Universidade Nova de Lisboa

Instituto de Higiene e Medicina Tropical

Disease correction and generation of isogenic induced pluripotent stem cell lines for disease modelling in patients with *DAND5* mutations

Micael de Jesus Almeida

DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS BIOMÉDICAS ESPECIALIZAÇÃO EM BIOLOGIA MOLECULAR EM MEDICINA TROPICAL E INTERNACIONAL

JANEIRO, 2019

Universidade Nova de Lisboa

Instituto de Higiene e Medicina Tropical

Disease correction and generation of isogenic induced pluripotent stem cell lines for disease modelling in patients with *DAND5* mutations

Autor: Micael de Jesus Almeida

Orientador: José António Belo

Coorientador: José Manuel Inácio

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ciências Biomédicas

ACKNOWLEDGMENTS

I would like to thank Professor José Belo for accepting me in his team and for the opportunity to develop this project, for all the support, guidance and important discussions during the last year. As well as to José Inácio for all the patience, help, clarification of doubts, availability and to teach me the experimental techniques and for the motivation given.

I would like to acknowledge all the other members of the team: Fernando Bonet, Fernando Cristo, Graça Rosas, Oriol Bover, Selin Pars and Sara Marques for having received me so well, for all availability and help.

Finally, I must express my very profound gratitude to my parents and to my sister, for providing me with unfailing support and continuous encouragement throughout my years of study. And my extended family for receive me so well in Lisbon.

RESUMO

As doenças cardiovasculares são a principal causa de morbidade e mortalidade, principalmente nos países desenvolvidos. Para melhorar as terapias já existentes para o miocárdio é necessário, em primeiro lugar, compreender claramente os mecanismos moleculares e as vias de sinalização envolvidas por trás da cardiomiogénese.

Utilizando um modelo de ratinho, foi possível observar o papel do *DAND5* em dois processos embrionários distintos, um relacionado com o estabelecimento de LR e outro relacionado com a proliferação de cardiomiócitos. Além disso, em um projeto de sequenciamento do exoma, desenhado para identificar mutações associadas à assimetria esquerda-direita (LR) em humanos identificou-se uma mutação missense (c.455G>A) no exão 2 do gene *DAND5* de dois doentes vivos independentes, apresentando características de assimetria LR e defeitos cardíacos associados. De acordo com a análise funcional *in vitro*, essa mutação leva a uma diminuição na função inibitória da proteína *DAND5*. No entanto, ainda há muito a aprender sobre a função de *DAND5* durante a diferenciação em relação aos cardiomiócitos em humanos.

Com isso em mente, iPSCs humanas (hiPSCs) derivadas de diferentes pacientes com mutação em *DAND5* foram estabelecidas. Mas, para estabelecer modelos para estudar a função deste gene e, em particular, a alteração c.455G> A, linhas isogénicas de células estaminais pluripotentes induzidas (iPSC) corrigidas em *DAND5* devem ser geradas. Este é um controlo importante para superar o ruído fenotípico causado por fundos genéticos variáveis de linhas PSC não relacionadas. Assim como, a geração de linhagens isogénicas de iPSC *DAND5* mutantes para esclarecer o efeito da completa perda de função de *DAND5* durante a diferenciação em cardiomiócitos

Aqui, usando a edição de genoma mediada por CRISPR, fomos capazes de corrigir a mutação em *DAND5* e gerar uma linha isogénica de hiPSC corrigida em *DAND5*, ou seja, é homozigótica para a base c.455G, confirmada por sequenciamento. Assim como de gerar uma isogénica hiPSC com o knockout de *DAND5*, após a utilização da edição do genoma mediada por CRISPR para inserir codões stop prematuros, confirmados por sequenciamento, tornando-se o gene *DAND5* inoperante. A sua análise mostra então que esta linha celular mantém as propriedades hiPSC: morfologia tipo célula estaminal embrionária (ESC); análise de imunofluorescência (IF) confirmou a expressão do fator de transcrição e autorrenovação OCT4 e do marcador de superfície SSEA4, marcadores característicos de células estaminais pluripotentes; diferenciação *in vitro* baseada no corpo embrionário (EB) seguida de análise IF do marcador endodérmico alfafetoproteína (AFP), do marcador mesodérmico actina alfa do músculo liso (SMA) e o marcador ectodérmico tubulina βIII (TUBB3) confirmaram a pluripotencia das hiPSC e a sua capacidade de se diferenciar em todas as três camadas germinativas.

Essas linhas de hiPSC são agora ferramentas úteis para a modelação de doenças e para entender o papel desse gene em pacientes com mutações em *DAND5*, rastreio de drogas e medicina regenerativa. Especialmente para o estudo dos mecanismos moleculares relacionados com a proliferação de cardiomiócitos.

Palavras-chave: Doenças cardiovasculares, *DAND5*, hiPSC, CRISPR-*Cas9*, Modelação de doença.

ABSTRACT

Cardiovascular diseases are the leading cause of morbidity and mortality, mainly in developed countries. To improve already existing therapies for myocardium is firstly necessary to clearly understand the molecular mechanisms and signalling pathways involved behind cardiomyogenesis.

Using a mouse model, it was possible to observe the role of *DAND5* in two distinct embryonic processes, one related with the LR establishment and the other is related with the cardiomyocyte proliferation. Besides, on an exome sequencing project designed to identify mutations associated to left-right (LR) asymmetry in humans, it was identified a missense mutation (c.455G>A) in exon 2 of the *DAND5* gene from two independent living patients presenting characteristic LR asymmetry-associated heart defects. According to functional analysis *in vitro*, this mutation leads to a decrease in the inhibitory function of the DAND5 protein. However, a lot remains to be learned about the function of *DAND5* during differentiation towards cardiomyocytes in humans.

With this in mind, human iPSCs (hiPSCs) derived from different patients with *DAND5* mutation have been established. But, to establish models to study the function of this gene and, in particular, the c.455G>A alteration, isogenic *DAND5*-corrected induced pluripotent stem cell (iPSC) lines must be generated. This is an important control to overcome the phenotypic noise caused by variable genetic backgrounds of unrelated PSC lines. Even as, the generation of isogenic full-mutant *DAND5* iPSC lines in order to clarify the effect of the complete *DAND5* loss-of-function during differentiation towards cardiomyocytes

Here, using CRISPR-mediated genome editing, we were capable to correct the defective *DAND5* mutation and generate an isogenic *DAND5*-corrected hiPSC line, that means, is homozygous for the c.455G base, confirmed by sequencing. As well as generate an isogenic *DAND5* knockout hiPSC, after using CRISPR-mediated genome editing to insert premature stop codons, confirmed by sequencing, made the *DAND5* gene inoperative. Their analysis show then this cell line maintains the hiPSC properties: embryonic stem cell (ESC)-like morphology; immunofluorescence (IF) analysis confirmed the expression of self-renewal transcription factor OCT4 and the surface marker SSEA4, characteristic markers of pluripotent ESCs; *in vitro* embryoid body (EB)-based differentiation followed by IF analysis of the endodermal marker α-feto protein (AFP), the mesodermal marker smooth muscle actin (SMA) and the ectodermal marker βIII-tubulin (TUBB3) confirmed the pluripotency of hiPSCs and their ability to differentiate into all three germ layers.

These hiPSC lines are now useful tools for disease modelling to understand the role of this gene in patients with *DAND5* mutations, drug screening and regenerative medicine. Specially to study the molecular mechanisms related to the cardiomyocyte proliferation.

Keywords: Cardiovascular diseases, *DAND5*, hiPSC, CRISPR-*Cas9*, Disease modelling.

INDEX

CHAPTER I: Introduction 1

1. Introduction 2

1.1. Early Human Development 3

1.2. *DAND5* Family 3

1.2.1. *DAND5/ Cerl2* aetiology 4

1.3. Stem cells 10

1.3.1. Induced Pluripotent Stem Cells 11

1.4. Genome editing techniques 13

1.5. Isogenic induced pluripotent stem cell lines as a disease model 17

1.6. Aims of the project 19

CHAPTER II: Materials and Methods 21

2. Materials and methods 22

2.1. Design of targeting components of CRISPR Tool 22

2.2. hiPSC Culture 24

2.2.1. hiPSC Culture Protocols: 24

2.3. hiPSC culture and transfection procedure 27

2.3.1. hiPSC Transfection Protocol: 27

2.4. Isolation of clonal cell lines 28

2.5. Detection of genomic alteration by PCR and sequencing 29

2.6. Internal controls of pluripotency and differentiation 32

2.6.1. hiPSC clones lines culture for pluripotency markers assay 32

2.6.2. EBs formation procedure for differentiation markers assay 33

2.6.3. Immunofluorescence analysis of Pluripotency and Differentiation Markers 34

CHAPTER III: Disease correction and generation of isogenic induced pluripotent stem cell lines for disease modelling in patients with DAND5 mutations 37

3.1. Design of targeting components of CRISPR Tool 38

3.2. hiPSC culture, transfection and isolation of clonal cell lines 43

3.3. Detection of genomic alteration by PCR and sequencing 44

3.4. Internal controls of pluripotency 48

3.5. Internal controls of differentiation 49

CHAPTER IV: Generation of a DAND5 KO isogenic induced pluripotent stem cell lines for disease modelling in patients with DAND5 mutations 53

4.1. Design of targeting components of CRISPR Tool 54

4.1.1. hiPSC culture, transfection and isolation of clonal cell lines 58

4.1.2. Detection of genomic alteration by PCR and sequencing 58

4.1.3. Internal controls of pluripotency 64

4.1.4. Internal controls of differentiation 65

CHAPTER V: Discussion and Conclusions 69

CHAPTER VI: References 73

LIST OF FIGURES

Figure 1.1 - DSB repair promotes gene editing. After the Cas9 (yellow) cut, the DSBs can be repaired by NHEJ or HDR pathway. A repair template could carry changed bases (blue) desired in the genomic DNA and, at the same time, can supply leverage the HDR pathway (figure adapted from Ran, F. Ann, et al., 2013). (31)……..14

Figure 1.2 – ZFN´s architecture. Targeted double-strand DNA cutting can be obtained by designing zinc fingers for specific sequences that flank the desired cleavage site. For example, 12 bp per ZFN are targeted with polypeptides containing four zinc-finger modules each (ZF-1 through ZF-4 and ZF-5 through ZF-8) (figure adapted from Carlson DF, Fahrenkrug SC, Hackett PB, 2012). (35)……...15

Figure 1.3 – TALEN´s architecture (figure adapted from Carlson DF, Fahrenkrug SC, Hackett PB, 2012). (35)……....16

Figure 1.4 – Schematic of the RNA-guided *Cas9* nuclease targeting genomic DNA by an sgRNA (blue) and a scaffold (red). The guide sequence pairs with the DNA target (blue bar on strand), directly upstream of a requisite 5′-NGG adjacent motif (PAM; pink). Cas9 mediates a DSB ~3 bp upstream of the PAM (red triangle) (figure adapted from Ran, F. Ann, et al., 2013). (31)……..17

Figure 3.1 - CRISPR design web tool print of gRNA sequences with guide 7 Cas9 cut score and suitable off-targets……...…….39

Figure 3.2 - Agarose gel results of sgRNA plasmid construction. M- molecular weight marker; 1 - Plasmid DNA from colony 1; 2 - Plasmid DNA from colony 2….……40

Figure 3.3 -DNA sequence confirming sgRNA#7.A sequence into pCAG-SpCas9-GFP-U6-gRNA. 1 - Plasmid DNA from colony 1; 2 - Plasmid DNA from Colony 2……41

Figure 3.4 - Scheme of CRISPR/*Cas9* system mode of action by HDR. The two loci of DAND5 gene exon 2 in the iUC10 cell line are represented with the gRNA 7.A and Template 7.3 homology places. The Cas9-cut at 3 bp from PAM sequence at 3-prime direction and the expected recombination region, after Cas9 cut, are signalized. The expected recombination regions, between the Template 7.3 and the iUC10 genome locus with the G>A alteration, there are before the c.455G>A alteration and after the enzymatic restriction altered base…….....42

Figure 3.5 - iUC10C cell line genotyping by PCR and enzymatic digestion. iUC10C and iUC10 PCR results present 399bp band. iUC10C digestion results present extra 217 and 182 bp bands when compared to the control, iUC10 cell line. M - molecular weight marker; NTC – PCR negative control……...…45
Figure 3.6 - DNA forward sequences confirming the c.455G>A variant correction (green arrow) in the iUC10C cell line, compared to iUC10. And the T>C base alteration to set enzymatic restriction sequence (blue arrow)……46
Figure 3.7 - DNA reverse sequences confirming the c.455G>A variant correction (green arrow) in the iUC10C cell line, compared to iUC10. And the T>C base alteration to set enzymatic restriction sequence (blue arrow)……47
Figure 3.8 - Morphology of the iUC10C line compared with iUC10. Scale bars: 100µm.……..……...48
Figure 3.9 - Immunodetection of pluripotency markers, Oct4 and SSEA4, of the iUC10C line compared with iUC10. Nuclei were stained with DAPI. Scale bars: 20µm……..…..49
Figure 3.10 - Embryoid Bodies, formed from the derived iUC10C cell line, differentiate spontaneously into the three germ layers, in suspension at left and plated at right. Scale bar: 200µm……...…50
Figure 3.11 - Immunofluorescence analyses of *in vitro* differentiation of EBs using specific antibodies against the AFP, TUBB3 and SMA for the iUC10C cell line. Nuclei were stained with DAPI. Scale bars: 20µm……..51
Figure 4.1 - Possible gRNA sequences, detached from CRISPR design web tool, with guide 1 Cas9 cut score and suitable off-targets……55
Figure 4.2 - Agarose gel results of sgRNA plasmid construction. M- molecular weight marker; 1- Plasmid DNA from colony 1; 2- Plasmid DNA from colony 2; 3- Plasmid DNA from colony 3……..56
Figure 4.3 - DNA sequence confirming sgRNA#1 sequence into pCAG-SpCas9-GFP-U6-gRNA from colony 1…….56
Figure 4.4 - Scheme of CRISPR/*Cas9* system mode of action by HDR for gene KO. The two loci of DAND5 gene exon 1 in the iUC10 cell line are represented with the gRNA 1.KO and Template 1 homology places. It is signalized the Cas9-cut at 3 bp from PAM sequence at 3-prime direction. As well as, the expected recombination region between the

Template 1 and the iUC10 genome loci, after Cas9 cut, which is before the three stop codons and after the last enzymatic restriction altered base……..…57

Figure 4.5 - hiPSC clones genotyping of the iUC10 derived cells for generation of DAND5 KO cell line by PCR and enzymatic digestion. hiPSC lines PCR positive results present 499bp band. hiPSC lines digestion positive results present 369 and 130 bp bands. M - molecular weight marker; hiPSC clones PCR results: ℕ; hiPSC clones' digestion results: ℕ*……...….59

Figure 4.6 - DNA sequences confirming the three stop codons (green arrow) insertion for the DAND5 KO in the clone 2 hiPSC line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow). And additionally, the point of mutation G>C (orange arrow)……..….60

Figure 4.7 - iUC10KO cell line genotyping by PCR and enzymatic digestion. iUC10KO and iUC10 PCR results present 499bp band. iUC10KO digestion results present the 499bp amplicons fragmentation in 369 and 130bp bands when compared to the control, iUC10 cell line…….61

Figure 4.8 - DNA forward sequences confirming the three stop codons (green arrow) insertion for the DAND5 KO in the iUC10 KO cell line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow)……...62

Figure 4.9 - DNA reverse sequences confirming the three stop codons (green arrow) insertion for the DAND5 KO in the iUC10 KO cell line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow)……...63

Figure 4.10 - Morphology of the iUC10KO cell line compared with iUC10. Scale bars: 100µm…….....64

Figure 4.11 - Immunodetection of pluripotency markers, Oct4 and SSEA4, of the iUC10KO cell line compared with iUC10. Nuclei were stained with DAPI. Scale bars: 20µm……65

Figure 4.12 - Embryoid Bodies, formed from the derived iUC10KO cell line, differentiate spontaneously into the three germ layers, in suspension at left and plated at right. Scale bar: 200µm……....66

Figure 4.13 - Immunofluorescence analyses of *in vitro* differentiation of EBs using specific antibodies against the endodermal marker AFP, ectodermal marker TUBB3 and mesodermal markers SMA, for the iUC10KO cell line. Nuclei were stained with DAPI. Scale bar: 20μm…….....67

LIST OF TABLES

Table 2.1 - Phosphorylate and anneal each pair of oligos..... 22
Table 2.2 - Conditions of set up ligation reaction between sgRNA and pCas9-GFP 23
Table 2.3 - Culture conditions for each culture vessel size..... 25
Table 2.4 - Primers nucleotide sequences for PCR to determine the cells' genotyping. 29
Table 2.5 - Conditions of set up PCR reaction..... 30
Table 2.6 - Conditions of set up enzymatic digestion..... 31
Table 2.7 - Conditions of set up controls of the enzymatic digestion reaction..... 31
Table 2.8 - Primary and secondary antibodies used for staining markers of pluripotency and germ layers´ differentiation..... 35

LIST OF ABBREVIATIONS AND ACRONYMS

AFP	α-fetoprotein
Amylase	alpha-amylase
Akt	Protein kinase B
Ankrd1	*Ankyrin Repeat Domain 1*
Anp	Atrial natriuretic peptide
APC	*Adenomatous polyposis coli*
α-Mhc	*Alpha Myosin Heavy Chain*
Baf60c	*BAF chromatin remodelling complex*
BMP	Bone morphogenetic protein
BMP-4	Bone morphogenetic protein 4
Bnp	Brain natriuretic peptide
BS	Blocking Solution
BSA	Bovine serum albumin
CHDs	Congenital heart diseases
Ck1	*Casein kinase 1*
CRISPR	Clustered regulatory interspaced short palindromic repeat
CRISPR/*Cas9*	Clustered regulatory interspaced short palindromic repeat/*Cas9*-
System	based RNA-guided endonuclease
cTnT	*Cardiac Troponin T*
CVDs	Cardiovascular diseases
cCer	*Chick Cerberus*
Cerl–2	*Cerberuslike-2*
DAND5	*DAN domain family member 5*
DAPI	4′,6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide Triphosphates
DSBs	Double-strand breaks
Dkk1	*Dickkopf 1*
DPBS	Dulbecco's phosphate-buffered saline
EBs	Embryoid bodies

ERE	Exfoliated renal epithelial
Erk	*Extracellular signal-regulated kinase*
ESCs	Embryonic stem cells
E8	Essential 8™ Flex Medium Kit
E. coli	*Escherichia coli*
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FGF	*Fibroblast growth factor*
FGF-8	*Fibroblast growth factor 8*
FISH	Fluorescence *in situ* hybridization
FLT1	*Vascular endothelial growth factor receptor 1*
Fwd	Forward
GATA4	*GATA binding protein 4*
gDNA	Genomic DNA
GFAP	Glial fibrillary acidic protein
GFP	*Green fluorescent protein*
gRNA	guide RNA
Gsk3	*Glycogen synthase kinase 3*
HDR	Homology directed repair
HESX1	*HESX Homeobox 1*
hiPSCs	Human induced pluripotent stem cells
hPSCs	Human pluripotent stem cells
IF	Immunofluorescence
iPSCs	Induced pluripotent stem cells
indels	insertion/deletion
Isl1	*Isl1 Transcription Factor, LIM Homeobox 1*
iUC10	UC-DAND5_455/10
iUC10C	UC-DAND5_455/10 with variant base c.455G>A of the DAND5 gene corrected
iUC10KO	UC-DAND5_455/10 with DAND5 gene knockout
KLF4	Kruppel-like factor 4
KO	Knockout

Lef	*Lymphoid enhancer-binding factor*
Lefty	*Left-right determination factors*
Lefty-1	*Left-right determination factor 1*
Lefty-2	*Left-right determination factor 2*
LIM1	*LIM homeobox 1*
LR	Left-right
LPM	Lateral plate mesoderm
L-LPM	Left-lateral plate mesoderm
L>R	Left to right
L<R	Right to left
Mef2	*Myocyte enhancer factor 2*
Mek	*Mitogen-activated protein kinase*
mRNA	Messenger ribonucleic acid
NHEJ	Non-homologous end joining
Nkx2	*Homeobox Protein NK-2*
Nkx2-5	*NK2 Homeobox 5*
NKX 3.2	*NK3 Homeobox 2*
Oct4	Octamer-binding transcription factor 4
Oct3/4	Octamer-binding transcription factor 3 or 4
OTX2	*Orthodenticle Homeobox 2*
PAM	Protospacer adjacent motive
Pax-6	*Paired box gene 6*
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PITX2	*Paired-like homeodomain transcription factor 2*
PI3K	*Phosphoinositide 3-kinase*
pre-crRNA	pre-CRISPR RNA
Raf	*RAF protein kinases*
REV	Reverse
RNA	Ribonucleic Acid
RNAi	Interference Ribonucleic Acid

RT-PCR	Reverse transcription polymerase chain reaction
RT	*Room temperature*
RVDs	Repeat variable disresidues
R-LPM	Right-lateral plate mesoderm
Sec. Ab.	Secondary antibody
SELI	Self-enhancement and lateral-inhibition
SF	Stem Flex Medium
sgRNA	Single guide RNA
SH	*Sonic hedgehog*
SMA	α-smooth muscle actin
Sox2	*SRY-Related High Mobility Group Box 2*
SSEA3/4	Stage-Specific Embryonic Antigen 3 or 4
SSEA4	Stage-Specific Embryonic Antigen 4
ssODNs	Single-stranded DNA oligonucleotides
TALENs	Transcription activator–like effector nucleases
Tcf	*T-cell factor*
TGF-β	*Transforming growth factor beta*
tracrRNA	Trans-activating CRISPR RNA
TUBB3	βIII-tubulin
TAE	tris-acetate-EDTA
Wnt	*Wingless/Integrated Family Members*
Wnt3	*Wingless/Integrated Family Member 3*
ZFNs	Zinc-finger nucleases

CHEMICAL SYMBOLS

H_2O	Water
ddH_2O	Double-distilled Water
CO_2	Carbon Dioxide
Cl	Chloride
K	Potassium Chloride
KH_2PO_4	Potassium dihydrogen phosphate
NaCl	Sodium Chloride
NaN_3	Sodium azide
$Na_2HPO_4.12H_2O$	Di-Sodium hydrogen phosphate dodecahydrate

UNITS

a.a.	Amino acids
bp	Base pairs
cm²	Square Centimetre
Kb	Kilobase Pair
mg	Milligram
min	Minutes
mL	Millilitre
ms	Millisecond
nt	Nucleotide
ng	Nanogram
rpm	Rotations per Minute
U	Units
V	Volt
w/v%	Weight/ Volume Percentage
xg	Times Gravity
µg	Microgram
µL	Microliter
µm	Micrometre
µM	Micromolar
%	Percentage
ºC	Degree Celsius

CHAPTER I: Introduction

1. Introduction

Cardiovascular diseases (CVDs) are the leading cause of morbidity and mortality, mainly in developed countries. It is estimated 17,7 million people died from CVDs, what is representing 31% of all global deaths and it is predictable then it is the cause of more than three-quarters of deaths worldwide in the upcoming decade. (1,2)

The ageing and growth of populations have led to an increase in the total number of cardiovascular deaths. Ischaemic heart disease, ischaemic stroke, and haemorrhagic stroke continue to cause most cardiovascular and circulatory deaths in almost all countries, which feed a huge market of medicines aimed at preventing or limiting the symptoms of the disease and restoring the quality of life. Restoring heart function in a patient diseased is of major importance once cardiac disease progression often involves cardiac stress resulting in dysfunction or loss of cardiomyocytes and the very limited regeneration capability of adult heart myocardial cells do not allow replace the damaged tissue (3–5).

Inside of the large group of CVDs is found too, the congenital heart diseases (CHDs), which is the most common cause of major congenital anomalies. The estimated birth prevalence of CHDs is of 8 per 1,000 live births is generally accepted as the best approximation. In adults, the prevalence of CHDs is estimated to be 4 per each 1,000 (6).

The aetiology of CHDs is incompletely understood because involve multiple genetic and environmental factors. Some of those diseases have been directly related to chromosomal anomalies or to defects in a single gene, which are at the origin of epistatic effects and polymorphic genes. Those effects affect different organs in time and space, during embryonic processes. Some of the genes responsible for those kinds of diseases appear to control the establishment of laterality as the *DAN domain family member 5* (*DAND5*) case. With especially relevancy between the third and eighth weeks of gestation, when most of the cardiovascular structures develop and when the integrity of the left-right (LR) axis is established (7–10).

