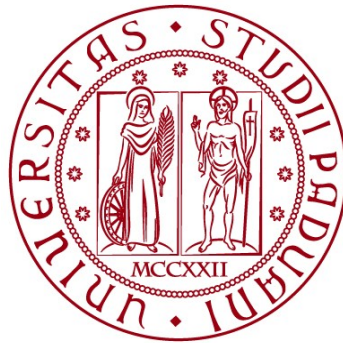**UNIVERSITÀ DEGLI STUDI DI PADOVA**

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea magistrale in Biologia Sanitaria

TESI DI LAUREA

hiPSC based models for Arrhythmogenic Cardiomyopathy: validation of RNA-seq data of hiPSC-CMs and set up of protocol for hiPSC-epicardial cells differentiation

Relatore: Prof.ssa Alessandra Rampazzo
Dipartimento di Biologia

Correlatore: Dott.ssa Martina Calore
Dipartimento di Biologia

Laureanda: Sabina Ferron

ANNO ACCADEMICO 2022/2023

TABLE OF CONTENTS

ABSTRACT

Arrhythmogenic cardiomyopathy is a rare genetic cardiac disease characterized by fibro-fatty replacement and sudden death. The majority of pathogenic mutations are found in genes encoding for components of the mechanical junctions of the heart. The pathogenic mechanisms are not clarified yet: however, among the potential sources of fibro-fatty tissue, epicardial cells are a valid candidate. Therefore to better understand the development and onset of the disease, in this thesis I optimized a protocol to generate hiPSC-epicardial cells to be used in the future as a platform to assess the molecular mechanism of ACM. In addition, a validation of the analysis of RNA sequencing data of hiPSCs carrying the p.Q558* mutation in DSG2, compared with the isogenic control, was performed to identify new pathways that may be involved in the disease. From this analysis, AREG, CCND3, CD35, OCLN, JAK2, NDUFA13, and TMEM65 genes were highlighted, providing possible new avenues to follow to understand disease development and to identify therapeutic targets.

1 INTRODUCTION

1.1 Cardiomyopathies

Cardiomyopathy is a family of heart muscle disease characterized by an alteration in the muscular and electrical function. It is defined by the American Heart Association as a disease of the myocardium mostly characterized by ventricular hypertrophy or dilatation [1]. Cardiomyopathies can be confined to the heart or be a part of generalized systemic disorders that can lead to cardiovascular death or progressive heart failure [2].

The classification divides the diseases into primary cardiomyopathies, that are limited to the heart and can be genetic, mixed or acquired, and secondary cardiomyopathies, that involve the heart as a part of a multi organ disease (Figure 1) [2].
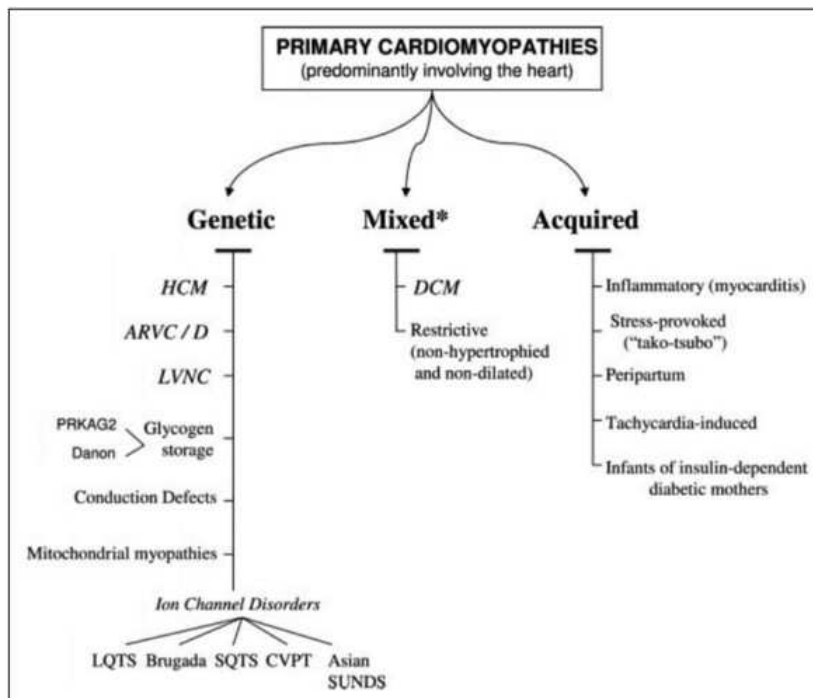
Figure 1: In this classification the different types of primary cardiomyopathy are divided into genetic, mixed and acquired [2].

In this complexity, arrhythmogenic cardiomyopathy (ACM), also known as arrhythmogenic right ventricular cardiomyopathy/ dysplasia (ARVC/D), is classified among genetic primary cardiomyopathies.

1.2 Arrhythmogenic cardiomyopathy

Arrhythmogenic cardiomyopathy is a genetic disease characterized by incomplete penetrance and variable expressivity. It manifests with fibro-fatty myocardial replacement, prominent ventricular arrhythmias and impairment of ventricular systolic function. It is one of the major causes of sudden death (SD) in young and in athletes: a post-mortem investigation in Veneto region on 60 young people with SD attested that 20% of them had ACM [3].

Historically, this disease is documented as a right ventricular dysplasia or Venetian disease (1982) [4] with the substitution of myocardium in the right ventricle (RV) with fibrofatty tissue. Further studies demonstrated that it had a genetic component and it was not present at birth and so it had been defined as cardiomyopathy instead of dysplasia [5]

Following the discoveries, also the involvement of the left ventricle (LV) had been registered, even if later and in a minor component. Eventually, some cases with predominant involvement of LV or of both ventricles had been discovered [3], so the name was changed into arrhythmogenic cardiomyopathy (ACM).

1.2.1 Epidemiology

Arrhythmogenic cardiomyopathy has an estimated frequency between 1:1000 and 1:5000. In some patients the first manifestation of the disease is sudden death (SD), so there could be more undiagnosed patients in the general population [3].

The disease has a higher frequency in males than in females (3:1) and also the severity is greater even if the mutation carriers are equally distributed among the two sexes [4]. There are different possible explanations related to the differences between the two sexes, in particular the difference of the sex hormones, such as estrogen, which may exert a protective role in women [3].

The disease is known all over the world, but it is historically known for a high presence in the North East of Italy, in Veneto region. Other cohorts, sharing the same mutations, were found in the Netherlands and in the rest of Europe. Also, other cases were identified in Asia and America while in South America a syndromic form had been found, the Carvajal syndrome, with wooly hair and palmoplantar keratoderma caused by homozygous mutation in the gene

DSP [3]. Another syndromic form was identified in Naxos, a Greek island, that gives the name to the syndrome itself. The Naxos syndrome is caused by homozygous mutation in the gene *JUP* [7].

The disease usually onsets after childhood between the second and fourth decade of life. There are some exceptions where it sets before puberty or when the patient has more than 70 years [6]. In particular, in the second case, ACM could be misdiagnosed because fibrofatty substitution can be mistaken for physiological adipose tissue, which is more common in the elderly population. The variable expressivity, also bound to the different age onset, can increase the percent of undiagnosed individuals.

1.2.2 Symptomatology and clinical features

Clinically, the disease is characterized by premature ventricular complexes (PVC), or ventricular tachycardia (VT) with left bundle branch block (LBBB) morphology and T-wave inversion in V1-V3 leads on basal electrocardiogram (ECG). Less-common presentations are RV or biventricular dilatation, with or without heart failure symptoms, mimicking dilated cardiomyopathy [6].

Clinically, the disease is divided into four phases. It starts with a subclinical phase, named concealed, characterized by subtle RV structural changes and generally there are no clinical symptoms, despite arrhythmias can be present and cardiac arrest can be the only manifestation. Secondly, the "overt electrical disorder" phase is characterized by palpitation and syncope. In this phase, ventricular arrhythmias of RV are present, mostly triggered by effort. Also, they can be life threatening leading to cardiac arrest. It is possible to detect these arrhythmic events with the presence of inverted T waves in a basal ECG that, after the age of 14 years old, are almost pathognomonic of ACM. Next, the "RV failure" phase is caused by the progressive loss of RV myocardium and the consequent impairment of the pump. Finally, there is the "biventricular failure" with the involvement of the LV. At this stage, ACM can mimic dilated cardiomyopathy, also because cardiac walls are so thin resulting in a dilatation of the heart [3],[6],[7].

1.2.3 Diagnosis

The diagnosis of ACM takes into account multiple criteria and they were decided in 1994 [8] and then updated in 2010 Marcus [8]. The diagnostic criteria are divided into groups that are: global or regional dysfunction and structural alterations; tissue characterization of wall; repolarization abnormalities; depolarization/conduction abnormalities; arrhythmias and family history [9]. In each of these groups the criteria are divided into major and minor. For the diagnosis, 2 major criteria, 4 minor, or 1 major and 2 minor from different groups have to be present [9].

The 2010 update of the diagnostic criteria was important because it added the genetic criteria to the clinical features, this increasing the specificity of the diagnosis.

These diagnostic criteria focus mostly on the diagnosis of the classic ARVC version of the disease, while the other presentations, involving primarily the left ventricle or both ventricles, are under-recognized. Therefore, more recently, new criteria, called "the Padua criteria" have been proposed, taking in consideration the variability of the expression of the disease. The used scoring system is the same as in the 2010 criteria, but with some implementation of tissue characterization by CE-CMR for detection of fibro-fatty myocardial replacement of both ventricles and the addition of new ECG criteria [7].

1.2.4 Therapy

The major treatments for patients with ACM consist in antiarrhythmic drugs (as sodium blockers, beta-blockers, sotalol, amiodarone, verapamil), catheter ablation, implantable cardioverter defibrillator and heart transplant [7]. In addition, sport disqualification for patients diagnosed with ACM, even If asymptomatic, is life-saving, since it is a major trigger of disease onset and progression. Screening of young athletes has an important role to increase the diagnosis and to reduce the mortality [6].

1.3 Genetics of Arrhythmogenic Cardiomyopathy

ACM is an autosomal dominant genetic disease characterized by incomplete penetrance and variable expressivity, also within the same family. In addition, the genetic characterization of the disease is made more complex by de novo and recessive mutations, compound heterozygosity, and digenetic inheritance [6].

The genetic etiology is established for more than 60% of the cases [8], and so it is recommended for the relatives of the patients to undergo a genetic test that can help for a prompt diagnosis. Most of the mutated known genes that are associated with ACM encode for proteins involved in desmosomes and area composita, the mechanical junctions of the intercalated discs [7]. In particular, causative mutations are mostly found in *PKP2*, encoding for plakophilin-2 [11], in *JUP*, encoding for plakoglobin [12], *DSG2* encoding for desmoglein-2 [13], *DSC2* encoding for desmocollin-2 [14], and *DSP*, which encodes for desmoplakin [15] (Table 1). For all of them the transmission can be dominant or recessive according to the mutation that they present. Most of the mutations known are in PKP2 gene.

GENE	ENCODED PROTEIN	REFERENCE
PKP2	Plakophilin-2	[11]
JUP or PG	Plakoglobin	[12]
DSG2	Desmoglein-2	[13]
DSC2	Desmocollin-2	[14]
DSP or DP	Desmoplakin	[15]

Table 1: Desmosomal genes involved in arrhythmogenic cardiomyopathy.

Less frequently, mutations have been reported also in non-desmosomal genes (Table 2). However, in many of these cases, functional studies are required to determine their pathogenic role in the disease.

GENE	ENCODED PROTEIN	LOCALIZATION	REFERENCE
CDH2	N-cadherin	Area composita	[16]
CTNNA3	αT-catenin	Area composita	[16]
TJP1	Tight junction protein-1	Tight junctions	[16]
DES	Desmin	Intermediate	[16]

		filaments	
LDB3	Cypher	Z-bands	[16]
ACTN2	A-actinin-2	Z-bands	[16]
SCN5A	Cardiac sodium voltage-gated channel subunit $\alpha 5$, $Na_v1.5$	Gap junctions	[16]
PLN	Phospholamban	SR membrane	[16]
LMNA	Lamin A/C	Nuclear envelope	[16]
TMEM43	Transmembrane protein 43	Nuclear envelope	[16]
LEMD2	Nuclear transmembrane protein-25	Nuclear envelope	[16]
TGFβ3	Transforming growth factor β 3	Secreted	[17]
TTN	Titin	Sarcomere	[16]
ILK	Integrin linked kinase	Focal adhesion	[16]
FLNC	Filamin-C	Z-bands	[16]

Table 2: Non-desmosomal genes that were found mutated in ACM patients.

1.4 Mechanical junctions and Desmoglein-2

Intercalated discs are points of contact between two adjacent cardiomyocytes, where mechanical and electrical junctions, as well as ion channel are present [18]. Mechanical junctions include desmosomes and area composita, crucial to keep the cells attached to each other, providing resistance to the mechanical stress. On the other hand gap junctions and ion channels allow the transmission of the electrical signal, thus contributing to the coordination of the contraction of the myocardium.

Desmosomes are multiprotein complexes that connect neighboring cells by a chain of interactions involving different proteins. Desmocollin-2 (DSC2) and desmoglein-2 (DSG2) are transmembrane proteins from the cadherin family that, with their extracellular domains, form homo- and heterodimers with the other cadherins protruding from the adjacent cells [18], ensuring their adhesion. These two proteins also present a cytoplasmic domain necessary

for the interaction with plakoglobin (PG) and plakophilin-2 (PKP2). In turn, PG and PKP2 are linked to desmin intermediate filaments through desmoplakin (DSP). This network is crucial to allow the propagation of the mechanical resistance of the cardiac tissue [17] (Fig 2).

Desmoglein-2, a calcium-binding glycoprotein of the cadherins family, is a key component of desmosomes in different types of tissues. The human genome encodes for four DSG (1 to 4) that share an overall domain organization. In particular, DSG2 is expressed in the heart, colon and all epithelial tissues containing desmosomes. In the heart, DSG2 has an important role because it is the only isoform expressed in cardiac desmosomes and area composita [19]. The structure is similar to other desmogleins consisting on an intracellular domain, a single-pass transmembrane region and five cadherin extracellular domains [19].

In this thesis, cells from a patient harboring the p.Q558X* mutation in desmoglein 2 were used. This nonsense mutation caused by the c.1672C>T substitution, produces a stop codon in the protein and it results in the loss of the transmembrane and intracellular domains, making it impossible to locate DSG2 in the plasma membrane (Fig 2).
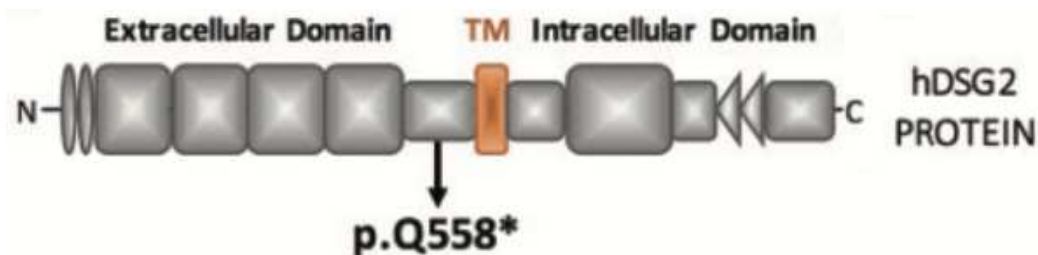
Figure 2: Scheme of DSG2 protein with the indication of the mutation Q558*, present in the cells used in the thesis.

In mammals, in the postnatal life, together with desmosome, also area composita (AC) can be found in the intercalated disc. It is a hybrid junction, consisting of proteins from both desmosomes and adherens junction (AJ). The AJ are constituted by N-cadherin (N-cad), which forms homodimers, with its extracellular domain, and interacts with β-catenin and plakoglobin with its intracellular domain. In turn, β-catenin interacts with αE-catenin that acts as a link with F-actin, a key component of the cytoskeleton. Finally the structure of area composita requires specific interactions between αT-catenin and plakophilin-2 [17] (Fig. 3).
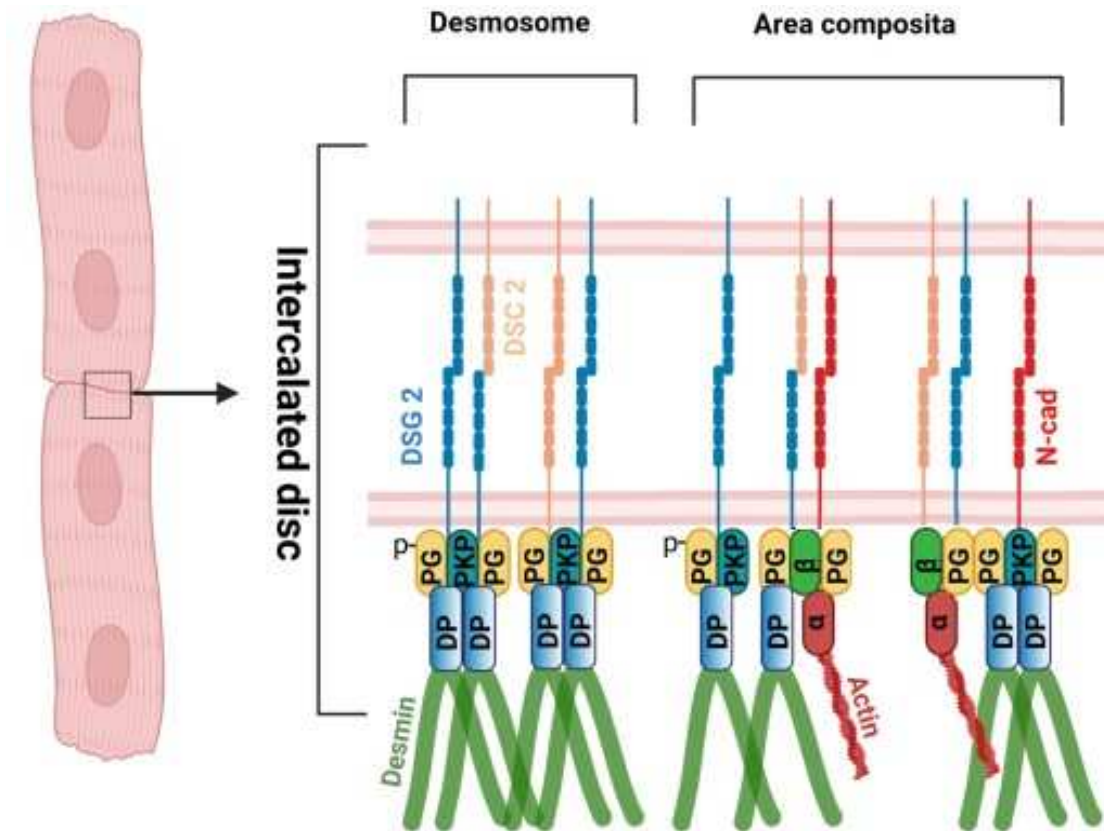
Figure 3: Representation of desmosome and area composita. DSC2, desmocollin-2; DSG2, Desmogelin-2; PG, plakoglobin know also as JUP or γ-catenin (γ); PKP2 plakophilin2; DP, desmoplakin; N-cad, N-cadherin, α, α-catenin; β, β-catenin. Adapted from [18]

1.5 Pathogenesis of Arrhythmogenic Cardiomyopathy

The hallmark of ACM is the loss of cardiomyocytes, either by necrosis or apoptosis, with fibro-fatty replacement. As a consequence, this leads to the impaired electrical conduction and the ventricular mechanical dysfunction typical of the disease [20]. In addition, inflammation has been observed in the postmortem tissue of ACM patients and seems to be implicated in pathogenesis [6].

Several theories have been formulated to explain ACM pathogenesis, taking into account the dysfunction of intercalated discs and the altered communication pathways. Indeed, the first hypothesis that was formulated, states that mutations in intercalated disc proteins cause loss of adhesion between cardiomyocytes, creating damage in the cardiac tissue, the progressive loss of ventricular myocardium and its replacement by fibro-fatty

tissue. This hypothesis was formulated by analyzing biopsies of patients with ACM, where it could be seen that cardiomyocytes had fewer desmosomes, which were also smaller in size, and increased intercellular distance [3],[7].

The role of desmosomes is not purely mechanical; they are also involved in both intra- and intercellular signal transduction pathways. Therefore, another theory takes this aspect into account [3], [6]. In particular, alterations in the Wnt/β-catenin/Tcf/Lef pathway were first observed in a *Dsp*-deficient mouse model and confirmed by other studies [6].

