

DEVELOPMENT OF A HUMAN CORONARY ARTERY DISEASE MODEL

CREATION OF A PATIENT-SPECIFIC ENDOTHELIAL MODEL TO
INVESTIGATE THE ROLE OF DHX34 IN EARLY-ONSET CORONARY ARTERY
DISEASE

By

ALEXANDRIA AFONSO, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the
Requirements for the Degree Doctor of Philosophy

McMaster University

© Copyright by Alexandria Afonso, April 2019

DOCTOR OF PHILOSOPHY (2019)

McMASTER UNIVERSITY

Biochemistry and Biomedical Sciences

Hamilton, ON

TITLE: Creation of a patient-specific endothelial model to investigate the role of DHX34 in early-onset coronary artery disease

AUTHOR: Alexandria Afonso, B.Sc.

SUPERVISOR: Dr. Karun Singh

NUMBER OF PAGES: xix, 243

LAY ABSTRACT

Cardiovascular diseases are the leading cause of death worldwide and is most commonly caused by coronary artery disease (CAD). CAD is characterized by a narrowing of arteries that supply blood to the heart as a result of the build-up of fat and cells in the inner portion of the vessel wall. The cells that line the inner part of the vessel (endothelial cells) are the first barrier against CAD. Through a collaboration with Dr. Paré, we identified new genetic variants in *DHX34* in early-onset CAD patients. We made stem cells from blood donated by these patients and then turned them into endothelial cells to determine if the variants had a role in their disease development. We found that the endothelial cells with the *DHX34* variants migrated more slowly and allowed for more blood cells to attach to them, both of which are processes involved in CAD development.

ABSTRACT

Coronary artery disease (CAD) is characterized by the development of atherosclerotic plaques in the coronary arteries and is one of the leading causes of global mortality. Despite the fact that CAD presents clinically as changes to myocardial function, it is ultimately a disease of the vessel wall. The endothelial cells that line the inner portion of arteries act as a primary defence system against CAD development and as such, were the focus of our work on early-onset CAD (EOCAD). Through a collaboration with Dr. Paré, exome sequencing was performed on EOCAD patients that resulted in the identification of novel, rare variants in *DHX34*, a component of the nonsense-mediated mRNA decay (NMD) pathway. Induced pluripotent stem cells (iPSCs) were generated from a small-number of peripheral blood monocytes from these patients using the Sendai virus. A protocol was generated that improved endothelial differentiation efficiency and was subsequently utilized to investigate the role of DHX34 in endothelial cell function. The patient-derived endothelial cells had decreased migratory capacity, increased leukocyte adhesion, and altered nitric oxide signaling, all of which are necessary in maintaining the anti-atherogenic environment of coronary arteries. Although further work is required to elucidate the specific mechanism through which the variants confer their impact, this work suggests that DHX34 plays a role in preventing CAD development.

ACKNOWLEDGEMENTS

I would like to first thank my supervisor, Dr. Karun Singh, for all of his guidance and help. Without all of the emotional and scientific support he has given me, especially over the last 6 months, I would not have made it to this point. I am so grateful for everything he has done to help me.

I would also like to thank the members of my supervisory committee, Drs. Guillaume Paré and Bernardo Trigatti. Your support and guidance over the years has been crucial in the advancement of my work and overall knowledge. I would also like to thank Kanwaldeep Singh who has always been supportive and extremely helpful.

To my friends – thank you for everything. To my lab friends, especially Nadeem Murtaza, Victor Gordon, and Vickie Kwan, for helping me troubleshoot experiments, for the emotional support in times of iPSC troubles, and for generally making my time spent in lab fun. To all my friends, thank you for all of the support and love, and for knowing when to, and more importantly, when not to ask how things in lab were going. Alicia Skrinjar, thank you for being my first friend and for keeping me sane since. To Sam Chan, thank you for being my science buddy in high school and then fellow grad school-er so we could commiserate together. Emily Taylor, thank you for helping me survive undergrad and for continuing to help me. Rachel Tomsic – thank you for always listening to me complain, being my TV show recommender, shopping partner, and most of all, for always being my cheerleader and best friend. I also have to thank Bart Ronde for everything he has done for me. Thank you for making sure I ate when I got wrapped up in lab work, for always passing me a tissue when I needed one, and for making me laugh when I needed it

most. Thank you for all of the love, and words of reassurance and support you have given over me over the years, even when I have not been my breezy self.

There are not enough words to express my gratitude towards my family. They have always supported me and helped in every way that they could. Thank you to my grandparents, and all of my cousins, aunts, and uncles for your love and for feeding me for weeks at a time with the leftovers from family functions. From helping with my projects in elementary school to editing this thesis, my parents, Maria and Florentino, have ensured my success in all of my academic endeavours and without them, I would not have made it. I would also like to thank my sisters, Samantha and Daniella, for always finding examples of other lab work disasters to make me feel better when my cells weren't cooperating. I also have to thank Julia, Brian, Karin, and Kieran (and Nathan) for the emotional support and for always checking in on my cells. And last but not least, to my cats, Ozzie and Chloe – thank you for your love and for hanging out with (and on) me while I wrote this thesis.

TABLE OF CONTENTS

LAY ABSTRACT	iii
ABSTRACT.....	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vii
TABLE OF FIGURES & TABLES.....	xi
LIST OF ABBREVIATIONS	xv
CHAPTER 1: INTRODUCTION.....	1
1.1 INDUCED PLURIPOTENT STEM CELLS	1
<i>1.1.1 Mechanism of cellular reprogramming</i>	<i>7</i>
<i>1.1.2 Use of iPSCs in disease modelling.....</i>	<i>12</i>
1.2 CORONARY ARTERY DISEASE	15
<i>1.2.1 CAD development</i>	<i>17</i>
<i>1.2.2 Risk factors for CAD development.....</i>	<i>23</i>
<i>1.2.3 Current models and their limitations.....</i>	<i>26</i>
1.3 SUMMARY OF INTENT.....	29
CHAPTER 2: DERIVATION OF AN EFFICIENT PROTOCOL FOR REPROGRAMMING PBMCS INTO IPSCS.....	38
<i>PREFACE</i>	<i>38</i>
2.1 INTRODUCTION.....	39

2.1.1 History of iPSC generation	39
2.1.2 Methods previously used for reprogramming.....	40
2.1.3 Somatic cell sources for reprogramming.....	44
2.1.4 Chapter objectives	48
2.2 MATERIALS AND METHODS	49
2.3 RESULTS	57
2.3.1 Lentiviral transduction did not successfully reprogram non-mobilized PBMCs.....	57
2.3.2 Sendai virus successfully reprograms non-mobilized PBMCs	58
2.3.3 Reduction in starting number of PBMCs does not decrease efficiency with Sendai virus.....	59
2.3.4 Sendai virus reprogramming of PBMCs is independent of CD34 ⁺ population size	60
2.3.5 Sendai reprogramming results in bona-fide iPSCs	61
2.4 DISCUSSION	62
CHAPTER 3: EFFICIENT DIFFERENTIATION OF PATIENT-DERIVED IPSCS INTO MATURE ENDOTHELIAL CELLS	83
PREFACE	83
3.1 INTRODUCTION	84
3.1.1 Role of endothelial cells in tissue homeostasis	84
3.1.2 Role of endothelial cells in disease and methods for studying	86
3.1.3 Existing endothelial differentiation protocols and their limitations.....	92
3.1.4 Chapter objectives	101
3.2 MATERIALS AND METHODS	102
3.3 RESULTS	105

3.3.1 Endothelial cell production with Lian protocol is lower than reported	105
3.3.3 Optimization of previously published protocols increases yield of endothelial progenitor cells	106
3.3.4 Endothelial cells obtained from optimized protocol are more mature	107
3.3.5 Newly established endothelial differentiation protocol produces cells that are functional	108
3.4 DISCUSSION AND FUTURE DIRECTIONS.....	109
 CHAPTER 4: RARE DHX34 VARIANTS ALTER ENDOTHELIAL FUNCTION IN THE CONTEXT OF EARLY-ONSET CORONARY ARTERY DISEASE.....	
PREFACE	125
4.1 INTRODUCTION.....	126
4.1.1 Role of endothelial cells in coronary artery disease development	126
4.1.2 Genetic component of coronary artery disease	131
4.1.3 DECODE project	135
4.1.4 Nonsense-mediated mRNA decay in cell biology and disease.....	137
4.1.5 Chapter objectives	141
4.2 MATERIALS AND METHODS.....	142
4.3 RESULTS.....	147
4.3.1 Reducing DHX34 in HUVECs alters tube formation and migration capacity	147
4.3.2 HUVEC KO cell pools have decreased migration and increased leukocyte adhesion	149
4.3.3 DHX34 KO and UPF1 KO pools have decreased NO production but no changes in eNOS	151

<i>4.3.4 Identification of DHX34 variants in EOCAD patients</i>	<i>153</i>
<i>4.3.5 DHX34 variants alter differentiation of iPSCs</i>	<i>154</i>
<i>4.3.6 Patient iPSC-derived endothelial cells have decreased DHX34 protein but no apparent changes in NMD</i>	<i>155</i>
<i>4.3.7 DHX34 variants impact migration and adhesion of endothelial cells</i>	<i>156</i>
<i>4.3.8 DHX34 variants alter NO production in endothelial cells</i>	<i>158</i>
4.4 DISCUSSION AND FUTURE DIRECTIONS.....	159
CHAPTER 5: CONCLUSIONS & FUTURE DIRECTIONS	202
5.1 PATIENT-SPECIFIC MODELS AS A MEANS TO BETTER STUDY THE GENETIC COMPONENT OF DISEASE ETIOLOGY AND TREATMENT	202
5.2 ROLE OF DHX34 IN ENDOTHELIAL FUNCTIONS ASSOCIATED WITH EARLY-ONSET CORONARY ARTERY DISEASE	204
5.3 SIGNIFICANCE AND FUTURE DIRECTIONS.....	207
BIBLIOGRAPHY	208

TABLE OF FIGURES & TABLES

CHAPTER 1

FIGURE 1. Common sources of human pluripotent stem cells and their potential applications.	32
FIGURE 2. Development and growth of an atherosclerotic plaque in coronary artery disease	34
TABLE 1. The role of cellular subtypes in CAD.....	36
TABLE 2. Animal models of atherosclerosis development.....	37

CHAPTER 2

FIGURE 1. Sendai virus efficiently reprograms a small number of PBMCs irrespective of CD34 ⁺ content	69
FIGURE 2. iPSC colonies generated with sendai virus are positive for classic pluripotency markers	71
FIGURE 3. Sendai-derived iPSCs demonstrate functional pluripotency through tri-lineage differentiation	73
FIGURE 4. Method of generating iPSCs with as little as 50 µl of non-mobilized blood using the sendai virus.....	75
SUPPLEMENTARY FIGURE 1. Immunofluorescence staining of iPSCs for Nanog .	76
SUPPLEMENTARY FIGURE 2. Immunofluorescence staining of iPSCs for Oct4	77

SUPPLEMENTARY FIGURE 3. SSEA3 plots of PBMC-derived iPSCs	78
TABLE 1. Previously used methods for cellular reprogramming	79
TABLE 3. Factors used for reprogramming	81
TABLE 3. qRT-PCR primers.....	82
 CHAPTER 3	
FIGURE 1. Using CHIR over BIO for GSk3 β antagonism improves yield of CD34 ⁺ cells and viability	116
FIGURE 2. Addition of common endothelial growth factors improves differentiation efficiency and expression of mature endothelial markers by day 5	117
FIGURE 3. Endothelial cells generated with new optimized protocol are functionally mature	118
TABLE 1. Summary of reported endothelial differentiation protocols efficiencies	120
TABLE 2. Endothelial-specific markers used for isolating and validating function of hPSC-derived endothelial cells	122
TABLE 3. qRT-PCR primers.....	123
TABLE 4. Optimization of Lian, <i>et al.</i> protocol parameters and output	124

CHAPTER 4

FIGURE 1. Nonsense-mediated decay recognition of PTC-containing mRNA transcript	172
FIGURE 2. DHX34 and UPF1 KO pools do not have altered levels of DHX34 protein or UPF1 phosphorylation despite differences in mRNA expression	174
FIGURE 3. Knocking out NMD components in HUVECs alters normal endothelial function	176
FIGURE 4. Differences in NO signalling in DHX34 KO HUVECs is not a result of differences in eNOS	178
FIGURE 5. Identification of rare variants in DHX34 in the DECODE study.....	180
FIGURE 6. Rare variants lower DHX34 mRNA levels in iPSCs without impacting differentiation capacity	181
FIGURE 7. Rare variants in <i>DHX34</i> decrease DHX34 protein levels without impacting NMD in endothelial cells	182
FIGURE 8. Patient endothelial cells with DHX34 variants have decreased migratory but increased adhesive capacities.....	184
FIGURE 9. Altered NO production of patient-endothelial cells is not a result of differences in eNOS	186
SUPPLEMENTARY FIGURE 1. Synthego analysis of the DHX34 KO pool	188
SUPPLEMENTARY FIGURE 2. Synthego analysis of the UPF1 KO pool.....	190

SUPPLEMENTARY FIGURE 3. Treatment with TNF α results in altered NMD signalling in the DHX34 KO pool	192
SUPPLEMENTARY FIGURE 4. NO production in patient-derived endothelial cells increases following TNF α treatment	193
SUPPLEMENTARY FIGURE 5. Western blots with HUVEC KO pool samples	194
SUPPLEMENTARY FIGURE 6. DHX34 western blots with iPSC-derived endothelial cells	196
SUPPLEMENTARY FIGURE 7. Phosphorylated UPF1 western blots with iPSC-derived endothelial cells	197
SUPPLEMENTARY FIGURE 8. Phosphorylated eNOS western blots with iPSC-derived endothelial cells	198
TABLE 1. Association of NMD with disease	199
TABLE 2. qRT-PCR primers	200
TABLE 3. Western blot antibodies	201
TABLE 4. Clinical presentation of patients	201
TABLE 5. Sanger sequencing primers	201

LIST OF ABBREVIATIONS

AcLDL	Acetylated low-density lipoprotein
ACS	Acute coronary syndrome
ANG1	Angiopoietin 1
ASC	Adult stem cell
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BAEC	Bovine arterial endothelial cell
BMP	Bone morphogenic protein
BOEC	Blood outgrowth endothelial cell
CAD	Coronary artery disease
CRISPR	Clustered regularly interspaced short palindromic sequences
CTD	C-terminal domain
CVD	Cardiovascular disease
DAF-FM	4-amino-5-methylamino-2',7'-difluorescein
DECID	Decay-inducing complex
DECODE	Genetic determinants of early-onset coronary artery disease
DHX34	DExH-Box helicase 34
EB	Embryoid body
ECM	Extracellular matrix
EJC	Exon junction complex
ELISA	Enzyme-linked immunosorbent assay
eNOS	Endothelial nitric oxide synthase
EOCAD	Early-onset coronary artery disease

EPC	Endothelial progenitor cell
EPS	Extreme phenotype sampling
eRF	Eukaryotic release factor
ESC	Embryonic stem cell (h – denotes human, m – denotes mouse)
FACS	Fluorescence associated cell sorting
FGF	Fibroblast growth factor
FH	Familial hypercholesterolemia
FLT-3L	Fms-like tyrosine kinase 3 ligand
FLT4	Fms related tyrosine kinase 4
gDNA	Genomic deoxyribonucleic acid
GWAS	Genome wide association study
HDAC	Histone deacetylase
HGF	Hepatocyte growth factor
HHEX	Hematopoietically expressed homeobox
HIF1 α	Hypoxia-inducible factor 1 α
HUVEC	Human umbilical vein endothelial cell
ICAM	Intercellular adhesion molecule
ICM	Inner cell mass
IG	Immunoglobulin
I κ B α	Inhibitor of nuclear factor kappa B
IL	Interleukin
iMEF	Irradiated mouse embryonic fibroblast
iNOS	Inducible nitric oxide synthase

IP	Immunoprecipitation
iPSC	Induced pluripotent stem cell
IVF	In vitro fertilization
KO	Knockout
KO-SR	Knockout serum replacement
LDL	Low-density lipoprotein
LIF	Leukemia inhibitory factor
MACS	Magnetic associated cell sorting
MAF	Minor allele frequency
MCP-1	Monocyte chemoattractant protein-1
MEF	Mouse embryonic fibroblast
MEF-CM	Mouse embryonic fibroblast-cultured media
MESA	Multi-ethnic study of atherosclerosis
MET	Mesenchymal to epithelial transition
MI	Myocardial infarction
MMP	Matrix metalloproteinase
mtDNA	Mitochondrial deoxyribonucleic acid
Nf- κ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NMD	Nonsense-mediated mRNA decay
nNOS	Neuronal nitric oxide synthase
NO	Nitric oxide
NSTEMI	Non-ST elevated myocardial infarction
NT-ESC	Nuclear transfer embryonic stem cell

OB	Oligonucleotide/oligosaccharide binding
OSKM	Oct4, Sox2, Klf4, cMyc
OxLDL	Oxidized low-density lipoprotein
PBMC	Peripheral blood mononuclear cell
PDGF	Platelet-derived growth factor
PGI ₂	Prostacyclin
PI3K	Phosphoinositide 3-kinase
PLGF	Placental growth factor
PROX1	Prospero homeobox 1
PRS	Polygenic risk score
PSC	Pluripotent stem cell
PTC	Pre-termination codon
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RecA1	Recombinase-like A1
ROS	Reactive oxygen species
SCCRI	Stem cell and cancer research institute
SCF	Stem cell factor
SCNT	Somatic cell nuclear transfer
SCOC	Stem cell oversight committee
SMC	Smooth muscle cell
SNP	Single nucleotide polymorphism
SSEA	Stage specific embryonic antigen
STEMI	ST elevated myocardial infarction

SURF	SMG1, UPF1, eRF1, eRF3
TCR	T-cell receptor
TGF β	Transforming growth factor β
TIMP	Tissue inhibitor of matrix metalloproteinase
TNF α	Tumour necrosis factor α
UPF	Up-frameshift suppressor
UTR	Untranslated region
VCAM	Vascular cell adhesion molecule
VEGF	Vascular endothelial growth factor
VEGFR2	Vascular endothelial growth factor receptor 2
vWF	von Willebrand factor
WT	Wildtype
7AAD	7-aminoactinomycin D

CHAPTER 1: INTRODUCTION

1.1 INDUCED PLURIPOTENT STEM CELLS

The single dizygotic cell that is formed through fertilization has the capacity to form all cells of both the embryonic and extra-embryonic tissues and is therefore termed totipotent¹. The single cell undergoes several rounds of division at which point the cells begin to spatially segregate to form the blastocyst. During the formation of the blastocyst, the cells differentiate to form an outer cell layer called the trophoblast, and the inner cell mass (ICM). The cells of the trophoblast become the extra-embryonic tissue such as the placenta¹, whereas the cells of the ICM become the embryo itself. The cells of the ICM therefore must form all cell types of the organism and as such are considered pluripotent stem cells (PSCs) (**Figure 1**).

PSCs are defined by two main features, namely the ability to proliferate without differentiation (self-renewal), as well as the ability to differentiate into the three germ layers (from which fully specialized cells arise i.e. ectoderm, mesoderm, and endoderm)². The ability of PSCs to form all mature cell types of an organism is their most attractive quality for regenerative medicine purposes, while their self-renewal capacity means that a single source of PSCs could be used for multiple applications. Possessing these abilities paves the way for them to be used for the growth and regeneration of tissues and organs.

The first type of PSCs that were studied were isolated from teratocarcinoma³. These tumours spontaneously form and include cell types and tissues across all three germ layers,

which suggested the presence of PSCs. In order to study these cells and their capabilities in the context of development, the cells from the ICM of a blastocyst were then isolated. To differentiate these PSCs from those derived from a teratocarcinoma, they became known as embryonic stem cells (ESCs). The first were isolated from mice (mESCs) in 1981^{4,5}. mESCs served as an extremely useful tool for setting standards for PSCs and in studying developmental pathways. mESCs can be implanted into a host blastocyst to form chimeric mice, which can be subsequently used to study developmental pathways as well as those established postnatally or in response to a stimulus^{6,7}. mESCs continue to be an extremely useful tool for researchers, especially those studying aspects of developmental biology that are not accessible for the human system.

Human ESCs (hESCs) were isolated and grown in culture 17 years after the first mESCs were collected². hESCs demonstrate the same self-renewal and differentiation capacity as mESCs, although the differentiation capacity and ability of hESCs to contribute to a chimeric animal is not testable for ethical reasons². Instead, the trilineage differentiation capacity of hESCs can be tested through a teratoma assay, whereby hESCs are injected into an immunodeficient mouse. This assay recapitulates the environment in which PSCs were first identified and allows the hESCs to spontaneously differentiate to cells of all three germ layers, functionally validating their pluripotent status^{1,2}. The use of hESCs is important to better understand human developmental biology as the majority of our understanding was based on other animal systems or sectioned embryos. Although useful, many of the early structures in embryo development differ greatly between humans and mice, which are the most widely used animal system for embryogenesis work.

Recapitulating findings from mESC work to the human system is important. Species differences likely account for many of the differences between mESCs and hESCs such as differences in surface marker expression and dependency on leukemia inhibitory factor (LIF). mESCs consistently express stage specific antigen embryonic antigen (SSEA)-1 but not SSEA-3; however, the opposite is true for hESCs, suggesting differences between the developmental pathway at an early stage². mESCs maintain their undifferentiated status when grown in the presence of LIF and spontaneously differentiate when LIF is removed from the culture medium⁸. This again does not hold true for hESCs as differentiation can occur in the presence or absence of LIF, suggesting a difference in the pathway that maintains their undifferentiated state². hESCs also differ from mESCs in their methylation status and state of X inactivation^{1,9}. mESCs tend to have a lower overall level of DNA methylation and have both X chromosomes active (in female samples) and are therefore considered ‘naïve’^{1,9}. hESCs, however, tend to have increased DNA methylation in comparison to mESCs and some of them have already gone through X-inactivation^{1,9}. This ‘primed’ status of hESCs is more comparable to the PSCs derived from the mouse epiblast rather than mESCs from the blastocyst¹⁰⁻¹². Again, these basic differences highlight the disparity between the murine and human model systems. Despite this, the value of mESCs is still high as they can be readily manipulated and used for chimeric studies which hESCs cannot. Studying selected cell types from hESCs, however, can be incredibly useful in studying direct developmental pathways as well as the mature cell type that is obtained at the end. As early as 1998 when they were first isolated, Thompson, *et al.* speculated on their use for regenerative medicine especially diseases where one single cell type is

affected². In this way, once the hESCs were isolated, optimizing high efficiency differentiation protocols became the next hurdle for stem cell biologists.

The potential use of hESCs in basic developmental research, drug screening, and regenerative medicine applications is immeasurable; however, ethical restrictions greatly limit their use. hESCs are obtained from the cleavage stage of human embryos that were frozen for in vitro fertilization (IVF)². The rules around obtaining these embryos differ across the world but in many countries the decision to donate the embryo must be made following the decision to destroy it. Embryo destruction is a by-product of hESC isolation and is a major ethical consideration and limitation in their derivation and/or use in many countries¹³. For many people, particularly those with religious affiliations, destruction of the embryo is akin to murder as they believe that because of its innate potential to form a full organism, a pre-implantation embryo is considered a person with rights. Several research advisory board reports have ruled that this reasoning does not prohibit the use of hESCs but strict rules, guidelines, and limitations still exist around their derivation and use¹³. In Canada, the *Assisted Human Reproduction Act* prohibits the generation of human embryos for strictly research purposes¹⁴. As such, the embryos from which hESCs are obtained must originally have been made for reproductive purposes. If the embryos are no longer needed, they may be donated for research purposes including the derivation of hESCs¹⁵. Explicit consent is required by the embryo donors, including individual gamete donors if they are not the same as the embryo donors, with full knowledge of all options available. The donors must re-consent at the time of hESC retrieval as there could be a considerable time delay from the original time of consent¹⁵. All proposals to isolate new hESC lines are reviewed by the Stem Cell Oversight Committee (SCOC), a division of the

Canadian Institutes of Health Research (CIHR) that was created in 2003 because of the complex ethical issues that exist around the use of hPSCs. SCOC maintains an electronically accessible registry of all hESC lines derived in Canada and is responsible for reviewing all research applications that involve their use. In situations where the proposed research involves the derivation of new hESC lines or the use of lines that have not been previously approved, a face to face review is required. Research that involves previously approved lines or their derivatives still require review; however, depending on the scope of the proposed work, this could vary from a teleconference review to a written notice to SCOC¹⁵. Irrespective of the standing legislation around their use, the ethical debate surrounding hESCs may ultimately limit their utility in regenerative medicine once the hurdles of differentiation and immune compatibility with the host are overcome. Therefore, an alternative source of hPSCs whose derivation and use is less controversial would mitigate the limitations of hESCs.

For a long time, it was believed that once a cell was specialized, its fate could not be altered. Gurdon disproved this theory when he initially demonstrated the ability to reprogram a somatic cell using somatic cell nuclear transfer (SCNT) in 1958¹⁶. In his seminal work, Gurdon transplanted the genetic information from the somatic cell of a *Xenopus laevis* into a host *Xenopus* oocyte and in doing so, created a PSC-like cell. These cells displayed the same key features of mESCs suggesting that the once somatic cell was now pluripotent in nature. This method was similarly replicated by other groups, but it wasn't until 2006 that reprogramming of somatic cells became prevalent.

In 2006, Yamanaka showed that with the introduction of four transcription factors, a mouse somatic cell could be reverted back to a pluripotent state¹⁷ (**Figure 1**). These cells,

termed induced pluripotent stem cells (iPSCs), possessed the self-renewal and differentiation capacity of ESCs without the ethical considerations of their derivation. One year later, he showed that the same four factors could be used to reprogram human adult fibroblasts thereby revolutionizing the landscape of human stem cell research¹⁸. The use of iPSCs is still heavily monitored by numerous governing agencies but their use exists in less of an ethical grey area than hESCs. In Canada, the use of iPSCs also falls under the jurisdiction of SCOC similar to hESCs; however, derivation of these lines does not necessitate a review unless grafting experiments outside of teratoma assays are proposed¹⁵.

iPSCs can now be generated through a plethora of techniques and somatic cell sources with factors other than those originally used by Yamanaka (see Chapter 2.1). Regardless of the method used to obtain them, their utility in research and their potential for clinical applications remains unmatched. iPSCs can be made from somatic cells easily obtained from willing donors or even frozen samples. This limits the toll on the donors and makes it more realistic to obtain sufficient iPSCs for research.

iPSCs have utility in many aspects of research including disease modeling, drug discovery, and potential avenues for regenerative medicine. To date, many hESC lines have been generated with abnormal karyotypes or mutations associated with known disease phenotypes and therefore could be used to model those specific diseases¹⁹. This is limited by the number of embryos obtained with those mutations and whether they would be donated for research purposes. With iPSC technology, cells with the same self-renewal and differentiation capacities as hESCs could be generated with little to no inconvenience to the patients. iPSCs can then be differentiated towards a specific cell type of interest and used with drug screening platforms or in disease modeling systems. Their use in treatment

discovery has been highlighted by the repurposing of drugs for diseases such as ALS that have bypassed animal models and gone to clinical trials²⁰⁻²². Drug screening will also be improved with iPSCs in using them to screen for cardiotoxicity by testing them *in vitro* on cardiomyocytes²³. Because the low drug attrition rates of today are largely due to unpredicted cardiotoxic effects, this will likely expedite the clinical trial process through better initial screening parameters and outcomes. Finally, one of the hurdles associated with hESCs in regenerative medicine applications is immune compatibility. Although the field has not reached the point of using iPSC-derived cells in the clinic, this issue could be theoretically overcome with the generation of autologous iPSC-derived cells for transplantation.

The application in basic scientific research as well as in the clinic of iPSCs is undeniable. As new approaches to their generation, culturing, and differentiation are optimized, the field becomes closer to realizing their full potential in personalized and regenerative medicine.

1.1.1 Mechanism of cellular reprogramming

Since the seminal work of Yamanaka, many other groups have developed alternative methods and factors to reprogram a somatic cell into an iPSC (discussed further in Chapter 2.1); however, the original factors used in the 2006 paper remain the most commonly used. In this paper, he showed that Oct4, Sox2, Klf4, and cMyc (OSKM), now termed the Yamanaka factors, were sufficient to reprogram somatic cells¹⁷. At the time, the method in which these 4 transcription factors achieved this was not understood. Through

the work of many groups, we now have a better understanding of the cellular mechanism of reprogramming.

In general, reprogramming of somatic cells to iPSCs is an inefficient process. This is most likely due to the number of epigenetic hurdles the cells have to overcome in a short amount of time. It is now suspected that even though the reprogramming process may be initiated in all of the infected cells, very few are able to acquire the necessary changes and pass them on to progeny^{24–26}. In the cells that are successfully reprogrammed, it is thought that there is a hierarchical set of steps that allow for the once somatic cells to gain a true pluripotent status²⁷. Most of the work to elucidate these steps has been done with mouse embryonic fibroblasts (MEFs) infected with the Yamanaka factors. As such, the process described below will differ depending on the factors used, the mechanism through which they are introduced, and the starting somatic cell type, but if successful, will ultimately result in the repression of lineage-specific markers with the simultaneous adoption of a pluripotent signature.

On a large-scale, the general steps that must be undertaken for MEFs to become iPSCs include increasing proliferation while also decreasing MEF-specific transcription, undergo mesenchymal to epithelial transition (MET), and finally activate pluripotency genes^{25,28–30}. Once inside the cell, the Yamanaka factors work to achieve these steps cooperatively and each step can individually impact the reprogramming process.

cMyc is a proto-oncogene that is often constitutively expressed in cancers. In the context of reprogramming it has been shown by multiple groups to be dispensable^{31,32}; however, it consistently increases the efficiency of the process^{9,33}. Once inside the cell, the targets of cMyc include those that regulate proliferation, cell metabolism, and biosynthetic

pathways²⁵, but cMyc also amplifies the transcription of the targets of the other factors^{32,34}. In MEFs specifically, cMyc has also been shown to decrease transforming growth factor β (TGF β) signaling which in turn increases MET and therefore reprogramming itself²⁹.

Oct4 and Sox2 form a heterodimer that regulates many aspects of pluripotency. Together with Nanog, Oct4 and Sox2 maintain pluripotency through sustaining self-renewal and preventing differentiation of the iPSCs once they are generated³⁵⁻⁴¹. At the onset of reprogramming, Oct4 and Sox2 have many of the same transcriptional targets as Klf4 that regulate the organization of a pluripotency network that stabilizes the transition through binding to their own promoters and creating a positive feedback loop^{40,41}.

One of the most important aspects of generating a stable pluripotency transcriptional profile is the loss of repressive chromatin marks on pluripotency genes that are found in somatic cells. During reprogramming, there is an overall shift in the euchromatin status of the cell, an area of extensive research³³. At early stages of the process, Oct4, Sox2 and Klf4 appear to act as pioneer factors of highly methylated distal promoter regions before the proximal promoters are open⁴². In a recent study, Schwarz *et al.* performed an assay for transposase-accessible chromatin using sequencing (ATAC-seq) to better understand how this process differs between the populations of MEFs that can and cannot reach the final pluripotent stage⁴³. They confirmed that the change in epigenetic status occurred quickly in all of the infected cells suggesting that the pioneer factor role of Oct4, Sox2, and Klf4 causes chromatin remodelling regardless of the iPSC fate of the cell. At a slightly later stage (d6 vs d3) there were distinct differences that separated those that were poised to reach iPSCs from those that were not⁴³. In these cells, there was a clear early wave of demethylation prior to the de novo methylation of lineage-specific regions,

suggesting that not only are these processes distinct in reprogramming, but de novo methylation is not necessarily vital for the initial silencing of the MEF-gene signature. Often these demethylated regions were in Oct4, Sox2, and Klf4 binding sites. They were also able to identify a more global state of hyperaccessibility in the cells that were destined to be iPSCs, that was specifically enriched in super-enhancer regions of pluripotency regulators, such as Nanog⁴³. Through this process, they concluded that adopting the overall chromatin status of an ESC during reprogramming is a product of accessible Oct4, Sox2, and Klf4 binding regions along with super-enhancers of core pluripotency factors. Once the pluripotency network is established, the chromatin and transcriptional profile of the cell is more similar to that of an ESC than the somatic cell it once was.

All aspects of the reprogramming process can ultimately impact the quality and type of iPSC that is generated at the end. The stoichiometry and method of introduction of the reprogramming factors can greatly impact the reprogramming trajectory as well as the iPSCs generated⁴⁴. Research has shown that large differences between iPSCs and ESCs can be attributed to residual expression of the reprogramming factors²⁴. There is a level of autoregulation for this as prolonged overexpression of certain factors, such as Oct4, can actually be detrimental to the cell⁴⁵. The importance of chromatin remodelling has been widely accepted for a long time and has led to the concept of iPSCs maintaining a form of epigenetic memory. In this theory, although the iPSC has obtained the overall epigenetic status of an ESC, not all of the regions responsible for the fate of the somatic cell are hypermethylated²⁵. As such, the iPSCs retain a ‘memory’ of their past fate and therefore makes it easier to differentiate them into that lineage^{46,47}. There has been a lot of work done that both supports and refutes this theory. Kim, *et al.* showed that the blood cell-derived

iPSCs had an improved differentiation efficiency towards cells of the blood lineage whereas the fibroblast-derived iPSCs displayed improved osteogenic differentiation capacity⁴⁶. This suggested that because the processes of demethylation and de novo methylation are slow, iPSCs retain a memory of their prior cell fate. However, this was done in low passage iPSCs and it has since been suggested that through additional *in vitro* culturing, this epigenetic memory is erased⁴⁷. In 2014, Rouhani, *et al.* demonstrated that multiple iPSC lines made from different cell types from one individual were more similar to each other than iPSCs generated from the same cell type but from different individuals⁴⁸. This suggests that genetic background, not epigenetic memory, plays a larger role in the iPSC's phenotype. Therefore, although it may not be of great concern for iPSCs that have undergone *in vitro* culture for some time, epigenetic memory may impact the differentiation capacity at early time points. Because extended culturing of iPSCs can impact the rate of chromosomal mutations that occur, extended culture periods must also be avoided in the absence of routine genetic screening.

Understanding the important role of each of the reprogramming steps has allowed for the identification of ways to increase the efficiency of the process. Considering the important role of MET, by using cells that already are epithelial-like, such as keratinocytes^{49,50}, or by increasing MET through TGF β inhibition²⁹, reprogramming efficiency can improve the process that OSKM initiate. Additionally, the importance of the epigenetic changes in the overall reprogramming mechanism can be seen through increases in reprogramming efficiency when chromatin remodelers or inhibitors of methyltransferases are used^{9,51–54}.

Another aspect of the somatic cell's function that alters during reprogramming is its metabolism. Similar to the Warburg effect described in cancerous cells, a metabolic switch occurs during reprogramming whereby ATP production through glycolysis is favoured over mitochondrial oxidation⁵⁵⁻⁵⁷. This provides yet another aspect of the reprogramming mechanism that can be chemically modulated to improve efficiency⁵⁵.

In just over 10 years since reprogramming was first accomplished, there have been many advancements in understanding the mechanism of cellular reprogramming and how it can be manipulated to increase efficiency. As bioinformatic and cellular technology improves, our knowledge about this process and how to best use it will improve not only our ability to generate iPSCs, but also the quality of the lines that are generated.

1.1.2 Use of iPSCs in disease modelling

As mentioned previously, one of the greatest advantages about iPSC generation is the potential for creating disease models, which was first done only 1 year after reprogramming was demonstrated in human cells⁵⁸. When embryos are created for IVF purposes, genetic screening can be done, especially for couples who are known carriers for inherited diseases. Should the screen turn up with a known mutation and the couple ultimately decides to donate that embryo to research, hESC lines for specific diseases can be generated. Although the resulting model would be similar to one created by iPSCs, there are a number of critical differences. Creating a model from iPSCs where the actual phenotype of the patient is known adds a level of complexity that would be impossible for one made by hESCs. Reprogramming technology also allows groups to create iPSC lines

from numerous controls, such as unaffected family members, which otherwise would not be possible in the context of hESCs. By creating control lines from unaffected members, the genetic background of the individual outside of the suspected disease-causing mutation can be more accurately accounted for. This again adds to the precision of the model as compared to one generated by hESCs because the specific role of the mutation would be better elucidated.

The utility of iPSCs in disease modelling has continued to expand with gene editing advancements. Clustered regularly interspaced short palindromic repeats/Cas9 (CRISPR/Cas9) technology has improved our ability to generate disease models in an unprecedented way. CRISPR/Cas9 is an easy way to design and create specific knockouts or knock-ins in the genome. This has been particularly important for elucidating the role of novel variants in disease. When a mutation is suspected to be involved in a disease, the gold-standard to prove its association with the observed phenotype would be through the introduction of the mutation into an unaffected control. Should the control develop the same disease phenotype as the patient, then an association between the mutation and the phenotype could be made with a high level of certainty. CRISPR/Cas9 can be used to create isogenic iPSC lines whereby the only difference in the lines is the mutation of interest. This method provides clarity on the role of the variant by removing other variables associated with the genetic background of the patient, similar to the use of unaffected family members as controls. This technique is especially important in genetically complex disorders to fully elucidate the underlying mechanism⁵⁹. Applying CRISPR/Cas9 to iPSCs specifically can also reduce the variability that would come from having to perform multiple transfections in terminally differentiated cells of interest. Currently, CRISPR/Cas9 editing is possible

but remains inefficient; however, new methods involving Cas9 variants and different modes of introduction continue to improve this efficiency^{60–63}. In 2018, Wen, *et al.* were able to improve upon previous methods to simultaneously reprogram and edit human fibroblasts into iPSCs^{64,65}. In their work, they reprogrammed PBMCs and were able to decrease the time in generating CRISPR/Cas9 editing in new iPSC lines from 3-4 months down to 1-2 by coupling the reprogramming and gene editing events⁶⁶. By transfecting peripheral blood mononuclear cells with episomal vectors carrying reprogramming factors and Cas9, along with a guide RNA, they were able to achieve high levels of homology-directed repair to introduce a specific knock-in in a shortened time frame in their iPSCs. Although it worked, the reprogramming efficiency was lower in the presence of gene editing and worked best only in certain loci of the genome that are more amenable to editing, therefore limiting its use for other regions⁶⁶. Nonetheless, this work stands as proof-of-principle that combining the two processes for generating disease models using both iPSCs and CRISPR/Cas9 may not only be feasible but more time-efficient.

To date, iPSCs have been used to generate hundreds of disease models especially those with known genetic components⁵⁹. The success of these models ultimately depends on the ability to differentiate the iPSCs to the affected cell type(s). Although it can be difficult in some cases, many differentiation protocols have been established. Using iPSCs to model diseases that include the cooperation of multiple cell types makes disease modelling not only possible but easier. From a single iPSC line, multiple terminally differentiated cells can be generated thereby reducing the requirement for obtaining different commercially available lines. This could be particularly advantageous when generating isogenic iPSC lines with CRISPR/Cas9. Differentiating multiple cell lineages

from one modified isogenic iPSC line would allow for a clear distinction of the variant across cellular compartments of the disease. This would be particularly useful in disease such as coronary artery disease, discussed below, where multiple cell types play a role and interact during disease development.

Using iPSC lines for disease modeling reduces the extent to which patient samples are needed to generate a multi-cellular disease model. As culturing techniques improve and begin to evolve through the addition of biomechanical engineering methodologies to incorporate cellular scaffolds to create 3D systems, better iPSC disease models can be generated that include structural and multi cellular components from a single iPSC line.

