous structures appeared, with stretched single cell morphology more pronounced than in protocol A. The morphology and organization of the cells at day 10 are shown in Figure 12.

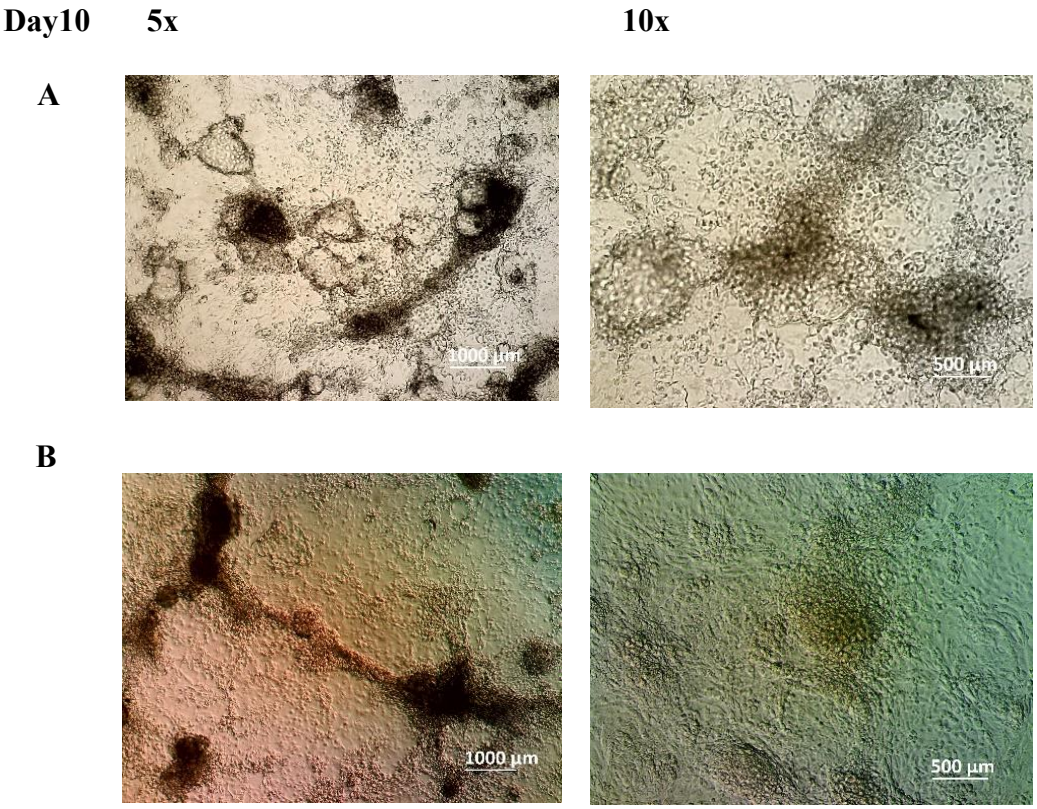


Figure 12 Light microscope images illustrating cells at day10 of protocol A (top) and protocol B (bottom), at 5X and 10X magnification.

At day 10, *TNNT2, GATA4* and *NKX2.5* expression levels were analysed as cardiac markers, using day 0 as control because of the low expression level of cardiac marker expected from the pluripotency state of the cells. Since expression level of the markers analysed at day 0 of protocol A was not significantly different from that of protocol B, day 0A was arbitrarily chosen as the reference for the analysis (figure 13).

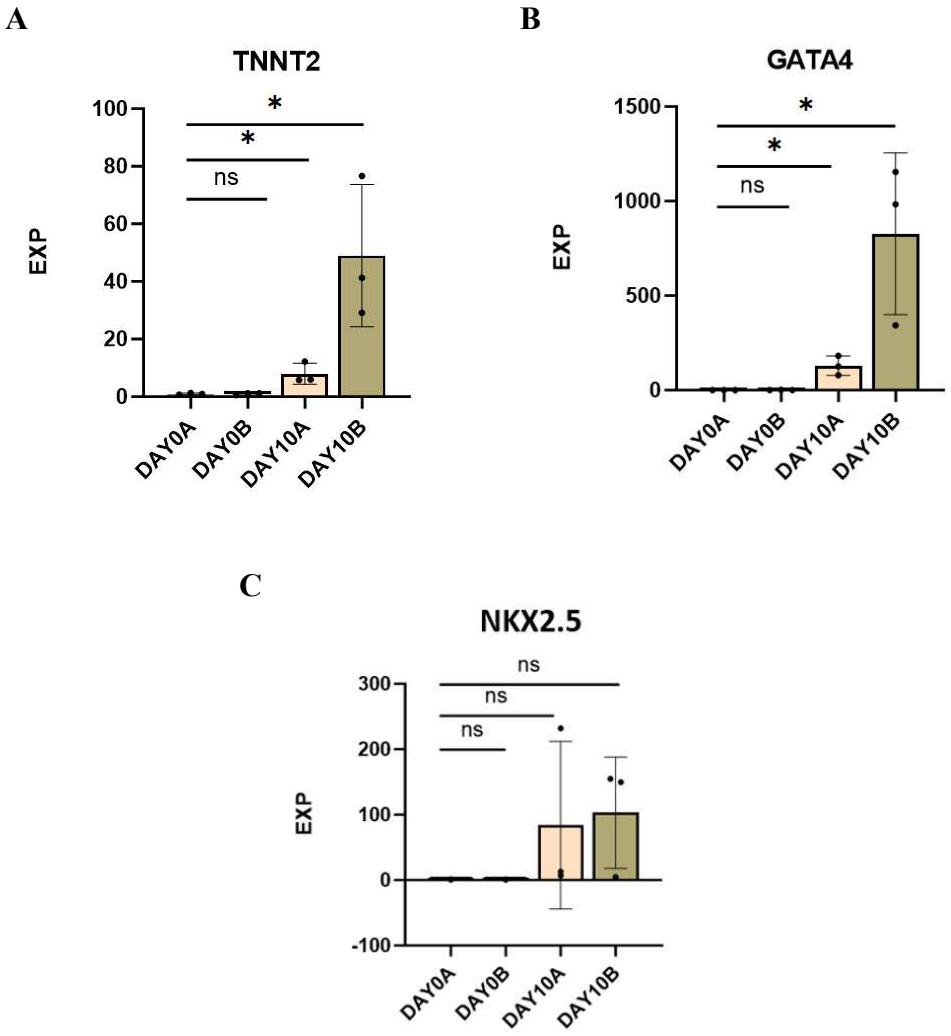


*Figure 13 Cardiac markers expression level analysis in cells at day 10 obtained with protocol A or B, using day 0A as control. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05.* ***A*** *TNNT2.* ***B*** *GATA4.* ***C*** *NKX2.5.*

Two of the three markers analysed (*TNNT2* and *GATA4*) showed a significantly increased expression level at day 10 compared with day 0 in both protocols, while the expression level of *NKX2.5* was not significantly different from control in either protocols. *TNNT2* and *GATA4* expression level appeared to be definitely higher in B than A.

4.1.2.4 DAY24: CARDIOMYOCYTES STATE

Metabolic selection was performed in protocol A from day 15 to day 20 using selection medium (SM), consisting of RPMI without glucose, sodium lactate (final concentration 2,2 mM) and 2-mercaptoethanol (final concentration of 0,1 mM), in order to increase the percentage of cardiomyocytes in culture by reducing the amount of other cell types. After selection, purified cardiomyocytes were cultured in BSFM(+) medium for 5 days (day 20 to day 24) for recovery.

Metabolic selection of protocol B encompassed days 14-9, in which cells were cultured with Cardiac Selection Medium (CSM), consisting of RPMI 1640 without glucose, 0.2% 2-mercaptoethanol, 1% P/S and 0.44% sodium lactate/HEPES solution (1 M). Finally, the medium was switched to CCM with weekly change until day 24 for recovery.

By day 24, cells from both protocols appeared to be organized into large and complex intercellular structures. In particular, the cells in protocol A appeared to form three-dimensional circular structures, joined together by filamentous segments. In protocol B, on the other hand, multiple filamentous segments creating a network-like pattern were observed. The morphology and organization of the cells at day 24 are shown in Figure 14.

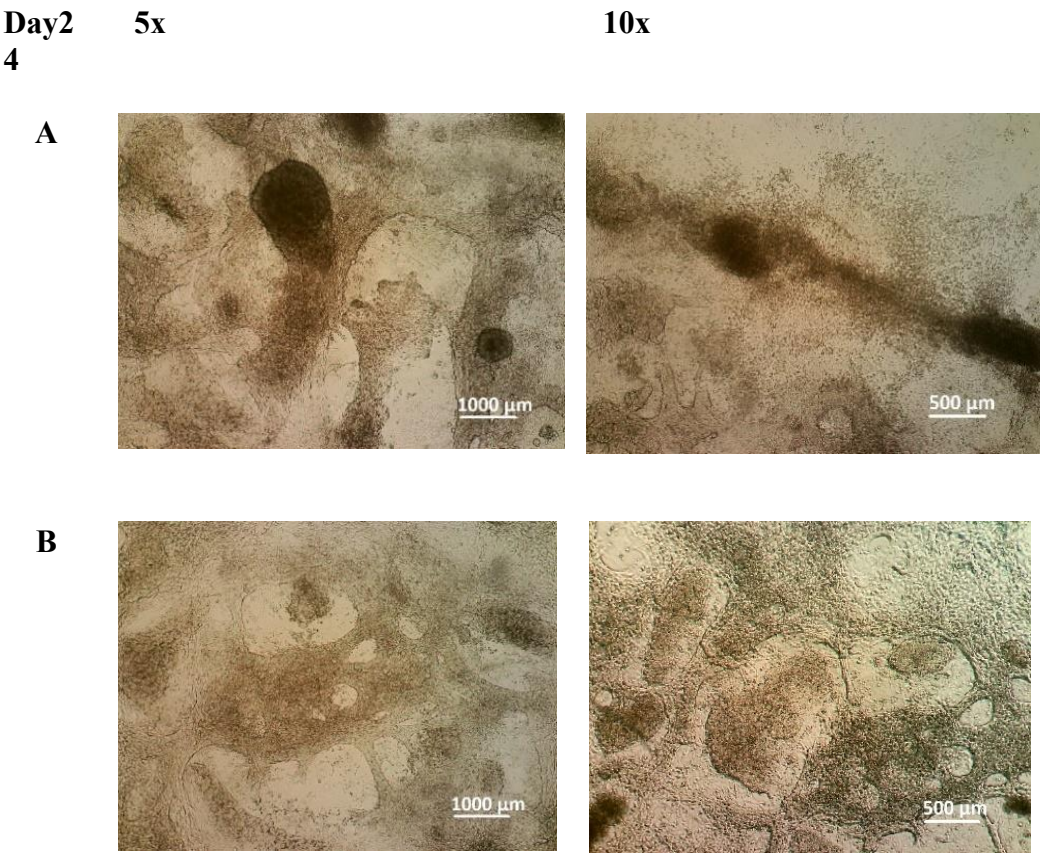


Figure 14 Light microscope images illustrating cells at day24 of protocol A and protocol B, at 5X and 10X magnification.

At day 24, *NR2F2* and *KCNJ3* expression levels were analysed as atrial cardiomyocytes markers, *IRX4* and MYL2 expression levels were analysed as ventricular cardiomyocytes makers. Cells at day 0 were used as control in the analysis, because of the low expression level of cardiomyocytes markers expected due to the pluripotency state. Since expression level of the markers analysed at day 0 of protocol A was not significantly different from that of protocol B, day 0A was arbitrarily chosen as the reference for the analysis (Figure 15).

As expected, all the four analysed markers showed significantly increased expression level at day 24 compared with day 0 in both protocols. Expression level of *NR2F2* marker appeared to be not significantly different between the two protocols; *KCNJ3* level resulted slightly higher in A; *IRX4* higher in A; *MYL2* definitely higher in B.

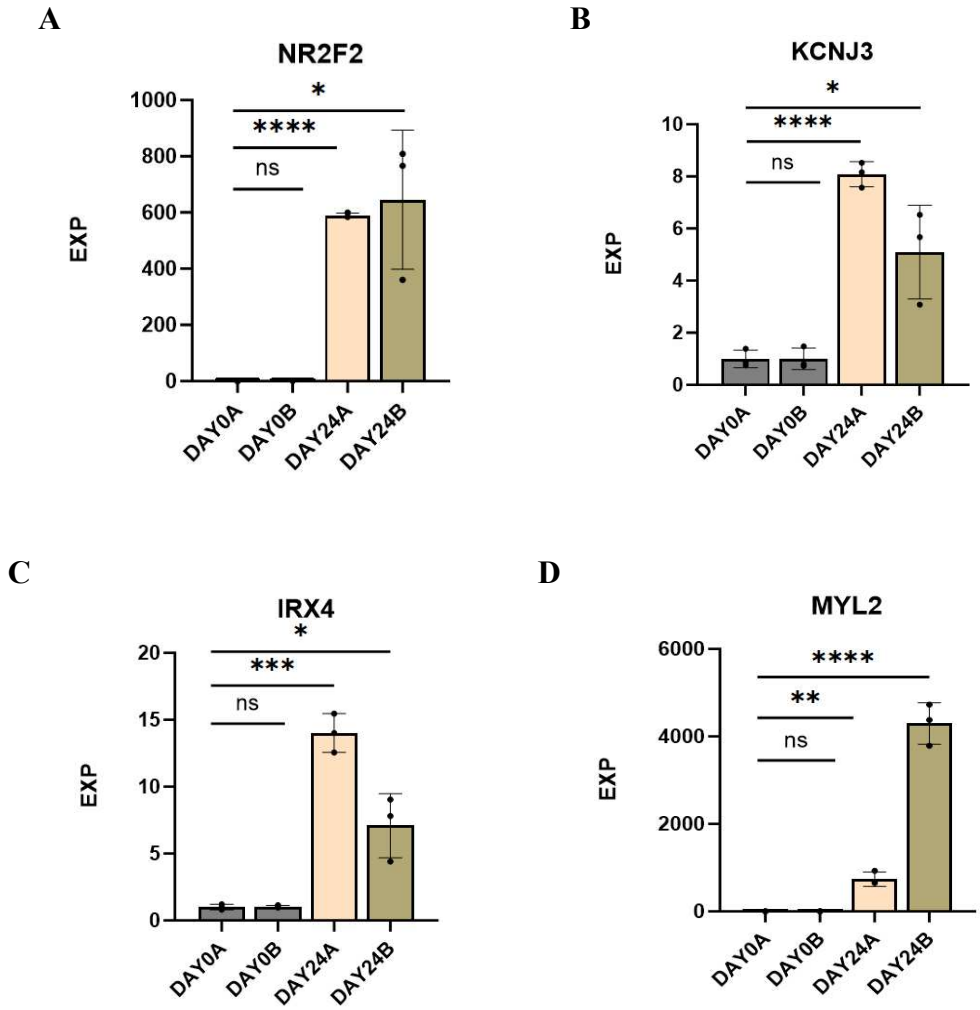


Figure 15 Atrial and ventricular cardiomyocytes markers expression level analysis at day 24 in cells obtained with protocol A or B, using day 0A as control. P-value of 0,05 was used as the threshold below which the data were considered statistically

*significant, with * indicating p-value<0,05, ** for p-value<0,01, *** for p-value<0,001 and **** for p-value< 0,0001. **A** NR2F2. **B** KCNJ3. **C** IRX4. **D** MYL2.*

4.1.3 BEATING BEHAVIOUR ANALYSIS

Beating behaviour was analysed with the MYOCYTER macro for ImageJ, manually selecting regions of interest (ROIs). Among the different parameters generated by the software running the macro, beat time, frequency, and peak time with thresholds of 10%, 50%, and 90% were selected. These thresholds represent the local minima plus 10, 50 or 90 % of the difference between the local minima and the successive maxima. These thresholds enable to highlight if amplitude of contraction is shifter or delayed, and they provide information about the width of the peak in the graph [28]. Graphic representation of the selected parameters is shown in figure 16.