In normal conditions, activation of the canonical Wnt pathway inhibits β-catenin degradation, which can then translocate into the nucleus, where it binds to T-cell factor/Lymphoid-enhancer binding factor (Tcf/Lef). Activation of this pathway normally leads to proliferation and differentiation of cells, including differentiation into cardiomyocytes [20] (Fig. 4, left side). Accordingly to studies on different ACM models, mutations in desmosomal proteins can lead to dissociation of JUP from the cell membrane and its translocation into the nucleus. JUP, also called γ-catenin, is β-catenin paralog and competes with β-catenin preventing its binding to Tcf/Lef, eventually turning the canonical Wnt pathway off. The result, in cardiomyocytes, is the promotion of adipogenesis, which therefore may also explain the accumulation of lipid droplets in the heart and differentiation into adipocytes [3], [20] (Fig. 4 right side). This variation seems to affect also PPARγ (peroxisome proliferator-activted receptor-γ) expression and that could link to the fibrofatty replacement, indeed PPARγ is a regulator of lipogenesis and adipocyte differentiation [20]. PPARγ seems to have also a direct effect on canonical Wnt signaling through its interaction with β-catenin and the degradation of the last one [21]. Another molecule involved in the Wnt pathway is GSK3β (glycogen synthase kinase 3β), which is a promoter of β-catenin degradation, thus it is a negative regulator of canonical Wnt pathway [21]. Both in patients with ACM and in mouse models, GSK3β mislocalizes to intercalated discs, and when inhibited, there is an improvement in cardiac function, cardiac histopathology and survivorship [3], [16], [20],

Hippo-YAP (Yes-associated protein) pathway is another pathway that was found dysregulated in ACM patients and two ACM mouse model, and it seems to be correlated with an increased adipogenesis [16]. Inactivation of the Hippo pathway causes a cascade of different kinases that ends with activation, by hypophosphorylation, of YAP and TAZ. YAP and TAZ are

paralogue transcriptional coactivators of transcriptional enhancer factor domain (TEAD) family transcription factors and when active can translocate into the nucleus. The TEAD complex, when activated, interacts with the β-catenin/TCF (T-cell factor) complex by enhancing its transcriptional activity. Thus, activation of Wnt and YAP/TAZ has the effect of increasing cell survival and growth through transcriptional regulation [6], [16] (Fig. 4, left side). In contrast, activation of the Hippo pathway, by neurofibromin-2, a protein involved in cell-cell and cell-matrix adhesions, activates the kinase cascade. MST1 and MST2 (mammalian STE20-like protein kinase 1 and 2) phosphorylate LATS1 and LATS2 (large tumor suppressor homologs 1 and 2). Finally, LATS kinase phosphorylates YAP and TAZ leading to their retention and degradation in cytosol [6], [20]. Aberrant activation of the Hippo pathway leads to retention of YAP in the cytosol, which can then interact with β-catenin, further inhibiting the Wnt pathway and YAP activity on TCF and TEAD and increase adipogenesis [6],[20]. Finally, in addition, JUP presence in the nucleus, due to the desmosomal structural alteration, inhibits the YAP dependent activation of β-catenin genes target [16], [21] (Fig. 4, right side).

Another hypotheses is the involvement of TGFβ signaling pathway with both the canonical and non canonical pathway [16], [21]. The canonical one is activated by the binding of TGFβ1 and TFGβ2 to receptors with the result of SMAD phosphorylation. SMAD is a receptor-associated transcription factor that drive the expression of extracellular matrix proteins and inhibitors of metalloproteinases. The non-canonical TGFβ pathway active mitogen-activated protein kinase (MAPK) with the generation of ROS [21].

Most of the current knowledge of ACM pathogenesis is derived from studies in animal models and cell culture models. In particular, both mouse and zebrafish models with different mutations in desmosomal genes have been used [16], [21]. Their use has been essential to better understand the pathophysiology and molecular pathways leading to the onset of ACM [22]. Studies in mice provided insight into ACM as a disease of cardiac desmosomes using transgenic, knock-in, and conditional heart-specific mouse models; the latter were developed because global knock-out of desmosomal genes led to embryonic death [22]. Further studies were conducted in *Danio rerio* (zebrafish) to better understand the role of desmosomal genes, but also to study the role of *SCN5A*, *LMNA*, *ACNT2* and *TTN*, which are also involved in the development of ACM [22]. The main limitation of these models is that they cannot reproduce the fibro-fatty

substitution and, in addition, the zebrafish heart is bicameral, which limits full visualization of the disease [22].
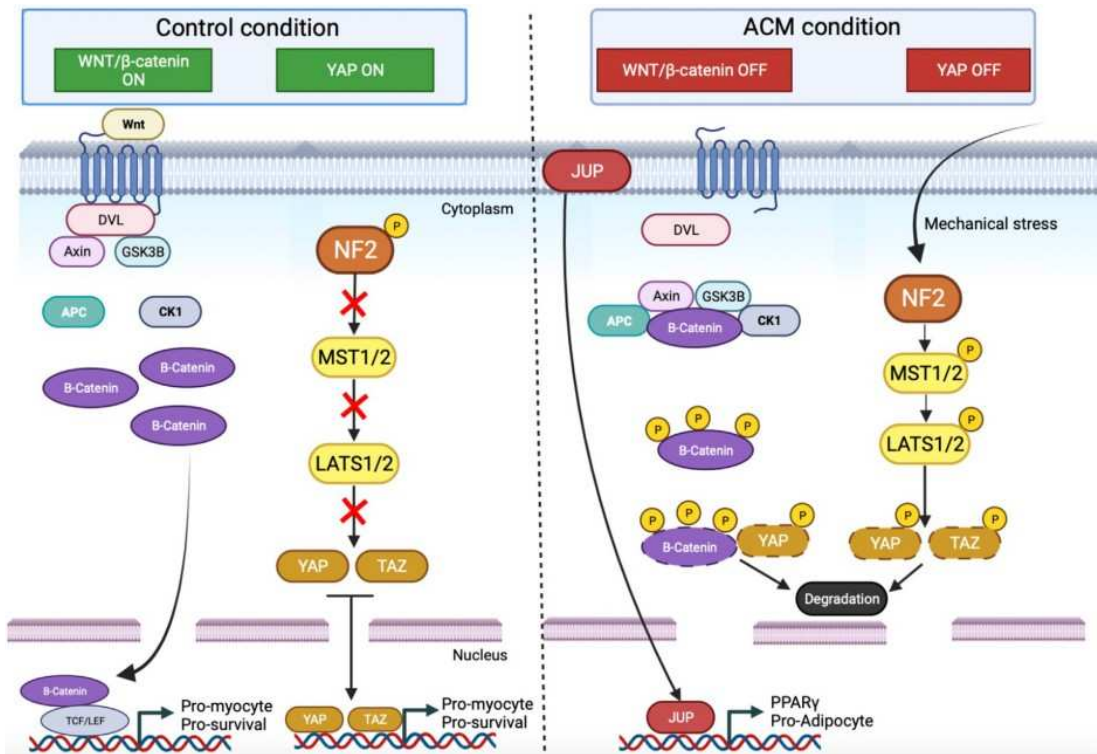
Figure 4: Wnt and Yap pathways in healthy condition (left) and ACM (right). On the left side, Wnt and YAP are active: For Wnt, activation of the Wnt receptor, through the Wnt ligand, triggers the assembly of a complex, which includes GSK3β. This induces the accumulation of free β-catenin in the cytosol that can enter the nucleus and bind TCF/LEF transcription factors. For YAP, NF2 is phosphorylated and inhibits the kinase cascade. YAP and TAZ are free to translocate into the nucleus. On the right are ACM conditions, when these pathways are misregulated: desmosomal disruption causes dissociation of JUP from other desmosomal proteins and its translocation into the nucleus, inhibiting Wnt/β-catenin signaling and leading to PPARγ transcription. The Hippo pathway is activated by NF2, resulting in activation of the kinase cascade; phosphorylation of YAP prevents its translocation into the nucleus and remains in the cytosol where it can interact with β-catenin leading to its degradation. GSK3β localizes at the level of intercalated discs [20].

An alternative is to focus on larger animals, in particular Boxer dogs, because they are a natural model of ACM and can better mimic the pathological features of the disease present in humans [22].

In addition, cell culture models have also been used to study the molecular mechanisms of ACM. The main problem with this technique is that primary cardiomyocytes cannot be maintained *in vitro* for long periods and they do not

divide. HL-1, an immortalized atrial cell line derived from transgenic mice, H9C2, a ventricular cell line, and more recently, human induced pluripotent stem cells (hiPSCs) have been used to overcome these difficulties. Using hiPSC technology, somatic cells from the patient can be reprogrammed to become pluripotent, which in turn can be induced to differentiate into several kinds of cardiac cells [21].

1.6 Fibro-adipogenic source

The hallmark of ACM is the fibro-fatty replacement of the myocardium following the loss of cardiomyocytes [21]. However, the origin of fibro-fatty tissue and inflammation is not clear yet [21].

At first, transdifferentiation of cardiomyocytes into adipocytes was proposed to explain the fibro-fatty replacement in ACM heart. However, different studies have shown that it is more likely that these cells undertake a process of lipogenesis, with lipid accumulation, rather than adipogenesis, differentiating into mature adipocytes [6],[21].

Another possible source of adipocyte, are cardiac progenitor cells of the second heart field, labeled by insulin gene enhancer protein (ISL1), from which cells of the ventricle derive [10]. The mechanism that is proposed to explain this process involves the suppression of the Wnt pathway in the canonical signaling [17].

A recent hypothesis is the involvement of cardiac fibro-fatty cell progenitors (cFAPs). These cells are characterized by the expression of PDGFRα (platelet-derived growth factor receptor α) and THY1 (Thy-1 cell surface antigen) [10]. It has been studied that a mutation in a desmosomal protein can lead to the differentiation of these cells into adipocytes [20]. In mouse models, it has been estimated that nearly 40% of adipocytes are derived from cFAPs [20], which suggests an important role of ACM pathogenesis, but may not be directly applicable in humans because of interspecies variation in the extent of adipogenesis [21].

Finally, it has been suggested that epicardial differentiation contributes to the fibro-fatty phenotype in patients with ACM. Cardiac pluripotent cells, specifically cardiac mesenchymal stromal cells, are normally involved in the structural management of the heart and localize in the epicardium [20]. The hypothesis is that, as in cardiomyocytes, the role of the desmosome is

15

essential for maintaining epicardial integrity and that loss of this structure may lead to differentiation of fibroblasts and fat cells [10]. These newly differentiated cells can then infiltrate the myocardium, also explaining the progressive involvement of the cardiac walls from the epicardium to the endocardium [21]. This hypothesis doesn't collide with the previous ones because from ISL-1+ progenitor cells can rise the epicardium and PDGFRA+ cells can rise from the epicardium [10]. Moreover, studies on hiPSC-epicardial cells demonstrate the EMT (epicardial-mesenchymal transition) phenotype as a consequence of suppression of desmosomal gene expression [10]. Further studies may better explain the real involvement of all these cell populations in the disease and their contribution to the fibro-fatty replacement in ACM [10].

2 Aim of the thesis

Arrhythmogenic cardiomyopathy is a genetic disease characterized by incomplete penetrance and variable expression. A major hallmark of the disease is fibrofatty myocardial replacement, which has been hypothesized to involve epicardial cells. However, the detailed mechanisms driving the disease haven't been clarified yet. It is therefore crucial to develop proper models to study the disease. In order to study the role of epicardial cell in ACM, in this thesis I aim to optimize the differentiation protocol of wild type hiPSC into epicardial cells. The protocol is different for 2 different concentrations of CHIR at the 2 different times of administration. The idea is to evaluate which condition can ensure more efficient differentiation in derived epicardial cells characterizing the cells at different time points and for the specific expression of markers.

In addition, in order to get novel insights into the molecular events occurring in ACM, I aim to validate an RNA sequencing dataset performed on hiPSC-CMs cells carrying p.Q558* nonsense mutation in DSG2 generated from an ACM patient.

3 Materials and methods

3.1 Cell cultures

In this thesis, I used hiPSC generated from bone marrow CD34+ cells from a healthy individual already available in the lab and stored in liquid N_2.

3.2 Coating

Cells were cultured in 6-wells plates. After thawing and during the first part of the differentiation process, 1 ml/well Geltrex (ThermoFisher Scientific, #A1413202) 0,16 mg/ml resuspended in DMEM F-12 (ThermoFisher Scientific, 11320033) was used as coating.

During the differentiation, after the first split (day 6), 1 ml/well 0,1% Gelatin dissolved in autoclaved MilliQ water was used. In both cases, the plates were kept 40 minutes at 37°C.

3.3 Procedure

3.3.1 hiPSC thawing

After thawing in a 37°C water bath, the cells were transferred into a 15 ml tube with 8 ml of E8++ medium composed by Essential 8 (ThermoFisher Scientific, A1517001) 2% E8 supplement, 0,5% of P/S and supplemented with 0,1% Thiazovivin (TZV). Then the mix was centrifuged at 200g for 5 min. Next, cells were resuspended in 1.5 mL E8 ++ medium with 0,1% TZV and seeded on Geltrex-coated 6-well plates (1,5 ml/well). TZV is beneficial for cell adhesion, however, after 24 hours it becomes toxic for cells so the medium has to be changed after 24 hours to E8++ medium without TZV.

3.3.2 hiPSC splitting

For cell expansion, after thawing, cells were divided using SDS-EDTA 0.5 mM: 1 ml was used for washing, after removing the medium, and 1.5 ml was used to detach the cells. Then it was necessary to incubate for 4 min or until the edges of the colonies start to detach. To seed the cells, it was necessary to remove the SDS-EDTA solution and add 1.5 ml of E8++ and shake gently to get patches of about 15 cells in suspension on the medium. Finally, divided the medium with the resuspended cells into 2 or 3 new wells and add enough

medium to reach the final volume of 1.5 ml. If necessary, the initial well can be maintained by adding new medium.

In preparation of the differentiation, when 80-100% confluency was reached, cells were split in a 1:2 to 1:6 ratio. After aspirating the old medium, cells were washed with 1 mL E8 medium and incubated for 5 min at 37°C with 800 µl of TrypLE. In order to dilute TrypLE, 1 ml E8 + medium with 0,5% P/S and with 0,1% TZV was added to each well and cells were collected in a 15 mL tube. After rinsing each well with 1 additional mL E8++ medium with 0,1% TZV to collect as many cells as possible, the samples were centrifuged for 3 min at 300g and the pellet was resuspended in the proper volume of E8++ medium with TZV chosen on the basis of the initial confluency. Finally, 1,5 mL cells were seeded in each Geltrex-coated well of a six-well plate.

After 24 hours, the media was changed the media with 1,5 ml E8++.

After the differentiation it was used a similar protocol for the splitting. The cells were splitted on day 6, if it was necessary, and after day 12 whenever was necessary.

In this case was used 1 ml of Versene (ThermoFisher Scientific, 15040066) and it was added with 2,5% Trypsin and neutralized with 2 ml of RPMI20 (RPMI 1640 with 20% FBS), after 5 min at 37°C and 5%CO2. Then the mix was centrifuged at 200g for 5 min and cells were resuspended in LaSR basal medium containing 2 µM of TZV (and 2µM TGF-β inhibitor after day 12).

3.3.3 hiPSC cryopreservation

For cryopreservation, the same protocol described for cell splitting was used. However, in this case, after centrifugation, the cells from each well were resuspended in 0,4 ml E8++ and were then added to 0,4 ml of cold cryopreservation medium. Cells were transferred into the cryovial (0,8 ml for each vial) and stored at -80°C using a Mr. Frosty Freezing Container, for 48h. Then the vials were transferred in liquid N_2.

3.3.4 hiPSC culture

hiPSC were cultured in 1,5 mL E8 ++ medium, which was changed daily, until 90-100%) confluence was obtained.

3.3.5 hiPSC differentiation into epicardial cells

The protocol used in this thesis was taken from Bao et. al and relies on the temporal modulation of Wnt/β-catenin signaling, which is a key pathway in the embryonic cardiogenesis [24].

hiPSC were cultured in 1 mL Epi differentiation medium composed by RPMI 1640 (ThermoFisher Scientific, 772400021) containing 0,2 mg/ml L-ascorbic acid (Sigma-Aldrich, A8960) with addition of 10 µM or 5 µM CHIR99021 (Sigma-Aldrich, SML1046) on day 0 and change 24 hour later in Epi differentiation medium alone. On day 2 Epi differentiation medium with addition of 5 µM IWP2 (Millipore, 681671), an inhibitor of Wnt Production 2. On day 4 it was again changed into Epi differentiation medium.

The medium was then changed to LaSR basal medium composed by Advanced Dulbecco's Modified Eagle's medium (ThermoFisher Scientific, 12634010), 1,3% of GlutaMAX (ThermoFisher Scientific, 35050061) and 0,1 mg/ml L-ascorbic acid, from day 6 with addition of 4 µM or 3 µM CHIR99021 on day 7 and 0,5 µM A83-01 (Stemgent, 04-0014), from day 12. A83-01 is an TGF-β inhibitor important to prevent spontaneous epithelial mesenchymal transition (EMT). From day 9 to 11 the medium was changed to LaSR basal medium daily and cells were splitted whenever they reached confluence with a ratio between 1:9 to 1:12.

3.4 Immunostaining

Using the same division procedure previously described, 1-well cells were resuspended in 1.2 mL and seeded 300µL for each Ibidi 4 chamber using LaSR + TZV (1:1000) + 0.5 A83-01. After 24 hours, the medium was changed to LaSR + 0.5 A83-01 and repeated until the cells reached confluence.

The preparation process required a 20-minute incubation at room temperature with 4% PFA. Then it was necessary to permeabilize the cells with 0.2% TritonX-100 in PBS-1%BSA for 1h at room temperature and finally incubate for 2h at room temperature with 1%BSA in PBS.

For immunostaining, cells were incubated with the primary antibody, 1:100 solution of 0.5mg/ml rabbit antibody, in 1%BSA PBS and were incubated overnight at 4°C, washed 2 times with 0.1%Tween-20 in PBS and 2 times with PBS, 5 min for each washing step. They were then incubated for 1.5 hours at room temperature with the secondary antibody, Goat anti-rabbit

Oregon green 488 1:1000 of 2 mg/ml, in 1%BSA in PBS in the dark, and the washing step was repeated again.

Hoescht's 1:1000 solution in PBS was used for nuclei staining and incubated for 20 min at room temperature. Finally, they were washed again 3 times with PBS.

3.5 RNA-sequencing data analysis

The data analysis has been done starting from a transcriptome dataset already available in the lab, from human induced pluripotent stem cells-cardiomyocytes (iPSC-CMs) harboring the heterozygous DSG p.Q558X mutation, generated from an ACM patient, and the relative isogenic control cells. To select the possible genes for the validation the analysis started from a dataset of GeneOntology (GO) terms, produced through an enrichment analysis using R already available in the lab.

Obtaining a list of significant GO terms, considered the literature on the disease and the previous knowledge, it is possible to select a list of interesting GO terms. For each GO term, all the genes that were given by the user related to that particular term are enlisted. Those genes can be confronted with a dataset of differentially expressed genes and a selection of significant genes can be done.

To select the genes, different parameters were used, first of all the comparison with false discovery rate (FDR). FDR is another way to adjust the p-value: in this case the output of 0,05 means that there are 5% chances that it is false positive but it studies only the significant result from p-value. A p-value with a score of 0,05 has 5% of false positives in all tests.

Starting from the GO, terms the genes involved are identified and then prioritized by the FDR (if over 0,05 they are not taken into account), p-value adjusted, prioritizing the lower score, and the log2 Fold Change.

The Fold Change represents how much the transcript's expression is changed in the comparison between the experimental and the control group, and it is reported on a logarithmic scale to base 2.The final choice was made considering the literature, focusing in the known pathway involved in the disease.

A list of gene was also select from different dataset containing genes that were differentially expressed in the cell model and in a mouse model with the

same mutation. In this dataset were present only genes that were significantly expressed in both models and the selection had been done considering the knowledge and literature.

3.6 Primer design

The design of the primer pairs for real time PCR was done using an online primer design program, Primer3 (https://primer3.ut.ee/). The input is the mRNA sequence of the gene of interest. In addition, the maximum difference in melting temperature between the two primers was set at 1°C, and the length of the amplicons was set between 120 and 205 nucleotides. Next to check for the presence of aspecific annealing regions in the DNA, BLAT (https://genome.ucsc.edu/cgi-bin/hgBlat?command=start) a UCSC Genome Browser tool, was used giving as input the sequence of both primers (forward and reverse).

Alternatively, another online primer design program was used, called PrimerBank (Harvard University, https://pga.mgh.harvard.edu/primerbank/). It uses a tested algorithm that favors the specificity against the chemical properties with higher relevance than in other programs. From this algorithm, the primers are experimentally tested throughout amplification of DNA and analysis of the amplicons, gel electrophoresis and NCBI Blast Analysis [23].

3.7 RNA extraction

RNA extraction was performed with the Zymo kit, Direct-zol™ RNA Miniprep. Specifically, it is necessary to remove the culture medium from the wells, wash with 1X DPBS and add 450 µL of TRIzol® (for a single well in a 12-well plate) to lyse the cells and leave for 3 minutes. Then, to stop the reaction, add the same volume of absolute ethanol and with a scraper detach all cells from the well. To purify RNA, transfer the mixture to a Zymo-Spin™ IICR column in a collection tube and centrifuge at 15000 g for 30 seconds. Discard the stream and wash with 400µL of RNA wash buffer and centrifuge at 15000g for 30 seconds. Then add 5µL of Dnase I (6 U/µL) and 75µL of DNA digestion buffer to digest the DNA and incubate at room temperature for 15 minutes. Wash 2 times with 400µL of Direct-zol™ RNA PreWash and centrifuge each time at 15000g for 30 seconds. Finally, add 700 µL of RNA Wash Buffer and centrifuge at 15000 g for 1 min to remove all wash buffer.

To elute RNA, move the column to an RNase-free tube, then add 25µL of DNase/RNase-free water to the column and centrifuge at 15000g for 30 seconds.