1.2 CORONARY ARTERY DISEASE

The utility of iPSCs in disease modelling is exemplified in those that involve multiple cell types such as coronary artery disease (CAD), which is a major contributing factor to global deaths associated with ischemic heart disease. Ischemic heart disease has been the leading cause of global mortality for the last 15 years⁶⁷. In 2016, it was attributed to the death of 9.4 million individuals which totals ~17% of all global deaths. This number has increased from just over 7 million (13% of all deaths) in 2010⁶⁷. The increased incidence of heart disease may be attributed in part to the increase in some of its risk factors such as diabetes and obesity. Nonetheless, heart disease remains a large burden on our healthcare system as well as a major determinant of life expectancy.

Ischemic heart disease is a broad category but more often than not is a result of CAD⁶⁷. CAD is an insidious disease, characterized by the formation of an atherosclerotic

plaque in the coronary arteries. These plaques grow slowly over the course of an individual's life and typically result in clinical presentation in the 6th or 7th decade of life. Patients with CAD can remain largely asymptomatic for the duration of the disease. As such, individuals typically learn of their CAD burden during an acute coronary syndrome (ACS) - angina, myocardial infarction, or even sudden cardiac death. CAD affects both sexes ; however, women show symptoms roughly 10 years later than men and ACSs are four times more frequent in men than in women⁶⁸.

CAD can be hard to detect in the general population because of its asymptomatic nature. Measuring blood lipid levels, blood pressure, and blood glucose are some techniques used by physicians to estimate a patient's risk of CAD. In patients that are judged to be at an intermediate risk of having an ACS, it has been suggested that measuring levels of C-reactive protein may represent a measure of the low-grade inflammatory aspect of CAD⁶⁹. More invasive techniques such as angiograms, intravascular ultrasound, or near-infrared spectroscopy may also be used to determine the burden of CAD in high-risk individuals^{70,71}. Lipid-lowering medication, such as statins, is one of the primary methods employed by physicians for patients with CAD, along with promoting lifestyle changes^{69,72}. If patients are found to have CAD following, or during, an ACS, more advanced measures may be taken such as an angioplasty or even bypass surgery depending on the severity of the disease.

Research around the development and treatment of CAD has been ongoing for many decades. The research has shown the importance of the risk factors in disease development, as well as the step-wise nature of plaque formation. Despite all of the advancements in both the pharmacological and surgical treatment of the disease as well as

information available about its prevention, the disease remains highly prevalent. Therefore, continued research into disease pathophysiology is necessary.

1.2.1 CAD development

CAD is characterized by the build-up of atherosclerotic plaques in the coronary arteries (**Figure 2**). These plaques grow slowly over many decades and may result in clinical presentation in the form of an ACS or may only be noted upon autopsy. The complexity of CAD presentation is mirrored in its development; disease development requires the coordination of many cell types (**Table 1**). The atherosclerotic plaques develop in the sub-endothelial compartment of the arteries, known as the intima, and are typically composed of a lipid pool, immune cells, smooth muscle cells, and often contains a necrotic core. As the plaque grows, blood flow to the surrounding tissue is compromised; however, the vessel wall remodels outward to compensate for plaque growth and to reduce the obstruction of blood flow. Outward remodelling will prevent vessel lumen thinning up to 40% of the plaque burden⁷³. At this point, the plaque begins to grow into the arterial passage and decrease the flow of oxygenated blood to the myocardium.

The prevailing ideology on CAD initiation is the “response to injury” model refined by Ross^{74,75}. He proposed that the development of CAD is innocent in its beginnings, but chronic injury turns the repair mechanism designed to respond to an insult into a disease itself⁷⁶. It is widely accepted that CAD initiation results from chronic minimal injury to the endothelium^{70,76–78}. The insult is typically a result of the environmental risk factors discussed below, but ultimately lead to a change in the function and phenotype of the

endothelial layer. The endothelial layer becomes more permeable, allowing for increased migration of low-density lipoprotein (LDL) particles from circulation into the intima⁷⁸. Once in the intima, the LDL can be oxidized by the endothelial cells forming oxLDL, which furthers endothelial dysfunction itself⁷⁹.

Endothelial dysfunction can encompass the loss of any of the normal functions of the endothelium but is often hallmarked by a decrease in nitric oxide (NO) bioavailability^{70,78–81}. NO is constitutively produced by endothelial nitric oxide synthase (eNOS), although its activity is regulated through phosphorylation in response to certain stimuli such as changes in shear force⁸². During the initial phases of CAD development, oxLDL or reactive oxygen species (ROS) that are present can sequester the available NO or uncouple the reaction by which eNOS produces NO^{81,83–85}. Uncoupling this reaction results in the further generation of ROS creating a feedback loop that continues to decrease available NO. NO is the most potent vasodilator but in endothelial cells it also plays an important role in the inhibition of the nuclear factor kappa-light-chain-enhancer of activated B cells (Nf- κ B). In its inactive state, Nf- κ B is in a complex with its inhibitor (I κ B α) and NO acts to stabilize this interaction⁸¹. Therefore, when NO is decreased, there is increased activity of this pathway, which results in transcription of adhesion markers, like vascular cell adhesion molecule (VCAM) and intercellular adhesion molecule (ICAM), as well as an increase in cytokine production⁸⁶. Both of these factors further contribute to endothelial dysfunction and result in a more activated endothelial state. Increased adhesion marker expression and cytokine production, especially monocyte chemoattractant protein-1 (MCP-1), recruits monocytes from circulation to the site of injury where they can adhere to the endothelium^{81,87,88}.

The adherent monocytes migrate across the endothelium and enter the intimal region. Once in the intima, monocytes differentiate into macrophages which are able to endocytose the oxLDL particles through their scavenger receptors^{89,90}. Macrophages hydrolyse oxLDL into free cholesterol and free fatty acids, with any excess cholesterol being re-esterified⁹¹. Through this process, the macrophages become laden with lipids and are termed foam cells. Generation and accumulation of foam cells in the intima forms the first visible plaque in CAD – the fatty streak. Fatty streaks can be seen as early as infancy or childhood but are capable of regression at this early stage if the insult to the endothelium is transient^{77,92}. This concept is the basis of the response to injury hypothesis – the injury to the endothelium increases intimal LDL which results in changes in endothelial function to recruit monocytes for its removal. Under CAD-promoting conditions, the injury to the endothelium is chronic and therefore, rather than rebounding from this dysfunctional state, the process continues.

As the insult to the endothelium continues, the extent of LDL accumulation increases thereby constantly recruiting monocytes from circulation and increasing the number of foam cells in the intima. The endothelial cells and foam cells secrete cytokines that continue the process of plaque development. Some of these cytokines, such as interleukin-1 (IL-1) and platelet-derived growth factor (PDGF)⁸⁷, promote the proliferation of smooth muscle cells and their migration from the media to the intima. During this process, it has been shown that the smooth muscle cells lose their contractile phenotype to gain a more synthetic phenotype whereby their ability to respond to stimuli for dilation or contraction is lost and instead, they begin to produce and excrete more growth factors and cytokines⁷⁹.

Smooth muscle cells play an important protective role in the development of CAD. Once in the intima, they migrate over the fatty streak and reside below the endothelial monolayer. Their adopted synthetic phenotype allows for increased production of an extracellular matrix (ECM) composed of collagen, proteoglycans, elastin, and glycoproteins⁷⁷. Together with the smooth muscle cells themselves, the ECM generates a fibrous protective cap over the developing plaque. As the plaque thickens, the fibrous cap can become vascularized which can play a critical role in terminal thrombosis events as discussed below⁷⁹.

This general process continues over the course of the disease with a growing number of monocytes and their derivatives (macrophages and foam cells), along with other immune cells, namely T cells, accumulating in the intima⁷⁶. Neovascularization of the inner core and the fibrous cap generates additional avenues by which inflammatory cells can enter the intima^{79,83,93,94}. Along with immune cells, the plaque also grows due to the deposition and release of more free cholesterol from macrophages processing the oxLDL as well as from those undergoing apoptosis⁹⁰. With many cells undergoing apoptosis in the growing plaque, efferocytosis may become disrupted and therefore a necrotic core can form alongside the lipid pool⁹⁵. The prothrombotic and lipid core creates a more complex inner plaque milieu that is highly pro-thrombotic.

Exposure of the prothrombotic core to the circulating blood can occur as a result of three types of events – (1) rupture of the fibrous cap, (2) erosion of the endothelium, and (3) calcified nodules⁹⁶. Calcified nodules account for only a small fraction of thrombosis events (2-7%) associated with CAD and as such, will not be covered here.

As previously mentioned, a large component of the fibrous cap is the ECM produced by the smooth muscle cells. The other cell types involved in plaque development, namely endothelial cells, monocytes, and macrophages, all produce matrix metalloproteinases (MMPs), which can degrade ECM components⁹⁷⁻¹⁰⁰. Under normal circumstances, the activity of MMPs is countered by the activity of tissue inhibitor of metalloproteinases (TIMPs)¹⁰¹. Together, MMPs and TIMPs are necessary for cellular turnover during normal growth and repair; however, in proatherogenic conditions, their activity is no longer balanced, resulting in weakening of the cap¹⁰¹. The aforementioned cells all can produce and activate different MMPs which work to degrade the ECM of the fibrous cap, thus weakening it. The cap is then susceptible to rupture due to nonlaminar and/or high-pressure flow of the blood^{90,102}. Exposure of the core to the circulation, along with the entry of additional platelets from the neovessels of the cap itself can cause a hemorrhage or thrombosis event^{77,90,94}.

Erosion of the endothelial layer accounts for 30-35% of terminal thrombosis events, and even more (>80%) in women under 50¹⁰³. The endothelium grows as a strict monolayer and as such, the cells are strictly regulated through contact inhibition^{79,104}. In areas over a developing plaque, it is likely that the endothelium is denuded from the disturbed blood flow through the artery⁷⁹. In this case, the endothelial cells on either side of the wound stretch out to fill the space and subsequently replicate to re-establish the barrier. Under proatherogenic conditions, it is likely that the injury occurs repeatedly in a small area. In this case, the cells on either side of the wound will have responded by undergoing replication enough times to become senescent. Although the remaining endothelial cells of the barrier may not have reached this critical point, because of contact inhibition they are

unable to migrate and close the gap, therefore leaving the area denuded of the endothelial barrier⁷⁹. Removal of the endothelium increases the exposure of procoagulant surfaces promoting the development of a thrombus.

Rupture of the fibrous cap or removal of the endothelial layer both result in the exposure of the inner milieu of the plaque to the circulation. Monocytes may have surface expression of tissue factor and their macrophage derivatives produce large amounts of it that pool in the core of the plaque⁶⁹. The decreased NO in areas of CAD also result in an endothelium that is more prone to platelet aggregation, and together with the increased collagen in the ECM produced by smooth muscle cells, provide a scaffold for thrombus formation⁸². Through either the interaction with tissue factor or collagen, platelets become activated¹⁰⁵. Activated platelets aggregate together and recruit activated, and inactivated platelets to the growing thrombus. The number of platelets that aggregate, and therefore the size of the thrombus itself, is a result of numerous factors but the amount of tissue factor sequestered in the plaque has been shown to be a major determinant of thrombus size^{106,107}. The important role of platelets in CAD has also been demonstrated through numerous animal studies that showed a lower atherosclerotic burden and decreased thrombotic events in animals with platelet count defects either as a result of von Willebrand disease or thrombocytopenia^{108,109}.

The extent of thrombosis along with the size of the underlying plaque dictates the clinical presentation, if any, of the event. In many cases, a thrombotic event will be clinically silent and will contribute to the progression of the disease⁷⁷. In fact, Davies found that in 17% of patients that died from causes other than an ACS, there was evidence of plaques that had undergone rupture, thrombus formation, resealing, and continued growth

without clinical manifestation¹¹⁰. The majority of plaques that result in an ACS upon rupture are small (stenosis of <75%) whereas larger, more stable plaques tend to result in either a small or silent event¹¹¹. Smaller plaques tend to be more lipid rich and have no, or a very small, fibrous cap making them more prone to rupture^{77,83}. In these cases, removal of the endothelium can result in immediate exposure of a procoagulant surface and result in thrombosis. In contrast, larger plaques tend to be more stable because of the presence of a thick, fibrous cap. It is also possible that these do not result in clinical presentation because of well-developed collateral arteries – natural bypass arteries to circumvent the obstruction⁷⁷.

CAD development is complex and requires the cooperation of many cell types. Currently, the main therapeutic strategies used in clinic aim to decrease the lipid burden through the use of statins or prevent thrombosis events by using anticoagulants or blood thinners in high-risk patients^{69,72,112,113}. The number of other cell types and pathways involved in the initiation and development stages of the disease may provide a platform for the development of novel pharmacological preventative or therapeutic measures.

1.2.2 Risk factors for CAD development

Genetic and environmental factors contribute roughly equally to CAD development risk in the general population. Both aspects can largely dictate the age at which an individual can present with CAD-associated symptoms, if at all. In this section, I will provide an overview of the environmental components that are well established, whereas the genetic aspect of CAD development will be reviewed in more detail in Chapter 4.1.2.

There are several major environmental risk factors for CAD development that are well validated and documented including (1) age, (2) hypercholesterolemia, (2) hypertension, (4) smoking, (5) diabetes mellitus, (6) obesity¹¹⁴. As was mentioned before, CAD develops slowly over the course of an individual's life, typically only clinically manifesting in later decades. As such, age itself remains a large risk factor for developing CAD.

As was previously discussed, CAD development largely relies on the accumulation of LDL molecules in the arterial wall. Patients that have hypercholesterolemia typically have increased levels of circulating LDL thereby increasing the risk of this aspect of the disease. Hypercholesterolemia may be a result of multiple genetic mutations, monogenic disease as is seen in individuals with familial hypercholesterolemia (FH), or may be a result of dietary choices. Patients with FH typically have a mutation in the LDL receptor, although mutations in other genes such as *ApoB* and *PCSK9* are also known to result in higher LDL cholesterol^{115–117}. For patients with FH as well as those with increased LDL levels as a result of diet, lifestyle changes and the use of lipid-lowering drugs, such as statins, are a major component of CAD prevention^{72,102}.

Hypertension results in changes to the normal laminar flow in the coronary arteries, thereby disturbing the repair process of the vessel wall¹⁰². The disturbed flow may also contribute to priming of the vasculature for CAD development (discussed in Chapter 4.1.1) or may result in fissuring of a developed atherosclerotic plaque⁷⁷. Hypertensive individuals may have increased levels of Angiotensin II and/or decreased NO which alters the surface of the endothelium such that it becomes more adhesive to circulating leukocytes thereby promoting their entrance to the intima^{102,118}.

Cigarette smoke contains many toxins that impact the vessel wall through mechanism that remain largely unknown¹⁰²; however, smokers have signs of platelet and endothelial dysfunction, decreased levels of anti-thrombotic factors, and higher overall inflammation¹¹⁹. Cigarette smoke increases ROS and alters endothelial permeability, both of which are indicators of endothelial dysfunction¹¹⁹. Additionally, it has been suggested that nicotine, which is present in cigarettes at high levels, can decrease the vasodilator prostacyclin which contributes to decreasing platelet aggregation¹¹⁸. The addictive nature of cigarettes means that these impacts are a chronic stressor on the vasculature, which distinguishes the normal repair process of the vessel wall from disease development. Despite not knowing the full details of the mechanism, the association between smoking and increased CAD development, and acute thrombosis events is well-documented⁹⁰. The smoking rate in Canada is decreasing (17.7% in 2015 to 16.2% in 2017)¹²⁰ but still remains a big contributor to overall health and CAD risk, which can be decreased by reducing or eliminating the number of cigarettes smoked.

The way in which diabetes mellitus and obesity contribute to CAD development is thought to be similar in nature to hypercholesterolemia. Individuals with diabetes mellitus and obesity typically have hyperglycemia, hypercholesterolemia, as well as higher levels of circulating triglycerides. Hyperglycemia can cause similar changes to the vasculature as hypercholesterolemia and increased triglycerides¹²¹. Combined with the chronic nature of these metabolic syndromes, a pro-atherosclerotic environment is fostered within the vasculature of these patients. For many of these patients, lifestyle changes that include decreasing dietary fats and sugars as well as increasing activity levels will contribute to a decrease in their CAD risk.

For some of the environmental risk factors mentioned above, the exact mechanisms remain elusive; however, their association with CAD development is well documented. As such, lifestyle changes to eliminate or lessen these risks remain a large part of CAD prevention and therapeutic strategies used today.

1.2.3 Current models and their limitations

Animal models are commonly used to study CAD and have been instrumental in our understanding of the disease. These systems typically require genetic manipulation of the lipid processing pathway to induce hyperlipidemia. This has been done most prevalently in mice through the creation of the *apoE*^{-/-} and *Ldlr*^{-/-} lines which effectively mimic human FH^{122,123}. This phenomenon is recapitulated in other animal systems such as the Watanabe rabbit; however, in the Watanabe rabbits, the mutation that effectively eliminates LDL binding to its receptor occurred naturally¹²⁴. These, along with other animal models (**Table 2**) have improved our understanding of disease etiology but are not without limitations¹²³.

Murine models of atherosclerosis are the most widely used because of the relatively low financial cost and the ease of manipulating the mouse genome¹²². In these systems, the mice are genetically modified to have dysfunctional cholesterol metabolism and are then put on a high fat diet for several weeks to induce atherosclerotic plaque development. Under these conditions, atherosclerotic plaques develop in the mice along a similar trajectory as in humans; however, the extent of hypercholesterolemia is higher than what is typically seen in the clinic. The other limitation with this experimental set-up is that the

plaque growth rate is significantly faster than in humans and therefore does not accurately depict the chronic nature of the disease¹⁰². Additionally, although certain murine models, such as *hdl^{-/-}* mice, have been shown to develop atherosclerosis in the coronary arteries, in the more commonly used systems, plaques tend to develop in the aorta and proximal great vessels as opposed to the coronary, carotid, and cerebral arteries as they do in humans¹²⁵. The smaller arteries in which the plaques develop in mice are structurally and hemodynamically different than the larger arteries that are more typically affected in humans. The murine atherosclerotic plaques also have a decreased propensity to rupture, which accounts for a large percentage of human clinical manifestations and therefore cannot recapitulate the terminal stages encountered by many patients^{102,126}.

The differences in the cellular subtypes in mice and humans contributes to the difficulty in translating murine models directly to the human system. Immune cells play a large and important role in CAD development in both systems; however, the specific roles of the immune cell types in humans and mice may differ¹²⁷. For example, the distinction between T_H1 and T_H2 is not as distinct as it is in humans, and neither are the markers of macrophage activation¹⁰². Both of these cell types are found in high concentrations in an atherosclerotic plaque thus making it important to better understand the differences in the two systems. Another cell type whose role may be inadequately represented in mouse models are smooth muscle cells. In the arteries affected in mice, the embryonic origin of the smooth muscle cells differs from the origin of those in the affected vessels of humans¹²⁸. In human coronary arteries, a population of smooth muscle cells can be found in the intimal region early in life which is lacking in the mouse system. Therefore, their role in the earlier stages

of the disease may be missed by using a mouse model that does not accurately depict their location pre-CAD induction¹⁰².

Discrepancies such as those mentioned above suggest that the etiology of the disease in humans and mice may differ. Despite this, using animal models provides an excellent tool for studying aspects of the disease that would otherwise be extremely difficult to model *in vitro*. By using a live animal model, the role of each individual cell type involved can be examined and potentially manipulated allowing researchers to gain better insight into their specific role. This may be improved for areas such as the immune aspect of the disease by using a humanized mouse model to determine if the role of the human immune cells differs from that of the mouse.

In order to better understand the human aspect of the disease, different methods have been employed. One of the earliest methods was to examine coronary arteries post-mortem^{92,110,111,129}. These autopsies provided invaluable information about plaque structure at different time points and highlighted their complexity. The amount of information that could be obtained through this method is limited to static views of the plaque. CAD development involves the dynamic interplay of many cells which is missed through the examination of post-mortem tissues. With the ability to culture human cells *in vitro*, groups started to use specific cellular subtypes involved in the disease to better understand their role. In fact, in 1984 Ross isolated smooth muscle cells directly from an atherosclerotic plaque and tried to culture them¹³⁰. Unfortunately, this method did not work as the cells were mostly senescent and thus did not provide an accurate model. Since then, different groups have used commercially available cells, such as human umbilical vein endothelial

cells (HUVECs), to better understand the individual role of these cells in disease development.

Researchers continue to use both animal models and more simplistic *in vitro* cellular platforms to understand aspects of CAD development and potentially identify new therapeutic avenues. With the advancement of stem cell technology, many groups have moved towards generating patient-specific cellular models. In 2010, Rashid *et al.* used this method to generate iPSCs from patients with FH which will help to advance our knowledge of lipid metabolism in these patients and potentially identify other cell types that may have been thus far overlooked¹³¹. This progression of the *in vitro* system allows for groups to examine the genetic component of the disease which is largely missed in using commercially available cells or animal model systems. Forming organoids that include multiple cell types and subsequently using them in xenograft models, as was recently done by Wimmer, *et al.* will combine the benefits of both animal models and *in vitro* human cell culture systems in studying CAD¹³².

1.3 SUMMARY OF INTENT

The overall objective of my PhD work was to create and use patient-specific endothelial cells to assess the impact of rare variants in individuals with early-onset coronary artery disease (EOCAD). The work presented here was done in collaboration with Dr. Guillaume Paré who spearheaded the Genetic Determinants of Early-Onset Coronary Artery Disease (DECODE) study. As mentioned in the subsequent chapters, the patient

samples that were necessary for the creation of the model were obtained through their enrolment into the DECODE study at the Hamilton General Hospital.

The first goal of my thesis was to reprogram patient samples into iPSCs. At the time that work began on this project, reprogramming of non-mobilized peripheral blood samples was not done in the Stem Cell and Cancer Research Institute (SCCRI). I was able to successfully reprogram numerous samples into iPSCs using the Sendai virus. The samples were collected at the Hamilton General Hospital following patient enrolment by a member of Dr. Paré's team. The isolation of the peripheral blood mononuclear cells and the subsequent reprogramming and validation of the iPSC lines was done by me. The iPSC lines were all generated as per the protocol outlined in Chapter 2 and were used for all work done in Chapters 3 and 4.

The endothelial cells are the first to experience disturbance during CAD development and therefore are of particular interest especially in patients who develop the disease several decades earlier than the average. The patient-specific iPSCs were differentiated to endothelial cells through a protocol that I optimized and validated across multiple iPSC lines. The optimized version of the protocol was able to achieve differentiation efficiencies up to 10-fold greater than the originally published protocol and generated more endothelial cells expressing mature markers. The protocol shown in Chapter 3 was used for all subsequent work in Chapter 4. This work was included as a figure in a review paper submission to *Stem Cell and Development* in August 2018 which I wrote as a means to outline an optimization process that can be applied to published differentiation protocols.

The ultimate goal of this work was to use both the iPSCs generated in Chapter 2 and the endothelial cells from Chapter 3 to determine if rare variants found in the DECODE population altered endothelial function. Through analysis of the exome sequencing data, 4 patients with variants in *DHX34* were selected for analysis. *DHX34* was recently identified as a component of the nonsense mediated mRNA decay (NMD) pathway but has not been widely studied in this context or in a cell-specific manner. To better understand the role of *DHX34* in endothelial cells, *DHX34* HUVEC knockout cell pools were purchased from Synthego as a basis for comparison to our patient-derived cells. In these cells and our patient-derived endothelial cells, we noticed several differences in endothelial functions related to CAD development such as migration, NO production, and monocyte adhesion. Although aspects of NMD show disruption, further research is required to determine which NMD targets may be responsible for the observed phenotype, as well as what other potential roles *DHX34* may have in endothelial cells that are independent of its role in NMD. The work presented here was done primarily by myself with the assistance of undergrad students (Nicole Yan, Shreya Jain) who helped with select cellular assays but mostly in performing western blots and RNA isolation for the qRT-PCRs.

The three main aims for my thesis were to (1) generate patient-specific iPSC lines from peripheral blood mononuclear cells, (2) optimize an endothelial differentiation protocol, and (3) to better understand the role of *DHX34* in endothelial function and how its variants may impact the development of CAD.

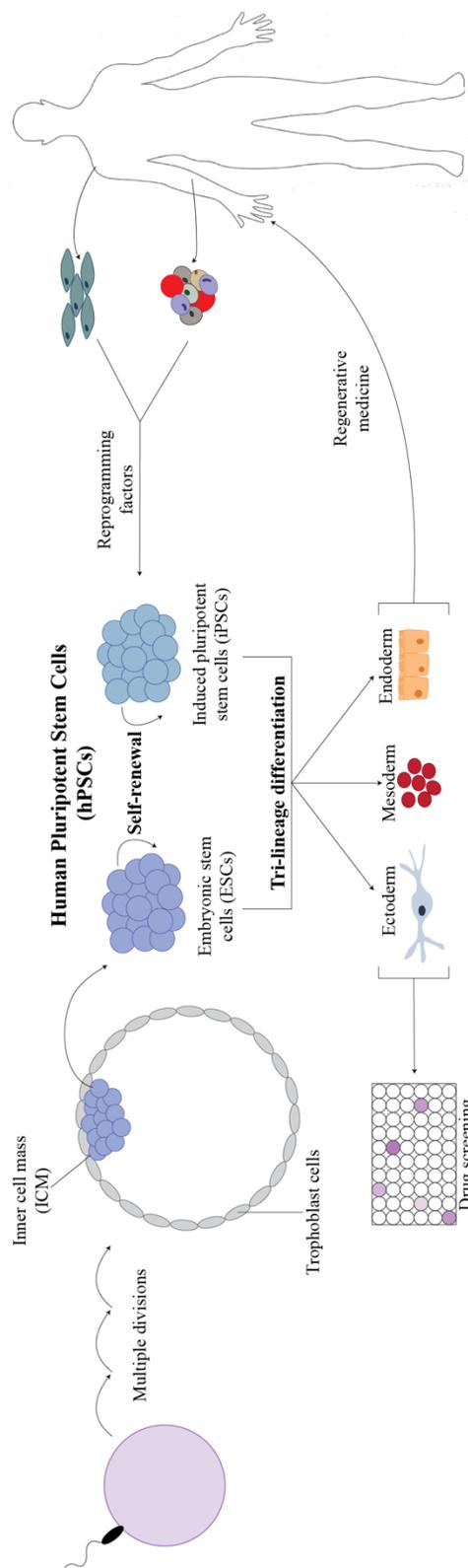


Figure 1. Common sources of human pluripotent stem cells and their potential applications. Following fertilization, the single dizygotic cell undergoes multiple rounds of division and begins to spatially segregate during blastocyst formation. At this point, the cells differentiate to those of the trophoblast lineage and the inner cell mass (ICM). The cells of the ICM are termed embryonic stem cells (ESCs). The ethical concerns surrounding ESC derivation and use have made it imperative to find alternative sources of human pluripotent stem cells (hPSCs). Yamanaka was the first to demonstrate that somatic cells, such as fibroblasts and mononuclear cells from the blood, can be reprogrammed into an induced pluripotent stem cell (iPSC) with the introduction of reprogramming factors. Both sources of hPSCs are capable of self-renewal as well as tri-lineage differentiation. The resulting differentiated cells from hPSC sources have potential applications in many aspects of biomedical research including drug screening. They also hold potential for future clinical applications in regenerative medicine with terminally differentiated cells or tissues from autologous iPSCs being used for transplant.

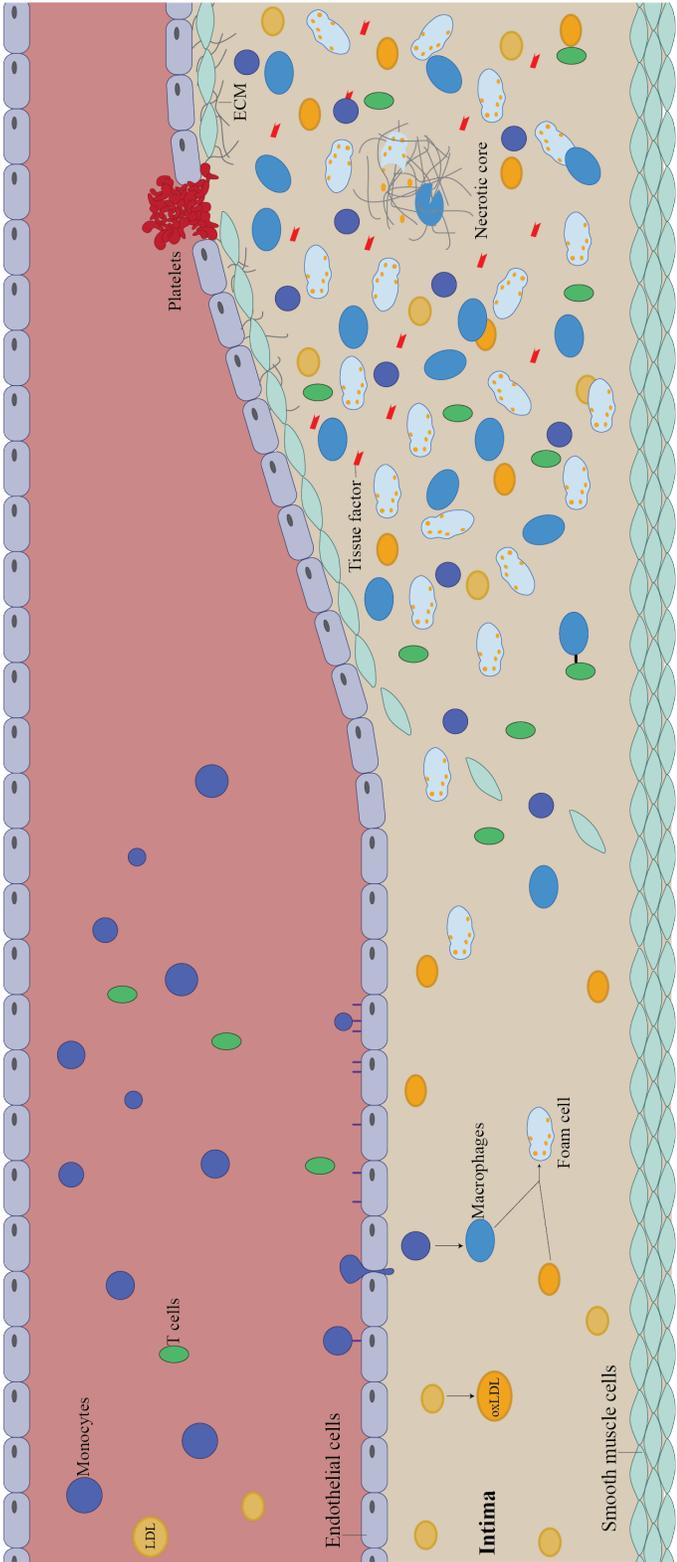


Figure 2. Development and growth of an atherosclerotic plaque in coronary artery disease. The development of coronary artery disease (CAD) involves the cooperation of many cell types over the course of several decades. Injury to the endothelium is widely accepted as an initiator of disease development. When the endothelium is injured, it becomes more permeable allowing for increased migration of LDL particles to the intima. In the intima, the LDL becomes modified to oxidized LDL (oxLDL). This advances injury to the endothelium, resulting in increased expression of adhesion markers and cytokines by the endothelium. This action recruits monocytes from the circulation to the site of injury, where they then transmigrate across the endothelium. In the intima, the monocytes differentiate to macrophages, which can engulf oxLDL particles thereby clearing it from the intima. In doing this, macrophages become laden with lipid and are termed foam cells. As this process continues, foam cells, their precursors, and other immune cells such as T cells continue to pool in the intima. As they accumulate, they continue to produce cytokines, some of which can promote the migration of smooth muscle cells from the media to the intima. The smooth muscle cells form a fibrous cap under the endothelium where they also secrete large amounts of an extracellular matrix (ECM) to further strengthen the cap. The high number of cells undergoing apoptosis in the inner core may impact the process of efferocytosis thereby creating a necrotic core in the plaque. This, along with the large amounts of tissue factor produced and sequestered in the core, creates a highly prothrombotic environment. If the inner core is exposed to the circulation as a result of endothelial erosion or rupture of the fibrous cap, platelets can adhere to the components of the ECM or cells that are expressing tissue factor themselves, to initiate the coagulation cascade, resulting in thrombus formation.

Table 1. The role of cellular subtypes in CAD

Cell type	Roles in CAD
Endothelial	<p>Dysfunction is considered initiator of CAD development</p> <p>Semi-permeable barrier becomes disrupted to allow increased LDL migration</p> <p>NO production decreases</p> <p>Surface becomes adhesive to both leukocytes and platelets</p> <p>Production of cytokines recruits immune cells</p> <p>Production/activation of MMPs can degrade fibrous cap</p> <p>Vascularize plaque and fibrous cap increasing immune cell invasion and risk of hemorrhage</p> <p>Denudation can cause terminal thrombosis events</p>
Smooth muscle	<p>Migrate to intima to form protective fibrous cap</p> <p>Phenotype changes from contractile to synthetic to increase ECM and cytokine production</p>
Monocytes/macrophages	<p>Monocytes differentiate into macrophages once in the intima</p> <p>Act as antigen-presenting cells to T cells</p> <p>Macrophages engulf oxLDL and become foam cells</p> <p>Apoptosis largely produces lipid core</p> <p>Production of cytokines recruits immune cells</p> <p>Production/activation of MMPs can degrade fibrous cap</p> <p>Produce large amounts of tissue factor rendering the core pro-thrombotic</p>
T cells	<p>Contribute to cell activation through the production of cytokines</p>
Platelets	<p>Initiate coagulation cascade through tissue factor interaction</p> <p>Produce thrombus responsible for ACS presentation</p>

Table 2. Animal models of atherosclerosis development¹²²

Animal system	Method of atherosclerosis induction*	Similarities with human disease	Differences with human disease
Mouse	LDLR ^{-/-} apoE ^{-/-}	Cellular subtypes involved Lesions in human FH cases are similar	Plaque location Complexity of plaque Time for disease development
Rabbit	New Zealand white - HFD Watanabe – natural 4bp deletion in LDLR	CAD development is spontaneous but accelerated by HFD Plaque location	Watanabe plaque composition is different to human
Pig	High fat diet apoB100 variants	CAD development is spontaneous but accelerated by HFD Complexity of plaque Plaque location	Different HDL subtypes

* Genetic mutations typically result in hypercholesterolemia so most accurately would reflect human familial hypercholesterolemia-induced CAD development; HFD = high-fat diet

CHAPTER 2: DERIVATION OF AN EFFICIENT PROTOCOL FOR REPROGRAMMING PBMCs INTO iPSCs

PREFACE

This chapter focuses on the reprogramming method I used to convert peripheral blood mononuclear cells into iPSCs for use in future work. Although more common place now, when I began this project, reprogramming of non-mobilized peripheral blood was not widely done as traditional methods were not very efficient. I was the first in the SCCRI to successfully convert non-mobilized PBMCs into iPSCs using the Sendai virus. This chapter begins with a more in-depth introduction into the field of reprogramming (Chapter 2.2) to supplement the overview given in Chapter 1.1. The methods highlighted in this chapter were used to generate all iPSCs used in both Chapter 3 and Chapter 4. All the data generated for this chapter was done by me with blood samples either collected at the Hamilton General Hospital as part of the DECODE study or PBMCs purchased from StemCell Technologies for the Healthy control. In all graphs, every data point indicates a separate replicate. For the qRT-PCR data, each replicate shown is the mean of three technical replicates. The data from the Genetic Analysis Kit is shown as the mean of three technical replicates from one passage of the indicated iPSC lines.

2.1 INTRODUCTION

2.1.1 History of iPSC generation

Reprogramming technology has allowed researchers to overcome the ethical debate surrounding the use of human embryonic stem cells (hESCs) and has significantly increased the possibilities for the regenerative and personalized medicine fields. Using patient-derived cell types allows for drug toxicity screening prior to clinical trial use, thereby providing a possible means to decrease the drug attrition rates^{22,23}.

In 2006, Yamanaka made his seminal discovery of cellular reprogramming for which he was awarded the Nobel Prize¹⁷. Yamanaka showed that through the introduction of just four factors, Oct4, Sox2, Klf4, and cMyc (OSKM), a somatic cell could be transformed into an ESC-like cell termed an induced pluripotent stem cell (iPSC). These factors, now known as the Yamanaka factors, were identified by screening 24 genes and were initially introduced via retrovirus based on the Moloney murine leukemia virus¹⁷. Over the course of 2 weeks, the murine fibroblasts transitioned into pluripotent stem cells whereby they were capable of both self-renewal and tri-lineage differentiation. This work was initially done in murine fibroblasts but in 2007, Yamanaka's group showed that the same four factors could be used on human fibroblasts, both adult and fetal, to similarly revert the somatic cell into an iPSC¹⁸. This discovery paved the way for patient-specific lines for use in disease modelling, which was first demonstrated only one year later by George Daley's group⁵⁸. They showed that patient-specific iPSC lines could be generated for use in disease investigation and drug development from genetic diseases with either Mendelian or

complex inheritance patterns. Since then, many disease-specific iPSC lines have been generated through a wide range of methods⁵⁹.

2.1.2 Methods previously used for reprogramming

It was originally believed that a somatic cell's lineage fate was final; however, Gurdon illustrated the plasticity of a somatic cell through his work. In 1958, Gurdon showed that through the transfer of a somatic cell's genetic information into an enucleated oocyte, the fate of the cell was reverted to an ESC-like state¹⁶. This original method of reprogramming, termed somatic cell nuclear transfer (SCNT), is achieved through undefined factors present in the oocyte. The pluripotent stem cells created through SCNT are similar to primed mouse ESCs (mESCs) in nature. The process of reprogramming through SCNT is fast – in under 22 hours, the cell is functionally pluripotent, and is more similar to an ESC than an iPSC generated through exogenous gene expression¹³³. Although SCNT can be a useful tool for reprogramming non-primate mammalian cells, it was only in 2013 that the technique was successfully used to reprogram human somatic cells using fetal fibroblasts, with up to 50% of resulting embryos giving rise to nuclear transfer ESCs (NT-ESCs)¹³⁴. The NT-ESCs were pluripotent as assessed functionally and molecularly. The method outlined by Tachibana, *et al.* is useful for proteomic and transcriptomic analysis of the creation of the pluripotency identity during reprogramming.

Though now possible, the feasibility of using SCNT for deriving hPSCs for regenerative medicine is uncertain. The efficiency of human oocyte SCNT reprogramming differs significantly across donors and the factors impacting this have not yet been

elucidated¹³⁴. Additionally, the work done in 2013 that demonstrated successful human SCNT did so with fetal fibroblasts. Whether the efficiency would be similar for adult somatic cells remains to be tested. SCNT requires that the donor undergo ovarian stimulation and oocyte retrieval which can take anywhere from 8-14 days¹³⁵. This is an invasive procedure that requires the donors to adhere to a stimulation protocol beforehand in order to obtain an adequate number of oocytes. Adherence to the protocol and the steps required to complete it may reduce the number of potential oocyte donors. Once the somatic cell is reprogrammed, only trace amounts of the donor mitochondrial DNA (mtDNA) can be detected in the NT-ESCs¹³⁴. This could be an advantage from the regenerative medicine standpoint in cases of mitochondrial disease, whereby replacing the diseased mitochondria in regenerated cells could be done with a healthy donor oocyte¹³⁵. Maintaining the oocyte mitochondria also ensures metabolically active hPSCs; however, for disease modelling purposes, having mtDNA from the donor oocyte may conceal metabolic phenotypes of the disease. Ultimately, in Canada the creation of hPSCs through SCNT is not permitted by the policy that covers the use of human biological materials, which is enforced and regulated by SCOC¹⁵. As such, alternative methods for deriving human PSCs that circumvent the ethical issues of both hESCs and those reprogrammed through SCNT are necessary.