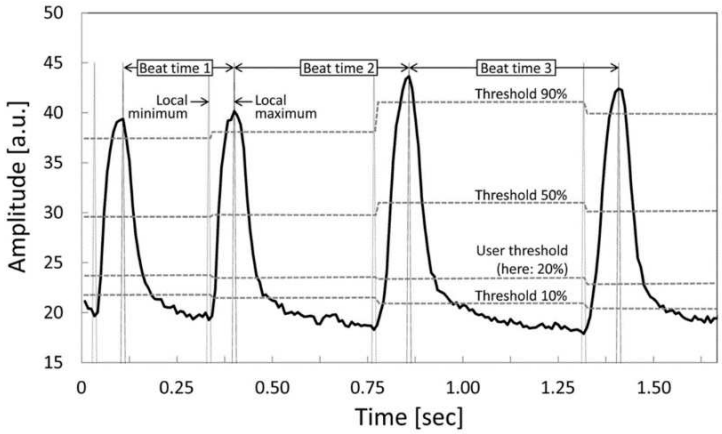


Figure 16 Graphic representation of the meaning of the parameters chosen for beating behaviour analysis: beat times, frequency, peak time threshold 10%, peak time threshold 50% and peak time threshold 90%. Figure from [28]

Cells obtained with protocol A started beating around day 10, with a ±1-day error among biological replicates. Cells obtained with protocol B cells started beating around day 8, also with a ±1-day error among biological replicates.

The analysis was performed at different timepoints, specifically immediately before metabolic selection (day 15 for protocol A and day 14 for protocol B), immediately after metabolic selection (day 20 for protocol A and day 19 for protocol B), and at the end of the protocol (day 24 for both protocols), with the expectation of appreciating some differences from the first, immature beat, to the stressful beat due to selection, and finally to the mature beat. For each timepoint and protocol, four videos were taken and analysed, one from each biological replicate. Videos of the cell beating of both protocols at the three chosen timepoints are reported in the QR codes in Figure 17.

Figure 17 Videos of beating cells generated with protocol A or protocol B beating at the different timepoints.

From the videos, it was possible to appreciate some differences in beating behaviour between the two protocols. In particular, the cells generated with protocol A appeared to beat only within the large three-dimensional intercellular structures, going from an immature and irregular beat at day 15, to an asynchronous beat between structures in the stressed cells at the end of metabolic selection (day 20), to a more synchronous beat at the end of the protocol (day 24). Even at this mature state, however, the beating was limited to areas of intercellular structures and not perfectly synchronous.

In cells obtained with protocol B, a different beating pattern can be appreciated. In particular, the cells appear to beat faster than those from protocol A and in a less spatially constrained manner, starting again from an immature and asynchronous beat at day 14, moving through a slower beat rate at the end of the metabolic selection due to the stress state of the cells, and arriving at day 24 at a synchronous, fast and expanded beat rate.

4.1.3.1 BEAT TIMES

Beat times were analysed at the three time points chosen for both protocols and represent the average time (seconds) for each complete beat, i.e., the time between a local amplitude maximum and the one immediately following it [28]. The beat times for each protocol, at the three time points, are shown in Figure 18.

The analysis of beat times in cells from protocol A showed no significant differences between the three time points, while the analysis of cells obtained with protocol B showed significantly different beat times between days 14 and 24 and between days 19 and 24, and not significantly different between days 14 and 19.

Before metabolic selection, beat time resulted to be about 2 seconds for both protocols, after the selection it appeared to reach almost 3 seconds for B while it remained around 2 seconds for A, at the end of the protocols it resulted about 1 second for B while, again, it stayed stable around 2 seconds for A.

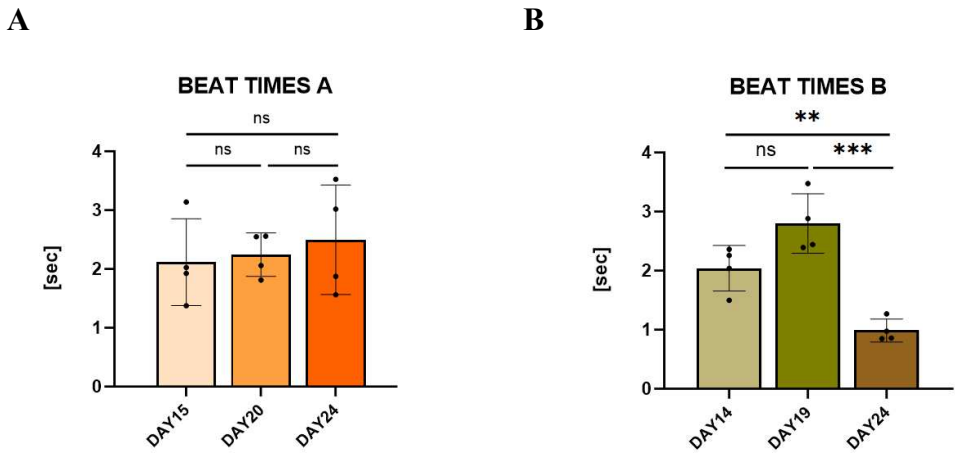


*Figure 18 Beat times of cells generated with protocol A (left) and B (right) before the metabolic selection (day 15 or 14), right after the metabolic selection (day 19 or 20) and at the end of the protocol (day 24). P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, ** for p-value<0,01, and *** for p-value<0,00. **A** Beat times analysis of protocol A. **B** Beat times analysis of protocol B.*

4.1.3.2 FREQUENCY

The frequency of the beating was analysed at the same three timepoints for both protocols. It indicates the number of beats in one second and shows an opposite trend from the beat times. The frequency of beating behaviour of cells obtained with protocol A or B at the chosen timepoints is shown in Figure 19.

In agreement with the beat times, the frequency measured in cells obtained with Protocol A did not differ significantly among the time points, while for Protocol B, significantly different frequencies could be appreciated between day 14 and day 24, and between day 19 and day 24, while day 14 and day 19 did not appear significantly different.

Before the metabolic selection, cells obtained with both protocols A and B showed beating frequency around 0,5 beat per second. After the selection, the beating frequency of cells obtained with both protocols A and B slightly decreased. Finally, at the end of the protocol, beating frequency of cells obtained with protocol A stayed stable around 0,5 beat per second, while beating frequency of cells obtained with protocol B increased to 1 beat per second.

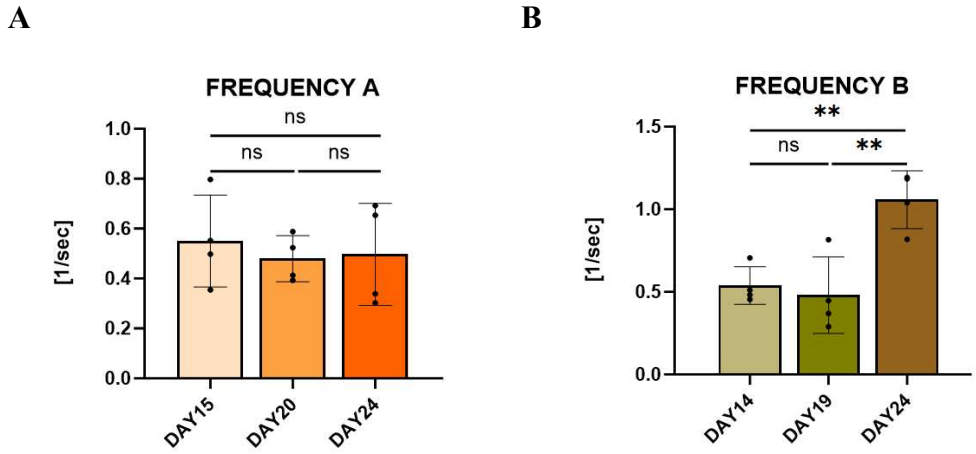


*Figure 19 Frequency of cells obtained with protocol A (left) or B (right) beating before the metabolic selection (day 14 or 15), right after the metabolic selection (day 19 or 20) and at the end of the protocol (day 24). P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, and ** for p-value<0,01. **A** Beat frequency analysis of protocol A. **B** Beat frequency analysis of protocol B.*

4.1.3.3 PEAK TIME THRESHOLD 10%

The 10% peak time threshold was analysed at the three time points for both protocols and represents the local minima plus 10% of the difference between the local minima and the successive maxima [28]. The 10% threshold peak times are shown in Figure 20 for both protocols at the chosen time points.

Cells obtained with protocol A showed no significant differences between the time points analysed, while cells obtained with protocol B showed significantly different 10% peak time thresholds between days 14 and 24 and between days 19 and 24, with no significant differences between days 14 and 19.

Before the metabolic selection, cells obtained with protocol A showed a peak time threshold 10% abour 0,7 seconds, and cells obtained with protocol B about 1,4 seconds. After the metabolic selection, this parameter remained almost constant in cells obtained with protocol A, while it increased to about 1,6 seconds in cells obtained with protocol B. Finally, at the end of the protocols, the 10% peak time threshold increased to about 1,2 seconds in cells obtained with protocol A, while it decreased to about 0,7 seconds in cells obtained with protocol B.

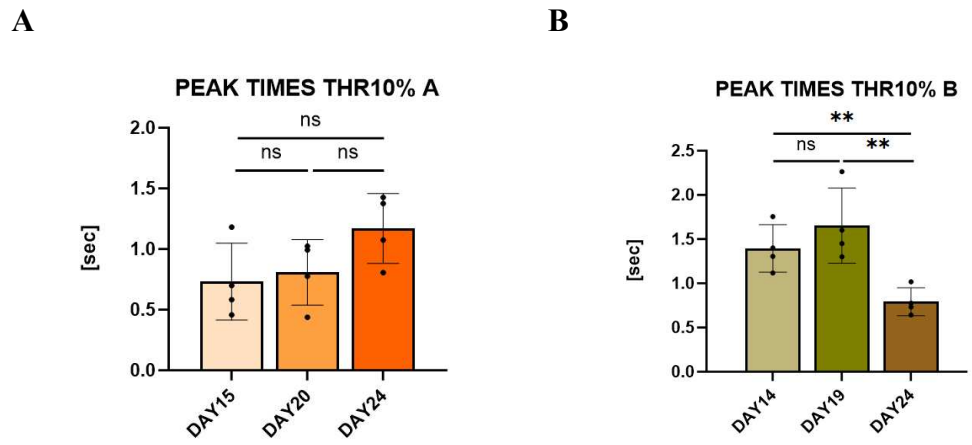


*Figure 20 Peak times threshold10% in cells obtained with protocol A (left) or B (right) before the metabolic selection (day 14 or 15), right after the metabolic selection (day 19 or 20) and at the end of the protocol (day 24). P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, and ** for p-value<0,01. **A** Peak times threshold10% analysis of protocol A. **B** Peak times threshold10% analysis of protocol B.*

4.1.3.4 PEAK TIME THRESHOLD 50%

The 50% peak time threshold was checked at the three chosen time points for both protocols and represents the local minima plus 50% of the difference between the local minima and the successive maxima [28]. The 50% threshold peak times are shown in Figure 21 in cells obtained with protocols A or B at the chosen time points.

In cells obtained with Protocol A, significantly different 50% threshold peak times can be appreciated between days 15 and 24, and between days 20 and 24, with no significant differences between days 15 and 20. Cells obtained with protocol B showed significantly different peak times threshold50% only between days 14 and 24.

Before the metabolic selection, the 50% peak time thresholds appeare to be around 0,2 seconds in cells obtained with protocol A, while it appeared to be around 0,8 seconds in cells obtained with protocol B. This parameter, after the metabolic selection, increased to 0,3 seconds in cells obtained with protocol B, while it remained stable in cells obtained with protocol B. Finally, at the end of the protocol, cells obtained with protocol A showed a peak time threshold 50% of about 0,6 seconds, while cells obtained with protocol B of about 0,5 seconds.

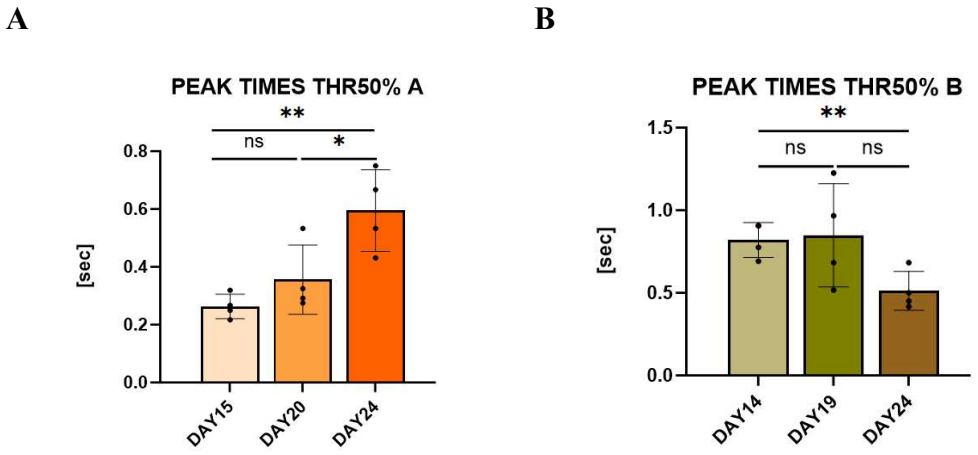


*Figure 21 Peak times threshold50% in cells obtained with protocol A (left) or B (right) before the metabolic selection, right after the metabolic selection and at the end of the protocols. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, and ** for p-value<0,01. **A** Peak times threshold50% analysis of protocol A. **B** Peak times threshold50% analysis of protocol B.*

4.1.3.5 PEAK TIME THRESHOLD 90%

The 90% threshold peak times were analysed at the three chosen time points for both protocols and represents the local minima plus 90% of the difference between the local minima and the successive maxima [28]. The 90% threshold peak times are shown in Figure 22 for protocols A and B at the chosen time points.

In cells obtained with Protocol A, significant differences can be appreciated in the 90% peak times between days 15 and 24, and between days 20 and 24, while cells obtained with Protocol B showed significant differences in the 90% threshold peak times only between days 14 and 24.

Before the metabolic selection, 90% peak time threshold appeared to be around 0,1 second in cells obtained with protocol A, while it appeared around 0,4 seconds in cells obtained with protocol B. After the selection, this parameter slightly increased in cells obtained with protocol A, while it remained almost stable in cells obtained with protocol B. Finally, at the end of the protocol, cells obtained with protocol A showed a 90% peak time threshold of about 0,3 seconds, while cells obtained with protocol B of about 0,2 seconds.