According to the last results, *DAND5* gene seems to be the master regulator of two embryonic processes. One related with the LR establishment, when in a precise time window, *DAND5* regulates *Nodal* signalling at the node and the transmission of the LR asymmetry information to the left-lateral plate mesoderm (L-LPM), what the under-regulation of this gene leads to a wide range of asymmetry defects defined as situs

ambiguous, and to cardiac defects, where some of them are laterality-associated. The other is related with the cardiomyocyte proliferation, when the increased cardiomyocyte mitotic index is associated with prolonged/increased *transforming growth factor beta* (*TGF-β)/Nodal* signalling in the heart due to *DAND5* deregulation (9–13).

1.1.Early Human Development

Human development begins with fertilization, the process by which the male and female gametes unite to give rise to a zygote. Once the zygote has reached the two-cell stage, it undergoes a series of mitotic divisions that results in an increase in cells which leads to blastomeres formation. After the eight-cell stage, blastomeres maximize their contact with each other, compaction, forming a compact ball of cells with inner and outer layers. Approximately 3 days after fertilization, compacted blastomeres divide to form a 16-cell morula. Inner cells of the morula constitute the inner cell mass, and surrounding cells compose the outer cell mass. The inner cell mass of the early blastocyst can self-renew indefinitely, and they are pluripotent, which means they can differentiate into all three germ layers. The inner cell mass gives rise to tissues of the embryo proper, and the outer cell mass forms the trophoblast, which contributes to the placenta. (14)

During the second and third week there is successive differentiation and formation of several cells' layers. In the third week of development the gastrulation begins with the appearance of the primitive streak, which has at its cephalic end the primitive node. In the region of the node and streak, epiblast cells invaginate to form the endoderm, mesoderm and ectoderm. The embryonic period or organogenesis occurs between the third and the eight-weeks (14).

1.2.*DAND5* Family

DAND5 related proteins have been identified in other vertebrate species: mouse *Cerberuslike-2* (C*erl–2*), chick *Cerberus* (*cCer*), *Xenopus Coco*, zebrafish *Charon*. *DAND5*-related genes belong to the cysteine-knot superfamily and encode for small secreted proteins (from 185 amino acids (a.a.), *Cerl-2*, to 272 a.a., *Xcer* and C*erl-1*; 189 a.a., *DAND5*) with a signal peptide at the N-terminal and a Cysteine-Rich Domain containing 9 cysteines at the C-terminal region. *Cerl-2* is the mouse homologue of human *DAND5* and it is an important study model. *DAN/Cerberus* genes encode a member of

the *TGF-β* family signalling antagonists. As an antagonist of *TGF-β* family, these genes play a role in negative regulation of *Nodal*, *Bone morphogenetic protein (BMP)* and *wingless/integrated family members* (*Wnt)* signalling pathways, which proteins have important performances during development, organogenesis, tissue growth and differentiation. (13,15)

1.2.1. *DAND5/ Cerl-2* aetiology

DAND5/ Cerl-2 role in the left-right establishment

Establishment of the body axes, anteroposterior, dorsoventral, and LR, takes place before and during the period of gastrulation. During embryonic development, there are three key stages in LR standardization. (14)

Cells in the hypoblast (endoderm) at the cephalic margin of the disc form the anterior visceral endoderm that expresses head-forming genes, including *Orthodenticle Homeobox 2* (*OTX2*), *LIM homeobox 1* (*LIM1*), and *HESX Homeobox 1* (*HESX1*) and the secreted factor Cerberus. These genes establish the cranial end of the embryo before gastrulation. In initial somithogenesis, at the posterior notochord end a structure, called node, the loss of the bilateral symmetry in the left to right direction (L>R) begins. (13,14)

The information for the LR, generated in the node begins to propagate to the mesodermal tissue at the periphery of the embryo, the lateral plate mesoderm (LPM), which will trigger an asymmetric cascade of LR gene expression. *Nodal*, a member of the *TGF-β* family of genes, is then activated and initiates and maintains the integrity of the node and streak. Once the streak is formed, a number of genes regulate formation of dorsal and ventral mesoderm, head and tail structures. Bone morphogenetic protein 4 (*BMP-4*) in the presence of *fibroblast growth factor* (*FGF*), ventralizes mesoderm during gastrulation and so that it forms intermediate LPM. *Chordin*, *noggin*, and *follistatin* antagonize *BMP-4* activity and dorsalize mesoderm to form the notochord and somitomeres in the head region. These structures formation in more caudal regions is regulated by the *Brachyury (T)* gene. (13–17)

The last stage begins with the integration of LR signalling by the primordial organs and the translation of this information at the cellular level for the asymmetric growth of tissues leading to the proper establishment of the morphology and position of the organ in the plane of the body. LR asymmetry is regulated by a cascade of genes. First,

fibroblast growth factor 8 (*FGF-8*), secreted by cells in the node and streak, induces *Nodal* and *left-right determination factor 2* (*Lefty-2*) expression on the left side. These genes upregulate *paired-like homeodomain transcription factor 2* (*PITX2*), a transcription factor responsible for left sidedness. Simultaneously, *left-right determination factor 1* (*Lefty-1)* is expressed on the left side of the floor plate of the neural tube and may act as a barrier to prevent left-sided signals from crossing over. *Sonic hedgehog* (*SH*) may also function in this role as well as serving as a repressor for left sided gene expression on the right in humans. The *Brachyury(T)* gene, another growth factor secreted by the notochord, is also essential for expression of *Nodal, Lefty-1*, and *Lefty-2*. In the right side the gene expression cascade are not as well defined, but the transcription factor *NK3 Homeobox 2 (NKX 3.2)* is restricted to the right-lateral plate mesoderm (R-LPM) and seems be the responsible for its regulation. (13–17)

In mice, C*erl-2* transcripts can be firstly detected in a horse-shoe shaped pattern in the perinodal region of the mouse embryo, during the early somitogenesis stage. Then cilia movement located at the posterior side of node cells cause unidirectional fluid flow (nodal flow) and it promotes asymmetric gene expression around the node. Before generation of nodal flow, *Cerl-2* is symmetric expressed around the node. When it starts to flow, in perinodal crown cells, *Cerl-2* starts to be asymmetric expressed from the right to left (L<R) plate side causing the *Nodal* inhibition. As result of that, L>R asymmetric pattern of *Nodal* activity is transmitted to the LPM. (11,15,16,18)

There the asymmetric activity of Nodal derived from the node is amplified by a self-enhancement and lateral-inhibition (SELI) system, resulting in an exclusively left-sided pattern of *Nodal* expression. A SELI system consisting of *Nodal* and *left-right determination factors (Lefty*) thus generates robust LR asymmetry in LPM by relying on the node-derived asymmetric signal. In the left side, where then *Cerl-2* let to be expressed, *wingless/integrated family member 3* (*Wnt3*)-positive feedback loop is relived, accentuating the asymmetry induced by nodal flow. That way, nodal flow followed by the expression and accumulation of *Wnt3* suggest being the responsible for *Cerl-2* mRNA degradation on the left side of the node. The action of both systems ensures that the initial small bias provided by nodal flow is gradually amplified and finally converted to exclusive asymmetry. (14,16,19)

The local reduction of *Cerl-2* mRNA in the perinodal cells on the left side of the node results in an increase of active Nodal signal at one-somite stage, what leads to an increase of phosphorylated Smad2/3 levels (pSmad2/3). This asymmetric distribution of Cerl-2 protein L<R causes the asymmetric L>R pSmad2 distribution. In *Cerl-2* Knockout (KO) embryos, *Cerl2* loss of function results in a symmetric distribution of pSmad2. (9,13)

At five-somite stage, pSmad2/3 begins to disappear on the left side, probably as result of the increase of Cerl-2 protein. Therefore, after preventing any activity of *Nodal* on the right side of the node, the presence of Cerl-2 protein on the left side suppresses the activity of *Nodal* in the entire node. (9,13,16)

In a while after an L>R difference in *Nodal* activity in the node is generated, *Nodal* without *Cerl-2* inhibition starts to be expressed on the left side of the L-LPM. *Nodal* might interact with *growth differentiation factor 1* (*Gdf1*), and that sulphated glucosaminoglycans in the extracellular matrix of the intervening tissue are most likely involved in the transport of this complex. Gdf1 bind various TGF-β receptors, like Nodal, leading to recruitment and activation of SMAD family transcription factors that regulate gene expression, ultimately regulate the downstream *TGF*-β*s/Nodal* target genes. (9,13,16,19)

During early somitogenesis, *Cerl-2* loss of function in KO mice leads to bilateral expression of the left side-specific genes *Nodal*, *Lefty-2* and *PITX2*, leading to bilateral expression of this genetic cascade. *Cerl-2* role seems to be limit *Nodal* activity to the left side of the node, preventing additional activation of *Nodal*, *Lefty*-2 and *PITX2* in the R-LPM. *Nodal and Lefty-2* expression in the LPM upregulate *PITX2,* a transcription factor responsible for establishing left sidedness. Simultaneously, *Lefty-1* is expressed on the left side of the floor plate of the neural tube and may act as a barrier to prevent left-sided signals from crossing over. (14,16,18,20)

When signalling pathways undergoes a change or break in its temporal and spatial activation, errors occurs during organism development. Those errors lead to abnormal phenotypes and decrease health conditions.

During mutant *Cerl-2* KO mice lines analysis it was identified a vast array of congenital cardiac malformations associated or not with extracardiac anomalies. 1/ 3 of the new-borns died within the first 48 hours after birth and displayed left pulmonary

isomerism, thoracic situs inversus and cardiovascular malformations. ¼ of the survival ones die between weaning age and 3 months old, most of them showing heterotaxia of the abdominal organs. These KO mice present thickening of the left ventricle and of the interventricular septum due to hyperproliferation of cardiomyocytes yet, independent of LR defects. (15)

In humans, the patient's phenotype clearly resembles typical defects related to LR asymmetry observed in *Cerl-2* KO mice, namely left isomerism, ventricular septal defects, pulmonary atresia, and Tetralogy of Fallot. These defects can be associated with early LR establishment impairment. (18,21)

DAND5 role in the cardiomyocyte differentiation

Beside the well-known relation between *DAND5* family with the LR establishment, the other signalling pathway is related with the cardiomyocyte proliferation. However, the precise mechanism that controls this difference in cardiac myocyte proliferation remains poorly understood.

Cardiomyocytes in the compact myocardium have a significantly higher proliferation rate than cardiomyocytes in the trabecular myocardium during normal cardiac development. In *Cerl-2* KO neonates was detect left ventricular hypertrophy with increased left ventricular mass and severe cardiac dysfunction, as too that increased left ventricular trabecular expansion. The results showed a significantly higher proliferation index in left ventricular walls, principally in the right ventricular anterior wall and right ventricular lateral wall, maybe as result of distinct cell-lineage sources. The mitotic index increase was not related with alterations in the levels of cellular death and *Cerl-2* KO neonatal hearts genotypes did not revealed differences, suggesting that did not display hypertrophy at the cellular level. These data demonstrate then increased myocardial thickness, seen mainly in the left ventricular is caused by cardiomyocyte hyperplasia and not by hypertrophy.(9,11)

Some of the reasons for those phenotypes were related with the expression of transcription factors important for normal heart development such as *Gata-4* and *homeobox protein NK-2* (*Nkx2*), genes involved in contractility like alpha *myosin heavy chain* (*a-Mhc*) and *cardiac troponin T* (*cTnT*), which are activated by the transcription factors *Gata-4* and *myocyte enhancer factor 2* (*Mef2*) and genes such as *atrial natriuretic*

peptide (*Anp*), *brain natriuretic peptide* (*Bnp*) and *ankyrin repeat domain 1* (*Ankrd1*), which are known to be involved in the hypertrophy program and cardiac stress, and they were affected in *Cerl-2* KO mice. During midgestation, there was detected reduction of the encoding contractile genes. Those alterations of transcription factors such as *Gata-4*, *NK2 Homeobox 5* (*Nkx2-5*) and *Mef2* and their target genes may compromise the cardiomyocyte differentiation program. (9)

The regulatory role of *Cerl-2* in the cardiomyocyte cell cycle machinery leads to increased expression of *Cyclin type D1* in *Cerl2* KO embryonic hearts causing cardiac hyperplasia, observed predominantly in the left ventricle. (9)

Cerl-2 is a *TGFbs/Nodal* antagonist and binding *of TGFbs/Nodal/Activin* to its receptors leads to phosphorylation of the intracellular proteins known as receptor regulated Smads (*Smad2* and *Smad3*). These Smads interact with the co-factor Smad4 forming a transcriptional complex, which will translocate to the nucleus to regulate the downstream *TGFbs/Nodal* target genes. *Cerl-2* emerges as an essential factor in the control of the *TGFbs/ Nodal*-signalling acting as a modulator of the SWI/SNF-like *BAF chromatin remodelling complex* (*Baf60c*) that takes place during embryonic cardiogenesis being this role essential for proper heart formation. First, through mesodermal and endodermal induction to promote cardiac induction and later, to control cardiomyocyte differentiation. Absence of *Cerl-2* leads to increased phosphorylated *Smad2* signalling and up-regulation of *Baf60c* in embryonic hearts. Once, there are observed then *Baf60c* RNAi KO embryos present severe heart defects, reduced myocardial proliferation in the ventricles and altered expression of cardiac markers, causing lethality. The increase of *Baf60c* expression seems to be consistent with the ventricular hyperplasia phenotype found in the *Cerl-2* null mutants.(9)

From another perspective, *Wnt* signalling diminishes B*rachyury* (T) expression and blocks early mesoderm cell differentiation. *Wnt* antagonists such as *crescent* and *dickkopf 1* (*Dkk1*) are expressed in the anterior endoderm and exert an inductive influence on the adjacent developing heart and promote the differentiation of precardiac mesodermal progenitors into cardiac progenitors. Similarly, during *in vitro* differentiation embryonic stem cells (ESCs), repression of *Wnt* signalling promotes mesoderm progenitors to differentiate along the cardiac lineage. Furthermore, *TGF-β/BMP* regulatory pathways,

especially known to influence growth and differentiation early on in cardiac specification and differentiation, may contribute to molecular changes causing cardiac maturation and/or redundancy of *Wnt* signals (11,12).

Wnt/β-catenin signalling of regulatory elements modulation seem be the most important signalling pathway to the cardiogenesis. The *Wnt/β-catenin* signalling pathway has previously been shown to be involved in early cardiac development and growth. In the absence of *Wnt* ligands, a *complex of axin, adenomatous polyposis coli* (*APC*), *casein kinase 1* (*Ck1*) and *glycogen synthase kinase 3* (*Gsk3*), mediates phosphorylation, ubiquitylation and degradation of *β-catenin*. For the other hand, *Wnt* ligands direct through inhibition of *Gsk3*, leading to nuclear accumulation of *β-catenin*, which associates with *T-cell factor/lymphoid enhancer-binding factor* (*Tcf*/*Lef*) and activates gene transcription. Cardiomyocyte differentiation was sensitive to the timing and dose of *Wnt* pathway modulation (11,12).

Although *Gsk3* is described as part of the canonical *Wnt* cascade, it also forms the cornerstone for molecular crosstalk. Signalling pathways orchestrated by *phosphoinositide 3-kinase*/ *protein kinase B* (*PI3K/Akt)* and *RAF protein kinases*/ *mitogen-activated protein kinase*/ *extracellular signal-regulated kinase* (*Raf*/*Mek*/*Erk*) indirectly inhibit *Gsk3* at the cytoplasmic level and thereby can act synergistically with activated *Wnt* signalling. It demonstrates that *Wnt* signalling can promote early cardiomyocyte proliferation during development, but the results do not rule out the possibility that *Akt* signalling may also converge on the *β-catenin* signalling pathway (11,12).

KO studies conclusively demonstrated that *β-catenin* is required in human pluripotent stem cells (hPSCs) differentiation to mesoderm and cardiac progenitors induced by Gsk3 inhibitors. *β-catenin* at the appropriate differentiation stage enhances the generation of cardiomyocytes during monolayer-based directed differentiation induced by ligands of the *TGF-β* superfamily, as it was *DAND5*/*Cerl-2* case. Although exogenous *TGF-β* superfamily growth factors are not necessary for cardiomyocyte differentiation, *activin/Nodal* and *BMP* pathway inhibitors resulted in a dramatic decrease in cTnT positive cells. As *DAND5/Cerl-2* is an *Nodal* and *BMP* antagonist, his modulation direct mesoderm and cardiac progenitors differentiation. (11,12).

In vivo, inactivation of *β-catenin* results in a partial loss of multipotent *transcription factor, LIM Homeobox 1 (Isl1)* progenitor cells that contribute to the right ventricle and outflow tract, during early embryonic development. Conversely, constitutive expression of *β-catenin* in these multipotent progenitors' results in the expansion of the cells contributing to the right ventricle and outflow tract. During *in vivo* murine development, the differential activation of *β-catenin* signalling promotes the preferential expansion of compact versus trabecular myocardium. Furthermore, cardiac loss-of-function of *β-catenin* results in an abrogation of myocyte proliferation. Conversely, activation of *β-catenin* and canonical *Wnt* signalling drives the proliferation of early ventricular cardiomyocytes. (11)

However, remains unclear the effect of the defective *DAND5* in the regulation of *Nodal* and *Wnt* signalling during cardiomyocyte differentiation/proliferation and inferring how different *DAND5* levels reflect in cardiomyocyte proliferation and *DAND5*-associated heart diseases.

1.3.Stem cells

ESCs are derived from the inner cell mass of mammalian blastocysts, can grow indefinitely while maintaining pluripotency, and are able to differentiate into cells of all three germ layers. This is why human ESCs technology has garnered worldwide attention for its therapeutic applications in vivo. However, for ethical and practical reasons, in many primate species, including humans, the ability of ESCs to contribute to the germ line is not a testable property. (22,23)

To overcome these drawbacks, attempts have been taken to achieve individual-specific pluripotent stem cells by somatic cell nuclear transfer, treatment with an extract of pluripotent cells and cell fusion. All these approaches of nuclear reprogramming are accompanied by frailties as a technical inconvenience, partial reprogramming, and tetraploid formation, respectively, which hamper their use in medical science and technology. (24)

Other solutions come from nonhuman models use for human diseases study, they are an invaluable resource but provide a limited representation of human pathophysiology. The problem of difference in species often nullifies the novel drugs and therapy study results. For example, almost one in three drugs are not used clinically

because of side-effects on the heart. Cardiotoxic drugs are often not detected in animal models or cultured cell lines expressing selected cardiac genes because their physiology differs from that of the human heart. (22,25,26)

One way to circumvent these issues is the generation of pluripotent cells directly from the patients' own cells. In 2006, Japanese researchers induced four transcription factors (octamer-binding transcription factor *3 or 4 (Oct3/4), SRY-related high mobility group box 2 (SOX2), c-MYC*, and Kruppel-like factor 4 (*KLF4*)) in murine fibroblasts with retroviruses and they generated the first induced pluripotent stem cells (iPSCs). Human iPSCs are highly similar to ESCs: they display ESC-like morphology, express pluripotency markers, share a similar level of gene expression and epigenetic states, and possess the capacity to develop into three germ layers in vitro and in vivo. (23,27)

1.3.1. Induced Pluripotent Stem Cells

The cells that emerged from the protocol devised by Yamanaka and Jamie Thomson became known as iPSCs and the human versions are called hiPSC. iPSCs are two particular properties, self-renewal to persist their undifferentiated state and differentiation potential to give rise to specialized cell types and even regenerate whole organism. *In vitro*, they are capable to maintain the same properties for a long period with stable developmental potential to form derivatives of all three embryonic germ layers even after prolonged culture. (22,24,28,29)

Moreover, the retroviral delivery of the Yamanaka factors into somatic cells has the potential risk of random integration of the vector-encoded genes into the host genome, which can result in changes in the expression of endogenous genes and unexpected mutations. To overcome such problems, new methods appear that can generate iPSCs without its involvement using non-integrating viruses, episomal plasmid vectors, synthetic modified mRNAs, or recombinant proteins have been developed. (28,29)

hiPSCs can be generated from a large variety of donor sources. iPSC can be generated from skin biopsy, hair follicle progenitor, muscle, bone marrows/mesenchymal stem cells, lymphocytes, epithelial cells from urinal track. However, the ideal cell source would be obtained non-invasively, simple and cost-effective to obtain, and obtained universally. (22,23,28)

Genetic and epigenetic memories from respected source of cells were issued in quality of iPSC, disease modelling, and cell-based transplantation. Bi-sulfide sequencing techniques revealed the vast differences in epigenetic markers between ESCs and iPSCs. (22,28)

Depending on the lineage and type of cells, both tissue memory containing (DNA rearrangement/epigenetics) and non-containing iPSCs can be generated. From this point of view, this makes iPSCs even better choice to perform disease modelling as well as cell-based therapy. Tissue memory containing iPSCs from mature leukocytes would be beneficial for curing cancer and infectious disease, for example. (22)

The discovery of iPSCs has opened up a promising avenue as a source of gene-therapy material for autologous stem cell transplantation. Although allogeneic ESCs/iPSCs derived differentiated tissues/cells has reduced level of MHC class II molecules, the allogeneic tissue rejection from ESCs mediated cell-based therapy is one of the major challenges. (22,30)

Creation of iPSC lines from patients with single-gene disorders allows experiments on disease phenotypes *in vitro*, and an opportunity to repair gene defects *ex vivo*. Those are genetically corrected by diverse gene therapy tools, such as recombinant viral vectors, zinc-finger nuclease, transcription activator-like effector nuclease and clustered regulatory interspaced short palindromic repeat (CRISPR)/*Cas9* endonuclease systems. The clonogenic potential of iPSCs makes it easier to apply gene-therapy tools than any other cell types. In the current iPSC-based gene therapy approaches they are generating iPSCs from patient derived somatic cells with gene mutation, and subsequently introduces gene-therapy tools into the manufactured iPSCs. The resulting cells, by virtue of their immortal growth in culture, can be extensively characterized to ensure that gene repair is precise and specific, thereby reducing the safety concerns. Once as the pluripotent cells can be differentiated into relevant somatic stem cell or tissue populations, further, they could allow endow personalized medicine for future therapy. They are expected to be used in a wide variety of applications such as in cell replacement regenerative therapies, developmental biology research, disease modelling, and drug screening. (22,23,29,30)

1.4.Genome editing techniques

Genomic editing refers to techniques that are able to change one or more nucleotides in a given gene using engineered nucleases. Several genome-editing technologies have emerged in recent years, including zinc-finger nucleases (ZFNs), transcription activator–like effector nucleases (TALENs) and the clustered regulatory interspaced short palindromic repeat/*Cas9*-based RNA-guided endonuclease (CRISPR/*Cas9* System). Genome editing in mammalian cells is the process of generating genetically modified cell lines using engineered sequence-specific nucleases. Custom-engineered nucleases are introduced into cells, in which they create targeted double-strand breaks (DSBs) at a genomic site of interest to creating insertion/deletion (indels) mutations or facilitating homologous recombination. These nucleases induce DNA DSBs in a sequence specific manner but rely on the endogenous cellular DNA repair machinery to repair the breaks following DNA cleavage. (2,31–33)

The first two technologies, ZFNs and TALENs, use a strategy of tethering endonuclease catalytic domains to modular DNA-binding proteins for inducing targeted DNA DSBs at specific genomic loci. By contrast, *Cas9* is a nuclease guided by small RNAs through Watson-Crick base pairing with target DNA, representing a system that is markedly easier to design, highly specific, efficient and well-suited for high-throughput and multiplexed gene editing for a variety of cell types and organisms. (31)

Upon cleavage, the target locus typically undergoes one of two major pathways for DNA damage repair, the error-prone, non-homologous end joining (NHEJ) or, the high-fidelity, homology directed repair (HDR) pathway. In the absence of a repair template, DSBs are re-ligated through the NHEJ process, which leaves scars in the form of indel mutations. (31,32)

NHEJ can be harnessed to mediate gene knockouts, as indels occurring within a coding exon can lead to frameshift mutations and premature stop codons. Multiple DSBs can additionally be exploited to mediate larger deletions in the genome. HDR is an alternative major DNA repair pathway. Although HDR typically occurs at lower and substantially more variable frequencies than NHEJ, it can be leveraged to generate precise, defined modifications at a target locus in the presence of an exogenously introduced repair template. The repair template can either be in the form of conventional double-stranded DNA targeting constructs with homology arms flanking the insertion

sequence, or single-stranded DNA oligonucleotides (ssODNs). The latter provides an effective and simple method for making small edits in the genome. Unlike NHEJ, HDR is generally active only in dividing cells, and its efficiency can vary widely depending on the cell type and state, as well as the genomic locus and repair template. (2,31,32)