3.8 Reverse Transcription

In order to perform a real time PCR, retrotranscription of extracted RNA was performed. Starting from RNA, it is possible to obtain cDNA using reverse transcriptase, oligo-dT used as primers and also hexamer oligos. This enzyme is encoded by the genetic material of retroviruses and it starts synthesizing the cDNA from the 3' end of the primers using the RNA sequences as template.

The reaction was done with the follow reagents:

	Volume per sample (µl)
5x buffer M-MLV	8
M-MLV RT	2
Oligo dT	2
Hexamer	2
RNAsin	2
dNTPs (10 mM)	2
RNA (1 µg)	22
Vol tot	**18**

Table 3: Dosage and reagent used in reverse transcription.

The used protocol was:

Phase	Temperature	Length
Reverse transcription	37°C	2 h
Denaturation enzyme	75°C	10 min

Table 4: Condition of reverse transcription

3.9 Polymerase Chain Reaction (PCR)

3.9.1 Gradient PCR

The gradient PCR had been used to test different temperatures of annealing in order to select the most suitable the real time PCR (Tab. 5).

The variation of the annealing temperature is given by the different characteristics of the pairs of primers, in particular length and the percentage of GC.

Target	Primer forward 5'→3'	Primer reverse 5'→3'	°C
ACTA2	TATCCCCGGGACTAAGACGG	CACCATCACCCCCTGATGTC	56.5
AKAP13	GTCAACGGGCACACTTTCAG	GGAGGCTAGACTTTCTCGGC	55
ANKRD1	TTTCTGAAGGCTGCTCTGGA	AGTGGATGGCTGTGGATTCA	60
AREG	CCCCAAAACAAGACGGAAAGT	CTCCGTGAATGCAGAAATTTTGA	60
BRA	TATGAGCCTCGAATCCACATAGT	CCTCGTTCTGATAAGCAGTCAC	59
CCND3	TACCCGCCATCCATGATCG	AGGCAGTCCACTTCAGTGC	60
CD36	GGTGCCATCTTCGAACCTTC	AGCCCCATAACAGTTCTCTCA	58
CFL1	TCCTACTAAACGGAAGGGGC	TCCTCTGGCGTTGAAGACTT	60
CHD7	TGCTCGGGAACTATCGGATTAA	CGAGGGGAAAACAAAACGCA	62
COL4A2	GTGACAGCCAACATCGGAAG	GCCCAAGGATGAACAAAGCT	65
CX3CL1	GCCACAGGCGAAAGCAGTA	GGAGGCACTCGGAAAAGCTC	58
CXCL16	GACATGCTTACTCGGGGATTG	GGACAGTGATCCTACTGGGAG	62
CYBRD1	GCTCCGCTTTCTCTCCGAG	TGTCAATCCCATAAGTGCTGTTG	58
DSC2	CGTTCCTGTAGATCGTGAGCA	CAGTAGTGCCCACTCTGCAAT	57
DSG2	TGCTGCTTCTCCTGATCTGC	CAGATCCTCTCCCTCCCGAA	57
DSP	CTCGGACGGCTACTGTCAAA	GCCTGGGCAAAACACTCATC	57
EGR1	CCACCACGTACTCCTCTGTT	CCTGGGAGAAAAGGTTGCTG	58
F2RL1	TCCTGGCCATGTACCTGATC	GCGTTCTTTGCATGATCCCT	60
FHL2	TGTCAGTCCTGGAGAAGCAG	GAGTTACCAGCCCCTTCCTT	58
GAPDH	TCGGAGTCAACGGATTTGGT	TTCCCGTTCTCAGCCTTGAC	55-65
GDF15	CAGCTACAATCCCATGGTGC	TCAGGAACCTTGAGCCCATT	62
GJA1	TCTGAGTGCCTGAACTTGCC	CCCTCCAGCAGTTGAGTAGG	57
H2BC4	AAGAAGGCAGTGACCAAAGC	CGATGCGCTCAAATATGTCG	58
IL17RD	TGGGCTGTACAACATCACCT	CACTTGGTCATGGCAAGCAT	65
ISL1	GCGGAGTGTAATCAGTATTTGGA	GCATTTGATCCCGTACAACCT	63.9
JAK2	AGTAAAAGTCCACCAGCGGA	AGGAGGGGCGTTGATTTACA	55
JUP	CGCATCTCCGAGGACAAGAA	TGGCATCCATGTCATCTCCAT	57

KRT18	TCGCAAATACTGTGGACAATGC	GCAGTCGTGTGATATTGGTGT	63.9
L7	AATTCGAATGGCGAGGATGGCAAGTGCAACACCTTTCGAACCTTTGGG		55-65
NANOG	TGTGTTCTCTTCCACCCAGC	ACCAGGTCTTCACCTGTTTG	63.9
NDUFA13	TCGACTACAAACGGAACTTGC	GGTCGGTTTCTGCCTGTAAC	62
NKX2.5	AAGTGTGCGTCTGCCTTTCCCG	TTGTCCGCCTCTGTCTTCTCCA	61
OCLN	CTCTCTCAGCCAGCCTACTC	GTTCCATAGCCTCTGTCCCA	58
OCT4	TCTTCAGGAGATATGCAAAGCAGA	GATCTGCTGCAGTGTGGGT	63.9
PECAM1	GGAAAGCAGATACTCTAGAACGGA	GGGATGTGCATCTGGCCTT	56.5
PKP2	CAGCTGGGGAATGCAGACAT	AAGCTTGAGGATGCCACGAA	57
PTGES	TCCTAACCCTTTTGTCGCCTG	CGCTTCCCAGAGGATCTGC	60
SERPINE1	AGTGGACTTTTCAGAGGTGGA	GCCGTTGAAGTAGAGGGCATT	62
SOX2	GGATAAGTACACGCTGCCCG	ATGTGCGCGTAACTGTCCAT	63.9
TBX18	GACGATCTTTCTCCCATCAAGC	CTATCTTCAGGCGAGTAATCTGC	58
TLR4	AGACCTGTCCCTGAACCCTAT	CGATGGACTTCTAAACCAGCCA	58
TMEM65	TGCTGCTTTGGGAAATCTTGT	AGCCAATAGTCACCCCAACA	62
TNFRSF10A	ACCTTCAAGTTTGTCGTCGTC	CCAAAGGGCTATGTTCCCATT	60
TNFRSF12A	GATCCAGTGACAATGTGCCC	GCAACCAGACACCTTGGAAG	55
TNNT2	AAGAGGATGCTGAAGCAGAG	TGGTTTGGACTCCTCCATTG	63.9
UBD	CTCTGTGTGCATGTCCGTTC	GGCTTCTCCGTGGCTTTAAG	62
WT1	TGTCAGCGAAAGTTCTCCCG	GCTGAAGGGCTTTTCACTTGT	56.5

Table 5: Primers used and the relative temperature

The used protocol was:

Phase	Temperature	Length	Repetition
Taq Activation	95°C	30 sec	1 cycle
Denaturation	95°C	30 sec	35 cycles
Annealing	From 55 to 68°C	30 sec	
Elongation	68°C	30 sec	
Final extention	68°C	10 min	1 cycle
Conservation	4°C	As needed	

Table 6: Condition of amplification for gradient PCR

The used reagents were:

MIX	µl x each well
F primer (10 µM)	0,5
R primer (10 µM)	0,5
LongAmp® Taq 2X Master Mix	6,25
cDNA	0,5
H$_2$O Milli-Q RNase/DNase free	4,75
Total volume	**12,5**

Table 7: Dosage and reagents used in gradient PCR

LongAmp® Taq 2X Master Mix is a mix that contains two types of DNA polymerase: Recombinant Taq DNA polymerase and Deep Vent® DNA Polymerase. Deep Vent® DNA Polymerase as 3'→5' exonuclease activity that increases the fidelity of the amplification.

The amplicons are visualized with electrophoresis run in 2% agarose gel, made with UltraPure™ Agarose in 1X TAE Buffer. The DNA is visualized using Ethidium Bromide Solution 10mg/ml, adding 5 µl/100ml.

3.9.2 Real Time PCR

Real Time PCR allows to follow the cDNA amplification thanks to the presence of SYBR Green in the reaction. SYBR Green intercalates in the stranded DNA (dsDNA) produced during the PCR amplification and it is possible to obtain the threshold cycle (Ct value) corresponding to the cycle in which the signal becomes strong enough to be detected. The threshold is the level of fluorescence, set by the software, above the background noise. The quantification of each gene is recorded during every DNA synthesis cycle.

Also at the end of the amplification, the quality of the melting curve was assessed for the absence of aspesifics.

Reagents:

PRIMER MIX	µl for well
iQ SYBR Green Supermix	12
Primer Forward 10µM	1
Primer Reverse 10µM	1

H$_2$0	10
Total volume	**24 µl**

Table 8: Composition of the primer mix.

For each well were add 1 µl of cDNA solution and 24 µl of the precedent mix.

The conditions of the reaction were:

Phase	Temperature	Length	Repetition
Taq Activation	95°C	10 min	1
Denaturation	94°C	15 sec	
Annealing	From 55 to 68°C	30 sec	44
Elongation	72°C	1 min	
Termination	95°C	1 min	1
Preparation	55°C	1 min	1
Melting curve	55°C to 94,5°C + 0,5°C		1

Table 9: Conditions of amplification for real time PCR

For each single gene, three biological samples in three technical replicates were used. In all the RT-PCRs, GAPDH was used as housekeeping gene.

3.9.3 Data analysis of real time PCR

The data analysis has been done using and 2-(ΔΔCt) method, where Ct is the cycle of the amplification reaction in which the signal intersect the threshold. The ΔCt is the difference between the interest gene and the housekeeping gene (in this was a normalization is possible). Finally ΔΔCt is difference between the ΔCt of the analyzed condition and the ΔCt of the control. The outcome is the fold change: in order to compare it with the RNA-sequencing data, log2 must be calculated. To assess the statistical significance of the obtained data, a t-test of the expression can be performed. The expression was obtained dividing each 2^(-ΔCt) for the the average of 2^(-ΔCt) of the control conditions.

3.10 Statistical analysis

The statistical analysis for the thesis has been done using software Prism v.10.0.0 (GraphPad Software) using t-test for pairwise comparison and ANOVA for multiple comparison. In addition with the ANOVA testing it was possible to check for the distribution and if it was not linear the value can be converted into logarithmic scale. For the pairwise comparison the q-test was applied to exclude possible outliers.

It has been considered significant when the p-value <0.05 and in particular in the graph has been draw * for p-value<0.05, ** for p-value<0.01, *** for p-value<0.001 and **** for p-value<0.0001.

4 Results

My thesis consists in two parts: first, I set up a protocol to generate epicardial cells from induced pluripotent stem cells (hiPSC), next, I validated a set of RNA-seq data obtained in iPSC-derived cardiomyocyte from an ACM patient harboring the p.Q558* mutation in DSG2.

4.1 Set up of the epicardial differentiation protocol

Based on the published literature [10], [24] it was possible to try to set in the lab a protocol for the epicardial differentiation of the available wild-type hiPSC, by modulating the Wnt/β-catenin signaling (Fig. 5). Different conditions were tested and cells were characterized at different time points by evaluating their morphology as well as marker gene expression (Fig. 5).
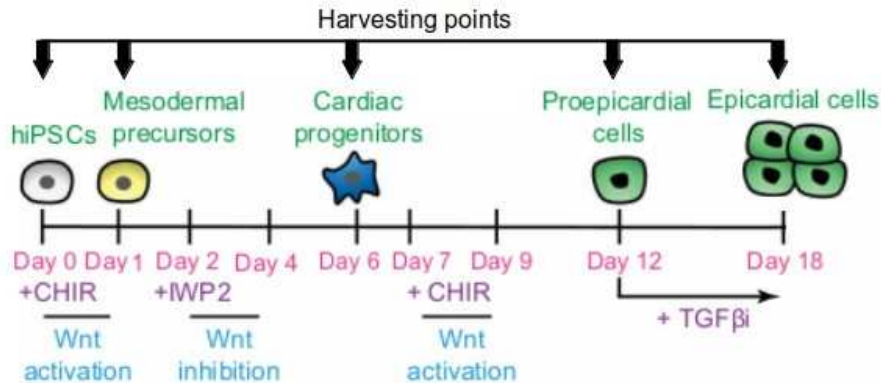
Figure 5: Schematic representation of the protocol used in this thesis with the indication of the different time points for the harvesting and differentiation. Adapted from [10].

The Wnt signaling is a crucial pathway for cadiogenesis in embryos. In order to induce the differentiation into the proper lineage, the Wnt signaling was modulated by using the GSK3β inhibitor CHR or the Porcupine inhibitor IWPs (Fig. 5). In particular, on day 0, Wnt activation using CHIR, promoted the differentiation into mesodermal precursors expressing Brachyury.

From day 2 to day 4, Wnt inhibition using IWP2 was necessary to direct hiPSC towards differentiation into cardiac progenitors expressing Nkx2.5 and Isl1.

Until differentiation into cardiac progenitors, RPMI medium was used, which was switched to LaSR medium after day 6.

From day 7 to day 9, further inhibition of Wnt, using CHIR, was necessary to direct the cells towards an epicardial fate. Finally, from day 12 cells were treated with the TGFβ inhibitor A83-01, to inhibit the epithelial-to-mesenchymal transition to promote self-renewal.

The different conditions are given due the different concentration of CHIR at the 2 time points:

- day 0
  - 10 µM
  - 5 µM
- day 7
  - 4 µM
  - 3 µM.

Day	Day 0
Name	Control

Day	Day 1/ Day 6	
Treatment done	+5 µM CHIR	+10 µM CHIR
Name	5µM	10µM

Day	Day 12/ Day 18			
Treatment done	+3 µM CHIR	+4 µM CHIR	+3 µM CHIR	+4 µM CHIR
Name	5+3µM	5+4µM	10+3µM	10+4µM

Table 10: Schematic representation of the used nomenclature

4.1.1 Day 1, mesodermal precursors

To assess the variation of the pluripotency, the quantification of the expression of the pluripotency markers *NANOG, OCT4* and *SOX2* was performed at each time point and compared with the initial condition (Day 0).

In addition, quantification of the expression of specific markers was performed for each time point.

As expected, all three markers were detected in all experimental settings. Their expression was comparable to that of day 0 when CHIR 10 µM was used, whereas, unexpectedly, it increased when CHIR 5 µM was used (Fig 6A, 6B and 6C). Mesodermal differentiation was assessed by the quantification of *BRA*, which was significantly increased in both protocols, but the difference was also significant between the two conditions, with higher expression in cells 10 µM, which could be in line with the results of pluripotency markers (Fig 6D).

Moreover, using both protocols, at day 1 hiPSCs showed the typical morphology of pluripotent cells, small and round, but there were more spots where cells detached (probably dead) in those treated with 10 µM (Fig 7 and 8).
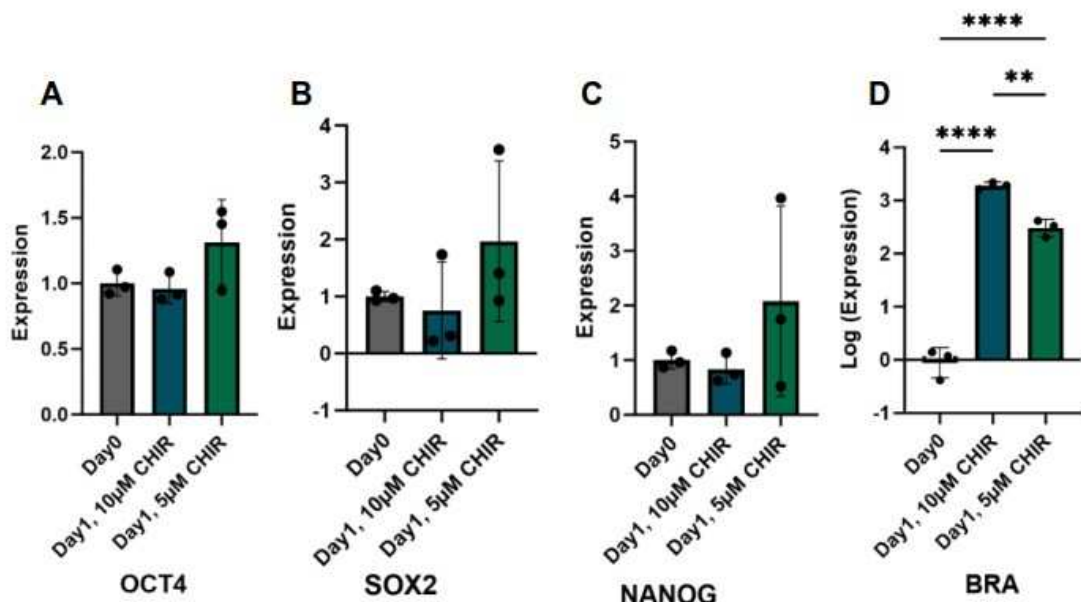
Figure 6: Expression level at day 1 in hiPSCs treated for 24 hours with either 10µM or 5µM CHIR: **A** *OCT4*, **B** *SOX2*, **C** *NANOG*, **D** *BRA*.
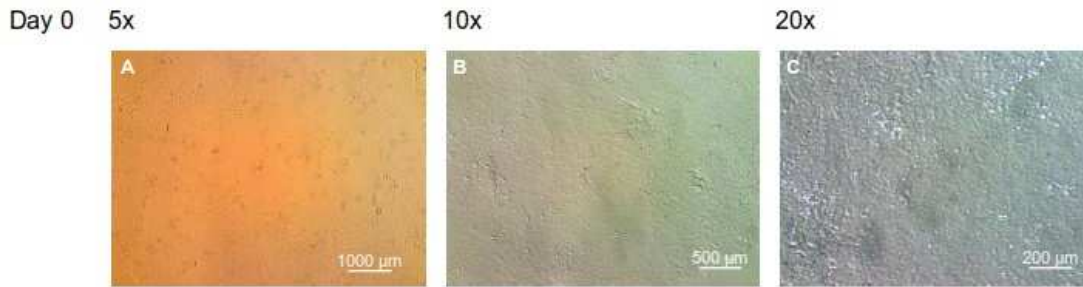
Figure 7: Light microscope images showing the morphology of hiPSCs at day 0: **A** 5x magnification, **B** 10x magnification, **C** 20x magnification.
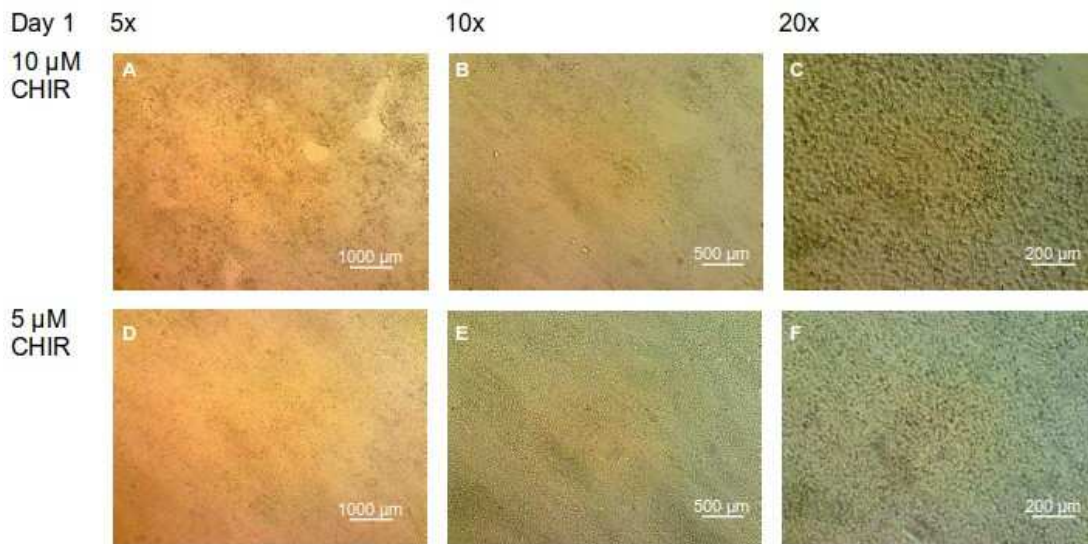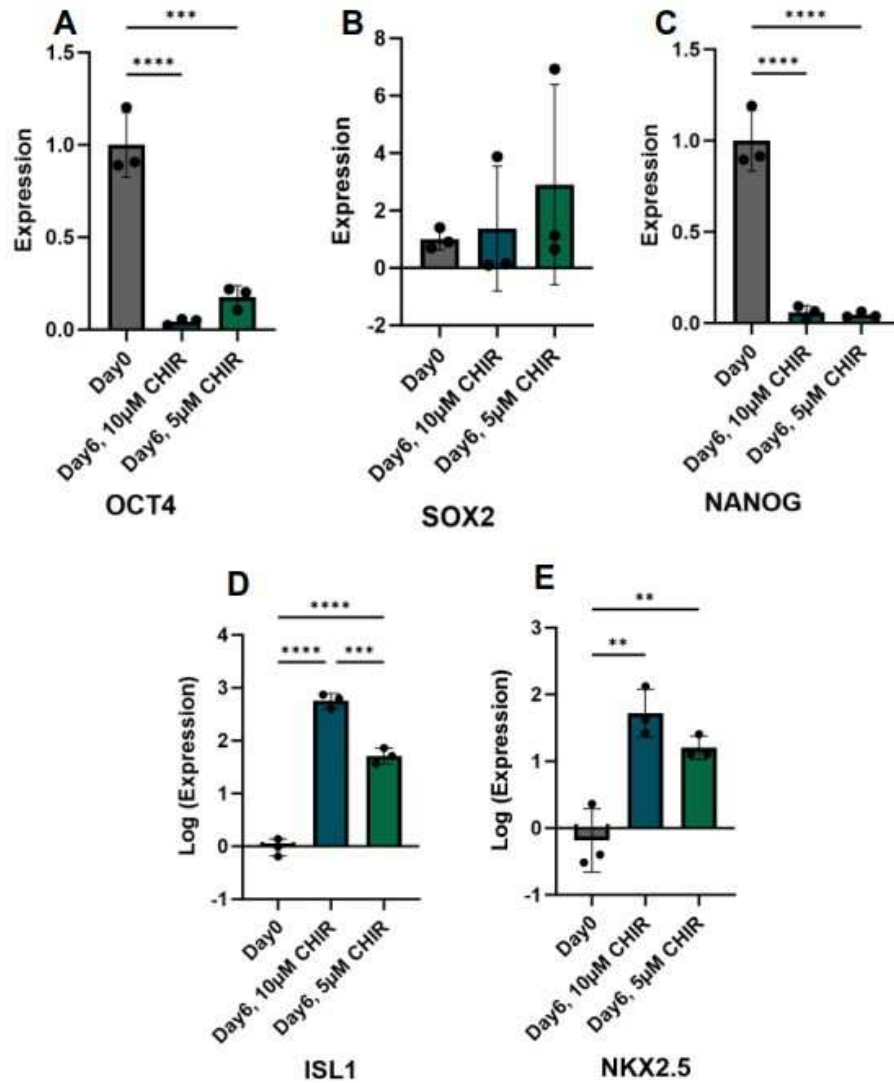

Figure 8: Light microscope images showing the morphology of hiPSCs at day 1: **A** 10 µM CHIR 5x magnification , **B** 10µM CHIR 10x magnification, **C** 10µM CHIR 20x magnification, **D** 5µM CHIR 5x magnification, **E** 5µM CHIR 10x magnification, **F** 5µM CHIR 20x magnification.