In 2007, Yamanaka's group filled the need for a source of hPSCs through their work showing retroviral reprogramming with OSKM resulting in 0.02% of human fibroblasts being reprogrammed into iPSCs¹⁸. Since then, many groups have worked towards improving efficiency of the reprogramming process in order to make it a more feasible component of regenerative medicine. One way that this has been accomplished is

by using alternate methods of introducing reprogramming factors (**Table 1**). Around the same time as Yamanaka's report on human fibroblast reprogramming, the Thompson group achieved a similar feat in fetal and neonatal fibroblasts, although with reduced efficiency (0.0095% vs 0.02%)¹³⁶. They showed that using Lin28 and Nanog in place of cMyc and Klf4 iPSCs could be generated, thus highlighting the potential for alternative reprogramming factors (**Table 2**).

It is important to confirm the efficacy of reprogramming genes or compounds in human cells if the experiments were initially done in the mouse system. Yamanaka's group initially reprogrammed murine fibroblasts in 2006 and showed in 2007 that the same four factors could be used for human fibroblasts without a significant decrease in efficiency^{17,18}. For example, when reprogramming murine somatic cells, Oct4 can be replaced by number of different factors including those that activate Oct4 target genes^{137,138}, inducing epigenetic changes at the *Oct4* locus^{139,140}, over-expression of *Sall4* or inducing *Sall4* expression with DZNep¹⁴¹, increasing MET with E-cadherin or TGFβ^{142,143}, or replacing Oct4 with endoderm specifiers that block ectoderm¹⁴⁴. When tried in human somatic cell reprogramming, the only methods that could replace Oct4 were introduction of GATA3-VP16¹⁴⁵ or mir302/367 and mir200c/302/369¹⁴⁶, albeit with a significant reduction in efficiency. In 2018, Mai, *et al.* used ATAC-seq, to identify the tumour suppressor NKX3-1 as having a role in Oct4-mediated reprogramming¹⁴⁷. Through identification of the IL6R-STAT3-NKX3-1-Oct4 signaling cascade, they were able to show that modulating this pathway upstream of Oct4 itself could result in similar levels of reprogramming with NKX3-1 in place of Oct4 in the reprogramming cocktail in both human and murine settings¹⁴⁷. Because human somatic cell reprogramming is necessary for regenerative

medicine and drug testing purposes, confirming murine methods in the human system is of the utmost importance.

Prolonged exposure of reprogramming factors is necessary to allow for chromatin modifications to occur such that endogenous expression of pluripotency genes can be turned on and the state maintained by cellular machinery⁹. To achieve this, early cellular reprogramming work that followed from Yamanka's discovery required transgene insertion for successful iPSC generation. Retrovirus was the first to be used for this process and although successful, yielded low efficiency^{17,18}. In an attempt to increase the rate of iPSC production, groups tried with lentivirus carrying the reprogramming factors^{136,148}. Similar to retrovirus, lentivirus results in insertion of the transgenes into the host genome; however, unlike retrovirus, lentivirus does not require the target cells to be actively dividing¹⁴⁸.

Lentiviral and retroviral methods were successful in reprogramming, but their potential for regenerative medicine applications is limited because the insertion sites of the transgenes are random and could potentially have unpredicted cellular impacts¹⁴⁹. In the initial examination of human iPSCs, Takahashi, *et al.* noted ~20 retroviral insertions in each line, similar to what was seen in the mouse iPSCs^{17,18}. In the mouse system, the ability of iPSCs to contribute to germline cells can be assessed through the formation of chimeric mice. When this was first done, however, 20% of these iPSC-derived mice, tumour development could be attributed in part to reactivation of cMyc¹⁵⁰. That same year, the therapeutic potential of iPSCs was shown with a humanized mouse model of sickle cell anemia, and although none of the mice in that study developed tumours, the authors noted that the potential remained and must be addressed before moving to human clinical

applications¹⁵¹. Therefore, it is likely that similar levels of transgene insertions in the human system pose the same risk of tumorigenesis. To eliminate these risks, integration-free methods are now commonly used for reprogramming¹⁴⁸. Many different methods have been used, including the use of transposons to insert transgenes which were subsequently removed following endogenous activation of pluripotency genes¹⁵², chemical modulators^{9,33,141,153}, adenovirus¹⁵⁴, and episomal vectors^{155–157}. All these methods have proved successful with varying levels of efficiency, as well as safety. For example, using transposons may reduce the risk of oncogene reactivation but activation of transposons themselves has the potential for multiple insertions and deletions to occur, thereby increasing the risk of indel formation^{148,152}. As we continue to understand the steps by which cellular reprogramming occurs, it becomes possible to use alternative and safer methods by which reprogramming can occur.

2.1.3 Somatic cell sources for reprogramming

When reprogramming, there are many factors to consider when selecting the starting somatic cell. The invasiveness, cost, and ease of the technique may limit, or in some cases prevent, patients from donating samples. When creating iPSCs for disease modelling or drug screening, it is essential that the genetic component of the disease remains intact, which is not only a consideration when selecting the reprogramming method but also in cell source selection. Additionally, it has been suggested that the starting cell may even impact the propensity of the cells to form teratomas, which may be even more of a consideration for downstream clinical applications¹⁵⁸.

Initial reprogramming experiments were conducted almost exclusively with fibroblasts and they still remain one of the most commonly used sources of somatic cells¹⁴⁸. Fibroblasts can be isolated from patients through a punch biopsy and subsequent culture of the sample. Culturing techniques for fibroblasts are well established, making them relatively easy to culture and expand; however, it can take up to 2-3 weeks for the fibroblasts to be confluent enough for their initial passage from the time of sample collection¹⁵⁹. Additionally, because they are isolated from the skin, there is the potential for additional unknown mutations to be present in the cells. Although this is the case for any cell of the adult body, the dermis is particularly susceptible to random mutations due to exposure to UV radiation, which may in turn cause genetic aberrations that are independent of the disease of the patient¹⁶⁰. Taken together, the length of time for fibroblast expansion *in vitro* and the increase in UV-mediated mutations makes fibroblasts an unattractive cell type in reprogramming, especially for regenerative medicine applications.

The reprogramming process requires that somatic cells take on the epigenetic, transcriptomic, and metabolomic profile of ESCs. Because of this, many people have also tried to use progenitors or adult stem cells (ASCs) for reprogramming purposes¹⁴⁸. These cells maintain many properties of ESCs, such as self-renewal and differentiation, although to a lesser extent. In this way, there would be fewer changes needed to the profiles of progenitors or ASCs, as compared to somatic cells, to revert to a pluripotent state¹⁴⁸. Similarly, some groups focused on reprogramming cell types that have endogenous expression of the Yamanaka factors in order to reduce the requirement for exogenous transgene expression. For example, Kim, *et al.* demonstrated in 2009 that neural stem cells could be reprogrammed with only Oct4 as they already have high endogenous expression

of cMyc and Sox2¹⁶¹. Similarly, Zhu, *et al.* showed that reprogramming efficiency was elevated in keratinocytes which endogenously express both Klf4 and cMyc¹⁶². In fact, many groups have reported that the reprogramming efficiency of ASCs and progenitors is higher than that of a bulk population of somatic cells^{161,163,164}. This supports the idea that these cells are more poised to be reprogrammed and has contributed to the notion that reprogramming somatic cells is inefficient because it is actually the very small progenitor/ASC population present in the bulk sample that is being reprogrammed. This idea was proposed by Yamanka's group in their first demonstration of reprogramming¹⁷. They suggested that because the frequency of multipotent stem cells in skin was similar to their overall reprogramming efficiency it was possible that they were targeting this poised population. They did, however, note that they were able to achieve similar efficiency with stromal cells which were less enriched in a stem-like population, thereby casting doubt on their previous hypothesis. Nonetheless, other groups attempted to reprogram stem and progenitor-like populations. In fact, the first demonstrations of reprogramming cells from the hematopoietic system were those that used CD34⁺ progenitors isolated from mobilized peripheral blood¹⁶⁵, bone marrow¹⁶⁶, or umbilical cord blood¹⁶³. Although progenitors and ASCs can be reprogrammed, the theory that the overall low efficiency of reprogramming was due to a small progenitor population has been refuted by groups that have shown that it is possible to reprogram terminally differentiated cells like islet cells¹⁶⁷ and those of the blood system^{164,168–170}. Instead, it is now believed that low reprogramming efficiency can be attributed to the difficulty of overcoming the epigenetic changes that are required during reprogramming, which also explains the higher efficiency rates in more progenitor-like populations^{25,148} (see Chapter 1.1.1)

Although reprogramming progenitors or ASCs may be efficient, these cells can be rare and therefore difficult to isolate from patients. The least invasive way to obtain human somatic cells is from peripheral blood or umbilical cord donation. Samples of non-mobilized peripheral blood are given regularly by patients and contain anywhere from 10^6 – 10^7 peripheral blood mononuclear cells (PBMCs) in every 10 mL (data collected through this project). On the other hand, although umbilical cord is a fully non-invasive supply of patient PBMCs and CD34⁺ progenitors, it is not available for every patient. When trying to isolate the CD34⁺ progenitors from peripheral blood, a mobilizing agent such as granulocyte-colony stimulating factor is typically given to patients¹⁶⁵. There is a risk of negative side effects, such as prolonged lethargy, bone pain, headaches, and may not be an option for some patients, including those that are pregnant or breastfeeding. Additionally, the cost and time associated with the isolation, expansion, and purification of CD34⁺ progenitors decrease the large-scale application of the technique. As such, using a non-mobilized PBMC source would be an ideal source for reprogramming because of the large number of cells that can be obtained from a small sample, the ease of acquisition, as well as the lower chance of having a chromosomal aberration due to the absence of environmental factors that could confound the genetic aspect of the disease in question.

In 2010, three groups simultaneously showed that human PBMCs could be similarly reprogrammed into iPSCs using lentivirus¹⁶⁸, Sendai virus¹⁶⁹, or polycistronic vectors¹⁷¹ carrying the Yamanaka factors. Working backwards, Loh, *et al.* and Staerk, *et al.* both demonstrated that the iPSCs generated came from a T cell origin, whereas Seki began with an isolated fraction of T cells. Although the methods were different, both techniques again refuted the concept that the progenitor population, specifically in the

hematopoietic compartment, is the one being reprogrammed. Additionally, by demonstrating the reprogramming of a bulk PBMC source, Loh, *et al.* and Staerk, *et al.* both independently confirmed that isolation and subsequent culture of a specific cell type within the population is unnecessary for successful reprogramming^{168,171}. In doing this, they highlighted the potential of PBMCs as a source of somatic cells for reprogramming in the context of disease modelling and regenerative medicine applications.

2.1.4 Chapter objectives

Creating patient-specific disease models is the first step in developing more personalized treatment options for patients. Often, treatments that work for many patients do not work for others. Although diseases may present clinically in a similar manner, the etiology may differ thus altering the efficacy of current treatment modalities. Creating patient-specific models of disease development allows us to investigate the role of their genetics in disease development in a manner more specific than existing mouse models. Reprogramming patient samples into iPSCs is an effective method to produce a large number of target cells with the least inconvenience to the patient. In certain cases, obtaining a skin biopsy, mobilized peripheral blood, or a large volume of non-mobilized peripheral blood is an obstacle because of the health of the patient. As such, the goal of this chapter is to highlight the method of reprogramming a small number of PBMCs isolated from non-mobilized peripheral blood that I developed in the lab. Using this method, as few as 50,000 PBMCs can be used to generate upwards of 20 iPSC colonies for validation and

maintenance. Using this technique, only small volumes of blood are required from patients for researchers to create and bank a large number of iPSCs for future work.

2.2 MATERIALS AND METHODS

2.2.1 Cell culture

Once collected, PBMCs were cultured in StemPro34 filtered following addition of the supplement, 2mM L-glutamine, 20ng/mL IL-3, 20ng/mL IL-6, 100ng/mL SCF, 100ng/mL Flt-3L (complete PBMC media). iPSCs were cultured in DMEM/F12, 10% KOSR, 10mM NEAA, 2mM L-glutamine, 7mM β -mercaptoethanol, 16 μ g/mL bFGF (iPSC media) on iMEFs. Alternatively, iPSCs were cultured on Matrigel-coated plates in MEF-CM with 8 μ g/mL bFGF or mTeSR1 (StemCell Technologies) for RNA or gDNA collection, immunofluorescence, and flow cytometry experiments. HEK cells were cultured in 90% DMEM, 10% FBS, 2mM L-glutamine. For lentiviral production, media was changed to high glucose DMEM, 10% FBS, 2mM L-glutamine, 0.1mM NEAA, 1mM sodium pyruvate (transfection media). All cells were maintained at 37°C and 5% CO₂.

2.2.2 PBMC isolation

10mL of non-mobilized peripheral blood was collected from patients with early-onset coronary artery disease. The blood was kept on ice for a maximum of 2 hours before isolation of the peripheral blood mononuclear cells took place. 10mL of blood was diluted

1:4 in DPBS, 1mM EDTA, 5% FBS (PEF). 8mL of the diluted blood was carefully layered onto 4mL of Ficoll. The tubes were centrifuged at 1500 rpm for 20 min, with acceleration set to 5 and the brake set to 0 on the centrifuge. Following centrifugation, the mononuclear cell layer was carefully removed from each tube into two separate clean tubes. These tubes were centrifuged at 1500 rpm for 5 min (acceleration:5, brake:5). The supernatant was aspirated, and the pellets were resuspended in 10mL cold ammonium chloride. The solutions were kept at 4°C for 5 min and then centrifuged at 1500 rpm for 5 min (acceleration:5, brake:5). The pellets were combined and resuspended in 1-4mL of PEF to count the cells. The isolated PBMCs were frozen down in aliquots of 1 or 2 million cells/vial in 90% FBS, 10% DMSO and kept in liquid nitrogen until future use.

2.2.3 Lentivirus production

Two types of lentivirus were generated for reprogramming purposes – one carrying Oct4/Sox2, and the second with cMyc/Klf4. Lentivirus was produced by transfecting HEK cells with one of the aforementioned vectors (11.5µg) in 1.5mL OptiMEM with 6.5 µg psPax2, 4.5 µg pMD2.6. Separately, 46µL of LTX was added to 1.5mL OptiMEM and left at room temperature for 4 mins. The LTX solution was added dropwise to the vector and incubated at room temperature for 20 min. Media on HEK cells was changed to the transfection media and the DNA:LTX media was added to the flasks and incubated overnight. Media was changed 24 hours later to transfection media and left for 48 hours. The lentivirus was harvested from the supernatant of each flask through ultra-

centrifugation for 2 hours at 25,000 rpm at 4°C. The viral pellet was resuspended in cold DPBS and stored at -80°C until future use.

2.2.4 Lentiviral reprogramming of PBMCs

Lentiviral reprogramming of PBMCs was attempted after 1, 2, and 7 days of culture in complete PBMC media. All the following steps were used for all conditions following the pre-culture period.

100,000 PBMCs were seeded onto Retronectin-coated non-TC 12 well plates in complete PBMC media. 24 hours after seeding, the Oct4/Sox2 and cMyc/Klf4 lentiviruses were added to complete PBMC media with 8µg/mL Polybrene and added to each well of PBMCs. The cells were spin-transduced at 1300 rpm for 90 min at 30°C. 24 hour after transduction, 1mL complete PBMC media was added to each well and left for 48 hours.

48 hours after transduction, PBMCs were transferred onto iMEFs in PBMC media, gradually changing to iPSC media over the next 48 hours. Cells were maintained on iMEFs in iPSC media until colonies appeared up to 2 weeks.

2.2.5 Sendai virus reprogramming of PBMCs

The CytoTune 2.0 kit was purchased from ThermoFisher for reprogramming purposes. The kit is composed of a three-vector system Klf4-Oct4-Sox2 (KOS), cMyc, and Klf4. The kit was initially thawed at 37°C for 30 seconds and then allowed to fully thaw at

room temperature before placing on ice. The contents of the kit were then aliquoted and frozen down for future use.

PBMCs were thawed one or two days prior to the Sendai transduction into complete PBMC media. On the day of transduction, the cells were counted and checked for viability. The volume required to obtain 50,000 or 100,000 cells was added to one well of a 24-well low-attachment plate along with additional complete PBMC media to a final volume of 350 μ L, with or without 4 μ g/mL Polybrene. Sendai virus was added to 150 μ L of complete PBMC media to obtain an MOI of 5, 5, and 3 for KOS, cMyc, and Klf4 respectively. The media with virus was added to the wells to a final volume of 500 μ L. The plate was sealed and centrifuged at 1500rpm for 90 min at room temperature. Following the spin, an additional 500 μ L of complete PBMC media was added to each well before placing the plate at 37°C overnight.

24 hours after transduction, the PBMCs were collected into tubes. Each well was gently washed with 500 μ L of media to ensure all cells were collected. The cells were centrifuged at 200 x g for 10 min to remove additional virus. The supernatant was aspirated, and the cells were resuspended in 500 μ L of complete PBMC media before placing back in the original wells of the 24-well plate.

48 hours after transduction (day 2), irradiated MEFs were plated in 1 well of a 12-well plate per sample transfected. On day 3, the 500 μ L of cell suspension was transferred to 1 well of a 12-well plate with irradiated MEFs containing 500 μ L of complete PBMC media. On day 4, half of the media was removed and was replaced with PBMC media without the cytokines (StemPro34 filtered after addition of the supplement, 2mM L-

glutamine). On day 5, half of the media was again changed to PBMC media without the cytokines. On day 7 onwards, half of the media was changed daily to iPSC media.

Once colonies were large enough (approximately day 15-17), each colony was transferred to 1 well of a 12-well plate coated with irradiated MEFs to maintain clonality of iPSC lines.

2.2.6 Immunofluorescence staining

iPSC lines reprogrammed with Sendai virus were maintained as clonal lines and passaged once every 7 days. At passage 2, a portion of each iPSC clone was passaged into an individual well of a 24-well plate for Tra1-60 staining. After 4 days in culture, the iPSCs were fixed with 4% PFA at room temperature for 10 min. The cells were washed twice with DPBS and then stained with Tra1-60 (R&D) overnight. Cells were visualized the following day and the two most Tra1-60 positive clones were maintained for future experiments.

For Oct4 and Nanog staining, the iPSCs were fixed at passage 15 using 4% PFA. The cells were washed twice with DPBS before permeabilization with methanol. Oct4 and Nanog antibodies (R&D) were added to individual wells overnight, diluted 1:200 and 1:100 in DPBS respectively. The following day, the cells were washed three times with DPBS before adding the secondary antibody (1:1000) at room temperature for 1 hour. The cells were washed again with DPBS before staining with DAPI (300mM) at room temperature for 10 min. The cells were finally washed again before being imaged. Brightness of the

brightfield images was adjusted for better viewing. Merged images were compiled using ImageJ.

2.2.7 Embryoid body differentiation

EB differentiation was performed as previously described¹⁷². In brief, iPSCs were passed using 1:3 diluted Accutase for 3 min at 37°C. The Accutase was carefully removed and replaced with DMEM with 10% KO-SR with 0.44 µL/mL Y-27632 for the first 48 hours. Cells were scraped into rounded clumps using a 5 mL pipette and transferred to a well of an ultra-low attachment plate. Media on the cells was changed every 48 hours. On day 7, the EBs were transferred to a gelatin-coated plate in DMEM + 10% FBS. Media was changed every 48 hours for an additional 7 days. On day 14, cells were scraped up and collected for RNA extraction.

2.2.8 qRT-PCR

iPSCs were collected at passage 13-15 for qRT-PCR purposes for by scraping up the colonies and allowing them to settle before aspirating the media. For EB differentiation, the EB differentiated cells were scraped up and centrifuged before aspirating the supernatant. RNA was extracted (Norgen) and made into cDNA (Bioline) following the protocols provided. Primers used are listed in **Table 3**.

2.2.9 Flow cytometry

All antibodies used for flow cytometry were purchased from BD Biosciences. Flow cytometry was performed on the LSRII and analysis performed with FlowJo.

Isolated PBMCs were initially analyzed through flow cytometry. PBMCs were collected and centrifuged at 1200 rpm for 5 min. Cells were resuspended in DPBS, 1mM EDTA, 5% FBS (PEF) and filtered through a 35µM filter. The following antibodies from BD Biosciences were used: CD45 (1:100), CD34 (1:200), CD3 (1:200), CD4 (1:100), CD2 (1:100), CD19 (1:100), CD20 (1:200), CD33 (1:100), CD14 (1:500), CD13 (1:500), CD15 (1:100). Cells were left to incubate with antibody for 30 min at room temperature before being washed once with PEF. 7AAD (1:50) was added once the cells were resuspended in PEF prior to running the cells on the LSRII.

iPSCs grown in mTeSR1 were used for flow cytometry experiments. In brief, the cells were singularized with Accutase at 37°C for 5 min and collected for centrifugation. The cells were resuspended in PEF and filtered. Cells were stained with SSEA3/4 (1:100) for 20 min at room temperature. Cells were then washed once with PEF before adding 7AAD (1:50) and running the samples on the MacsQuant.

Data collected from both flow cytometers was analyzed and compiled using FlowJo v10.

2.2.10 iPSC genetic analysis

The hPSC Genetic Analysis Kit was purchased from StemCell Technologies and modified for use in a 96-well format as opposed to the suggested 384-well format. Genomic DNA (gDNA) was isolated from the iPSCs using the QIAmp DNA kit (Qiagen) and was stored at -30°C until use.

Briefly, the gDNA and kit contents were thawed on ice protected from light. 580ng of gDNA and the 2x Master Mix were combined with nuclease free water to a final concentration of 5 ng/μL. The primer-probe sequences were resuspended in 33μL TE buffer and used to create primer-probe mixes with double the volume indicated in the table provided with the kit. 14μL of the DNA sample to wells of the 96-well plate followed by 6μL of the appropriate primer-probe mixes. All samples were done in triplicate. The plate was run on the QuantStudio3 machine by Applied Biosystems following the standard cycling parameters provided with the kit without the addition of ROX. Analysis was done with the online tool provided by StemCell Technologies associated with the kit (https://stemcell.shinyapps.io/psc_genetic_analysis_app/). Briefly, ΔC_t was calculated by normalizing each locus from the sample against the internal ch4p control locus (ΔC_{tX}) and for the control sample as well (ΔC_{tC}). These values were used to calculate $\Delta\Delta C_t$ ($\Delta C_{tX} - \Delta C_{tC}$). Copy number (C_n) was calculated by taking two to the negative exponent of $\Delta\Delta C_t$ and multiplying it by two to signify the diploid state ($C_n = 2 \times 2^{-\Delta\Delta C_t}$).

2.2.11 Data Analysis

Data are expressed as mean \pm SD. Analysis was performed using t-tests and two-way ANOVAs as indicated with $p < 0.05$ as statistically significant.

2.3 RESULTS

2.3.1 Lentiviral transduction did not successfully reprogram non-mobilized PBMCs

Introducing reprogramming factors through lentivirus is widely used for human fibroblasts, and some hematopoietic cells. Reprogramming of human blood cells with lentivirus has only been accomplished using isolated populations of cells, namely T-cells¹⁶⁸ or CD34⁺ cells either from cord blood¹⁶³ or mobilized blood¹⁶⁵. Reprogramming PBMCs from non-mobilized peripheral blood using lentivirus has not been previously shown. We attempted to reprogram non-mobilized PBMCs using the same method that had been used in the Stem Cell and Cancer Research Institute (SCCRI) to reprogram CD34⁺ cells from cord blood (**Figure 1a**). The PBMCs were cultured prior to transduction with a combination of two lentiviruses – Oct4/Sox2 and cMyc/Klf4. After 2 weeks in culture on iMEFs, no colonies had appeared. Absence of any Tra1-60⁺ cells confirmed that there were no iPSCs that we had missed through gross examination of the wells (**Figure 1b**). The same method was attempted multiple times with various PBMC samples with the same results.

To confirm that viability of the PBMCs prior to transduction was not the reason we didn't see any iPSC colonies form, we tested 3 lengths of pre-culture. Regardless of the culture length prior to transduction, the lentiviral method did not yield any iPSC colonies from the non-mobilized PBMCs (**Figure 1c**).

2.3.2 Sendai virus successfully reprograms non-mobilized PBMCs

Successful reprogramming of non-mobilized PBMCs without requiring enrichment for a specific cell population has been done with Sendai virus^{169,173}. As such, we tried the CytoTune2.0 kit (ThermoFisher) to reprogram our PBMCs. The protocol that came with the kit suggested pre-culturing the PBMCs for 3 days prior to transduction; however, to increase viability and likelihood of success, we cultured for only 1-2 days (**Figure 1d**). On day 8, small colonies became visible and by day 12, had classic iPSC morphology and stained positive for Tra1-60 (**Figure 1e**) indicating successful reprogramming of the non-mobilized PBMCs with Sendai virus. This method successfully worked for 15 different PBMC samples (**Figure 1g**).

Traditional lentiviral reprogramming requires the use of Polybrene to aid in viral entry into the cells. Polybrene (hexadimethrine bromide) is a cationic polymer that increases retroviral and lentiviral transduction efficiencies by neutralizing the charge differential between the viral particles and sialic acid on the cell membrane. Because Sendai virus enters mammalian cells through binding to the sialic acid surface receptor¹⁷⁴, we wanted to determine whether the addition of Polybrene would increase our yield of PBMCs through Sendai reprogramming. We utilized the same method (**Figure 1d**) with

and without polybrene. There was a dramatic decrease in the number of PBMC colonies that we were able to obtain when we added polybrene to the viral mixture before adding to the well (**Figure 1f**), and as such, continued without the addition of Polybrene.

2.3.3 Reduction in starting number of PBMCs does not decrease efficiency with Sendai virus

One of the most common uses for iPSCs is for disease modelling using patient-specific samples. Depending on the disease, the patients may present clinically in acute distress thereby making it impossible to obtain large blood samples for reprogramming purposes. This may also be the case in situations where the disease of the patient limits their ability to donate in other ways, or simply if they find it intrusive to provide large samples. As such, it's critical to be able to generate iPSCs with limited starting material, especially if additional assays can/must be done on the donated blood sample. To see whether the reprogramming protocol could be used for small sample numbers, we tried both 100,000 and 50,000 PBMCs for reprogramming. The original CytoTune2.0 kit suggests using 300,000 PBMCs/well - therefore we were reducing the number by a factor of 3 and 6 respectively. The same size wells were used for both 100,000 and 50,000 and the cells were all re-plated into 12 well plates (1 well each) coated with iMEFs. When transduced in the presence of Polybrene, we did notice a reduction in the number of colonies; however, in the absence of Polybrene there were >100 colonies in all wells. 100 iPSC colonies from a single donor are more than is needed for research purposes and so we concluded that 50,000 PBMCs were sufficient to obtain iPSCs using Sendai virus.

50,000 PBMCs in the absence of Polybrene at the time of transfection were used for all subsequent experiments

2.3.4 Sendai virus reprogramming of PBMCs is independent of CD34⁺ population size

In one of the first successful demonstrations of reprogramming hematopoietic cells with the Sendai virus, one group demonstrated that the CD34⁺ fraction were the only ones to be successfully infected and reprogrammed¹⁷⁵. Although in other cases of hematopoietic reprogramming this has not held true^{168,169,171}, we were interested to see whether the number of colonies we were able to achieve was dependent on the size of the CD34⁺ progenitor population of the PBMC samples. When the blood samples were obtained for PBMC isolation, no mobilization agent was given to the donors; however, small numbers of CD34⁺ populations can be routinely found in circulation.

We performed flow cytometry on 15 PBMC samples to assess the percentage of CD34⁺ cells that could be found in the non-mobilized blood samples that were isolated. When we then reprogrammed the same 15 samples, we found that there was no correlation between the CD34⁺ population and the number of colonies that we generated through Sendai reprogramming (**Figure 1g**; $R^2 = 0.03608$).

2.3.5 Sendai reprogramming results in bona-fide iPSCs

The colonies we picked from our Sendai transfection were chosen based on Tra1-60 expression and having a similar morphology to hESCs. We then had to confirm that the selected colonies were truly iPSCs. All cells in the iPSC colonies expressed Nanog (**Figure 2a, Supplementary Figure 1**) and Oct4 (**Figure 2b, Supplementary Figure 2**). The cells also expressed SSEA3 (**Figure 2c, Supplementary Figure 3**) as assessed by flow cytometry. At the transcript level, the iPSC colonies expressed a wider panel of pluripotency markers to a similar level as H9 hESCs (**Figure 2d**, n=1 for H9 hESCs). At this time, there was no sign of Sendai virus as assessed by qRT-PCR (data not shown). Taken together, the gene signature of the reprogrammed colonies corresponded to that of a reprogrammed cell.

The reprogramming process requires that the somatic cell undergo epigenetic and transcriptional changes. It is common that other genetic aberrations can occur simultaneously with the necessary changes to take on the pluripotent phenotype as well as through prolonged *in vitro* culture. The assay, from StemCell Technologies, showed that for the majority of the iPSCs generated had a copy number of 2 for the tested chromosomal regions (**Figure 2e**). In two of the iPSC lines, L175P and R947C, a duplication in chromosome 20q was detected. Patient L175P is female but only had a copy number of 1 for the X chromosome suggesting a deletion in that region as well.

Transcriptionally and morphologically, the iPSCs were similar to hESCs; however, pluripotency must also be functionally verified. To do this, we used an *in vitro* embryoid body (EB) formation assay (**Figure 3a**). This assay allows us to assess the spontaneous

differentiation of the iPSCs into the three primary germ layers – endoderm, mesoderm, and ectoderm. The EBs were collected and used for qRT-PCR to measure transcript levels of global germ layer markers. In doing, this, we found that iPSCs generated through Sendai reprogramming were capable of forming cells of the three germ layers (**Figure 3b-h**) although not all of the markers were increased in the EB samples compared to the iPSC lines. Certain markers appeared enriched more frequently in the EB samples, such as AFP and MAP2, whereas others showed no difference between the EB and iPSC samples across all lines, such as FOXA2, SOX17, and MSX1.

2.4 DISCUSSION

The ability to reprogram adult somatic cells into iPSCs has transformed the field of biology, providing unprecedented ways to create *in vitro* human disease models, screen drugs, and pave the way for regenerative and personalized medicines. Here, I outline the method that I employed to create patient-specific iPSC lines from a small number of non-mobilized PBMCs using Sendai virus carrying the Yamanaka factors (**Figure 4**). The iPSCs that were generated show transcriptional and functional evidence of pluripotency.

In 2010, Loh, *et al.* reprogrammed PBMCs into iPSCs using lentivirus; however, in my study, lentiviral reprogramming of PBMCs did not yield any iPSCs¹⁶⁸. The reason for the discrepancy could be due to the methods that were used. In the 2010 paper, Loh, *et al.* used a 2-hit approach in order to generate iPSCs¹⁶⁸. Here, I tried only a 1-hit lentiviral approach because the viability of the PBMCs following the first hit was already so low that a second hit would not be reasonable. Despite the inability to create iPSCs using lentivirus,

we were successful in creating the lines required using the Sendai virus. When tested with qRT-PCR there was no evidence of any remaining levels of Sendai transcript; however, in 2015, Choi, *et al.* demonstrated that the transcriptional pattern in iPSCs may still be affected by Sendai after viral loss¹⁷⁶. This was not assessed in this study but is an important factor to consider for future work, especially if using the Sendai-generated iPSCs for any downstream clinical or drug-screening applications.

There was no correlation between the reprogramming efficiency and the size of the CD34⁺ population in the samples, contrary to the original report on Sendai virus PBMC reprogramming¹⁷⁵. Several groups in the past have suggested that terminally differentiated T cells are the cells that get reprogrammed in bulk PBMC populations^{168,171}, although all have shown the capacity to be reprogrammed to some extent¹⁶⁴. In this work, we did not separate any subpopulations of the PBMCs during the reprogramming experiments, so we did not do a targeted reprogramming method as was first done with Sendai virus¹⁶⁹. However, it is possible to replicate the work done by Staerk, *et al.* and Loh, *et al.* and check what the cell of origin for the iPSCs was^{168,171}. This can be done by retrospectively examining the T cell receptor (TCR) and immunoglobulin (IG) loci of the genome. In T cell development, the TCR locus of the genome is rearranged, whereas in B cell development, the IG locus is. Because these loci are irreversibly changed in the genome, PCR analyses with specifically designed primer sets can determine if the cell of origin for reprogramming was a T cell (TCR rearrangement), a B cell (IG rearrangement) or a myeloid cell (neither TCR or IG rearrangement)¹⁷¹.

One of the advantages to using PBMCs in reprogramming work is the decreased risk of having genetic aberrations unrelated to the patient's disease because of genetic

factors. In using Sendai virus, we can maintain this because it is a non-integrating method of reprogramming, whereas lentiviral methods require transgene insertion. By using lentivirus on PBMCs, the patient's genetic landscape may be further altered such that the benefit of using PBMCs is decreased; however, genetic mutations may arise through *in vitro* culturing of iPSCs. In 2012, Cheng, *et al.* demonstrated that most of these variants are in line with mitosis rates so although pre-existing mosaicism in somatic cells may account for part of this, long-term culture is a major factor^{24,177}. Some of the most common chromosomal changes that occur in the derivation and culturing of iPSCs can be tested with the Genetic Analysis kit (StemCell Technologies). In using this kit, I was only able to confirm that 3 out of 5 lines generated had no or a minimal chance of having an amplification or deletion in 1 of the tested regions; however, 2 of the lines had a definite duplication in chromosome 20q. This amplification is commonly seen in iPSCs and usually involves a minimal amplification in the centromeric region of the q-arm (StemCell Technologies). This specific region of duplication is not commonly found through typical G-band karyotyping because the resolution is not high enough. The passages at which the iPSCs were tested was not consistent (range from passage 19 to 41) and should have been done routinely to detect when/if these mutations arise which is now more possible with the use of the Genetic Analysis kit. The two iPSCs that had definite duplications were at passages 24 (R947C) and 41 (L175P) whereas R1000Q had no signs of the common abnormalities and was at passage 28. This suggests that although with every replication and passage, the risk of incurring a mutation increases, routine checking of at least the most common chromosomal mutations of iPSCs should be widely implemented at all time points of culturing. This is reinforced by the presence of a deletion in the X chromosome in patient

L175P. Deletions in the X chromosome in iPSCs is common and often includes deletion of the entire chromosome. For groups studying X-linked diseases, this could be of extra importance as it would severely impact their model and the read-out from any assays, again highlighting the importance of routine iPSC genetic analysis. To further analyze these iPSCs and confirm the results seen through this screen, full scale karyotyping should be used to identify any less common chromosomal changes that were not screened by the kit or ones that occur in less than 20-30% of the culture as that is the detection limit of the kit.

iPSCs must be functionally validated on top of the transcriptional analysis that is typically performed. Here, I showed the use of an EB assay to assess tri-lineage differentiation capacity that has been previously used¹⁷². To validate this, various markers from each lineage were chosen based on the fact that they have been previously used for this purpose and/or they are considered early markers of that lineage (such as Brachyury). The N819S iPSC line showed increased expression of at least one marker of each line in the EBs compared to the iPSCs, whereas the Healthy iPSCs showed no enrichment in any of the tested markers. The other lines showed increased expression in the EB samples in at least one or more of the tested markers, but further analysis must be done to validate the tri-lineage potential of all lines. In many other studies, an EB Taqman scorecard was used that screened 92 genes encompassing all three germ layers^{172,178,179}. This scorecard improved the ability to detect tri-lineage differentiation capacity as some of the early markers, such as some of those used here including Brachyury, Sox17, FoxA2, and Pax6, are also expressed to a certain level in hPSCs¹⁷⁹. The kinetics of the expression of various markers also changes over the course of EB differentiation which may explain the variability in marker expression¹⁷⁹. In the study from which the EB protocol was taken, the

authors noted that there was variability across the iPSC lines in their ability to differentiate into all three germ layers based on gene expression of certain markers, but all did express some markers from the three germ layers¹⁷². Thus, using a larger number of genes would improve the read out of the EB assay significantly as the genes tested may not have been as widely expressed at the end point. Using only 2-3 markers for each lineage, even though they were general markers for that lineage, limits the cell types that can be identified. Additionally, the gold standard teratoma assay would be ideal to further validate these findings. Through histological analysis of teratomas, it is possible to identify specific mature or progenitor cell types that may be otherwise missed in the method used here. Bisulfite sequencing or ATAC-seq would also be useful to show that the reprogramming process resulted in permanent chromatin remodeling to mimic that of an ESC.

Another advantage to using PBMCs is the ease of sample collection from patients. Drawing non-mobilized blood is a routine procedure that requires little time and is minimally invasive and inconvenient to the donors. In this work, I was able to generate a substantial number of iPSCs from only 50,000 PBMCs. The PBMCs were isolated using the simple and standard Ficoll separation method and were not enriched for any subpopulation. The cells were cultured for only 1-2 days prior to transfection and resulted in visible iPSC colonies within 1-2 weeks of transfection, thereby significantly reducing the amount of time from sample collection to iPSC generation compared to fibroblast reprogramming techniques. In 10 mL of peripheral blood, it is possible to get anywhere from 10^6 - 10^7 PBMCs, meaning that on average only $\sim 50\mu\text{L}$ of blood would be required to obtain a usable number of iPSC colonies from a patient. The remainder of the PBMCs could be banked and used for measuring metabolites, primary analysis on the PBMCs, or

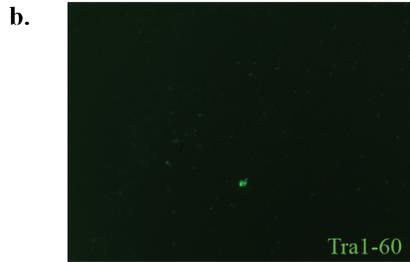
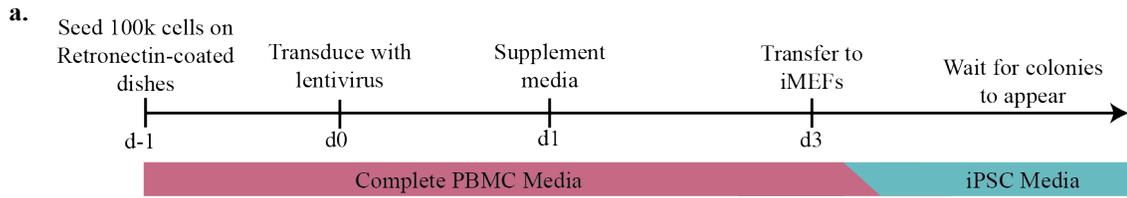
future work thereby increasing the therapeutic avenues from the same original donated sample. Alternatively, even smaller samples could be taken from patients to reduce any stress that the patients may experience during a blood draw. Additionally, by only requiring a small number of cells, creating iPSCs from banked samples where there may only be limited samples is possible. This would be particularly advantageous for reprogramming archived samples from patients with rare diseases. In doing this, larger samples sizes could be generated and therefore improve the power of any experimental analysis, which would allow for more a better biological understanding of rare genetic diseases.

Despite the success of the aforementioned protocol, the use of these cells for therapeutic applications remains uncertain. The use of a transgene free method of reprogramming significantly improves upon the need for integration of oncogenes, where reactivation of these transgenes is of concern; however, a transgene-free approach without the use of oncogenes in general would be ideal. In 2010, Zhu, *et al.* showed that it was possible to reprogram human somatic cells (keratinocytes, HUVECs, and amniotic fluid derived cells) only using Oct4 from Yamanaka's original cocktail¹⁶². Through the addition of chemicals that activate PDK1 and inhibit histone deacetylases (HDACs) as well as the TGF β and MAPK/ERK pathways, iPSC colonies could be generated. Although they were successful, this process took anywhere from 5-8 weeks and was extremely inefficient, generating only 4-6 colonies from 1,000,000 seeded keratinocytes¹⁶². Generating iPSCs from the adult keratinocytes also required the supplementation of CHIR 99021 and parnate, implying that the adult cell reprogramming process was more inefficient. This protocol requires additional optimization to be more widely employed; however, this work stands

as a useful proof-of-principle that alternatives to the oncogenic Yamanaka factors could be used to improve the clinical outlook of patient-derived iPSCs.