These tools can thus be used to generate large-scale deletions, gene disruption, DNA addition or single-nucleotide changes. There is an array of potential applications for genome editing in hiPSCs, for example, including the following: knocking out a gene of interest by inducing an indel mutation, knocking in a disease-associated DNA variant, generating reporter cell lines, and correcting a causal mutation in a patient-specific iPSC line. (2,33)

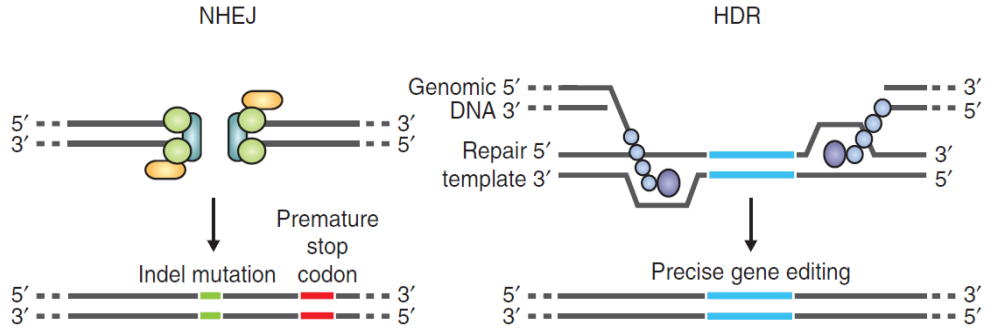

Figure 1.1 - DSB repair promotes gene editing. After the Cas9 (yellow) cut, the DSBs can be repaired by NHEJ or HDR pathway. A repair template could carry changed bases (blue) desired in the genomic DNA and, at the same time, can supply leverage the HDR pathway (figure adapted from Ran, F. Ann, et al., 2013). (31)

Zing-Finger nucleases (ZFNs)

The specificity of genomic editing depends on the DNA binding domains of the tools of choice. ZFNs are artificial enzymes which consist of a non-specific type II endonuclease *FokI* DNA cleavage domain fused with a series of zinc finger polypeptides, each recognizing specific target DNA sequences. The zinc finger domain is the most abundant DNA recognition domain in eukaryotes, and as such the diversity of known domains enables the targeting of a broad range of sequences. Each zinc-finger domain, which are approximately 30 a.a. in length and contain two b-sheets and an α-helix that coordinate to stabilize a zinc ion and is able to recognize 3 DNA base pairs (bp). Each domain binds three sequential nucleotides in the major groove of double stranded DNA through its unique α-helical motif. (32,34,35)

As multiple zinc finger domains are needed to confer ZFN specificity, the DNA binding domains of ZFNs are often composed of six to seven zinc finger domains that target sites of 18–21 nucleotides (nt). (32,34,35)

However, there are limitations to their applications. While the DNA binding regions of zinc finger domains are engineered to recognize most of the 64 possible nucleotide triplets, there are specific nucleotide triplets that do not have corresponding zinc finger domains because zinc finger domains preferentially bind guanine-rich sequences. A well designed ZFNs should only target a single site in the genome but it is frequent appear undesired mutations and chromosome aberrations caused by off-target DSB induction by ZFNs. Additionally, linking together multiple zinc finger domains causes interactions between the domains that reduce their DNA binding specificity. (32,34,35)

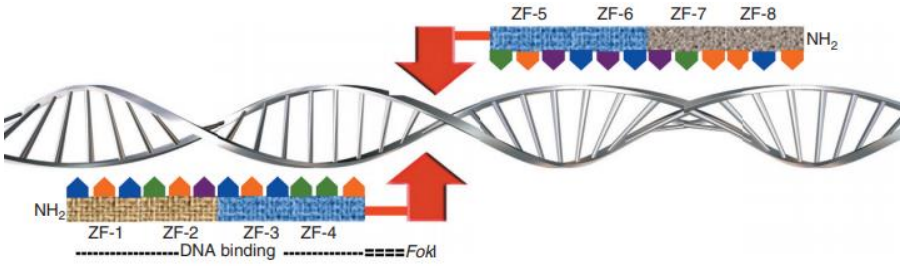

Figure 1.2 – ZFN´s architecture. Targeted double-strand DNA cutting can be obtained by designing zinc fingers for specific sequences that flank the desired cleavage site. For example, 12 bp per ZFN are targeted with polypeptides containing four zinc-finger modules each (ZF-1 through ZF-4 and ZF-5 through ZF-8) (figure adapted from Carlson DF, Fahrenkrug SC, Hackett PB, 2012). (35)

Transcription Activator-Like Effector Nucleases (TALEs)

On the other hand, TALENs represent a new genomic editing system that is considered a convenient alternative to ZFNs. The specificity of genomic editing depends on the DNA binding domains of the tools of choice. Once TALEN binding sites are chosen, custom forward and reverse TALENs are constructed that are specific to the 15 bp binding sites, each in the 5′ to 3′ direction. Each TALEN consists of a sequence-specific 15mer DNA binding domain as well as a nuclear localization signal (NLS), invariant N-term and C-term domains, and a *FokI* nuclease domain, which is able to cleave nucleotides in a nonspecific manner. This last characteristic is shared with ZFNs. (32–35)

TALE's DNA binding domain is a series of 33–35 a.a. tandem repeats. Two amino acids termed repeat variable disresidues (RVDs) at the centre of each repeat are variable and this variability controls base pair recognition. RVDs are found at positions 12 and 13

of the amino acid chain repeat sequence and confer sequence specificity to the TALENs. There are four RVDs—Asn/Asn (NN), Asn/Ile (NI), His/ Asp (HD), and Asn/Gly (NG)—that recognize the nucleotides guanine, adenine, cytosine, and thymine respectively. TALENs work in pairs to target a genomic locus, like ZFNs, but maintain a 14–18 bp spacer between the two binding domains. Also, like ZFNs, TALE repeats are linked together to recognize stretches of DNA sequences, but TALEs are easier to link together and linking multiple TALEs do not alter binding specificities, allowing TALENs to have longer domains with greater sequence specificity. (32,34,35)

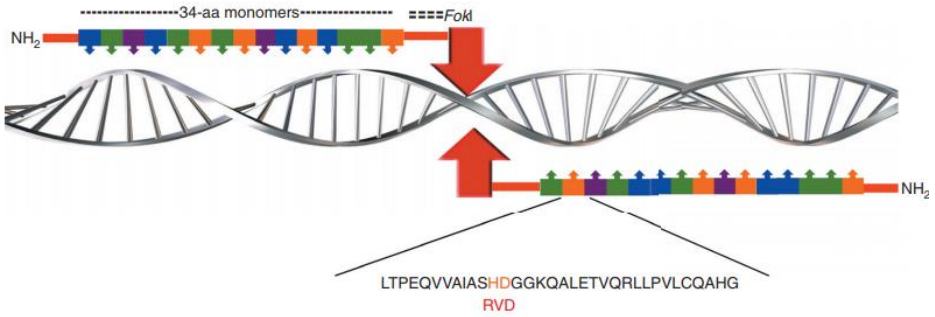

Figure 1.3 – TALEN´s architecture (figure adapted from Carlson DF, Fahrenkrug SC, Hackett PB, 2012). (35)

Clustered Regulatory Interspaced Short Palindromic Repeat/Cas9-based RNA-guided endonuclease (CRISPR/Cas9 System)

CRISPR system is an easy and efficient strategy for the generation of gene-modified cells and organisms. Three types of CRISPR systems have been identified across a wide range of bacterial and archaeal hosts, wherein each system comprises a cluster of CRISPR-associated (*Cas*) genes, noncoding RNAs and a distinctive array of repetitive elements. These nucleases arose during evolution as an adaptive immune system in bacteria designed to recognize and cleave specific sequences of foreign DNA. The CRISPR-*Cas* mechanism incorporates alien DNA or RNA into CRISPR repeat cassettes and transcribes them into pre-CRISPR RNA (pre-crRNA). pre-crRNA is being processed to crRNA and forms a complex with the Cas protein. The Cas-crRNA complex binds complementary pathogenic DNA and the endonuclease cleaves and silences the pathogenic DNA sequence-specifically. (31,32,36,37)

The Type II CRISPR system is one of the best characterized and it has a *Cas9* nuclease which give rise to the CRISPR/*Cas9* system. This system consists in the *Cas9* nuclease, the crRNA sequence encoding the target sequence, the single guide RNA

(sgRNA), and a trans-activating crRNA (tracrRNA), required for processing the crRNA into functional units. Each crRNA consists of a 20 nt sequence. In the target genome the corresponding 20nt sequence needs to be flanked by a NGG protospacer adjacent motive (PAM) to be recognised as invading DNA sequence. The Cas9 protein possesses a nuclease activity and can induce a DNA DSB in any genomic sequence guided by a guide RNA (gRNA), provided that a PAM sequence exists in the target locus. The DSB could be then repaired by NHEJ or HDR according with the desired modifications. Since Cas9 has no DNA recognition specificity except for the PAM sequence, simple multiplication of gRNAs along with a common Cas9 protein results in multiplex genome engineering. The DSB could be then repaired by NHEJ or HDR according with the desired modifications. (2,31–33,36,37)

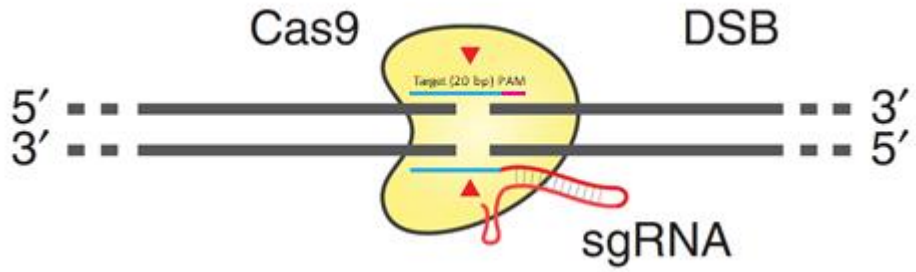

Figure 1.4 – Schematic of the RNA-guided *Cas9* nuclease targeting genomic DNA by a sgRNA (blue) and a scaffold (red). The guide sequence pairs with the DNA target (blue bar on strand), directly upstream of a requisite 5′-NGG adjacent motif (PAM; pink). Cas9 mediates a DSB ~3 bp upstream of the PAM (red triangle) (figure adapted from Ran, F. Ann, et al., 2013). (31)

1.5.Isogenic induced pluripotent stem cell lines as a disease model

iPSCs can be derived from somatic cells by gene transfer of reprogramming transcription factors. Its discovery has opened up a promising avenue as a source of gene-therapy material for autologous stem cell transplantation. The clonogenic potential of iPSCs makes it easier to apply gene-therapy tools than any other cell types. Because the robust proliferation of iPSCs is readily able to form cell clumps from single cells that can be genetically corrected by diverse gene therapy tools, such as ZFN, TALEN and CRISPR/*Cas9* systems. (2,22,28,30,32)

In the following, there is four technical challenges. First, the generation of reprogramming factor-free iPSCs to minimize or eliminate genetic alterations in the derived iPSC lines, gene-targeting strategies to generate markers for differentiation and gene correction, establishing disease-relevant phenotypes *in vitro* and establishing

disease-relevant phenotypes *in vivo*. Thus, a fundamental premise of current iPSC-based gene therapy approach is that iPSCs are able to maintain the pluripotency status, well as single-cell-dissociated are able to preserve the clonogenic potential to grow up as gene corrected iPSC clones. (2,22,28,30,32,38)

After the generation of reprogramming factor-free iPSCs, the steps necessary to accomplish genome editing in iPSCs starts with the design and construction of a gene therapy tool throw the genome target analysis. There are many design tools to screen computationally off-target sites and possible off-target effects. After transfection, clonal isolation of cell lines can be achieved by isolating single cells through either fluorescence-activated cell sorting (FACS) or serial dilutions, followed by an expansion period to establish a new clonal cell line. Afterward, the modified clones can be identified either by the SURVEYOR nuclease assay or by sequencing. (31,33,37,38)

During the cell lines establishment must be extensively tested to proof that they are indeed fully reprogrammed to a pluripotent stage, and if the gene editing does not change their properties. The first observation is their morphology, a typical small, round shape, and tightly packed ESC-like morphology with a high nucleus/cytoplasm ratio with prominent nucleoli. Follow by pluripotency and differentiation tests. (21,22,32,38,39)

For early screening of pluripotency, alkaline phosphatase activity could be measured as indicator of successful reprograming. Others characterization methods used are: stem cell markers detection by immunostaining of the keratan sulfate antigens TRA-1-60 and TRA-1-81, the glycolipid antigens *stage-specific embryonic antigen 3 or 4* (*SSEA4*) and the transcription factors *octamer-binding transcription factor* (*Oct4*), Sox2, NANOG and REX-1; Analyses of pluripotency markers expression by reverse transcription polymerase chain reaction (RT-PCR) and protein assays for detection of the set of pluripotent genes and proteins; Characterization of the epigenetic status of promoter as indicator of gene expression caused by the H3 histones and methylation status; And gene expression profiling by DNA microarray for global gene expression patterns as a relative for ESCs and parental cell expression profile. (40–42)

Differentiation of iPSC in all three germ layers are crucial as proof of pluripotency and for that could be done the following tests: Differentiation tests *in vitro* by embryoid bodies (EBs) formation (using the methods: liquid suspension culture, hanging drop culture or hydrogel culture systems) following by immunohistochemistry or

immunofluorescence (IF) assays to detect three germ layers markers (endodermal markers: α-fetoprotein (AFP), GATA binding protein 4 (GATA4) or alpha-amylase (amylase); ectodermal marker: βIII-tubulin (TUBB3), glial fibrillary acidic protein (GFAP) or paired box gene 6 (Pax-6); and mesodermal markers: α-smooth muscle actin (SMA), Brachyury or vascular endothelial growth factor receptor 1 (FLT1); Direct *in vitro* differentiation, verifying the iPSCs ability to differentiate into specific tissue and confirming by immunostaining and RT-PCR analysis, although this test does not prove the full pluripotency of iPSC; The most rigorous test is the differentiation *in vivo* by teratoma formation following by immunohistochemistry and histopathological analyses for detection of the three germ layers. Additionally telomerase activity, doubling time, cross-contamination test, karyotyping, fluorescence in situ hybridization (FISH) and genome-wide single-nucleotide polymorphism array analyses can be done. (22,38–41)

After they be extensively tested could be used in unlimited ways since regenerative therapies, developmental biology research, disease modelling, and drug screening. (22)

1.6.Aims of the project

To improve already existing therapies for myocardium is firstly necessary to clearly understand the molecular mechanisms and signalling pathways involved behind cardiomyogenesis. On an exome sequencing project designed to identify mutations associated to LR asymmetry in humans, it was identified a missense mutation in *DAND5* from two independent living patients presenting characteristic LR asymmetry-associated heart defects. According to functional analysis *in vitro*, this mutation leads to a decrease in the inhibitory function of the *DAND5* protein. Since there are no well-established models to study the function of this mutation in humans.

hiPSCs derived from different *DAND5* mutation patients have been recently established.

Using one of these cell lines and the latest tool in genome-editing, CRISPR technology, we have now:

- Corrected the single point mutation in iPSCs derived from *DAND5* patient.
- Inserted a premature triad of stop codons in that same iPS cell line.

The first cell line is an important control to overcome the phenotypic noise caused by variable genetic backgrounds of unrelated hPSC lines. Even as, the generation of

isogenic full-mutant *DAND5* iPSC lines in order to clarify the effect of the complete *DAND5* loss-of-function during differentiation towards cardiomyocytes. All to better understand the disease mechanism of patients with mutations in *DAND5*, and similar pathologies, with the ultimate purpose to find new therapeutics related with drug screening and regenerative medicine.

CHAPTER II: Materials and Methods

2. Materials and methods

The following protocols was redacted based in the articles “Genome engineering using the CRISPR-Cas9 system”, by Ran, F. Ann, et al., “Genome editing in human pluripotent stem”, by Peters, D.T., Cowan, C.A. and Musunuru, and “Multiplex genome engineering in human cells using all-in-one CRISPR/*Cas9* vector system”, by .K.; Sakuma, T., Nishikawa, A., Kume, S., Chayama, K.&Yamamoto, T. (31,33,36)

2.1. Design of targeting components of CRISPR Tool

The CRISPR/*Cas9* vector used in this protocol was pCAG-SpCas9-GFP-U6-gRNA, available through Addgene (#79144).

Specifics sgRNA were designed using the CRISPR design web tool in *crispr.mit.edu* and the genome targets chose. The selected sequence serves as template for design a pair of primers. The top oligo sequence must be preceded by a 5´ CACCG sequence and the bottom oligo must be preceded by a 5´ AAAC and followed by a 3´ C sequences. After pair of primers were designed, the oligos were ordered from Sigma-Aldrich. That primers were prepared in a sgRNA expression construct, using pCAG-SpCas9-GFP-U6-gRNA with Bbs1 cloning sites for insertion of the guide sequence. For that needed to phosphorylate and anneal each pair of oligos, first doing a mixture like described in the Table 2.1. And next, annealing in a thermocycler (MyCycler^{TM} thermal cycle, BIO-RAD) at 37ºC for 30 minutes (min.), followed by 95ºC for 5 min. and then ramp down to 25ºC at 5ºC/min.

Table 2.1 - Phosphorylate and anneal each pair of oligos

Components	Amount (µL)
oligo 1 (100µM)	1
oligo 2 (100µM)	1
5X T4 Ligation Buffer (Invitrogen)	2
H_2O milliQ	5
T4 PNK (NEB)	1
Total	10

Cloning the sgRNA into the plasmid vector for co-expression with *Cas9* was the next step. For that was prepared a mixture as described in the Table 2.2, set up the ligation

reaction and incubated at room temperature for 60 min..

Table 2.2 - Conditions of set up ligation reaction between sgRNA and pCas9-GFP

Components	Amount (µL)
BbsI (Thermo Scientific) digested pCAG-SpCas9-GFP-U6-gRNA (50ng)	2
phosphorylated and annealed oligo duplex (1:250 dilution)	1
5X Ligation Buffer (Invitrogen)	2
ddH2O	4
Ligase (Invitrogen)	1
Total	10

Finished the ligation reaction, a competent *Escherichia coli* (*E. coli*) strain was transformed. For that, 100µL of *E. coli* suspension was placed into a microtube and 2,5µL of the plasmid construct from the previous reaction were added. The mixture was incubated on ice for 40 min., heat-shocked at 42°C for 1,5 min. and returned immediately to ice for 5 minutes. At flame, 1 mL of Luria Broth medium (Sigma) was added and next incubated at 37°C for 30 min.. After recovery period, the cells were centrifuged at 3000rpm for 5 min.. At flame, 1 mL of supernatant was removed and resuspend in the remaining volume. All the contend was added to a LB Agar (Sigma) Petri plate (Thermo Scientific) containing 100 µg/mL of ampicillin (Thermo Scientific) and incubated overnight at 37°C.

For *E. coli* and plasmid vector expansion, Falcon tubes (Sarstedt) with 5 mL of Luria Broth medium and 100 µg/mL ampicillin were inoculated, utilizing a sterile pipette tip to pick a single survival *E. coli* colony coming from stage before, for each tube. The cultures were inoculated and shacked at 37 °C overnight.

For each colony, plasmid DNA extraction was done with nzyMiniprep kit (NZYtech) following the manufacturers´ instructions.

The plasmid DNA samples were then separated on a 1% (w/v%) agarose (Grisp Research Solutions) gel dissolved in tris-acetate-EDTA (TAE) 1x buffer. To intercalate between DNA chains, 3µL of GreenSafe Premium (NZYTech) was added for a final volume of 50 mL of agarose gel. Agarose gel electrophoresis was performed in TAE 1x buffer, at a constant voltage of 70 V, running in parallel with the samples a 1Kb Plus

DNA ladder (Alfagene). Then, the separated plasmid DNA products were visualized using Chemidoc (Bio-Rad) by UV-light exposure.

Upon verification of the DNA purity, 1μL of each sample was placed in the spectrophotometer (Nanodrop 2000, Thermo Scientific) to determine the DNA quantity and quality of the extracted samples from the transformed *E. coli* colonies.

To confirm the sgRNA sequence cloned into a plasmid pCAG-SpCas9-GFP-U6-gRNA vector sequence, the plasmid DNA samples from each *E. coli* colony were analysed by Sanger sequencing from the U6 promoter using the U6- Forward (Fwd) primer (U6-Fwd primer sequence: GAGGGCCTATTTCCCATGATTCC), ordered from STAB VIDA.

With the sgRNA oligos cloned into a plasmid pCAG-SpCas9-GFP-U6-gRNA vector for co-expression with *Cas9*, the first CRISPR Tool component was done. The second component is the ssODN repair template. This element was designed manually carrying the bases which were pretend be seen altered in the genome: the variant base c.455G>A corrected for the wild type base used for disease correction or rewriting the genome using three stop codons in a row used for the *DAND5* KO; and rewriting the genome for neutral enzymatic restriction sequence for a preliminary clone selection. The homology arms design was based on the distance to the *Cas9* cut side in the genome sequence (3bp from the PAM sequence) had as reference 5´ 40 /60 3´ nucleotides ratio. The ssODN repair templates was ordered from Sigma-Aldrich.

2.2.hiPSC Culture

The cell line used in the project is a hiPSC line generated from exfoliated renal epithelial (ERE) cells, designed as UC-DAND5_455/10 (iUC10), and described in the article “Generation of human iPSC line from a patient with laterality defects and associated congenital heart anomalies carrying a DAND5 missense alteration” by Cristo et all. in 2017.

2.2.1. hiPSC Culture Protocols:

The hiPSC were cultured in Essential 8™ Flex Medium Kit (E8) (Alfagene) or Stem Flex Medium (SF) (Alfagene) and incubated at 37 °C and 5% CO_2.

Table 2.3 - Culture conditions for each culture vessel size

Reagents \ Culture vessel (surface area)	Tissue Culture dish (60cm²) (Biotectomica)	6 well plate (10 cm²) (Cito+Cell)	12 well plate (4 cm²) (Cito+Cell)	24 well plate (2 cm²) (Cito+Cell)	96 well plate (2 cm²) (Cito+Cell)	Incubation Time
1. Geltrex® hESC-Qualified, Ready-To-Use (Alfagene)	8	1	0,5	0,3	0,05	60 (at 37°C) + 60 (at RT)
2. DPBS 1x (Gibco® Life Technologies)	10	3	1	0,5	0,1	6 (at 37°C)
3. TrypLE™ Select 1x (Gibco® Life Technologies)	1	0,5	0,2	0,1	0,05	3 (at 37°C)
4. Resuspension Medium (E8 or SF)	9	4,5	1	0,5	0,1	
5. Culture Medium (E8 or SF)	10	3	1	0,5	0,1	
	To passage and defrost add 1% RevitaCell Suplement 100x (Alfagene)					
6. CryoStor® cell cryopreservation media (Sigma-Aldrich)	10	1,5	0,5	-	-	
Units	Millilitres					Minutes
7. Centrifuge condition	1000rpm (200xg) for 5 min. at 25°C					

Defrost protocol (For a vial)**:**

A well of 6 well plate was coated with Geltrex® and incubated. E8 (or SF) was warmed at room temperature for an hour. The cells frosted in CryoStor® cell cryopreservation media were defrosted so fast as possible (at 37°C) and resuspended with 10 times more volume of E8 (or SF). The cell suspension was transferred to a Falcon tube and centrifuged. At the meantime, Culture Medium was prepared and 2mL were added to a coated 6 well plate, after the remaining fluid was removed. After the centrifuge, the supernatant was removed, and the cell pellet dissociated by pipetting up and down gently with 1mL of Culture Medium. The cell suspension was added to the 6 well plate and incubated at 37 °C and 5% CO_2. The medium was changed every day with Culture Medium at room temperature (without RevitaCell).

Passage protocol (For a well)**:**

The cells were ready for passage at 70-90% confluency. For that, a plate was coated with Geltrex® and incubated. E8 (or SF) was warmed at room temperature for an hour.

Next, the medium was removed, and the cells were rinsed once adding dulbecco's phosphate-buffered saline (DPBS) 1x preheated at 37°C and incubating them. Then DPBS 1x was removed and TrypLE™ Select 1x preheated at 37°C was added and incubated. Resuspension Medium (at room temperature) was added to inactivate the TrypLE™ Select 1x, the cells were resuspended, transferred to a Falcon tube and then centrifuged. In the meantime, a Falcon tube with Culture Medium was prepared and the necessary volume of this was added to the coated 6 well plate so as to make up the final volume after the cell suspension volume has been added on the next step, but first the remaining fluid was removed from the plate. After the centrifuge, the supernatant was removed, and the cells were dissociated by pipetting them up and down gently with the volume necessaire of warmed E8 (or SF) for a right dilution (usually 1:4 or 1:8). To desired cells number, the corresponding volume was added to the coated plate (where the Culture Medium was added) and it was incubated at 37 °C and 5% CO_2. The medium was changed every day with Culture Medium at room temperature (without RevitaCell).