4.1.2 Day 6, cardiac progenitors

At day 6, the expression of the pluripotency markers is decreased but it is significant only for *NANOG* and *OCT4,* in both conditions (Fig 9A and 9C). Unexpectedly *SOX2* expression increased, despite it is an unsignificant way, in both conditions. This might be influenced by the high variability of gene expression among the samples of the same group (Fig. 9B). On the other hand, in both protocol settings, *ISL1* and *NKX2.5,* markers of cardiac progenitors, were significantly under-regulated compared to day 0, indicating the progression of the differentiation process. However, for *ISL1*, a significant difference was detected also when comparing the 10µM and 5µM conditions,

suggesting that the former is a stronger trigger of this particular differentiation step (Fig. 9D, 9E).

Moreover, for both protocols, cells showed irregular edges, and appeared less round than before. However, cells grew on the wells differently, in terms of uniformity: some areas have higher confluence, which is more evident for cells initially treated with 10 µM CHIR (Fig. 10A, 10B and 10C)

Figure 9: Expression level at day 6 in hiPSCs treated for 24 hours with either 10µM or 5µM CHIR and for 48 hours with 4µM IWP2: **A** *OCT4,* **B** *SOX2,* **C** *NANOG,* **D** *ISL1,* **E** *NKX2.5.*
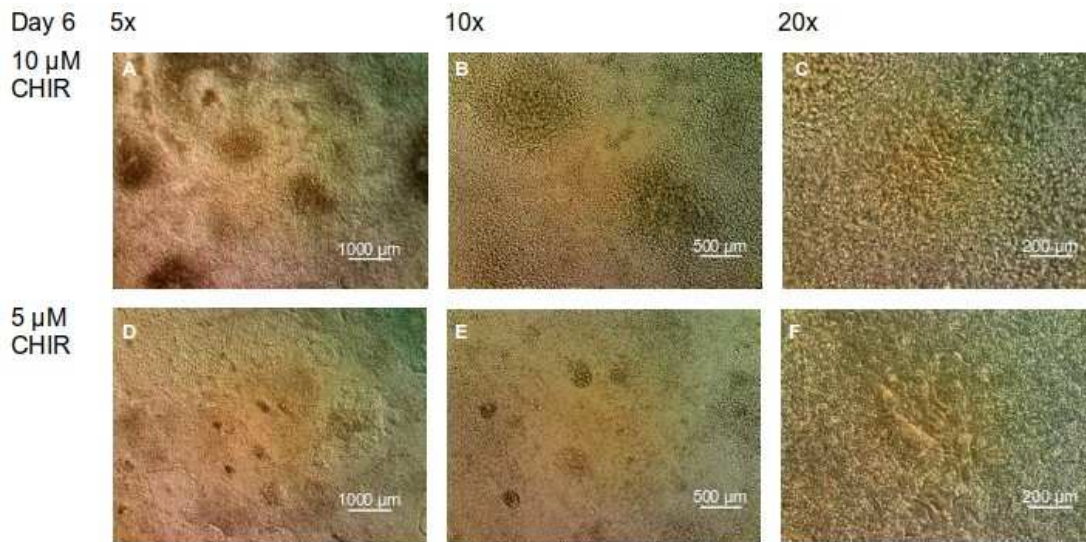
Figure 10: Light microscope images showing the morphology of hiPSCs at day 6: **A** 10 µM CHIR 5x, **B** 10µM CHIR 10x, **C** 10µM CHIR 20x, **D** 5µM CHIR 5x, **E** 5µM CHIR 10x, **F** 5µM CHIR 20x.

4.1.3 Day 12, pro-epicardial cells

At day 7, cells were treated again with CHIR for 48 hours to activate the Wnt signaling to induce the proepicardial differentiation. Setting the proper concentration of this compound also at this stage is critical for an efficient differentiation; therefore two different concentrations, 3 µM and 4µM, were tested.

As expected, the expression of all the pluripotency markers, including *SOX2*, significantly decreased for all conditions (Fig. 11A, 11B, 11C) at day 12. For *SOX2,* the line treated with 10+3µM has a lower significant expression also if compared with 5+3µM (Fig. 11 B).

As result of the differentiation process, proepicardial cells were obtained at day 12, as shown by the increased expression levels of *the* epicardial markers *WT1, TBX18* and *K18* that is an epithelial marker (Fig. 11D, 11E, 11F). Notably, *WT1* expression increased for all four conditions, but resulted statistically significant only for the two conditions 10 µM CHIR when compared with day0. Furthermore, the expression of both conditions was significantly different when compared with the cells initially treated with 5 µM and with each other, where the ones 10+3 µM had a higher *WT1* expression (Fig. 11D). A similar consideration can be done for *TBX18* which expression was significantly different for each condition compared with day 0 as well as

34

for both cell cultures treated with 10µM, if compared with both conditions initially treated with 5µM of CHIR (Fig. 11E). Finally, *K18* significantly increased, comparably for all conditions, with the exclusion of 10+3µM, which, however showed the same trend (Fig. 11F).

At this time point, in addition, other possible outcomes of the differentiation process were taken into account to be tested to characterize the cells and evaluate the differentiation protocols. In particular, the expression of *ACTA2*, a marker of smooth muscle cells, was generally increased but, notably, is significantly increased only for the cells initially treated with CHIR 10 µM compared with day0 and the ones treated with CHIR 5µM (Fig. 11G). Similarly, the expression of *TNNT2*, a cardiomyocyte marker, resulted increase in all conditions (Fig. 11I). The opposite can be said for *PECAM1*, a marker of endothelial differentiation, which was decreased in all conditions but not significant (Fig. 11H).

In general, the results are in line with what was expected, with a decrease in pluripotency markers expression and an increase in the expression of differentiation markers. In line with these results, at this time point cells showed elongated and tapered shape with larger dimensions, very different from the small, round shape of the previous stages (Fig 12).
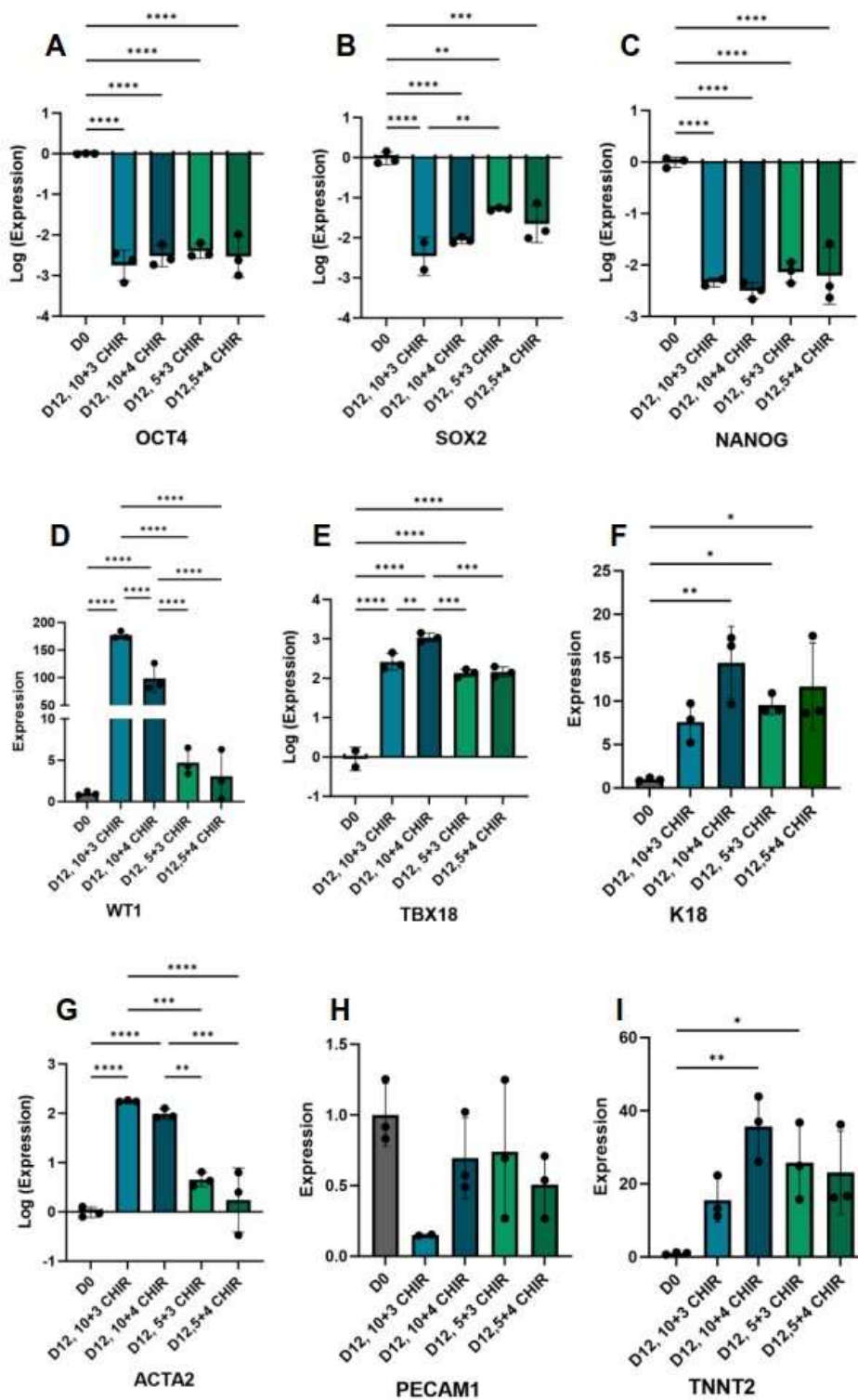
Figure 11: Expression level at day 12 in hiPSCs treated for 24 hours with either 10µM or 5µM CHIR, 48 hours with 4µM of IWP2 and 48 hours of either 3µM or 4 µM CHIR: **A** *OCT4,* **B** *SOX2,* **C** *NANOG,* **D** *WT1,* **E** *TBX18,* **F** *K18,* **G** *ACTA2,* **H** *PECAM1,* **I** *TNNT2.*
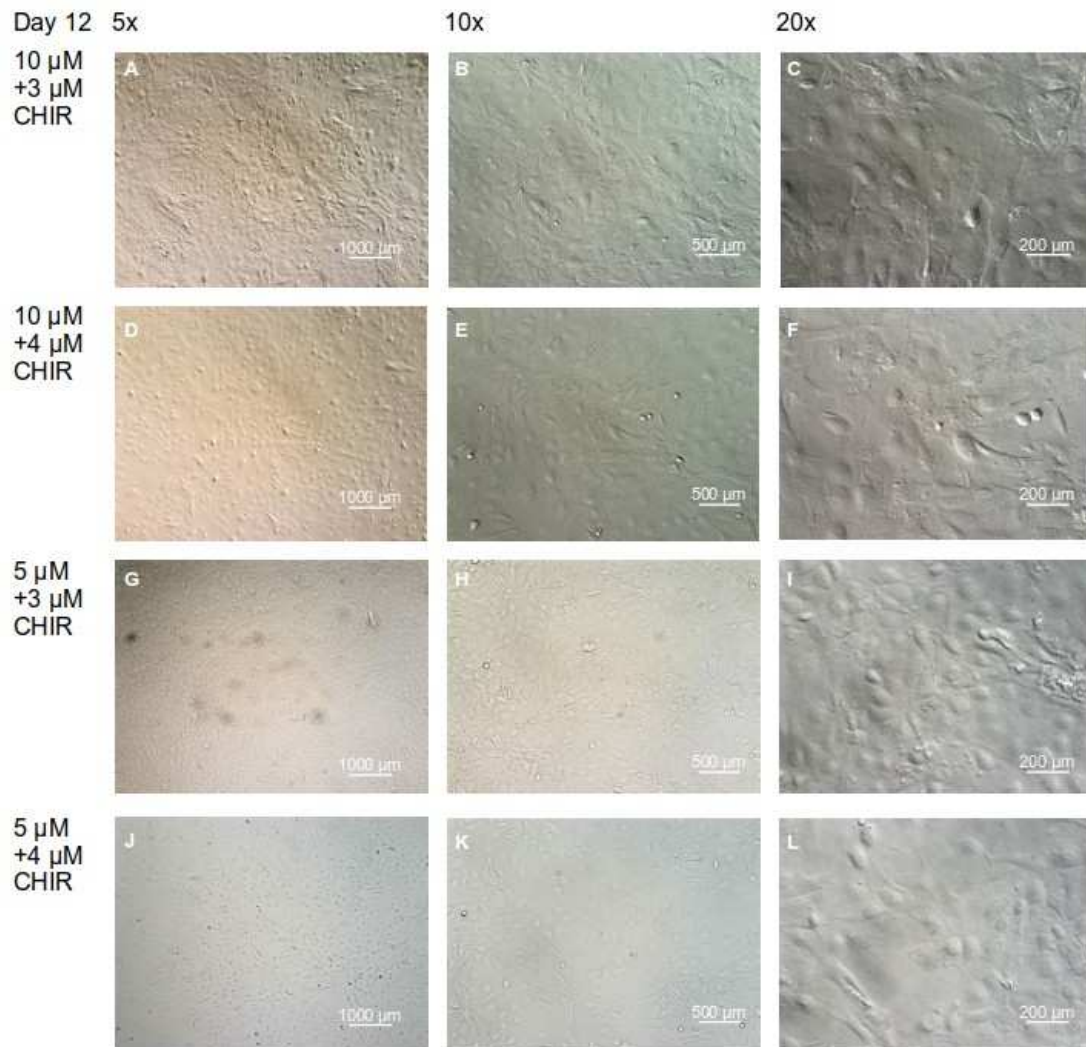
Figure 12: Light microscope images showing the morphology of hiPSCs at day 12: **A** 10+3 µM CHIR 5x, **B** 10+3 µM CHIR 10x, **C** 10+3 µM CHIR 20x, **D** 10+4 µM CHIR 5x, **E** 10+4 µM CHIR 10x, **F** 10+4 µM CHIR 20x, **G** 5+3 µM CHIR 5x, **H** 5+3 µM CHIR 10x**, I** 5+3 µM CHIR 20x, **J** 5+4 µM CHIR 5x, **K** 5+4 µM CHIR 10x, **L** 5+4 µM CHIR 20x.

4.1.4 Day 18, epicardial cells.

At the last time point, where the cells are expected to be fully differentiated into epicardial cells. The expression of *OCT4* and *NANOG* was significantly decreased for each condition compared with day0 (Fig. 13A, 13C), the reduction in expression was also present for *SOX2*, but it was significant only for 10+4μM and 5+3μM.

Regarding epicardial markers, *WT1* and *TBX18* showed the same pattern within all conditions with a significantly increased expression of the marker compared to the control condition. Moreover for both genes, the expression was also significantly increased in cells initially treated with 10μM CHIR compared to both ones treated with 5μM CHIR (Fig. 13D and 13E). As far as the expression of *K18* is concerned, as observed in the previous time point, it resulted increased in all cases compared with the control condition excluding 10+3μM (Fig. 13F).

Finally, for *ACTA2* and *TNNT2* expression was more uniform compared with day 12. In particular, *TNNT2* expression was significantly increased in each condition compared to day 0, and a similar observation can be made for *ACTA2* expression, with the exception of 10+3μM, which resulted slightly lower compared to the other conditions, despite it was not significantly different (Fig. 13G, 13I). The expression pattern of *PECAM1* was more complex: both samples initially treated with 10μM CHIR showed reduced expression compared with day 0, significant only for 10+3μM, but both have significantly lower expression when compared with both conditions initially treated with 5μM CHIR. The latter are also significantly more expressed when compared with the control condition (Fig. 13H).

The Fig. 14 shows that the differentiated cells are, as for the day 12, more elongated and tapered, if compared with the classic hiPSC shape.
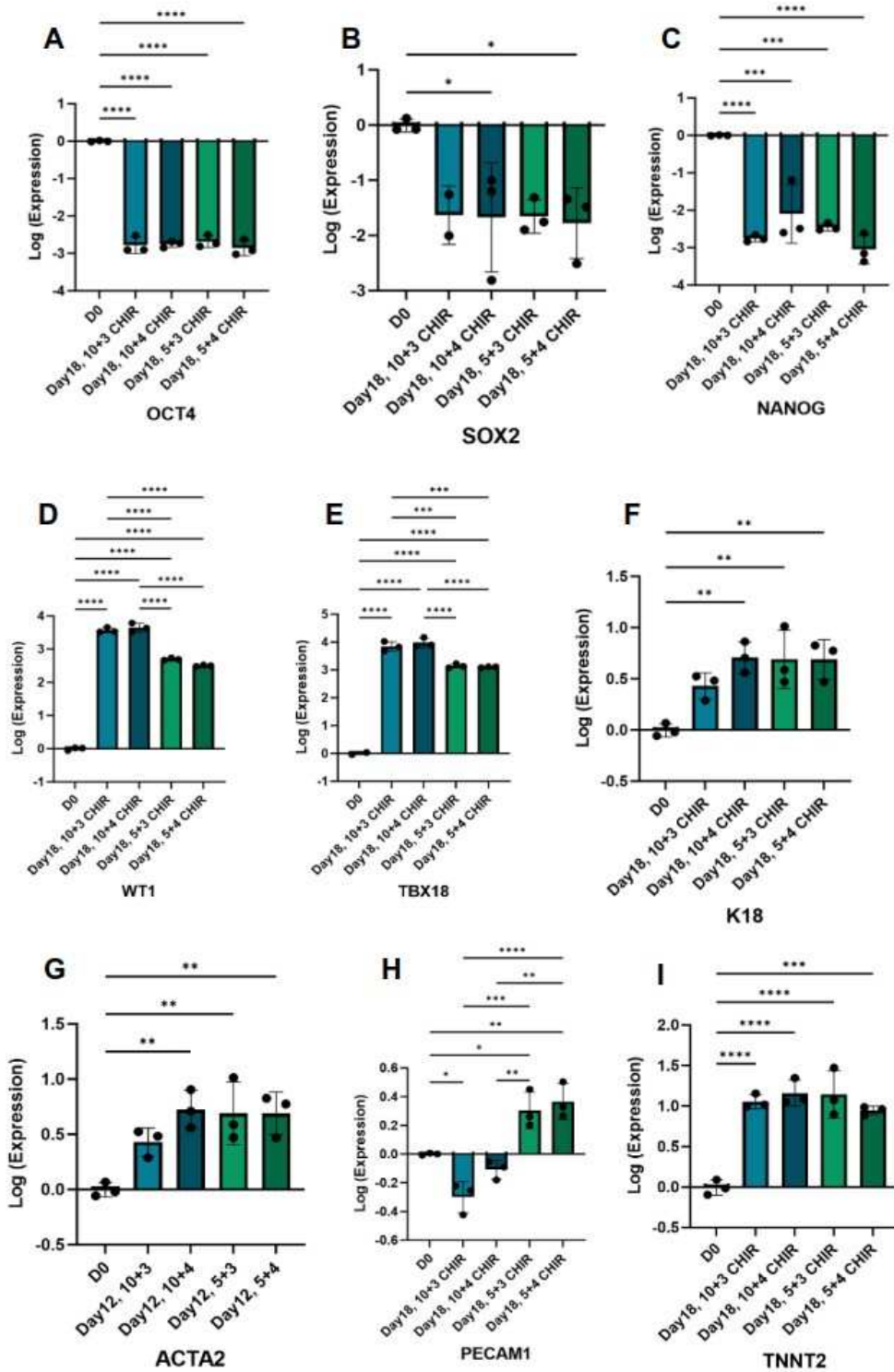
Figure 13: Expression level at day 18 in hiPSCs treated for 24 hours with either 10μM or 5μM CHIR, 48 hours with 4μM of IWP2 and 48 hours of either 3μM or 4 μM CHIR: **A** *OCT4,* **B** *SOX2,* **C** *NANOG,* **D** *WT1*, **E** *TBX18*, **F** *K18*, **G** *ACTA2*, **H** *PECAM1*, **I** *TNNT2.*