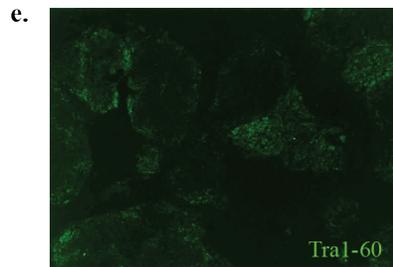
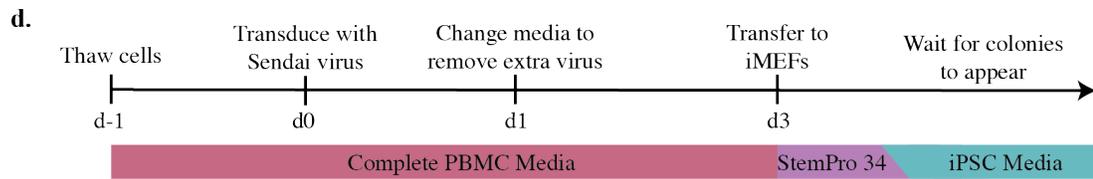
As more work is done to elucidate the stepwise process through which a somatic cell is reprogrammed into an iPSC, it will be possible to use alternatives to the Yamanaka factors. Recently, Bidy, *et al.* (2018) demonstrated that through sequential cell tagging, it is possible to reconstruct the reprogramming timeline with improved accuracy¹⁸⁰. In their work, they did direct lineage reprogramming of MEFs into induced endoderm progenitors; however, there is a clear application of this work in the reprogramming of somatic cells into iPSCs. If this method could be combined with a similar approach as the one used by Mai, *et al.* more specific steps in the reprogramming process would be elucidated^{147,180}. Using this method could potentially identify upstream regulators of the Yamanaka factors, which could then be used in place of the traditional oncogenic factors. Together with more closely studying the factors present in oocytes that make SCNT possible in a short time frame could improve the clinical outlook of iPSCs and their derivatives.

The protocol presented here demonstrates that a small number of PBMCs isolated from a small sample of non-mobilized peripheral blood requires no enrichment or significant *in vitro* culture conditions to create insertion-free iPSCs in as little as 2 weeks. The iPSCs generated can be used for the creation of *in vitro* disease models or drug-screening platforms.



c.

Days cultured before transduction	Number of iPSC colonies observed
1	0
2	0
7	0



f.

	+ Polybrene	- Polybrene
50,000 PBMCs	5	>100
100,000 PBMCs	10	>150

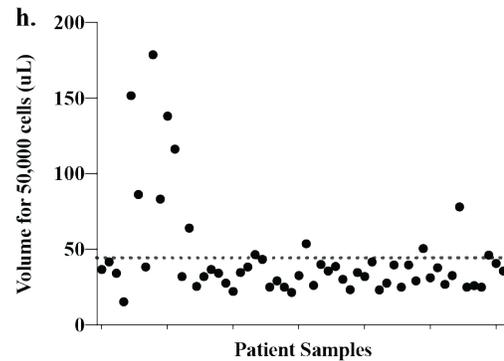
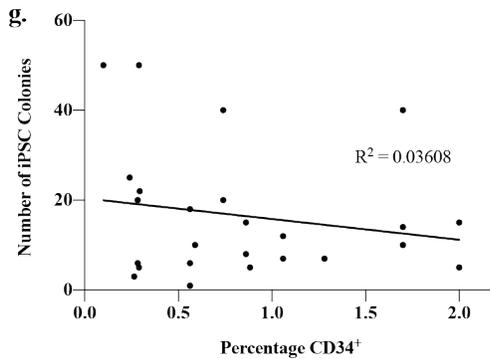


Figure 1. Sendai virus efficiently reprograms a small number of PBMCs irrespective of CD34⁺ content. A lentiviral reprogramming method (a) previously reported to successfully reprogram CD34⁺ cord blood was used with non-mobilized PBMCs but did not yield any Tra1-60⁺ colonies (b). Three different pre-transduction culturing lengths were used (c) but none of them improved the efficiency of reprogramming. Instead, the Sendai virus reprogramming method (d) was adopted and successfully reprogrammed the PBMCs as seen through the generation of Tra1-60⁺ colonies (e). The reprogramming efficiency was improved in the absence of Polybrene and was not dependent on initial PBMC number (f), or percentage of CD34⁺ cells in the original samples (g; $R^2 = 0.03608$). Based on the efficiency we obtained through using 50,000 cells, we determined that across the 56 patient samples collected, the volume needed to generate iPSC colonies is $44.46 \mu\text{L} \pm 31.99 \mu\text{L}$ (h; mean \pm SD; dotted line on graph represents average volume).

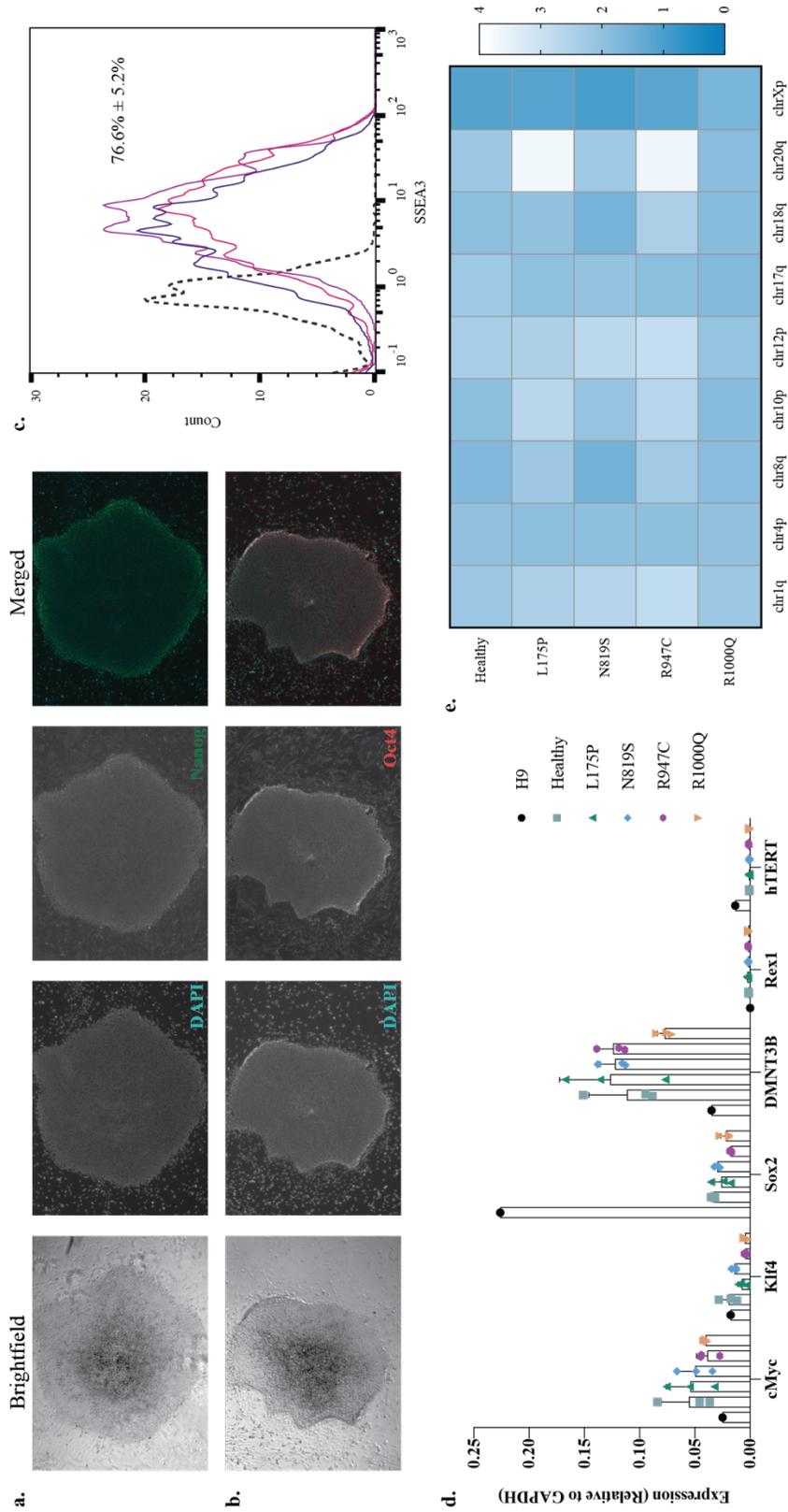


Figure 2. iPSC colonies generated with Sendai virus are positive for classic pluripotency markers. Immunofluorescence staining was performed on the iPSC colonies for Nanog (**a**, images from L175P) and Oct4 (**b**, images from R1000Q) to validate pluripotency. All DAPI stained cells of the colonies appeared positive for both markers (ImageJ was used to merge DAPI and Oct4 or Nanog images for the generation of this figure). Surface expression of SSEA3 was confirmed via flow cytometry (**c**, plots of the Healthy iPSC line). The expression of common pluripotency markers was similar across the iPSC lines and were comparable to that measured in the H9 ESC line, although the H9 line showed elevated Sox2 and decreased DMNT3B expression compared to the Sendai iPSCs as measured through qRT-PCR (**d**). Some of the iPSC colonies generated with Sendai showed signs of containing some of the most commonly identified genetic abnormalities in iPSCs as tested with the Genetic Analysis Kit from StemCell Technologies (**e**). The heat map in this figure shows copy number of the locus tested with 2 indicating no abnormalities. Lines L175P and R1000Q are female and should therefore have a copy number of 2 for chrXp whereas N819S and R947C are male and therefore a copy number of 1 for chrXp is considered normal.

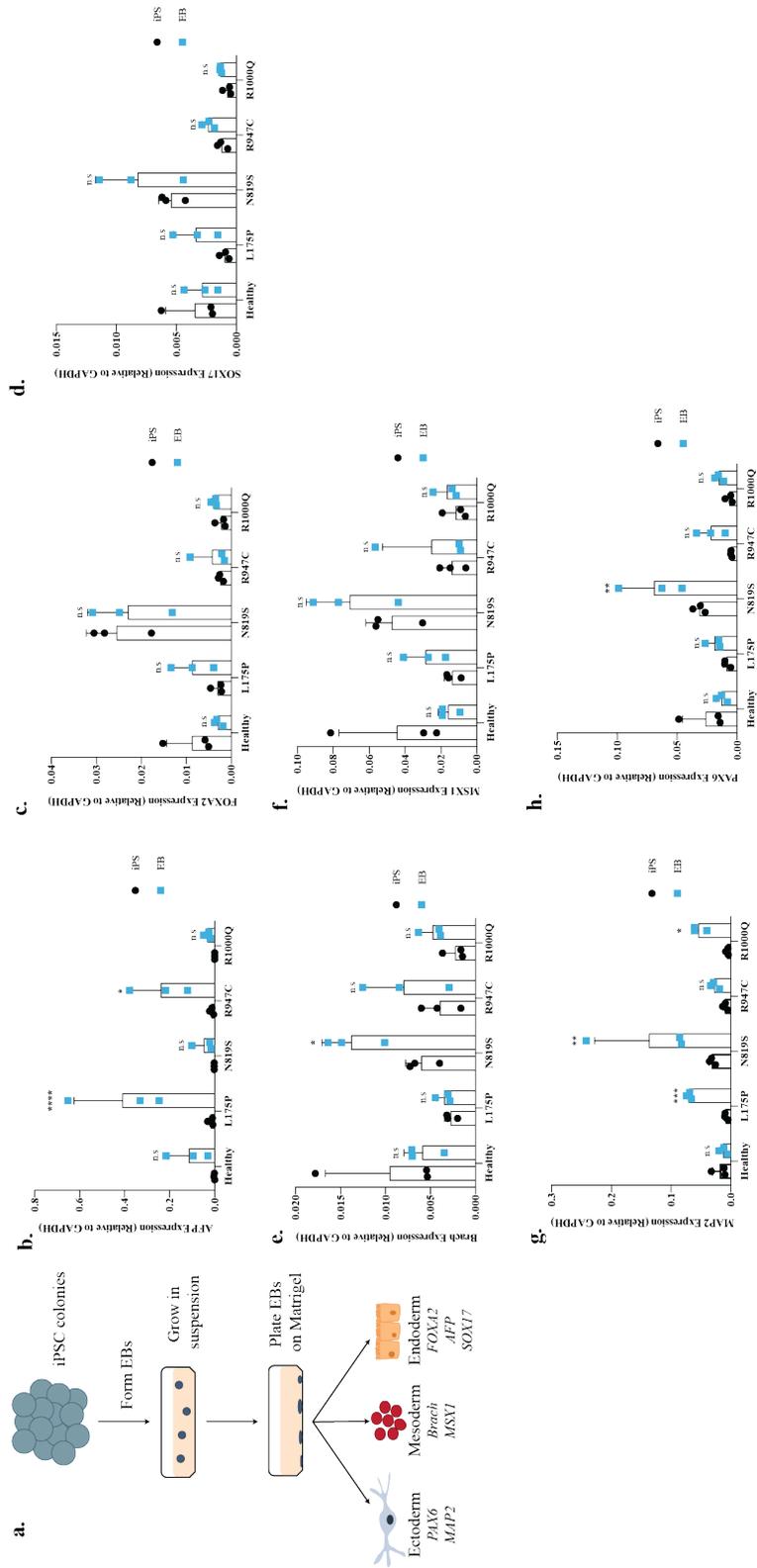


Figure 3. Sendai-derived iPSCs demonstrate functional pluripotency through tri-lineage differentiation. Pluripotency is defined through transcriptional and functional abilities of the cells. To confirm tri-lineage differentiation capacity, an EB assay was performed, where generation of the three germ layers were confirmed through qRT-PCR analysis of common markers (**a**). With this method, markers of endoderm (**b-d**), mesoderm (**e-f**), and ectoderm (**g-h**) were detected in the EB samples to varying degrees. AFP and MAP2 were enriched in 3 of the iPSC samples showing the highest detection in the EB over iPSC samples; however, some of the markers were also measured in the starting iPSC lines, suggesting further screening is needed to confirm the efficiency of the EB differentiation assay used. (Significance calculated between EB sample and iPSC sample for each marker using two-way ANOVA; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$)

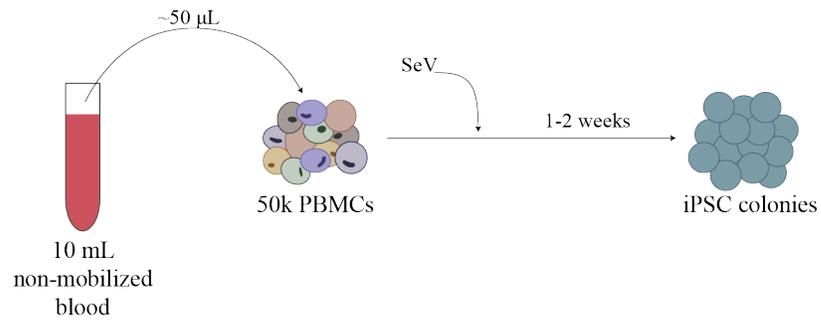
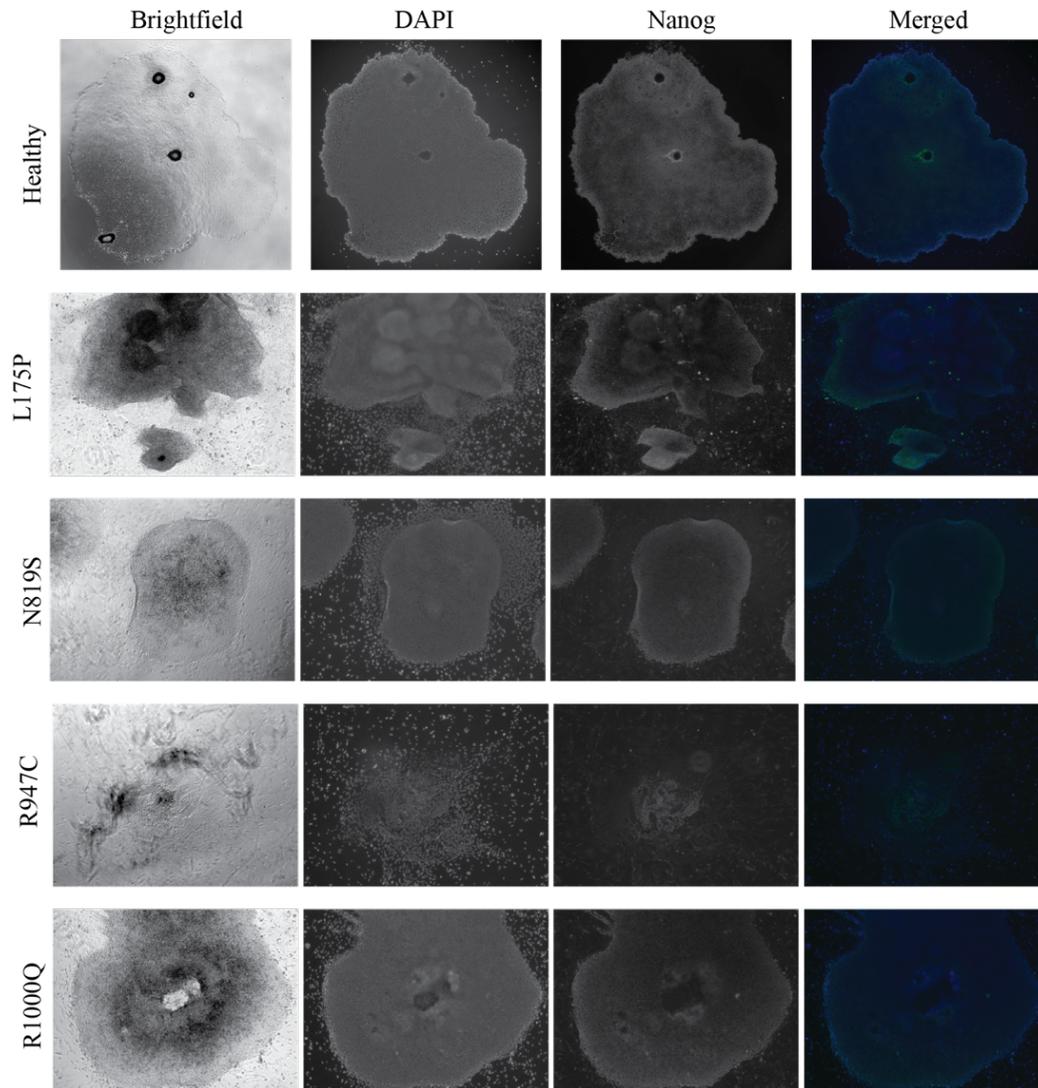
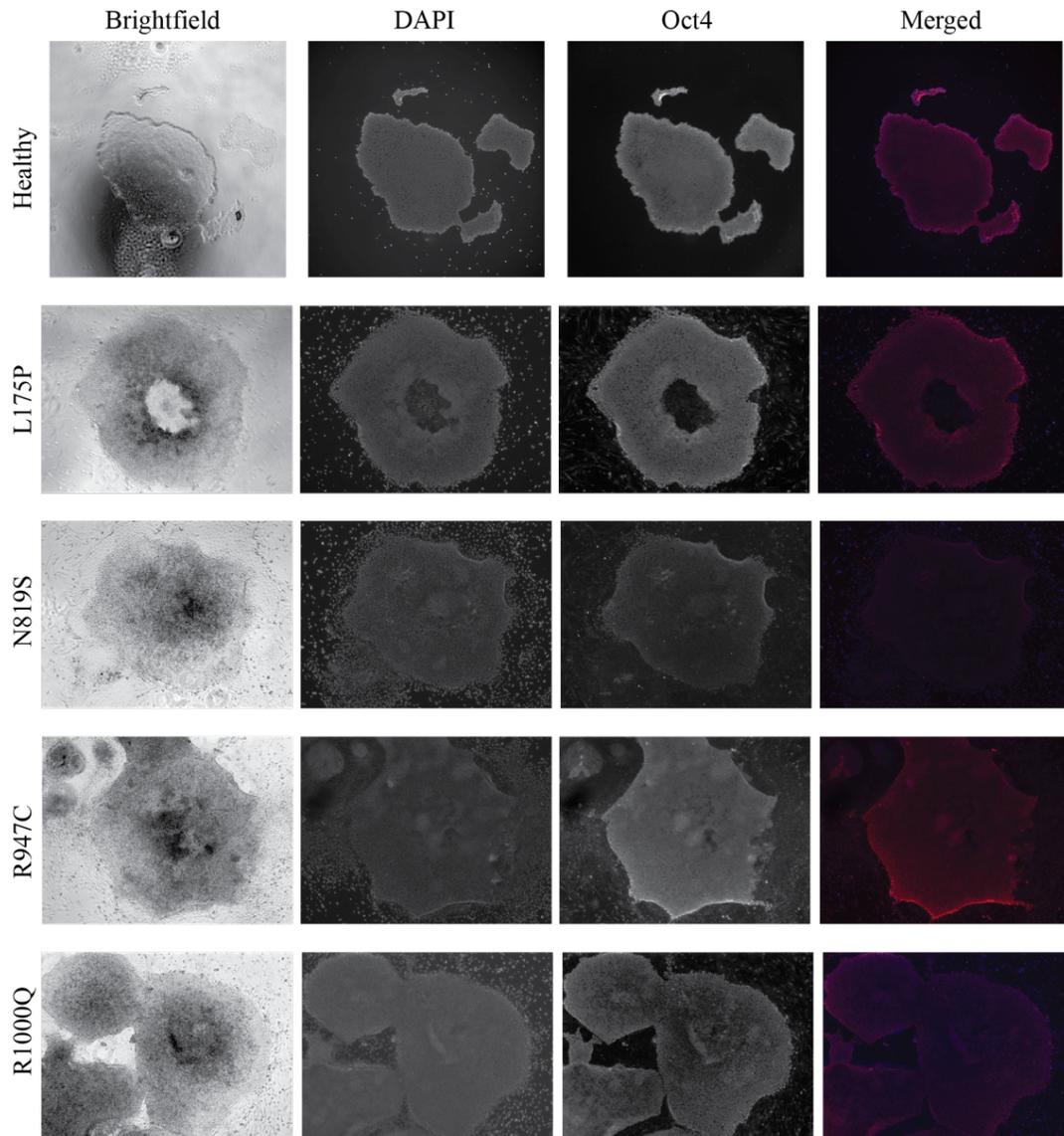


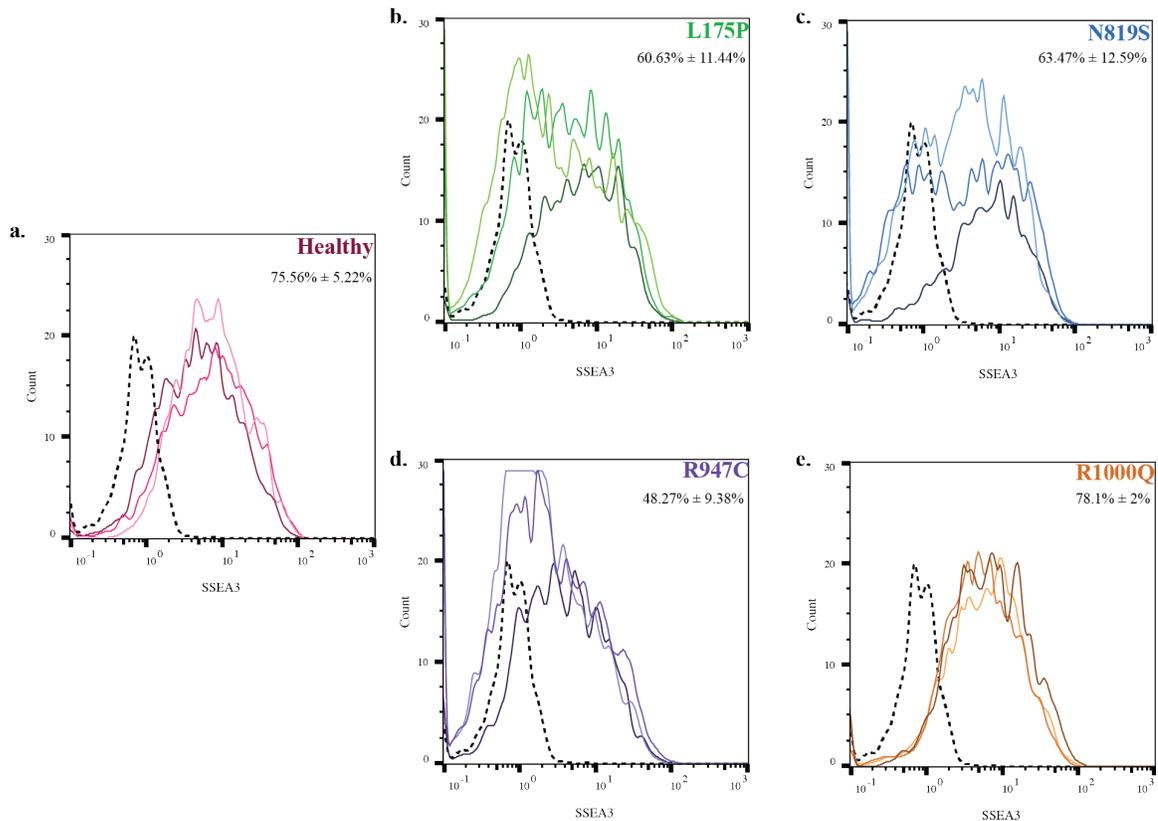
Figure 4. Method of generating iPSCs with as little as 50 µL of non-mobilized blood using the Sendai virus. From the starting 10 mL of non-mobilized blood, roughly ~ 50 µL is needed to isolate 50,000 PBMCs, which is sufficient to generate iPSCs within 2 weeks.



Supplementary Figure 1. Immunofluorescence staining of iPSCs for Nanog. Nanog expression was assessed through immunofluorescence for all lines. The Healthy iPSC images were taken of the colonies cultured in mTeSR1 and the remainder (L175P, N819S, R947C, R1000Q) were cultured in MEF-CM. Brightfield images were taken of the colonies, followed by images for the DAPI nuclear stain and Nanog all at the same magnification. ImageJ was used to merge the DAPI and Nanog images to generate this figure.



Supplementary Figure 2. Immunofluorescence staining of iPSCs for Oct4. Nanog expression was assessed through immunofluorescence for all lines. The Healthy iPSC images were taken of the colonies cultured in mTeSR1 and the remainder (L175P, N819S, R947C, R1000Q) were cultured in MEF-CM. Brightfield images were taken of the colonies, followed by images for the DAPI nuclear stain and Oct4 all at the same magnification. ImageJ was used to merge the DAPI and Oct4 images to generate this figure.



Supplementary Figure 3. SSEA3 plots of PBMC-derived iPSCs. iPSCs from all created lines were tested for SSEA3 expression via flow cytometry – Healthy (a), L175P (b), N819S (c), R947C (d), R1000Q (e). Each coloured line represents a separate replicate from the same iPSC line. The lines showed variable, but positive expression of SSEA3 compared to the unstained control (dotted black line in each histogram). Images were created using FlowJo v10. Values indicate percentage of cells positive for SSEA3 and are represented as mean ± SD.

Table 1. Previously used methods for cellular reprogramming

	Factors Used	Cell Type Used	Method of Introduction
Takahashi	Oct4, Sox2, Klf4, cMyc	MEFs Human fibroblasts (adult)	Retrovirus (MMLV-based)
Yu	Oct4, Sox2, Nanog, Lin28	Human fibroblasts (fetal & foreskin)	Lentivirus
Kaji	Oct4, Sox2, Klf4, cMyc	MEFs Human fibroblasts (fetal)	Linearized 2A-peptide-based polycistronic vector PiggyBac transposon
Stadtfield	Oct4, Sox2, Klf4, cMyc	Mouse hepatocytes and fibroblasts	Replication-defective adenovirus (PHIHG-Ad2)
Zhou	Oct4, Sox2, Klf4, cMyc	Human fibroblasts (fetal)	Adenovirus (pShuttle-CMV)
Fusaki	Oct4, Sox2, Klf4, cMyc	Human fibroblasts (fetal & adult)	Sendai virus
Yu	Oct4, Sox2, Klf4, cMyc, Nanog, LIN28, SV40LT	Human fibroblasts (foreskin)	oriP/EBNA1 vector

Warren	Oct4, Sox2, Klf4, cMyc, LIN28	Human fibroblasts (fetal & foreskin)	RNA delivery
Zhou	Oct4, Sox2, Klf4, cMyc, VPA	MEFs	Poly-arginine transduction domain- fused proteins
Weltner	Oct4, Sox2, Klf4, cMyc, Lin28, Tp53 shRNA	Human fibroblasts (neonatal, adult)	CRISPRa targeting

Table 2. Factors used for reprogramming

Genes		Small Molecules		
		<i>HDAC Inhibitors</i>	<i>DNA Methyltransferase Inhibitors</i>	<i>Pathway Modulators</i>
Oct4	Lin28	VPA	AZA	MEK inhibitors (PDK1 activator)
Sox2	Nanog	TSA	DZNep	GSK3 inhibitors (CHIR990291)
GATA3-VP16	SV40LT	SAHA		Dnmt1 inhibitors
mir200c/302/369	NKX3-1	NaB		TGFβ inhibitors (616542)
	hTERT			Forskolin
	Esrrb			
	Sall4			
	Nr5a2			

Table 3. qRT-PCR primers

Gene	Forward Primer 5' → 3'	Reverse Primer 5'→3'
hTERT	TGTGCACCAACATCTACAAG	GCGTTCCTGGCTTTCAGGAT
REX1	TCGCTGAGCTGAAACAAATG	CCCTTCTGAAGGTTTACAC
SOX2	CCCAGCAGACTTCACATGT	CCTCCATTTCCCTCGTTTT
MYC	TGCCTCAAATTGGACTTTGG	GATTGAAATTCTGTGTAAGTGC
KLF4	GATGAACTGACCAGGCACTA	GTGGGTCATATCCACTGTCT
DNMT3 B	ATAAGTCGAAGGTGCGTCGT	GGCAACATCTGAAGCCATTT
SeV C	GGATCACTAGGTGATATCGAG	ACCAGACAAGAGTTTAAGAGATATGTA TC
SeV- KOS	ATGCACCGCTACGACGTGAGC GC	ACCTTGACAATCCTGATGTGG
SeV-Klf4	TTCCTGCATGCCAGAGGAGCCC	AATGTATCGAAGGTGCTCAA
SeV-Myc	TAACTGACTAGCAGGCTTGTCG	TCCACATACAGTCCTGGATGATGATG
GAPDH	TCCCTGAGCTGAACGGGAAG	GGAGGAGTGGGTGTGCGCTGT

CHAPTER 3: EFFICIENT DIFFERENTIATION OF PATIENT-DERIVED iPSCS INTO MATURE ENDOTHELIAL CELLS

PREFACE

This chapter focuses on the creation of an endothelial specific model for use in Chapter 4. The iPSCs used for the differentiation steps are those developed through the methods highlighted in Chapter 2. At the time I joined the lab, there wasn't an established efficient protocol for endothelial cell differentiation being used in the SCCRI. The protocol I optimized was first published in Stem Cell Reports in 2014 by Lian *et al.*¹⁸¹. The improved protocol highlighted in this chapter was used for all subsequent differentiation to an endothelial lineage in the lab and is the protocol used to generate the cells in Chapter 4. The information presented in this chapter, in both the introduction (3.1) and the results (3.3) was submitted to Stem Cells and Development as part of a Review paper on the efficiency of current endothelial differentiation protocols. I was responsible for writing the manuscript of the review as well as generating all figures and tables that were also submitted. All of the work shown in this chapter was done by me with the exception of the sorting process which was done by one of Zoya Shapovalova, Minomi Subapanditha, or Hong Liang.

3.1 INTRODUCTION

3.1.1 Role of endothelial cells in tissue homeostasis

Endothelial cells are some of the most abundant cell types, with roughly 10^{12} cells in the human body¹⁸². Initially, endothelial cells were viewed as having a passive role in organ function by simply allowing for the transport of oxygen, nutrients, and inflammatory cells^{183,184}. Since then, their role in tissue homeostasis has come to be better appreciated. Endothelial cells can also function to guide regeneration and repair of adult tissues both directly as well as through their ability to sustain tissue-resident stem cell populations¹⁸².

Endothelial cells line the vessels of the vascular network including large arteries, veins, and capillaries, as well as the vessels of the lymphatic system. The function and morphology of endothelial cells differs from one vessel type to another^{99,185}. During development, endothelial cells arise from mesodermal cells termed angioblasts¹⁸⁶. Angioblasts coalesce into the primitive vascular plexus where underlying genetic differences and hemodynamic forces result in the specification of arterial and venous lineages¹⁸⁷. Subsequent differentiation of the venous lineage gives rise to the lymphatic vessels. Angiogenic remodelling of the arterial and venous endothelium continues throughout development to generate tissue-specific endothelium that have unique functional capacities required of their niche¹⁸⁷.

Endothelial cells of major large arteries have been heavily studied as they have a major role in the development of coronary artery disease – the leading cause of global mortality⁶⁷. Endothelial cells of arteries are important for transport of nutrients and oxygen

and are also critical for the maintenance of vascular tone and a thromboresistant surface¹⁸⁸. Through the production of vasodilators such as prostacyclin (PGI₂) and nitric oxide (NO), endothelial cells initiate a signal cascade in the neighbouring smooth muscle cells causing them to relax, leading to a decrease in overall blood pressure. In contrast, endothelial cells also produce both angiotensin II and endothelin-I, which are potent vasoconstrictors¹¹⁸. Changes in the ability of endothelial cells to produce either type of vasoactive compound can result in vascular changes. The production of both PGI₂ and NO can be stimulated by shear stress-induced intracellular increases in calcium levels or induction of cyclooxygenase 1 and eNOS^{188,189}. Endothelial produced PGI₂ acts locally on the smooth muscle cells to cause dilation of the blood vessels, whereas NO has a more systemic impact on arterial relaxation¹¹⁸. At the same time, both of these factors reduce platelet aggregation on the endothelium^{188,190}. In cases of endothelial dysfunction, as seen in hypertension and coronary artery disease, the level of NO is significantly impaired. This impairment results in an inability to decrease blood pressure as a result of its impact on smooth muscle cells, as well as a decrease in the thromboresistant feature of the endothelium.

Endothelial cells of capillaries are further specialized based on the tissues in which they reside. For example, the endothelial cells of the blood-brain barrier form a very impermeable layer to protect the brain¹⁹¹, whereas the endothelial cells of the kidneys are more permeable to allow for blood filtration¹⁹². The intercellular interactions are not the only tissue-specific change that can be seen in endothelial cells. Morphological differences, such as the presence of specialized fenestrae on the surface of endothelial cells in endocrine organs, also play a critical role in their ability to regulate organ function¹⁹². Organotypic endothelial cells also differ in the angiocrine factors that they produce, which are important

for homeostasis of the organ as well as tissue patterning during development or subsequent repair¹⁸². The exact mechanism by which further specialization occurs in these cells is not fully understood. It is likely, however, that both extrinsic and intrinsic factors play a role¹⁸². Extrinsic biophysical factors, such as shear stress, differ across body tissues and can be sensed by the endothelial cells. The signals are transduced by a mechanosensory complex, which may alter the transcriptional signature and therefore angiocrine factor production and overall morphology of endothelial cells from different sites.

As technology advances, our ability to dissect the specific roles and differences between endothelial cells of various types increases. In this way, our ability to then model and develop these cell types will also increase thereby strengthening and improving current *in vitro* endothelial cell generation protocols.

3.1.2 Role of endothelial cells in disease and methods for studying

Dysfunctional endothelial cells have been linked to several diseases. Their roles in metabolic and genetic disorders, as well as cancers, exemplifies the importance of studying them. In fact, the leading cause of global mortality, cardiovascular disease, is primarily caused by CAD which is ultimately a disease of the arterial wall. Modelling disease states using human endothelial cells requires the use of the correct endothelial subtype for the disease of interest¹⁸². This can be accomplished by using *in vivo* systems, which are commonly used and allow researchers to examine the cell in the proper biological context. Studying endothelial cells in this way is extremely helpful but can be limiting if dynamic changes are to be examined because when using animal models, it is typically done post

mortem. In the 1970s, when *in vitro* culturing of endothelial cells became possible, their diverse functions and therefore potential role in disease development, became more apparent⁸². By growing endothelial cells *in vitro*, more precise manipulations to the system can be made such that changes in function or morphology can be better analyzed following either an acute or chronic disruption. However, for many disease types this can be difficult as obtaining the primary cells, as well as maintaining them in culture for long periods, is complex and limits the potential downstream applications due to cell numbers. In terms of *in vitro* modelling, this problem has been mitigated in two ways; through the use of non-primate cell sources, and through alternative sources of human endothelial cells. It is important to note that with all *in vitro* studies, aspects of the environment such as shear stress, are lost and changes in the cell types may occur. For example, when endothelial cells are cultured *in vitro*, they lose their quiescent state and become more activated in terms of replication rate⁸². Nonetheless, the ability to study and culture endothelial cells *in vitro* has been incredibly useful to the field as a whole.

To model arterial diseases, such as atherosclerosis, bovine arterial endothelial cells (BAECs) have often been used^{193–200}. These cells have the characteristic arterial gene signature making them useful for assessing molecular changes associated with risk factors; however, intrinsic differences between bovine and human cells makes translation of the results to a clinical setting difficult.

Similarly, the zebrafish and mouse systems have been a major asset for researchers in the endothelial field^{201–208}. Transgenic mice have been indispensable tools for learning about endothelial specification during development as well as *in vivo* response to disease¹²². This topic, along with the use of other animal models for atherosclerosis

specifically, has been extensively explored by Getz and Reardon, and was summarized Chapter 1.2.3 for CAD specifically¹²². Similar to issues that arise from the use of BAECs, these systems have intrinsic flaws for modelling human diseases that are primarily based on species differences^{102,123}. Size discrepancies between humans and the animals used in research equates to differences in heart rates and blood vessel diameter, therefore changing blood vessel hemodynamics^{102,122}. Hemodynamics are a major environmental factor that impacts endothelial function especially in the context of disease. Using larger animal models such as non-human primates or pigs reduces the variability between the human and model system; however, the cost and ethical considerations in using these animals, particularly for non-human primates, limits their utility in endothelial research¹²². In the endothelial field, like many others, mouse models remain a major contributor to advancing our current knowledge. The murine system is not without fault and differs from human pathophysiology in many cases¹²³. This issue is highlighted in the atherosclerotic mouse model systems. For mice to develop significant atherosclerosis spontaneously, they must possess deletions of genes involved in cholesterol metabolism such as *apoE* or *LDLR*. Although *apoE*^{-/-} mice can develop atherosclerotic plaques on a normal chow diet, it is common for them to be placed on a high-fat diet, similar to what is necessary to generate significant plaques in *LDLR*^{-/-} mice¹²². Although this may recapitulate a small fraction of human disease, it does not allow for the examination of genetic risk factors, which is believed to account for ~40% of disease risk²⁰⁹. Mice are also significantly smaller than humans and have an accelerated heart rate and smaller blood vessels, altering the hemodynamic environment^{102,123}. These differences impact the extent to which the information gathered from murine models can be applied to the human system. Thus,

although current non-human mammalian model systems have proven to be extremely useful in endothelial research, human models remain critical to improve disease models and therapeutic screening platforms.

Primary human endothelial cells can be difficult to obtain for disease-specific models, which has made it critical to find alternative sources. A number of human endothelial cell lines including arterial^{210–214}, venous, and microvascular^{215–217} from different tissues, are commercially available and/or can be isolated from patients, making it possible to use a more physiologically relevant endothelial subtype in experiments; however, the use of these has not become widespread in research. Many of these subtypes are difficult to maintain in culture as they undergo de-differentiation *in vitro*²¹⁸ and/or can be difficult to obtain from donors. Although some disease-specific cell lines are available, such as those for types I and II diabetes (Lonza), the etiology of the disease or genetic-background of the donors is not well-documented for other commercially available cells. As such, the genetic contribution to disease development would be ignored limiting our understanding of cellular behaviour in that context. To examine endothelial function in human disease, some groups have also used post-mortem or surgical tissue for histological analysis^{219–222}. This analysis provides information on human- and endothelial subtype-specific changes in the context of disease; however, it only provides a static view of the endothelium. Accordingly, an alternative strategy can be the use of well-defined endothelial cell models or patient specific endothelial cells. To this end, human umbilical vein endothelial cells (HUVECs) and blood outgrowth endothelial cells (BOECs) have been commonly used.