Frost protocol:

The cells were ready to frost at 70-90% confluency. For that, the medium was removed, and the cells were rinsed once adding DPBS 1x preheated at 37°C and incubating them. Then DPBS 1x was removed and TrypLE™ Select 1x preheated at 37°C was added and incubated. Resuspension Medium (at room temperature) was added to inactivate the TrypLE™ Select 1x, the cells were resuspended, transferred to a Falcon tube and then centrifuged. After the centrifuge, the supernatant was removed, and the cells were dissociated by pipetting them up and down gently with CryoStor® cell cryopreservation media. 0,5mL of cell suspension, for 6 or 12 well plates, and 1mL of cell suspension, for Petri dishes, were put in each vial - BRAND® cryogenic tube with screw cap (Sigma) - and placed at -80°C overnight and next stored at -150°C.

Note: Versene Solution (Alfagene) could be used as TrypLE™ Select 1x substitute in the same conditions, but without needed to inactivate it with medium and centrifuge the cells. In this case, it was just removed the Versene Solution before the cells was resuspended and then they were resuspended directly with the necessaire volume of warmed E8 (or SF) for a right dilution.

2.3.hiPSC culture and transfection procedure

hiPSC were maintained in culture as described before in “hiPSC Culture” until the transfection day. For transfection was used the components obtained from “Design of targeting components of CRISPR Tool” protocol, the plasmid construct and the ssODN repair template. Those two components are part of CRISPR tool, which were used for *DAND5* variation correction and *DAND5* gene knockout in the hiPSC line.

2.3.1. hiPSC Transfection Protocol:

The transfection protocol was performed using the Neon® Transfection System (Thermo Scientific). following the manufactory’s instructions.

The cells have been plated in a 6 well plate cultured with E8 and on transfection day, they were at 70-90% confluency. SF was warmed at RT for an hour and the cells medium was changed, at least an hour before the transfection protocol starts, for SF with RevitaCell, and incubated at 37 °C and 5% CO_2. The cells were cultured as described in the “hiPSC Culture” protocols with SF.

Firstly, two wells of a 6 well plate were coated with Geltrex® and incubated.

Next, the medium was removed, and the cells were rinsed once adding DPBS 1x and incubating them. Then DPBS 1x was removed and TrypLE™ Select 1x was added and incubated. Resuspension Medium (at room temperature) was added to inactivate the TrypLE™ Select 1x, the cells were resuspended, transferred to a Falcon tube and then centrifuged. The supernatant was removed, and the pellet was resuspended in 5mL of DPBS 1x preheated at 37°C. The cells were counted and centrifuged two million equally divided in two different Falcon tubes. One was used for transfection and the other was used as transfection control.

In the meantime, a Falcon tube with 6mL of Culture Medium was prepared and 3mL were used to resuspend the transfection control cells after the supernatant have been removed. The cell suspension was then added to a coated well from the 6 well plate, after the remaining fluid have been removed. And the remaining 3 mL of Culture Medium was added to the other well, after the remaining fluid have been removed. Next, 2µg of plasmid DNA and 1µM of ssODN repair template (DNA volume never can overcome

10% of total volume), from “Design of targeting components of CRISPR Tool” step, were prepare a microtube.

The for-transfection cells supernatant was removed after last centrifuge, the pellet was dissociated by pipetting up and down gently with 100µL of Buffer R (Resuspension Buffer) and then joined and mixed with the DNA previously prepared in the microtube. Neon Tube was set up with 3mL of Electrolytic Buffer. The cell suspension mixed with DNA were introduced in the Neon Tip and place it in Neon – Transfection system and executed the electroporation in the next conditions: voltage: 1100V; width: 30ms; pulses: 1 pulse. After that, the product of electroporation was plated in the remaining well from the coated 6 well plate and incubated.

The cells were checked after 24 hours for transfection efficiency using fluorescence microscope by green fluorescent protein (GFP) gene reporter expression. If cells weren´t more than 70% of confluency, the medium was changed every day using SF (without RevitaCell) at RT.

When cells reached more than 70% of confluency, was used flow cytometry to apply the FACS for isolation of clonal cell lines.

2.4. Isolation of clonal cell lines

When the transfected cells reached more than 70% of confluency, was used flow cytometry to apply the FACS for cells with GFP expression and Propidium Iodide (Molecular Probes™) negatives.

To apply FACS technique, the cells were cultured as described in the “hiPSC Culture” protocolls with SF at 37 °C and 5% CO_2 in this step.

At least an hour before starting the FACS procedure, the medium was removed from the transfected and control cells to be sorting and was added SF with RevitaCell. A 96 well plate was coated with Geltrex® and incubated. In the same time, SF was warmed at RT for an hour. After that was prepared in a Falcon tube SF with 1% of RevitaCell and 1% of Penicillin-Streptomycin (10,000 U/mL) (Alfagene). To each coated 96 well plate well was added 110µL of it, after the incubation time and have been removed remaining fluid, and it was incubated again at 37 °C until the cell sorting time.

Next, the cultured 6 well plate medium was removed, and the cells were rinsed once adding DPBS 1x preheated at 37°C and incubating them. Then DPBS 1x was removed

and Versene Solution preheated at 37°C was added and incubated. The Versene Solution was removed, and the cells were resuspended in 5 mL of SF. Then they were transferred to a Falcon tube and centrifuged. After that, the supernatant was removed, the cells resuspended in 5 mL of DPBS 1x at RT and centrifuged again. The supernatant was removed, the cells resuspended in 500µL of DPBS 1x at RT and after filtering them through a strainer tube, the transfected cells were added to a microtube with 0,5µL of Propidium iodide and the control cells were added to other empty microtube.

The transfected single cells were then sorted by FACS Aria III (BD Biosciences) to the prepared 96 well plate, using the control cells as a procedure control. The cells were incubated at 37°C for around 2 weeks in resting, allowing them to expand, changing the medium every other day.

One week after sorting, the colonies were inspected for "clonal" appearance, and there was rejected the wells with more than a single cell.

When the hiPSC clones were more than 60% confluent, all the cells were passed to 12 well plates following the protocols described in "hiPSC Culture" with SF at 37 °C and 5% CO_2. At the next passage to another 12 well plate, 20% of the resuspended cells were plated into the replica well to keep the clonal lines and the remaining 80% resuspended cells were used for DNA isolation and genotyping.

2.5.Detection of genomic alteration by PCR and sequencing

Genomic DNA (gDNA) was extracted from hiPSCs clones using the IsolateII Genomic DNA kit (BioPortugal) following the manufacturers´ instructions. The DNA samples were evaluated relating to quantity and quality using a spectrophotometer (Nanodrop 2000, Thermo Scientific). Subsequently, amplification by PCR of the exons 1 or 2 of *DAND5* gene was carried out using the primers (ordered from Sigma-Aldrich) listed in Table 2.4 and a thermocycler (MyCycler^{TM} thermal cycle, BIO-RAD).

Table 2.4 - Primers nucleotide sequences for PCR to determine the cells' genotyping.

Gene		Primers
DAND5	Exon 1	FWD 5´- GTCGACTGCTAGTGACCTTGAG -3´
		REV 5´- TCAGGTGGAGGATACAGGACTT -3´
	Exon 2	FWD 5´- GGAAGTGGACAGGTGATTATCC -3´
		REV 5´- CACGTCTTTCTTGGTCCATCTC -3´

For *DAND5* Exon 1 PCR was followed the next protocol: Denaturation step at 98ºC for 60 seconds, followed by 35 cycles of denaturation at 98ºC for 15 seconds, annealing step at 66,5ºC for 30 seconds and extension at 72ºC for 30 seconds. The last extension step was done at 72ºC for 300 seconds. The reactions were prepared for a final volume of 50µL, composed by the reagents described in the Table 2. 5.

For *DAND5* Exon 2 PCR was followed the next protocol: Denaturation step at 98ºC for 60 seconds, followed by 35 cycles of denaturation at 98ºC for 15 seconds, annealing step at 65ºC for 30 seconds and extension at 72ºC for 20 seconds. The last extension step was done at 72ºC for 300 seconds. The reactions were prepared for a final volume of 50µL, composed by the reagents described in the Table 2.5.

Table 2.5 - Conditions of set up PCR reaction

Components	Amount (µL)
1. gDNA	V_{gDNA} = 20-50µg/µL
2. MilliQ H_2O	29,5 - V_{gDNA}
3. Primer FWD	2,5
4. Primer REV	2,5
5. dNTPs (Thermo Scientific)	5
6. 5x Phusion HF Buffer (Thermo Scientific)	10
7. Phusion DNA Polymerase (Thermo Scientific)	0,5
Final Volume	50

10µL of each PCR product were then separated on a 1% (w/v%) agarose (Grisp Research Solutions) gel dissolved in TAE 1x buffer. To intercalate between DNA chains, 3/6/9µL of GreenSafe Premium (NZYTech) was added for a final volume of 50/100/200mL of agarose gel. Agarose gel electrophoresis was performed in TAE 1x buffer, at a constant voltage of 90/110/140 V, running in parallel with the samples a 1Kb Plus DNA ladder (Alfagene). Then, the separated plasmid DNA products were visualized using Chemidoc (Bio-Rad) by UV-light exposure.

The remaining PCR products, which have been resulted in clear bands with 499bp fragments size for *DAND5* Exon 1 PCR and 399bp fragments size for *DAND5* Exon 2

PCR after electrophoresis, were purified using NZYGelpure (nzytech) following manufacturers´ instructions, and then digested for preliminary positive clone selection.

For enzymatic digestion was prepared a mixture as described in the Table 2.6 and the samples were placed in a drying oven at 37°C for 45 minutes. Each DNA sample from *DAND5* Exon 1 PCR products were digested by enzyme FastDigest *BamHI* (Thermo Scientific) and each DNA sample from *DAND5* Exon 2 PCR products were digested by enzyme FastDigest *BfoI* (Thermo Scientific).

Table 2.6 - Conditions of set up enzymatic digestion

Components	**Amount (µL)**
DNA	10
10x FastDigest Green Buffer (Thermo Scientific)	2
Enzyme	1
MilliQ H_2O	7
Total	20

Each sample were then separated on a 2% (w/v%) agarose gel dissolved in TAE 1x buffer. Each enzymatic digestion sample result was directly added to agarose gel and the remaining PCR product from each sample run in parallel after was prepared the mixture described in the Table 2.7. To intercalate between DNA chains, 3/6/9µL of GreenSafe Premium was added for a final volume of 50/100/200mL of agarose gel. Agarose gel electrophoresis was performed in TAE 1x buffer, at a constant voltage of 90/110/140 V, running in parallel with the samples a 1Kb Plus DNA ladder. Then, the separated plasmid DNA products were visualized using Chemidoc by UV-light exposure.

Table 2.7 - Conditions of set up controls of the enzymatic digestion reaction

Components	**Amount (µL)**
DNA	10
10x FastDigest Green Buffer (Thermo Scientific)	2
MilliQ H_2O	8
Total	20

After each sample were separated, the agarose gel was observed and were looked for the positive ones. That means, the samples which have been resulted in clear bands

with 369bp, 130bp and maybe 499bp (this band presence could mean heterozygosity or incomplete digestion) fragments sizes for *DAND5* Exon 1 PCR, and 399bp, 217bp and 182bp fragments sizes for *DAND5* Exon 2 PCR after electrophoresis.

After the positive clones have been selected. A new PCR have been performed from the genomic DNA, purified using NZYGelpure following manufacturers´ instructions, and sent to sequencing by Sanger sequencing from the *DAND5* Exon 1 FWD and REV primers or *DAND5* Exon 2 FWD and REV primers, ordered from STAB VIDA.

The hiPSC clones, which the pretended genome sequence have been altered, were then amplified in tissue culture dishes using "hiPSC culture" protocols and tested for internal controls of pluripotency and differentiation.

2.6. Internal controls of pluripotency and differentiation

The correctly edited hiPSC clones were tested for the pluripotency markers by IF analysis to confirm the expression Oct4 and SSEA4. Even as, after they in vitro EB-based differentiation were tested by IF analysis for the differentiation markers using specific antibodies against the endodermal marker AFP, ectodermal marker TUBB3 and mesodermal markers SMA. That way was confirmed or denied the pluripotency of iPSCs and their ability to differentiate into all three germ layers.

2.6.1. hiPSC clones lines culture for pluripotency markers assay

For the pluripotency tests were used 24 well plates and put inside the wells for each cell line in test: one coverslip for each test control and for each pluripotency marker in test. As antibody positive control, the process was repeated for the control cell line iUC10, from where the clones were originally generated.

The cells were cultured in the conditions described in the "hiPSC Culture" protocolls but changing de culture medium from E8 to SF.

The cells cultured in a 6 well plate with SF were ready for the procedure at 70-90% confluency. For that, each well of the 24 well plate prepared before with cover glass (VWR) were coated with Geltrex® and incubated. SF was warmed at room temperature for an hour.

Next, the cultured 6 well plate medium was removed, and the cells were rinsed once adding DPBS 1x preheated at 37°C and incubating them. Then DPBS 1x was removed

and Versene Solution preheated at 37°C was added and incubated. The Versene Solution was removed, the cells were resuspended in 1 mL of SF and transfer them to a microtube.

In the meantime, Culture Medium with RevitaCell was prepared in a Falcon tube. The remaining fluid was removed from the coated 24 well plate and were added 0,5mL of the Culture Medium with RevitaCell prepared before to each well. Then, 20µL of cell suspension were added to each prepared well.

The medium was changed every day with SF at room temperature (without RevitaCell). When the cells have formed small colonies (~2 days), the IF analysis started.

2.6.2. EBs formation procedure for differentiation markers assay

The undifferentiated hiPSC cells in culture were used to test the pluripotency potential and spontaneous differentiation by the hanging droplet method. Firstly, EB media I and EB medium II were prepared for hiPSC cells culture to the differentiation procedure by EB formation. EB media I was prepared dissolving 4mg/mL of Poly(vinyl alcohol) (Sigma-Aldrich) in E8 for an hour at RT and next it was filtered with 0,2µm filter in laminar flow cabinet. EB medium II was prepared mixing 50% E8 and 50% Differentiation Media (DMEM (Gibco® Life Technologies) with 20% fetal bovine serum (FBS) (Sigma), 1% glycine (Gibco® Life Technologies), 1% Non-Essential Amino Acids (Gibco® Life Technologies) and 0,02% β-mercaptoethanol (Gibco® Life Technologies) and next it was filtered with 0,2µm filter in laminar flow cabinet.

The undifferentiated hiPSC cells have been plated in a 6 well plate with SF and they were readied for EB formation procedure at 70-90% confluency. Cells in culture were resuspended using the culture conditions from the Table 2.3. SF was warmed at room temperature for an hour. Next, the medium was removed, and the cells were rinsed once adding DPBS 1x preheated at 37°C and incubating them. Then DPBS 1x was removed and TrypLE^{TM} Select 1x preheated at 37°C was added and incubated. Resuspension Medium (at room temperature) was added to inactivate the TrypLE^{TM} Select 1x, the cells were resuspended and transferred to a Falcon tube. Using a haemocytometer, the number of cells in the suspension were counted and 400000 cells corresponding volume were centrifuged in a Falcon tube. The supernatant was removed, and the pellet was resuspended in 4mL of EB medium I (at RT). Using a multichannel pipette, two 10cm² anti-adherent tissue culture dish were cover with 20µL droplets (2000 cell each) arranged

in parallel lines. It was inverted into the lid to allow growing of EBs in drop suspension (day 0). Also, 3 mL of DPBS 1x was added to the lid to provide a humidified environment. The cells were incubated at 37ºC and 5% CO_2 for 48 hours (days 1 and 2). At day 2, the DPBS was removed on the cover and the dishes turned around to the original position. Then, 10 mL of EB medium II (at RT) were added to the droplets from the both dishes and joined in the same dish. EBs were cultured in suspension and incubated at 37ºC and 5% CO_2 for 72 hours (days 3, 4 and 5). At day 5, a 24 well plate (TPP) was coated with Geltrex® and incubated. The remaining fluid was removed from the plate and to each well, 500µL of Differentiation Media and 2/3 EBs from the dish were added. EBs were incubated at 37ºC and 5% CO_2 and the medium was renewed every other day until day 19.

At day 19, immunohistochemical analysis of *in vitro* differentiation of EBs were done, using specific antibodies against the endodermal marker α-fetoprotein (AFP), ectodermal marker βIII-tubulin (TUBB3) and mesodermal markers α-smooth muscle actin (SMA). Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) (ThermoFisher Scientific).

2.6.3. Immunofluorescence analysis of Pluripotency and Differentiation Markers

The medium from undifferentiated or differentiated clone cell lines was removed and were fixed with 500µL of 4% paraformaldehyde (PFA) (Sigma) in 1x phosphate-buffered saline (PBS) for 10 minutes, at RT. The PFA was removed, the cells were permeabilized with 300µL of 1x PBS (MilliQ H_2O; 8% NaCl (VWR); 0,2% KCl (Sigma); 3,58% $Na_2HPO_4.12H_2O$ (Merck); 0,24% KH_2PO_4 (Sigma)) containing 0,2% Triton™ X-100 (Sigma) and incubate at RT for 30 minutes. The Triton solution was removed, cells were incubated at RT with 500µL of Blocking Solution (BS) (1x PBS; 7,5% glycine (Sigma); 2% bovine serum albumin (BSA) (PAA Laboratories, Inc); 0,2% NaN_3 (Merck)) for 30 minutes to prevent nonspecific antibody-binding sites. The BS was removed and 300µL of each primary antibody dilution, listed on the Table 2.8, was added to each well. To the test control cells were added 300µL of BS. The cells were then incubated at RT for 2/3 hours (or ON at 4°C).

Next, the antibodies and the BS were removed and followed by three washes with 300µL of BS, incubating for 3 minutes between them at RT. 300µL of each secondary antibody dilution, listed on the Table 2.8, were added respectively to the cells and incubated at RT for an hour (or ON at 4°C). Each secondary antibody was added too to the negative test control cells. The antibodies were removed, and the cells washed three times with 300µL of BS, incubating for 3 minutes between them at RT. Ensuing, 300mL of BS were added to the cells.

For the hiPSC clones lines cultured for the pluripotency markers assay: the coverslips were removed from the 24 well plate and put in microscope slides (Thermo Scientific) with 4µL of VECTASHIELD® Mounting Medium with DAPI (Vector Laboratories) and finally sealed. Then, they were stored at 4°C until the cell images were acquired.

For the EBs formation procedure for differentiation markers assay: the BS were removed from the 24 well plate and 300µL of 1:1000 PBS 1x diluted DAPI were added and incubated at room temperature for 20 minutes. Then the solution was removed and three washes with 300µL of PBS 1x were performed, incubating each one at room temperature for 10 minutes. The 24 well plate were stored at 4°C until the cell images were acquired.

All the cell images were acquired with Zeiss LSM 710 confocal microscope using a 40x water immersion objective. Images were taken in sequential mode and posteriorly adjusted in ZEN 2.3 SP1 FP1 program.

Table 2.8 - Primary and secondary antibodies used for staining markers of pluripotency and germ layers´ differentiation

	Antibody	Dilution	Source
Pluripotency Markers	Rabbit anti-OCT4	1:400	Abcam
	Mouse anti-SSEA4	1:200	Abcam
Differentiation Markers	Mouse anti-Human TUBB3	1:400	Sigma-Aldrich
	Mouse anti-Human SMA	1:600	Dako
	Rabbit anti-Human AFP	1:200	Dako
Secondary antibodies	Alexa Fluor 488-conjugated Donkey anti-Mouse IgG (H + L)	1:1000	Jackson ImmunoResearch Labs
	Alexa Fluor 488-conjugated Donkey anti-Rabbit IgG (H + L)	Pluripotency tests - 1:500 Differentiation tests - 1:1000	Jackson ImmunoResearch Labs

CHAPTER III: Disease correction and generation of isogenic induced pluripotent stem cell lines for disease modelling in patients with *DAND5* mutations

The hiPSC line used in this project, UC-DAND5_455/10 (iUC10) was previously generated from ERE cells isolated from a urine sample obtained from a 7-year old male child. The patient was clinically diagnosed with ventricular septal defect with overriding aorta, right ventricular hypertrophy and pulmonary atresia (a case of extreme tetralogy of Fallot phenotype). Genetically, the patient carries a heterozygous non-synonymous variant in exon 2 of *DAND5* gene (c.455G>A). Here it was used CRISPR-Cas9-based genome editing tool to correct the single point of mutation in this cell line by HDR.(21)

3.1.Design of targeting components of CRISPR Tool

Target gene: DAND5 allele sequence with G>A alteration

5´ – ATGCTCCTTG GCCAGCTATC CACTCTTCTG TGCCTGCTTA
GCGGGGCCCT GCCTACAGGC TCAGGGAGGC CTGAACCCCA GTCTCCTCGA
CCTCAGTCCT GGGCTGCAGC CAATCAGACC TGGGCTCTGG GCCCAGGGGC
CCTGCCCCCA CTGGTGCCAG CTTCTGCCCT TGGGAGCTGG AAGGCCTTCT
TGGGCCTGCA GAAAGCCAGG CAGCTGGGGA TGGGCAGGCT GCAGCGTGGG
CAAGACGAGG TGGCTGCTGT GACTCTGCCG CTGAACCCTC AGGAAGTGAT
CCAGGGGATG TGTAAGGCTG TGCCCTTCGT TCAGGTGTTC TCCCGGCCCG
GCTGCTCAGC CATACGCCTC CGAAATCATC TGTGCTTTGG TCATTGCTCC
TCTCTCTACA TCCCTGGCTC GGACCCCACC CCACTAGTCC TGTGCAACAG
CTGTATGCCT GCTCACAAGC GTTGGGCACC CGTGGTCCTG TGGTGTCTCA
CTGGCAGCTC AGCCTCCCGT CGACGGGTGA AGATATCCAC CATGCTGATC
GAGGGGTGTC ACTGCAGCCC AAAAGCATGA – 3´, *DAND5* exon 1 at grey, *DAND5* exon 2 and G>A alteration at green.