A **B**

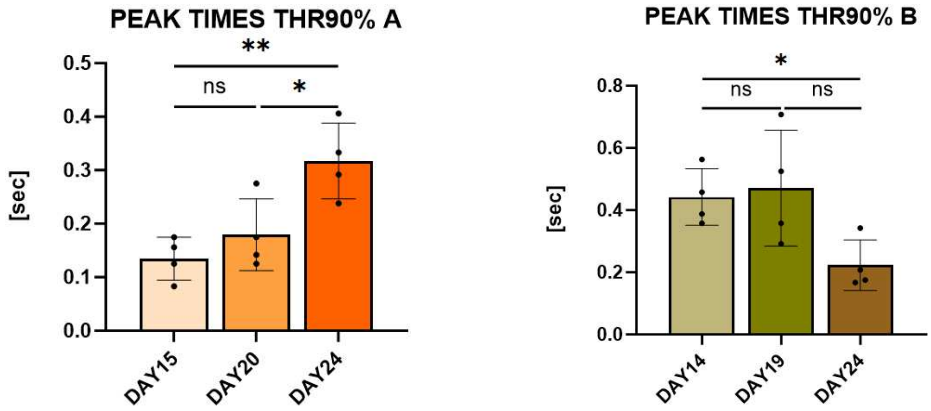


*Figure 22 Peak times threshold90% in cells obtained with protocol A (left) or B (right)before the metabolic selection (day 15 or 14), right after the metabolic selection (day 20 or 19) and at the end of the protocol (day 24). P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, and ** for p-value<0,01. **A** Peak times threshold90% analysis of protocol A. **B** Peak times threshold90% analysis of protocol B.*

4.2 RNA-SEQ DATASET VALIDATION

4.2.1 BACKGROUND

The transcriptomic dataset already available in the lab was obtained from RNA-seq performed in the heart of 3- and 6-months old (m.o.) knock-in mice harboring the p-Q653* mutation, corresponding to the human p-Q558* DSG2 mutation associated with ACM. RNA samples from both 3 m.o. heterozygous and homozygous knock-in mice were used, while only samples from 6 m.o. heterozygous were available. Initially, three different transcriptomic datasets were generated: the first included differentially expressed genes (DEG) found in 3 m.o. heterozygous mice compared with their wild type controls, the second reported DEG found in 3 m.o. homozygous samples compared with their wild type controls, and the third included DEG found in 6 m.o. heterozygous samples compared with their wild type controls. The final dataset was obtained by overlapping the previous three. For each transcriptomic dataset, a list of GeneOntology (GO) terms associated with the DEG was available and it was used in this thesis as the first tool for gene selection for validation by real-time-PCR.

4.2.2 SELECTION OF GENES TO BE VALIDATED

The final list of DEG obtained merging the two datasets from 3 m.o. animals and the one including DEG found in 6 m.o. heterozygous mice reported few and not-well characterized sequences. Therefore, the selection of genes to be validated was

first based on lists of GeneOntology (GO) terms associated with the genes reported in the 3 m.o. datasets, focusing on those most relevant for the pathology of the disease. In particular, we selected classes of interest that may be interesting for the disease, such as mitochondria, apoptosis, inflammation, heart development and cell adhesion. The selected GO terms and relative significances and classes are shown in Table 12.

PATHWAY	PADJ	CLASS
GOCC_RESPIRASOME	3,19156E-16	Mitochondrion
GOBP_OXIDATIVE_PHOSPHORYLATION	1,35783E-16	Mitochondrion
GOCC_MITOCHONDRION	5,58162E-29	Mitochondrion
GOCC_MITOCHONDRIAL_ENVELOPE	4,62136E-23	Mitochondrion
GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX	9,93209E-18	Mitochondrion
GOBP_AEROBIC_RESPIRATION	3,88246E-16	Mitochondrion
GOBP_MITOCHONDRION_AUTOPHAGY	0,035593654	Mitochondrion
GOBP_BIOLOGICAL_ADHESION	5,69411E-05	Adhesion
GOBP_HOMOTYPIC_CELL_CELL_ADHESION	2,27634E-05	Adhesion
GOBP_VENTRICULAR_CARDIAC_MUSCLE_TISSUE_DEVELOPMENT	0,001130543	Development
GOBP_HEART_DEVELOPMENT	0,001724443	Development
GOBP_NEGATIVE_REGULATION_OF_CELL_DEATH	0,005158058	Apoptosis
GOBP_NEGATIVE_REGULATION_OF_INTTRINSIC_APOPTOTIC_SIGNALING_PATHWAY	0,017949626	Apoptosis
GOBP_PROGRAMMED_CELL_DEATH	2,12783E-07	Apoptosis
GOBP_INFLAMMATORY RESPONSE	0,000227936	Inflammation
GOBP_RESPONSE_TO_OXYGEN_CONTAINING_COMPOUND	1,36691E-05	Inflammation

Table 12 Selected GO terms and classes of interest. PADJ, p-value adjusted. GOBP, GO term biological process. GOCC, GO term cellular component.

Starting from the selected GO terms, the choice of genes for real-time-PCR validation was based on the different expression level compared with the wild type controls reported in the transcriptomic dataset. Specifically, genes were selected first based on the False discovery rate (FDR), using FDR<0.05 as the threshold above which genes were discarded. The FDR is defined as an indicator of the proportion of false positives (type I error) in the test. Then, a second selection was performed based on the expression level predicted by the dataset, listing the genes based on the logFC of their expression (from higher to lower) compared to wild type controls.

The third and final criteria for gene selection took into account literature information regarding function, expression location, and collocation or involvement of the encoded protein in pathways of potential interest for ACM pathology.

At the end of this selection, 25 genes were chosen for real-time-PCR validation.

Moreover, two additional genes (*Cfl1* and *Col4a2*), were selected from an allele-specific expression analysis based on RNA sequencing data to determine the relative expression of individual alleles due to the cis-acting regulatory system causing allelic imbalance [29]. The final list then consisted of 27 genes for real-time-PCR validation.

4.2.3 VALIDATION OF SELECTED DEG BY RT-PCR

As reported in Table 13, with the exception of *Dsg2*, *Des* and *Tmem88*, all other genes showed opposite trends compared with the one reported in the dataset in at least one of the three sample groups.

GENE	HETEROZYGOUS				HOMOZYGOUS	
/	3 M.O. (Log2FoldChange)		6 M.O. (Log2FoldChange)		3 M.O. (Log2FoldChange)	
/	RNA-seq	qPCR (P)	RNA-seq	qPCR (P)	RNA-seq	qPCR (P)
Dsg2	-0,78	-0,71 (0,035)	-0,92	-1,11 (0,01)	-2,46	-1,89 (0,0022)
Arhgap5	0,62	0,01 (0,92)	-0,07	0,38 (0,112)	0,70	0,8 (0,045)
Cfl1	-0,35	0,03 (0,805)	-0,03	-0,32 (0,805)	-0,38	1,01 (0,013)
Col4a2	-0,47	-0,41 (0,403)	-0,15	-1,21 (0,008)	-0,44	1,61 (0,019)
Des	-0,50	-0,01 (0,945)	-0,12	-0,18 (0,479)	-0,53	-0,08 (0,79)
Enc1	-0,17	0,19 (0,58)	0,22	-1,52 (0,012)	-0,83	1,39 (0,008)
Etfb	-1,18	-0,3 (0,54)	0,03	-0,81 (0,024)	-1,36	0,11 (0,57)
Gpx1	-0,48	0,35 (0,243)	0,03	-0,55 (0,023)	-0,57	0,92 (0,043)
Hrc	-0,09	-0,13 (0,28)	0,06	-0,9 (0,0002)	-0,83	0,61 (0,044)
Isg20	-1,85	-0,91 (0,192)	0,14	-1,59 (0,034)	-2,09	-0,67 (0,257)
Lars2	-10,81	1,23 (0,027)	0,68	1,01 (0,019)	-14,15	3,09 (0,0049)
Litaf	-0,64	0,02 (0,91)	-0,24	-2,07 (0,027)	-0,74	0,74 (0,122)
Loxl1	0,96	-0,17 (0,483)	-0,08	-0,48 (0,096)	-0,64	-0,32 (0,136)
Mbd3	-0,43	-0,02 (0,948)	0,04	-0,88 (0,002)	-0,68	0,55 (0,0035)
Myl2	-0,51	-0,25 (0,37)	-0,04	0,06 (0,64)	-0,69	-0,28 (0,037)
Ndufb10	-0,72	-0,13 (0,845)	0,002	-0,95 (1,24E-5)	-1,00	0,30 (0,285)
Noc2l	-0,46	-0,21 (0,39)	0,09	-0,58 (0,051)	-0,48	0,76 (0,005)
Nol3	-0,59	0,05 (0,697)	0,04	-0,42 (0,0264)	-0,61	0,51 (0,02)
Ppp1r13l	-0,61	0,33 (0,0474)	0,02	-0,44 (0,006)	-0,64	1,33 (0,0098)
Rb1cc1	0,40	-0,11 (0,719)	-0,05	-0,25 (0,0062)	0,49	0,61 (0,0085)
Sap18	-0,45	0,01 (0,901)	-0,03	-0,88 (0,0012)	-0,57	0,51 (0,0014)
Sik1	-0,33	1,46 (0,272)	0,25	0,07 (0,731)	-1,36	-0,22 (0,86)
Slc2a3	-1,89	-0,6 (0,276)	-0,6	0,63 (0,217)	-1,24	-0,64 (0,302)
Sod1	-0,46	-0,13 (0,225)	0,13	0,33 (0,144)	-0,63	0,64 (0,0093)
Tgfb1	-0,54	-0,06 (0,855)	-0,18	-1 (0,0369)	-0,62	0,16 (0,736)
Tmem88	-0,70	-0,84 (0,032)	-0,13	-1,81 (0,043)	-0,81	-1,07 (0,013)
Txn2	-0,32	-0,12 (0,383)	-0,06	-0,85 (0,153)	0,52	0,9 (0,0009)

Table 13 Expression level (Log2FoldChange) obtained from the transcriptomic dataset and by real-time PCR of the selected genes, divided by sample groups: 3m.o. heterozygous vs WT mice, 6 m.o. heterozygous vs WT mice, 3 m.o. homozygous vs WT mice. P, p-value.

Results are therefore presented based on the statistical significance in the heterozygous samples, highlighting disease progression between 3 m.o. and 6 m.o. mice. Specifically, the genes were divided into three groups: (i) genes deregulated in both 3- and 6-mo. Mice (ii) genes significantly deregulated only in 6 m.o. mice (iii) genes that did not show a statistically significant different expression level compared to WT controls at either 3 or 6 months of age.

4.2.4 GENES DIFFERENTIALLY EXPRESSED IN 3 M.O. HETEROZYGOUS AND HOMOZYGOUS MICE

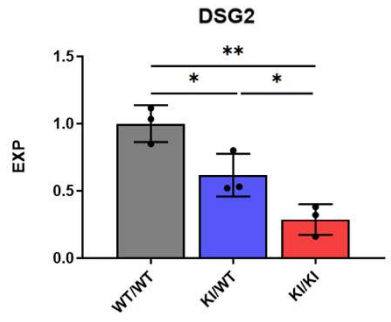
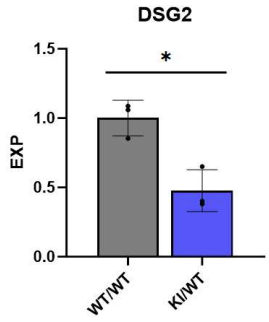
qPCR results of DEG analysed in samples from 3 m.o. heterozygous mice are indicated in Figure 23.

As expected, because of the non-sense mutation, the *Dsg2* gene showed significantly lower expression levels in both heterozygous and homozygous samples in the 3 m.o. set and in 6 m.o heterozygous mice, compared to WT, confirming the expression trend reported in the dataset.

Lars2 resulted overexpressed compared to WT in all three samples groups: 3 mo. homozygous, 3 m.o. heterozygous, and 6 m.o. heterozygous. The dataset reported under-expression in both homozygous and heterozygous 3 m.o. samples and over-expression in heterozygous 6 m.o. ones.

Ppp1r13l resulted overexpressed in both homozygous and heterozygous 3 m.o. samples, while it appeared under-expressed in heterozygous 6 mo. samples. All three cases showed opposite tren compared to the one reported in the dataset.

Tmem88 resulted under-expressed in all three samples groups: 3 mo. homozygous, 3 m.o. heterozygous, and 6 m.o. heterozygous. All three cases confirmed the trend reported in the dataset.