The isolation of HUVECs is non-invasive²²³ thereby making them one of the most easily accessible human endothelial cell and has resulted in it becoming the most commonly used human endothelial cell type in research^{195,200,210,224–230}, meaning that their phenotype and assay parameters have been well characterized. There is a limited ability to isolate HUVECs from patients with vascular abnormalities as they would have to be identified *in utero* and therefore their use in disease modelling often requires transfection if a genetic modification is necessary. As a primary cell, HUVECs are still difficult to transfect with high efficiency but many protocols and cell-specific reagents have been designed to mitigate this including the HUVEC-specific Nucleofection reagents by Lonza²³¹. As such, it is possible to perform knockdowns^{232–234}, over-expression²³³ and recently, CRISPR-mediated gene editing²³⁵, on HUVECs to create more accurate disease models *in vitro*.

As previously mentioned, using the proper endothelial subtype for modelling disease is critical for assessing phenotypic and genetic changes. Using HUVECs to model arterial or lymphatic phenotypes presents an issue as the gene signature between venous, arterial, and lymphatic endothelial cells differ^{99,185}. As with many primary cell types, longevity *in vitro* is another consideration with the use of HUVECs. They have a limited lifespan and are known to undergo phenotypic drift making it necessary to rely on early-passage data only²³⁶. Another issue with the use of HUVECs for modelling endothelial diseases is that HUVECs are from a fetal source and therefore express a more immature genetic signature compared to adult endothelial cells. Assessing adult diseases in a tissue or cell type that carries a fetal signature makes the model incomplete. This is an issue for

in vitro disease modelling that has to be addressed in most fields. To circumvent these issues, some researchers have begun to use blood outgrowth endothelial cells (BOECs).

BOECs can be derived from non-mobilized peripheral blood²³⁷. BOECs are circulating endothelial cells, which differ from early endothelial progenitor cells (EPCs); the genetic fingerprint of BOECs is that of an endothelial cell whereas that of the EPC more closely resembles a hematopoietic cell²³⁶. Since BOECs can be isolated from adult donors, their genetic signature is that of adult endothelial cells rather than the fetal signature possessed by HUVECs. In fact, BOECs more closely resemble microvascular and large-vessel ECs than other endothelial cell types at the genetic level²³⁶. BOECs express the endothelial markers CD31, vascular endothelial growth factor receptor 2 (VEGFR2), CD144, and von Willebrand factor (vWF) making them a useful tool for studying endothelial function *in vitro*^{226,236–238}. Another advantage to the use of BOECs is that it allows for patient-specific disease modelling *in vitro* since they can be obtained from patients with endothelial specific diseases²³⁹.

The use of BOECs has not become widespread, however, due to a number of disadvantages. One such disadvantage is that BOECs are low in number in circulation making it necessary to obtain large volumes of blood to isolate sufficient BOECs to perform *in vitro* analysis. Although BOECs have been reported to have a high proliferative capacity despite their late outgrowth, only approximately 0.34 colonies appear per 1 mL of peripheral blood²³⁶. This could pose as a problem for patients who present to the clinic and become enrolled in studies at the height of their disease. Typical protocols for isolating BOECs require 50 to 100 mL of peripheral blood resulting in the generation of 10^{19} endothelial cells following expansion²³⁶. Although this is a large number of cells, the

volume of blood required from patients can be onerous depending on their condition. Additionally, the number of BOEC colonies that are derived per 1 mL of blood is patient-specific. It was found that the number of endothelial outgrowth colonies is decreased in patients with a higher risk of cardiovascular disease, and also, the BOECs isolated from these patients have a higher rate of senescence *in vitro*^{240,241}. Therefore, for modelling cardiovascular disease from these patients, it would be necessary to obtain additional or larger samples to acquire sufficient BOECs for experimental purposes, which could again be difficult given the condition of the patients.

In order to minimize the dependency on sample abundance for the generation of patient-specific models, and to mitigate the issues associated with the general use of primary cells, researchers have turned to induced pluripotent stem cells (iPSCs) as a source of endothelial cells.

3.1.3 Existing endothelial differentiation protocols and their limitations

Pluripotent stem cells (PSCs) have become an invaluable tool for studying endothelial cells *in vitro*. The use of PSCs has allowed researchers to better elucidate the molecular cues for endothelial specification during human embryonic development. Moreover, PSCs are an indispensable tool for *in vitro* studies as they can be continuously used to generate endothelial cells from the same source. Typically, differentiation protocols aim to mimic *in vivo* development through the timed addition and removal of growth factors and antagonists of the main signalling pathways used in cell specification. For

endothelial cells, many different protocols have been established with varying efficiencies, some of which are outlined in **Table 1**.

Embryonic stem cells (ESCs), whether human or mouse, typically grow as colonies in a monolayer *in vitro*. Through the removal of LIF for mESCs or through manual scraping of hESCs, cystic-like structures, termed EBs, are formed. EBs undergo spontaneous differentiation to cells of all three germ layers²⁴². Using mESCs, vascular-like channels containing endothelial cells were identified on the surface of EBs, suggesting that the isolation of endothelial cells through EB differentiation was possible^{243–247}. Although this process was first observed through the use of ESCs, iPSCs mimic the results of the hESCs making it possible to perform the same EB differentiation assays with iPSCs.

The identification of endothelial-specific markers was instrumental in developing more sophisticated differentiation protocols (**Table 2**). Through early EB methods, the isolated vascular cells were typically a more homogenic endothelium precursor; a subpopulation of CD144⁺ and CD31⁺ cells in EBs were associated with hematopoietic commitment¹⁸⁶. In order to separate endothelial and hematopoietic cells in EBs, a combination of markers and functional outputs from both cell types is required. On day 10 of EB differentiation, a CD45⁻CD31⁺/VEGFR2⁺ population can be isolated²⁴⁸. When these cells are cultured in hematopoietic conducive conditions²⁴⁹, hematopoietic cells emerge; however, when cultured in endothelial conditions, vWF⁺AcLDL⁺eNOS⁺ cells emerge indicating the presence of mature endothelial cells²⁴⁸. This data reinforces the importance of both marker- and function-specific requirements for isolating a population of endothelial cells especially through EB differentiation.

In 1996, Vittet, *et al.* tracked the expression of endothelial markers across an 11 day EB differentiation protocol and found that endothelial precursors arose starting on day 3¹⁸⁶. To increase the extent of vascular differentiation in the EBs, Vittet, *et al.* supplemented the traditional EB protocol with angiogenic growth factors, which increased the final percentage of CD31⁺ cells. For the first time, this suggested that growth factors could be used to drive EB differentiation towards the generation of more specific endothelial cells rather than allowing for entirely spontaneous differentiation¹⁸⁶. The length and media used during EB differentiation are additional aspects that various groups have attempted to optimize for endothelial cell specification. In 2002, Levenberg, *et al.* tracked the expression of endothelial markers over the course of EB differentiation, and found that expression peaked during days 13-15²⁵⁰. This allowed them to isolate the cells at that stage and perform functional assessments on them as well as expansion by using endothelial-specific media. In 2007, Ferreira, *et al.* examined the difference in differentiation when either KO-SR or FBS were used²⁵¹. FBS was found to increase the yield of CD34⁺ endothelial cells as compared to KO-SR, although efficiencies differed depending on the cell line used. In the same year, another group developed an entirely serum-free protocol to avoid batch differences between lots of serum that may contribute to changes in differentiation efficiency²⁵². Mitigating factors that impact differentiation efficiency is an important consideration for the use of iPSCs as a source of endothelial cells for disease modelling.

The nature of EB differentiation allows it to be optimized to increase the yield of endothelial cells; however, it does not produce a pure population of any specific cell type.

Using a monolayer differentiation approach has been the next development in endothelial cell generation and appears to result in a more directed differentiation method.

Similar to the mechanism of EB differentiation, monolayer protocols attempt to mimic *in vivo* endothelial cell development. The structure of EBs promotes the crosstalk of cells and signalling cascades similar to what is experienced in the developing embryo; however, when the PSCs are plated as a monolayer, the structural cues are lost. To compensate for this, growth factors and antagonists can be added to the differentiation media at discrete time points to mimic development. Although the 3D environment which is more similar to what occurs *in vivo*, the even distribution and therefore exposure of the monolayer of PSCs to the differentiation media is thought to result in a more homogenous differentiation pathway.

In 2007, Wang, *et al.* were the first to demonstrate that endothelial differentiation could be accomplished through a 2D monolayer approach²⁵³. This method resulted in a similar percentage of CD34⁺ endothelial precursors as the EB method; however, the absolute number of cells that were collected was significantly higher. In creating a serum-free method, this group supplemented their standard differentiation media with angiogenic factors and saw an increase in the percentage of CD34⁺CD31⁺ cells, suggesting that the addition of angiogenic factors in a 2D differentiation system increases efficiency²⁵³.

Accurately mimicking *in vivo* development requires the addition of both agonists and antagonists of key signalling pathways required for cell specification. In 2010, James, *et al.* employed a transforming growth factor β (TGF β) inhibitor, SB431542, to increase the yield of endothelial cells²⁵⁴. The cells were grown in suspension with the addition of bone morphogenic protein (BMP), fibroblast growth factor (FGF), and TGF β pathway

activators. The cells were then switched to an adherent protocol where TGF β pathway activation is removed and VEGF activation is added²⁵⁴. After 7 days of differentiation, the conditions switch to only include VEGF and FGF stimulation, and TGF β inhibition with SB431542. The timed inhibition of TGF β increased the ratio of PSCs:endothelial cells from 1:0.2 to 1:7.4 after the 14-day protocol. Inhibition of TGF β in this system reduced the maturation of the progenitors and as such, increased the proliferation capacity of the cells thereby increasing the final yield of endothelial cells²⁵⁴. Similarly, in 2014, Sahara, *et al.* performed a 2 phase differentiation over the course of 6 days²⁵⁵. In the first stage, BMP and Wnt signalling pathways were activated by BMP4 and GSK3 β inhibition respectively, both of which were done to direct differentiation towards the mesodermal lineage. The second phase was designed for endothelial specification and thus required the addition of vascular endothelial growth factor (VEGF), placental growth factor (PLGF), and hepatocyte growth factor (HGF), as well as the suppression of Notch through the addition of DAPT, a γ -secretase inhibitor²⁵⁵. Using this method, 1 PSC generated 2.5 endothelial cells; however, after one week of culture the ratio increased to 1:20 and subsequently 1:140 one week later. Similarly, Patsch, *et al.* modulated BMP and Wnt signalling to commit the PSCs to mesoderm and then used VEGF and protein kinase A activation by the addition of forskolin to direct the cells to the vascular lineage²⁵⁶. In 2016, Nguyen, *et al.* performed similar differentiation optimization work to determine if isolation of the cells at different times along with additional pharmacological modulation would impact both the percentage and maturity of the cells²⁵⁷. Their group found that although both isolation protocols produced a similar ratio of PSCs:endothelial cells (1:4.25 for protocol I and 1:5.03 for protocol II), the cells isolated from protocol II had a more mature phenotype than those

isolated from protocol I²⁵⁷. Collectively, this work clearly demonstrates that optimizing differentiation protocols through modulation of the BMP, Wnt, TGF β , and VEGF pathways in a time-wise manner, more accurately recapitulates *in vivo* development, and can result in more physiologically relevant cells for disease modelling. Additionally, timing of the process can impact the maturity of the resulting cells. These factors are important considerations when selecting a differentiation method to develop an *in vitro* disease model.

Performing *in vitro* differentiation eliminates many of the environmental cues that exist *in vivo* and as such, some groups have examined the impact of changing these factors during differentiation. In 2014, Zhang, *et al.* hypothesized that differentiating PSCs to endothelial cells in a controlled 3D environment would increase the differentiation efficiency²⁵⁸. Unlike EBs where the structure is formed spontaneously and is cystic in nature, Zhang, *et al.* differentiated PSCs on a specified 3D fibrin scaffold patch over 14 days, which included pharmacological modulation of the TGF β and VEGF pathways as the protocols mentioned above. The use of 3D scaffold increased the efficiency of generating CD31⁺ cells from 5% in a monolayer to about 44%, suggesting that differentiating PSCs in a specified 3D environment can increase efficiency in a significant way²⁵⁸.

An important environmental condition that is often ignored in *in vitro* differentiation protocols is O₂ levels. Hypoxia increases the production of VEGF, bFGF, platelet-derived growth factor (PDGF) and angiopoietin 1 (ANG1), which promote vascular cell recruitment and differentiation and as such, Tsang *et al.* sought to determine whether lowering O₂ *in vitro* would result in an increase in differentiation to vascular cells²⁵⁹. The PSCs differentiated under hypoxic conditions (1% O₂) had double the amount

of CD31⁺CD144⁺ cells as compared to the normoxic conditions (21% O₂). They determined that hypoxia increased hypoxia inducible factor 1 α (HIF1 α) and in turn *Etv2* expression, which was responsible for the increased endothelial cell differentiation levels²⁵⁹.

Together this data demonstrates that continued optimization of differentiation protocols in the future should entail not only pharmacological modulation but also environmental regulation as well.

The protocols mentioned above differ in many aspects but typically result in the need to isolate the endothelial cells from the differentiation conditions. Methods such as magnetic- or flow-associated cell sorting (MACS or FACS respectively) are employed to harvest a pure endothelial population. These techniques rely on the surface expression of endothelial cell-specific markers (**Table 2**) to ensure that the resulting population is pure for downstream applications. Subsequent culturing of the isolated cells in endothelial-specific media is used to allow the cells to maintain their cellular identity and to further mature in culture.

As mentioned above, using cells for disease models *in vitro* requires the use of physiologically relevant cells. For endothelial cells, whether the cells are more venous, arterial, or lymphatic can greatly impact the relevance of the phenotypes to the disease at hand. Many published protocols generate a spectrum of endothelial cells rather than one specific subtype^{181,254,255}. Within these populations of isolated cells, there are often both venous and arterial cells, although lymphatic endothelial cells are rarely reported. The co-existence of arterial and venous endothelial cells in culture are likely due to the mechanisms used to generate them, whereas lymphatic endothelial cells arise from a subset of venous

endothelial cells and require cues from VEGFC, fms related tyrosine kinas 4 (FLT4), and prospero homeobox 1 (PROX1), the key lymphatic transcription factor, to induce their differentiation^{260–262}. Recently, hematopoietically expressed homeobox (HHEX) was found to be the transcriptional regulator of this signalling axis, making it an ideal target for future lymphatic endothelial cell differentiation protocols²⁶³.

Endothelial differentiation protocols typically begin with the induction of mesoderm progenitors from the PSCs, and then begin to specify towards endothelial cells through modulation of the TGF β , VEGF, and Wnt pathways. These pathways are important for the function of both venous and arterial endothelial cells, providing an explanation for the heterogeneous population. To generate one subtype of endothelial cell specifically requires additional work to mimic the specification steps that occur *in vitro*. Notch signalling is an important pathway in vascular development²⁶⁴, and its differential regulation is known to be important for arterio-venous specification^{218,265}. Although it was originally believed that hemodynamic factors contributed to this specification, knockout studies in mice showed that the arterio-venous specification is genetically determined prior to blood flow^{266–270}. As such, further investigating these pathways and modulating them would aid in the creation of a more homogeneous endothelial population. During development, lymphatic endothelial cells are derived from the cardinal vein and subsequently differentiation towards the lymphatic lineage, although much of the signaling pathway remains unknown²⁶⁶. Therefore, furthering this knowledge is the first step to creating a specific lymphatic endothelial differentiation protocol. Creating homogenous endothelial populations is extremely important not only for the improvement of *in vitro*

disease models but also for personalized and regenerative medicine applications that would require endothelial cells with tissue-specific functions²⁶⁶.

Arterial endothelial cells, specifically those of the coronary artery, are arguably the most important endothelial subtype for disease modelling and regenerative medicine purposes because coronary artery disease is the leading cause of cardiovascular-related deaths worldwide. Studying the role of endothelial cells in disease prevention and management will hopefully result in the generation of efficient treatment methods for patients who must manage the disease long-term.

Generating an enriched population of arterial endothelial cells was done in 2017 by tracking arterial endothelial cells from hPSCs using an *EFNB2*-tdTomato and *EPHB4*-GFP reporter line²¹⁸. The expression of *EFNB2* and *EphB4* are higher in arterial endothelial cells and as such, allowed the group to optimize the differentiation protocol to enrich for arterial cells. The reporter line allowed them to determine that VEGF-A, resveratrol (a NOTCH agonist), FGF2, L690 (an IMPase inhibitor), and SB431542 were necessary for arterial over venous specification. Interestingly, insulin, BMP4 and PDGF-BB, which are common in other endothelial differentiation pathways, specifically inhibit arterial endothelial specification although they do not alter pan-endothelial differentiation efficiencies²¹⁸. This work demonstrates the importance of examining differences in the genetic profiles of endothelial cells in order to best design protocols for the desired subtype. It again underlines the importance of examining the growth factors used in generating a disease model to ensure that the most physiologically relevant model is produced.

3.1.4 Chapter objectives

Endothelial cells underlie the pathology of the most common causes of death worldwide. Not only is CAD the leading cause of mortality, but many individuals live with the disease and side-effects from the medications needed to manage it and prevent further complications. The need to develop more comprehensive preventative and treatment methods necessitates the ability to study *human* cells in disease states. Using patient-derived iPSCs allows us to generate disease- and cell-specific models to better elucidate the molecular mechanisms underpinning cellular dysfunction. Deriving endothelial cells *in vitro* is an invaluable tool to the field of regenerative medicine. Not only is the vascularization of organoids and/or tissues grown *in vitro* necessary for their continued growth, but engineered blood vessels may come to be extremely helpful in the treatment of many diseases. Although many protocols have been published, optimization of these protocols is typically required to account for innate iPSC line differences as well as those imparted by the culture conditions. With this in mind, this chapter highlights the process by which a published protocol was significantly improved on to efficiently generate endothelial cells from PBMC-derived iPSCs.

3.2 MATERIALS AND METHODS

3.2.1 Cell culture

iPSCs were cultured in 90% DMEM/F12, 10% KOSR, 10mM NEAA, 2mM L-glutamine, 7mM β -mercaptoethanol, 16 μ g/mL bFGF (iPSC media) on iMEFs until a week before differentiation. iPSCs were passaged onto Matrigel coated plates and cultured in MEF-CM with 8 μ g/mL bFGF. Once isolated, endothelial cells were cultured in EBM2 media (Lonza) on Matrigel- or gelatin-coated plates. HUVECs were purchased from Lonza and cultured in EBM2 media on tissue-culture treated plates. Once confluent, HUVECs were passaged using TrypLE and seeded at a density of 2500 cells/cm².

3.2.2 Differentiation of iPSCs to endothelial cells

iPSCs were differentiated to endothelial cells using a protocol optimized from a previously published group¹⁸¹. In brief, iPSCs were singularized with Accutase at 37°C for 5 min. The MEF-CM supplemented with 8ng/mL bFGF and 5 μ M Y27632. 50,000 cells/cm² were plated on Matrigel-coated plates. Media was changed daily to MEF-CM + 8ng/mL bFGF. On day 0, media was changed to DMEM/F12, 2.5mM L-glutamine, 60 μ g/mL ascorbic acid (termed LaSR media) supplemented with 6 μ M CHIR99021 or 0.6 μ M BIO. 48 hours later, the media was changed to just LaSR for the remaining 3 days before FACS isolation of the CD34⁺ cells. During the optimization phase of this protocol, BMP4 (10-50ng/mL), Activin A (3-10ng/mL), and VEGF-A (10-50ng/mL) were added to

the media in varying concentrations after removal of the CHIR99021 or BIO. CD34⁺ cells were isolated through FACS and cultured on Matrigel-coated plates in EBM2 media.

3.2.3 Isolation of endothelial cells

Endothelial cells isolated through both protocols were cultured on Matrigel-coated plates in EBM2 (Lonza). On day 5, cells were dissociated with either Accutase for 10 min at 37°C or TrypLE for 3 min at 37°C. Once dissociated, cells for either protocol were centrifuged and resuspended in PEF before filtering through a 35µM filter. Anti-CD144 or anti-CD34 fluorescently conjugated antibodies (both at 1:100 from BD) were added to the cell suspension and incubated on ice for 30 min. Cells were washed once with PEF and filtered again before adding 7AAD (1:50). Cells were collected in EBM2 media and kept on ice until being plated onto Matrigel-coated plates in EBM2.

3.2.4 qRT-PCR

When indicated, cells were activated with 50 ng/mL TNF α for 24 hours prior to cell collection. Endothelial cells were collected for qRT-PCR using TrypLE at 37°C for 3-5 min. The cells were centrifuged at 1350rpm for 5 min and then frozen at -30°C until RNA extraction. RNA was extracted (Norgen) and made into cDNA (Bioline) following the protocols provided. Primers used for qRT-PCR are listed in **Table 3**.

3.2.5 Flow cytometry

Endothelial cells were disassociated using TrypLE at 37°C for 3-5 min before being centrifuged at 1350 rpm for 5 min. Cells were resuspended in PEF and filtered through a 35µM filter. Antibodies used included CD34 (BD Biosciences, 1:100) and CD31 (BD Biosciences, 1:100). Cells were incubated with antibody at room temperature for 30 min before being washed once with PEF. 7AAD (BD Biosciences) was then added at 1:50. Flow cytometry was performed on the LSRII and analysis was performed with FlowJo v10.

3.2.6 Tube formation assay

Endothelial cells were disassociated with TrypLE at 37°C for 3-5 min and then counted. 50,000 endothelial cells were used per well of a 48-well plate. Each well was pre-coated with Matrigel (diluted 1:2) and left at room temperature for 10 min before being placed at 37°C for 30 min to solidify. 50,000 cells were plated in 250µL of EBM2 total volume and left undisturbed at 37°C for 24 hours before being imaged. Tube formation was quantified based on number of tubes and nodes per 1000 cells plated.

3.2.7 AcLDL uptake assay

Endothelial cells were incubated with 10µg/mL dil-AcLDL (Life Technologies) for 4 hours at 37°C. The cells were washed twice with DPBS and then fixed using 4% PFA at room temperature for 10 min.

3.2.8 Nitric oxide production

Once confluent, endothelial cells were treated with 10 μ M DAF-FM (Cayman Chemicals) in EBM2 for 1 hour to allow for dye-loading. After 1 hour, the cells were gently washed 2 times with DPBS and incubated with 1mM L-arginine for 5 min. Cells were disassociated with TrypLE and collected. The cells were resuspended in PEF and passed through a 35 μ M filter before performing flow cytometry to measure mean intensity fluorescence. Flow cytometry was performed on the LSRII and analysis done on FlowJo.

3.2.9 Data Analysis

Immunofluorescence images were compiled using ImageJ. Brightness was adjusted when needed to improve the ability to see the staining in the figures. Data are expressed as mean \pm SD. Analysis was performed using t-tests and two-way ANOVAs as indicated with $p < 0.05$ as statistically significant.

3.3 RESULTS

3.3.1 Endothelial cell production with Lian protocol is lower than reported

In the original publication, Lian, *et al.* reported that this protocol resulted in ~60% of the cells co-expressing CD34 and CD31 on day 5 of the protocol¹⁸¹. Initially, modifications were made to account for media differences – MEF-CM was used rather than

mTeSR1, and L-glutamine rather than GlutaMAX – to use the protocol with the rest of the parameters remaining the same (**Figure 1a**). The authors of the original manuscript tested two GSK3 β antagonists – BIO and CHIR. Similar to their findings, the efficiency of deriving CD34⁺ cells was drastically increased with CHIR over BIO (**Figure 1b**). This increase was likely in part due to an increase in the viable population at the time of sorting (**Figure 1c**). In the original report, Lian, *et al.* reported that they were able to obtain up to ~60% CD34⁺CD31⁺ cells¹⁸¹; however, we were only able to get ~0.1%. Although the yield of endothelial cells was lower than reported, the Lian protocol presented an opportunity for further optimization to increase the yield of endothelial cells.

3.3.3 Optimization of previously published protocols increases yield of endothelial progenitor cells

In our study, the yield of CD34⁺ cells was significantly lower than the reported yield. To improve the efficiency, we set out to optimize the protocol. The published protocol uses GSK3 β antagonism to promote mesodermal differentiation. Through their own protocol optimization, they found that the LaSR media was sufficient to promote endothelial specification following 48 hours of GSK3 β antagonism. As mentioned earlier, many other protocols used additional growth factors to promote endothelial specification from a mesodermal precursor. Although Lian, *et al.* did report that no significant difference was seen in their efficiency with the addition of VEGF¹⁸¹, we developed a scheme for optimization based on key regulators of endothelial specification (**Figure 2a**).

The timing and concentration of growth factors added to the differentiation media varied across each condition to determine the optimal method. The differentiations were carried out in parallel using the same iPSC line, in the same plate, and with the same basal media to rule out any confounding factors. On day 5, flow was performed on the wells to assess the yield of CD34⁺ cells. All methods tested increased the percentage of 7AAD⁻CD34⁺ cells (**Table 4**), although we did observe a modest decrease in the viability compared to the original protocol. The addition of Activin A (7 ng/mL), BMP4 (20 ng/mL) and VEGF (25 ng/mL) for 3 days following CHIR treatment resulted in the least amount of cell death and the highest percentage of 7AAD⁻CD34⁺ cells, with roughly a 10-fold increase from the original protocol (**Figure 2b-c**).

3.3.4 Endothelial cells obtained from optimized protocol are more mature

To validate our optimized protocol, we used it to differentiate four iPSC lines derived from PBMCs. When compared to the original protocol, the new protocol consistently resulted in a higher differentiation efficiency, with no significant difference across the iPSC lines tested (**Figure 2c**).

The protocol indicates that the cells are to be sorted based on CD34 expression. CD34 marks progenitor cells of multiple cell types, including those of smooth muscle and hematopoietic origin. However, the publication by Lian, *et al.* demonstrates that the CD34⁺ cells lose their CD34 expression as they gain CD31 over passages in endothelial-specific media¹⁸¹. To further test our optimized protocol, we looked at whether the cells sorted on day 5 also expressed the more mature endothelial marker CD31. We repeatedly saw that

when compared to the original protocol, our optimized protocol yielded more cells that co-expressed CD34 and CD31 (**Figure 2d**), suggesting that the cells are more committed to the endothelial lineage upon sorting than those derived from the original protocol.

3.3.5 Newly established endothelial differentiation protocol produces cells that are functional

Surface marker expression alone is not sufficient to demonstrate functionality or maturity of cells. To confirm that the cells we were isolating through the differentiation protocol were in fact endothelial cells, we performed *in vitro* functional assays and compared their function to HUVECs, a “gold-standard” for endothelial assays. The iPSC-derived endothelial cells showed that they were able to form tubes, take-up AcLDL, and express vWF and VE-cadherin to a similar extent as HUVECs (**Figure 3a-d**).

After several passages, the iPSC-derived endothelial cells appeared to lose their CD34 expression, while more cells appeared to gain CD31 expression (**Figure 3e**), suggesting that as was reported, continuous culture in endothelial conditions promotes further maturation of the endothelial cells.

Two main functions of endothelial cells *in vivo* are their ability to produce vasoactive compounds and maintain an anti-adhesive barrier during basal conditions but respond adequately in times of inflammation to allow for proper tissue repair. These functions were tested in the iPSC-derived endothelial cells through a NO production assay and through measuring adhesion marker expression. The iPSC-derived endothelial cells produced NO to a similar extent as HUVECs (**Figure 3f**). Under both conditions, the

expression of VCAM and ICAM in the iPSC-derived endothelial cells was similar to that of HUVECs (2-way ANOVA, not significant; $p > 0.05$) and both cell types follow the increase in VCAM and ICAM transcript expression following TNF α treatment (Figures 3g-h, t-test; * $p = 0.029$, ** $p = 0.008628$, *** $p = 0.000697$) although the increase in ICAM for the iPSC-endothelial cells did not reach significance ($p = 0.06$).

Taken together, the *in vitro* functional assays suggest that the optimized protocol developed from the original protocol by Lian, *et al.* produced functionally mature endothelial cells that are comparable to HUVECs.

3.4 DISCUSSION AND FUTURE DIRECTIONS

The role that endothelial cells play in disease development as well as in tissue repair and regeneration distinguishes them as a cell type of high interest in drug discovery and regenerative medicine applications. Efficiently deriving these cells from hPSCs is therefore of utmost importance. Here, I've presented an optimization method that could be applied to any published protocol in order to improve the efficiency of endothelial cell generation. The protocol that I generated increased the yield of endothelial cells up to 10-fold from the original protocol when used in our culture system, and the resulting iPSC-derived endothelial cells were more mature in surface marker expression than their counterparts derived from the original protocol. Generating endothelial cells that were more mature at the time of sorting will ultimately decrease the time in which the cells need to be cultured before they could be used for downstream applications. This is beneficial from a time and financial perspective, but also, like any other somatic cell, the chance of mutation and

senescence of the endothelial cells increases with extended passaging. By developing a differentiation protocol that produces more mature endothelial cells at the same time point as the original, we have decreased the chance for multiple mutations as less expansion is needed to acquire enough mature cells for assays.

Compared to HUVECs, the iPSC-derived endothelial cells generated from this protocol are functional as measured through the assays tested. For a long time, HUVECs have been considered the gold standard for *in vitro* endothelial assays. Their utility for disease modelling and regenerative medicine applications is limited, as previously mentioned; however, their ease of isolation and culturing maintain them as a useful tool for verifying the function of *in vitro* derived endothelial cells.

Many *in vitro* differentiation methods aim to mimic the *in vivo* developmental path for generating that cell type. For endothelial cells specifically, mesodermal specification is therefore often the initial step in protocols. Here, GSK3 β antagonism through the use of CHIR directs the cells to a mesoderm progenitor. From there, Activin A and BMP4 act together to induce specification towards a more hemangioblast-like progenitor population^{250,271,272}. The hemangioblast population *in vivo* is believed to be the source of both endothelial and hematopoietic cells for the developing embryo. To push this population towards a more endothelial-like cell, VEGF is also used in the differentiation cocktail. Through the addition of these growth factors to the original protocol provided by Lian, *et al.* we are guiding the mesodermal cells towards a hemangioblast population and from there, endothelial specification.

Many factors can impact the differentiation efficiency of hPSCs towards a certain lineage. Some iPSC lines have an increased ability to form cells of one germ layer over the

other, for reasons that are still not well known but are likely a result of epigenetic differences between iPSC lines. Because of this, often researchers will move forward with lines that generate larger numbers of their desired cell. This could then impact the reported levels of differentiation, which may not be achievable with all iPSC lines. The protocol I developed here worked to similar extents in all iPSC lines tested, none of which were preferentially chosen for this ability.

The protocol outlined here for differentiating the iPSCs to endothelial cells was an improvement from the original protocol but was not 100% efficient. The marker used to isolate the cells, CD34, is also expressed on multiple cell types which suggests that even by isolating the CD34⁺ cells, they may not all be endothelial precursors. This is especially true for the protocol used here as the publication by Lian, *et al.* highlighted the bipotent capacity of the CD34⁺ cells¹⁸¹. Culturing the isolated cells in endothelial-specific media was the method that we used as per the published protocol's to further push the isolated progenitors towards the endothelial lineage. This method worked as seen through the CD31/CD34 flow that we performed; however, it was not done for each round of differentiation at every passage. This means that the population of cells may have been heterogeneous at times which may have impacted the read out of the assays. Using a combination of markers, such as CD34 and CD31 or other endothelial-specific markers such as CD144 at the time of isolation may partially mitigate this issue. Further optimization of the protocol may be done by increasing the time prior to isolation in the presence of a TGF β inhibitor, such as SB431542. Inhibition of the TGF β pathway promotes endothelial proliferation and has been used by other groups to a similar end²⁵⁴. By combining the methods used by other groups to expand the size and number of a mature

endothelial population prior to isolation with the optimized protocol here, a more homogeneous and mature endothelial population can be isolated. Additionally, a thorough time course analysis on the cells following isolation should be performed to better understand the maturation process that occurs *in vitro*. In doing this, additional factors that may be required to promote overall maturation, and in turn result in a more homogeneous cell population can be identified and adopted in culturing techniques.

The trends that we observed with the HUVECs were mimicked by the iPSC-derived endothelial cells; however, the reaction to inflammatory stimuli, as seen through changes to adhesion marker transcript levels, was not as robust. For ICAM specifically, although there was a noticeable increase, it did not reach significance likely because half of the replicates showed no real change in expression. This inconsistency is something that should be addressed in future work. In this work, TNF α stimulation was done for 24 hours²¹³; however, the transcript levels of both ICAM and VCAM peak at different times (12 hours and 6 hours respectively), although they have been shown to persist for ~72 hours²⁷³. As such, a more robust reaction could be caught following a shorter length of stimulation. The variation in marker expression may also be in part due to a heterogenous endothelial cell subtype population as they respond differently to the same stimuli in terms of their expression of adhesion markers²⁷³. Testing additional adhesion markers, such as those of the selectin family, or additionally testing cytokine production following stimulation would also provide additional metrics whereby the cells could be compared. Overall, by adopting a more endothelial-specific isolation strategy, as mentioned above, and adding other TNF α -mediated responses would allow a more thorough comparison of the cells.

As the field moves toward more clinical applications, using clearly defined media is important. At the time of this protocol generation, mTeSR1 was not widely used in the field. As such, modifying the protocol to accommodate for iPSC culture conditions (growing on iMEFs), meant that using MEF-CM allowed us to improve viability upon differentiation initiation without having to condition the cells to grow in mTeSR1. Applying the optimized protocol to iPSCs culture in mTeSR1 may also result in higher levels of CD34⁺ cell generation making this protocol applicable to more defined culture conditions as well.

To fully appreciate which disease model systems these iPSC-derived endothelial cells should be used for, additional analysis should be conducted. Although the surface marker expression does indicate that these cells have committed to the endothelial lineage and they are functionally mature, more analysis should be done. Performing single cell RNA sequencing on these cells would allow us to identify which endothelial subtype most of the cells belong to and therefore, what diseases they would be best suited to model. The distribution of endothelial subtype generation by differentiation protocols varies greatly and usually results in generation of both venous and arterial subtypes. In this case, the original publication of the protocol did not investigate the expression of subtype-specific markers and neither did we. By identifying the endothelial subtype(s) present in the cells generated through differentiation, we may be able to better understand the discrepancies between the iPSC-derived endothelial cell phenotypes and those seen by the HUVECs. Because HUVECs are venous in nature, using an additional primary endothelial line that represents another subtype, such as a human arterial line, may strengthen the validation of the endothelial cells as well as act as an additional control. This could be especially useful

for situations where the exact subtype distribution is unknown, as well as for further confirmation that one subtype was generated over the other if a more directed differentiation protocol is utilized. In the same way that Zhang, *et al.* used a reporter assay to determine the requirements for arterial specification, a similar method could be used for venous or lymphatic endothelial cells as well and be used to supplement protocols to better guide differentiation to one specific lineage²¹⁸. As a whole, more effort should also be made to distinguish endothelial subtypes by their tissue of origin. By examining the transcriptomic profile of endothelial cells from different tissues in the body we may be able to understand how their final stages of specification occur and mimic this process *in vitro*. This would not only present huge opportunities for disease modelling and drug screening platforms, but the advancement in tissue repair and organ regeneration applications would be unprecedented. For endothelial cells specifically, it has been demonstrated that those generated through *in vitro* differentiation of iPSCs have a similar immune response to those from a primary graft²⁷⁴. Although more work needs to be done to determine the specifics of this process²⁴, this work serves as an important proof of principle that endothelial allogenic grafts are feasible thus further highlighting the importance in generating more specialized differentiation protocols.

Blood vessels are comprised of several cell types that are all necessary for proper function. One of the major limitations of the method developed here, as well as in many published protocols, is the culturing of endothelial cells alone. *In vivo*, endothelial cells are one of the few cell types that do exist essentially only as a monolayer. In this respect, culturing them *in vitro* in a 2D system doesn't alter their normal configuration; however, developing more 3D models for disease modelling is important. Recently, Wimmer, *et al.*

(2019) showed a method to create blood vessel organoids¹³². These organoids are composed of both endothelial cells and supportive pericytes. The organoids were transplanted into mice and used to study diabetic vasculopathy. In this system, not only are the endothelial cells in a more realistic configuration on their own, but having the interaction with pericytes makes the system significantly more clinically relevant¹³². Adopting this technique for other model systems is the next step in making *in vitro* work more translatable.

Researchers interested in many other cell types have also adopted the organoid method for making their systems more biologically relevant. At the moment, a huge limitation in this field is the vascularization of organoids which ultimately limits their size and utility for fully recapitulating human organ function. Taking advantage of information that could be gathered from single cell RNA sequencing of developing endothelial cells, compared to primary endothelial cells from certain tissues, we could ultimately drive their generation. Incorporating the tissue-specific endothelial cells into an organoid system could improve their function and further the regenerative medicine field as a whole.

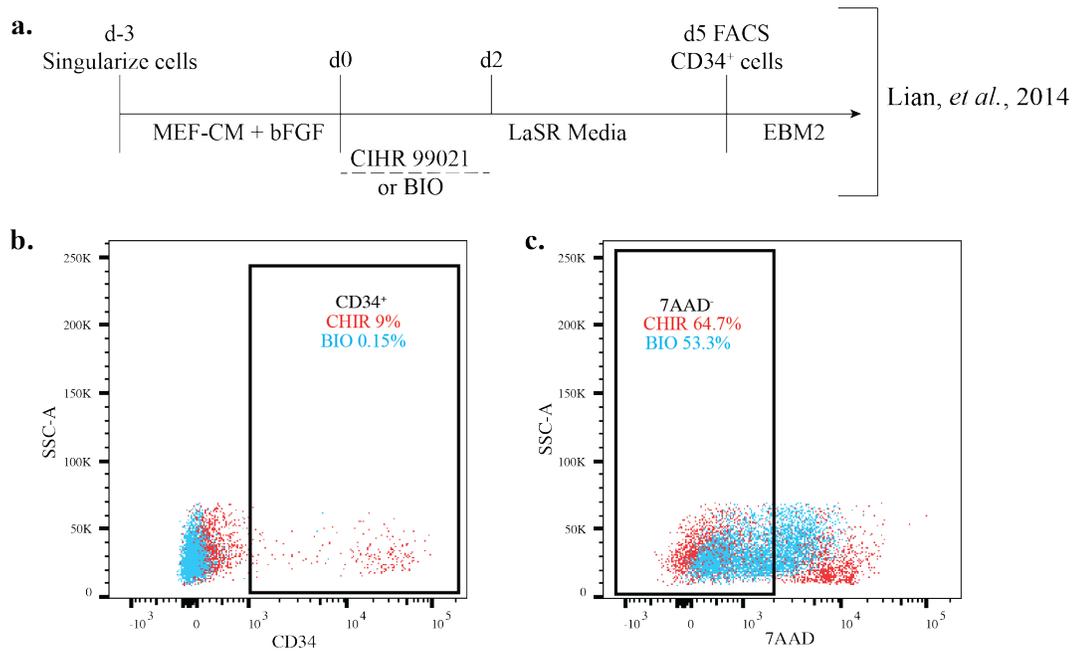


Figure 1. Using CHIR over BIO for GSK3 β antagonism improves yield of CD34+ cells and viability. The method used by Lian, *et al.*¹⁸¹ to differentiate iPSCs to endothelial cells was adopted in our lab by substituting the use of mTeSR1 for MEF-CM (a). In testing which GSK3 β antagonist to use to induce mesoderm differentiation, we saw an increase in both CD34+ cell production (b) and viability (c) with the use of CHIR 99021 over BIO and was therefore used for all subsequent experiments.