39
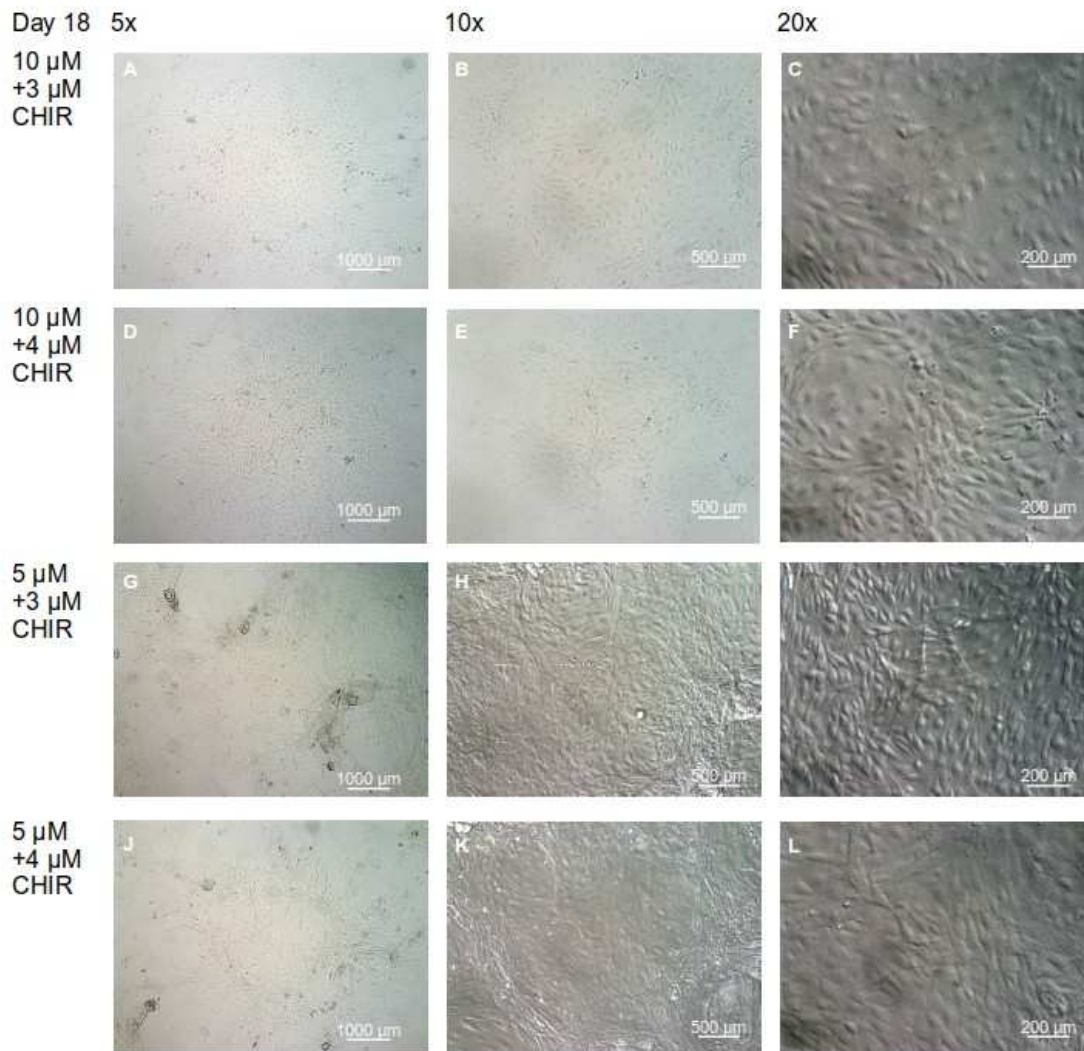
Figure 14: Light microscope images showing the morphology of hiPSCs at day 18: **A** 10+3 µM CHIR 5x, **B** 10+3 µM CHIR 10x, **C** 10+3 µM CHIR 20x, **D** 10+4 µM CHIR 5x, **E** 10+4 µM CHIR 10x, **F** 10+4 µM CHIR 20x, **G** 5+3 µM CHIR 5x, **H** 5+3 µM CHIR 10x, **I** 5+3 µM CHIR 20x, **J** 5+4 µM CHIR 5x, **K** 5+4 µM CHIR 10x, **L** 5+4 µM CHIR 20x.

4.2 Expression of intercalated disc genes in hiPSC-epicardial cells

Specific information on epicardial expression of desmosomal proteins has been difficult to find in the literature. This could provide important information about whether the mutation of desmosomal genes may affect only cardiomyocytes or also epicardial cells, so I quantified RNA expression of pivot desmosomal genes (*JUP, PKP2, DSP, DSC2, DSG*) by qPCR.

On Fig. 15, it can be seen that all desmosomal mRNAs were expressed at day 18. *JUP* was more expressed in the differentiated samples but not significantly (Fig. 15 C), and a similar consideration can be made for *PKP2*, where the difference was significant only between 10+4µM and 5+4µM compared with day 0 (Fig. 15 A). *DSP* was significantly expressed in all conditions of differentiated cells compared with day 0, but the difference was significant comparing 10+4µM with both conditions initially treated with 5µM CHIR and between 10+3µM and 5+3µM (Fig. 15 E). For *DSC2,* the expression was significantly increased in the two conditions initially treated with 10µM CHIR compared with all other conditions, whereas the cells initially treated with 5µM CHIR showed an increase in expression, but not significant (Fig. 15 B).
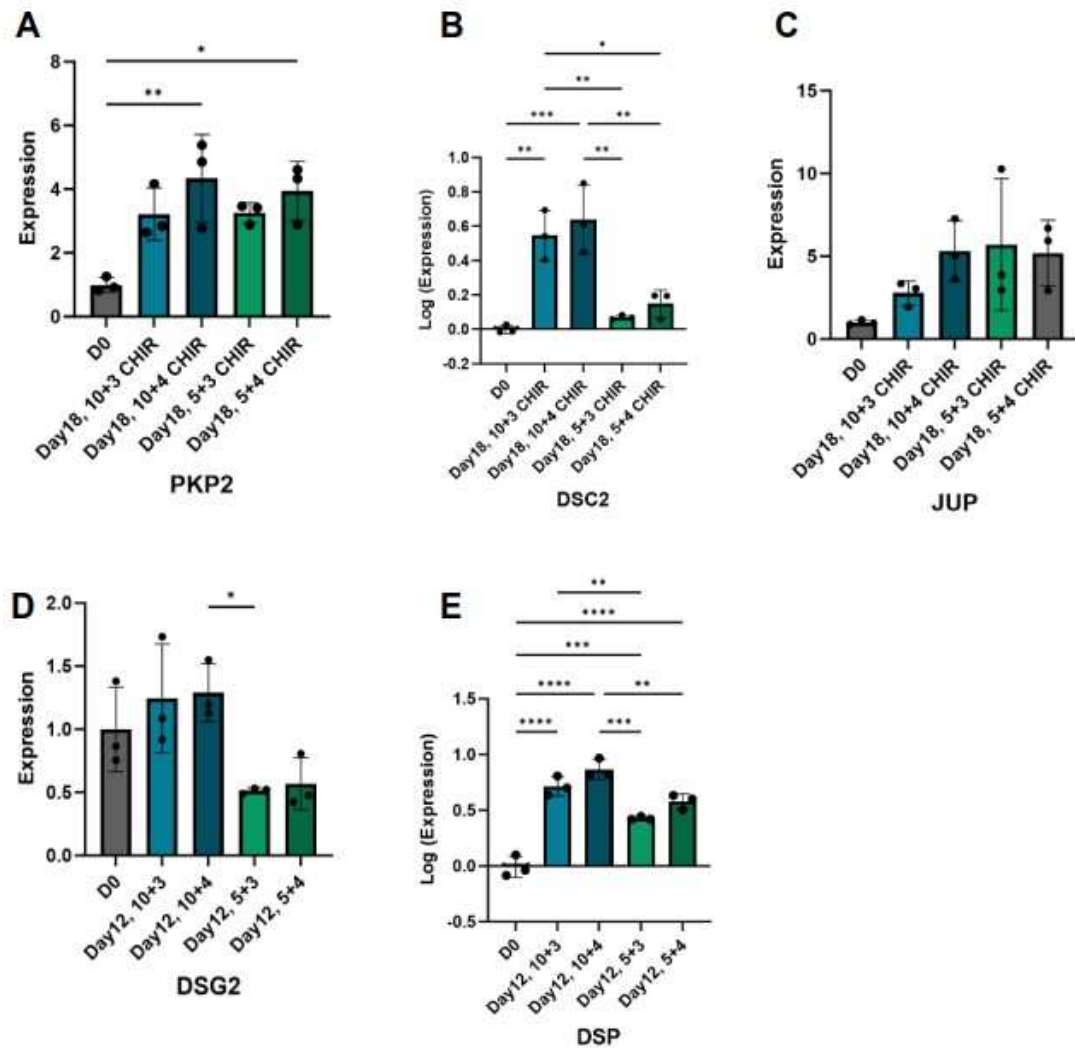
Figure 15: Expression level at day 18 in hiPSCs treated for 24 hours with either 10μM or 5μM CHIR, 48 hours with 4μM of IWP2 and 48 hours of either 3μM or 4 μM CHIR: **A** *PKP2,* **B** *DSC2,* **C** *JUP,* **D** *DSG2*, **E** *DSP.*

DSC2 Ct values:

Day 0	10+3μM	10+4μM	5+3μM	5+4μM
25.19	24,8	23,71	25,33	25,59
26.29	23,25	23,73	25,89	25,8
25.91	23,38	24,49	25,3	24,97

Table 11: Ct value of *DSC2* at day 18.

DSG2 Ct values:

Day 0	10+3μM	10+4μM	5+3μM	5+4μM
25.14	25.30	24.32	25.84	26.26
26.18	25.35	24.13	26.02	25.66
24.59	24.69	25.21	25.37	25.21

Table 12: Ct value of *DSG2* at day 18.

DSP Ct values:

Day 0	10+3μM	10+4μM	5+3μM	5+4μM
25.75	23.97	22.60	24.02	24.45
25.76	22.82	22.52	23.76	23.62
25.60	23.42	23.22	23.81	23.41

Table 13: Ct value of *DSP* at day 18.

JUP Ct values:

Day 0	10+3μM	10+4μM	5+3μM	5+4μM
30.34	28.98	28.29	28.00	28.13
30.52	28.28	28.31	27.79	27.41
29.76	28.75	28.34	27.85	28.80

Table 14: Ct value of *JUP* at day 18.

PKP2 Ct values:

Day 0	10+3μM	10+4μM	5+3μM	5+4μM
28.05	26.16	25.48	25.17	25.26
27.60	24.78	24.94	25.13	25.54
27.92	24.97	25.83	24.96	24.64

Table 15: Ct value of *PKP2* at day 18.

Surprisingly, *DSG2* showed an irregular expression pattern among the different conditions: the cells initially treated with 10μM CHIR showed higher

expression, compared with day0, while the ones treated with 5µM CHIR have lower expression. The only comparison that is significantly different is only between 10+4µM and 5+3µM (Fig. 15D).

In order to make the data more meaningful, Tables 11-15 show the Ct value obtained from the qPCR analyses. The values are never higher than 31, which means that the expression can be considered good.

The differentiated cells express *GJA1,* which resulted significantly reduced if compared with day0. Moreover its expression was higher in cells initially treated with 10µM CHIR and showed a significant difference between the 10+3µM and 5+4µM conditions.

Using an antibody specific for Cx43, available in the lab, was possible to conduct immunostaining analysis to evaluate the expression of this gap junction protein in the differentiated cells at day 18.

In agreement with the gene expression, Fig. 17 shows that Cx43 is expressed in the differentiated cells and, as expected, its localization is mainly at the membrane level, concentrated in spots, with some accumulation in the nucleus.
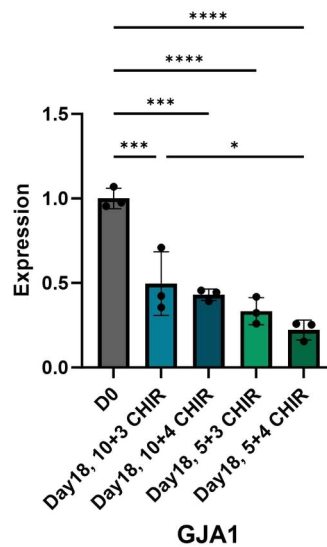
Figure 16: Expression level of *GJA1* at day 18 in hiPSCs treated for 24 hours with either 10µM or 5µM CHIR, 48 hours with 4µM of IWP2 and 48 hours of either 3µM or 4 µM CHIR.
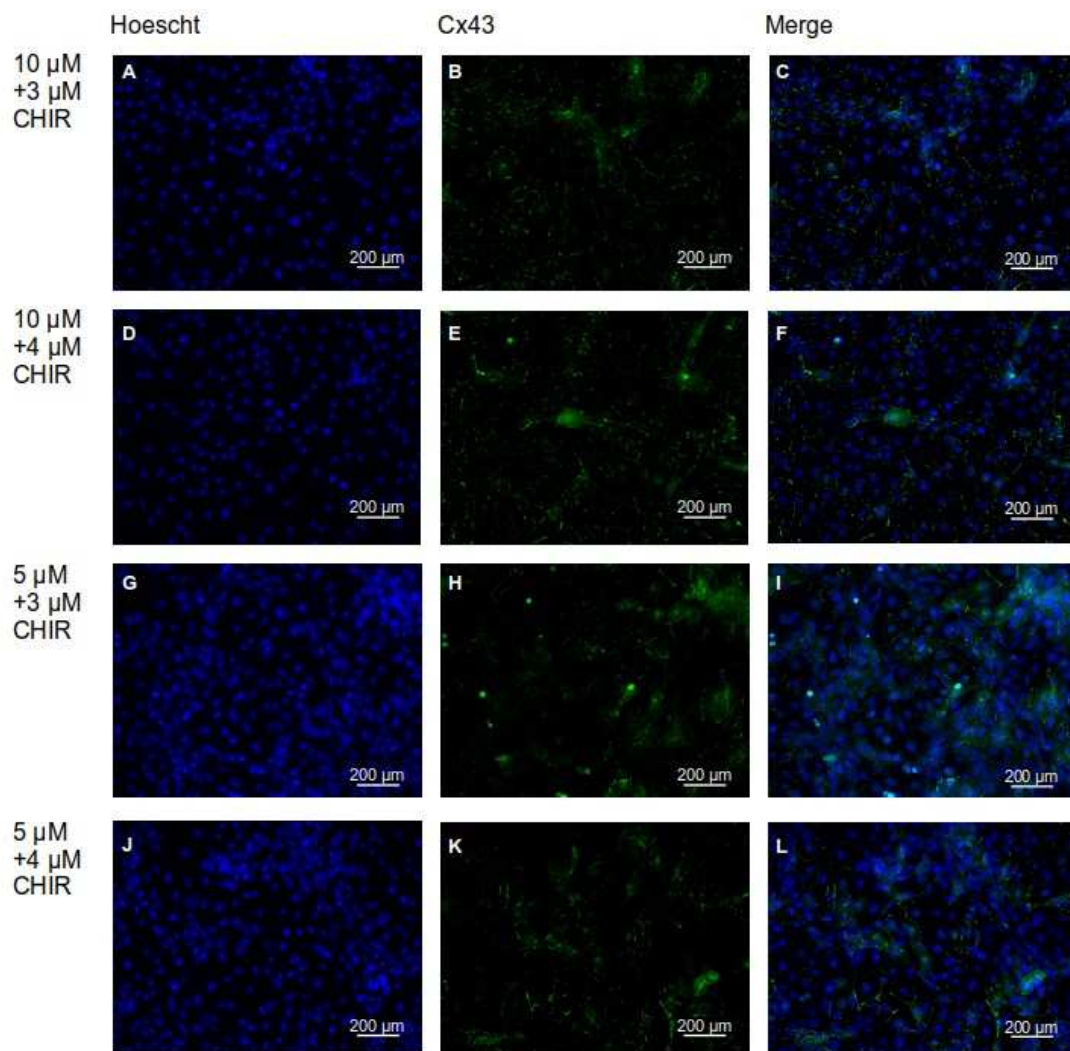
Figure 17: Fluorescent microscope images showing the localization of nuclei(A, D, G and J), Cx43 (B, E, H, K) and the merged image (C, F, I, L) at the 20x magnification for the condition: **A-C** 10µM+3µM CHIR, **D-F** 10µM+4µM CHIR, **G-I** 5µM+3µM CHIR, **J-L** 5µM+4µM CHIR.

4.3 RNA-sequencing data analysis

In the lab where I performed my thesis, data obtained from RNA-seq on iPSC-CMs harboring the p.Q558* mutation and their isogenic controls kept on culture until day 30, were available. A preliminary analysis was already performed and produced a list of significantly enriched GO terms, that I further analyzed to select relevant genes to be validated. My analysis led to the selection of 22 statistically enriched GO terms involved in inflammation, apoptosis and immune response, which are hallmark of ACM (Tab. 16).

Pathway	p-adj
GOMF_CHEMOKINE_RECEPTOR_BINDING	0.002787
GOMF_CYTOKINE_ACTIVITY	0.004016
GOBP_LEUKOCYTE_DIFFERENTIATION	0.006429
GOBP_RESPONSE_TO_CYTOKINE	0.007142
GOBP_POSITIVE_REGULATION_OF_CELL_DEATH	0.007492
GOBP_NEGATIVE_REGULATION_OF_MEGAKARYOCYTE_DIFFERENTIATION	0.009796
GOMF_CHEMOKINE_ACTIVITY	0.012771
GOMF_CXCR_CHEMOKINE_RECEPTOR_BINDING	0.015991
GOBP_MYELOID_LEUKOCYTE_ACTIVATION	0.016137
GOBP_EXTRINSIC_APOPTOTIC_SIGNALING_PATHWAY_VIA_DEATH_DOMAIN_RECEPTORS	0.017471
GOBP_LEUKOCYTE_MEDIATED_IMMUNITY	0.019447
GOBP_REGULATION_OF_CELL_DEATH	0.022555
GOBP_REGULATION_OF_EXTRINSIC_APOPTOTIC_SIGNALING_PATHWAY_VIA_DEATH_DOMAIN_RECEPTORS	0.023719
GOBP_CYTOKINE_MEDIATED_SIGNALING_PATHWAY	0.024067
GOBP_LEUKOCYTE_MIGRATION	0.030062
GOBP_POSITIVE_REGULATION_OF_LEUKOCYTE_DEGRANULATION	0.031215
GOBP_POSITIVE_REGULATION_OF_MAST_CELL_ACTIVATION	0.031549
GOBP_INTERLEUKIN_1_BETA_PRODUCTION	0.033728
GOBP_EXTRINSIC_APOPTOTIC_SIGNALING_PATHWAY	0.039703
GOBP_INTERFERON_GAMMA_MEDIATED_SIGNALING_PATHWAY	0.048948
GOBP_MYELOID_LEUKOCYTE_DIFFERENTIATION	0.049133
GOBP_APOPTOTIC_PROCESS	0.049416

Table 16: List of the selected enriched GO terms. GOMF, GO term molecular function; GOBP, GO term biological process.

A list of about 500 genes was generated from these GO terms. Genes to be validate were first selected based on $\log_2$Fold change and adjusted p-value (p<0.05). Next, genes were selected based on their relevance for the disease, accordingly to previous knowledge and literature. Eventually, 21 genes were selected for the validation (Tab. 17).

Gene	LOG$_2$(Fold Change)	p-value
AKAP13	0.9906	1.920E-08
ANKRD1	3.8324	2.020E-218
AREG	6.5561	8.612E-10
CCND3	-1.4956	2.360E-12
CD36	2.5523	2.028E-14
CX3CL1	4.3474	8.329E-07
CXCL16	3.2475	3.446E-10
CYBRD1	-2.9162	6.125E-20
EGR1	3.0891	2.269E-13
F2RL1	4.5573	2.643E-22
FHL2	1.5111	2.913E-13
GDF15	3.4535	1.835E-09
IL17RD	1.6530	3.305E-15
JAK2	1.8468	3.596E-10
OCLN	3.0802	3.845E-17
PTGES	-1.6624	1.996E-06
SERPINE1	2.3398	1.633E-09
TLR4	0.6411	6.692E-04
TNFRSF10A	4.8609	8.577E-08
TNFRSF12A	2.1806	1.385E-25
UBD	8.2718	2.756E-07

Table 17: Genes selected from the list generated from the analysis of the GO terms.