3 M.O. SAMPLES	6 M.O. SAMPLES
	

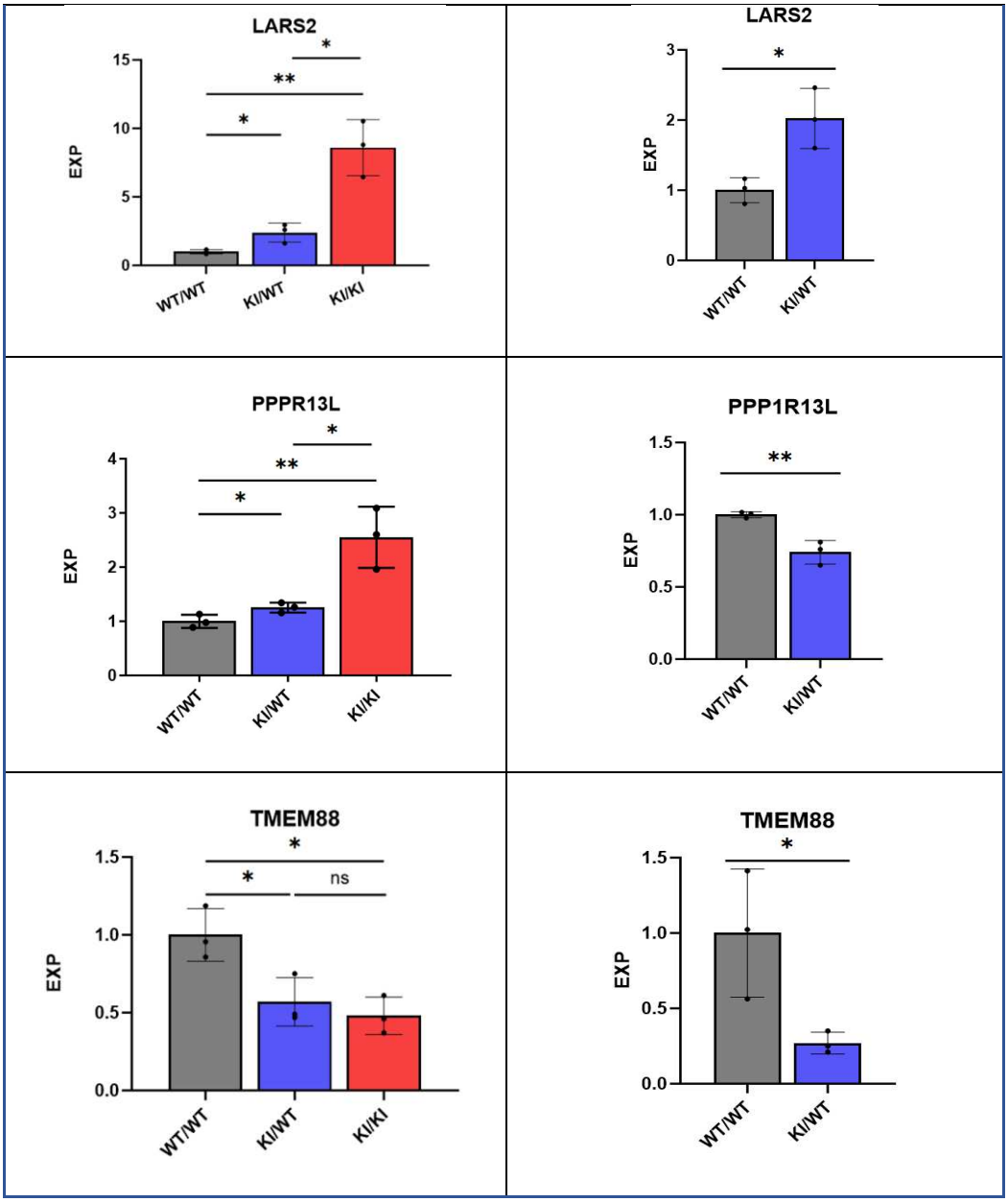


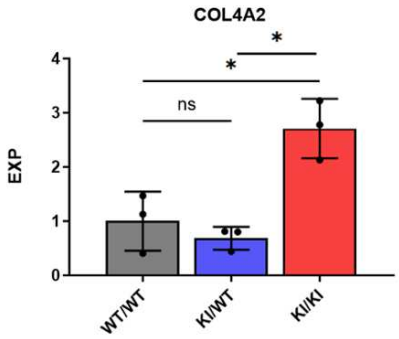
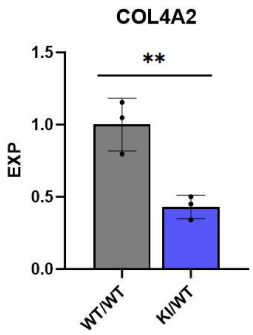
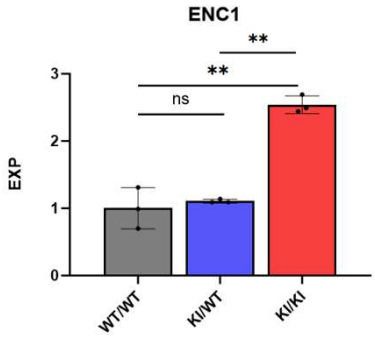
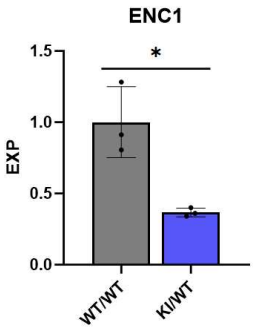
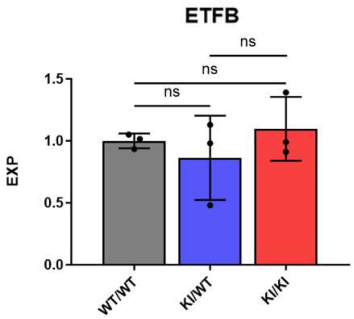
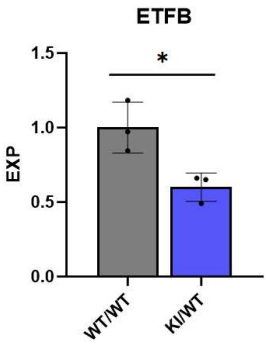
*Figure 23 Expression levels of DEG analysed in 3m.o. heterozygous and homozygous mice (left), as well as in 6m.o heterozygous mice (right) compared with WT controls. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, ** for p-value<0,01, *** for p-value<0,001 and **** for p-value< of 0,0001.*

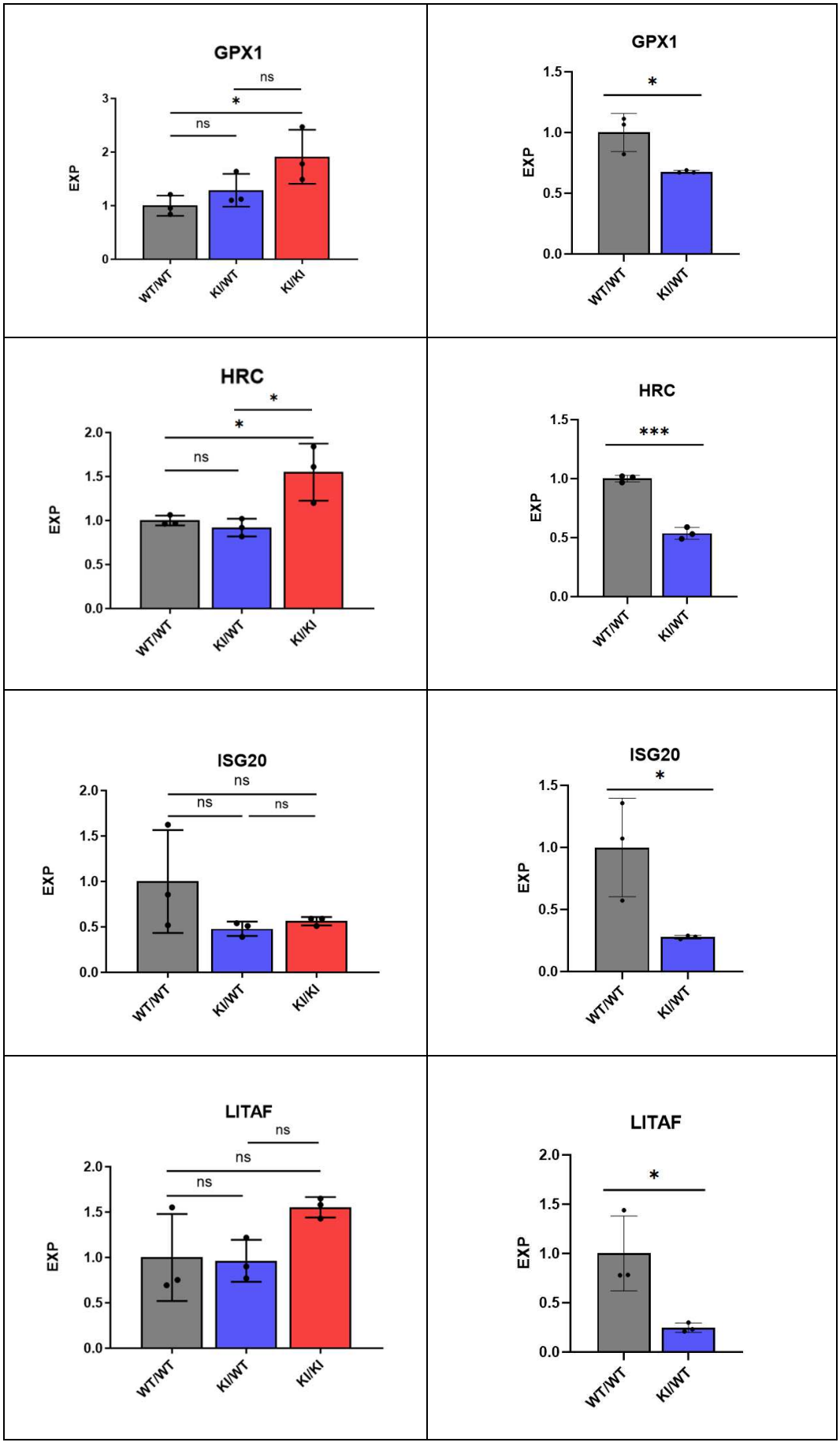
4.2.5 GENES SIGNIFICANTLY DEREGULATED IN 6 M.O. HETEROZYGOUS MICE

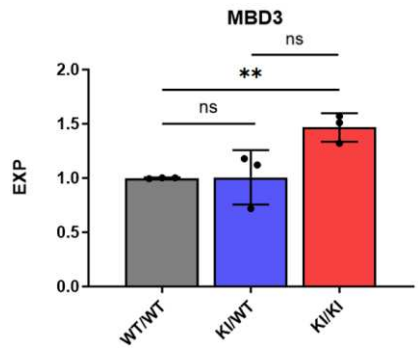
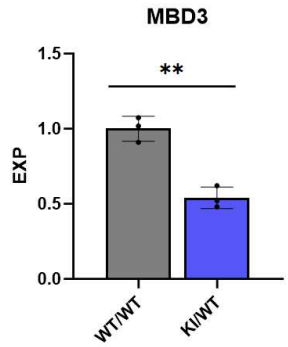
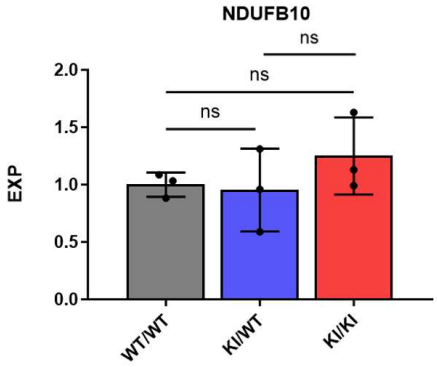
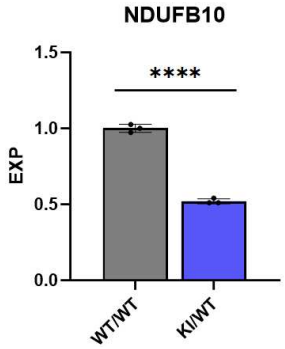
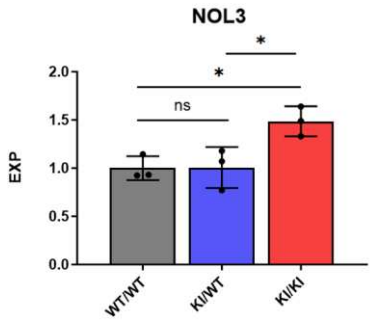
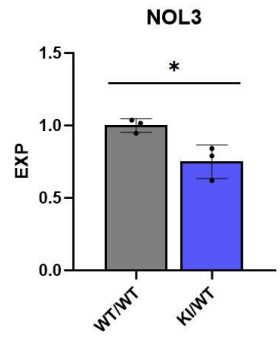
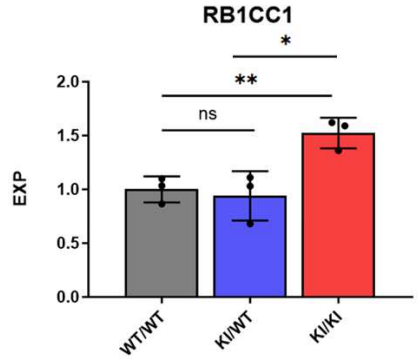
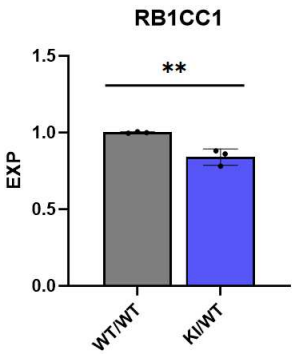
13 genes were not significantly deregulated in 3 m.o. heterozygous mice, but resulted significantly altered in 6m.o. heterozygous mice, compared to WT controls, namely *Col4A2, Enc1, Etfb, Gpx1, Hrc, Isg20, Litaf, Mbd3, Ndufb10, Nol3, Rb1cc1, Sap18, Tgfb.* They are shown in Figure 24.

Some of these genes showed a non-significant over-expression in heterozygous 3 m.o. samples and resulted under-expressed in heterozygous 6 m.o. ones, namely *Enc1, Gpx1*, and *Tgfb.* All these 13 genes resulted to be under-expressed in heterozygous 6 m.o. samples.

Col4A2, Tgfb resulted confirmed the expression trend reported in the dataset with exception made for homozygous 3 m.o. samples. *Rb1cc1* confirmed the trend with exception made for heterozygous 3 m.o. samples. *Hrc, Isg20* confirmed the dataset with exception made for heterozygous 6 m.o. samples. *Litaf, Sap18* confirmed the trend only for heterozygous 6 m.o. samples. *Etfb, Mbd3, Ndufb10* confirmed the dataset only for heterozygous 3 m.o. samples. *Enc1, Gpx1, Nol3* in all three samples groups showed opposite trend than the reported one.

3 M.O. SAMPLES	6 M.O. SAMPLES
	
	
	


GPX1
ns
*
ns
EXP
3
2
1
0
WT/WT
KI/WT
KI/KI
GPX1
*
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
HRC
*
*
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
HRC
***
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
ISG20
ns
ns
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
ISG20
*
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
LITAF
ns
ns
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
LITAF
*
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT

 MBD3 ns ** ns EXP 2.0 1.5 1.0 0.5 0.0 WT/WT KI/WT KI/KI	 MBD3 ** EXP 1.5 1.0 0.5 0.0 WT/WT KI/WT
 NDUFB10 ns ns ns EXP 2.0 1.5 1.0 0.5 0.0 WT/WT KI/WT KI/KI	 NDUFB10 **** EXP 1.5 1.0 0.5 0.0 WT/WT KI/WT
 NOL3 * * ns EXP 2.0 1.5 1.0 0.5 0.0 WT/WT KI/WT KI/KI	 NOL3 * EXP 1.5 1.0 0.5 0.0 WT/WT KI/WT
 RB1CC1 * ** ns EXP 2.0 1.5 1.0 0.5 0.0 WT/WT KI/WT KI/KI	 RB1CC1 ** EXP 1.5 1.0 0.5 0.0 WT/WT KI/WT