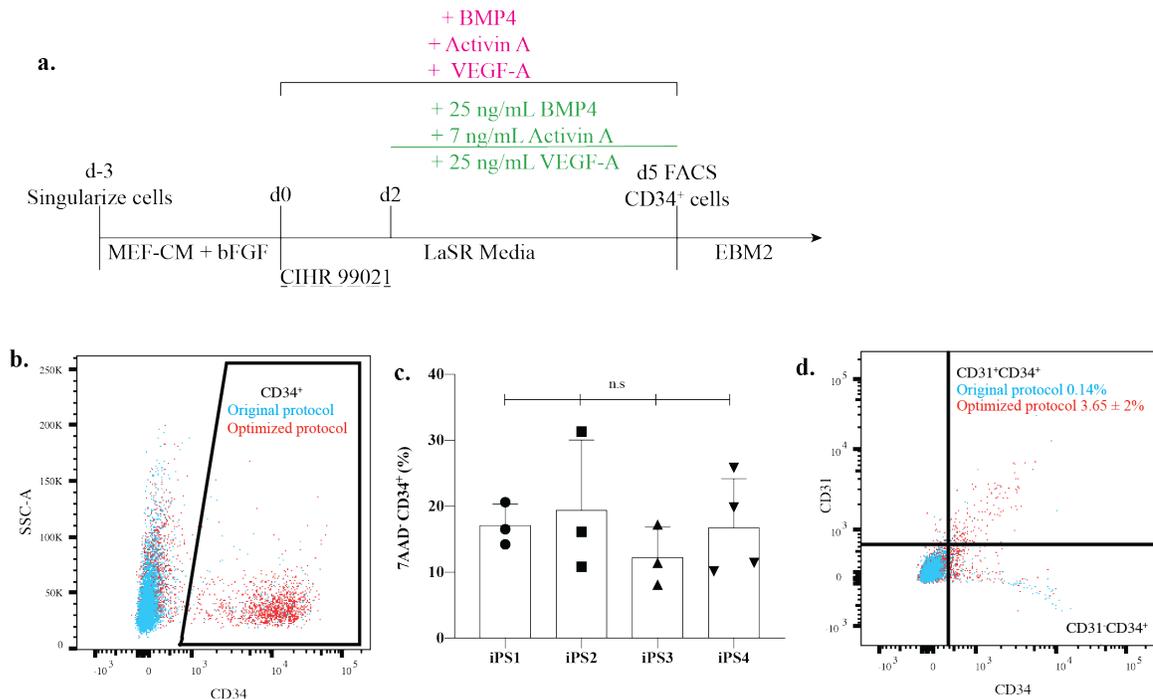


Figure 2. Addition of common endothelial growth factors improves differentiation efficiency and expression of mature endothelial markers by day 5. The original protocol presented by Lian, *et al.*¹⁸¹ was optimized through the addition of BMP4, Activin A, and VEGF-A at different concentrations and for different time periods during the differentiation process (**a**). The optimized protocol (**a**, in green), dramatically increased the yield of CD34⁺ cells by day 5 (**b**) and worked consistently across multiple iPSC lines tested, with no difference between the lines (**c**, n.s. p>0.05). CD34 marks an endothelial progenitor population, while CD31 expression is found on more mature endothelial cells. With the optimized protocol, the CD31⁺CD34⁺ population was more than 10-fold higher than the original protocol (**d**, value represents mean ± SD).

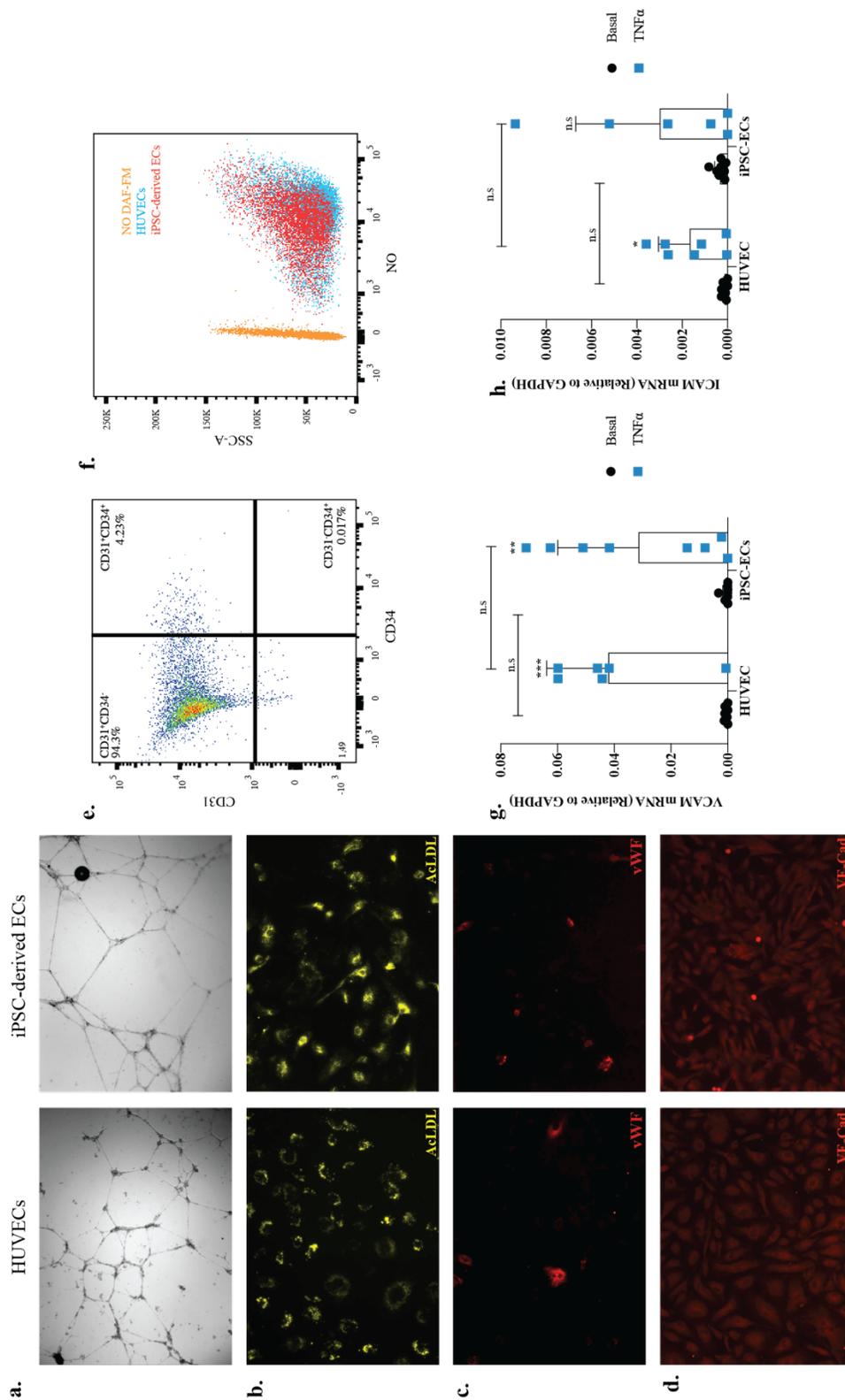


Figure 3. Endothelial cells generated with new optimized protocol are functionally mature. To confirm functional competence, iPSC-derived endothelial cells were compared to HUVECs for functional assays. The iPSC-derived endothelial cells were able to form tube networks (a), take up AcLDL (b), and expressed vWF (c) and VE-Cadherin (d) to a similar extent as HUVECs. As the iPSC-derived endothelial cells remained in culture, the cells were found to mostly express CD31 and not the progenitor marker CD34 (e). The iPSC-derived endothelial cells were also able to produce NO at a similar level as HUVECs as measured through the use of DAF-FM (f). The iPSC-derived endothelial cells also responded similarly to HUVECs (no difference between samples, 2-way ANOVA $p > 0.05$) in response to the inflammatory cytokine $\text{TNF}\alpha$ in their ability to upregulate the adhesion markers VCAM (f, t-test; *** $p = 0.000697$, ** $p = 0.008629$) and ICAM (g, t-test; * $p = 0.029$, n.s $p = 0.06$).

Table 1. Summary of reported endothelial differentiation protocols efficiencies

Source of PSCs	2D or 3D	Length of Protocol	Reported Efficiency	Reference
mESCs	3D	11 days	73% \pm 3% CD31 ⁺	Vittet, <i>et al.</i> ¹⁸⁶
Mesenchymal stem cells	3D	7 days	Not specified	Oswald, <i>et al.</i> ²⁷⁵
hESCs	3D	13 days	78.3% CD31 ⁺	Levenberg, <i>et al.</i> ²⁵⁰
hESCs	3D	10 days	Not specified	Wang, <i>et al.</i> ²⁴⁸
hESCs	3D	10 days	10-15% CD34 ⁺	Ferreira, <i>et al.</i> ²⁵¹
hESCs	2D	10 days	9.6% CD34 ⁺	Wang, <i>et al.</i> ²⁵³
hESCs Human iPSCs	2D	8 days	~3 – 6% CD31 ⁺ CD43 ⁻	Choi, <i>et al.</i> ²⁷⁶
hESCs Human iPSCs	3D + 2D	14 days	2 – 6% CD31 ⁺	James, <i>et al.</i> ²⁵⁴
Human iPSCs	3D	10 days	14 – 22% CD31 ⁺ CD144 ⁺	Adams, <i>et al.</i> ⁸⁸
hESCs Human iPSCs	2D	8 days	58% CD34 ⁺	Lian, <i>et al.</i> ¹⁸¹
hESCs Human iPSCs	2D	6 days	64% CD31 ⁺ CD144 ⁺	Sahara, <i>et al.</i> ²⁵⁵
Human iPSCs	3D	14 days	44% CD31 ⁺	Zhang, <i>et al.</i> ²⁵⁸
hESCs Human iPSCs	2D	5 days	~88 – 93% CD31 ⁺	Wu, <i>et al.</i> ²⁷⁷
hESCs Human iPSCs	2D	6 days	~61 – 88% CD144 ⁺	Patsch, <i>et al.</i> ²⁵⁶

hESCs	2D	7 days	62% VEGFR2 ⁺ CD31 ⁺ CD144 ⁺	Nguyen, <i>et al.</i> ²⁵⁷
mESCS hESCs	2D	8 days	38% CD31 ⁺ CD144 ⁺	Tsang, <i>et al.</i> ²⁵⁹
hESCs Human iPSCs	2D	6 days	95.5% CD31 ⁺ CD144 ⁺	Zhang, <i>et al.</i> ²¹⁸

Table 2. Endothelial-specific markers used for isolating and validating* function of hPSC-derived endothelial cells

Platelet Endothelial Cell Adhesion Molecule 1 (PECAM1) CD31	Surface marker	Intercellular junction component
Vascular Endothelial Cadherin (VE-Cad) CD144	Surface marker	Intercellular junction component
CD34	Surface marker	Expressed on endothelial progenitors
Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) KDR Flk-1	Surface marker	Receptor for VEGF-A
von Willebrand Factor (vWF)	Surface marker	Glycoprotein produced and stored in Weibel-Palade bodies of mature endothelial cells
Dil-Ac-LDL	Uptake requirement	Scavenger receptor pathway uptake of AcLDL is restricted to macrophages and endothelial cells

*The thick border separates those used for isolating and validating from those used strictly for validation.

Table 3. qRT-PCR primers

Gene	Forward Primer 5' → 3'	Reverse Primer 5' → 3'
GAPDH	TCCCTGAGCTGAACGGGAAG	GGAGGAGTGGGTGTCGCTGT
VCAM	GTCTCCAATCTGAGCAGCAA	TGGGAAAAACAGAAAAGAGGTG
ICAM	AGGGTAAGGTTCTTGCCAC	TGCTATTCAAACCTGCCCTGA

Table 4. Optimization of Lian, et al.¹⁸¹ protocol parameters and output

Condition	7AAD ⁻ (%)	7AAD ⁻ CD34 ⁺ (%)
Original protocol	97.0	1.97
d0 - d5: low BMP4, low Activin A, low VEGF	75.0	6.62
d0 - d5: med BMP4, med Activin A, med VEGF	66.1	13.5
d0 - d5: high BMP4, high Activin A, high VEGF	71.7	17.1
d1 – d5: low BMP4 d2 – d4: low Activin A d2 – d5: low VEGF	75.2	5.02
d1 – d5: high BMP4 d2 – d4: high Activin A d2 – d5: high VEGF	73.8	14.4
d1 – d2: med BMP4, med Activin A, med VEGF	78.3	7.15
d3 – d5: med BMP4, med Activin A, med VEGF	83.9	27.7
d0 – d5: low BMP4, low VEGF	77.0	14.8
d0 – d5: med BMP4, med VEGF	77.7	17.8
d0 – d5: high BMP4, high VEGF	82.6	10.2

*Concentrations used can be found in Chapter 3.2 Materials and Methods.

CHAPTER 4: RARE DHX34 VARIANTS ALTER ENDOTHELIAL FUNCTION IN THE CONTEXT OF EARLY-ONSET CORONARY ARTERY DISEASE

PREFACE

This chapter focuses on the work I've done to elucidate the role of DHX34 in endothelial cells and how this translates to rare variants identified in patients enrolled in the DECODE study. The idea for the project was generated by Dr. Guillaume Paré and although the work done in this chapter was primarily generated by me, it would not have been possible without the collaboration with the Paré lab. The exome sequencing data was analyzed by Ricky Lali. The western blots and analysis were done by myself and Shreya Jain. The iPSCs used are those that were generated and validated with methods shown in Chapter 2. The endothelial cells were generated with the protocol highlighted in Chapter 3. Some of the endothelial-specific methods used in this chapter have been previously outlined in Chapter 3 as well but for ease of read, they have been repeated in this chapter's Materials and Methods (4.2) section. The bar graphs in this chapter display all individual replicates within the graphs. For the qRT-PCR and tube formation data, all replicates shown are the mean of 3 technical replicates. This is also true for the data shown in the scratch assay, although individual replicate points are not shown in those graphs. For all assays, each replicate was derived from the same iPSC line but through a different round of differentiation.

4.1 INTRODUCTION

4.1.1 Role of endothelial cells in coronary artery disease development

Coronary artery disease (CAD) is a complex disease that develops slowly over several decades through the concerted effort of many cell types (covered in Chapter 1.2.1). The individual role of the cell types in CAD development has been elucidated through both *in vivo* and *in vitro* work. The endothelium, the innermost section of the arterial wall, is involved in all stages of CAD development^{82,190}. Initial work to demonstrate its general importance in CAD development was done by showing that endothelial denudation in coronary vessels is associated with advanced atherosclerotic lesion formation in pigs²⁷⁸, and further, endothelial damage and loss is a common feature of human CAD¹²⁹. Since then, many groups have illustrated the role that proper endothelial function plays in all stages of CAD development from priming an area for plaque development to the rupture of the plaque.

Endothelial cells can “prime” an area of the arterial network for CAD development²⁷⁹. Human CAD tends to develop in specific areas of the arteries, specifically those of disturbed blood flow. This nonlaminar flow is detected by the endothelial cells through a mechanosensory complex comprised of CD31, CD144, and vascular endothelial growth factor receptor 2 (VEGFR2)²⁸⁰. The force of the blood flow is detected by CD31, which signals through CD144 to VEGFR2, ultimately activating phosphoinositide 3-kinase (PI3K). This signal transduction results in the phosphorylation of the inhibitor of $\text{Nf-}\kappa\text{B}$, I κ B α , and thereby allowing for the translocation of the p65 and p50 subunits of $\text{Nf-}\kappa\text{B}$ to

the nucleus^{86,280}. Once in the nucleus, these subunits activate the transcription of pro-atherogenic genes such as the adhesion markers ICAM and VCAM, and cytokines such as IL-6, and several matrix metalloproteinases (MMPs)⁸⁷. In 2000, Hajra, *et al.* showed that in mice, the endothelium of areas subjected to disturbed flow had an increase in the extent of nuclear p65 and phosphorylated I κ B α compared to those in areas of laminar flow²⁸¹. In areas that are at a lower risk for plaque development, the Nf- κ B pathway is downregulated once they are acclimatized to the flow; however, this is not what appears to happen in atheroprone regions of the vasculature. In this way, the endothelium dictates the probability of plaque development by “priming” a certain area with nuclear Nf- κ B accumulation.

During initiation of CAD, the endothelium is the first to experience functional and phenotypic changes. Endothelial dysfunction is a hallmark of CAD development and is characterized by an increase in endothelial permeability, and a decrease in nitric oxide (NO) bioavailability^{79,82}. The decrease in NO levels can be a result of a number of factors including uncoupling of endothelial nitric oxide synthase (eNOS), and/or an increase in reactive oxygen species (ROS)^{80,84,85,282}. In healthy patients, an influx of acetylcholine, which is an activator of eNOS, results in arterial dilation; however, this is lost in CAD patients because of endothelial dysfunction⁸². NO is the most potent vasodilator, but it also has an important role in inhibiting Nf- κ B signaling⁸¹. With the reduction of NO, there is a concomitant increase in nuclear Nf- κ B. The influx of p65 to the nucleus results in endothelial activation as characterized by the increased level of adhesion marker expression and cytokine production^{84,283}. The combined effect of endothelial dysfunction and activation is the recruitment of circulating leukocytes to an area that is more adhesive and permeable to the cells thereby increasing the risk of CAD development. In primed

endothelial cells, this response may occur more rapidly as Nf- κ B can already be found to some extent in the nucleus.

The upregulation of inflammatory genes by Nf- κ B in the endothelium can be interpreted as a normal response to injury and to a certain extent is self-regulating. Nf- κ B not only regulates the transcription of pro-atherogenic genes, but also those that oppose the process such as IL-10 and tissue inhibitors of MMPs (TIMP-1)⁸⁷. It also increases transcription of I κ B α to create a negative feedback loop and prevent continuous Nf- κ B signaling⁸¹. In CAD, the injury that the endothelium experiences, whether through environmental or genetic factors, is chronic. In these cases, even though inhibitors of Nf- κ B signaling may also be transcribed, the initial response is sustained and allows for the development of an atherosclerotic plaque.

As the plaque grows and CAD develops, the endothelium continues to play an important role. As the disease progresses, neovessels comprised almost exclusively of endothelial cells, appear within the plaque⁷⁹. These newly formed microvessels support the persistent influx of inflammatory cells to the sub-endothelium allowing for continued growth of the plaques⁸³. The microvessels further destabilize the plaque as they lack supportive cells such as pericytes and are therefore susceptible to leakage and subsequent hemorrhage⁹⁴.

Plaque rupture as a result of the loss of the fibrous cap is responsible for 55-65% of acute clinical symptoms in CAD patients⁹⁶. An atherosclerotic plaque tends to rupture in areas that are less stable – typically those that have a higher macrophage content and a weaker fibrous cap^{77,90}. When the plaque ruptures, the prothrombotic core is exposed to the circulating blood resulting in the formation of a thrombus. When the thrombus is large

enough, this can partially or fully occlude the artery and prevent adequate blood flow to the myocardium. Clinically, this can manifest as an acute coronary syndrome (ACS) which includes non-ST elevated myocardial infarction (NSTEMI), ST-elevated myocardial infarction (STEMI), or sudden cardiac death⁹⁷. The cap, composed primarily of smooth muscle cells, lies directly below the endothelium. The smooth muscle cells of the cap produce an extracellular matrix (ECM) composed of collagen, proteoglycans, elastin, and glycoproteins⁷⁷. This ECM can be degraded and remodelled, in part by endothelial cells, through the activity of MMPs and TIMPs. One particular MMP produced by endothelial cells is MMP-14. MMP-14 activates MMP-2, which degrades type IV collagen, thereby contributing to the weakening of the fibrous cap and rendering the plaque more susceptible to rupture⁹⁷. In normal conditions, the activity of MMPs can be regulated through TIMPs; however, in susceptible plaques, the balance of the two is disrupted, thereby promoting more degradation of the ECM and weakening the cap.

The second most common cause of clinical manifestation of CAD is erosion of the endothelium. Approximately 20-25% of terminal thrombosis events occur as a result of a denuded endothelium^{94,96}. Endothelial cells grow as a monolayer and as such are strictly regulated by contact inhibition. In normal vessel homeostasis, when endothelial cells are wounded or removed from the vessel wall, the cells closest to the wound stretch to fill the gap and then replicate to re-establish the endothelial barrier. In cases such as CAD where injury occurs chronically in the same area, after so many replications the endothelial cells reach a point of senescence. At this point, they are no longer capable of replicating to re-establish the monolayer. Although the cells farther from the wound itself may be capable of performing this function, however, due to contact inhibition they are not able to reach

the area which becomes denuded of the endothelium^{79,110}. If this occurs in CAD, the subendothelial ECM produced by smooth muscle cells is exposed to the circulation. The ECM, which is rich in collagen and acts as an ideal location for platelet adhesion, along with high amounts of tissue factor produced and sequestered in the inner core can initiate the coagulation cascade and result in thrombus formation^{105,129}. The correlation between the integrity of the endothelium and clinical symptom severity emphasizes the importance of endothelial cells in the late stages of the disease.

In many cases, the aforementioned changes to the endothelium are thought to be a result of the classic environmental risk factors associated with increased CAD risk, namely smoking, hypertension, hypercholesterolemia, obesity, and diabetes. Cigarette smoke produces a large amount of ROS throughout the body¹¹⁹. In the endothelium, the ROS can contribute to a decrease in NO, which as mentioned above, is a major component of endothelial dysfunction. Hypertension can further disrupt blood flow in areas that are atheroprone, thus initiating the pro-atherogenic signal cascade through the endothelial mechanosensory complex¹⁰². Hypercholesterolemia, obesity, and diabetes all act in part by increasing the relative abundance of LDL and triglycerides in circulation^{102,284}. High amounts of LDL can enter the arterial wall through the endothelium where it can be modified to oxidized LDL (oxLDL). OxLDL can contribute to the production of ROS and uncoupling of eNOS, thereby resulting in endothelial dysfunction and again, if chronic, can result in the formation of an atherosclerotic plaque⁸².

In the development of CAD, the normal function of the endothelium is disrupted. Typically, the endothelium serves as a non-adhesive barrier to the underlying vessel wall. However, with the upregulation of adhesion markers, cytokines, and procoagulant tissue

factor, the endothelium is transformed into a surface that allows for leukocyte attachment. As these changes in endothelial phenotype persist, they can also result in irregularities in endothelial organization and morphology in areas overlying the plaque, which negatively impacts barrier function⁸². Because of disrupted NO production, not only is the endothelium more adhesive, but the normal vasodilatory role of endothelial cells is also decreased. This results in a more constricted blood vessel, thereby impacting blood flow to the myocardium.

When taken together, it is clear that the endothelium plays a pivotal role in all areas of CAD progression. As such, it should be considered as a therapeutic target in disease prevention and treatment throughout all stages of CAD development. Because of the potential differences in the specific functions of endothelial cells across species and blood vessel types, it is imperative that we develop and utilize human endothelial models that most accurately replicate the *in vivo* environment of CAD development.

4.1.2 Genetic component of coronary artery disease

Epidemiological studies have long identified the important role of genetics in CAD development. In 1951, Gertler, Gran, and White were the first to note that an individual had an increased risk for having a myocardial infarction (MI), a common clinical symptom of CAD, if there was parental history of the condition²⁸⁵. Twin and familial studies of CAD were the first genetic studies to be conducted and were instrumental in determining that CAD heritability is ~40-50%²⁸⁶. In 2007, four independent groups identified the first specific locus to contribute to this heritability through GWAS²⁸⁷⁻²⁹⁰. The 9p21 locus was

found to be associated with an increased risk for CAD with MI, and still remains the locus that confers the largest increase in risk for the first MI event associated with CAD¹¹⁴. Despite its importance in CAD predisposition, it was only recently that this risk locus was found to control a wide number of genes that predispose vascular smooth muscle cells to pro-atherogenic phenotypes²⁹¹. Since the identification of the 9p21 risk locus, four large genome wide analysis studies (GWAS) consortia (CARDIoGRAM, MIGen, C4D Genetics, and UK BioBank) were formed and have been successful in identifying 163 loci which independently increase CAD risk¹¹⁴.

Most of the knowledge gathered on the genetic susceptibility factors for CAD has come from GWAS and meta-analyses of these studies. The patients used in these studies vary in disease severity and age; however, the resulting information has been pivotal in expanding our understanding of the overall disease etiology. GWAS analysis typically uses parameters that enrich for common variants, i.e. those with a minor allele frequency (MAF) of >5%, because they are believed to account for the majority of heritable risk for complex diseases²⁹². This idea was first proposed in 1970 by Morton, *et al.* who suggested that the majority of diseases caused by multiple genes are due to the small effect of many genes, and only a small number of cases are due to single genes with large effects²⁹³. This idea has been put into practice by calculating a polygenic risk score (PRS). A weighted PRS can be determined by summing the number of risk alleles that one individual carries, weighted by the impact of the particular variant¹¹⁴. Multiple groups have shown the potential PRS has to stratify CAD patients based on their risk²⁹⁴⁻³⁰⁰. In 2018, Theriault, *et al.* showed that a high PRS based on 182 single nucleotide polymorphisms (SNPs) from the CARDIoGRAMplusC4D Consortium was correlated with an elevated risk of CAD²⁹⁴.

In fact, in their sampling of an extreme population (males ≤ 40 and females ≤ 45), 1 in 53 had a PRS corresponding to an increase in risk similar to that of familial hypercholesterolemia (FH) patients. This rate is higher than the incidence of FH, highlighting the large role genetics play in the young CAD cohort²⁹⁴. In the same year, Khera, *et al.* performed a similar analysis whereby the PRS for CAD patients ≤ 55 was calculated²⁹⁵. Similarly, their work showed that high PRS scores were associated with a drastic increase in CAD risk. In this study, they dissected their results for each racial group tested and although all showed a similar correlation, the highest increase in risk was estimated for those of white European descent²⁹⁵. Similarly, Hajek, *et al.* found that a 46-SNP PRS for patients enrolled in the Multi-Ethnic Study of Atherosclerosis (MESA) was significantly associated with a cardiac event in white males, but not females²⁹⁶. Both of these latter studies suggest that while PRS may help to identify populations that may benefit the most from preventative treatments, the current library of SNPs may not be predictive for all populations. To rectify this difference, the field as a whole should strive to include a more diverse population of individuals in subsequent sequencing initiatives. As it stands, European Caucasians represent the largest group of sequenced individuals which biases the available genomic information²⁹⁵. Increasing the available data for other ethnic groups will provide a means by which a similar analysis can be conducted based on SNPs.

Many identified risk alleles have benefited the scientific community's understanding of the disease despite the fact that the majority of identified SNPs (~75%) have been in non-coding promoter/enhancer regions of the genome¹¹⁴. These loci have been annotated by the Encyclopedia of DNA Elements (ENCODE) project based on the proximity; however, it has been estimated that only ~27% of these SNPs actually interact

with the nearest gene promoter³⁰¹. This suggests that the genes listed may not be those that are actually associated with the elevated risk. In 2017, Gupta, *et al.* used fine mapping of the 6p24 risk locus, which confers an increased risk for vascular diseases including CAD, to identify the causal gene³⁰². This risk locus was annotated as being the phosphatase and actin regulatory protein 1 (*PHACTRI*) based on proximity because the prominent risk variants lie within the 3rd intron of *PHACTRI*. Despite this, the causal gene at this locus had not been definitively proven to be *PHACTRI*. Their group identified the SNP in the locus with the strongest association with CAD/MI and introduced it into healthy iPSCs using CRISPR/Cas9. The edited iPSCs were subsequently differentiated to endothelial cells, as vascular-specific effects were predicted to be associated with the SNPs as assessed by data provided by unbiased phenome-wide association studies. Through their work, they noted that the identified SNP is a distal regulator of endothelin-1, which is located 600 kb upstream of *PHACTRI*³⁰². Although numerous other groups had previously confirmed the association of the risk locus and *PHACTRI* expression, the work done by Gupta, *et al.* suggests that the SNP in question increased susceptibility to CAD through modulating endothelin-1 function, not *PHACTRI* expression. This study highlights the complexity of elucidating the mechanism through which non-coding variants work and underscores the possibility that in many cases, the genes that have been associated with the variants may not actually be those that impact disease progression. There is more work to be done to fully elucidate the causal role of the SNPs identified through GWAS.

Despite the amount of work that remains to be done, the information gathered thus far has already greatly impacted our understanding and treatment of CAD. The importance of regulating LDL and total cholesterol in preventing CAD, as seen through FH-associated

variants and those in LDL-C, has been elucidated by the effectiveness of statins as preventative and therapeutic medications. Currently, there are a number of clinical trials for more targeted therapies against identified causal genes such as *APOC3*, and *ANGPT3*³⁰³. Although therapies may originate through the identification of specific variants or risk factors possessed by a subset of the population, they do not have to be limited to those that carry the variants, as is seen in statin use^{113,114}. In this way, by furthering our understanding of a small subset of CAD risk loci, we stand to identify novel therapeutics that would benefit the CAD patient cohort as a whole, regardless of the variant penetrance.

4.1.3 DECODE project

Identifying the genetic target of the SNPs identified by GWAS is one of the major hurdles associated with the workflow required to elucidate their mechanism of action. One method used to identify risk variants that have a better-defined causal gene is through exome sequencing. The exome is the protein-coding region of the genome and therefore, the gene impacted by the variant is well-defined. Even though the exome only comprises ~1.5% of the human genome, it is predicted to contain 85% of all disease-causing mutations thus highlighting its importance³⁰⁴. *In silico* protein-predictive software can be helpful in predicting the impact of identified variants on protein structure and function but performing direct biological analysis on samples would allow for a more definitive output.

CAD GWAS have an upper age limit for patient enrollment, based on the principle that at an earlier age of onset, it is unlikely that the exposure to environmental factors has

been sufficient to induce disease phenotypes. This idea can be expanded to identify more deleterious variants by conducting extreme phenotype sampling (EPS), whereby individuals with a more extreme disease phenotype are recruited. Selecting patients through EPS enriches for causative variants^{305,306}. GWAS with large sample sizes can identify low-frequency (MAF<0.1%) variants; however, they typically identify SNPs that are common (MAF >5%)¹¹⁴. Although similar variants can be identified through EPS, because the disease phenotype is more severe in these patients, it is likely that rare variants that confer a larger risk per allele can be identified.

The Genetic Determinants of Early-Onset Coronary Artery Disease (DECODE) study aims to combine both exome sequencing with EPS to identify rare, causative variants for early-onset coronary artery disease (EOCAD). Previous studies have shown that heritability of CAD is ~50%; however, for EOCAD the heritability has been estimated to be >60% therefore exceeding the contribution of environmental risk factors⁷². As was mentioned above, a PRS based on common SNPs is much higher in EOCAD patients as compared to the general population²⁹⁴ underscoring the impact that genetic variants have on this cohort. This population stands as an ideal tool to elucidate risk variants for CAD development in general, especially those that confer a larger risk per allele. These variants were identified and screened by Dr. Paré's group and then used as the basis for further analysis through the use of patient-derived iPSCs to assess biological impact in the context of disease development.

Patients were enrolled in the DECODE study Hamilton Health Sciences Heart Investigation Unit and Lipid Clinic if they met the inclusion criteria:

1. Angiographically proven CAD with at least 70% stenosis in one coronary vessel at <40 years for male or <45 for females
2. Absence of co-morbidities including chronic kidney disease, type I diabetes mellitus, insulin-dependent type II diabetes mellitus, chronic hepatitis, HIV, vasculitis, systemic autoimmune disease, and/or chronic use of amphetamines or steroids

Patients that were successfully enrolled in the study provided peripheral blood samples for exome sequencing analysis, macrophage analysis, and biological assays as done through the use of iPSCs. The results of macrophage analysis are not discussed here. The exome sequencing analysis will be briefly covered below as it pertains to the biological work with iPSCs, which is the focus of this work.

4.1.4 Nonsense-mediated mRNA decay in cell biology and disease

Nonsense-mediated mRNA decay (NMD) is an evolutionarily conserved pathway originally believed to strictly degrade mRNAs that contain a premature termination codon (PTC). Through its regulation of PTC-containing mRNAs, NMD safeguards cells by preventing the translation of partially functional proteins that may impair normal cellular function. Approximately one third of all disease-causing genetic mutations result in the introduction of PTC, thereby targeting them for NMD (**Table 1**)³⁰⁷. A prime example of the role of NMD in disease pathology is β -thalassemia where patients harbour mutations in the β -globin gene that result in the production of truncated β -globin chains³⁰⁸. If produced, the truncated β -globin chains form toxic precipitations that result in the clinical

symptoms associated with β -thalassemia. In patients with the NMD-sensitive mutation, the PTC-containing transcript is degraded through NMD thereby preventing the production of truncated β -globin and therefore the aggregation of the dysfunctional protein³⁰⁹. Patients with this form of β -thalassemia tend to have milder symptoms than those with the NMD-insensitive mutation. Individuals that are heterozygous for these variants are typically asymptomatic because the unaffected allele can compensate for the truncated β -globin that is degraded through NMD³⁰⁹. The role of NMD in disease pathophysiology can also be seen in other cases where disease severity is impacted by the location of the PTC in the transcript (**Table 1**). The location of the PTC may impact the severity of the disease as it may impact the extent to which the protein is partially translated due to delays in initiating NMD thereby altering the phenotype. This is further evidenced by data that shows that preventing NMD activity may reduce a patient's symptoms by allowing for partial activity of the protein as is a current area of research for cystic fibrosis therapy³⁰⁸. In cystic fibrosis, some of the most common mutations result in a PTC in the cystic fibrosis transmembrane conductance receptor (CFTR) gene. By preventing NMD activity, the shortened CFTR transcript is translated and although it is not fully formed, in this case, partial restoration of the function is sufficient to alleviate symptoms in some patients³⁰⁸.

It was traditionally thought that NMD only regulated transcripts with PTCs; however, it is now widely accepted that NMD can regulate 3-20% of all naturally occurring cellular transcripts^{308,310,311}. The mechanism by which the transcripts are identified for degradation is still largely unknown although it may in part be based on structural elements of transcripts such as a long 3' untranslated region (UTR)³¹². NMD may also regulate the level of splice variants as alternative splicing of certain genes may result in the inclusion

of a PTC, thereby marking it for degradation through NMD. Together, this highlights the important function of NMD in transcriptional regulation.

There are multiple branches of NMD that differ on their dependence on the Up-Frameshift Suppressor (UPF) proteins, but all identified pathways ultimately result in the phosphorylation of UPF1. The full mechanism by which mRNAs are funnelled through different NMD pathways has not been elucidated but likely relies in part on the feature that targets it for degradation, i.e. PTC or 3'UTR³¹². The most heavily studied pathway of NMD is initiated by the presence of a PTC that is recognized during the initiation round of translation (**Figure 1**)³¹³. In mammalian translation, the ribosome pauses when it reaches a termination codon. The codon is recognized as a PTC because of the existence of an exon-junction complex (EJC) located 50-55 nucleotides downstream of the termination codon^{312,314}. EJCs are deposited during post-transcriptional mRNA processing and their presence downstream of a stop codon suggests that there are additional exons following the stop, demarcating it as a PTC. When the ribosome pauses, the eukaryotic release factors eRF1 and eRF3 are recruited^{308,314-316}. If the termination codon is not premature, eRF1 and eRF3 interact with the poly(A) tail binding proteins signaling the termination of translation^{308,312}. If it is a PTC, or there is a long 3'UTR, eRF1 and eRF3 are unable to bind to the poly(A) tail and instead recruit UPF1, DHX34, and the SMG1 complex (SMG1, SMG8, SMG9)³¹⁷. These proteins form the transient SURF complex (SMG1, UPF1, eRF1/3), which is bound to the stalled ribosome at the PTC. Once the SURF complex is formed, DHX34 catalyzes the hydrolysis of ATP to ADP, resulting in a conformational change that allows the ribosome, along with eRF1 and eRF3, to disassociate from the mRNA at the PTC^{317,318}. This new conformation, known as the DECID complex, favours

the phosphorylation of UPF1 by SMG1 with DHX34 acting as a scaffold for this pivotal step of NMD³¹⁸. Once UPF1 is phosphorylated, DHX34 disassociates from the complex. Phosphorylated UPF1 inhibits further translation of the mRNA and recruits SMG5, SMG6, and SMG7 to the mRNA. SMG6 is an endonuclease and cleaves near the PTC, creating an unstable intermediate lacking either a 5'cap or 3' poly(A) tail, making it susceptible to exonucleases³¹⁴. The SMG5-SMG7 complex is associated with the recruitment of deadenylase and decapping complexes such as CCR4-NOT and DCP2/DCP1 α . SMG5 and SMG6 also recruit phosphatase 2a to the NMD complex to dephosphorylate and recycle UPF1, thus regulating the extent of NMD occurring in the cell³¹⁴.

NMD is a tightly regulated cellular process and the disruption of this regulation has the potential to result in disease. In the mouse system, deletion of UPF1 is embryonic lethal suggesting it has a crucial role in development³¹⁹. However, UPF1 also functions in maintaining telomere length which likely plays a larger role in the embryonic lethality of this knockout over its role in NMD³⁰⁸. UPF3b is involved in certain NMD pathways as a means to connect the EJC with the DECID complex. It has been demonstrated that UPF3b and its paralog UPF3a have antagonistic functions whereby UPF3b activates NMD and UPF3a inhibits the process³²⁰, which is important during developmental processes such as differentiation of stem cells into somatic cells³¹⁰. Interestingly, variants in UPF3b have been associated with mental retardation and a number of psychiatric disorders³²¹, again demonstrating the importance of proper NMD function in developmental pathways. As mentioned above, SMG6 functions in the degradation aspect of NMD as compared to UPF1 and UPF3b which play a larger role in identifying the mRNA for degradation³¹². SNPs associated with SMG6 were identified through GWAS to increase risk of CAD

development³²². The exact mechanism through which the identified SNPs in SMG6 impact the progression of CAD or whether they alter SMG6 or another gene entirely, remains unknown. Nonetheless, the fact that variants in all stages of NMD have the potential to increase disease risk highlights the impact of this pathway in normal cellular function and disease prevention.

4.1.5 Chapter objectives

The role of NMD in CAD development has been suggested through the identification of SNPs in SMG6 through GWAS³²²; however, no cellular biology has been conducted to confirm this association. Through exome sequencing, Dr. Paré's lab has identified novel, rare variants present in proteins similarly involved in the NMD pathway. These rare variants are present in a more severe form of CAD and as such, it is possible that these changes have resulted in a more exaggerated cellular phenotype. By studying these variants we can determine the potential impact of NMD dysregulation in certain cellular compartments associated with CAD development. Specifically, because CAD is a disease of the arterial wall and the endothelium acts as the initial deterrent to plaque development, studying the role of NMD in the initial phase of the disease would contribute to the overall understanding of EOCAD etiology.

Animal models have been used extensively to model CAD and have contributed substantially to our knowledge of the disease; however, their utility in therapeutic discoveries is limited due to biological differences. To mitigate these differences, I used a system for creating an *in vitro* human model of EOCAD. DHX34 variants were enriched

in the EOCAD cohort and were selected for follow-up through the biological assays. In this chapter, I first examined the impact of reducing DHX34 in HUVECs through the use of a knockout cell pool generated by Synthego. I then used iPSCs derived from patients (Chapter 2) with rare variants present in DHX34 to determine whether they may result in changes in the endothelial compartment such that CAD may progress at a more accelerated rate. We used endothelial cells differentiated from the iPSCs (Chapter 3) to model the endothelium of the patients in question along with a number of biological and molecular assays to elucidate any differences in functionality.