In addition from the same RNA-seq dataset the genes that were significantly differently expressed in the *in vitro* hiPSC-CMs model were compared with the those differently expressed in the heart of a 3 month-old knock-in mouse model for ACM harboring the same mutation. From this intersection, a list of 54 genes was produced (Fig. 18). Applying the same parameters for $\log_2$Fold change, p-value and literature, three genes were selected (Tab. 18).
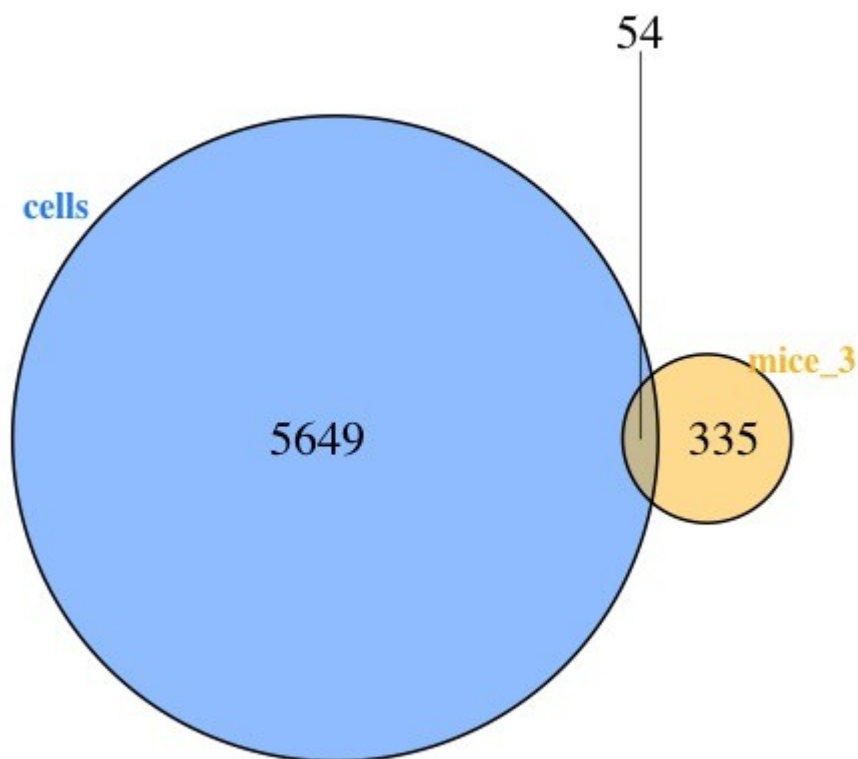
Figure 18: Diagram of the intersection of the 2 groups of significant genes from the 2 RNA-sequencing analyses, already available in the laboratory.

Gene	Log$_2$(Fold Change) from RNA-seq	p-value from RNA-seq
H2BC4	-0.7816	1.224E-02
NDUFA13	0.8001	1.683E-02
TMEM65	0.8579	3.767E-05

Table 18: Genes selected from the intersection of the genes differently expressed in both models. Value of Log$_2$Fold Change and p-value from cells RNA-sequencing.

Therefore, overall I selected 24 genes to be validated by qPCR.

4.4 Selection of the housekeeping gene

First, two different housekeeping genes, *L7* and *GAPDH,* were tested by qPCR on mutant and control hiPSC-CMs. Their expression was compared among the biological replicates and the most stable housekeeping gene was selected for the subsequent analyses (Fig 19).
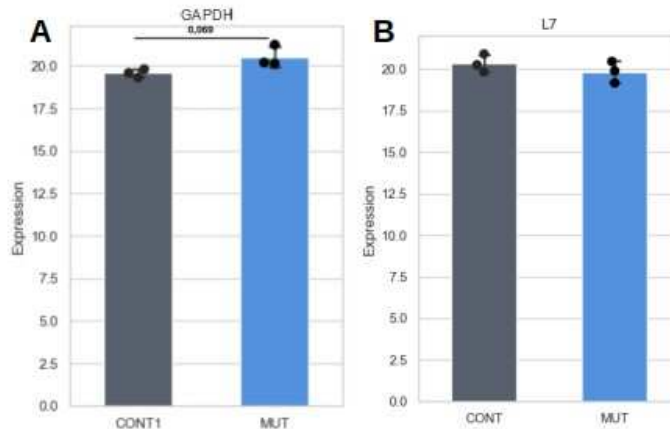
Figure 19: A Average value of the biological replicates of mutant and controls for *GAPDH* (p-value=0,069), **B** Average value of the biological replicates of mutant and controls for *L7*. CONT isogenic control; MUT, mutant cells harboring the p.Q558* mutation.

Figures 19A and 19B show the mean of the biological replicates and, as shown, the difference between the two groups in *GAPDH* expression has a p value of 0.069. For these reasons, I chose to use *L7* as the housekeeping gene (p=0,366): the values in the two conditions are more similar to each other with the same range of variation, while the range of *GAPDH* is higher in the mutant samples.

4.5 RNA-sequencing data validation by qPCR

The 21 genes that were selected in the analysis of the transcriptome dataset from hiPSC-CMs only were quantified using qPCR from RNA isolated from mutant and control cells kept in culture until day 30 available in the lab (Fig 20, Tab.19).

The expression level of the three genes differently expressed in both ACM iPSC-CMs and heterozygous mice was also assessed in samples harvested at culture day 30.

As can be seen from Fig. 20 and Fig. 21 most of the genes analyzed were not confirmed by qPCR analysis. In particular, from Fig 20 A, 20B, 20F-U and Fig 21A expression is not significantly different, and, in addition, 10 of these genes have a Log_2Fold Change reversed than the one obtained from RNA-sequencing (Tab. 19). These results may be affected by the high variability of expression within samples.

However, 5 of these genes are of interest: *AREG*, *CCND3*, *CD36* (Fig 20C, 20D and 20E) and *NDUFA13*, *TMEM65* (Fig 21B and 21C). In particular *AREG*, *CCND3* and *NDUFA13* have almost significant expression (less than 0.07), but for *CCND3* the trend is a decrease in expression, as observed in RNA-sequencing, while for the other two genes the trend is opposite to that one obtained from RNA-sequencing (Tab. 19).
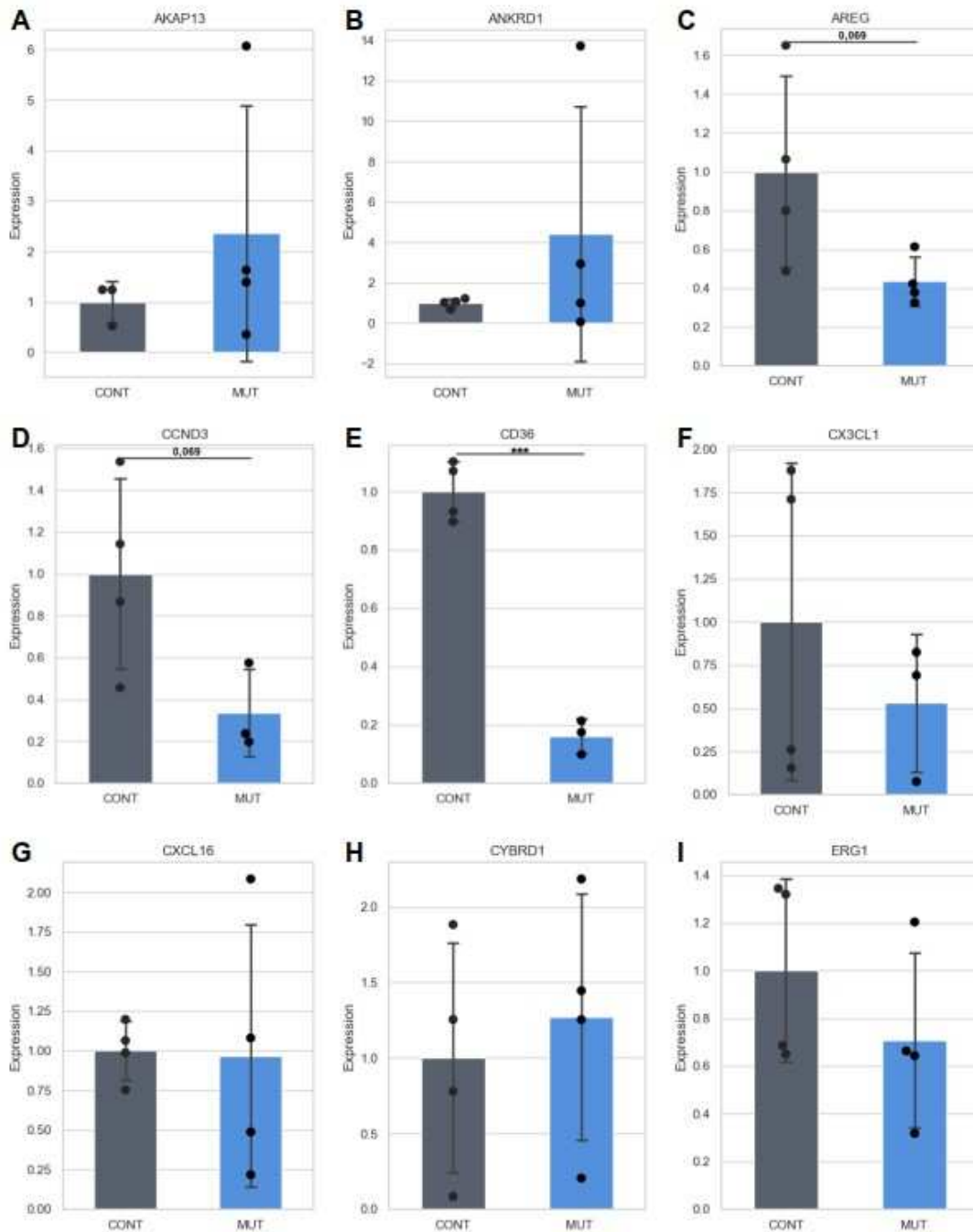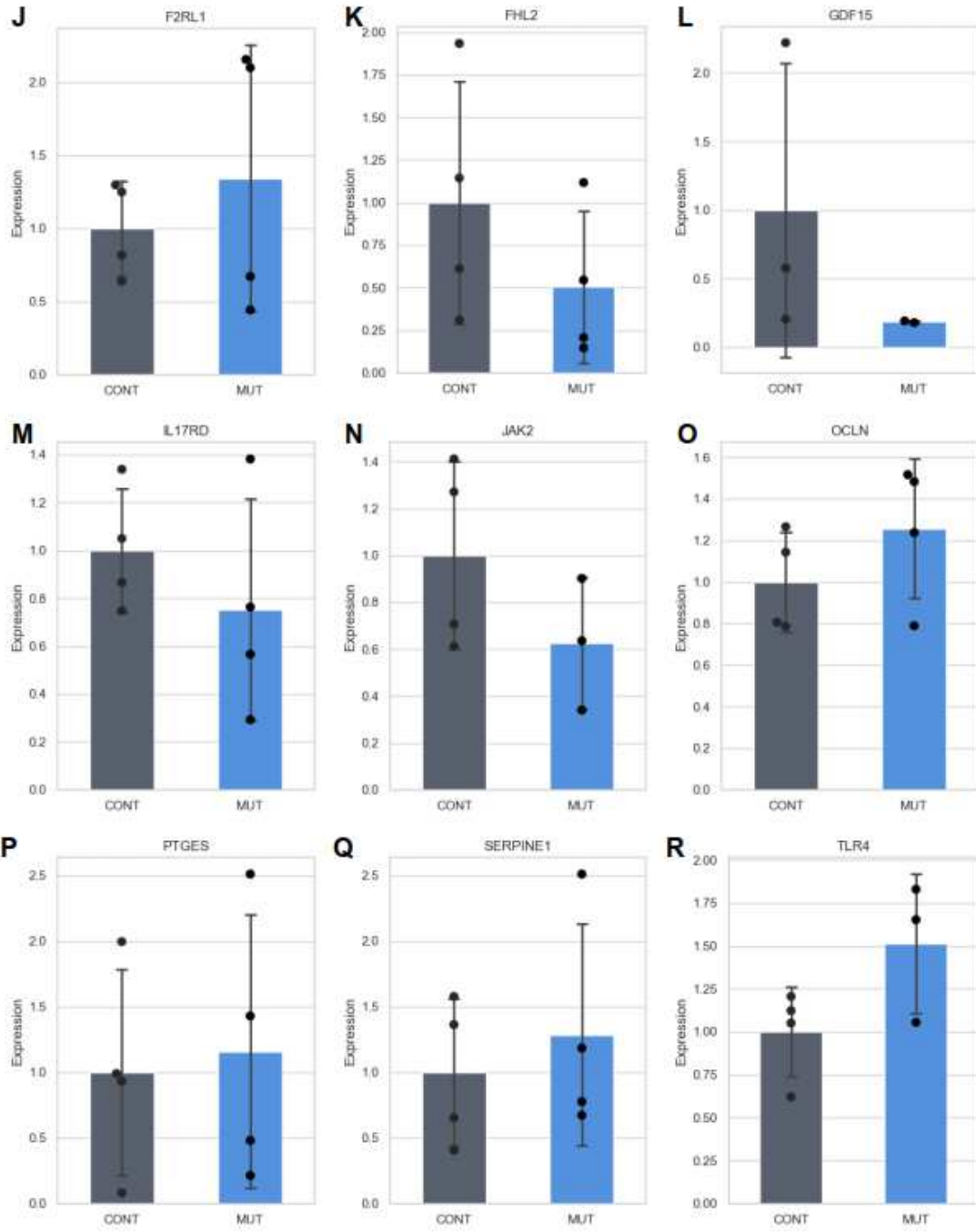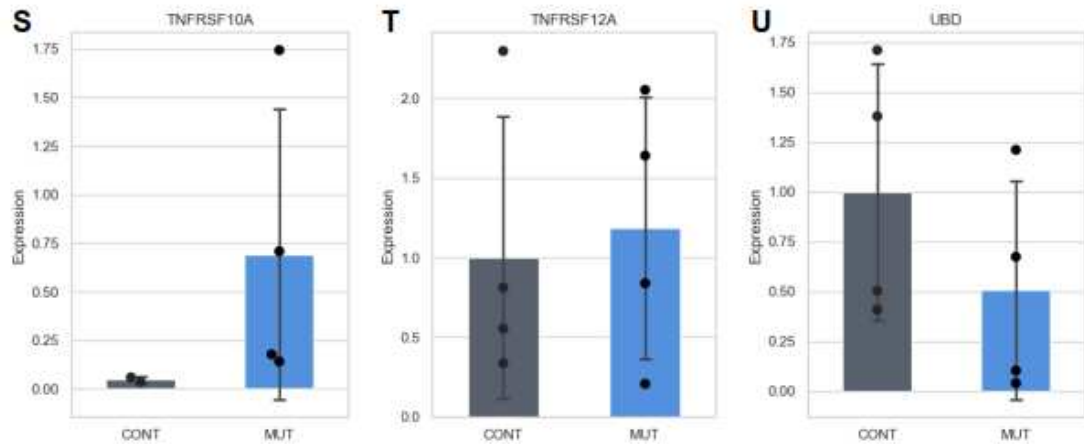
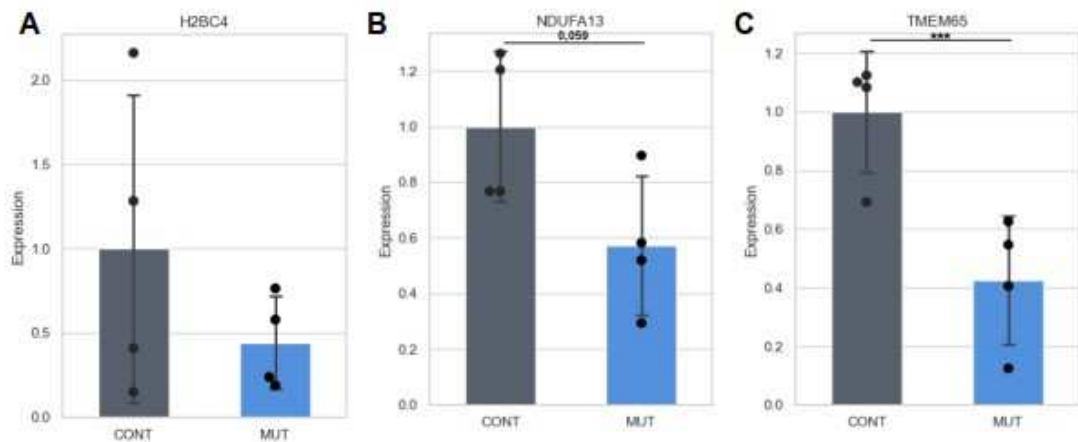**Figure 20:** Expression of selected genes at day 30: **A** *AKAP13*, **B** *ANKRD1*, **C** *AREG* (p-value 0,068), **D** *CCND3* (p-value 0,069), **E** *CD36* (p-value <0,001), **F** *CX3CL1*, **G** *CXCL16*, **H** *CYBRD1*, **I** *ERG1*, **J** *F2RL1*, **K** *FHL2*, **L** *GDF15*, **M** *IL17RD*, **N** *JAK2*, **O** *OCLN*, **P** *PTGES*, **Q** *SERPINE1*, **R** *TLR4*, **S** *TNFRSF10A*, **T** *TNFRSF12A*, **U** *UBD*. CONT, isogenic control, MUT, mutant cells harboring the p.Q558* mutation.

Figure 21: Expression of the genes selected from the intersection with mouse model at day 30: A H2BC4, B NDUFA13 (p-value 0,059), C TMEM65 (p-value<0,001). CONT, isogenic control, MUT, mutant cells harboring the p.Q558* mutation.

GENE	LOG$_2$(Fold Change) from RNA-seq	LOG$_2$(Fold Change) at day 30	LOG$_2$(Fold Change) at day 60
AKAP13	0.9906	0.8786	
ANKRD1	3.8324	2.1782	1.6889
AREG	6.5561	-1.0643	-0.7789

Gene	Day 30	Day 60	Expected
CCND3	-1.4956	-1.4408	
CD36	2.5523	-2.6202	1.2014
CX3CL1	4.3474	0.2196	
CXCL16	3.2475	-0.0283	
CYBRD1	-2.9162	1.0240	
EGR1	3.0891	-0.4154	0.4297
F2RL1	4.5573	0.4855	0.8826
FHL2	1.5111	-0.6755	
GDF15	3.4535	0.2358	
H2BC4	-0.7816	-0.5383	-0.9826
IL17RD	1.6530	-0.3791	-0.0566
JAK2	1.8468	-0.5846	-0.0013
NDUFA13	0.8001	-0.7634	0.2842
OCLN	3.0802	0.3612	1.1943
PTGES	-1.6624	0.8992	
SERPINE1	2.3398	0.5662	
TLR4	0.6411	1.9180	
TMEM65	0.8579	-1.3737	-0.0325
TNFRSF10A	4.8609	1.9180	
TNFRSF12A	2.1806	0.6241	
UBD	8.2718	-0.7187	0.2963

Table 19: Log_2(Fold Change) comparison between the obtained value of Log_2(Fold Change) at day 30 and day 60 and the expected value from RNA-sequencing at day 30.

Also for *CD36* (Fig. 20E) and *TMEM65* (Fig. 21C), the only two genes significantly differentially expressed at day 30, the trend is opposite to the expected one (Tab. 19).

For some genes, it was possible to asses the expression level also in hiPSC-CMs kept in culture until day 60 already available in the lab (Fig. 22 and Fig. 23). Unfortunately, the cDNA or RNA available was not sufficient to assess the expression of all the selected genes.