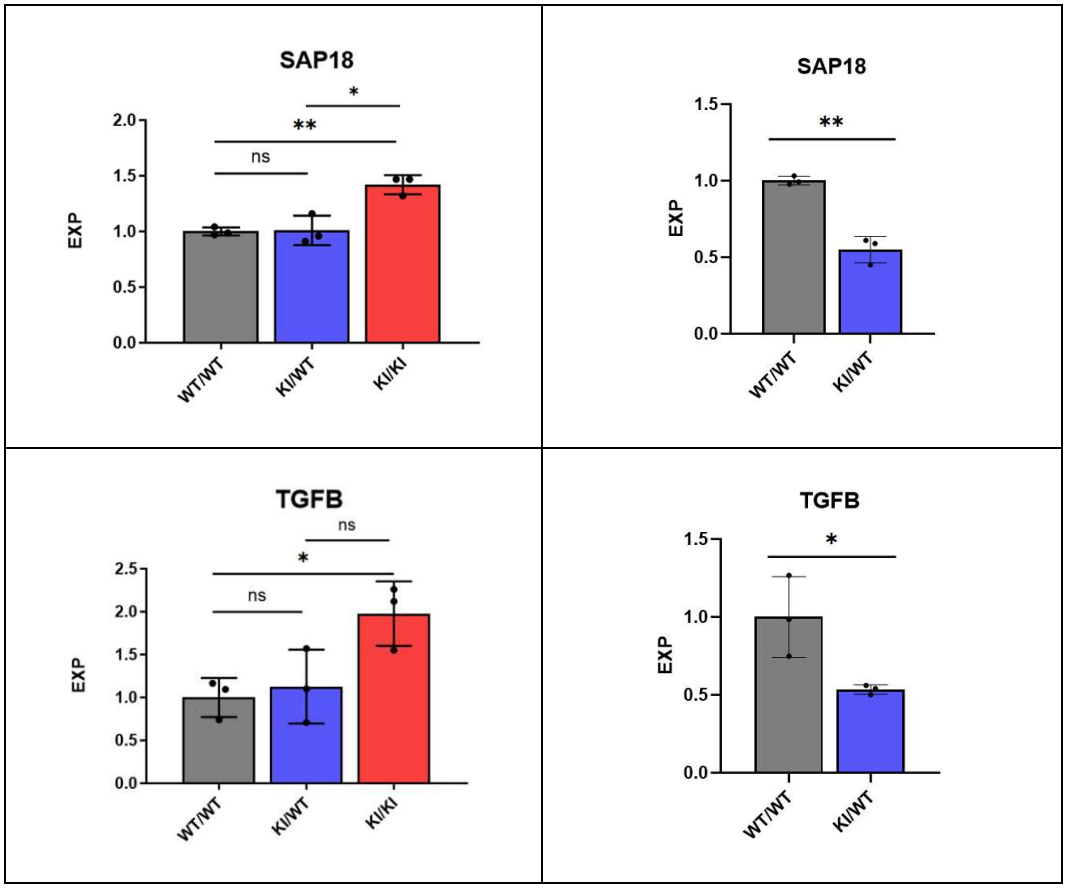


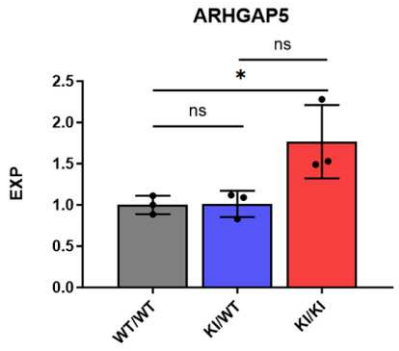
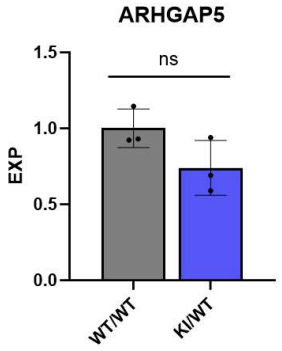
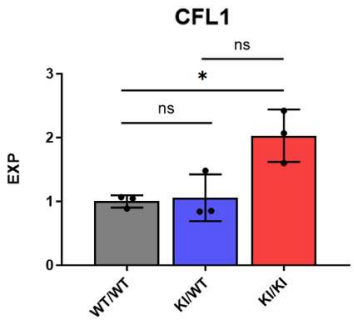
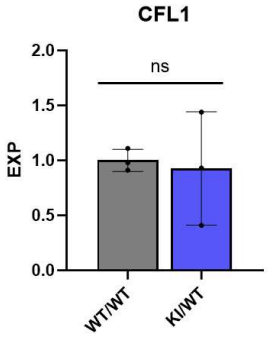
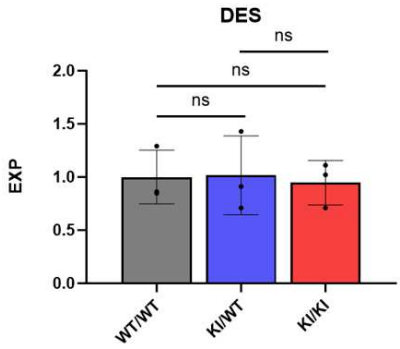
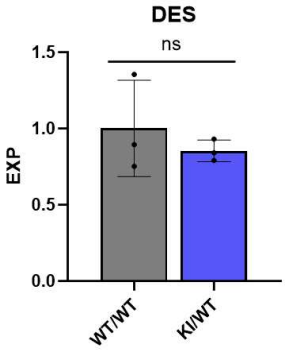
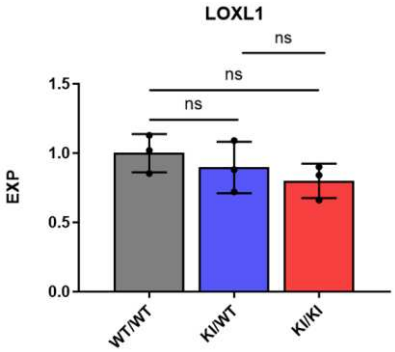
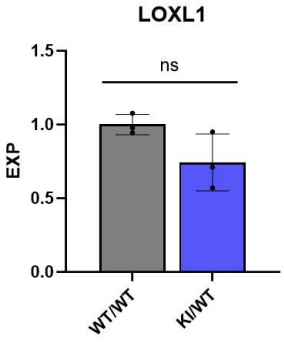
*Figure 24 Expression levels of genes that were not significantly deregulated in 3m.o. heterozygous mice (left) but resulted significantly altered in 6m.o. heterozygous mice (right), and their respective expression level in 3 m.o. homozygous mice. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, ** for p-value<0,01, *** for p-value<0,001 and **** for p-value< of 0,0001.*

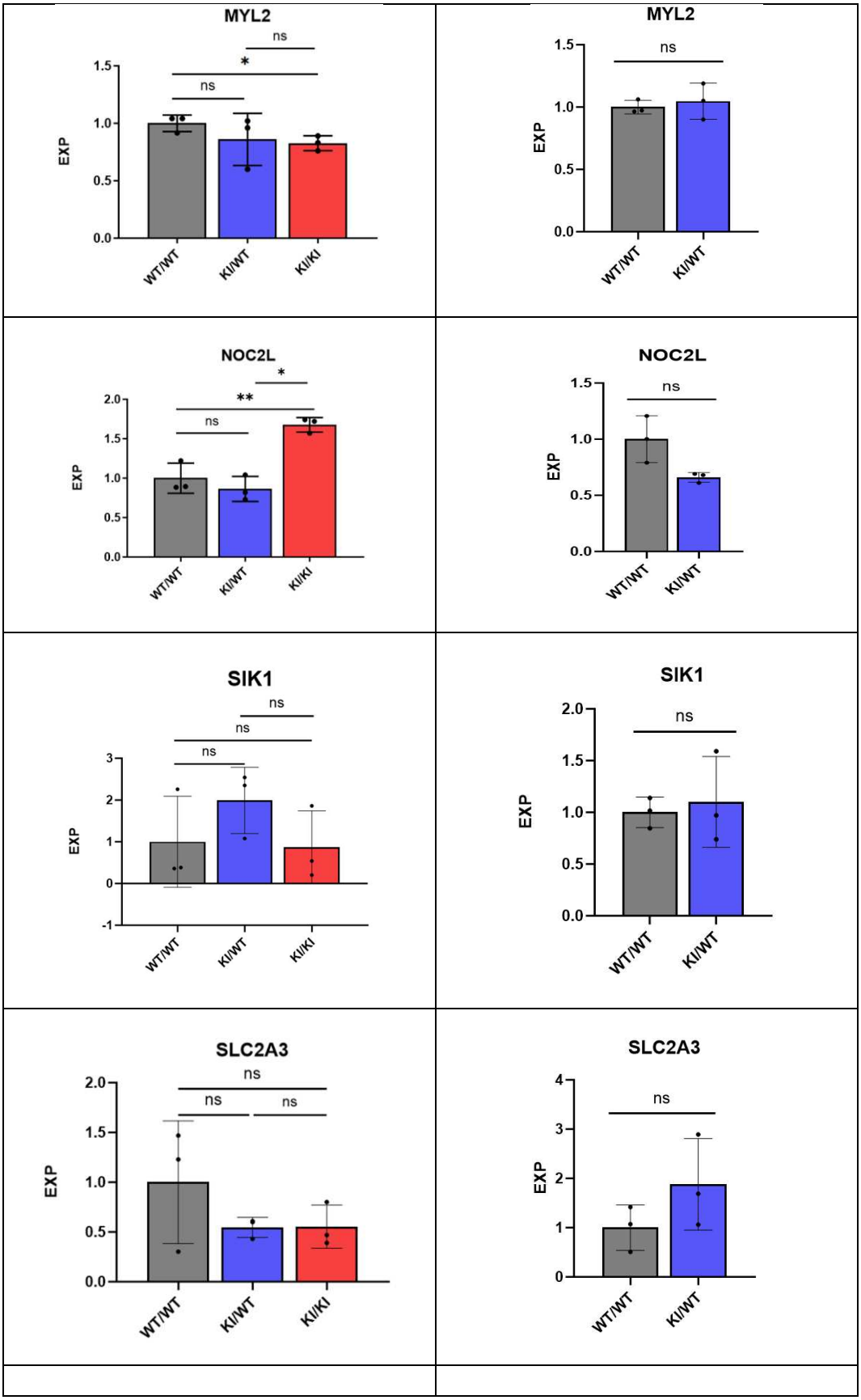
4.2.6 GENES NOT SIGNIFICANTLY DEREGULATED

10 genes did not show significantly different expression levels in the heterozygous samples compared with their wild type control at 3 nor 6 months, namely *Arhgap5, Cfl1, Des, Loxl1, Myl2, Noc2l, Sik1, Slc2a3, Sod1, Txn2*.They are shown in Figure 25.

Some of these genes showed under-expression in heterozygous 3 m.o. samples, while over-expression in heterozygous 6 m.o. ones, namely *Myl2, Slc2a3, Sod1*.

Des confirmed in all three samples group the dataset, but in a non-significant way. *Arhgap5, Sod1* confirmed the expression trend reported in the dataset with only exception made for homozygous 3 m.o. samples. *Loxl1, Sik1, Txn2* confirmed the trend with only exception made for heterozygous 3 m.o. samples. *Myl2, Slc2a3* confirmed the dataset with only exception made for heterozygous 6 m.o. samples. *Noc2l* confirmed the dataset only in heterozygous 3 m.o. samples. *Cfl1* confirmed the dataset only in heterozygous 6 m.o. samples.

3M.O. SAMPLES	6M.O. SAMPLES
	
	
	
	


MYL2
ns
*
ns
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
MYL2
ns
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
NOC2L
*
**
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
NOC2L
ns
EXP
1.5
1.0
0.5
0.0
WT/WT
KI/WT
SIK1
ns
ns
ns
EXP
3
2
1
0
-1
WT/WT
KI/WT
KI/KI
SIK1
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
SLC2A3
ns
ns
ns
EXP
2.0
1.5
1.0
0.5
0.0
WT/WT
KI/WT
KI/KI
SLC2A3
ns
EXP
4
3
2
1
0
WT/WT
KI/WT

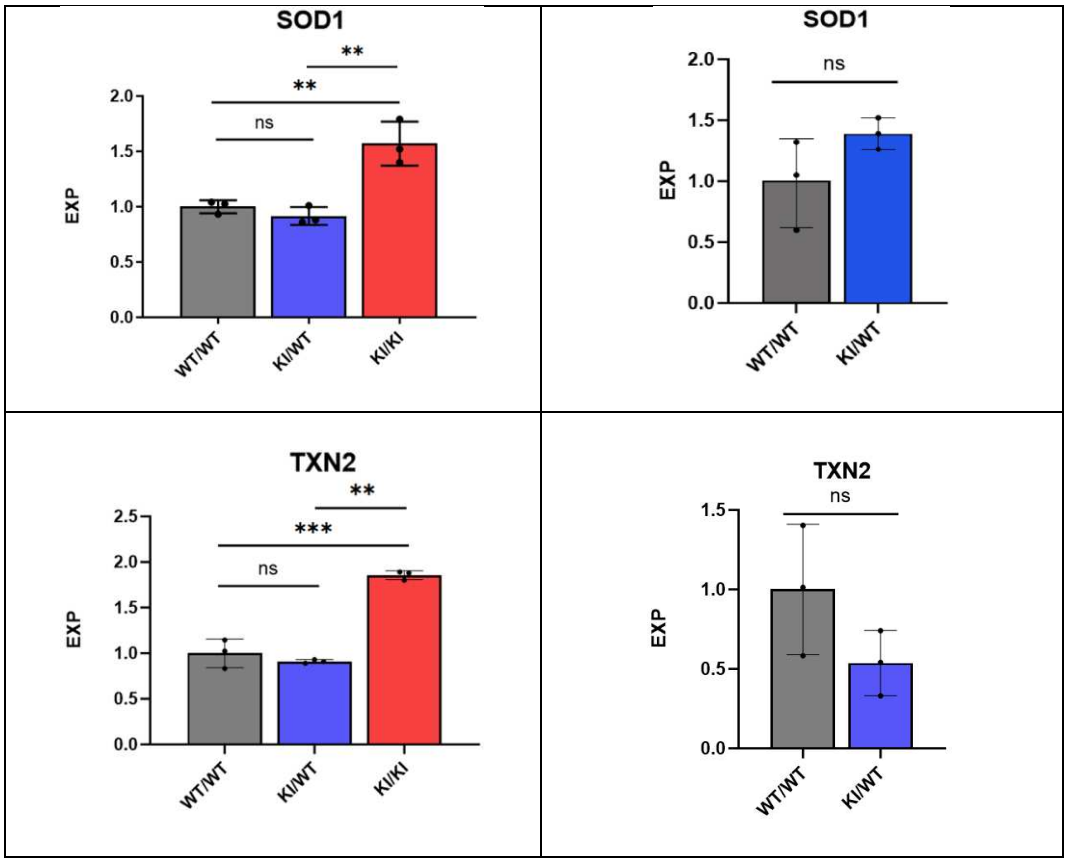


*Figure 25 Expression level of genes that resulted not significantly deregulated in either 3 m.o. heterozygous mice (left) nor 6 m.o. heterozygous mice (right), and their respective expression level in 3 m.o. homozygous mice, compared to WT controls. P-value of 0,05 was used as the threshold below which the data were considered statistically significant, with * indicating p-value<0,05, ** for p-value<0,01, *** for p-value<0,001 and **** for p-value< of 0,0001.*

5 DISCUSSION

In the first part of this thesis, two protocols for hiPSCs differentiation to hiPSC-CMs were used and compared, in order to provide information regarding their robustness and reproducibility, and to offer a useful tool for choosing the most appropriate one for the purpose of the study.

In the second part of this thesis, an RNA-seq dataset obtained from murine heart samples harbouring the DSG2 p.Q563* mutation, corresponding to the human DSG2 p.Q588* mutation previously associated with ACM, was validate, to gain new insights on the molecular mechanisms of its pathogenesis.

5.1 COMPARISON OF hiPSC-CM DIFFERENTIATION PROTOCOLS

hiPSC-derived cardiomyocytes (hiPSC-CMs) represent a powerful tool to improve understanding of the pathophysiology of genetic heart disease, such as arrhythmogenic cardiomyopathy, and to find potential therapeutic targets. Thanks to the patient's genetic background, they offer a unique platform to study developmental pathways and patient-specific mutations. To use this human *in vitro* model to analyse molecular regulation and cell-cell interactions in healthy and mutated samples, it is necessary to use well-defined and characterized differentiation protocols from hiPSCs state. [30]

Differentiation of hiPSCs into cardiomyocytes has progressed from complex, uncontrolled systems to relatively robust, specific, and defined protocols, leveraging the knowledge gained on each step from more than two decades of work. However, further improvements are still needed in scalability, time and cost, and increased robustness, with the minimum nutritional requirements still to be found [26]. In addition, cardiomyocytes differentiation protocols still show incoherence and different levels of effectiveness and reproducibility [24]. Comparing available protocols is therefore a useful tool in this regard, also to choose the most suitable one for the purpose of the study, and to reach solid conclusions based on the comparison of data from different studies.

In the first part of this thesis, two protocols for differentiation of hiPSCs into hiPSC-CMs were used and compared through analysis of the expression level of differentiation markers at different timepoints, comparison of cell morphology and organization, and study of beating behaviour.