4.2 MATERIALS AND METHODS

4.2.1 Cell culture

All iPSCs were maintained on iMEFs in media composed of DMEM/F12, 10% KO-SR, 10mM NEAA, 2mM L-glutamine, and 7mM β -mercaptoethanol, supplemented with 16ng/mL bFGF added fresh. iPSCs were passaged every 6-7 days by manual picking. Endothelial cells were maintained on Matrigel-coated plates in EBM2 media (Lonza) and passaged once confluent at a ratio of 1:3. DHX34 and UPF1 HUVEC KO pools were purchased from Synthego and were maintained in EBM2 media (Lonza) in tissue-culture treated plates. When confluent, they were also passed at a ratio of 1:3 or 1:4. When indicated, both iPSC-derived endothelial cells and HUVECs were treated with 50 ng/mL TNF α for 24 hours. THP1 cells (ATCC) were cultured in RMP1 1640 with 10% FBS with passaging occurring every 5-6 days at a ratio of 1:5.

4.2.2 Differentiating iPSCs into endothelial cells

After passage 10, iPSCs were used for differentiation towards an endothelial lineage using a modified version of the protocol published by Lian, *et al.*¹⁸¹, as outlined in Chapter 3. Briefly, prior to differentiation, iPSCs were passaged onto Matrigel-coated plates in MEF-CM supplemented with 8ng/mL bFGF. Once confluent, the cells were singularized with Accutase at 37°C for 5 min. MEF-CM supplemented with 8ng/mL bFGF and 5µM Y-27632 was added to the wells before seeding the cells in 12-well plates coated with Matrigel at a density of 50,000 cells/cm² (day -3). The cells were then cultured for two additional days in MEF-CM supplemented with 8ng/mL bFGF. On day 0, the media was changed to the DMEM/F12, 2.5mM L-glutamine, and 60µg/mL ascorbic acid (LaSR media) supplemented with 6µM CHIR. 48 hours later, the media was changed to LaSR supplemented with 7ng/mL Activin A, 20ng/mL BMP4, and 25ng/mL VEGF-A. Half of the media was changed every 24 hours for 3 days.

4.2.3 Isolation and growth of endothelial cells

On day 5 of differentiation, CD34⁺ cells were isolated using FACS. The cells were disassociated using TrypLE Express at 37°C for 3 min. The TrypLE was neutralized using DPBS, 1mM EDTA, 5% FBS (PEF). The cells were centrifuged at 1350 rpm for 3 min and resuspended in PEF prior to filtering through a 35µm cell strainers (Falcon Corning). The cells were stained with CD34 APC (BD Biosciences, 1:100) for 30 min on ice. The cells were washed once with PEF before the addition of 7AAD (BD Biosciences, 1:50). CD34⁺

cells were collected in EBM2 media (Lonza) and subsequently cultured on Matrigel-coated plates in EBM2 media.

4.2.4 qRT-PCR

Endothelial cells were collected for qRT-PCR using TrypLE at 37°C for 3-5 min. The cells were centrifuged at 1350rpm for 5 min and then frozen at -30°C until RNA extraction. RNA was extracted (Norgen) and made into cDNA (Bioline) following the protocols provided. Primers used for qRT-PCR are listed in **Table 2**.

4.2.5 Western blot

Proteins from endothelial cell lysates were run on an 8% SDS-PAGE gel at 100V for 30 min followed by 130V for 1 hour before being transferred onto a PVDF membrane. Membranes were blocked in TBST (50mM Tris-HCl pH 7.4, 150mM NaCl, 0.1% Tween-20) with either 5% skim-milk powder or 5% BSA at room temperature for 1 hour. Primary antibodies were added to 2.5% skim-milk powder in TBST and added to the membranes to incubate overnight at 4°C. Membranes were washed 3 times for 5 min each with TBST and then incubated at room temperature with horseradish-peroxidase-conjugated antibodies for 1 hour (polyclonal rabbit anti-mouse IgG H&L conjugated to HRP Abcam cat. 97046; polyclonal goat ant-rabbit IgG H&L conjugated to HRP Abcam cat. 6721). Membranes were then washed again 3 times for 5 min with TBST. Band intensity was detected with

chemiluminescence using the ImageLab software of the ChemiDoc MP (Bio Rad). All antibodies used and their concentrations are listed in **Table 3**.

4.2.6 Tube formation assay

Endothelial cells were disassociated with TrypLE at 37°C for 3-5 min and then counted. 50,000 endothelial cells were used per well of a 48-well plate. Each well was pre-coated with concentrated Matrigel and left at room temperature for 10 min before being placed at 37°C for 30 min to solidify. 50,000 cells were plated in 250µL of EBM2 per well and left undisturbed at 37°C for 24 hours before being imaged. Whole well images were taken and images were manually stitched together. Tube formation was quantified based on number of tubes per 1000 cells plated.

4.2.7 THP1 adhesion assay

Endothelial cells were seeded on Matrigel-coated plates and allowed to reach ~80% confluency in a 48 well plate. 500,000 THP1 cells were incubated with 10µM Calcein-AM (Abcam) or 300µM DAPI in 1mL of IMDM with 10% FBS for 1 hour at 37°C. Dye loading of the THP1 cells was stopped with the addition of 1mL cold IMDM with 10% FBS. Cells were centrifuged at 1500 rpm for 5 min and resuspended at 1×10^6 cells/mL. 1×10^4 to each well. The plate was placed at 37°C on a shaking incubator for 1 hour. Cells were gently washed twice to remove non-adherent THP1 cells with DPBS and then imaged. Adherent THP1 cells were counted using ImageJ.

4.2.7 Scratch assay

All scratch assays were done in 12 well Matrigel-coated plates in triplicate. Endothelial cells were cultured until 100% confluent. A scratch was made using a 5mL serological pipette down the centre of the well. The scratch was imaged at 3 locations/well at the time of the scratch, 8 hours, 12 hours, and 24 hours post-scratch in the same locations. The cells were kept at 37°C, 5% CO₂ in between images. Average scratch width was taken across the three images per/well and migration was assessed as the percentage of the scratch that remained at each time point.

4.2.8 Nitric oxide production

Once confluent, endothelial cells were treated with 10µM DAF-FM (Cayman Chemicals) in EBM2 for 1 hour to allow for dye-loading. After 1 hour, the cells were gently washed 2 times with DPBS and incubated with 1mM L-arginine for 5 min. Cells were disassociated with TrypLE and collected. The cells were resuspended in PEF and passed through a 35µM filter before performing flow cytometry to measure mean intensity fluorescence. Flow cytometry was performed on the LSRII and analysis done on FlowJo.

4.2.9 Data Analysis

Data are expressed as mean ± SD. Analysis was performed using t-tests and two-way ANOVAs as indicated with p<0.05 as statistically significant.

4.3 RESULTS

4.3.1 Reducing DHX34 in HUVECs alters tube formation and migration capacity

It was only recently that DHX34's role in NMD was identified. DHX34 was shown to be recruited to the SURF complex during the initial phase of NMD³¹⁷. In this complex, DHX34 hydrolyzes ATP to promote reconfiguration into the DECID complex (**Figure 2a**), where it acts as a scaffold to allow for the SMG1-mediated phosphorylation of UPF1³¹⁸. Since then, very little work has been done to elucidate how DHX34 may function in different arms of the pathway, and what its function, outside of NMD, in various cell types may be. To elucidate its specific role in endothelial cells, we ordered knockout (KO) cell pools from Synthego for both DHX34 and UPF1. The latter was ordered to serve as a control for inhibiting all NMD functions so that NMD-independent functions of DHX34 may be better identified.

The KO cell lines provided displayed no gross morphological differences from the wildtype (WT) HUVECS (**Figure 2b**) although the UPF1 KO cells did display slower growth compared to the other lines. The KO pools were generated with CRISPR/Cas9 and Synthego reported high levels of editing efficiency as confirmed through sequencing (**Supplementary Figure 1-2**). However, when DHX34 mRNA expression was measured in the three lines, there was only a slight decrease in mRNA expression in the DHX34 KO line (**Figure 2c**, $p=0.0528$) and no significant difference in the DHX34 protein expression as measured through western blot analysis (**Figure 2d**). Interestingly however, there was a

significant increase in the DHX34 mRNA level in the UPF1 line as compared to the WT control (**Figure 2c**, * $p=0.0145$).

We also measured the mRNA expression of UPF1 through qRT-PCR in the WT, DHX34 KO and UPF1 KO. Again, there was only a slight but insignificant decrease in the UPF1 KO pool ($p=0.1912$), but a significant decrease in the DHX34 KO sample (**Figure 2e**, ** $p=0.0072$).

As a means to determine if either of these lines had a difference in the NMD efficiency as a result of the gene editing, we measured phosphorylated UPF1 through western blot but saw no difference between the KO lines and the WT control (**Figure 2f**). Despite the fact that no differences in phosphorylated UPF1 could be detected, there was significantly less mRNA of the known NMD targets SMG1 (**Figure 2g**, * $p=0.0182$, ** $p=0.0095$) as well as ATF3 and GADD45B (**Figure 2h**, * $p<0.05$, ** $p=0.0012$, *** $p=0.0006$) in both the DHX34 KO and UPF1 KO compared to WT thereby suggesting an alteration in NMD activity.

Although the lines did not display a reduction at the protein level for DHX34, the significant changes at the mRNA level in DHX34 and UPF1 for the UPF1 KO pool and DHX34 KO pool respectively, as well as the changes in NMD targets, were the reason we decided to further test the lines functionally.

4.3.2 HUVEC KO cell pools have decreased migration and increased leukocyte adhesion

Tube formation is an *in vitro* assay that can be used to assess angiogenic capacity of endothelial cells. Angiogenesis is an extremely important function of endothelial cells and plays a role in CAD with the formation of microvessels in the developing plaque. *In vitro*, when endothelial cells are plated on an extracellular matrix such as Matrigel, they spontaneously form vessel-like structures. All of the HUVEC lines purchased from Synthego maintained their tube formation capacity (**Figure 3a**). When the number of tubes per well was quantified, we noted a significant increase in the DHX34 KO wells as compared to the WT (**Figure 3b**, * $p=0.0179$), but no difference in the UPF1 KO pool ($p=0.1483$) suggesting increased angiogenic capacity of the DHX34 KO pool.

Migration of endothelial cells also plays a role in blood vessel formation and in the maintenance of the endothelial barrier of an existing vessel. When the cells are confluent, a scratch can be made in the monolayer and the closure of the gap of cells over 24 hours is quantified as a proxy for migration (**Figure 3c**). We did this with the three HUVEC lines and quantified the percentage of the original scratch remaining at 0 hours, 8 hours, 12 hours, and 24 hours. The UPF1 KO pool showed decreased migration as the percentage of the scratch remaining was significantly greater all time points tested when compared to the WT (**Figure 3d**, * $p<0.05$). The DHX34 KO pool showed no difference to the WT line at early timepoints, but at 24 hours, a significantly larger portion of the original scratch remained when compared to WT (**Figure 3d**, ** $p=0.0071$), thus suggesting the changes in DHX34 and UPF1 expression may impact the migratory capacity of endothelial cells.

The basal state of the endothelium is such that it is not adhesive to circulating leukocytes. This changes under inflammatory conditions, such as those promoted by $\text{TNF}\alpha$, where the expression of the adhesion markers VCAM and ICAM increases due to $\text{Nf-}\kappa\text{B}$ activation^{86,87}. In all of the HUVEC lines, the transcript levels of VCAM increased significantly following $\text{TNF}\alpha$ treatment (**Figure 3e**; $**p<0.01$) with no difference between the lines themselves. For ICAM levels, there was an increase following $\text{TNF}\alpha$ treatment in the WT and DHX34 KO pool (**Figure 3f**; $*p<0.05$) but not in the UPF1 KO pool. Interestingly, the ICAM transcript levels in the $\text{TNF}\alpha$ -treated DHX34 KO pool was significantly greater than in the wildtype $\text{TNF}\alpha$ -treated sample (**Figure 3f**; $**p=0.01$).

Both VCAM and ICAM play a role in monocyte adhesion to the endothelium. As such, the function of these adhesion markers was tested with the THP1 adhesion assay. THP1 cells are monocytes and can therefore bind to the adhesion markers expressed on endothelial cells. Under basal conditions, the endothelium maintains an anti-adhesive phenotype, which alters in an inflammatory environment such as the one under which CAD develops. By dye-loading the THP1 cells with Calcein-AM, we can measure the number of adherent THP1 cells in the presence of endothelial cells under both basal conditions and when activated with $\text{TNF}\alpha$ (**Figure 3g**). For each line, the addition of $\text{TNF}\alpha$ resulted in an increase in the number of adherent THP1 cells (**Figure 3h**, $*p<0.05$), although the increase did not reach significance for the DHX34 KO pool ($p=0.056$). Under $\text{TNF}\alpha$ -activated conditions, there were significantly more adherent THP1 cells on the DHX34 KO monolayer as compared to the WT monolayer (**Figure 3h**; $*p=0.0291$), with no difference in the number of adherent THP1 cells between the DHX34 KO and UPF1 KO lines. Taken

together, this suggests that the UPF1 KO and DHX34 KO pools have altered adhesive properties. The more advanced phenotype seen in the DHX34 KO over the UPF1 KO pool suggests a difference in target genes in the presence of DHX34 and/or a potential NMD-independent role for DHX34.

4.3.3 DHX34 KO and UPF1 KO pools have decreased NO production but no changes in eNOS

NO is the most potent vasodilator and is constitutively produced by eNOS in endothelial cells⁸². NO production can be measured *in vitro* through the use of 4-amino-5-methylamino-2',7'-difluorescein (DAF-FM) diacetate. DAF-FM diacetate can passively cross the cell membranes and once inside the cell, becomes de-esterified to DAF-FM, which is very weakly fluorescent. The fluorescence increases significantly (~160-fold) once it reacts with NO. The change in fluorescence can be detected and quantified via flow cytometry. When we measured basal NO production in the HUVEC KO pools, we noticed a distinct loss of the more strongly positive population (NO^{high}) in the DHX34 KO pool (**Figure 4a**). There was not only a difference in the NO⁺ population, but also a decrease in the percentage of NO⁺ cells in both the DHX34 KO and UPF1 KO pool compared to WT (**Figure 4b**, **p<0.01). We then quantified the mean intensity fluorescence of the NO⁺ population which showed no difference at the basal level between lines (**Figure 4c**) suggesting that the cells that were producing NO, were doing so to the same extent.

Because CAD occurs under inflammatory conditions, we tested whether NO production was impacted by TNF α stimulation. Interestingly, the production of NO in

response to TNF α was opposite in the DHX34 KO pool when compared to both the WT and UPF1 KO pool (**Supplementary Figure 4**). In the WT and UPF1 KO pool TNF α -stimulated samples, both the NO^{high} and NO^{low} populations persisted; however, there was a decrease in the median fluorescence intensity because of the shift towards the NO^{low} population which was significant in the WT line (**Figure 4c** *p=0.039; and **Supplementary Figure 4**). In contrast, following TNF α stimulation, a NO^{high} population appeared in the DHX34 KO pool, resulting in an increase in the median fluorescence intensity (n.s, p=0.088), which was significantly higher than what was measured in the WT (##p=0.001) and UPF1 KO (##p=0.0042) lines upon TNF α treatment (**Figure 4c**).

To determine if this was a result in a difference in eNOS, we measured the levels of both the transcript and protein. When analyzed through qRT-PCR (**Figure 4d**) there was no difference in eNOS transcript levels in the WT, DHX34 KO or UPF1 KO cells in either the basal or TNF α -stimulated state, and all three lines showed significant decrease in eNOS after TNF α treatment (**Figure 4d**, *p<0.05, **p=0.008). When activated, eNOS becomes phosphorylated on several residues including Ser1177. As such, we measured the level of phosphorylated eNOS in relation to total eNOS (**Figure 3e-f**) but noticed no significant difference across the samples (**Figure 3g**, p>0.05) suggesting differences in eNOS transcript and/or activation are not responsible for the decreased number of NO-producing cells in the DHX34 KO and UPF1 KO pools.

eNOS is not the only protein that can produce NO as there are both inducible and neuronal nitric oxide synthases (iNOS and nNOS respectively). nNOS is not largely expressed outside of neurons but iNOS is expressed across multiple cell types. We

measured iNOS mRNA transcript levels through qRT-PCR in the HUVECs and saw that the basal transcript level in both the DHX34 KO and UPF1 KO pools was significantly lower than in the wildtype cells (**Figure 4h**, *** $p=0.0004$ and ** $p=0.0043$), which may be able, in part, to explain the discrepancy in NO production across the three HUVEC lines.

4.3.4 Identification of DHX34 variants in EOCAD patients

Using knockout HUVEC cell pools suggested that DHX34 may play a role in aspects of endothelial function that are impacted during CAD development. As part of the DECODE study, patients with EOCAD provide blood samples that are used for exome sequencing. Through this effort, novel rare variants in *DHX34* were identified in 4 patients (**Table 4**).

At the time of the identification of the DHX34 variants, 55 patients had been enrolled in the DECODE study; 40 were male (age 35.6 ± 3.9) and 15 were female (age 39.7 ± 4.1). Exome sequencing was performed at the Genetics and Molecular Epidemiology Laboratory at the David Braley Cardiac, Vascular and Stroke Research Institute using the Ion Proton™ and Ion S5XL™ platforms (Life Technologies). Three levels of quality control (sex check, ethnicity check, and genotypic concordance) were used for the outputs before analysis was done to identify variants that were enriched in the DECODE population (**Figure 5a**). The identified variants were characterized based on allele frequency, amino acid change, predicted disruptiveness, and conservation across the cohort. DHX34 was identified through an additive model of analysis of the exome sequencing data with nominal association ($p=0.007$). Four unique variants were identified

in separate patients. Of the four, two were male and two were female with average ages representing the general DECODE population (37.5 ± 0.5 and 38.5 ± 6.5 respectively). The identified variants were individually validated through Sanger sequencing (**Figure 5b**) using primers listed in **Table 5** and were predicted to be damaging through PolyPhen2 and SIFT (data not shown) based on the resulting amino acid change.

The identified variants are found in three separate domains of DHX34; the recombinase A (RecA)1-like domain, the oligonucleotide/oligosaccharide binding (OB) domain, and the C-terminal domain (CTD) (**Figure 5c**)³¹⁸. Although not well defined, the RecA1 region contains the ATP hydrolysis and putative RNA-binding domains. The OB domain was found to contain sequences that regulate the ability of DHX34 to bind to UPF1 in the unphosphorylated form, whereas the CTD has been shown to be necessary for binding to SMG1³¹⁸. Due to the unique functions of each domain, impact of the variants on DHX34 function were assessed separately compared to the control sample.

4.3.5 DHX34 variants alter differentiation of iPSCs

PBMCs isolated from the patients were used to reprogram iPSCs using the Sendai virus. There was no difference in the reprogramming efficiencies of any of the patient samples when compared to the healthy control (**Figure 6a**; $n=2$ and $p>0.05$ for all). In the generated iPSCs, there was lower DHX34 mRNA expression in all lines compared to the healthy iPSC control except for the N819S line (**Figure 6b**, $*p<0.05$).

NMD has been previously shown to impact the differentiation of hPSCs^{310,323}; however, when the iPSCs were functionally validated through an embryoid body assay,

there was no difference in their ability to form cells of the three lineages compared to the healthy iPSCs (**Figure 6c**; $p>0.05$ for all). Interestingly, when a more targeted approach was used for differentiating the iPSCs to endothelial cells, the N819S iPSCs had a reduced efficiency compared to the healthy iPSCs whereas the R1000Q line had an increased efficiency (**Figure 6d**, $*p<0.05$).

The R1000Q iPSCs had no growth defects, and no obvious morphological differences to the other iPSC lines. Despite the increased endothelial differentiation capacity, the resulting endothelial cells grew extremely slowly. After sorting, the cells grew normally for 1-2 passages and then appeared to reach a point of senescence as they could no longer be expanded. As such, the data from this patient's endothelial cells will not appear in the sections to follow.

4.3.6 Patient iPSC-derived endothelial cells have decreased DHX34 protein but no apparent changes in NMD

To determine if the DHX34 variants impacted the patient endothelial cells in a similar way to what was seen in the HUVEC KO pools, the iPSCs were differentiated to endothelial cells as per the protocol in Chapter 3. In the endothelial cells, there was no difference in DHX34 mRNA expression in any of the patient lines as compared to the healthy endothelial cells (**Figure 7a**, $p>0.05$ for all), unlike what was seen at the iPSC level. However, all patient endothelial cells had significantly less DHX34 protein compared to the healthy cells (**Figure 7b**, $*p=0.014$, $**p<0.01$), suggesting the variants did impact either the translation or stability of the protein.

Again, we wanted to examine if the difference in DHX34 impacted NMD in the iPSC-derived endothelial cells. We measured UPF1 mRNA in the cells as it is itself regulated by NMD but noticed no difference across the samples (**Figure 7e**, $p > 0.05$ for all). Since DHX34 was found to play a role in regulating UPF1 phosphorylation, we wanted to see if the decreased DHX34 protein levels in our patient samples impacted the level of phosphorylation of UPF1. We measured phosphorylated UPF1 through western blot analysis and noticed no significant differences between the healthy control and patient lines (**Figure 7d**).

To further examine the impact on NMD, we looked at SMG1 mRNA levels as it is another NMD factor known to be regulated by the process itself³¹². Again, we saw no difference between the samples and the healthy control (**Figure 7e**, $p > 0.05$ for all). This was also the case when we examined the levels of ATF3 and GADD45B, two other known targets of NMD³¹⁰ (**Figure 7f**, $p > 0.05$ for all). As such, although the variants did appear to impact the extent of DHX34 protein expression in endothelial cells, NMD was not impacted as a result.

4.3.7 DHX34 variants impact migration and adhesion of endothelial cells

The process of NMD was not impacted in the endothelial cells generated from the patient iPSCs. However, we wanted to determine if the decrease in DHX34 protein expression impacts endothelial function in a similar way to what we saw in the HUVECs. The endothelial cells themselves did not appear to have any morphological differences

when grown in a monolayer (**Figure 8a**). They were all capable of forming tube-like structures with no significant differences across the samples (**Figure 8b-c**).

We assessed migration with the same scratch assay that was performed with the HUVECs (**Figure 3c**). When the migration was quantified, we detected significantly impaired migration in the L175P and N819S endothelial cells at the early time points compared to the healthy endothelial cells (**Figure 8d**, 8 hours and 12 hours $*p<0.05$). By the 24-hour timepoint, this difference was no longer significant.

Because of the role of inflammatory cytokines in regulating adhesion marker expression in CAD development, we measured the level of VCAM and ICAM mRNA in the patient-specific endothelial cells. The VCAM mRNA expression level increased in all samples following $\text{TNF}\alpha$ treatment (**Figure 8e**, $*p<0.05$, $***p=0.000253$). Interestingly, the VCAM mRNA expression following $\text{TNF}\alpha$ stimulation in the N819S and R947C endothelial cells was significantly elevated compared to the control (**Figure 8e**, $##p=0.0043$, $###p=0.0001$). There was again an increase in ICAM levels following $\text{TNF}\alpha$ treatment, although the increase was only significant for the R947C sample (**Figure 8f**, $*p=0.04$) with no difference across the samples under either basal or $\text{TNF}\alpha$ -stimulated conditions.

We again employed the THP1 adhesion assay to assess if the differences in adhesion marker expression at the mRNA level resulted in functional changes in leukocyte adhesion to the endothelium. For all samples, we saw a slight increase in the number of adherent THP1 cells following $\text{TNF}\alpha$ treatment (**Figure 8g**), although it did not reach significance for any of the samples (**Figure 8h**). However, the number of adherent cells under the inflammatory condition was elevated in both the L175P ($*p=0.0227$) and the

N819S cells (n.s p=0.0508) compared to the control, suggesting that the endothelium may be more adhesive for leukocytes in the presence of the DHX34 variants.

4.3.8 DHX34 variants alter NO production in endothelial cells

As was mentioned before, NO is an important signaling molecule in endothelial cells and is a hallmark of endothelial dysfunction. When we measured NO production through DAF-FM addition in the patient endothelial cells, we noticed no obvious differences in the population distribution of NO⁺ cells by flow cytometry (**Figure 9a**) which was quantified as the percentage of NO⁺ cells from each sample (**Figure 9b**, p>0.05 for all). However, when further examined, there was a significant increase in the mean intensity fluorescence of the NO⁺ populations from the L175P and N819S samples compared to the healthy endothelial cells (**Figure 9c**, *p=0.041, **p=0.0091). Similar to what we did with the HUVECs, we also measured NO production following treatment with TNF α (**Supplementary Figure 3**) but none of the samples showed a significant shift in NO production under the inflammatory conditions (**Figure 9c**, p>0.05 for all).

We then moved to examine whether the difference in the basal mean intensity fluorescence could be explained by differences in eNOS. We detected no difference between the eNOS mRNA level between our samples and the healthy endothelial cells (**Figure 9c**, p>0.05 for all), nor in the level of phosphorylated eNOS (**Figure 9d**, p>0.05 for all). The antibody for total eNOS did not work with these samples and as such, only phosphorylated eNOS was measured. In the HUVECs, there was a detectable difference in iNOS mRNA in the KO pools but this trend was not seen in our patient endothelial cells as

there was no difference between the DHX34 variant endothelial cells and the control (**Figure 9e**). Therefore, the increased mean intensity fluorescence observed in the L175P and N819S endothelial cells is a result of differences outside of eNOS and iNOS production.

4.4 DISCUSSION AND FUTURE DIRECTIONS

NMD plays an important role in disease development by regulating the level of PTC-containing transcripts; however, dysregulation of NMD proteins themselves has only been found to result in clinical symptoms in the context of UPF3b³²¹. In this work, I've demonstrated for the first time that manipulating DHX34, a member of the NMD pathway, can alter endothelial function by decreasing migration, increasing adhesiveness of the cells, and altering NO production.

To date, very little work has been published on DHX34. Its role in NMD was identified in 2014³²⁴ and more closely examined in 2016 where it was found to catalyze the SMG1-mediated phosphorylation of UPF1, thereby activating NMD³¹⁸. In this work, we used KO HUVEC pools from Synthego in an attempt to determine how reducing endogenous expression of both DHX34 and UPF1 could impact endothelial cell function specifically. The UPF1 KO pool was used to assess the difference between reducing DHX34 specifically and reducing the activity of NMD more globally as UPF1 is necessary for NMD. The KO pools were generated with CRISPR/Cas9 editing techniques and were validated by Synthego through sequencing (**Supplementary Figure 1-2**). The results suggest that the UPF1 KO worked to a greater extent than the DHX34 KO as per the r^2

value calculated from the discordance analysis. Although there were no significant morphological differences, a reduced growth rate of the UFP1 KO pool was noted by both myself and Synthego which has been reported to be an impact of NMD inhibition³²⁵. Despite the fact that Synthego reported that the editing had been successful, when we tested the level of expression at both the protein and mRNA level, we only noted a slight (but insignificant) decrease in DHX34 mRNA and no change at the protein level. One replicate of the DHX34 KO sample had a much higher expression than the other two while the WT cells had much more variable DHX34 expression. Because of this high level of variability, it is not possible to actually determine whether the KO worked and therefore, the western blot analysis should be repeated. The variability could be due to the fact that the cells were a KO pool rather than a clonally isolated and expanded KO cell. As was indicated in the analysis provided by Synthego, the KO score was between 50 for the DHX34 line and 42 for the UPF1 line suggesting that the many of the indels they detected did not result in a frameshift. As such, further work should be done with a clonally isolated KO line to validate the reported results.

Despite the lack of obvious decrease in the desired protein expression of both KO pools, the UPF1 KO had a significant increase in DHX34 mRNA whereas the DHX34 KO had a significant decrease in UPF1 mRNA. Both of these lines also had significantly reduced mRNA expression of SMG1, ATF3, and GADD45B, all of which are known targets of NMD^{310,312}. This result was contrary to what was expected in the UPF1 KO line, as NMD should have been fully inhibited and therefore resulted in an increase in the expression of the targets. This was seen in previous work by Hug, *et al.* where the knockdown of UPF1 in HEK293T cells led to an increase in these transcripts as measured

through microarray³²⁴. However, when they knocked down DHX34 with the same system, there was only an increase in SMG1 levels, suggesting alternative pathways of regulation. Moreover, because of the role NMD has on regulating naturally occurring transcripts^{308,311} in a cell-specific manner^{308,311,326}, it is possible that in HUVECs, the role of UPF1 and DHX34 in regulating these transcripts is different than in HEK293T cells. It is also possible that NMD inhibition occurred, as indicated by the decreased growth rate of the UPF1 KO pools, and as such overall transcription was decreased in these cells resulting in lower mRNA levels. Therefore, using an NMD complementation assay that has been used by multiple groups to determine its activity would be helpful. The reporter consists of the T-cell receptor β minigene containing a PTC and another reporter without the PTC³²⁷, thus allowing for the activity of NMD to be directly assessed by comparing the levels of the WT and PTC-containing transcript levels. This would eliminate any differences based on transcription rates of the cells.

Through functional assays, it was clear that the editing process had an impact on endothelial phenotype in the HUVECs. Both KO pools showed reduced levels of migration, which was more consistent in the UPF1 KO pool than the DHX34 KO pool, although this pattern was the reverse for the adhesion assays. The difference in phenotypic severity between the DHX34 KO and UPF1 KO pools again highlights the possibility that DHX34 may have a role outside of NMD in endothelial cells. The elevated ICAM in the DHX34 KO pool following TNF α treatment is likely the cause of the elevated THP1 adhesion under the same conditions as there was no difference in the VCAM level compared to either the WT or UPF1 KO. Both VCAM and ICAM are under the transcriptional control of Nf- κ B, which is in part regulated through NO inhibition^{81,87}.

However, under the conditions used here, NO signaling is not likely responsible for the differences in adhesion marker expression and function. There was no difference in the eNOS mRNA levels or phosphorylation status across the samples, reinforcing this idea. To further confirm this, additional western blots should be done to determine whether the other eNOS regulation sites differ. Here, we performed a western blot for the activating phosphorylation mark on Ser1177; however, eNOS can also be phosphorylated on Thr459, which decreases activation²⁸². By analyzing all of these, together with total eNOS protein level, a more decisive conclusion can be made as to whether eNOS plays a role in the altered NO production in the DHX34 KO pool.

Although there didn't appear to be a difference in basal eNOS expression to explain the reduced NO⁺ population in the DHX34 KO pool, there was a significant decrease in the iNOS mRNA. eNOS constitutively produces NO in endothelial cells but iNOS is the highest output isoform of the three NOSs and contributes to the endothelial production of NO under certain conditions such as in the presence of ROS³²⁸. Therefore, although the lower iNOS mRNA may explain the decrease in total NO⁺ cells under basal conditions, more work must be done to determine if iNOS protein expression is also altered and how, or if, this changes under inflammatory conditions. This would also be important to clarify a potentially contrasting relationship between ATF3 and iNOS in endothelial cells and in macrophages. In 2015, Jung, *et al.* reported that in macrophages, a reduction in ATF3 led to the activation of Nf-κB and thereby an increase in NO through enhancement of iNOS³²⁹. This relationship appears to be different in the HUVEC system used here. We noted a significant reduction in both ATF3 mRNA and iNOS mRNA, as well as a decreased number of NO⁺ cells under basal conditions in the KO pools compared to WT. This

discrepancy again highlights the importance of studying NMD target regulation in a cell-specific context in order to better understand how manipulation of NMD may impact observed phenotypes.

Using the HUVEC KO pool system provided some insight into the potential role of DHX34 in endothelial cells; however, the system itself was not ideal. The extent of knockout that we could measure was not significant despite the sequencing results. This could be in part due to the fact if the editing conferred a disadvantage to the cells, the WT cells in the pool could take over. Additionally, because HUVECs are a primary cell line, the length of *in vitro* culture must be considered when using them for experimentation²³⁶. As such, it is ideal to use primary HUVECs at a low passage number, but the KO cell pools provided from Synthego were at passage 7. Using the cells for multiple assays at higher passages could change the readout based on the changes that the cells undergo after repeated divisions. It is also possible that off-target effects that were not identified were incurred through the editing process and may have impacted the phenotype of the cells. Additionally, although the KO system is extremely useful for elucidating the function of a protein, naturally occurring knockouts in DHX34 have not been reported. In order to further elucidate the role of DHX34 in endothelial cells and its potential role in disease, we examined the impact of *DHX34* variants found in EOCAD patients in endothelial cells using iPSCs.

NMD has been previously shown to have a role in regulating pluripotency and differentiation of pluripotent stem cells. Findings reported in 2014 by Lou, *et al.* placed NMD as an important pathway in stem cell maintenance³²³. Their findings suggest that UPF1 plays an important role in establishing and maintaining pluripotency through

regulation of SMAD7 as well as through the degradation of transcripts that oppose proliferation and the undifferentiated state. In our patient-derived iPSC lines, we did not detect any difference in the reprogramming efficiency of the DHX34 variant samples when compared to the healthy control, suggesting that re-establishing the pluripotency network was not impacted by these variants. There was a difference in the targeted differentiation capacity of two of the patient lines, N819S and R1000Q; however, no differences were noted from the EB assay results. This is not surprising as the samples were collected from adults with no known congenital disorders and therefore, none of these variants could have caused embryonic lethality or significant congenital deformities. Despite this, it is possible that small differences in lineage, or even cell specific, preferences could be detected if a larger number of genes had been used to screen the EB assay (covered in Chapter 2.4). Expanding these findings would be interesting and could contribute to the current model whereby NMD inhibition is necessary to regulate differentiation^{310,323}.

All of the patient samples used in this work had unique variants that fell in different domains of DHX34. The RecA1 domain, where L175P is found, has not been explicitly studied. This domain harbours the ATP hydrolysis domain and could therefore theoretically play a role in the conversion of the SURF complex to the DECID complex through impacting ATP hydrolysis by DHX34. The OB domain, in which both the N819S and R947C variants lie, was found to impact the binding of DHX34 to UPF1 by Melero, *et al*¹⁸. In their work, removing the OB domain led to increased DHX34-UPF1 interaction, but its specific impact on NMD was not determined. The R1000Q variant is located in the CTD, which is necessary for SMG1 binding and therefore for UPF1 phosphorylation. When this domain was deleted, Melero, *et al*. saw a reduction in both phosphorylated UPF1

levels and in NMD activity, suggesting that the domain is vital for this process to continue³¹⁸. The iPSCs harbouring the R1000Q variant had no growth defects and had increased differentiation efficiency towards an endothelial lineage; however, the endothelial cells had noticeable problems expanding and were therefore not able to be used for this study. To determine if this was a result of the R1000Q variant itself, creating an isogenic iPSC line and differentiating it towards an endothelial lineage would be crucial.

Unlike the iPSCs, the patient endothelial cells did not show any decrease in the extent of DHX34 mRNA expression; however, the level of DHX34 protein was significantly lower than what was measured in the control. This suggests that the patient variants do not impact transcription of DHX34 or stability of the mRNA but rather translation of the transcript or protein stability. Additional work, however, is needed to determine which of these processes is responsible for the decrease. Despite the decreased DHX34 protein, no impact on NMD was detected in the endothelial cells. Performing an immunoprecipitation (IP) assay on these endothelial cells would be necessary to fully determine if and how they are impacting binding to UPF1 and/or SMG1. This would be especially interesting to see if the variants located in the OB domain impact DHX34 binding to UPF1 and how that may impact its phosphorylation. At this time, we have not found any evidence that the variants impact the binding of DHX34 to UPF1; however, performing an IP along with measuring phosphorylated UPF1 in relation to total UPF1 is a necessary future direction for this work.

The patients had no significant changes in their tube formation capacity, which was not unexpected since serious vascular abnormalities detected by a loss of tube formation function would have caused additional clinical symptoms at a younger age. Although the

lines did form tubes, they were not as adept at it as the HUVECs. This could be a result of the maturity of the cells at the time of tube formation. Angiogenesis is initiated *in vivo* through a number of ways including in response to elevated VEGF. As such, the addition of VEGF to the media during the tube formation assay will promote maximal angiogenesis³³⁰, but this method was not used in this work. Performing the same assays with the addition of VEGF may help to increase tube formation of all lines and allow for better analysis of any potential differences across the lines.

From this work it is clear that the DHX34 variants impact components of the anti-atherogenic profile of endothelial cells. Maintenance of the anti-adhesive endothelial barrier is extremely important in the prevention of CAD. Here we show that the migration of the patient endothelial cells is reduced when compared to the healthy control, suggesting that they are slower to respond to an initial injury than their healthy counterpart. In the context of CAD development, this would equate to prolonged periods where the vessel is denuded of the endothelium, therefore rendering the vessel more susceptible to migrating LDL and leukocytes, as well as thrombus formation. Under inflammatory conditions, the patients also displayed elevated VCAM expression as well as elevated leukocyte adhesion, which has been suggested to have an important role in early stages of CAD development³³¹. Although the increase in adhesion marker expression is part of the normal response to injury and is necessary in times of acute injury, excessive expression of adhesion markers and leukocyte adhesion are more likely to cause CAD development at an accelerated rate. Additional work is required to determine the full mechanism of this process as not all patients that had elevated adhesion marker mRNA levels also had increased THP1 adhesion. Examining additional adhesion markers such as those in the selectin family may

help to explain this discrepancy. Cytokine production is another important aspect of endothelial activation that contributes to leukocyte migration and adhesion to the site of injury. Measuring cytokine production by the patient endothelial cells, such as MCP-1 and IL-8, through enzyme-linked immunosorbent assay (ELISA) is necessary. In terms of leukocyte adhesion, examining the impact that the DHX34 variants in patient monocytes have on this process as well is important. Similar to what has been previously done by Hu, *et al.*³³², the same adhesion assay shown here could be performed with patient monocytes collected upon enrolment. Using the patient cells on a healthy endothelial monolayer and vice versa would allow us to better dissect the contribution of both cell types to the process that occurs *in vivo*.

A decrease in NO bioavailability is the hallmark of endothelial dysfunction but in our patient cells, we saw an increase in the production of NO compared to the control. This was seen in the absence of any differences in eNOS or iNOS mRNA levels or eNOS phosphorylation on the Ser1177 residue. Unfortunately, the total eNOS antibody used with the HUVEC samples did not work well with the patient samples and so, further optimization must be done to confirm the extent of eNOS phosphorylation with respect to total eNOS. Additionally, measuring iNOS protein levels should be done, especially as it produces higher levels of NO than eNOS³²⁸ and could therefore be playing a role in the elevated NO levels in our patients.

The relationship between NO and the adhesion marker expression program that we saw in the patients is another aspect that needs further investigation. NO acts as an inhibitor of Nf- κ B in the endothelium which means that when NO bioavailability decreases, the Nf- κ B subunits translocate to the nucleus and increase transcription of genes, such as those for

adhesion markers⁸⁶; however, there was no decrease in basal adhesion marker expression despite increased NO in the patient samples. On the other hand, TNF α treatment on its own also results in the translocation of the p65 subunit of Nf- κ B to the nucleus⁸⁶, resulting in increases in VCAM and ICAM mRNA as was seen here. Measuring Nf- κ B activity by measuring the nuclear versus cytoplasmic p65 fraction under basal conditions is therefore necessary to better understand the impact of the elevated NO levels on downstream pathways.

In both the HUVEC KO pools and the patient-derived endothelial cells, we saw a distinct impact on migration rate and leukocyte adhesion. These findings suggest a novel role for DHX34 in maintaining endothelial functions important for CAD prevention. The difference in NMD activity as measured by target transcript expression, could be a result of alternate NMD pathways being induced in the absence of DHX34. Because various NMD pathways differentially depend on certain proteins^{333,334}, it is possible that another pathway not reliant on DHX34 was able to compensate in the patient cells. In the HUVECs, we saw what appeared to be an overactive NMD phenotype, although the lower transcript levels may be related to the decreased growth rate as a result of NMD inhibition as mentioned above. Overall, performing larger scale RNA sequencing and/or using an NMD complementation reporter would provide a better assessment of NMD function in both the patient endothelial cells and the HUVEC KO pools.