The results of the analysis at day 60 are in line with previous results: most of the RNA-sequencing data were not validated (Fig. 22 A, 22C-F, I and Fig. 23). Only 3, out of 9 genes, have a Log2Fold variation trend inconsistent with what is seen in RNA sequencing (Tab. 19). Furthermore, 2 of these 9 genes

change trend when compared with the analysis at day 30, following RNA-sequencing results (Tab. 19).
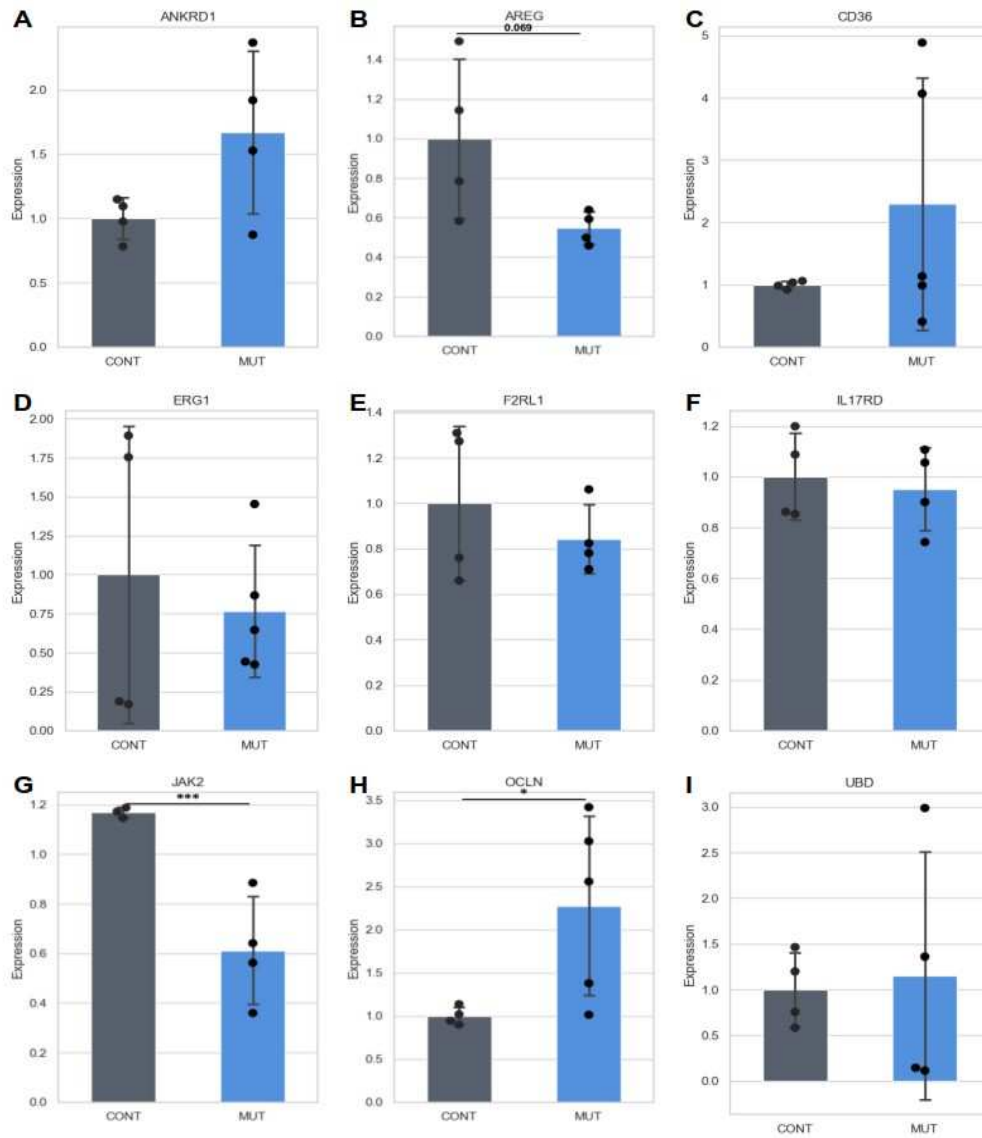
Figure 22: Expression of the genes at day 60: **A** *ANKRD1*, **B** *AREG* (p-value 0,069), **C** *CD36*, **D** *ERG1*, **E** *F2RL1*, **F** *IL17RD*, **G** *JAK2* (p-value <0,001) , **H** *OCLN* (p-value<0,05)**, O** *UBD*. CONT, isogenic control, MUT, mutant cells harboring the p.Q558* mutation.
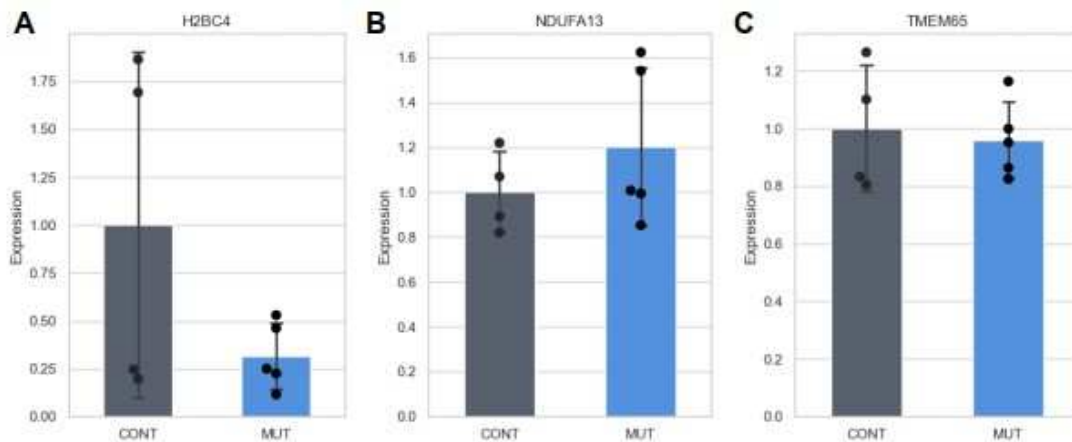
Figure 23: Expression of the genes selected from the intersection with mouse model at day 60: **A** *H2BC4*, **B** *NDUFA13*, **C** *TMEM65*. CONT, isogenic control, MUT, mutant cells harboring the p.Q558* mutation.

In agreement with data obtained at day 30, at day 60 *AREG* (Fig. 22B) resulted downregulated (p-value=0,069), still showing the opposite trend compared to RNA-seq data (Tab. 19). Moreover, also in the case of *JAK2* and *OCLN* (Fig. 22G, 22H) expression is significantly down- and upregulated, respectively, as observed at day 30, but only *OCLN* Log$_2$Fold Change is coherent with the RNA-seq data (Tab. 19).

4.6 Assessment of *TNNT2* expression in hiPSC-CMs

From the histograms in the previous paragraph, it is possible to observe a great variability in the expression of the selected genes within both the control and mutated groups. One possible cause of this broad variability might be due the differences in the differentiation efficiencies. To assess whether the cells used for the RNA-seq validation differentiated into hiPSC-CMs in a comparable way (both at day 30 and at day 60), I evaluated the expression of the cardiac marker *TNNT2* by qPCR (Fig. 24).

From the Fig. 24 is possible to see that there is a wide variation but that is in a similar range with no significant difference between the two conditions, at day 30. For the second time point, the variation is wider for the control condition but 6 sample of isogenic controls were available for this test and only 3 of the mutated condition.
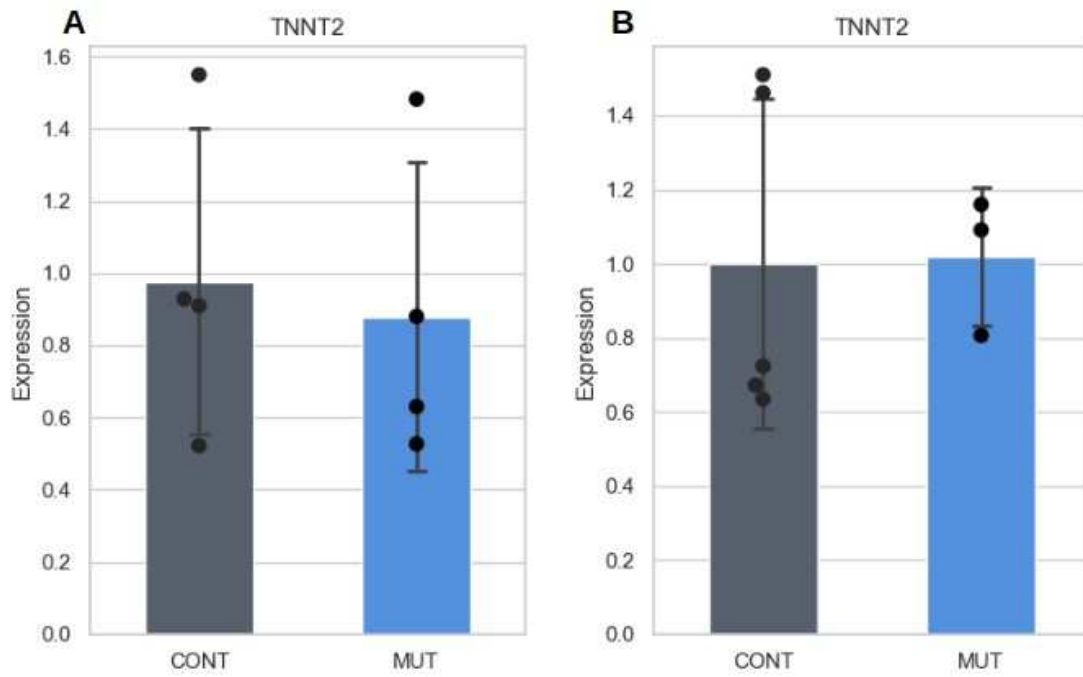
Figure 24: Expression of *TNNT2* at day 30 (**A**) and day 60 (**B**). CONT, isogenic control, MUT, mutant cells harboring the p.Q558* mutation.

5 DISCUSSION

Arrhythmogenic cardiomyopathy is a genetic disease characterized by incomplete penetrance and variable expressivity. It manifests itself with fibrofatty myocardial replacement that can origin from different sources, possibly including epicardial cells [10]. In particular, cells derived from the epicardium are involved in cardiac remodeling in pathological conditions [20] and studies have shown that hiPSC-epicardial cells can display spontaneous fibro-fatty differentiation in ACM, which was not observed in hiPSC-CMs [10]. In addition, similar behavior was observed following the silencing of PKP2, DSP or JUP expression, highlighting the possible effect of these mutations in epicardial cells and the consequent role in ACM [10].

Considering the relevance of setting an *in vitro* system assess the involvement of epicardial cells in ACM, in this thesis, based on the protocol of Bao et al [24], I tested different conditions and optimized the procedure to differentiate hiPSC into hiPSC-epicardial cell in our lab.

5.1 Differentiation protocol

CHIR is an inhibitor of glycogen synthase kinase 3 (GSK-3) that acts a Wnt activator, that is needed at day 0 and at day 7.It binds secreted Wnt proteins to their receptor and is considered highly specific and safe. Because CHIR concentration is crucial for the optimal differentiation process, I tested four different conditions in hiPSCs.

All the differentiation protocols succeeded in generating hiPSC-epicardial cells as observed by elongated morphology of the cells at day 18, comparable while Bao et al. describe it as cobblestone morphology [24]. However the presence of other type of cardiac cells cannot be completely excluded.

From the analysis of the epicardial and epithelial markers, that should be both increased in the differentiated cells, it is possible to say that *WT1* and *TBX18,* the two epicardial markers, have the expected result for all conditions while, the epithelial marker *K18*, is not significantly more expressed in the 10+3μM setting (Fig 11F and Fig 13F). For this reason the protocol might not have an high differentiation efficiency [24]. Bao protocol advises to work with a cell population expressing, for more of the 90% of it, *BRA* on day 1, and more than 50% *ISL1* and *NKX2.5* on day 6, to obtain a final population that is for

more than 90% of it *WT1+* [24]. In addition the *K18 and TBX18* expression is expected to be always high during the long therm maintenance so those analyses should be repeated after day 18 to assess if the expression is maintained [24]. The lack of cells after day 18 is a limitation of the study because it prevents the long-term maintenance and behavior of the culture from being evaluated.

The expression of possible markers of cardiac cells can give important information regarding the possibility that the cells differentiate in alternative type of cells [25]. As the figure 25 shows from the same cardiac progenitor expressing *NKX2.5* and *ISL1*, with a different treatment, is possible to obtain smooth muscular cells (*ACTA2*), endothelial cells (*PECAM1*) and cardiomyocyte (*TNNT2*). For the same reason Bao et al studied the TNNT2 expression by flow cytometry, in the differentiated cell and, results shows the expression of TNNT2 is present, similar to the isogenic control and less then cardiomyocyte [24]. This result is different from the ones obtained because the expression a day 12 and 18 is higher if compared with the control (Fig. 13). ACTA2 and PECAM1 weren't quantified in Bao et al, but a better understanding of their behaviour could be important to characterize the cells in terms of future differentiation into fibroblasts or smooth muscle cells.
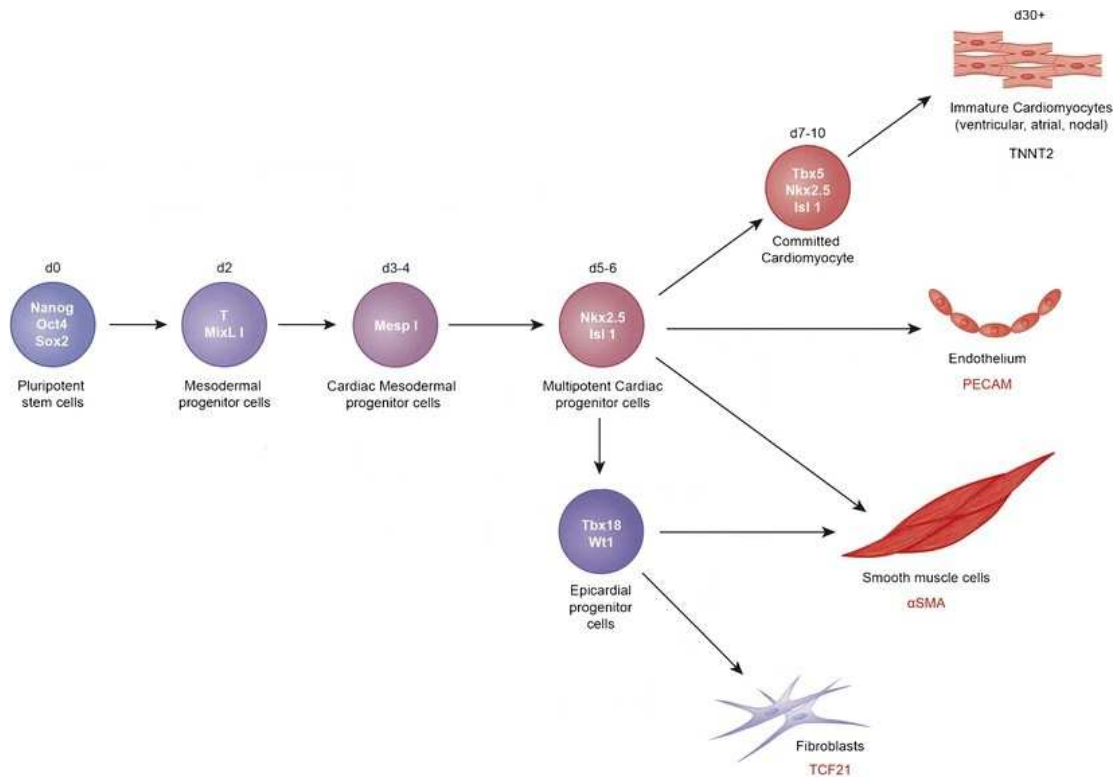
Figure 25: Differentiation line of cardiac cells (different protocol) with the expressed markers. Adapted from [25]

According to the results, the cells generated with all four protocols increased expression of both smooth muscle cell and cardiomyocyte markers, with some difference between days 12 and 18. On the other hand, endothelial differentiation marker expression increased only in cells treated with 5µM CHIR initially, suggesting differentiation in this direction. This result may suggest that there is nevertheless lateral differentiation in other cardiac cell types, and to quantify this it would have been possible to use flow cytometry to sort the cells according to the markers expressed, which would have allowed better quantification and characterization of the cells. Furthermore, the same method could have been used during the differentiation process to work with a purer culture; the lack of this application is a major limitation of the study. This may be important in the future because, as shown in Figure 25, the possibility of further differentiation of the derived epicardial cells into endocardial cells and fibroblasts, starting from a pure culture, may help to obtain a better differentiation of these cells. Finally, the ability to derive all types of cardiac cells is an important step that may lead to the generation of co-culture and 3D models in the future, that have the advantage to allow a better maturation of hiPSC-CMs [26].

Taken together, the results do not allow us to identify a better protocol to generate hiPSC-epicardial cells, but 10+4µM is the condition that best meets expectations in terms of expression of differentiation markers, even though the expression of ACTA2 and TNNT2 is significantly increased compared to the control condition.

5.2 Intercalated disc gene expression in hiPSC-epicardial cells

Quantification shows that there is also expression of desmosomal genes and GJA1 in hiPSC-epicardial cells generated using the four protocols. In addition, the expression of protein Cx43 was confirmed by immunocytochemistry and it is possible to spot it in the membrane and in the nucleus. The literature reports that in the nuclei, Cx43 can bind β-actin that stops, slows down Cx43 transport [27]. An important limitation of the study is that, in order to have actual information on the desmosomal composition of epicardial cells, it would have been necessary to perform immunostaining

with the application of a confocal microscope, which was not available at the time of this thesis, but which will be interesting to study in the future.

In any case, this information allows us to state that the five genes that are most affected by mutations in patients with ACM are also expressed in hiPSC-epicardial cells, making this cell type very interesting to plan future studies on the disease.

5.3 RNA-sequencing data validation

Besides fibro-fatty replacement, one of the features reported in different studies focused on ACM is inflammation, that can precede fibrosis [21] and may be a primary or secondary event in ACM, although this is yet to be verified. Specifically, the theory supporting it as a primary event is based on data demonstrating recruitment of inflammatory cells due to desmosomal dysfunction resulting in myocardial remodeling and contributing to apoptosis and fibrosis. The theory supporting it as a secondary event, on the other hand, assumes that the involvement of inflammatory cells occurs only after necrosis, thus after damage [21]. For these reasons, in the prioritization of GO terms, I focused on inflammation but also apoptosis and immunitary response. Of note, the genes that I subsequently selected where listed in more than 1 GO term and from different pathways.

In addition to those, the confront between genes differently express in both mouse and human model, allowed the selection of of 3 genes, based only on the $\log_2$Fold change, previous knowledge and literature but not based on the role of inflammation.

Out of a total of 24 genes analyzed, 14 followed the same expression change observed in RNA-seq. Specifically: *AKAP13, ANKRD1, CD36* (day 60 only), *CX3CL1, F2RL1, GDF15, NDUFA13* (day 60 only), *OCLN, SERPINE1, TLR4, TNFRSF10A* and *TNFRSF12A* showed overexpression, while *CCND3* and *H2BC4* were underexpressed, confirming the RNA-sequencing. But only *ANKRD1, CD36, F2RL1, H2BC4, NDUFA13* and *OCLN* were analyzed at both time points.

The only gene that showed the same deregulation seen in RNA-seq and that was significant (p-value<0.05) is *OCLN,* encoding for occludin, a protein of tight junctions, at day 60. Interestingly, *OCLN* expression in the cardiomyocyte is not well documented and appears to be limited to

59

endothelial and epithelial cells [28]. It is known to interact with TJP1, which is one of the proteins found to be mutated in patients with ACM (Tab. 2), so a possible alteration in *OCLN* could have an effect on TJP1 and contribute to the disease. Finally, it is known that inflammation has a disruptive effect on TJ and causes a decrease in *OCLN* expression [28]; this information is in contrast to the qPCR and RNA-seq results. It would be interesting to understand whether this gene has a remarkable functions also in cardiomyocyte and, if so, its role in ACM pathogenesis.

qPCR performed on samples harvested at day 30, revealed *CCND3* was found downregulated, despite not significant (p value=0.069), recapitulating the same trend seen in RNA-seq. In this case, Ct values showed broad distribution, which might have impacted the statistical significance of the analyses. Increasing the sample size or a better selection of the samples could reveal whether this gene is really deregulated in ACM hiPSC-CMs.

CCND3 encodes for cyclin D3, a cyclin involved in the regulation of G1/S transition. In particular, it was found underexpress in hiPSC-CMs cells. There is not much information about the role of *CCND3* in the heart, given also the fact that the cardiomyocyte remains quiescent and has no active proliferation conditions. However, the underexpression may be of interest because overexpression of *CCND3* has been linked to cardiac hypertrophy, so it may have a role in the heart that further studies may help to better understand [29].

The 10 remaining genes plus the 2 genes that were not validated at day 30 but only at day 60 are: *AREG*, *CD36* (only at day 30), *CXCL16, ERG1, FHL2, IL17RD, JAK2, NDUFA13* (only at day 30), *TMEM65*, and *UBD* should have been overexpressed, but validation showed an opposite trend, while *CYBRD1* and *PTGES* were overexpressed while in the RNA-seq they were underexpressed. From those genes only *AREG, CD36, ERG1, IL17RD, JAK2, NDUFA13, TMEM65* and *UBD* were analyzed at both time points.

The expression was significantly different for 3 of those: *CD36* and *TMEM65*, at day 30 and *JAK2*, at day 60.

CD36, also known as fatty acid translocate, is a membrane protein that acts as receptor for a broad spectrum of molecules, including fatty acid (FA). This gene resulted overexpressed in RNA-seq analysis but the validation showed an unexpected significant underexpression at day 30. In particular, CD36 promotes the expression of PPARγ, which is already overexpressed under

ACM conditions due to Wnt turn-off, and CD36 overexpression leads to increased apoptosis, which can be reduced by knocking down CD36. Finally, CD36 is important for the transport of FA within cells [20], [30]. All these data suggest that, given the characteristics of ACM, CD36 may be overexpressed under ACM conditions, as found by RNA-seq.

TMEM65 encodes for transmembrane protein 65. This is an important protein for the proper localization of Cx43 at the level of the intercalated discs of the heart, that is known to be involved in the disease [31]. *TMEM65* was found upregulated in ACM samples by RNA sequencing, but qPCR revealed the underexpression of this gene at both time points with a significant difference only at day 30. It would be interesting to study and confirm this data as well the interaction of TMEM65 with Cx43 and the role in intercalated discs [31].

JAK2, Janus Kinase 2, is underexpressed at both time points but it is only significant at day 60. It encodes for is a non-receptor kinase that is involved in different cellular processes, in particular it was selected in this study for its role in inflammation and apoptosis [32]. Its underexpression seems to have a protective role [32] so it is not validated but also shows the opposite of what was expected from the disease.

In addition to these genes, there are other two, *AREG* and *NDUFA13,* that despite not significant (p-value<0,07) are worth of consideration.