The entire CRISPR-*Cas9* system was delivered in the plasmid pCAGSpCas9- GFP-U6-gRNA. CRISPR-*Cas9* plasmid contains two main expression cassettes: an RNA Pol III promoter driving expression of a chimeric, sgRNA and a constitutive RNA Pol II promoter driving the expression of the *Cas9* gene. This protocol uses specifics sgRNA designed using the CRISPR design web tool in *crispr.mit.edu* and the genome target is the *DAND5* gene exon 2 with c.455G>A variant. The results were sequences of 20 nt followed by a PAM sequence (NGG) and the possible off-target matches throughout the

selected genome. From those suitable target sites were selected the guide with the best score, less likely off-targets and closer to the point of mutation, increasing the probability of *Cas9* cut.

all guides

scored by inverse likelihood of offtarget binding
mouse over for details ... show legend

	score	sequence	
Guide #1	91	CAGCTCAGCCTCCCGTCGAC	GGG
Guide #2	91	CAGATGATTTCGGAGGCGTA	TGG
Guide #3	88	GCAGCTCAGCCTCCCGTCGA	CGG
Guide #4	82	AGCTGTTGCACAGGACTAGT	GGG
Guide #5	76	CAGCTGTTGCACAGGACTAG	TGG
Guide #6	76	GCTGTTGCACAGGACTAGTG	GGG
Guide #7	76	GGTGCCCAACGCTTGTGAGC	AGG
Guide #8	75	GTTGCACAGGACTAGTGGGG	TGG
Guide #9	75	CAGGCATACAGCTGTTGCAC	AGG
Guide #10	74	CAAAGCACAGATGATTTCGG	AGG
Guide #11	73	GTTGGGCACCCGTGGTCCTG	TGG
Guide #12	70	TTGCACAGGACTAGTGGGGT	GGG
Guide #13	70	GTGGTCCTGTGGTGTCTCAC	TGG
Guide #14	68	GTGAGACACCACAGGACCAC	GGG
Guide #15	67	CTCCGAAATCATCTGTGCTT	TGG
Guide #16	67	GAGGCGTATGGCTGAGCAGC	CGG
Guide #17	66	AGCTGCCAGTGAGACACCAC	AGG
Guide #18	66	GTATGCCTGCTCACAAGCGT	TGG

guide #7 **quality score: 76**

guide sequence: GGTGCCCAACGCTTGTGAGC AGG
on-target locus: unknown
number of offtarget sites: 65 (0 are in genes)

top 20 genome-wide off-target sites ☐ show all exonic

sequence	score	mismatches	UCSC gene	locus
GGTGCCCAACGCTTGCGAGCAAG	17.2	1MMs [16]		chr8:-127673010
GGTTCCTTACGCTTGTGAGCTGG	1.7	3MMs [4:7:8]		chr17:+76492714
GATGGCACACGCTTGTGAGCAGG	0.9	4MMs [2:5:7:8]		chr2:-85161082
AGGGTCCAACTCTTGTGAGCCAG	0.8	4MMs [1:3:5:11]		chr14:+95062112
GAGGCCCCACCCTTGTGAGCGGG	0.8	4MMs [2:3:8:11]		chrY:+25007376
GCAGCCCAGTGCTTGTGAGCGAG	0.8	4MMs [2:3:9:10]		chr10:+94346808
GGAACCCAGAGCTTGTGAGCAGG	0.8	4MMs [3:4:9:10]		chr20:+61183002
CCTCCCCAACGCATGTGAGCAGG	0.6	4MMs [1:2:4:13]		chr1:+209742796
AGTTCCCAAGGCCTGTGAGCGGG	0.5	4MMs [1:4:10:13]		chr6:-147677677
TGTGCCCCAGGCCTGTGAGCGGG	0.5	4MMs [1:8:10:13]		chr1:+183403593
GCTGCCCACATCTTGTGAGCTAG	0.4	4MMs [2:9:10:11]		chrX:+31389876
AGTGCCCCACCTTTGTGAGCTAG	0.4	4MMs [1:8:11:12]		chr18:+77384289
GGTTCCCATTGATTGTGAGCTGG	0.4	4MMs [4:9:10:12]		chr16:+22009612
GGCGCGGAGCGCTTGTGAGCGGG	0.3	4MMs [3:6:7:9]		chr5:+5910119
AGTGCCCAGGGCATGTGAGCCGG	0.3	4MMs [1:9:10:13]		chrX:-305065
GATGCCCAGGGCGTGTGAGCTGG	0.3	4MMs [2:9:10:13]		chr22:-41354265

Figure 3.1 - CRISPR design web tool print of gRNA sequences with guide 7 Cas9 cut score and suitable off-targets.

The sequence GGTGCCCAACGCTTGTGAGC was selected and it serves as template for design a pair of primers (Figure 3.1). This sequence was score for *Cas9* cut

of 76 and 65 off-targets. However, the *Cas9* cut score is decrease, because when the off-targets were analysed, the first sequence to appear is the same sequence without the point of mutation skewing the results. All the others off-targets sequences have low probability of occurring, what means it was a good candidate.

Then the specific sgRNA (gRNA#7.A) were designed, the top oligo sequence must be preceded by a 5´ CACCG sequence - gRNA#7.A – CACCGGGTGCCCAACGCTTG TGAGC - and the bottom oligo must be preceded by a 5´ AAAC and followed by a 3´ C sequences - gRNA#7.A – AAACGCTCACAAGCGTTGGGCACCC - to allow the plasmid vector arms recognition after digestion. After pair of primers were designed, the primers were annealed and prepared in a sgRNA expression construct, using pCAG-SpCas9-GFP-U6-gRNA with Bbs1 cloning sites for guide sequence insertion, for co-expression with *Cas9*.

Analysis of plasmid DNA samples extracted from two different *E. coli* colonies in an agarose gel, did not show presence of strange genome and the expected plasmid size (Figure 3.2). The DNA concentration was determined and the insertion of the sgRNA into the plasmid verified by sequencing. The concentration results were 419,1 and 509,0 ng/µL, respectively for the *E. coli* colonies 1 and 2. The sequencing results demonstrated the sgRNA inserted in the plasmid DNA of the two samples (Figure 3.3). In the following procedures only the plasmid DNA from the colony 1 have been used.

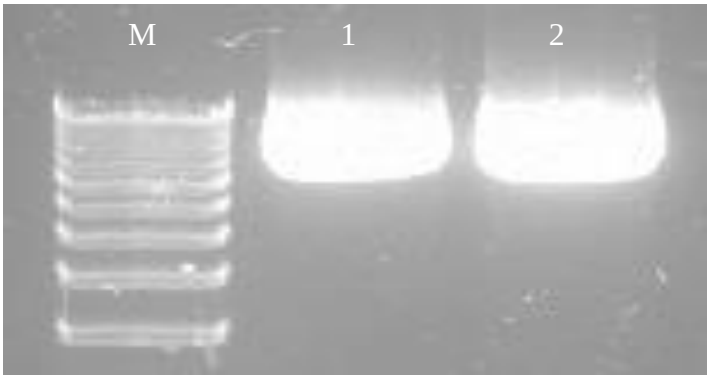

Figure 3.2 - Agarose gel results of sgRNA plasmid construction. M- molecular weight marker; 1 - Plasmid DNA from colony 1; 2 - Plasmid DNA from colony 2.

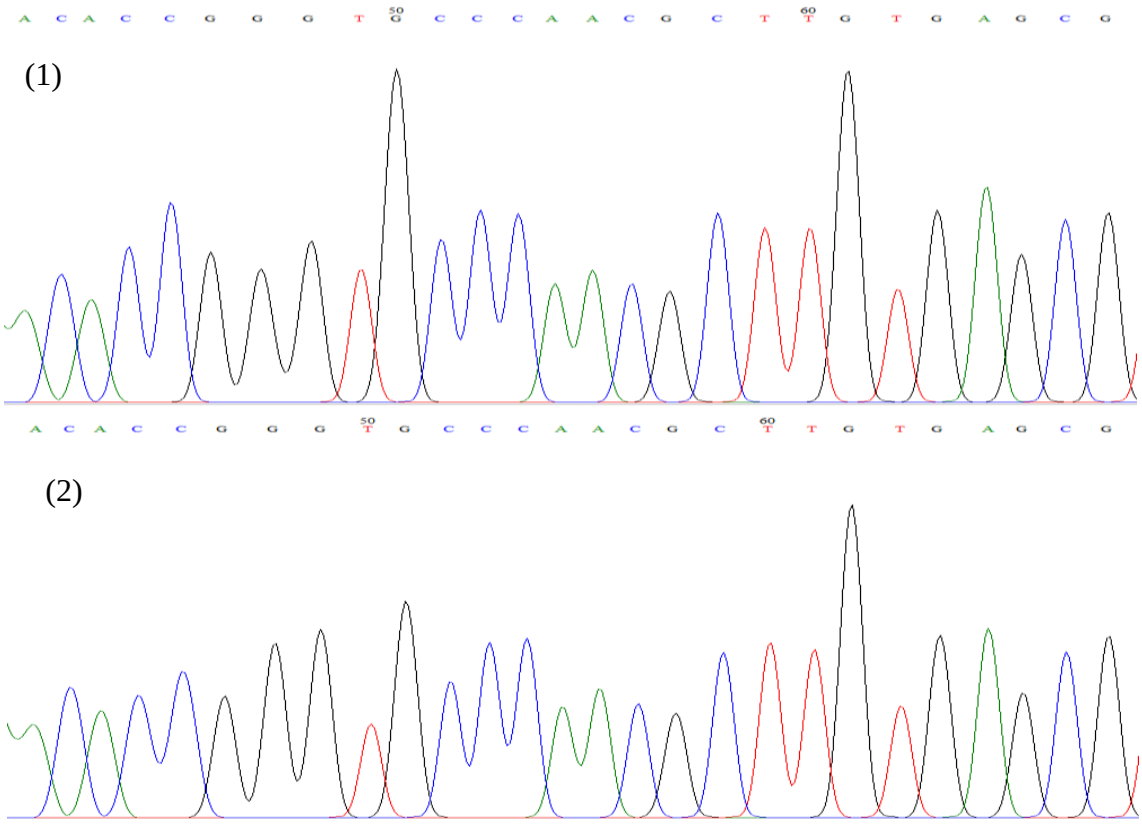

Figure 3.3 -DNA sequence confirming sgRNA#7.A sequence into pCAG-SpCas9-GFP-U6-gRNA. 1 - Plasmid DNA from colony 1; 2 - Plasmid DNA from Colony 2

ssODN repair template consists in a single stranded DNA sequence which carries the alterations which would like to see integrated in the genome. This component is a second part of the CRISPR system. Co-delivery of a homologous template during genome editing significantly increases homology repair mechanisms efficiencies, promoting the homology directed repair. This element was designed manually carrying the bases which were pretend be seen altered in the genome. In this case, it was intended rewriting the c.455G>A variant, correcting this point of mutation and generate a *DAND5* corrected hiPSC line. In addition, it was inserted a neutral enzymatic restriction sequence, avoiding additional mutations in the protein but allowing a preliminary clone selection. ssODN repair template homology arms design was based on the distance to the Cas9 cut side in the genome sequence (three pair bases from the PAM sequence) and with reference 5´ 40 /60 3´ nucleotides ratio. The FastDigest BfoI restriction enzyme was used and recognises 5´-RGCGCY-3´sequence. That way, the ssODN repair template designed was: Template 7.3 – ACCCCACCCCACTAGTCCTGTGCAACAGCTGTATGCCTGCTCGCAAGCG CTGGGCACCCGTGGTCCTGTGGTGTCTCACTGGCAGCTCAGCCTCCCGTCGA CGGGTGAAG.

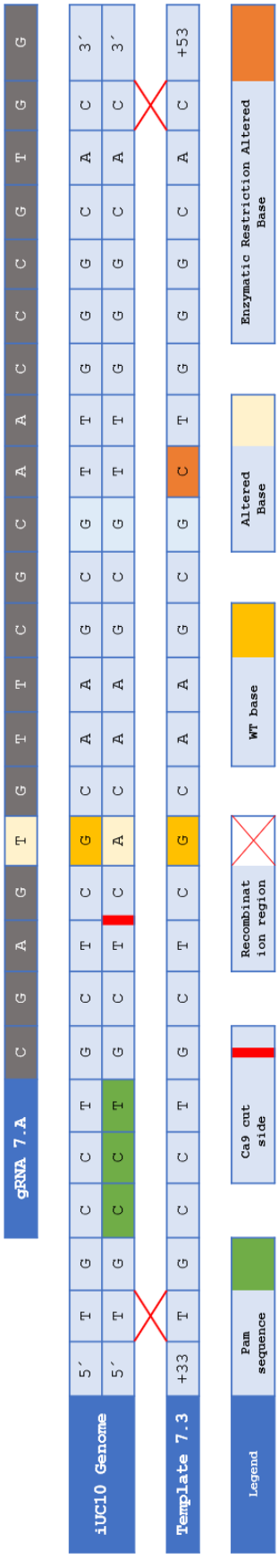

Figure 3.4 - Scheme of CRISPR/*Cas9* system mode of action by HDR. The two loci of *DAND5* gene exon 2 in the iUC10 cell line are represented with the gRNA 7.A and Template 7.3 homology places. The Cas9-cut at 3 bp from PAM sequence at 3-prime direction and the expected recombination region, after Cas9 cut, are signalized. The expected recombination regions, between the Template 7.3 and the iUC10 genome locus with the G>A alteration, there are before the c.455G>A alteration and after the enzymatic restriction altered base.

In this template, the PAM sequence was not modified, once its alteration causes an amino acid change, what make increase the probability of Cas9 re-cutting events. These events could be a problem when do not analysed carefully, namely by sequencing. That way, it was expected using CRISPR system an only Cas9 cut at three prime from the PAM sequence and the recombination by HDR between the sequence in the ssODN repair template and the genomic *DAND5* gene sequence. This way c.455G>A variant is altered, generating a homozygotic cell line for the wild type base, as well as the enzymatic restriction altered base, which generate a heterozygotic cell line for this base (Figure 3.4).

3.2.hiPSC culture, transfection and isolation of clonal cell lines

The cell line used in the project is a hiPSC line, cultured with supplemented E8 or SF, which are capable of maintain hiPSCs properties including pluripotency, karyotype, and trilineage differentiation potential. The cells should be replaced at 70-90% confluency to prevent its spontaneous differentiation. The stem cell differentiation is a preoccupation during its culture and could be detected by morphological analysis. hiPSC display a typical small, round shape, and tightly packed ESC-like morphology with a high nucleus/cytoplasm ratio with prominent nucleoli.

Before the transfection, the cells medium should be refreshed allow to the cell a better and small period of recuperation. This new medium, SF, is already supplemented with Revitacell, it has Rho-associated protein kinase (ROCK) inhibitor which coupled with antioxidants and free radical scavengers, helping to minimize the impact of stress from single-cell passaging, enabling maximal cell viability and more efficient recovery of cells. The cells in culture were transfected with 2μg of gRNA#7.A expression construct and 1μM of the designed ssODN repair template, Template 7.3. After transfection the cells need to recover of stress caused by electroporation, mostly because of membrane damage, and restore its normal metabolic activity. For this reason, the cells were cultured again in SF with Revitacell.

Once the cells establish their normal function, they were checked for transfection efficiency using fluorescence microscope by GFP gene reporter expression. The successful transfected cells have green fluorescence, that means, the CRISPR tool have been introduced successfully inside the cells by transfection. The GFP protein is then detected by flow cytometer and FACS was applied for isolation of clonal cell lines. The

cells with GFP expression and propidium iodide negatives was selected. The plasmid pCAGSpCas9-GFP-U6-gRNA vector expresses a reporter gene GFP in parallel with the sgRNA cloned co-expressed with *Cas9* and, propidium iodide can not cross the membrane of live cells, making it useful to differentiate healthy from necrotic or apoptotic cells. In flow cytometry, propidium iodide is used as a DNA stain to evaluate cell viability and that way isolate the clones with the best chances of survive to the procedure inside of the cohort of GFP expression cells.

Before the FACS procedure the transfected cells medium was refreshed, being cultured in SF with Revitacell, to prepare the cells to the procedure. And after the procedure, the isolated clones were cultured in SF with Revitacell and Penicillin-Streptomycin to prevent contaminations arising from the procedure.

The isolated clones expanded after sorting, the colonies were inspected for "clonal" appearance, and the wells with more than a single colony was rejected to prevent cell isolation with distinct genotype, as well as the ones whose did not display ESC-like morphology. After the expansion the clones were used for DNA isolation and genotyping.

3.3.Detection of genomic alteration by PCR and sequencing

hiPSCs clones and iUC10 gDNA was extracted and evaluated its quantity and quality to subsequently, amplification by PCR of the exon 2 of *DAND5* gene. The iUC10 serve as reactions control. The amplicons from each PCR sample of each clone were separated, the agarose gel was observed and were looked for the samples which have been resulted in a clear band with 399bp. However, this selection did not detect genomic alteration in amplicons with the same size.

The ssODN repair template has an enzymatic restriction sequence and when the clone's genome is edited by HDR, that sequence is incorporated. That sequence is recognised by the specific enzyme, which are used to select the clones whose have recombined with the ssODN repair template and which also has the wild type base. In some occasions, the enzymatic restriction could not mean the c.455G>A base correction, as result of recombination between the wild type base and the enzymatic restriction sequence place. This problem is solved after sequencing.

The clones´ PCR samples were digested and separated in parallel with the no digested portion in agarose gel. The samples which have been resulted in clear bands with

399, 217 and 182 bp fragments sizes for *DAND5* exon 2 PCR after electrophoresis are positives. After those steps only 1 in 44 clones analysed was positive (Figure 3.5).

This clone was analysed by Sanger sequencing and the results confirmed that the c.455G>A variant base have been corrected. This hiPSC line, designed iUC10C, is homozygotic for the wild type base and heterozygotic for the enzymatic restriction alternated base, as intended (Figure 3.6 and 3.7).

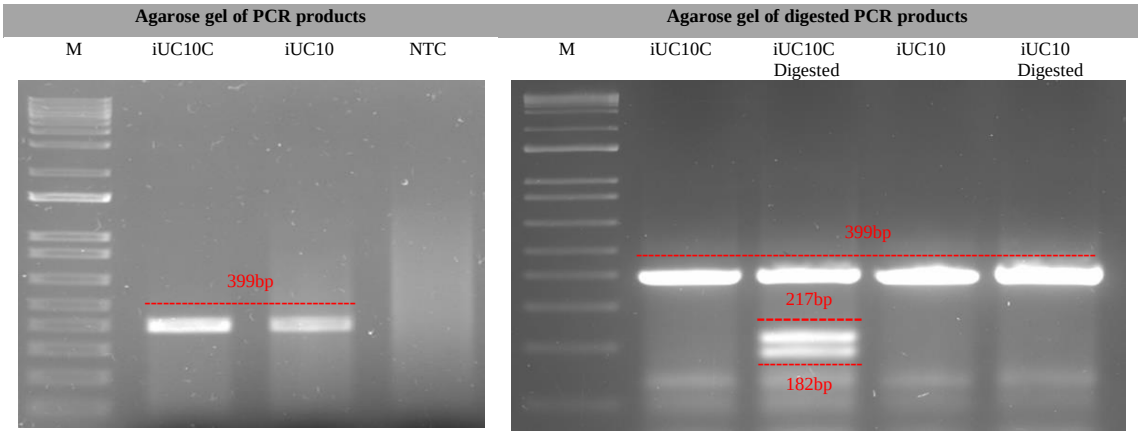

Figure 3.5 - iUC10C cell line genotyping by PCR and enzymatic digestion. iUC10C and iUC10 PCR results present 399bp band. iUC10C digestion results present extra 217 and 182 bp bands when compared to the control, iUC10 cell line. M - molecular weight marker; NTC – PCR negative control.

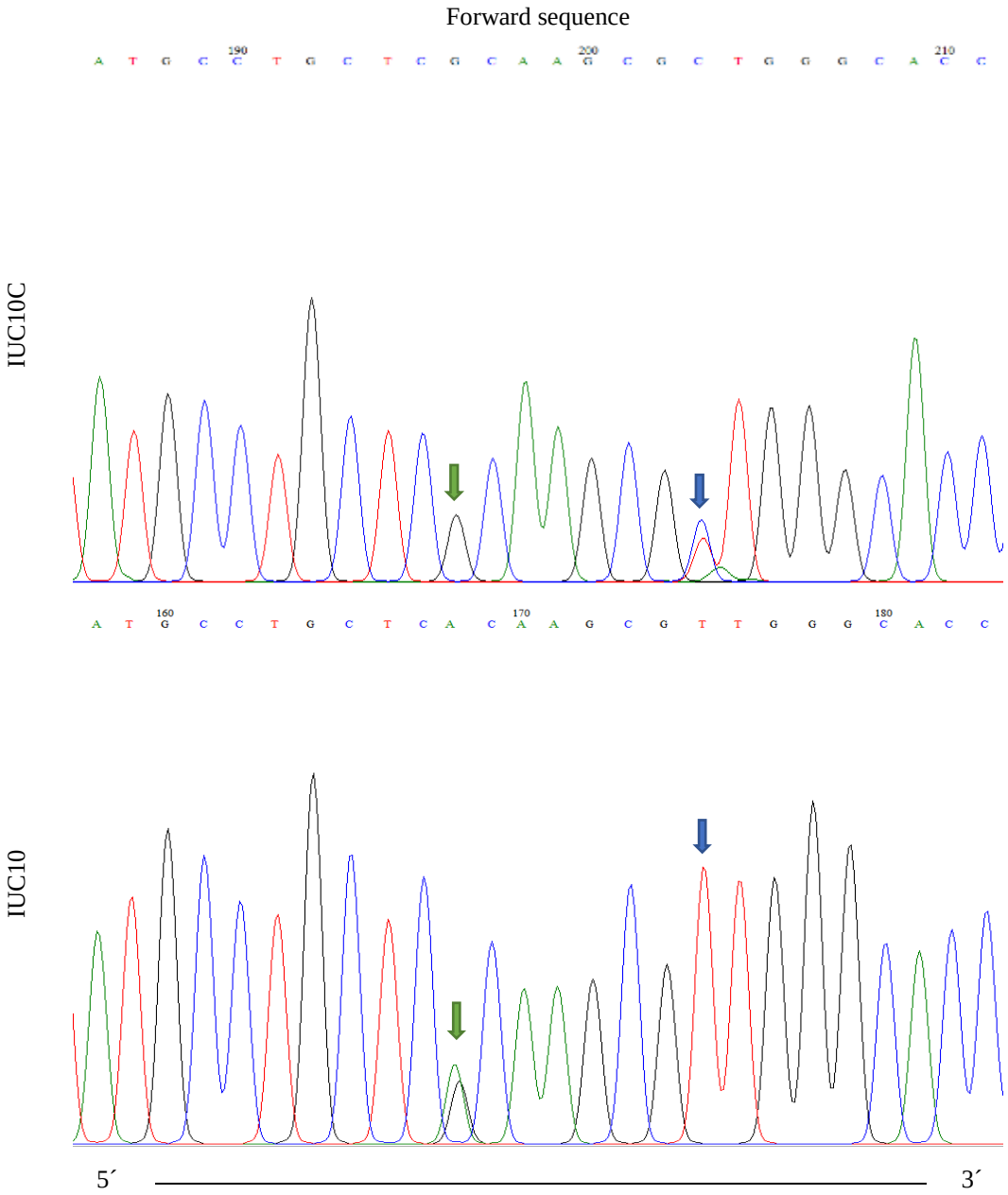

Figure 3.6 - DNA forward sequences confirming the c.455G>A variant correction (green arrow) in the iUC10C cell line, compared to iUC10. And the T>C base alteration to set enzymatic restriction sequence (blue arrow).

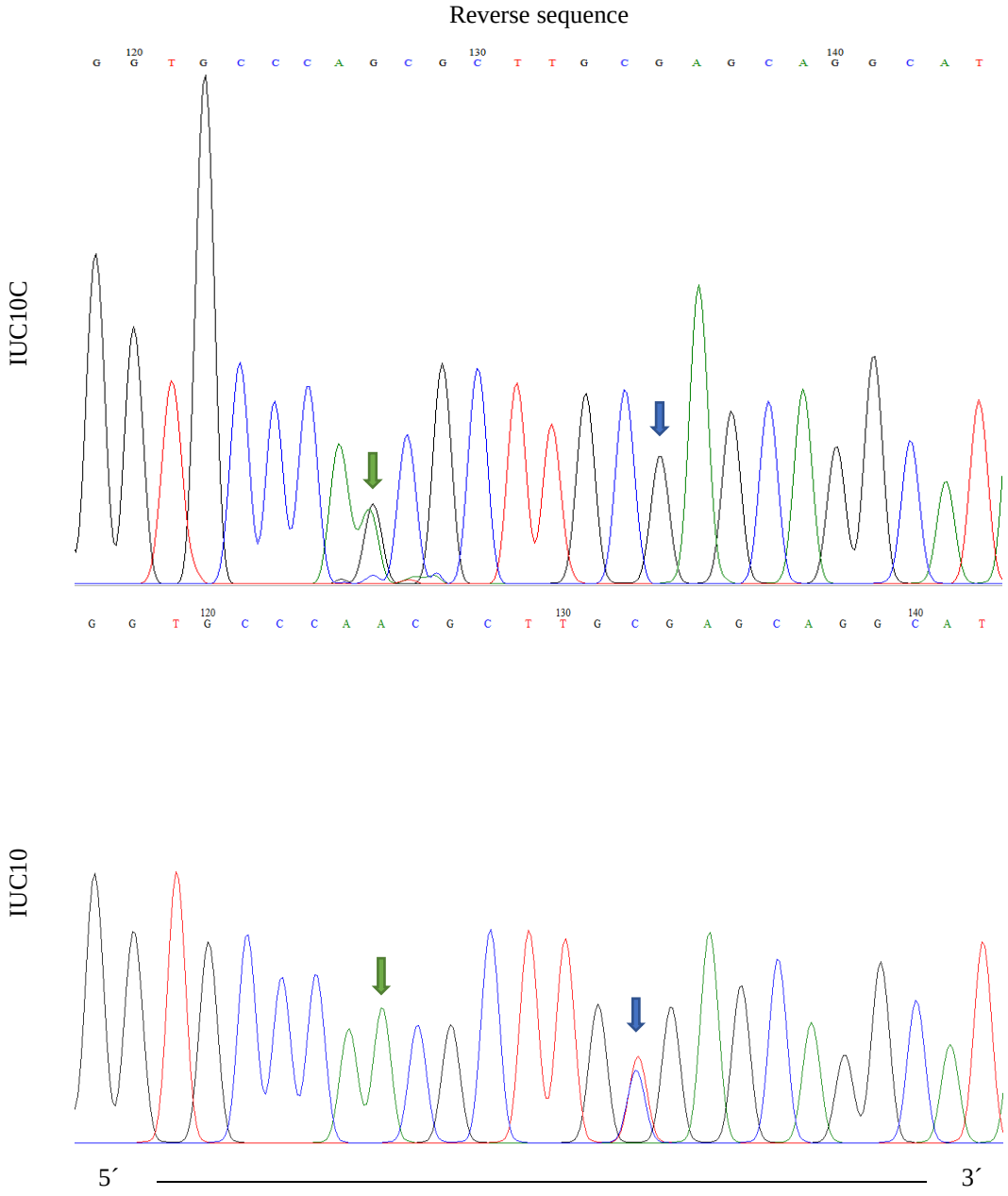

Figure 3.7 - DNA reverse sequences confirming the c.455G>A variant correction (green arrow) in the iUC10C cell line, compared to iUC10. And the T>C base alteration to set enzymatic restriction sequence (blue arrow).