5.1.1 DIFFERENCES IN REAGENTS USED IN THE TWO PROTOCOLS

Starting from day 0, the first difference between the two protocols that can be seen is the concentration of the B27 supplement used in RMPI medium, specifically 0,4% in protocol A against 2% in protocol B. It is clear from the literature that the most common B27 concentration is 2% [25], [31], which also represents the

appropriate concentration indicated by the ThermoFischer company. A significantly lower level could lead to immature or less efficient differentiation, due to lower level of its major component essential for cardiac differentiation when using RPMI medium, specifically ascorbic acid and albumin [26]. Perhaps it contributed to cause the limited spatial beating behaviour observed in the cells of protocol A even in the mature state. In addition, insufficient levels of retinoic acid, present in B27 supplement, may influence subtype specification [25].

During the first five days of differentiation, insulin has been found to decrease cardiomyocytes yield [31]. Accordingly, both protocols used B27 supplement without insulin for the first three days, with protocol B continuing to use it until the end of the cardiac phase (day 8), while protocol A started using insulin-containing medium as early as the beginning of the cardiac phase (day 3). This may be another cause of the limited spatial beating behaviour and high mortality rate during metabolic selection observed in protocol A.

Other differences between the two protocols, such as CHIR concentration and mesodermal induction molecules, can be seen at this stage. CHIR represents an activator of the Wtn signalling pathway, particularly it functions as selective inhibitor of glycogen synthase kinase 3 (GSK-3) [32]. Its use at day 0 significantly increases cardiac differentiation due to the promotion of mesodermal formation given by this pathway [30]. The concentration of CHIR should be optimized for each cell line and protocol to be sure to achieve the expected effect; in this case CHIR concentration of protocol B is ten times the one of protocol A (10 μM vs 1 μM), possibly impacting cardiac differentiation efficiency [25] and explaining higher mortality rate observed during metabolic selection in protocol A.

The Wtn signalling pathway represents a biphasic element because of its impact in both mesodermal and cardiac differentiation phases. Indeed, in the latter, inhibition of this pathway increases specification toward cardiac progenitor cells [30]. Small molecules that inhibit Wtn pathways are therefore typically used in the cardiac phase [34]: IWP4, used in protocol A, and IWP2, used in protocol B, are examples. They induce paracrine signalling pathways, such as TFG-β, promoting cardiomyocytes differentiation [24]. The use of different Wtn inhibitors has been shown to impact hiPSC-CM differentiation and maturation to a lesser extent than the chosen differentiation schedule and timing, demonstrating that the maturation phase of hiPSC-CMs is regulated by the temporal control of the Wtn pathway [32]. Protocol A and protocol B uses the same Wtn inhibition timing.

The two protocols also differ in the use of specific molecules during the mesodermal induction phase; specifically BMP4, ActivinA, and bFGF were used in protocol A, while human albumin was used in protocol B. The cross-talk between MBP4, Activin A and Wnt/β-catenin pathways is considered to represent the essential element for cardiac differentiation [26], while albumin-containing media seem to support low-density hiPSC culture well, but apparently they allow partially reprogrammed cells and spontaneous differentiation to occur, leading to final cell population composed of different cell types [26]. Proteins of the BMP cytokine

super-family, promote cardiomyogenesis; an example is represented by TGF-β. This family, together with Wnt and FGFs, is thought to control the early phase of mesodermal formation and cardiogenesis [24]. Recent studies revealed that the presence of BMP4 during cell differentiation results in late mesodermal derivatives formation such es primitive haematopoietic precursors [24]. High doses of Activin A, instead, are though to guide the differentiation towards endodermal derivatives [24]. Based on the mesodermal induction reagents used, Protocol A seems to be more supported by literature data.

Cardiac differentiation phase length is probably the first big discrepancy that can be noticed observing the two protocols; particularly, it spans 9 days in protocol A, while only 5 days in protocol B. This almost double cardiac phase time may have caused the organization of protocol A cells in the big intercellular spheric structures that can be seen at the end of the protocol, against a net-like organization typical of cells of protocol B that have been treated with the cardiac differentiation medium for almost half of the time. Important is to keep in mind though, that different reagents may need different time of reaction.

Cardiac differentiation does not proceed with 100% efficacy; a variable non-differentiated, partially differentiated, or non-desired cell type amounts are therefore still present in culture at the end of the protocol [24]. During metabolic purification, other cell types besides cardiomyocytes are eliminated due to the replacement of glucose with lactate in the culture medium: cardiomyocytes are the only cell type capable of utilizing the TCA cycle for ATP production [25] In this phase, differences can be appreciated between protocol A and protocol B in 2-mercaptoethanol and sodium-lactate concentrations. Protocol A uses higher 2-mercaptothanol and lower sodium-lactate concentrations than protocol B, with 0,4 mM against 0,1 mM for the first one and 2,2 mM against 4,4 mM for the second one. Despite these differences and the knowledge that the microenvironment plays a vital role in the differentiation of hiPSC-CMs, variability in the surviving level among biological replicates was observed for both protocols, possibly due to passage number together with other factors [25]. Important is to mention though that, by eye observation, the average cell mortality rate of protocol A seemed to be higher than that of protocol B, although higher doses of sodium-lactate could lead to a stronger selection, resulting in higher purified cardiomyocytes but in a more stressed state.

The higher cell mortality rate and the beating behaviour limited only to the three-dimensional inter-cellular structures of protocol A, suggest that protocol B represents a better way for hiPSC-CMs differentiation.

5.1.2 COMPARISON OF MARKER GENES EXPRESSION IN CELLS PRODUCED WITH THE TWO PROTOCOLS

The two protocols were compared also by analysing the expression of marker genes at different differentiation stages in cells generated by the two protocols.

On day 0, the expression level of *SOX2, OCT4* and *NANOG* as pluripotency markers was checked. Specifically, *SOX2* represents an intron-less protein-coding gene whose product is required for stem cell maintenance. It forms a trimeric complex with *OCT4* and interacts with the Wnt pathway. *OCT4* and *NANOG* genes encode for transcription factors important for embryonic development and stem cell pluripotency [33].

All these markers were found to have a significantly higher expression level in samples at day 0 of both protocols than in samples at day 10 (control), highlighting the pluripotency state at day 0 and the differentiated state at day 10. In particular, the expression level of *SOX2* was slightly higher in protocol A than in protocol B, with high variability between biological replicates for both protocols. The expression level of *OCT4* was also slightly higher in protocol A than in protocol B, while the *NANOG* one in protocol A appeared to be more than double that of protocol B. Collectively, these results suggest the higher pluripotency markers expression level of protocol A compared to protocol B, which still showed significance compared with the control.

On day 3, *BRACHYURY, TBX5* and *EOMES* expression level was analysed as mesodermal markers. *BRACHYURY* gene product regulates the transcription of genes required for mesodermal differentiation, while *TBX5* and *EOMES* genes products are transcription factors important for developmental processes such as heart formation [33].

Despite the differences observed between the two protocols in the mesodermal phase, all these three markers analysed showed significantly higher expression level in samples collected on day 3 compared to samples harvested on day 0, with almost same level in protocol A and B for *BRACHYURY* expression, both with high variability between biological replicates, higher level in protocol A for *TBX5* expression and higher level in protocol B for *EOMES* expression. These results indicate the reliability of both protocols regarding mesodermal differentiation.

Analysis of *TNNT2, GATA4* and *NKX2.5* expression level on day 10 confirmed that the cardiac phase was reached for both protocols. *TNNT2* gene codes for the cardiac form of troponin T, the tropomyosin-binding subunit of the troponin complex, important for muscle contraction. *GATA4* gene product represents a transcription factor that regulates the expression of genes involved in embryogenesis and myocardial differentiation. *NKX2.5* gene codes for a transcription factor that plays a central role in heart formation and development; mutations in *NKX2.5* gene cause atrial septal defect [33].

Despite the differences that the two protocols showed, both resulted to have significantly higher expression level of *TNNT2* and *GATA4* markers at day 10

compared to samples harvested on day 0 and non-significantly higher expression level of *NKX2.5* compared with the control. It is relevant to notice though, that protocol B showed higher expression level of the first two analysed markers than protocol A, suggesting more efficient cardiac differentiation.

In order to assess whether the obtained hiPSC-CMs expressed atrial and/or ventricular markers, on day 24 levels of *NR2F2* and *KCNJ3* (atrial markers) and *IRX4* and *MYL2* (ventricular markers) were measured. The atrial *NR2F2* gene product represents a ligand-dependent transcription factor activated by high level of retinoic acid, while *KCNJ3* codes for an inward-rectifier potassium channel important for heartbeat regulation. The ventricular *IRX4* gene product is predicted to enable DNA-binding transcription factor activity, playing an important role in ventricular differentiation, while *MYL2* gene codes for myosin light chain 2, a major sarcomeric protein central for heart muscle structure and function [33].

Even in this case, despite the differences reported by the two protocols, expression level of cardiomyocyte markers resulted to be mixed, with one atrial marker higher in its expression level in protocol A (*KCNJ3*) and one higher in protocol B (*NR2F2*), and one ventricular marker higher in its expression level in protocol A (*IRX4*) and one higher in protocol B (*MYL2*). All the four markers resulted in a significantly higher expression level compared to control (day 0), highlighting the mixed population of atrial and ventricular cardiomyocytes present in both protocols. Interesting to notice is that protocol B seems to present higher variability between biological replicates than protocol A at this stage.

Taken together, this markers analysis highlighted efficiency and reliability of both protocols for cardiomyocytes differentiation from hiPSCs.

5.1.3 ANALYSIS OF BEATING BEHAVIOUR

Evaluation of phenotypic features of hiPSC-CMs is one of the key factors in the effectiveness of an in vitro hiPSC-CM disease model, with contractility representing one of the most relevant of these features for studying pathologic changes, aging, and pharmacological effects [28].

The beating behaviour of the cells was analysed with the MYOCYTER macro for ImageJ by manually selecting regions of interest (ROIs). Among the different parameters generated by the software running the macro, beat time, frequency, and peak time with thresholds of 10%, 50%, and 90% were selected. The analysis was performed at different timepoints: before the metabolic selection, immediately after the metabolic selection and at the end of the protocol, with four videos analysed at each timepoint for both protocols.

Cells generated with protocol A started beating around day 10, with a ±1-day error among biological replicates. Cells generated with protocol B started beating around day 8, also with a ±1-day error among biological replicates.

From the results, it can be clearly appreciated how the cells in protocol B evolved in their beating behaviour during differentiation: from a slow immature beat before the metabolic selection (day 14), to an even slower beat right after the metabolic selection due to possible stressed state of the cells (day 19), to a much faster beat rate at the end of the protocol, characterizing mature cardiomyocytes (day 24).

In protocol A this evolution is not so clear, given the high variability in beating behaviour of the cells among the videos at each timepoint and the not significant differences in beat time and frequency at the analysed timepoints.

The 10, 50, 90 % peak time thresholds are defined as the local minima plus 10, 50, 90% of the difference between the local minima and the next following maxima, providing information about the width of the peak in the graph and allowing tracking of shifted amplitude [28]. Regarding these thresholds, the protocols showed almost opposite trends, with thresholds levels obtained in protocol A increasing with time, while thresholds levels generated by protocol B decreasing with time. These results, together with the near-constant beat frequency information, suggest an increase in the width of the beat on the graph for protocol A during time. Cells from protocol A also showed increase in beat amplitude with time, indicating the difference between every single frame and a reference image, which typically represents the resting phase. Amplitude can therefore be defined as the deviation from the resting phase [28].

Protocol B resulted in increased beat frequency and decreased thresholds analysed over time, resulting in an almost equal beat width in the graph of the different timepoints. The beat amplitude slightly decreased at the end of the protocol, when the beat appeared regular.

This different behaviour may depend on single protocol features, such as duration of each differentiation phase, the reagents and their concentration. Overall, it can be said that the cells in protocol B more closely mirrored the expected beating behaviour than those in protocol A, in terms of increasing or decreasing parameters over time. It is possible that the high variability in the beating behaviour among the videos analysed for protocol A led to this result.

5.2 RNA-SEQ DATASET VALIDATION

The second part of this thesis consisted of validating a transcriptomic dataset already available in the lab, obtained from three- and six-month-old knock-in mice heart for the p.Q563* *DSG2* mutation previously associated with arrhythmogenic cardiomyopathy (ACM). At the timepoint of three months, samples from both heterozygous and homozygous mice were available, while at six months, only samples from heterozygous mice were available.

Arrhythmogenic cardiomyopathy is a rare familiar cardiac disease, mainly characterized by ventricular arrhythmias, fibrofatty replacement of the myocardium and sudden cardiac death [20]. It is now confirmed that the most common cause of

ACM is desmosome mutations, which cause loss of adhesion between cardiomyocytes, leading to detachment and death fibrofatty replacement [6]. However, the detailed pathogenesis of the disease is not fully understood, but the involvement of element such as mitochondrial dysfunction [20], inflammation [6], [34], and cardiomyocytes death [6] is evident. Therefore, mitochondria, inflammation, apoptosis, and cell adhesion were therefore chosen as the classes of interest, along with heart development, as the main parameters for the choice of GO terms in the selection of gene for validation from the transcriptomic dataset.

Most of the genes analysed did not confirm the expression trend observed in the RNA-seq dataset, with the only exceptions being the *Dsg2, Des* and *Tmem88* genes. The limited number of biological replicates (3) used for qPCR validation, inhomogeneity of the samples or the accuracy of the dataset may explain this result. The inhomogeneity of the samples could also explain the high variability between biological replicates observed in some genes, possibly hiding the difference in expression level compared to the WT. Further studies, such as analysis at the protein level with Wester Blot technique on independent samples or qPCR with increased number of biological replicates, should be performed to clarify the difference in the expression obtained by the two analyses.

However, the results may be useful to study disease progression in heterozygous samples between three and six months of age. Specifically, 4 of the selected genes (*Dsg2, Lars2, Ppp1r13l, Tmem88)* showed a significantly different expression level as early as three months in the heterozygous samples compared to WT, while 13 genes (*Col4a2, Enc1, Etfb, Gpx1, Hrc, Isg20, Litaf, Mbd3, Ndufb10, Nol3, Rb1cc1, Sap18, Tgf-β)* were significantly deregulated only at six months. Moreover, 10 genes (*Arhgap5, Cfl1, Des, Loxl1, Myl2, Noc2l, Sik1, Slc2a3, Sod1, Txn2*) did not show significantly changed at either three or six months. These results could be the starting point of future studies to identify other altered pathways.

Important to mention is that in homozygous mice some genes did show significantly different expression level compared to WT, already at three months of age, although they weren't deregulated in heterozygous animals. This suggests that in these mice the presence of a single functional sequence is sufficient to maintain the expression level of these genes comparable to WT mice. Therefore, it would have been interesting to repeat the analysis of disease progression for homozygous samples as well. Important is to keep in mind though, that the deregulation of these genes could be a secondary effect, rather than directly caused by the mutation.