AREG, encoding for amphiregulin, is particularly interesting because it is nearly significant at both time points, possibly influenced by the high variability of expression in the control samples. While this gene was found overexpressed qPCR revealed the opposite trend. Amphiregulin is a member of the epidermal growth factor family with different functions but, in the cardiomyocyte, it seems to be involved in Cx43 localization and phosphorylation, that is necessary for its function [33]. On the other hand it underexpression seems to reduce fibrosis and apoptosis in mice harboring Angiotensin II induced cardiac hypertophy [34]. Further studies can understand the role of *AREG* in ACM.

Finally *NDUFA13* was found underexpressed at day 30 with a p-value of 0,059, with the opposite trend of RNA-seq and day 60. The result may be influenced by the variability of the samples. Interestingly, this gene was found underegulated also in ACM murine models RNA-seq. *NDUFA13* encodes for NADH: ubiquitinon oxidoreductase subunit A13, that is part of the complex I of the respiratory chain, involved in the transport of electrons from NADH to

the respiratory chain. This could be interesting because recent studies discovered the presence of alteration in the mitocondria in ACM cells due to the alteration and the depletion of ATP [35].

It is possible that the variation in the expression of many of the genes selected in this study may be due to the use of dyshomogenous cell populations, as suggested by the broad variability of expression of the cardiomyocyte marker *TNNT2* (Fig. 24). This variability can had affect the analysis masking or altering the mean and making the data difficult to interpret. In addition, the difference between day 30 and day 60 qPCR results may be due to the progression of the disease in the cells, but also to variability in protocol efficiencies, making it difficult to compare the data.

A possible solution to this could be the use of flow cytometry to divide the cells according to TNNT2 expression and the possible expression of other markers to discriminate the presence of other cell types, testing only cardiomyocytes. Alternatively, cells could have been analysed for relevant markers prior to validation of the selected genes. Important limitations of the study are the variability of expression and the lack of a pure culture.

REFERENCE

1 Brieler J, Breeden MA, Tucker J. Cardiomyopathy: An Overview. *Am Fam Physician*. 2017;96(10):640-646.

2 Maron BJ, Towbin JA, Thiene G, 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. 2006;113(14):1807-1816. doi:10.1161/CIRCULATIONAHA.106.174287

3 Corrado D, Basso C, Judge DP. Arrhythmogenic Cardiomyopathy. Circ Res. 2017;121(7):784-802. doi:10.1161/CIRCRESAHA.117.309345

4 Marcus FI, Fontaine GH, Guiraudon G, et al. Right ventricular dysplasia: a report of 24 adult cases. Circulation. 1982;65(2):384-398. doi:10.1161/01.cir.65.2.384

5 Thiene G, Nava A, Corrado D, Rossi L, Pennelli N. Right ventricular cardiomyopathy and sudden death in young people. N Engl J Med. 1988;318(3):129-133. doi:10.1056/NEJM198801213180301

6 Thiene G, Corrado D, Basso C. Arrhythmogenic right ventricular cardiomyopathy/dysplasia. Orphanet J Rare Dis. 2007;2:45. Published 2007 Nov 14. doi:10.1186/1750-1172-2-45

7 Pilichou K, Thiene G, Bauce B, et al. Arrhythmogenic cardiomyopathy. Orphanet J Rare Dis. 2016;11:33. Published 2016 Apr 2. doi:10.1186/s13023-016-0407-1

8 Marcus FI, McKenna WJ, Sherrill D, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. Eur Heart J. 2010;31(7):806-814. doi:10.1093/eurheartj/ehq025

9 Corrado D, Perazzolo Marra M, Zorzi A, et al. Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria. *Int J Cardiol*. 2020;319:106-114. doi:10.1016/j.ijcard.2020.06.005

10 Kohela A, van Kampen SJ, Moens T, et al. Epicardial differentiation drives fibro-fatty remodeling in arrhythmogenic cardiomyopathy [published correction appears in Sci Transl Med. 2023 Feb 8;15(682):eadg7534]. Sci Transl Med. 2021;13(612):eabf2750. doi:10.1126/scitranslmed.abf2750

11 Gerull B, Heuser A, Wichter T, et al. Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy [published correction appears in Nat Genet. 2005 Jan;37(1):106]. Nat Genet. 2004;36(11):1162-1164. doi:10.1038/ng1461

12 McKoy G, Protonotarios N, Crosby A, et al. Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease). Lancet. 2000;355(9221):2119-2124. doi:10.1016/S0140-6736(00)02379-5

13 Pilichou K, Nava A, Basso C, et al. Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy. Circulation. 2006;113(9):1171-1179. doi:10.1161/CIRCULATIONAHA.105.58367

14 Syrris P, Ward D, Evans A, et al. Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2. Am J Hum Genet. 2006;79(5):978-984. doi:10.1086/509122

15 Rampazzo A, Nava A, Malacrida S, et al. Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. Am J Hum Genet. 2002;71(5):1200-1206. doi:10.1086/344208

16 Gerull B, Brodehl A. Insights Into Genetics and Pathophysiology of Arrhythmogenic Cardiomyopathy. Curr Heart Fail Rep. 2021;18(6):378-390. doi:10.1007/s11897-021-00532-z

17 Rampazzo A, Calore M, van Hengel J, van Roy F. Intercalated discs and arrhythmogenic cardiomyopathy. Circ Cardiovasc Genet. 2014;7(6):930-940. doi:10.1161/CIRCGENETICS.114.000645

18 Yeruva S, Waschke J. Structure and regulation of desmosomes in intercalated discs: Lessons from epithelia. J Anat. 2023;242(1):81-90. doi:10.1111/joa.13634

19 Harrison OJ, Brasch J, Lasso G, et al. Structural basis of adhesive binding by desmocollins and desmogleins. Proc Natl Acad Sci U S A. 2016;113(26):7160-7165. doi:10.1073/pnas.1606272113

20 Reisqs JB, Moreau A, Sleiman Y, Boutjdir M, Richard S, Chevalier P. Arrhythmogenic cardiomyopathy as a myogenic disease: highlights from cardiomyocytes derived from human induced pluripotent stem cells. *Front Physiol*. 2023;14:1191965. Published 2023 May 11. doi:10.3389/fphys.2023.1191965

21 Austin KM, Trembley MA, Chandler SF, et al. Molecular mechanisms of arrhythmogenic cardiomyopathy. *Nat Rev Cardiol*. 2019;16(9):519-537. doi:10.1038/s41569-019-0200-7

22 Gerull B, Brodehl A. Genetic Animal Models for Arrhythmogenic Cardiomyopathy. *Front Physiol*. 2020;11:624. Published 2020 Jun 24. doi:10.3389/fphys.2020.00624

23 Spandidos A, Wang X, Wang H, Dragnev S, Thurber T, Seed B. A comprehensive collection of experimentally validated primers for Polymerase Chain Reaction quantitation of murine transcript abundance. *BMC Genomics*. 2008;9:633. Published 2008 Dec 24. doi:10.1186/1471-2164-9-633

24 Bao X, Lian X, Qian T, Bhute VJ, Han T, Palecek SP. Directed differentiation and long-term maintenance of epicardial cells derived from human pluripotent stem cells under fully defined conditions. Nat Protoc. 2017;12(9):1890-1900. doi:10.1038/nprot.2017.080

25 Doyle MJ, Lohr JL, Chapman CS, Koyano-Nakagawa N, Garry MG, Garry DJ. Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes as a Model for Heart Development and Congenital Heart Disease. Stem Cell Rev Rep. 2015;11(5):710-727. doi:10.1007/s12015-015-9596-6

26 Giacomelli E, Meraviglia V, Campostrini G, et al. Human-iPSC-Derived Cardiac Stromal Cells Enhance Maturation in 3D Cardiac Microtissues and Reveal Non-cardiomyocyte Contributions to Heart Disease. *Cell Stem Cell*. 2020;26(6):862-879.e11. doi:10.1016/j.stem.2020.05.004

27 Epifantseva I, Shaw RM. Intracellular trafficking pathways of Cx43 gap junction channels. *Biochim Biophys Acta Biomembr*. 2018;1860(1):40-47. doi:10.1016/j.bbamem.2017.05.018

28 Feldman GJ, Mullin JM, Ryan MP. Occludin: structure, function and regulation. *Adv Drug Deliv Rev*. 2005;57(6):883-917. doi:10.1016/j.addr.2005.01.009

29 Li JM, Poolman RA, Brooks G. Role of G1 phase cyclins and cyclin-dependent kinases during cardiomyocyte hypertrophic growth in rats. *Am J Physiol*. 1998;275(3):H814-H822. doi:10.1152/ajpheart.1998.275.3.H814

30 Xu L, Chen W, Ma M, et al. Microarray profiling analysis identifies the mechanism of miR-200b-3p/mRNA-CD36 affecting diabetic cardiomyopathy via peroxisome proliferator activated receptor-γ signaling pathway. J Cell Biochem. 2019;120(4):5193-5206. doi:10.1002/jcb.27795

31 Sharma P, Abbasi C, Lazic S, et al. Evolutionarily conserved intercalated disc protein Tmem65 regulates cardiac conduction and connexin 43 function. *Nat Commun*. 2015;6:8391. Published 2015 Sep 25. doi:10.1038/ncomms9391

32 Jiang L, Zhang L, Yang J, et al. 1-Deoxynojirimycin attenuates septic cardiomyopathy by regulating oxidative stress, apoptosis, and inflammation via the JAK2/STAT6 signaling pathway. *Biomed Pharmacother*. 2022;155:113648. doi:10.1016/j.biopha.2022.113648

33 Sugita J, Fujiu K, Nakayama Y, et al. Cardiac macrophages prevent sudden death during heart stress. *Nat Commun*. 2021;12(1):1910. Published 2021 Mar 26. doi:10.1038/s41467-021-22178-0

34 Ji M, Liu Y, Zuo Z, Xu C, Lin L, Li Y. Downregulation of amphiregulin improves cardiac hypertrophy via attenuating oxidative stress and apoptosis. *Biol Direct*. 2022;17(1):21. Published 2022 Aug 22. doi:10.1186/s13062-022-00334-w

35 van Opbergen CJM, den Braven L, Delmar M, van Veen TAB. Mitochondrial Dysfunction as Substrate for Arrhythmogenic Cardiomyopathy: A Search for New Disease Mechanisms. *Front Physiol*. 2019;10:1496. Published 2019 Dec 10. doi:10.3389/fphys.2019.01496

ITALIAN SUMMARY

La cardiomiopatia aritmogena (ACM) è una malattia genetica rara caratterizzata da penetranza incompleta ed espressività variabile, che fa parte della famiglia delle cardiomiopatie primarie. La sua frequenza è di circa 1:1000-1:5000, con una prevalenza di 3:1 nel sesso maschile rispetto a quello femminile e, generalmente, si presenta tra i 10 e i 40 anni. Clinicamente la progressione viene divisa in 4 fasi: la fase subclinica con modificazioni strutturali del ventricolo destro (RV) ma senza manifestare sintomi, se non l'arresto cardiaco; la fase di alterazione elettrica con aritmie; la terza fase con la perdita di funzionalità del ventricolo destro e scompenso nella contrazione cardiaca; infine la perdita di funzione di entrambi i ventricoli con il coinvolgimento del ventricolo sinistro.

La diagnosi si basa su dei criteri stabiliti nel 1994 ed aggiornati nel 2010 che tengono conto di disfunzioni, alterazioni strutturali, caratterizzazione delle pareti, alterazioni di conduzione e storia famigliare dividendoli i criteri maggior e criteri minori. Per porre una diagnosi è necessario soddisfare 2 criteri maggiori oppure 4 criteri minori o 2 criteri minori e 1 maggiore. Recentemente sono stati proposti i "Criteri di Padova" che cercano di essere più comprensivi tenendo conto di diverse presentazioni della malattia. Le terapie, ad oggi disponibili, sono principalmente farmaci anti-aritmici o, nei casi più gravi, defibrillatore o trapianto di cuore.

Varianti patogene nei geni noti vengono identificate in circa 60% dei casi, in particolare, nella maggior parte di questi, le mutazioni coinvolgono i geni codificanti le proteine dei desmosomi e dell'area composita. Entrambe queste strutture sono delle giunzioni meccaniche presenti a livello dei dischi intercalari dei cardiomiociti, assieme a *gap junctions* e canali ionici che permettono la trasmissione del segnale elettrico. In particolare, i desmosomi, sono costituiti da desmogleina-2 (DSG2) e desmocollina-2 (DSC2), che interagiscono tra di loro, tra cellule diverse, permettendo l'adesione. Grazie a dei domini transmembrana sono ancorati alla cellula e all'interno interagiscono con la placoglobina (PG o JUP) e la placofilina-2 (PKP2). A loro volta, queste due proteine sono ancorate ai filamenti intermedi grazie alla desmoplachina (DSP). In questo progetto di tesi si è lavorato con cardiomiociti derivati da cellule pluripotenti indotte (hiPSC-CMs) presentanti la mutazione non senso p.Q558* nel gene codificante la proteina DSG2. Le areae composite (AC), invece, sono costituite da proteine presenti sia nei desmosomi che nelle giunzioni aderenti (AJ), inclusa DSG2.

La malattia è caratterizzata dalla perdita di cardiomiociti e la conseguente apoptosi o necrosi, con la sostituzione in tessuto fibro-adiposo, ma anche dallo sviluppo di infiammazione. Ci sono diverse teorie che cercano di spiegare la patogenesi della malattia: innanzitutto, l'alterazione dei desmosomi causerebbe una perdita di adesione e danno ai cardiomiociti, ma in aggiunta, si pensa che i desmosomi siano coinvolti anche in vie di segnalazione. L'alterazione dei desmosomi causa quindi un'alterazione delle vie di segnalazione Wnt/β-catenin/Tcf/Lef e Hippo-YAP, ma anche la via canonica e non di TGFβ.

Gli studi sulla malattia sono stati permessi, principalmente, dall'utilizzo di topi e *Danio rerio*, i quali però non presentano una forma spontanea della malattia per questo sono utilizzate tecniche di *knock-in* e organismi transgenici. In alternativa alcuni studi sono stati fatti su culture cellulari usando HL-1, H9C2 e hiPSCs.

Un meccanismo che, ad oggi, risulta essere poco chiaro, è la generazione di tessuto adiposo a livello cardiaco. Le possibili fonti individuate si focalizzano principalmente su: transdifferenziazione dei cardiomiociti in adipociti o, più probabilmente, l'accumulo di lipidi all'interno di essi e il ruolo delle cellule dell'epicardio, in quanto dalle cellule ISL1+, che sembrano essere coinvolte in questo processo, possono derivare le cellule dell'epicardio e a loro volta dall'epicardio possono generare le cellule progenitrici cardiache fibro-adipose (cFAP). In aggiunta, si è studiato che le cellule dell'epicardio, a seguito di soppressione dell'espressione dei geni desmosomiali possono andare incontro a transizione epitelio mesenchimale (EMT).

Lo scopo della tesi è di ottimizzare un protocollo di differenziamento di hiPSC in cellule dell'epicardio e di validare un data set ottenuto da RNA-sequencing in hiPSC-CM con mutazione p.Q558* in DSG2, cercando di ottenere nuove informazioni riguardo i processi molecolari coinvolti nella malattia.

Per la prima parte della tesi, seguendo il protocollo, sono state testate quattro diverse condizioni di coltura. In particolare, per ottenere cellule dell'epicardio è stato necessario modulare il *pathway* Wnt/β-catenina con CHIR, ai giorni 0 e 7, e IWP2 al giorno 2. Le diverse concentrazioni di CHIR usate (5 o 10µM, al giorno 0 e 3 o 4µM, al giorno 7) hanno permesso di ottimizzarne il dosaggio. Per studiare quale condizione portasse ad ottenere in modo più efficiente cellule hiPSC-epicardiche si è fatta la quantificazione dell'espressione di marker specifici ai giorni 1, 6, 12 e 18 confrontandoli con il giorno 0. In particolare, ad ogni *time point* si sono quantificati, tramite qPCR, i

marker di pluripotenza più marker specifici: al giorno 1 *BRA* (precursore mesodermale), al giorno 6 *NKX2.5* e *ISL1* (progenitore cardiaco) e ai giorni 12 e 18 *WT1*, *TBX18* e *K18* (cellule pro-epicardiche ed epicardiche). In aggiunta agli ultimi due *time point* si sono quantificati anche marcatori per cellule muscolari lisce (*ACTA2*), cellule endoteliali (*PECAM1*) e cardiomiociti (*TNNT2*).

Dall'analisi di questi dati è possibile osservare una diminuzione d'espressione per i marcatori di pluripotenza e l'aumento, ai corretti *time point*, dei marcatori di differenziamento, come atteso.

In conclusione, sembra che tutte e 4 le condizioni abbiano portato alla generazione di cellule dell'epicardio derivate da hiPSC, anche se, tenendo conto l'espressione dei diversi marker, 10+4μM sembra essere quella più efficiente per le hiPSCs wild type disponibili nel laboratorio.

In aggiunta, con le stesse cellule, è stato possibile verificare l'espressione dei geni desmosomiali nelle cellule epicardiache derivate da hiPSC. Inoltre, è stata anche studiata l'espressione di *GJA1* con qPCR e la localizzazione della proteina Cx43 grazie all'immunofluorescenza, confermandone non solo l'espressione ma anche la localizzazione a livello della membrana anche nelle cellule generate. Queste informazioni permettono di confermare l'espressione dei geni desmosomiali e di GJA1 anche a livello delle cellule dell'epicardio, fornendo importanti informazioni per il loro studio e per un possibile ruolo nello sviluppo della malattia.

La seconda parte della tesi, invece, si focalizza nella validazione di un dataset di RNA-sequencing, già disponibile nel laboratorio, eseguito su hiPSC-CM con mutazione p.Q558* nella DSG2 confrontate con i rispettivi controlli isogenici. Analisi preliminari di arricchimento con R, già svolte dal laboratorio, hanno permesso di ottenere una lista di geni differenzialmente espressi che appartengono a diverse categorie di *Gene Ontology* (GO), di queste ne ho selezionate 22 coinvolte in infiammazione, apoptosi e risposta immunitaria decidendo poi di analizzare 21 geni. Ulteriori 3 geni sono stati selezionati dal confronto degli stessi dati RNA-sequencing e quelli ottenuti dall'analisi di cellule cardiache di topi *knock-in* eterozigoti, portatori della stessa mutazione confrontati con rispettivi controlli *wild type*.

Dall'analisi con qPCR è risultato che per la maggior parte dei geni, analizzati in cellule da colture bloccate al giorno 30 ed in alcuni casi al giorno 60, la differenza tra le due condizioni non risultasse essere significativa, nella

maggior parte dei casi. Ma, curiosamente, dei totali 24 geni analizzati, 12 geni + 2 geni ad un solo time point (giorno 60) hanno seguito lo stesso andamento dell'RNA-sequencing mentre 10+2 geni ad un solo time point (giorno 30) avevano andamento opposto.

Tra quelli il cui andamento è stato confermato ci sono, in particolare, *OCLN* che è risultata essere significativamente sovraespresso nelle cellule con la mutazione al giorno 60. Questo dato è particolarmente interessante perché non ci sono informazioni riguardo la sua espressione nei cardiomiociti, ma questo gene è conosciuto per il coinvolgimento nelle giunzioni occludenti, che si conosce essere alterate in condizioni di infiammazione. Anche *CCND3,* ciclina D3, coinvolta nella regolazione della transizione G1/S nella divisione cellulare, in particolare il suo ruolo a livello cardiaco non è conosciuto, conferma l'andamento osservato nell'RNA-sequencing. Presenta una diminuzione dell'espressione al giorno 30 (manca analisi al giorno 60), che non risulta però essere significativa (p=0.069).

Tra i geni in cui il trend di espressione è risultato essere opposto rispetto a quello dell'RNA-seq di particolare interesse sono: *CD36*, che è risultato essere sottoespresso in modo significativamente diverso solamente al giorno 30, è coinvolto nel trasporto intracellulare di acidi grassi e nella promozione dell'espressione di PPARγ; *TMEM65* che è risultato essere sottoespresso in modo significativo al giorno 30, potrebbe essere importante perché interagisce con Cx43; *JAK2* sottoespresso in modo significativo al giorno 60 ma la sua sottoespressione è stata studiata, in letteratura, come fattore protettivo per infiammazione e apoptosi.

In aggiunta a questi sono anche interessanti: *AREG* che risulta essere sottoespresso in modo quasi significativo (p<0,07) ad entrambi i time point. Nei cardiomiociti, questo gene, risulta essere coinvolto nella localizzazione e fosforilazione di Cx43 e nei processi di fibrosi e apoptosi; *NDUFA13* invece risulta essere sottoespresso al giorno 30 con p-value di 0,059 e questa sottoespressione è stata osservata anche nell'analisi di *RNA-seq* dei cardiomiociti murini. Codifica per una proteina mitocondriale che fa parte del complesso I della catena respiratoria, che è stato trovato alterato anche in alcuni studi su modelli murini di ACM.

In conclusione, l'espressione dei geni è risultata essere altamente variabile, come anche l'espressione di *TNNT2*. Questa variabilità potrebbe avere l'effetto di mascherare possibili alterazioni esistenti. Anche la differenza di significatività o di andamento, registrata in alcuni casi nei due *time point*

potrebbe essere effetto della variabilità dei campioni oppure della progressione del danno nei modelli *in vitro*. Una soluzione potrebbe essere l'utilizzo della citometria a flusso per valutare la presenza di cellule presentanti marcatori di popolazioni cellulari contaminanti ed ottenere una coltura di hiPSC-CMs più pura.