The iUC10C, which the pretended genome sequence has been altered, were then extended and tested for internal controls of pluripotency and differentiation.

3.4. Internal controls of pluripotency

After expansion, iUC10C cell line, continued to display a typical small, round shape, and tightly packed ESC-like morphology with a high nucleus/cytoplasm ratio with prominent nucleoli (Figure 3.8).

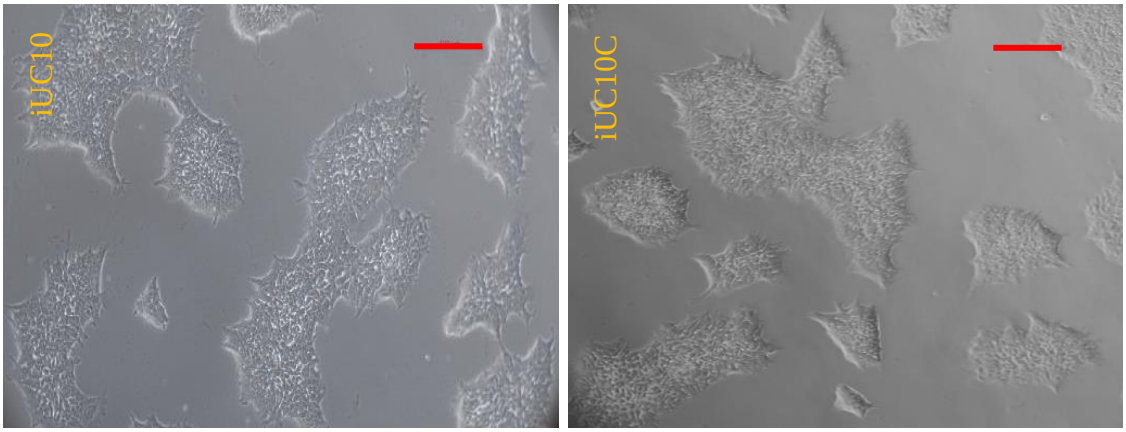

Figure 3.8 - Morphology of the iUC10C line compared with iUC10. Scale bars: 100µm.

To validate pluripotency properties of the generated cell line, the correctly edited hiPSC clone was tested for the protein expression of OCT4 and SSEA4, using IF analysis. The positive control used was the iUC10 cell line used to generate the in-test cell line.

These markers were similarly expressed between the two lines (iUC10 and iUC10C), without a notorious difference in pluripotency properties. They were confirmed the expression of self-renewal transcription factors OCT4 and the surface marker SSEA4, characteristic markers of pluripotent ESCs (Figure 3.9). As a control of cellular quantification, cells were stained with DAPI that exclusively marks DNA in the nucleus, more specifically the A-T rich regions. The secondary antibody (Sec. Ab.) test demonstrate his specificity to the primary antibody.

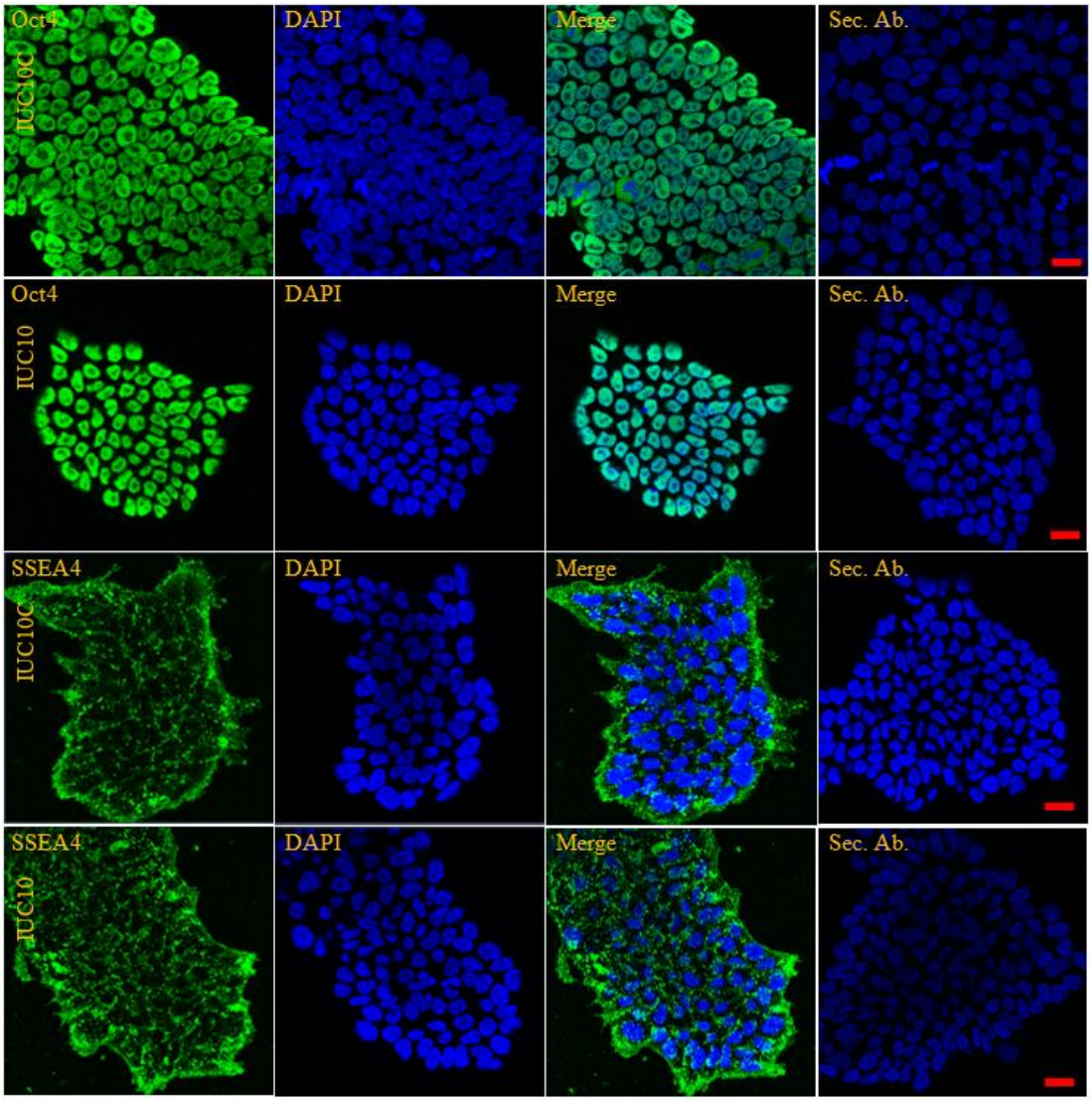

Figure 3.9 - Immunodetection of pluripotency markers, Oct4 and SSEA4, of the iUC10C line compared with iUC10. Nuclei were stained with DAPI. Scale bars: 20µm.

3.5. Internal controls of differentiation

The undifferentiated hiPSCs in culture were used to test the pluripotency potential and spontaneous differentiation by the hanging droplet method. Test the capacity of spontaneous differentiation in the three germ layers is a crucial validation step to assurance the pluripotency properties of the generated hiPSC line. The medium used in the procedure are not capable of maintain hiPSC properties, allowing is spontaneous differentiation. The hiPSC cultured in suspension tend to aggregate to each other's forming a spheroid body composed by several layers of cells and through cellular signalling, endoderm, mesoderm and ectoderm started to differentiate (Figure 3.9). Then

the EBs from the dish were added to a coated plate to allow the cell dispersion followed by IF analysis (Figure 3.9).

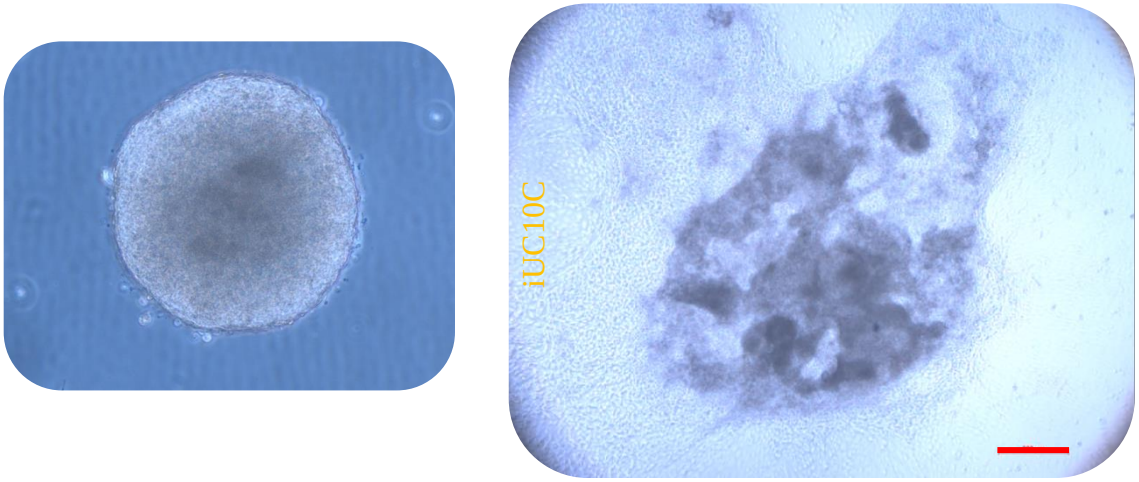

Figure 3.10 - Embryoid Bodies, formed from the derived iUC10C cell line, differentiate spontaneously into the three germ layers, in suspension at left and plated at right. Scale bar: 200µm.

In vitro EB-based differentiation followed by IF analysis of the endodermal marker AFP, mesodermal marker SMA and the ectodermal marker TUBB3 confirmed the pluripotency of iUC10C and their ability to differentiate into all three germ layers (Figure 3.11). As a control of cellular quantification, cells were stained with DAPI. The secondary antibody test demonstrates his specificity to the primary antibody.

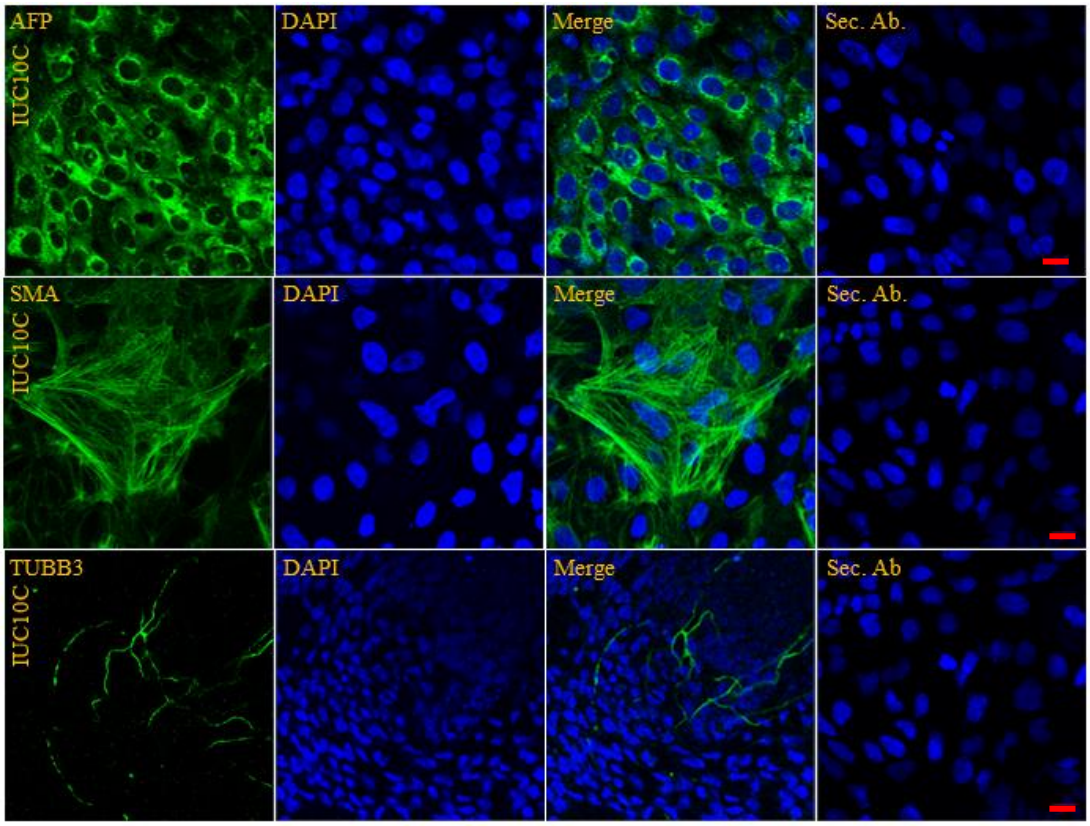

Figure 3.11 - Immunofluorescence analyses of *in vitro* differentiation of EBs using specific antibodies against the AFP, TUBB3 and SMA for the iUC10C cell line. Nuclei were stained with DAPI. Scale bars: 20µm.

CHAPTER IV: Generation of a *DAND5* KO isogenic induced pluripotent stem cell lines for disease modelling in patients with *DAND5* mutations

The hiPSC line, previously, generated from ERE cells of a patient affected with CHDs and laterality defects, designed iUC10, carries a heterozygous non-synonymous variant, c.455G>A, in exon 2 of *DAND5* gene. Here it was used CRISPR-*Cas9*-based genome editing tool to insert premature stop codons in iPSCs derived from *DAND5* patient by HDR. The premature stop codons should be altered in the gene sequence beginning, at 5 prime of *DAND5* gene exon 1, to disrupt completely his protein and do not allow the translation of non-functional protein residues. In order to complete *DAND5* loss-of-function, *DAND5* KO, both loci must be edited, the sequence with the genetic variation as well as the one that does not have it, since it is a heterozygotic iPSC line.

4.1.Design of targeting components of CRISPR Tool

Target gene: DAND5 allele sequences

5´ – ATGCTCCTTG GCCAGCTATC CACTCTTCTG TGCCTGCTTA GCGGGGCCCT
GCCTACAGGC TCAGGGAGGC CTGAACCCCA GTCTCCTCGA CCTCAGTCCT
GGGCTGCAGC CAATCAGACC TGGGCTCTGG GCCCAGGGGC CCTGCCCCCA
CTGGTGCCAG CTTCTGCCCT TGGGAGCTGG AAGGCCTTCT TGGGCCTGCA
GAAAGCCAGG CAGCTGGGGA TGGGCAGGCT GCAGCGTGGG CAAGACGAGG
TGGCTGCTGT GACTCTGCCG CTGAACCCTC AGGAAGTGAT CCAGGGGATG
TGTAAGGCTG TGCCCTTCGT TCAGGTGTTC TCCCGGCCCG GCTGCTCAGC
CATACGCCTC CGAAATCATC TGTGCTTTGG TCATTGCTCC TCTCTCTACA
TCCCTGGCTC GGACCCCACC CCACTAGTCC TGTGCAACAG CTGTATGCCT
GCTCA(/G)C AAGCGTTGGG CACCCGTGGT CCTGTGGTGT CTCACTGGCA
GCTCAGCCTC CCGTCGACGG GTGAAGATAT CCACCATGCTG ATCGAGGGGT
GTCACTGCAG CCCAAAAGCA TGA – 3´, DAND5 exon 1 at grey, DAND5 exon 2 and G>A alteration at green.

For CRISPR tool design is necessary define what the sequence target in the genomic DNA, design the sgRNA and the ssODN repair template. The specific sgRNA were designed using the CRISPR design web tool in *crispr.mit.edu* and the genome was the *DAND5* gene exon 1.

all guides

scored by inverse likelihood of offtarget binding
mouse over for details ... show legend

	score	sequence	
Guide #1	89	TAGGCAGGGCCCCGCTAAGC	AGG
Guide #2	80	CTTAGCGGGGCCCTGCCTAC	AGG
Guide #3	78	TCTTCTGTGCCTGCTTAGCG	GGG
Guide #4	62	ACTCTTCTGTGCCTGCTTAG	CGG
Guide #5	61	GAAGAGTGGATAGCTGGCCA	AGG
Guide #6	61	GGTCGAGGAGACTGGGGTTC	AGG
Guide #7	61	GGGGCCCTGCCTACAGGCTC	AGG
Guide #8	60	GGCACAGAAGAGTGGATAGC	TGG
Guide #9	57	CTCTTCTGTGCCTGCTTAGC	GGG
Guide #10	53	CCTCCCTGAGCCTGTAGGCA	GGG
Guide #11	52	GGGCCCTGCCTACAGGCTCA	GGG
Guide #12	49	TCAGGCCTCCCTGAGCCTGT	AGG
Guide #13	41	GCTAAGCAGGCACAGAAGAG	TGG
Guide #14	39	GCCTCCCTGAGCCTGTAGGC	AGG
Guide #15	28	CCCTGCCTACAGGCTCAGGG	AGG

guide #1 **quality score: 89**

guide sequence: TAGGCAGGGCCCCGCTAAGC AGG
on-target locus: chr19:-12969692
number of offtarget sites: 113 (0 are in genes)

top 20 genome-wide off-target sites ☐ show all exonic

sequence	score	mismatches	UCSC gene	locus
GAGGCAGAGCCCCGCTGAGCGGG	1.3	3MMs [1:8:17]		chr7:-79431008
CAGGCAGCGCCCTGCTAAGCGGG	1.2	3MMs [1:8:13]		chr19:+50205433
AAATCAGGGCCCAGCTAAGCGAG	0.6	4MMs [1:3:4:13]		chr6:+152724673
CAGGCAGCTCCCCGCTAAGAAGG	0.4	4MMs [1:8:9:20]		chr7:+17804097
TGGGCATGGGCCCGCTGAGCTAG	0.4	4MMs [2:7:10:17]		chrX:+17884662
GGGGCTGGGCCCCGCTGAGCAAG	0.4	4MMs [1:2:6:17]		chr4:+185945989
TAGGATGAGCCCAGCTAAGCAGG	0.3	4MMs [5:6:8:13]		chr1:-19425090
AATGCAGGGCCACGCTCAGCGGG	0.3	4MMs [1:3:12:17]		chr7:-35596359
CGGGCAGGGCCGGGCTAAGCTGG	0.3	4MMs [1:2:12:13]		chr18:-54181267
GGGGCAGGGCCCAGCTAAGGGGG	0.3	4MMs [1:2:13:20]		chr5:-62959822
GAGGCAGGGCCGCGCGAAGCAAG	0.3	3MMs [1:12:16]		chr10:-23284250
GAGGCTGGTCCTCGCTAAGCCAG	0.3	4MMs [1:6:9:12]		chr3:-123593634
CAGGCAGCGCCCTGCTCAGCCAG	0.2	4MMs [1:8:13:17]		chr6:+138649253
TTGGCAGAGCCCAGCTCAGCAAG	0.2	4MMs [2:8:13:17]		chr17:-41332018
CAGGCAGGGCCCCACTCAGCGAG	0.2	3MMs [1:14:17]		chr4:+40829711
TACGCAGGATCCCGCTAGGCCAG	0.2	4MMs [3:9:10:18]		chr1:-238390491

Figure 4.1 - Possible gRNA sequences, detached from CRISPR design web tool, with guide 1 Cas9 cut score and suitable off-targets.

The sequence CACCGTAGGCAGGGCCCCGCTAAGC was selected and it serves as template for design a pair of primers (Figure 4.1). This sequence was score for Cas9 cut of 89 and 113 off-targets, but all with low probability of occurring. The top oligo sequence must be preceded by a 5´ CACCG sequence - gRNA#1.KO – CACCGTAGG CAGGGCCCCGCTAAGC - and the bottom oligo must be preceded by a 5´ AAAC and followed by a 3´ C sequences - gRNA#1.KO – AAACGCTTAGCGGGGCCCTGCCT AC. The pair of primers were annealed after designed and inserted in pCAG-SpCas9-GFP-U6-gRNA with Bbs1 cloning sites for, ultimately, prepare the sgRNA expression construct.

Analysis of plasmid DNA samples extracted from three different *E. coli* colonies in an agarose gel show the expected plasmid size and, did not expose any contaminants (Figure 4.2). The DNA concentration is determined and the insertion of the sgRNA into the plasmid is verified by sequencing. The concentration results were 608,8, 595,3 and 766,9 ng/µL, respectively for the *E. coli* colonies 1,2 and 3. However, after the sequencing results, just the colony 1 was right to be use, displaying the sgRNA insertion in the plasmid vector (Figure 4.3).

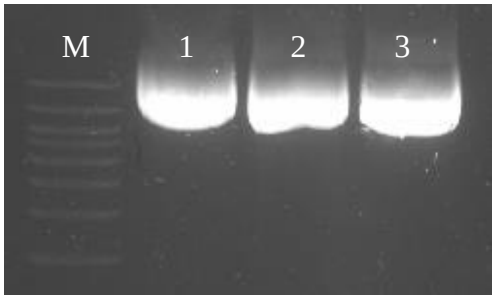

Figure 4.2 - Agarose gel results of sgRNA plasmid construction. M- molecular weight marker; 1- Plasmid DNA from colony 1; 2- Plasmid DNA from colony 2; 3- Plasmid DNA from colony 3.

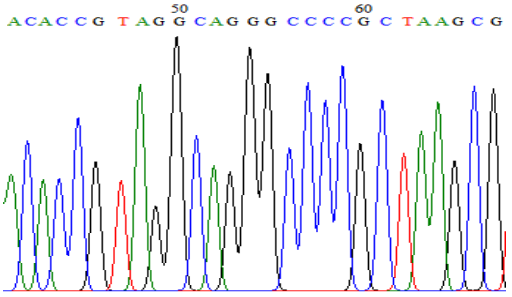

Figure 4.3 - DNA sequence confirming gRNA#1 sequence into pCAG-SpCas9-GFP-U6-gRNA from colony 1.

The other CRISPR system component is the ssODN repair template. This element carries the bases to be altered in the genome. In this case, it was proposed insert in the genome three stop codons in a row for the *DAND5* KO, promoting protein disruption. Besides that, it was inserted a neutral enzymatic restriction sequence for a preliminary clone selection, more specifically 5´-GGATCC-3´sequence, which is recognized for the FastDigest BamHI restriction enzyme. That way, the ssODN repair template designed was: Template 1 – ATGCTCCTTGGCCAGCTATCCACTCTTCTGTGCCTGCTTAG CTAATAATAACCTACAGGATCCGGGAGGCCTGAACCCCAGTCTCCTCGACC TCAGTCCTGGGCTGCAG.

The PAM sequence was not modified in this template, what make increase the probability of Cas9 re-cutting events. CRISPR system could be used without ssODN repair template and in this case, the Cas9 cut will cause indel mutation by NHEJ repair, what ultimately cause a protein disruption, as pretended. Although, this procedure is quite imprecise and unpredictable. That way, it was used CRISPR system with ssODN repair template, where Cas9 re-cutting events could be a problem.