The mitochondria-related genes (*Gpx1, Nol3, Sod1, Ndufb10, Rb1cc1, Etfb, Col4a2, Lars2, Txn2, Des*), among those analysed, showed an interesting trend: in homozygous samples harvested at three months, their expression level was opposite to that obtained in RNA-seq. On the other hand, in samples from 3 m.o. heterozygous the same genes showed the same expression trend reported in the RNA-seq data, despite not-significantly. Interesting, *GPX1, Nol3, Ndufb10, Rb1cc1, Etfb, Col4A2, Lars2*, were significantly deregulated in 6 m.o. heterozygous mice. The under-expression found in some of the mitochondria-related genes may

be related to the metabolic switch observed under pathological conditions, which makes gene expression and metabolism more like fetal heart than adult one, relying more on glucose and less on fatty-acids oxidation as energy source [20].

Apoptosis-related genes (*Gpx1, Ppp1r13l, Nol3, Sap18, Rb1cc1, Litaf, Tgf-β, Col4a2*) were found significantly deregulated only in 6 m.o. heterozygous mice, confirming the trend observed in the RNA-seq data. Most of these genes resulted over-expressed in the three samples group, sustaining the finding of apoptotic death of cardiomyocytes during ACM pathogenesis, possible due to detachment between each other based on desmosomal mutations [6].

Most of genes related to cytoskeleton and cell adhesion (*Arhgap5, Loxl1, Myl2, Rb1cc1, Tgf-β, Cfl1, Col4a2, Des*) confirmed the RNA-seq data in 3 m.o. samples, but none of them significantly in heterozygous samples. In 6 m.o. samples, they completely confirmed the RNA-seq data, despite only few of them in a significant way. The under-expression of genes related to cytoskeleton and cell adhesion may lead to a decrease in cardiomyocytes adhesion between each other, leading to detachment and apoptotic death [6].

Inflammation-related genes (*Gpx1, Isg20, Litaf, Tfg-β, Cfl1, Slc2a3, Tmem88)* showed an increase in both confirmation of expression trend and statistical significance from three to six months. Even in 6 m.o. samples, however, only about half of them significantly confirmed the expression trend predicted by the dataset. Most of inflammation-related genes resulted under-expressed, in contrast with the evidence of its involvement in ACM pathogenesis sustained by the finding of inflammatory infiltration in patient's samples [34].

It is also interesting to note how genes involved in pathways related to ACM behaved in 3 m.o. and 6 m.o. samples. In particular, *Enc1* and *Tmem88* are known to be involved in or to interact with the Wnt pathway [33]. While it wasn't significantly deregulated in 3 m.o. heterozygous mice, *Enc1* was significantly altered in 6 m.o. mice, as observed in RNA-seq. On the other hand, *Tmem88* significantly confirmed the RNA-seq data already in 3 m.o. heterozygous mice. As both genes were significantly under-expressed, it will be interesting to assess whether this may be related to the canonical Wnt/*β*-catenin signalling pathway observed in ACM pathogenesis, leading to the induction of expression of pro-adipogenic genes [18].

Three genes, *Rb1cc1, Loxl1* and *Tfg-β*, are involved in the transforming growth factor-β (*Tgf-β*) signalling pathway. In 3 m.o. mice they showed an expression trend opposite to the RNA-seq data in homozygous mice and confirmed the RNA-seq data but non-significantly in heterozygous mice. In 6 m.o. mice, however, they significantly confirmed the trend. All these genes were under-expressed in 6 m.o. mice, in contrast to the evidence of *Tgf-β* pathway involvement in promoting fibrosis in response to cardiomyocyte injury in ACM pathogenesis [35].

Four genes, *Gpx1, Nol3, Ppp1r13l* and *Tfg-β*, represent mitogen-activated protein kinase (*MAPK*) signalling pathway, though to be activated by non-canonical *Tgf-β*

pathway [35]. Most of these genes showed opposite expression trend than the one reported in the RNA-seq data in both three- and six-months old samples, but with an increase in significance between the two groups. The under-expression level of these genes is in contrast with the founding of myocardial fibrosis, apoptosis, and hypertrophy stimulation by *MAPK* pathway [35]

5.3 CONCLUSIONS

In conclusion, the first part of this thesis highlighted, despite robustness and reliability of both protocols, how cells obtained with protocol B exhibit higher beating properties to those originated with protocol A. For this reason, along with cost-effectiveness, protocol B should be preferred when choosing which one to use. The hiPSC-CMs originated with protocol B may be a useful tool for future studies, such as two- and three-dimensional co-cultures, to mimic the hole heart environment and allow a better understanding of the interactions between different cell types, or personalized medicine approaches, based on patient's specific cellular environment and genetic background.

The second part of this thesis highlighted a discrepancy between the data obtained by RNA-seq and qPCR, with only exception made for 3 out of 27 analysed genes. In order to confirm a set of data, further analysis should be performed, starting with the study of protein level by Western Blot technique on a group of independent samples, along with qPCR analysis with increased number of biological replicates to achieve significance in the results. However, this analysis has provided useful information on the temporal evolution of ACM, that may represent a starting point for future investigations to better understand the molecular mechanisms driving the pathogenesis, starting with those genes that belong to more than one class of interest, such as *Col4A2, Gpx1, Myl2, Nol3, Ppp1r13, Tgf-β*, given that they may play a central role in the process.

REFERENCES

[1] "Cardiomyopathy: An Overview - PubMed." Accessed: Oct. 23, 2023. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29431384/

[2] K. Pilichou *et al.*, "Arrhythmogenic cardiomyopathy," *Orphanet J Rare Dis*, vol. 11, no. 1, Apr. 2016, doi: 10.1186/S13023-016-0407-1.

[3] B. J. Maron *et al.*, "Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention," *Circulation*, vol. 113, no. 14, pp. 1807–1816, Apr. 2006, doi: 10.1161/CIRCULATIONAHA.106.174287.

[4] B. Gerull and A. Brodehl, "Insights Into Genetics and Pathophysiology of Arrhythmogenic Cardiomyopathy," *Curr Heart Fail Rep*, vol. 18, no. 6, pp. 378–390, Dec. 2021, doi: 10.1007/S11897-021-00532-Z.

[5] J. B. Reisqs, A. Moreau, Y. Sleiman, M. Boutjdir, S. Richard, and P. Chevalier, "Arrhythmogenic cardiomyopathy as a myogenic disease: highlights from cardiomyocytes derived from human induced pluripotent stem cells," *Front Physiol*, vol. 14, 2023, doi: 10.3389/FPHYS.2023.1191965.

[6] D. Corrado, C. Basso, and D. P. Judge, "Arrhythmogenic Cardiomyopathy," *Circ Res*, vol. 121, no. 7, pp. 785–802, Sep. 2017, doi: 10.1161/CIRCRESAHA.117.309345.

[7] G. Thiene, D. Corrado, and C. Basso, "Arrhythmogenic right ventricular cardiomyopathy/dysplasia," *Orphanet J Rare Dis*, vol. 2, no. 1, 2007, doi: 10.1186/1750-1172-2-45.

[8] F. I. Marcus *et al.*, "Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria," *Eur Heart J*, vol. 31, no. 7, pp. 806–814, 2010, doi: 10.1093/EURHEARTJ/EHQ025.

[9] D. Corrado *et al.*, "Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria," *Int J Cardiol*, vol. 319, pp. 106–114, Nov. 2020, doi: 10.1016/J.IJCARD.2020.06.005.

[10] B. Gerull and A. Brodehl, "Genetic Animal Models for Arrhythmogenic Cardiomyopathy," *Front Physiol*, vol. 11, Jun. 2020, doi: 10.3389/FPHYS.2020.00624.

[11] B. Gerull *et al.*, "Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy," *Nat Genet*, vol. 36, no. 11, pp. 1162–1164, Nov. 2004, doi: 10.1038/NG1461.

[12] G. McKoy *et al.*, "Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease)," *Lancet*, vol. 355, no. 9221, pp. 2119–2124, Jun. 2000, doi: 10.1016/S0140-6736(00)02379-5.

[13] K. Pilichou *et al.*, “Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy,” *Circulation*, vol. 113, no. 9, pp. 1171–1179, Mar. 2006, doi: 10.1161/CIRCULATIONAHA.105.583674.

[14] P. Syrris *et al.*, “Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2,” *Am J Hum Genet*, vol. 79, no. 5, pp. 978–984, 2006, doi: 10.1086/509122.

[15] A. Rampazzo *et al.*, “Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy,” *Am J Hum Genet*, vol. 71, no. 5, pp. 1200–1206, 2002, doi: 10.1086/344208.

[16] M. Calore *et al.*, “A novel murine model for arrhythmogenic cardiomyopathy points to a pathogenic role of Wnt signalling and miRNA dysregulation,” *Cardiovasc Res*, vol. 115, no. 4, pp. 739–751, Mar. 2019, doi: 10.1093/CVR/CVY253.

[17] A. Rampazzo, M. Calore, J. Van Hengel, and F. Van Roy, “Intercalated discs and arrhythmogenic cardiomyopathy,” *Circ Cardiovasc Genet*, vol. 7, no. 6, pp. 930–940, Dec. 2014, doi: 10.1161/CIRCGENETICS.114.000645.

[18] A. Lorenzon, M. Calore, G. Poloni, L. J. De Windt, P. Braghetta, and A. Rampazzo, “Wnt/β-catenin pathway in arrhythmogenic cardiomyopathy,” *Oncotarget*, vol. 8, no. 36, pp. 60640–60655, 2017, doi: 10.18632/ONCOTARGET.17457.

[19] K. M. Austin *et al.*, “Molecular mechanisms of arrhythmogenic cardiomyopathy,” *Nat Rev Cardiol*, vol. 16, no. 9, pp. 519–537, Sep. 2019, doi: 10.1038/S41569-019-0200-7.

[20] C. J. M. van Opbergen, L. den Braven, M. Delmar, and T. A. B. van Veen, “Mitochondrial Dysfunction as Substrate for Arrhythmogenic Cardiomyopathy: A Search for New Disease Mechanisms,” *Front Physiol*, vol. 10, Dec. 2019, doi: 10.3389/FPHYS.2019.01496.

[21] O. J. Harrison *et al.*, “Structural basis of adhesive binding by desmocollins and desmogleins,” *Proc Natl Acad Sci U S A*, vol. 113, no. 26, pp. 7160–7165, Jun. 2016, doi: 10.1073/PNAS.1606272113.

[22] K. Sallam, K. Kodo, and J. C. Wu, “Modeling Inherited Cardiac Disorders: a Cell is Worth a Thousand Genes.”

[23] C. J. Chua, J. Morrissette-McAlmon, L. Tung, and K. R. Boheler, “Understanding Arrhythmogenic Cardiomyopathy: Advances through the Use of Human Pluripotent Stem Cell Models,” *Genes (Basel)*, vol. 14, no. 10, Oct. 2023, doi: 10.3390/GENES14101864.

[24] J. Lewandowski, T. J. Kolanowski, and M. Kurpisz, “Techniques for the induction of human pluripotent stem cell differentiation towards cardiomyocytes,” *Journal of Tissue Engineering and Regenerative Medicine*, vol. 11, no. 5. John Wiley and Sons Ltd, pp. 1658–1674, May 01, 2017. doi: 10.1002/term.2117.

[25] P. W. Burridge *et al.*, “Chemically defned generation of human cardiomyocytes,” *Nat Methods*, vol. 11, no. 8, pp. 855–860, 2014, doi: 10.1038/NMETH.2999.

[26] D. M. Lyra-Leite, Ó. Gutiérrez-Gutiérrez, M. Wang, Y. Zhou, L. Cyganek, and P. W. Burridge, "A review of protocols for human iPSC culture, cardiac differentiation, subtype-specification, maturation, and direct reprogramming," *STAR Protoc*, vol. 3, no. 3, Sep. 2022, doi: 10.1016/J.XPRO.2022.101560.

[27] X. Wang, A. Spandidos, H. Wang, and B. Seed, "PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update," *Nucleic Acids Res*, vol. 40, no. Database issue, Jan. 2012, doi: 10.1093/NAR/GKR1013.

[28] T. Grune, C. Ott, S. Häseli, A. Höhn, and T. Jung, "The 'MYOCYTER' – Convert cellular and cardiac contractions into numbers with ImageJ," *Scientific Reports 2019 9:1*, vol. 9, no. 1, pp. 1–13, Oct. 2019, doi: 10.1038/s41598-019-51676-x.

[29] D. van Beek *et al.*, "Allele-specific expression analysis for complex genetic phenotypes applied to a unique dilated cardiomyopathy cohort," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/S41598-023-27591-7.

[30] M. J. Doyle, J. L. Lohr, C. S. Chapman, N. Koyano-Nakagawa, M. G. Garry, and D. J. Garry, "Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes as a Model for Heart Development and Congenital Heart Disease," *Stem Cell Rev Rep*, vol. 11, no. 5, pp. 710–727, Oct. 2015, doi: 10.1007/S12015-015-9596-6.

[31] X. Lian *et al.*, "Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions," *Nat Protoc*, vol. 8, no. 1, pp. 162–175, Jan. 2013, doi: 10.1038/nprot.2012.150.

[32] C. Tsoi *et al.*, "Temporal Control of the WNT Signaling Pathway During Cardiac Differentiation Impacts Upon the Maturation State of Human Pluripotent Stem Cell Derived Cardiomyocytes," *Front Mol Biosci*, vol. 9, p. 24, Mar. 2022, doi: 10.3389/FMOLB.2022.714008/FULL.

[33] G. Stelzer *et al.*, "The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses," *Curr Protoc Bioinformatics*, vol. 54, no. 1, pp. 1.30.1-1.30.33, Jun. 2016, doi: 10.1002/CPBI.5.

[34] V. Meraviglia, M. Alcalde, O. Campuzano, and M. Bellin, "Inflammation in the Pathogenesis of Arrhythmogenic Cardiomyopathy: Secondary Event or Active Driver?," *Front Cardiovasc Med*, vol. 8, p. 784715, 2021, doi: 10.3389/FCVM.2021.784715.

[35] K. M. Austin *et al.*, "Molecular mechanisms of arrhythmogenic cardiomyopathy," *Nat Rev Cardiol*, vol. 16, no. 9, pp. 519–537, Sep. 2019, doi: 10.1038/S41569-019-0200-7.

ITALIAN SUMMARY

La cardiomiopatia aritmogena (ACM) è una rara malattia genetica del muscolo cardiaco classificata come cardiomiopatia primaria legata alle giunzioni cellulari, caratterizzata da progressiva sostituzione del miocardio ventricolare con tessuto fibro-adiposo e conseguente compromissione della funzione ventricolare. Un pattern di trasmissione autosomica dominante è stato identificato per questa patologia, caratterizzato da dominanza incompleta ed espressività variabile. La manifestazione clinica dell'ACM sta nella presenza di aritmie ventricolari significative, da cui la definizione di cardiomiopatia "elettrica" oltre che "strutturale". Essa rappresenta una delle principali cause di morte cardiaca improvvisa, soprattutto tra i giovani atleti, sostenendo la teoria che vede l'esercizio fisico intenso come elemento di accelerazione per la patogenesi nei soggetti portatori di mutazioni. L'ACM ha una prevalenza di 1:1000-1:5000 con presenza globale, probabilmente una sottostima del reale dato in quanto la morte cardiaca improvvisa ne può rappresentare la prima manifestazione. La frequenza è più elevata nel maschio, in rapporto 3:1 rispetto alla femmina, probabilmente grazie all'effetto protettivo di ormoni come gli estrogeni, e le manifestazioni spesso iniziano nella seconda-quarta decade d'età. La patogenesi viene divisa in quattro fasi: subclinica, priva di manifestazioni sintomatiche ma con alterazioni strutturali del ventricolo destro (RV); fase con aritmie ventricolari dovute ad alterazioni strutturali; perdita di funzionalità del RV; alterazioni anche nel ventricolo sinistro, con finale perdita in entrambi della funzionalità. Tutte le fasi sono caratterizzate dalla possibile presenza di arresto cardiaco.