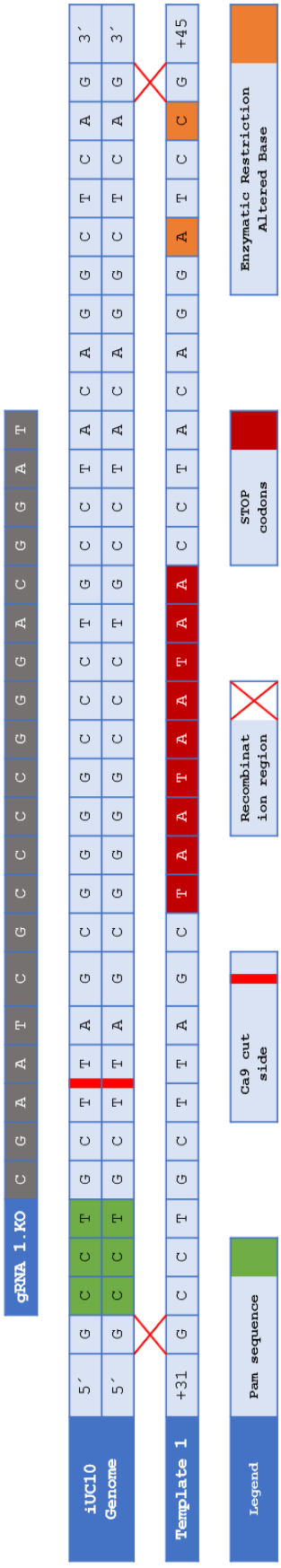

Figure 4.4 - Scheme of CRISPR/*Cas9* system mode of action by HDR for gene KO. The two loci of *DAND5* gene exon 1 in the iUC10 cell line are represented with the gRNA 1.KO and Template 1 homology places. It is signalized the Cas9-cut at 3 bp from PAM sequence at 3-prime direction. As well as, the expected recombination region between the Template 1 and the iUC10 genome loci, after Cas9 cut, which is before the three stop codons and after the last enzymatic restriction altered base.

Anyway, it is expected that Cas9 cuts in both loci, followed recombination between the ssODN repair template and the genomic *DAND5* gene sequence by HDR. That way, the two loci of *DAND5* gene exon 1 in the iUC10 cell line will be edited, inserting in the genome three stop codons in a row as well as the two bases for the sequence recognised by the restriction enzyme (Figure 4.4).

4.1.1. hiPSC culture, transfection and isolation of clonal cell lines

The cell line used was the same used in the Chapter III, the iUC10 cell line. It was cultured with supplemented E8 or SF, which are capable of maintain characteristic ESC-like properties during hiPSC culture. During that time cells were checked, and its medium changed every day to prevent spontaneous differentiation.

Before and after the procedures which implicate much manipulation, as cell transfection or FACS, the medium was refreshed or changed and added Revitacell. In the procedures where likely contamination, it was added Penicillin-Streptomycin to prevent it, as it is the FACS procedure case. Analysed in detail in the Chapter III.

The cells in culture were transfected with 2µg of gRNA#1.KO expression construct and 1µM of the specific ssODN repair template, Template 1. After that, it was allowed cells recovery and FACS was performed. The cells with GFP expression and propidium iodide negatives was selected. The clones were enabled to expand and were used for DNA isolation and genotyping, as explained before.

4.1.2. Detection of genomic alteration by PCR and sequencing

The gDNA from hiPSCs clones and iUC10 cell line were extracted and their quantity and quality were evaluated. After that, the exon 1 of *DAND5* gene was amplified by PCR and the iUC10 DNA was used as reactions control, for all the clone hiPSC lines.

The amplicons from each PCR sample were separated in the agarose gel and the samples which had resulted in a clear band with 499bp were selected. The clones without genomic alteration were selected too, once then amplicons with the same size are indistinguishable (Figure 4.5). This problem could be solved by enzymatic digestion, once then the enzymatic restriction sequence is in ssODN repair template and it is incorporated in the edited cells genome. For that, the PCR samples were digested and separated in parallel with those have not been digested. The samples which have been

resulted in clear bands with 369, 130 and maybe 499bp fragments sizes for *DAND5* exon 1 PCR after electrophoresis, are positives. After those steps 4 in 25 clones analysed seems to be positives (Figure 4.5). The enzymatic digestion results of the 2, 9 and 14 clones presented separation in three bands. But while the clone 2 has a strong 369bp band, probably as result of incomplete digestion, the others two have a 499bp strong band, probably as result of just one locus has been edited. In relation with the clone 17 there was no doubt about the two loci have been edited, as initially intended for the generation of a *DAND5* KO hiPSC line (Figure 4.7).

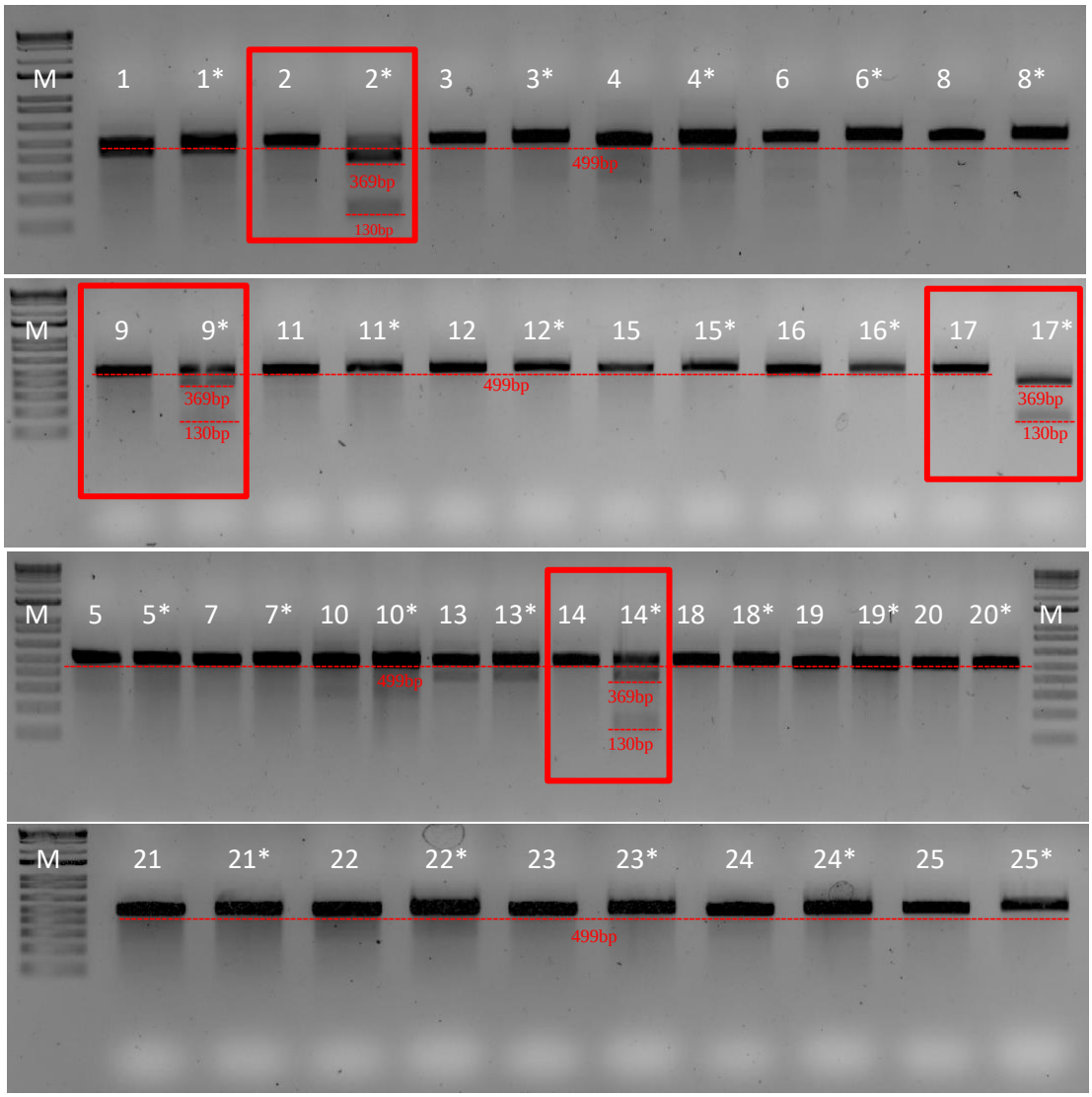

Figure 4.5 - hiPSC clones genotyping of the iUC10 derived cells for generation of DAND5 KO cell line, confirmed by PCR and enzymatic digestion. hiPSC lines PCR positive results present 499bp band. hiPSC lines digestion positive results present 369 and 130 bp bands. M - molecular weight marker; hiPSC clones PCR results: ℕ; hiPSC clones´ digestion results: ℕ*.

The ssODN repair template has the three stop codons in a row and it is expected, that at the time of recombination, the edited clones have the enzymatic restriction altered bases, as well as the stop codons. Once, it is dependent on the recombination site, the previous selection is not enough because that not always occurs properly and must be verified by sequencing.

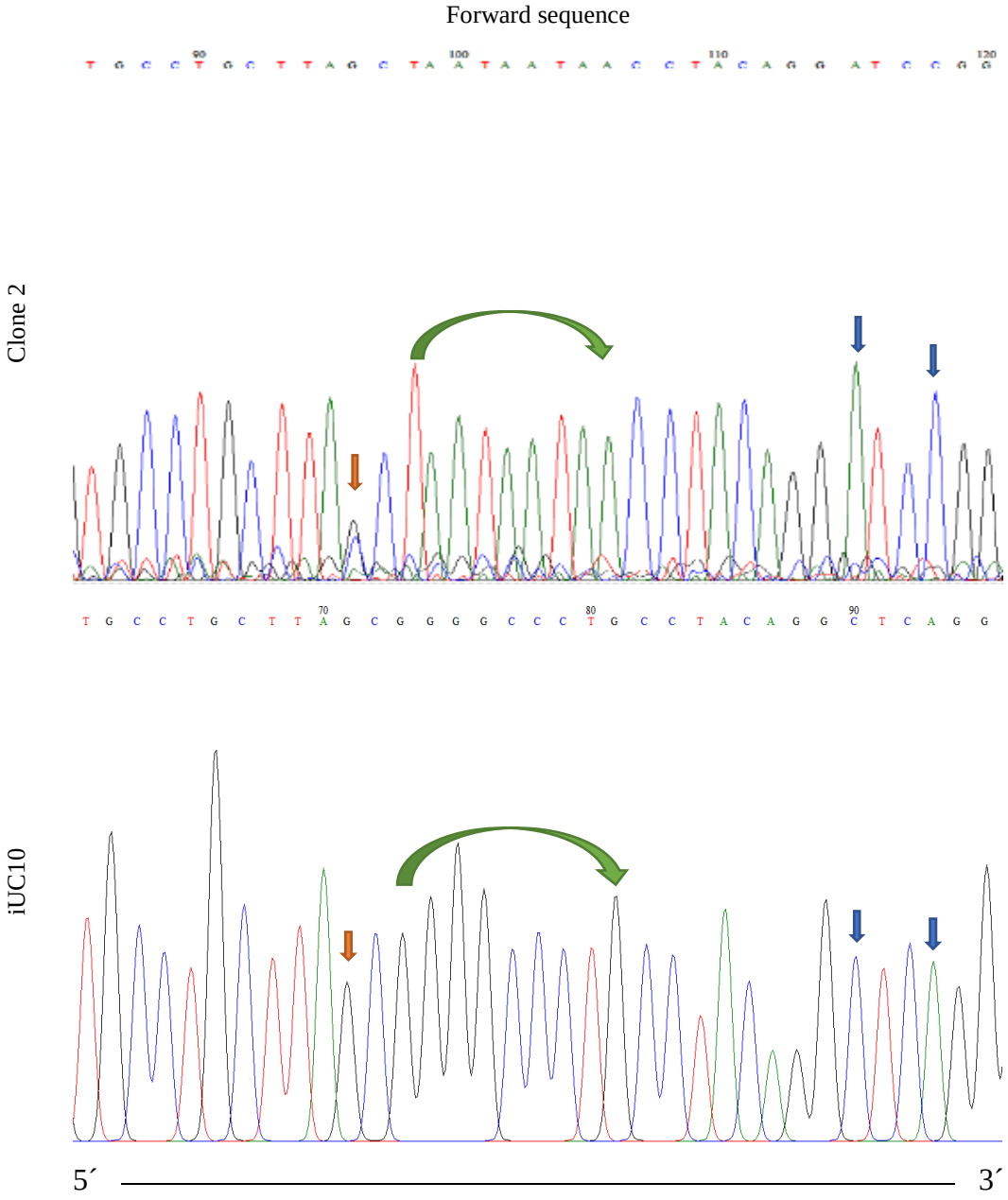

Figure 4.6 - DNA sequences confirming the three stop codons (green arrow) insertion for the *DAND5* KO in the clone 2 hiPSC line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow). And additionally, the point of mutation G>C (orange arrow).

The clones 2 and 17 were analysed by Sanger sequencing and the results confirmed that in both clones the *DAND5* gene was edited in both loci, resulting in a KO gene (Figure 4.6 and 4.8). The clone 2 chromatogram shows the three stop codons in a row and the two bases for the completed enzymatic restriction sequence, but beside it presents a mutation before the stop codons and near Cas9 cut (Figure 4.6). The clone 17 chromatogram analyses showed a correct gene editing, with the three stop codons in a row and the two bases for the completed enzymatic restriction sequence, exactly as intended (Figure 4.8). That way just this clone was used for further analyses. The clone 17, designed now iUC10KO, has the intended gene alterations, and the generation of a *DAND5* KO cell line goal have been achieved.

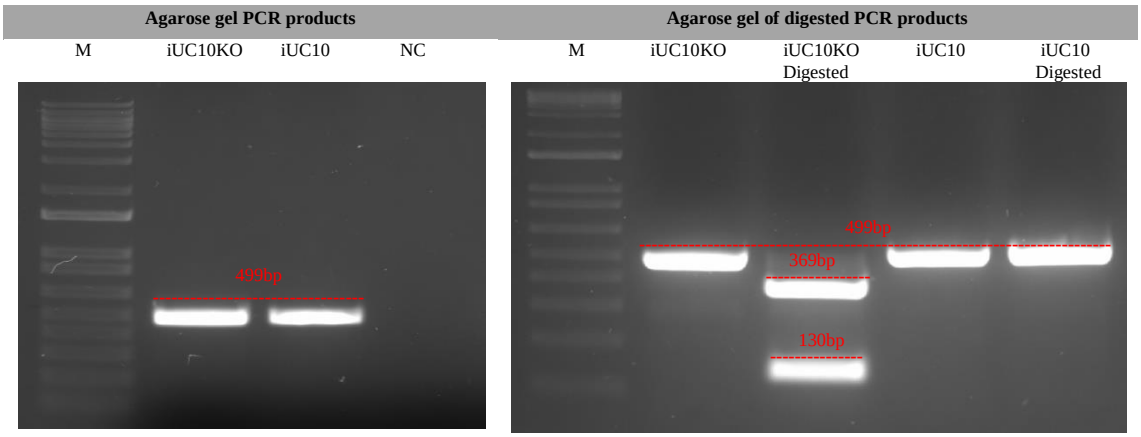

Figure 4.7 - iUC10KO cell line genotyping by PCR and enzymatic digestion. iUC10KO and iUC10 PCR results present 499bp band. iUC10KO digestion results present the 499bp amplicons fragmentation in 369 and 130bp bands when compared to the control, iUC10 cell line.

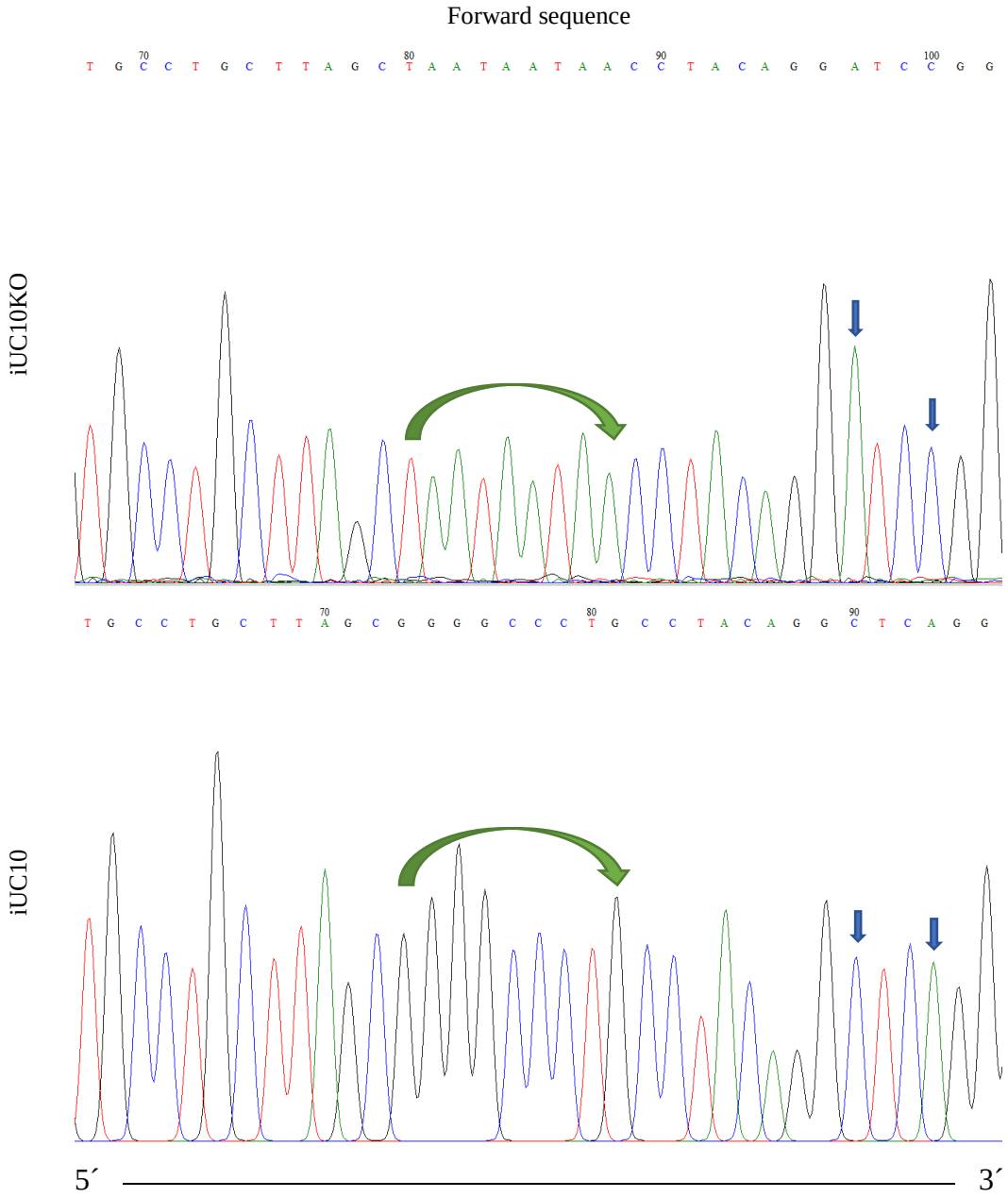

Figure 4.8 - DNA forward sequences confirming the three stop codons (green arrow) insertion for the DAND5 KO in the iUC10 KO cell line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow).

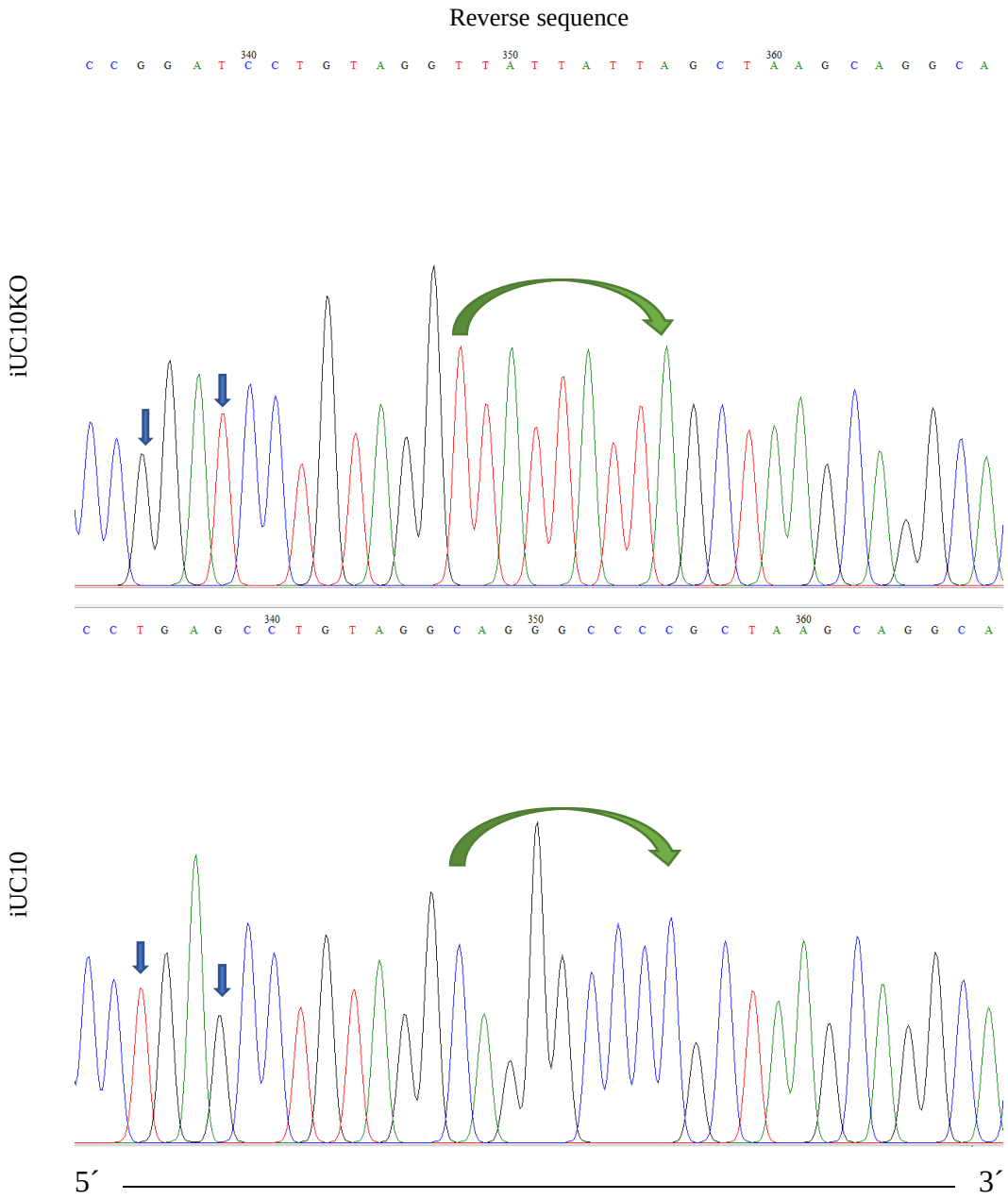

Figure 4.9 - DNA reverse sequences confirming the three stop codons (green arrow) insertion for the DAND5 KO in the iUC10 KO cell line, compared to iUC10. As well as, the C>A and A>C bases alterations to set enzymatic restriction sequence (blue arrow).

As have been done before for the iUC10C, the iUC10KO cell line was then amplified and tested for internal pluripotency and differentiation controls.

4.1.3. Internal controls of pluripotency

After expansion, iUC10KO cell line continued to display a typical ESC-like (Figure 4.10).

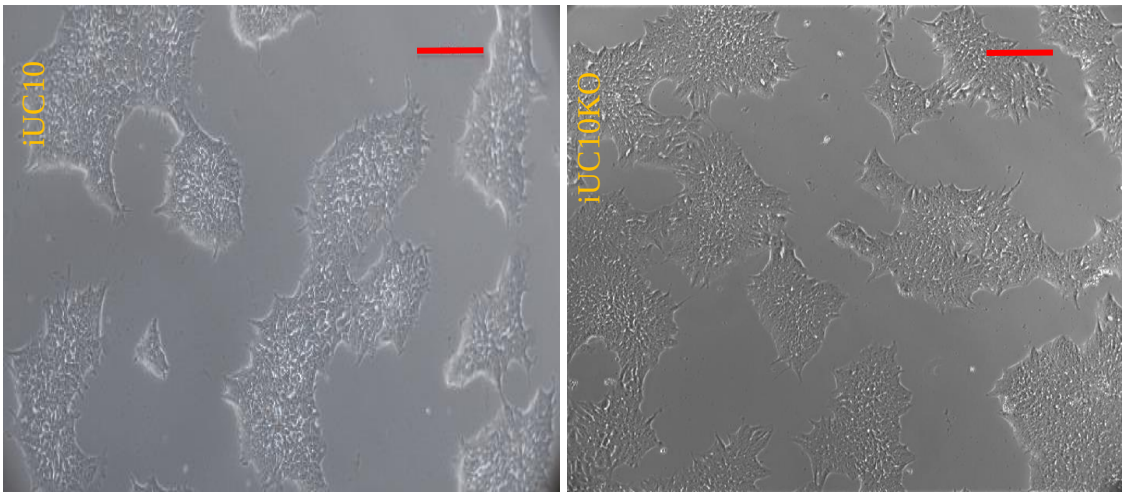

Figure 4.10 - Morphology of the iUC10KO cell line compared with iUC10. Scale bars: 100µm.

To test the ESC-like properties of this cell line, firstly the pluripotency markers, OCT4 and SSEA4, were used to test its pluripotency by IF technique and as positive control the iUC10 cell line was used. As transcription factor, OCT4 is expressed in the nucleus and the Merge image of Figure 4.11 shows the conjugation of blue from DAPI and green from OCT4. On the other hand, the surface marker SSEA4 in the Merge image of the same figure shows green fluorescence around the nucleus, where the membrane would supposed be, though, as expected, in clusters due to the permeabilization step. Sec. Ab. test demonstrated its specificity to the primary antibody once there was not green fluorescence signal. In general, the results showed similar expression of the markers in both iUC10 and iUC10KO cell lines, confirming its pluripotency state. (Figure 4.11)

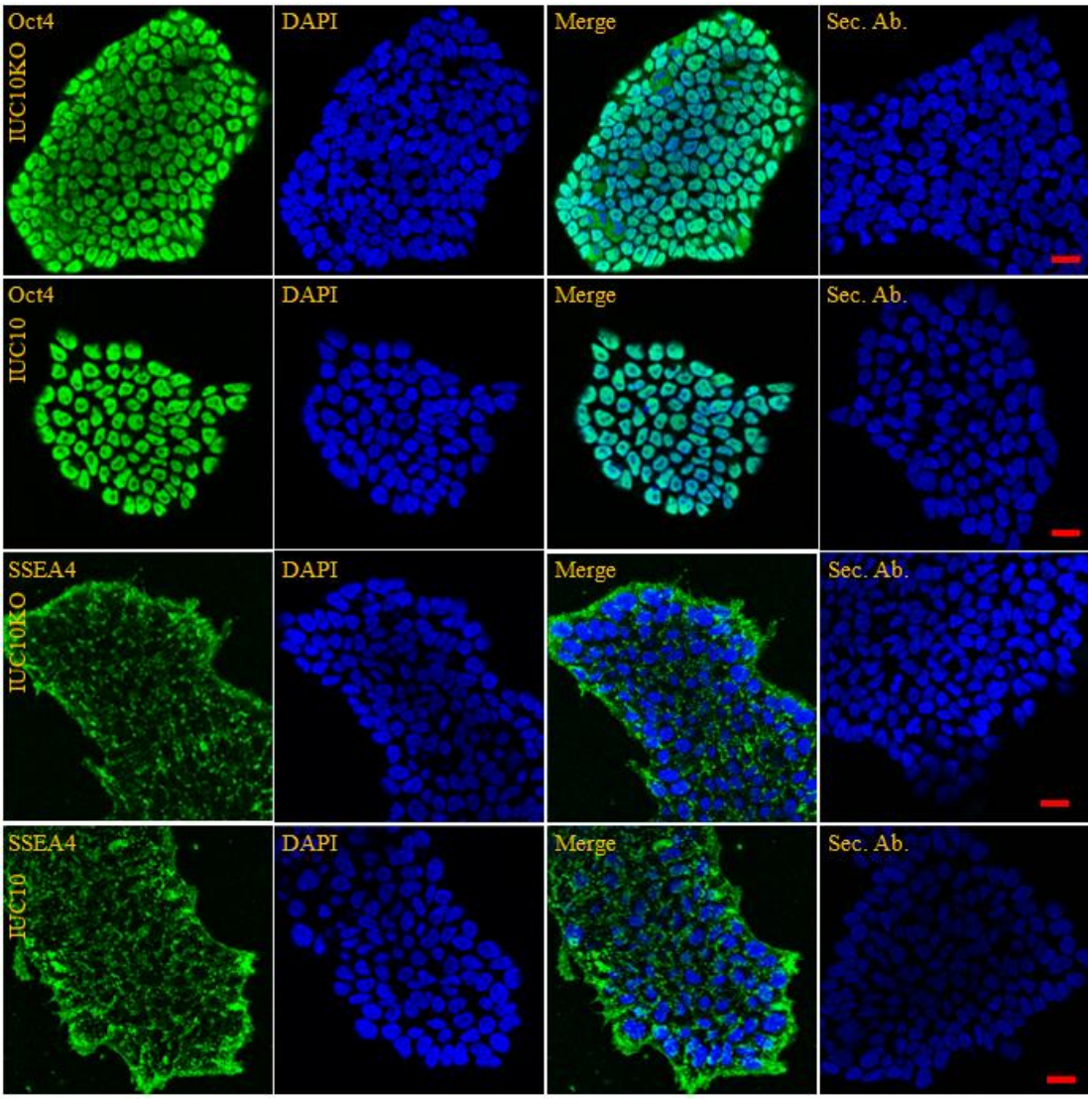

Figure 4.11 - Immunodetection of pluripotency markers, Oct4 and SSEA4, of the iUC10KO cell line compared with iUC10. Nuclei were stained with DAPI. Scale bars: 20μm

4.1.4. Internal controls of differentiation

The undifferentiated iUC10KO cells in culture were differentiated by the hanging droplet method to test its pluripotency potential. In the Figure 4.12, the spheroid EB (at left) is a cell aggregate composed for several layers of cells in differentiation. When iUC10KO EBs in suspension were added to a coated plate, the cells dispersed across the plate from the clustered cells bodies (Figure 4.12 at right).

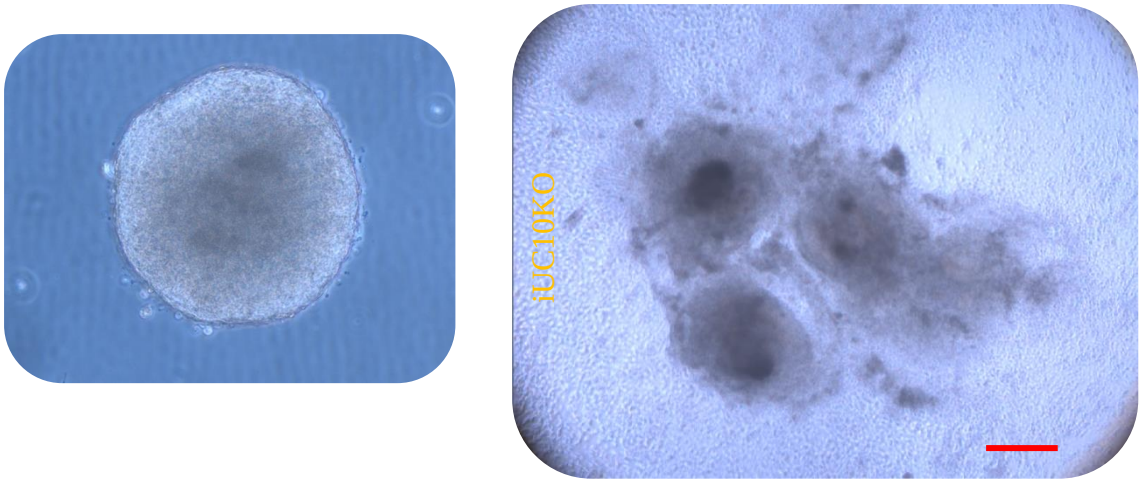

Figure 4.12 - Embryoid Bodies, formed from the derived iUC10KO cell line, differentiate spontaneously into the three germ layers, in suspension at left and plated at right. Scale bar: 200µm.

After 20 days it is expected spontaneous differentiated cells. The *in vitro* EB-based differentiation was followed by IF analysis and the results are presented in the Figure 4.13. The results showed the presence of green fluorescence signal for the endodermal marker (AFP), mesodermal marker (SMA) and ectodermal marker (TUBB3) tests, confirming the iUC10KO cell line capacity to differentiate in the three germ layers and its pluripotency potential.

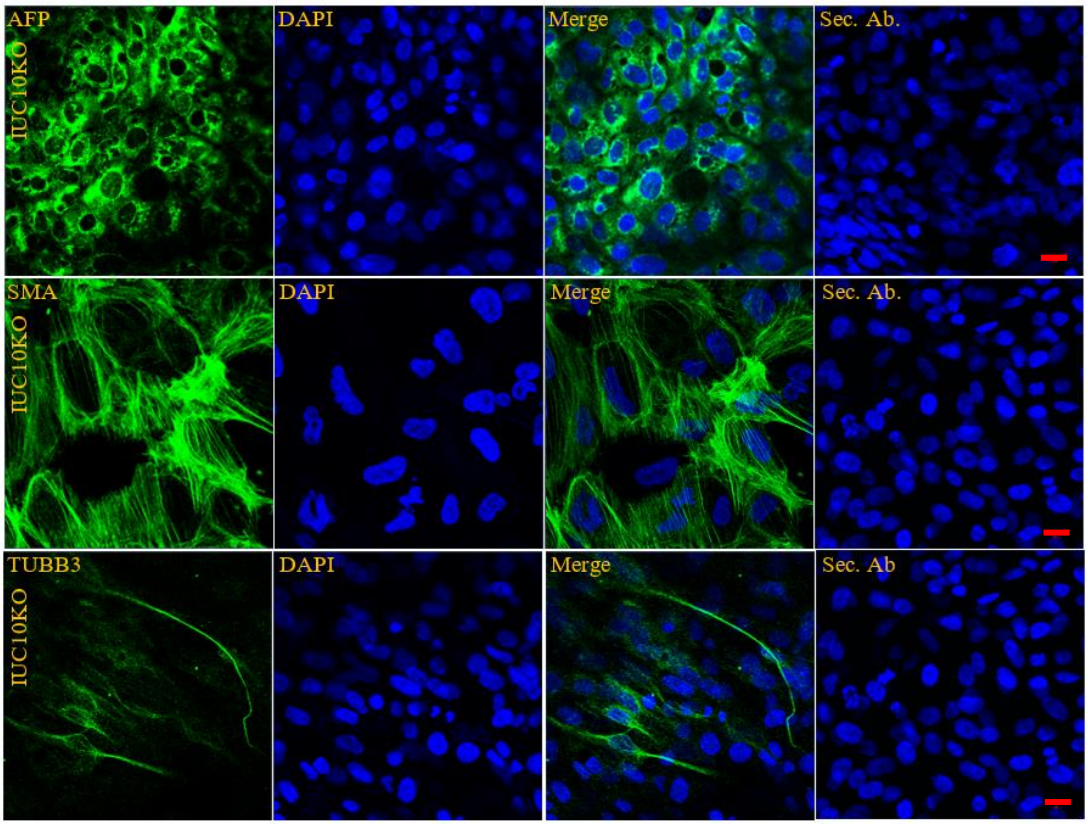

Figure 4.13 - Immunofluorescence analyses of *in vitro* differentiation of EBs using specific antibodies against the endodermal marker AFP, ectodermal marker TUBB3 and mesodermal markers SMA, for the iUC10KO cell line. Nuclei were stained with DAPI. Scale bar: 20µm.

CHAPTER V: Discussion and Conclusions

hiPSC technology has an enormous potential in the treatment of various human diseases. Before iPSCs, ESCs has been studied in many parts of regenerative medicine as well as disease modelling. However ethical issues and immunological rejection, which ESCs treatment faces, discourages their use. iPSC, on the other hand, can easily be generated from almost any patient cell source, what provides us unlimited access to fresh disease samples. Other option was animal studies and immortalized cell studies, but the problem of difference between species often nullifies the novel drugs and therapy study results. Immortalized cell-based studies ignore pathology development issue in real patient yet.(22)

This project is an example how iPSCs can be an important source of information using the CRISPR-*Cas9* system to generate a disease model of patients with *DAND5* mutations. The hiPSC line used was generated from ERE cells isolated from a urine sample obtained from a 7-year old male child. The patient was clinically diagnosed with ventricular septal defect with overriding aorta, right ventricular hypertrophy and pulmonary atresia (a case of extreme tetralogy of Fallot phenotype). Genetically, the patient carries a heterozygous non-synonymous variant in exon 2 of *DAND5* gene (c.455G>A). This variation is partially responsible for the disease, but is not fully responsible, since the child´s mother does not present any apparent symptom of the disease, despite having the same genetic variation. (21)

To know the importance of the gene in the observed phenotype is necessary to evaluate the effects caused by the mutation, correcting the mutation, and using a *DAND5* KO to test a more severe phenotype in contrast to the corrected line. Additionally, replicate the mutation site of *DAND5* in a control iPSC line may provide valuable information on the influence of the mutation of interest in different genetic backgrounds. It has already been described that the mutation causes a loss of protein function and in turn less inactivation of *Nodal*. In this way we can test the effect of total loss of protein, as well as what is the epigenetic importance for the phenotype development. (2,10)

Using CRISPR-mediated genome editing, we were capable to correct the defective *DAND5* mutation and generate an isogenic *DAND5*-corrected stem cell line, that means, is homozygous for the c.455G base, confirmed by sequencing. The generated isogenic *DAND5*-corrected stem cell line, iUC10C, has ESC-like morphology. iUC10C cell line expresses the pluripotency markers SSEA4 and OCT4, and after EBs formation by

hanging droplet method it expresses the differentiation markers AFP (endodermal marker), SMA (mesodermal marker) and TUBB3 (ectodermal marker) by IF assays. That means then that this cell line maintains the ESC-like properties observed in the original iUC10. The same results were obtained for the generated isogenic *DAND5* KO stem cell line, iUC10KO, after using CRISPR-mediated genome editing to insert premature stop codons, making the *DAND5* gene inoperative.

These hiPSC lines are now useful tools for disease modelling to understand the role of this gene in patients with *DAND5* mutations, drug screening and regenerative medicine. Specially to study the molecular mechanisms of cardiomyocyte proliferation, once, like ESCs, iPSCs are multipotent and have the ability to differentiate into cardiac and vascular cells. Different studies have established protocols to differentiate into myocytes with the structural and functional properties of cardiomyocytes. *In vitro* cardiomyocyte differentiation reproduces the process of cardiogenesis with the initial induction of mesoderm and then of cardiac precursors. To improve the efficiency of cardiomyocyte differentiation from iPSCs, culture conditions have been optimized with several cytokines and growth factors, such as BMP4 and activin/Nodal, during its differentiation. And the emerging interest in the generation of cardiospheres as an alternative approach for therapeutic application of the treatment of CVDs bring the opportunity to study this process. (2,29,43)

However, iPSC may present a few problems. Comparing with ESCs, iPSCs present differences in epigenetic markers, gene expression level and genome methylation status, including an epigenetic memory of the original cell source. One of the most important issues in using epigenetic reprogramming and site-specific gene correction concern to ensure the integrity of the chromosomal DNA which could be affected by the reprogramming process, the prolonged *in vitro* culture, and/or the gene correction process itself. Consequently, bi-sulfide sequencing techniques, whole genome expression array analysis, whole genome sequencing, karyotyping or other kind of tests may be performed to ensure the procedures accuracy. (22,29,41,44)

The same techniques could be used to attest the iPSC faithfulness after the use of gene editing tools as the CRISPR-*Cas9* system, which off-targets are the principal problem. The high Cas9 efficiency has demonstrated significant increase in the off-target frequency when compared with ZFNs or TALENs, leading to indel mutations in

unwanted sites due to the leniency of CRISPRs toward mismatched base-pairing between gRNA and target sequences, which can disrupt gene function and cause genome instability and therefore influence research and therapeutic applications. The double nicking system was reported with reduced off-target activity and may be used as alternative method. (2,32)

Further work goes through characterize the effect of the defective *DAND5* in the regulation of *Nodal* and *Wnt* signalling during cardiomyocyte differentiation/ proliferation and inferring how different DAND5 levels reflect in cardiomyocyte proliferation and *DAND5*-associated heart diseases.

CHAPTER VI: References

1. WHO | Cardiovascular diseases (CVDs) [Internet]. WHO. World Health Organization; 2017 [cited 2018 Jan 18]. Available from: http://www.who.int/mediacentre/factsheets/fs317/en/
2. Jeziorowska D, Agathe K, Salem J, Fish K, Hulot J-S. Generating patient-specific induced pluripotent stem cells-derived cardiomyocytes for the treatment of cardiac diseases. Expert Opin Biol Ther. 2015;15(10):1399–409.
3. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe : epidemiological update. Eur Heart J. 2013;34:3028–34.
4. Townsend N, Wilson L, Bhatnagar P, Wickramasinghe K, Rayner M, Nichols M. Cardiovascular disease in Europe: epidemiological update 2016. Eur Heart J. 2016;37:3232–45.
5. Czubryt MP. Common threads in cardiac fibrosis , infarct scar formation , and wound healing. Fibrogenesis Tissue Repair. 2012;5(19):1–11.
6. van der Linde D, Konings EEM, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJM, et al. Birth Prevalence of Congenital Heart Disease Worldwide. J Am Coll Cardiol. Journal of the American College of Cardiology; 2011;58(21):2241–7.
7. Pediatric Cardiac Genomics Consortium PCG, Gelb B, Brueckner M, Chung W, Goldmuntz E, Kaltman J, et al. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. Circ Res. 112(4):698–706.
8. Icardo JM, Rincón JMG, Ros MA. Congenital heart disease, heterotaxia and laterality. Revista española de cardiología. 2002;55(9):962–74.
9. Araújo AC, Marques S, Belo JA. Targeted inactivation of Cerberus like-2 leads to left ventricular cardiac hyperplasia and systolic dysfunction in the mouse. PLoS One. Public Library of Science; 2014;9(7):e102716.
10. Cristo F, Inácio JM, de Almeida S, Mendes P, Martins DS, Maio J, et al. Functional study of DAND5 variant in patients with Congenital Heart Disease and laterality defects. BMC Med Genet. 2017;18(77):1–9.
11. Buikema JW, Mady AS, Mittal N V, Atmanli A, Caron L, Doevendans PA, et al. Wnt / β -catenin signaling directs the regional expansion of first and second heart field-derived ventricular cardiomyocytes. Stem Cells Regen. 2013;140(20):4165–76.

12. Lian X, Hsiao C, Wilson G, Zhu K, Hazeltine LB, Azarin SM, et al. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. PNAS. 2012;109(27):E1848–E1857.
13. Belo JA, Marques S, Inácio JM. The Role of Cerl2 in the Establishment of Left-Right Asymmetries during Axis Formation and Heart Development. J Cardiovasc Dev Dis Multidisciplinary Digital Publishing Institute; 2017;4(23):1–10.
14. Sadler TW. Langman´s Medical Embryology. 9th ed. Koogan G, editor. Montana: Lippincott Williams & Wilkins; 2013. 1-115 p.
15. Belo JA, Silva AC, Borges A-C, Filipe M, Bento M, Gonçalves L, et al. Generating asymmetries in the early vertebrate embryo : the role of the Cerberus-like family. Int J Dev Biol. 2009;53:1399–407.
16. Nakamura T, Saito D, Kawasumi A, Shinohara K, Asai Y, Takaoka K, et al. Fluid flow and interlinked feedback loops establish left–right asymmetric decay of Cerl2 mRNA. Nat Commun. 2012;3(1322):1–13.
17. Inácio JM, Marques S, Nakamura T, Shinohara K, Meno C, Hamada H, et al. The Dynamic Right-to-Left Translocation of Cerl2 Is Involved in the Regulation and Termination of Nodal Activity in the Mouse Node. PLoS One. 2013;8(3):e60406.
18. Marques S, Borges AC, Silva AC, Freitas S, Cordenonsi M, Belo JA. The activity of the Nodal antagonist Cerl-2 in the mouse node is required for correct L / R body axis. Genes Dev. 2004;18:2342–7.
19. Nakamura T, Mine N, Nakaguchi E, Mochizuki A, Yamamoto M, Yashiro K, et al. Generation of Robust Left-Right Asymmetry in the Mouse Embryo Requires a Self-Enhancement and Lateral-Inhibition System. Dev Cell. 2006;11:495–504.
20. Oki S, Kitajima K, Marques S, Belo JA, Yokoyama T, Hamada H, et al. Reversal of left-right asymmetry induced by aberrant Nodal signaling in the node of mouse embryos. Development. 2009;136(23):3917–25.
21. Cristo F, Inácio JM, Rosas G, Carreira IM, Melo JB, de Almeida LP, et al. Generation of human iPSC line from a patient with laterality defects and associated congenital heart anomalies carrying a DAND5 missense alteration. Stem Cell Res; 2017;25:152–6.
22. Kim C. Disease modeling and cell based therapy with iPSC : future therapeutic option with fast and safe application. Blood Res. 2014;49(1):7–14.

23. Shi L, Cui Y, Luan J, Zhou X, Han J. Urine-derived induced pluripotent stem cells as a modeling tool to study rare human diseases. Intractable Rare Dis Res. 2016;5(3):192–201.
24. Rawat N, Singh MK. Induced pluripotent stem cell : A headway in reprogramming with promising approach in regenerative biology. Vet World. 2017;10(6):640–9.
25. Park I, Arora N, Huo H, Maherali N, Ahfeldt T, Shimamura A, et al. Disease-specific induced pluripotent stem (iPS) cells. Cell. 2009;134(5):877–86.
26. Van Meer BJ, Tertoolen LGJ, Mummery CL. Concise Review: Measuring Physiological Responses of Human Pluripotent Stem Cell Derived Cardiomyocytes to Drugs and Disease. Stem Cells. 2016;34:2008–15.
27. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell. 2006;126:663–76.
28. Majlinda L, Armstrong L, Stojkovic M. Induced Pluripotent Stem Cells : It Looks Simple but Can Looks Deceive ? Stem Cells. 2010;28:845–50.
29. Tanaka A, Yuasa S, Node K, Fukuda K. Cardiovascular Disease Modeling Using Patient-Specific Induced Pluripotent Stem Cells. Int J Mol Sci. 2015;16:18894–922.
30. Kim B, Jeong S, Lee S-Y, Lee SM, Gweon EJ, Ahn H, et al. Concurrent progress of reprogramming and gene correction to overcome therapeutic limitation of mutant ALK2-iPSC. Exp Mol Med. Nature Publishing Group; 2016;48:1–12.
31. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281–308.
32. Kim HS, Bernitz JM, Lee D, Lemischka IR. Genomic Editing Tools to Model Human Diseases with Isogenic Pluripotent Stem Cells. Stem Cells Dev. 2014;23(22):2673–86.
33. Peters DT, Cowan CA, Musunuru K. Genome editing in human pluripotent stem cells. StemBook, ed. 2013;1(94):1–14.
34. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-Free Zinc-Finger Nuclease Engineering by Context-Dependent Assembly (CoDA). Nat Methods. 2011;8(1):67–9.
35. Carlson DF, Fahrenkrug SC, Hackett PB. Targeting DNA With Fingers and

TALENs. Mol Ther Acids. 2012;1(e3):1–5.

36. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep. 2014;4(5400):1–6.
37. Giacalone JC, Sharma TP, Burnight ER, Fingert JF, Mullins RF, Stone EM, et al. CRISPR-Cas9-Based Genome Editing of Human Induced Pluripotent Stem Cells. Curr Protoc Stem Cell Biol. 2018;44:5B.7.1–5B.7.22.
38. Saha K, Jaenisch R. Technical Challenges in Using Human Induced Pluripotent Stem Cells to Model Disease. Cell Stem Cell. Elsevier Inc.; 2009;5:584–95.
39. Carpenter MK, Rao SM. Concise Review: Making and Using Clinically Compliant Pluripotent Stem Cell Lines. Stem Cell Transl Med. 2015;4:381–8.
40. Sheridan SD, Surampudi V, Rao RR. Analysis of Embryoid Bodies Derived from Human Induced Pluripotent Stem Cells as a Means to Assess Pluripotency. Stem Cells Int. 2012;2012:1–9.
41. Lewandowski J, Kurpisz M. Techniques of Human Embryonic Stem Cell and Induced Pluripotent Stem Cell Derivation. Arch Immunol Ther Exp (Warsz). 2016;64(5):349–70.
42. Heart G, Alliance H, Medical I, Alliance MH. Live Fluorescent RNA-Based Detection of Pluripotency Gene Expression in Embryonic and Induced Pluripotent Stem Cells of Different Species. Stem Cells. 2015;33:392–402.
43. Xu J, Lee Y, Ran X, Liao S-Y, Yang J, Au K-W, et al. Generation of Induced Cardiospheres via Reprogramming of Skin Fibroblasts for Myocardial Regeneration. Stem Cells. 2016;34:2693–706.
44. Garate Z, Davis BR, Quintana-bustamante O, Segovia JC. New Frontier in Regenerative Medicine: Site-Specific Gene Correction in Patient-Specific Induced Pluripotent Stem Cells. Hum Gene Ther. 2013;24(6):571–83.

2019

Micael de Jesus Almeida

Disease correction and production of isogenic induced pluripotent stem cell lines for disease modelling in patients with DAND5 mutations