Un set di criteri divisi in maggiori e minori, stabili nel 1994, aggiornati nel 2010 e recentemente migliorati dai "Criteri di Padova", rappresenta la base per la diagnosi della malattia. Essi riguardano alterazioni strutturali, disfunzioni, caratterizzazione delle pareti cardiache, caratteristiche di conduzione elettrica e storia familiare del paziente. La diagnosi viene stabilita davanti all'identificazione di una fra tre possibili combinazioni di criteri: 2 maggiori, 4 minori o 2 minori e 1 maggiore. Al giorno d'oggi, la terapia rimane sintomatica, con farmaci antiaritmici come prima linea, e defibrillatore o trapianto di cuore nei casi più gravi.

La principale causa della patologia è data da mutazioni in geni codificanti proteine desmosomiali e appartenenti all'area composita, gli elementi che assicurano la trasmissione del segnale elettrico e garantiscono aderenza meccanica tra i cardiomiociti insieme a canali ionici e gap junctions. Mutazioni anche in geni non-desmosomiali sono state identificate e associate alla patologia, arrivando a comprendere in totale il 60% dei casi.

Desmogleina-2 e desmocollina-2 rappresentano i componenti desmosomiali che, interagendo tra cellule adiacenti, ne assicurano l'adesione grazie alla loro ancora intracellulare ai filamenti intermedi citoscheletrici, per tramite di elementi come desmoplachina, placoglobina e placofilina-2. Alcune tra queste proteine desmosomiali, insieme a componenti delle giunzioni aderenti, vanno a comporre inoltre le aree composite.

Varie teorie sono state proposte per spiegare la patogenesi della malattia, la quale comprende la morte dei cardiomiociti per apoptosi e necrosi, conseguente infiammazione e sostituzione fibro-adiposa. La più accreditata di queste teorie si basa sulle modificazioni strutturali dei desmosomi e la conseguente perdita di adesione, la quale porterebbe alla morte dei cardiomiociti. Visto il ruolo dei desmosomi anche nella trasmissione di segnali intracellulari come Wnt/β-catenin/Tcf/Lef, Hippo-YAP e TGF-β, la loro alterazione potrebbe anche portare alla sostituzione fibro-adiposa e al fenotipo pro-aritmico. Altra teoria accreditata parte dall'alterato metabolismo cardiaco del calcio presente nei pazienti e vede il ruolo dei mitocondri centrale nella patogenesi, con però relazioni causalità o consequenzialità ancora da chiarire.

La maggior parte delle informazioni sulla patologia derivano da studi in vivo ed in vitro. Diversi modelli transgenici, knock-in e knock-out di topo e zebrafish hanno infatti permesso di comprendere la natura desmosomiale della maggior parte delle mutazioni associate e di identificare parte dei meccanismi molecolari alla base della patogenesi. I loro limiti stanno però nell'incapacità di rappresentare una delle principali manifestazioni cliniche, ovvero la sostituzione adiposa, e nelle differenze rispetto al modello umano, come il cuore bicamerale di zebrafish. Le colture cellulari forniscono invece il vantaggio di partire da cellule umane, con il principale limite dato dall'incapacità di mantenere i cardiomiociti primari in coltura per molto tempo e le loro scarse abilità duplicative. A questo scopo è stata introdotta la tecnologia delle hiPSCs, cellule staminali pluripotenti indotte umane che permette di riprogrammare una cellulare somatica del paziente, passando per lo stato di pluripotenza, verso la linea cellulare desiderata, senza limiti di tempo di coltura e capacità replicative. Per il differenziamento in cardiomiociti, si è passati da sistemi complessi e poco controllati a protocolli specifici e definiti, i quali però al mostrano giorno d'oggi diversità tra loro in termini di reagenti utilizzati ed efficienza di differenziamento, e limitazioni comuni date dal fenotipo immaturo dei cardiomiociti ottenuti e incapacità di rappresentare un cuore nella sua interezza. Diverse strategie di maturazione dei cardiomiociti sono state ideate, comprendendo coltura a lungo termine, co-coltura, approcci meccanici, elettrici e tridimensionali, per rendere il fenotipo più simile al cardiomiocita adulto rispetto a quello fetale.

Lo scopo di questa tesi è, nella sua prima parte, il confronto di due protocolli per il differenziamento di hiPSCs in cardiomiociti, comparando morfologia cellulare, organizzazione intercellulare, livello di espressione di marcatori del differenziamento a diversi timepoints e proprietà di battito. Con questa prima parte si vuole guidare nell'eventuale scelta del protocollo da utilizzare e saggiare la solidità degli studi eseguiti con questi protocolli. La seconda parte di questa tesi ha lo scopo di fornire nuove informazioni riguardo i meccanismi molecolari della patogenesi della ACM, tramite la validazione di un dataset trascrittomico ottenuto da campioni di cuore di topo knock-in per la mutazione p.Q563*, corrispondente alla mutazione umana p.Q558*, nella desmogleina-2 precedentemente associata alla patologia.

Nella prima parte di questa tesi un adattamento del protocollo usato nel Medical Centre di Goettingen, indicato con "A", e un protocollo in uso nel laboratorio della Dr.ssa Calore all'Università di Maastricht, indicato con "B", sono stati svolti in parallelo e confrontati. Le principali differenze tra i due stanno in: durata della fase cardiaca (9 giorni per A contro 5 per B), concentrazione di B27 (0,4% in A contro 2% in B), concentrazione di CHIR (1 μM contro 10 μM), reagenti per l'induzione mesodermale (BMP4, ActivinA, bFGF contro albumina umana), inibitore di Wnt (IWP4 contro IWP2), concentrazioni di 2-mercaptoetanolo (0,4 mM contro 0,1 mM) e di lattato di sodio (2,2 mM contro 4,4 mM) nella selezione metabolica.

Entrambi i protocolli mostrano chiaro cambiamento nella morfologia cellulare nelle fasi avanzate rispetto allo stato pluripotente, mentre differenze possono essere apprezzate tra i due a livello di organizzazione intercellulare. In particolare, il protocollo A genera cellule organizzate in grosse strutture intercellulari tridimensionali, quasi sferiche, mentre le cellule del B mostrano un'organizzazione a rete bidimensionale con maglie sempre più fitte. La tridimensionalità delle strutture in A è risultata comportare distacco di alcune di esse a seguito dell'inizio del battito dei cardiomiociti. Osservando l'organizzazione intercellulare è anche possibile apprezzare una riduzione del numero di cellule a seguito della selezione metabolica superiore in A rispetto a B, dovuta alla maggior mortalità cellulare.

L'analisi dei marcatori del differenziamento ha mostrato solidità ed efficienza per entrambi i protocolli. In particolare, l'analisi dei marcatori di pluripotenza *SOX2, NANOG* e *OCT4* al giorno 0 è risultata significativamente più alta rispetto al controllo (giorno 3) sia per A che per B, così come l'analisi dei marcatori mesodermali *BRA, TBX5* e *EOMES* al giorno 3 rispetto al controllo (giorno 0). Dei tre marcatori cardiaci invece, analizzati al giorno 10, *TNNT2* e *GATA4* hanno mostrato significatività rispetto al controllo (giorno 0), mentre *NKX2.5* non l'ha mostrata in nessuno dei due protocolli. Infine, al giorno 24 sono stati analizzati due marcatori dei cardiomiociti atriali, *NR2F2* e *KCNJ3*, e due dei cardiomiociti ventricolari, *IRX4* e *MYL2*. Tutti e quattro hanno mostrano livello significativamente più alto nei campioni rispetto al controllo (giorno 0), delineando la presenza di una popolazione mista di cardiomiociti atriali e ventricolari. L'analisi ha dunque confermato, per tutti i marcatori ad eccezione di *NKX2.5*, livelli d'espressione significativamente più alti nei campioni alle varie fasi del differenziamento rispetto ai controlli per entrambi i protocolli.

Il confronto delle proprietà di battito ha utilizzato l'analisi di Myocyter macro per ImageJ di video delle cellule due protocolli presi a tre timepoints, ovvero immediatamente prima della selezione metabolica (giorno 15 per A e 14 per B), immediatamente dopo la selezione metabolica (giorno 20 per A e 19 per B), ed alla fine del protocollo (giorno 24 per entrambi). Tra i parametri forniti in output, il tempo di battito, la frequenza e i thresholds del 10%, 50% e 90% del tempo di picco. I dati del battito del protocollo B hanno confermato l'atteso, hanno infatti mostrato un battito lento perché immaturo al primo timepoint, un rallentamento nel secondo timepoint dovuto allo stato stressato delle cellule dopo la selezione metabolica, ed una accelerata significativa alla fine del protocollo confermando la ripresa delle

cellule e la maturità. L'analisi del battito del protocollo A non ha invece mostrato l'atteso. Il tempo di battito, dunque la velocità è risultata pressocché costante nei tre timepoints, con un leggero, ma non significativo, rallentamento. Anche i dati riguardo la pendenza della curva che descrive l'ampiezza di battito sul tempo forniti dai tre thresholds scelti mostrano un battito in calando col tempo, probabilmente dovuto al fatto che i cardiomiociti dopo la selezione metabolica non sono stati in grado di riprendersi e dunque di andare verso il battito veloce maturo.

Concludendo, la prima parte di questa tesi ha sottolineato come, nonostante entrambi i protocolli siano affidabili e solidi nel differenziamento, le cellule ottenute con il B abbiano proprietà di battito superiori rispetto a quelle generate con l'A. Per questo motivo e per convenienza economica, il protocollo B dovrebbe essere considerato come prima scelta. I cardiomiociti differenziati con esso potrebbero rappresentare un buon punto di partenza per studi futuri di co-coltura di più tipi cellulari bi o tri dimensionale, per mimare l'ambiente cardiaco e analizzare le interazioni tra tipi cellulari, oppure per approcci di medicina personalizzata basata sull'ambiente cellulare ed il background genetico del paziente.

La seconda parte di questa tesi ha visto la validazione di un dataset trascrittomico ottenuto da campioni di cuore di topo knock-in per la mutazione p.Q563* nella desmogleina-2 associata alla ACM. Il dataset si compone di campioni omozigoti a tre mesi, eterozigoti a tre mesi, ed eterozigoti a sei mesi, mostrando, per ognuno di essi, l'elenco dei geni con il rispettivo livello d'espressione rispetto al wild type. La validazione è stata eseguita su campioni indipendenti ed è partita dalla selezione dei geni da analizzare tramite PCR quantitativa. Il primo parametro di selezione è stato un elenco di Gene Ontology (GO)-terms associati ai geni riportati nel dataset, concentrandosi sui termini interessanti per la patogenesi, riguardanti in particolare: sviluppo del cuore, citoscheletro ed adesione cellulare, mitocondrio, infiammazione e apoptosi. Dopo aver scremato i geni associati ai GO-terms scelti in base alla significatività, l'analisi della letteratura ha permesso di selezionarne 25 la cui espressione e funzionalità fossero interessanti per la malattia. Attenzione particolare è stata posta sui geni appartenenti a pathways di segnalazione noti per essere coinvolti nella patogenesi, come Wnt/β-catenina, Tfgβ e MAPK, allo scopo di confermare o meno il loro coinvolgimento e l'eventuale co-esistenza di alterazioni in più pathways.

Due geni invece, *COL4A2 e CFL1*, sono stati selezionati partendo da uno studio di espressione allele-specifica su dati di RNA-requencing di una condizione di disquilibrio allelico dovuto ad un meccanismo di regolazione cis-acting.

Dall'analisi in PCR quantitativa solamente 3 geni, *Dsg2, Des* e *Tmem88*, hanno confermato il trend d'espressione riportato nel dataset, dunque sovra-espressione o sotto-espressione rispetto al wild type. Tutti gli altri geni hanno mostrato, in almeno uno dei tre campioni (eterozigoti a tre mesi, omozigoti a tre mesi ed eterozigoti a sei mesi), trend di espressione opposto rispetto a quello ottenuto nell'RNA-seq. Il limitato numero di repliche biologiche (3) potrebbe essere alla base della non-significatività del livello d'espressione ottenuto rispetto al wild type di alcuni geni,

mentre fattori legati ai campioni utilizzati per la validazione o alla generazione del dataset potrebbero spiegare i trend d'espressione opposti ottenuti.

I risultati mostrano però una chiave di lettura interessante, ovvero la progressione della malattia da tre a sei mesi. Osservando gli eterozigoti infatti, 13 dei 27 geni analizzati mostrano un livello d'espressione che a tre mesi non è significativamente diverso rispetto al wild type, ma lo diventa nei campioni a sei mesi. Questo andamento potrebbe indicare come l'evoluzione della malattia nel lasso di tempo che intercorre tra i due campioni abbia comportato un cambiamento dell'espressione genica tale da far acquisire significatività. Questi 13 geni appartengono principalmente a termini riguardanti il mitocondrio, l'apoptosi e MAPK pathway. Ad avere livello significativamente diverso dal wild type nell'eterozigote a tre mesi e a mantenerlo a sei mesi sono invece 4 dei 27 geni, mentre a non mostrare significatività nell'eterozigote rispetto al wild type né a tre né a sei mesi sono 10. Interessante da sottolineare è come alcuni di questi ultimi hanno mostrato significatività nel loro divello di espressione rispetto al wild type nei campioni omozigoti a tre mesi, indicando come la presenza di una sequenza priva della mutazione potrebbe essere, in alcuni casi, sufficiente a mantenere il livello di espressione comparabile con quello wild type. In alternativa, questo fenomeno potrebbe essere un effetto secondario, non direttamente causato dalla mutazione.

Concludendo, la seconda parte della tesi ha mostrato discrepanza tra i risultati ottenuti con l'RNA-seq e la qPCR, se non per 3 su 27 geni analizzati. Per confermare uno dei due set di dati può essere utile eseguire in futuro studi a livello proteico su campioni indipendenti, insieme a ripetere l'analisi in qPCR aumentando il numero di repliche biologiche allo scopo di acquisire significatività. Questo studio ha comunque conferito informazioni utili per l'evoluzione temporale della ACM che potranno rappresentare il punto di partenza per studi futuri, in particolare concentrandosi sui geni appartenenti a più classi d'interesse, come *Col4A2, Gpx1, Myl2, Nol3, Ppp1r13l e Tgf-β,* in quanto potenzialmente centrali nella patogenesi.

RINGRAZIAMENTI

Un pensiero a chi mi ha seguita in questo percorso formativo e nei precedenti, dandomi gli strumenti necessari per trovare la mia strada.

Ringrazio chi mi ha aiutata ad arrivare fino a qui, pronto ad accompagnarmi anche verso i prossimi traguardi. Il vostro entusiasmo è stato supporto, i vostri consigli sono stati guida.

Ringrazio chi è pronto a condividere la vita, chi da sempre con me l'ha condivisa, chi vi ha lasciato traccia seppur di passaggio.

Questo spazio è dedicato a voi.