The disparate results seen between the HUVECs and the iPSC-derived endothelial cells could partially be a result of differences in endothelial subtype. As mentioned in Chapter 3.4, further analysis on the iPSC-derived endothelial cells is necessary to determine which subtype they are most similar to. If the patient iPSCs are closer to arterial

than to venous endothelial cells, some of the variance associated with NO signaling and potentially cell-specific NMD functions could be attributed to these differences.

Despite the fact that the mechanism remains elusive, this work demonstrates that mutations in *DHX34*, impacted NO signaling in a potentially NMD- and UPF1-independent manner. Alterations in NO occur *in vivo* through multiple ways including the uncoupling of eNOS, whereby the protein level remains unchanged, but the output of NO is decreased, or by activating iNOS which produces higher amounts of NO. Therefore, performing additional analysis on ROS levels and iNOS activity in these cells could help explain the observed phenotypes.

Despite the differences we noted in endothelial function, it is important to consider the fact that CAD development requires the cooperation of many cellular subtypes. Therefore, differentiating the iPSCs generated for this study towards other lineages that also contribute to CAD development, such as smooth muscle cells, monocytes, and macrophages, would be extremely important in elucidating the overall role of *DHX34* in EOCAD development.

Future work on elucidating the mechanism at play here should also focus on the use of additional controls. For a disease such as CAD that is prevalent and polygenic, the use of multiple healthy control lines becomes extremely important. The patient iPSC lines and the endothelial cells derived from them were compared to an iPSC line derived from PBMCs from one healthy control. Deriving similar iPSC lines from familial controls or other unaffected controls will clarify the significance of the results reported here and aid in reducing the variability associated with other genetic variants present in the patients.

With CRISPR/Cas9 technology, it is now possible to create isogenic iPSC lines. Despite the methods that have been developed to promote homology-directed repair^{335–338}, which is required to induce specific point mutations, insertions, or deletions, the efficiency of this process is still quite low. Nonetheless, the utility of isogenic iPSC lines in studies such as these is immeasurable. In the patients used for this study alone, multiple rare variants that were enriched in the EOCAD population were identified through exome sequencing. Therefore, it is possible that the contribution of the DHX34 variant to the observed phenotype is less than what we perceive. Although the work with the HUVEC KO pools supports the conclusion that DHX34 does impact endothelial function, creating isogenic iPSC lines is currently the gold-standard to determine genetic contribution to an observed phenotype. By inserting the specific DHX34 variants into an iPSC line derived from unaffected individual and differentiating it as was done here, we can more definitively determine how that specific variant contributes to the phenotype while negating the impact of other factors in the genetic background. To further validate the findings, we could also correct the variant in the patient line to return the DHX34 gene to the wildtype state. In performing the same analysis as was done here with endothelial cells from the corrected isogenic iPSC line in conjunction with the mutated isogenic iPSC line, a definitive conclusion on the contribution of a specific DHX34 variant on endothelial function could be drawn. Once created, the isogenic lines could also then be used for differentiation towards other cell types as previously mentioned, so that the role of DHX34 in other aspects of CAD development could be ascertained.

The work presented here demonstrates for the first time that DHX34 has an important role in modulating endothelial function associated with CAD development,

namely migration and leukocyte adhesion. This work demonstrates the utility in using patient-derived cell lines for elucidating the impact of novel variants in disease development. With follow-up from the future studies suggested here to elucidate the mechanism, the novel role of DHX34 in maintaining anti-atherogenic endothelial function could be seen as a potential therapeutic target for CAD patients.

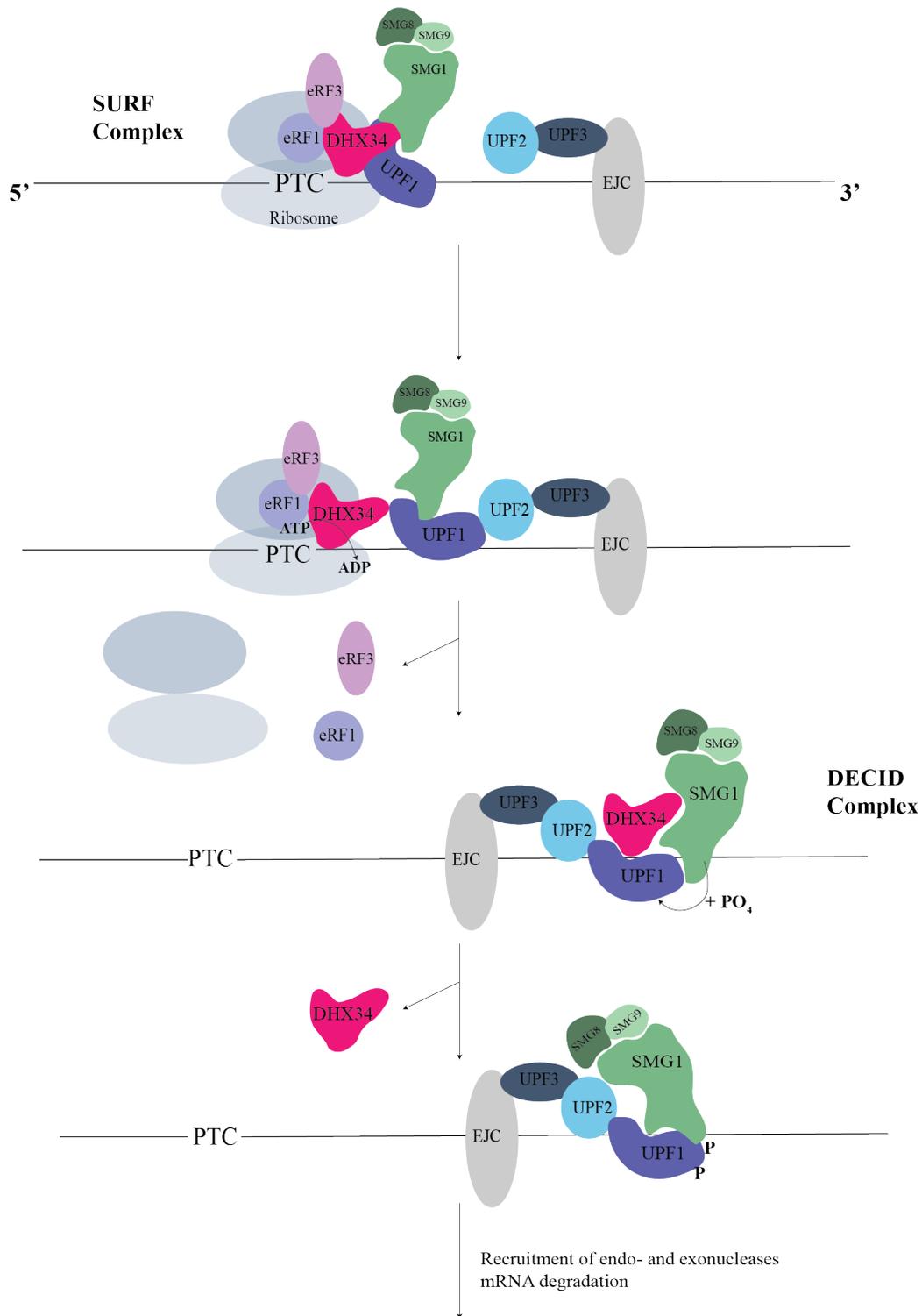


Figure 1. Nonsense-mediated decay recognition of PTC-containing mRNA transcript.

During translation, the ribosome pauses when it reaches a termination codon and recruits eRF1 and eRF3 to the stalled ribosome. When the termination codon is premature, eRF1 and eRF3 cannot bind to proteins in the poly(A) tail and as a result, recruit UPF1, DHX34, and the SMG1 complex (SMG1, SMG8, SMG9), forming the SURF complex. At this point, DHX34 hydrolyzes ATP to ADP resulting in the release of the ribosome and the elongation factors from the transcript and reconfiguring the proteins to the DECID complex. In this conformation, DHX34 acts as a scaffolding protein to promote the phosphorylation of UPF1 by SMG1. Once UPF1 is phosphorylated, DHX34 disassociates from the transcript and NMD is activated. Exo- and endonucleases are recruited to phosphorylated UPF1 and act on the mRNA transcript. The exo- and endonucleases also recruit phosphatases, such as phosphatase 2a which dephosphorylates UPF1. Dephosphorylating UPF1 acts to both recycle UPF1 and to limit the extent of NMD. Image adapted from Hug, *et al.* 2014³²⁴.

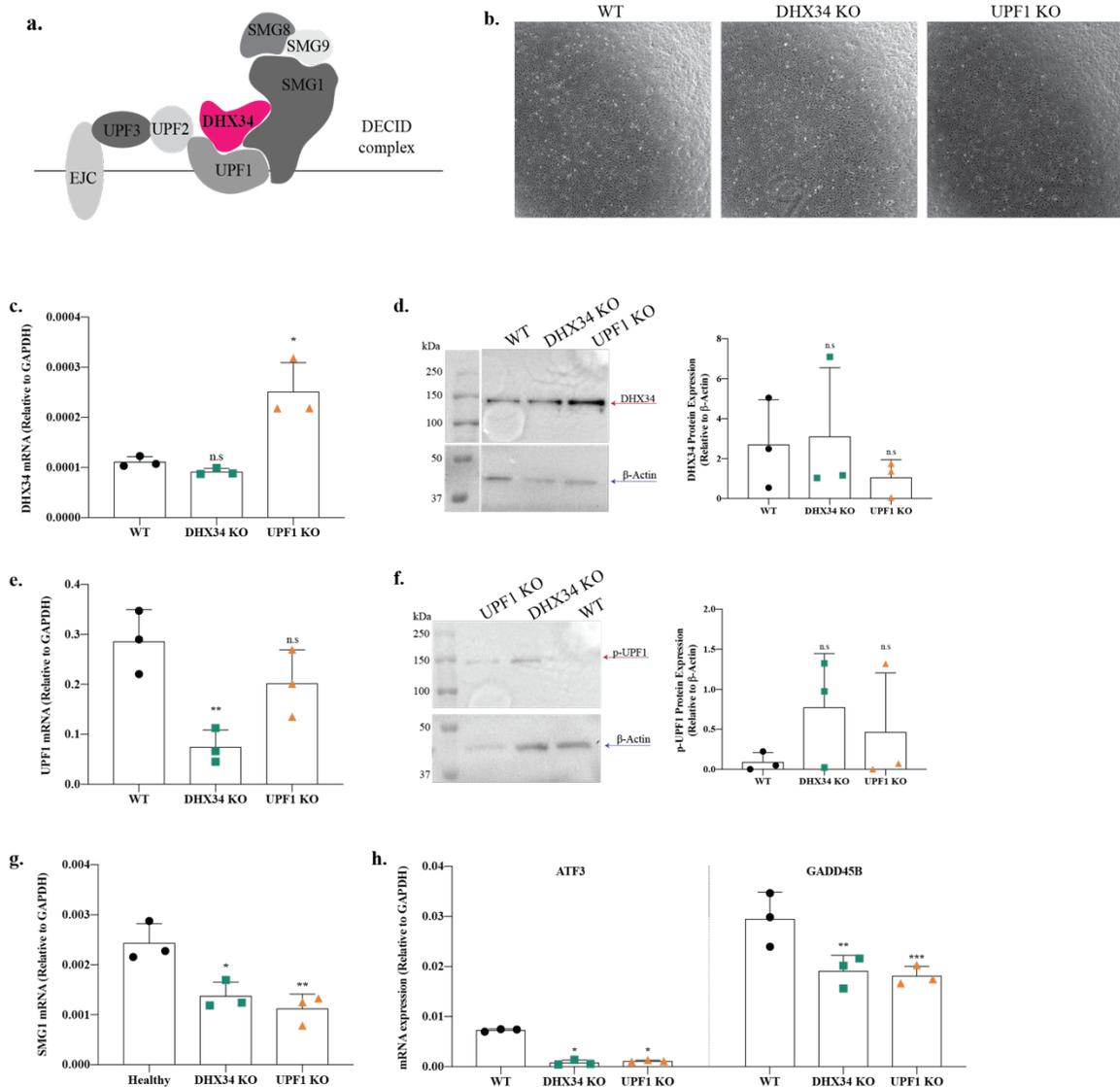


Figure 2. DHX34 and UPF1 KO pools do not have altered levels of DHX34 protein or UPF1 phosphorylation or despite differences in mRNA expression. It was recently demonstrated that DHX34 acts as a scaffolding protein to allow for the SMG-mediated phosphorylation of UPF1, a critical step in NMD activation (**a**, figure adapted from Hug, *et al.* 2014³²⁴). KO pools of both DHX34 and UPF1 were purchased from Synthego. The

cells displayed no gross morphological differences when grown in a monolayer (**b**). The cells were tested for mRNA expression of DHX34 and while the reported DHX34 KO had only a slight decrease in DHX34 mRNA ($p=0.0528$), the UPF1 KO had significantly greater DHX34 mRNA compared to the control (**c**, t-test; $*p=0.0145$). Protein levels were tested using western blot with β -Actin as the loading control (predicted molecular weight: DHX34 – 128 kDa, β -Actin – 42 kDa). This showed no difference in DHX34 protein expression across the samples (**d**). When UPF1 mRNA expression was assessed, the reported UPF1 KO showed no difference ($p=0.1912$); however, the DHX34 KO pool did (**e**, t-test; $**p=0.0072$). When NMD status was checked in these cultures by probing for phosphorylated UPF1 (predicted molecular weight p-UPF1 – 125 kDa), no difference was detected (**f**). NMD activity was also measured by detecting mRNA levels of known targets including SMG1 (**g**, t-test; $*p=0.0182$, $**p=0.0095$), ATF3, and GADD45B (**h**, 2-way ANOVA; $*p<0.05$, $**p=0.0012$, $***p=0.0006$) which were all lowered in the KO pools.

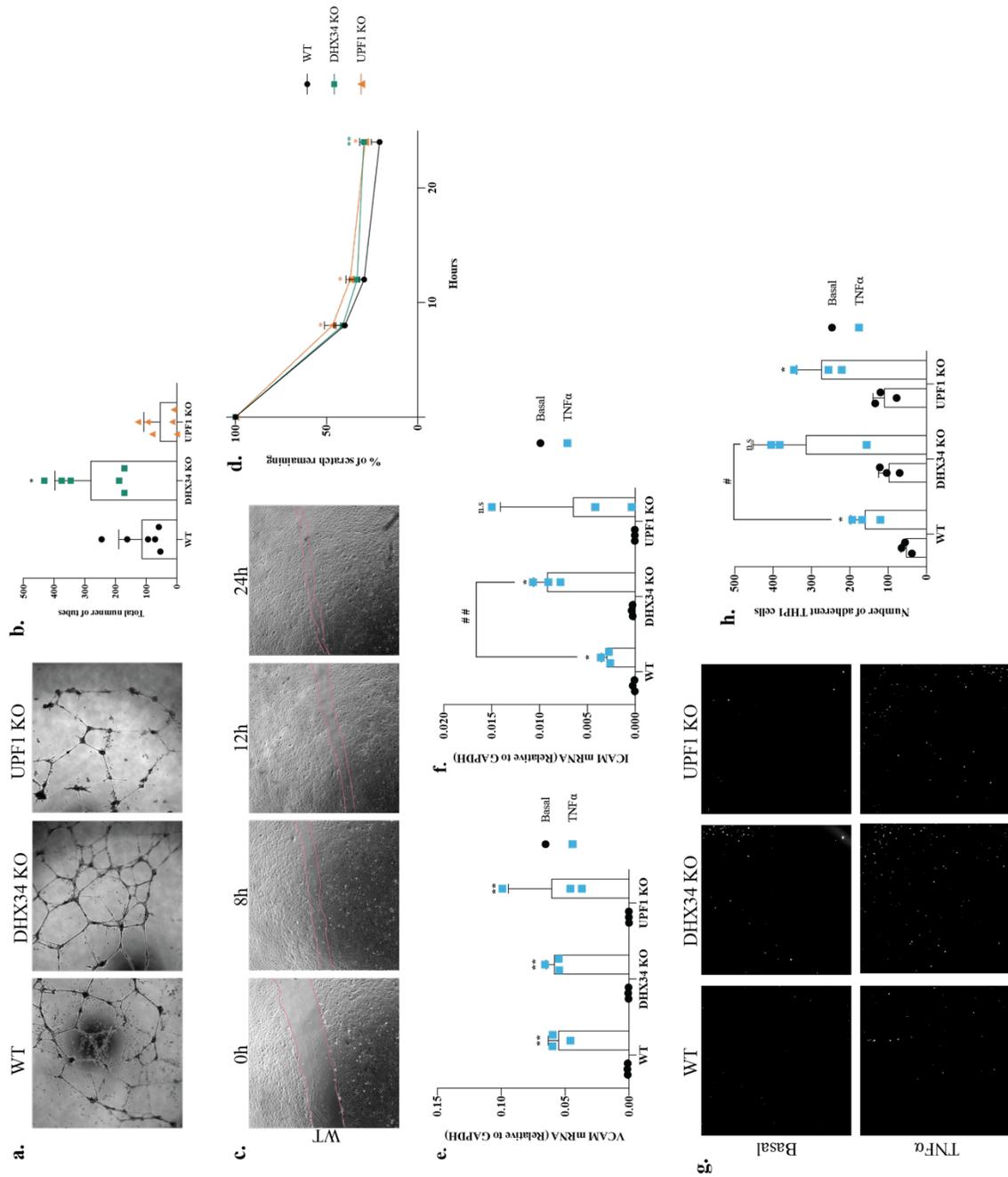


Figure 3. Knocking out NMD components in HUVECs alters normal endothelial function. All HUVEC lines were capable of forming tubes (**a**) with the DHX34 KO line having an increased number of total tubes compared to the wildtype (**b**, t-test; * $p=0.0179$). A scratch assay was used to measure migration over an area denuded of the monolayer (**c**) over 24 hours. The UPF1 KO line had consistently slower migration rate at every time point measured (**d**, 2-way ANOVA; * $p<0.05$) whereas there was no difference between wildtype and DHX34 KO until 24 hours (2-way ANOVA; ** $p=0.0071$). VCAM levels in all lines increased significantly (**e**, 2-way ANOVA; ** $p<0.005$) following $TNF\alpha$ treatment with no difference between the lines. ICAM levels also increased significantly in all but the UPF1 KO line following $TNF\alpha$ treatment (**f**, t-tests; * $p<0.05$). The ICAM mRNA level in the $TNF\alpha$ -treated DHX34 KO line was significantly greater than what was measured in the wildtype line (**f**, t-test; ## $p=0.01$). The function of the adhesion markers was measured through the THP1 adhesion assay where adhered fluorescently labelled THP1 monocytes were visualized (**g**) and then quantified (**h**). Following $TNF\alpha$ treatment, all lines showed an increase in adherent THP1 cells, although not significant in the DHX34 KO line (**h**, * $p<0.05$, n.s $p=0.056$). Again, the number of adherent THP1 cells in the $TNF\alpha$ -treated DHX34 KO line was significantly greater than what was measured in the wildtype line (**h**, # $p=0.0291$).

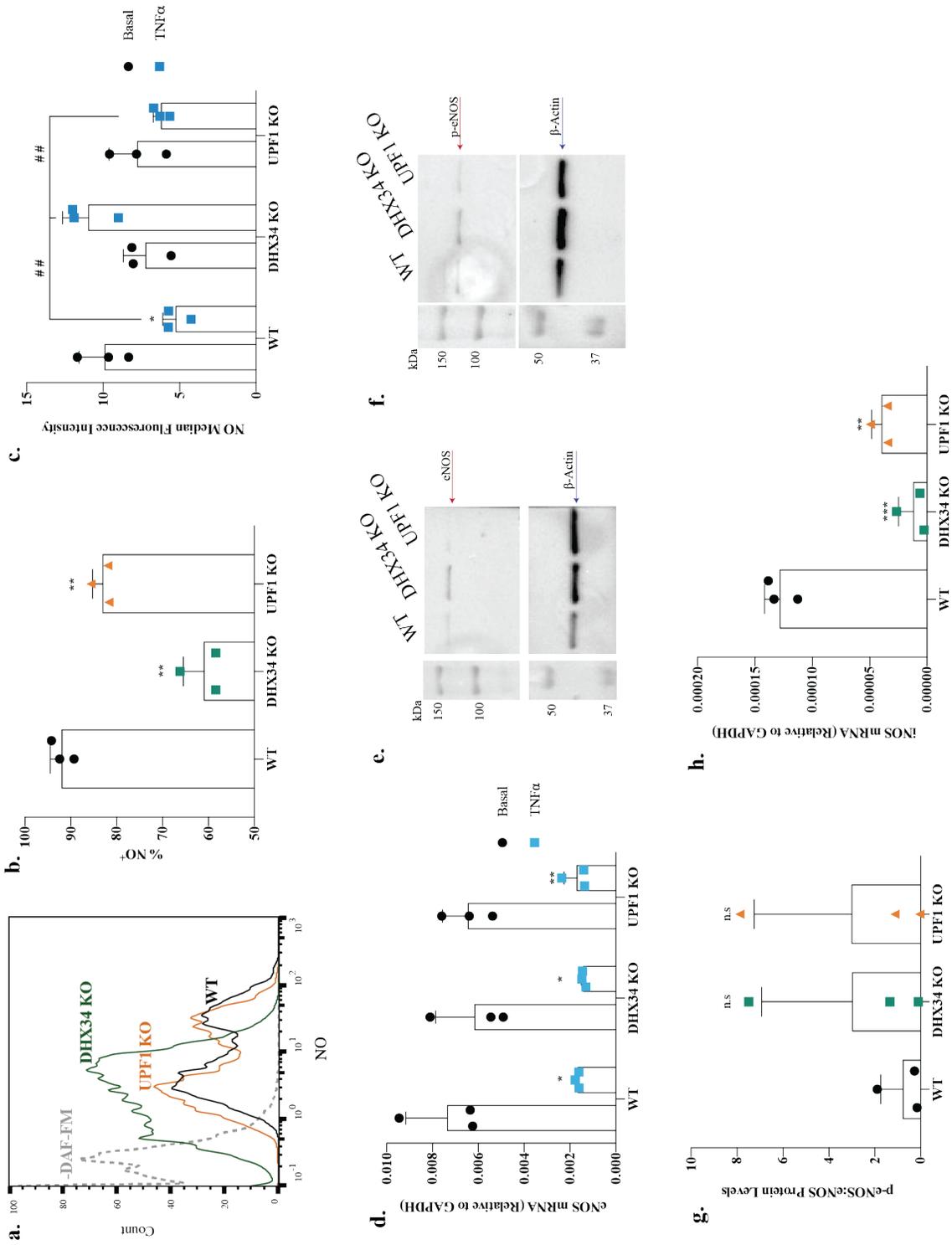


Figure 4. Altered NO signalling in DHX34 KO HUVECs is not a result of differences in eNOS. NO production was measured in HUVECs through the addition of DAF-FM to the culture media and quantified with flow cytometry. There were two distinct populations in both the WT and UPF1 KO lines, whereas only one in the DHX34 KO pool (**a**). When the percentage of NO⁺ cells was quantified, there was a significant decrease in the DHX34 KO pool as well as the UPF1 KO pool compared to the WT line (**b**, t-test; **p<0.01). When the median intensity fluorescence of the NO⁺ cells was measured, there was no difference between the lines (**c**). Following TNF α treatment, the median intensity fluorescence decreased in the WT but did not change across the other samples (**c**, t-test; *p=0.039). However, the median intensity fluorescence in the TNF α -treated DHX34 KO line was significantly greater than in the wildtype and UPF1 KO pool (**c**, ## p<0.01). There was no difference across the samples in eNOS mRNA levels under either basal or TNF α -treated conditions with all samples showing significantly less eNOS mRNA when treated with TNF α (**d**; *p<0.05, **p=0.008). Western blot analysis was performed for total eNOS (**e**) and for phosphorylated eNOS (**f**, predicted molecular weight: eNOS – 140 kDa, β -Actin – 42 kDa). There was no difference in phosphorylated eNOS with respect to total eNOS across the samples (**g**). There were however significantly lower levels of iNOS in both the DHX34 KO and UPF1 KO pools compared to the wildtype under basal conditions (**h**; 2-way ANOVA, **p=0.0043, ***p=0.0004).

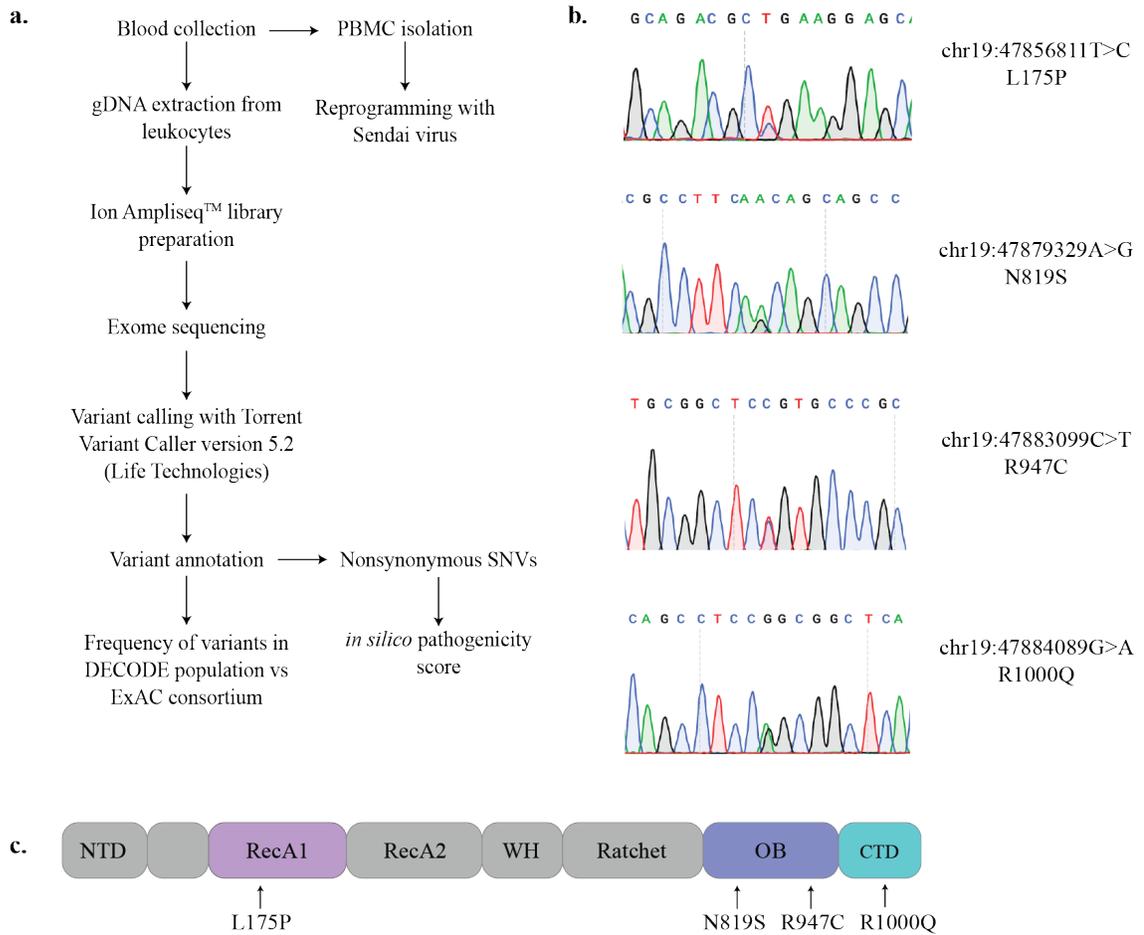


Figure 5. Identification of rare variants in DHX34 in the DECODE study. Exome sequencing and subsequent analysis was performed on patient samples collected from the DECODE cohort by Ricky Lali from Dr. Paré’s lab (a). This resulted in the identification of four novel rare variants present in DHX34; L175P, N819S, R947C, and R1000Q. All of the patients were heterozygous for the different point mutations as validated through Sanger sequencing (b). The four variants are located in three different domains of the DHX34 protein (c).

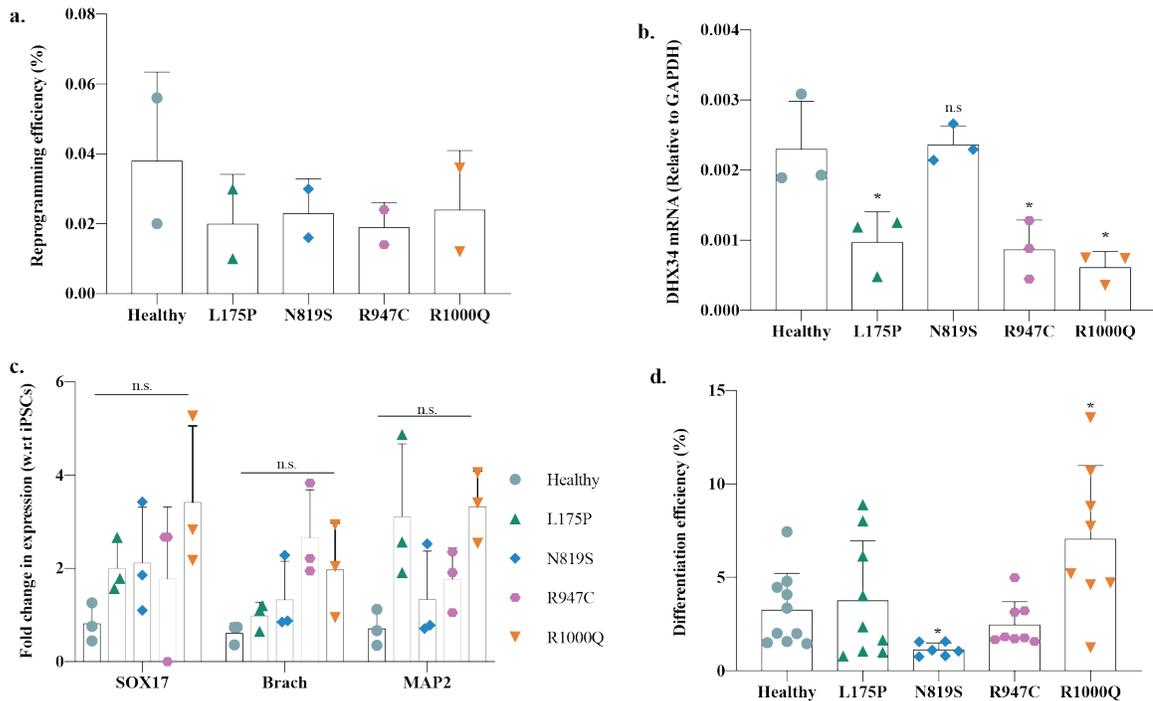


Figure 6. Rare variants lower DHX34 mRNA levels in iPSCs without impacting differentiation capacity. Patient samples along with a healthy PBMC control purchased from StemCell Technologies were reprogrammed to iPSCs using the same protocol with no difference in the reprogramming efficiency (a). In the iPSCs, DHX34 mRNA levels were lower in the patient lines, with the exception of the N819S line (b, t-test $*p < 0.05$). When the iPSCs were used in an EB assay, no differentiation bias was noted as there was no difference in any marker between the samples (c, $p > 0.05$ for all). When the iPSCs were used in a directed differentiation protocol towards endothelial cells, the N819S line showed a reduction whereas the R1000Q line showed increased efficiency compared to the healthy line (d, t-test; $*p < 0.05$).

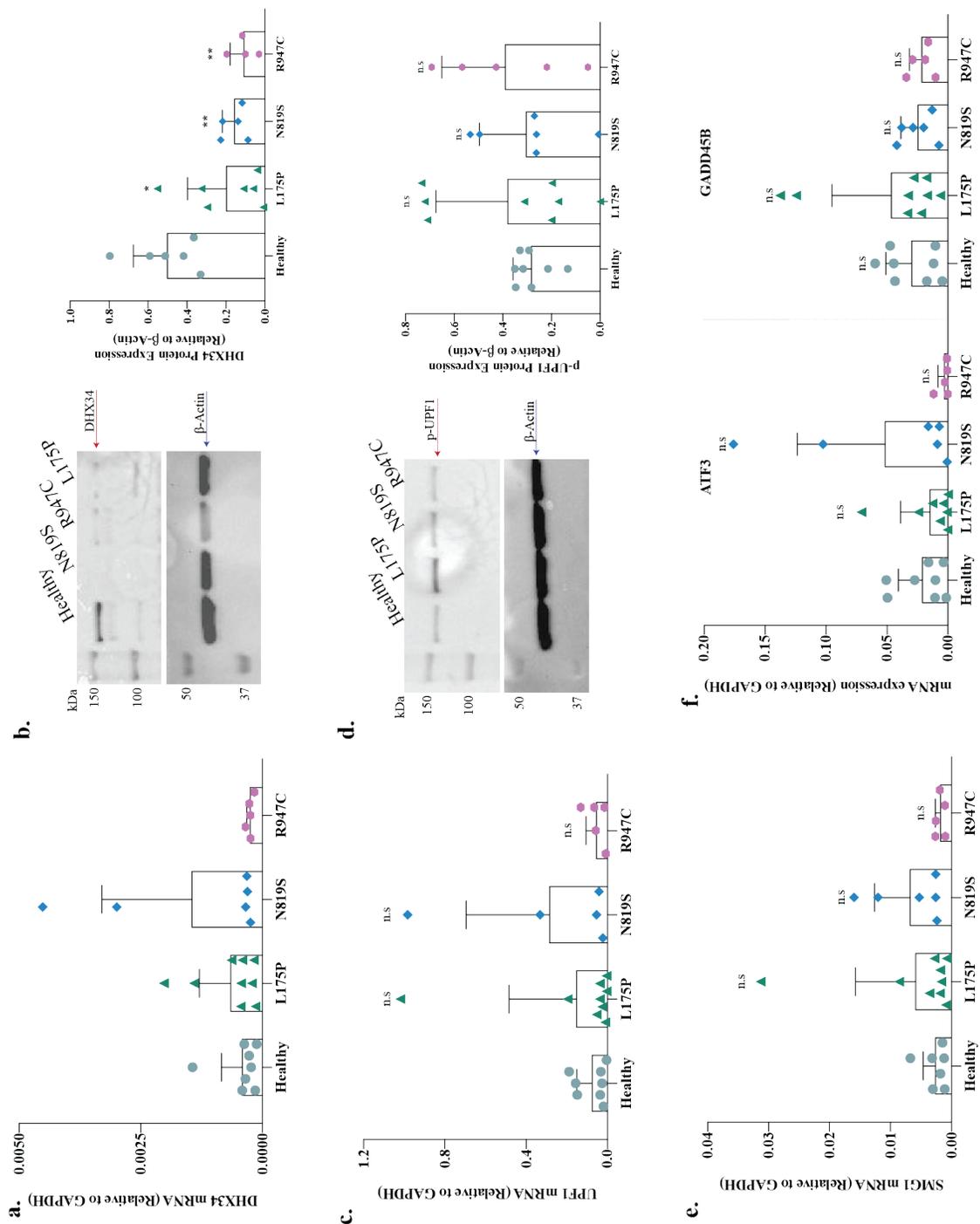


Figure 7. Rare variants in DHX34 decrease DHX34 protein levels without impacting NMD in endothelial cells. Patient-specific endothelial cells generated from iPSCs were used to test if the variants impacted DHX34 and NMD. Through qRT-PCR, no difference was detected in the level of DHX34 (**a**), but there was a significant decrease in protein levels as assessed by western blot (**b**, predicted molecular weight: DHX34 – 128 kDa, β -Actin – 42 kDa, t-test; * $p=0.014$, ** $p<0.01$). DHX34 is known to interact with UPF1, which itself is regulated by NMD. No changes were measured in the UPF1 mRNA levels in the patients (**c**), or in phosphorylated UPF1 (**d**, predicted molecular weight p-UPF1 – 125 kDa). The NMD factor SMG1 is also known to be regulated through the pathway itself, although no differences in expression of its mRNA (**e**) or two other known NMD targets – ATF3 and GADD45B (**f**), were detected between the patient samples and the healthy control.

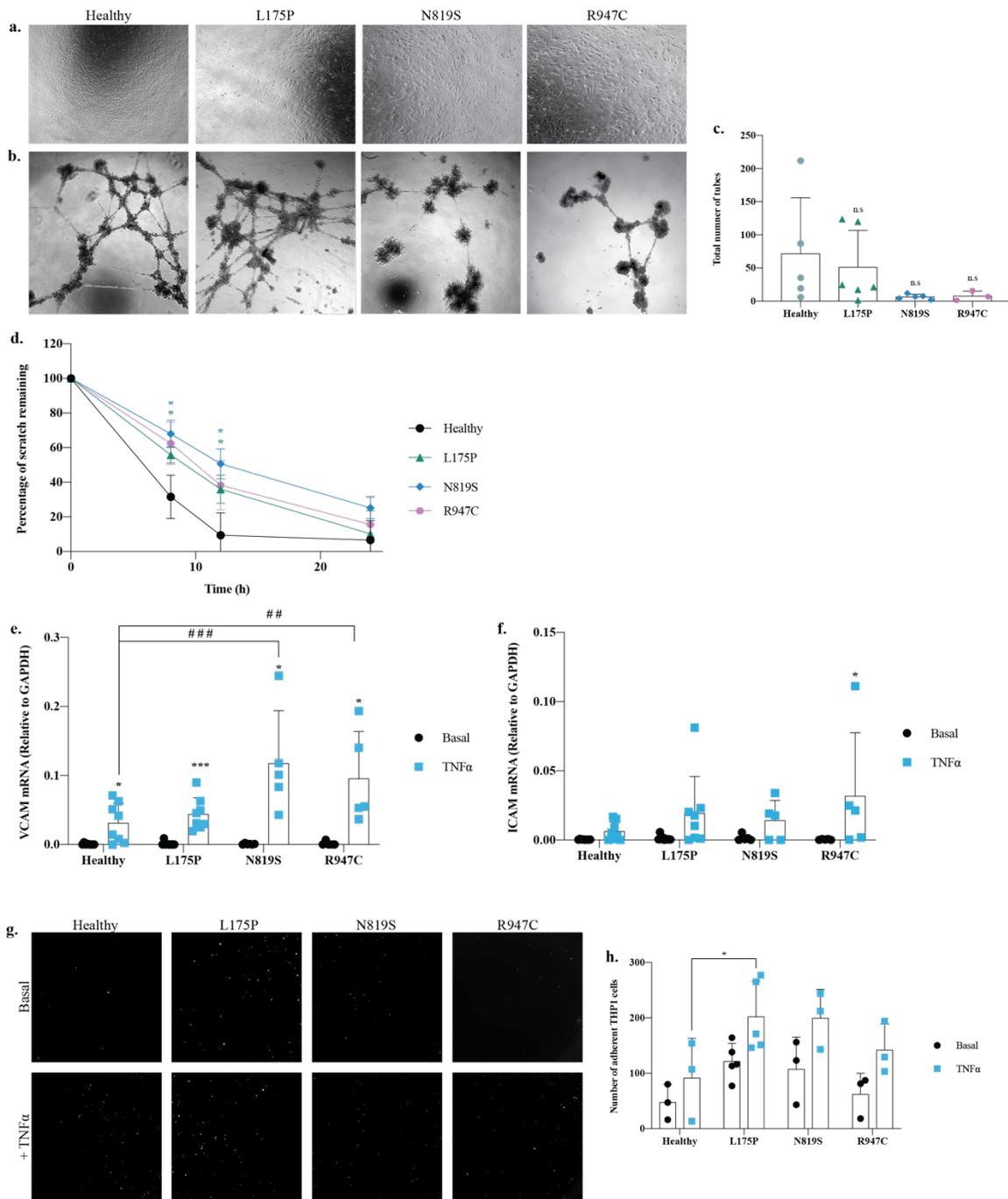


Figure 8. Patient endothelial cells with DHX34 variants have decreased migratory but increased adhesive capacities. The endothelial cells derived from patient iPSCs showed no major morphological differences when compared to those from the healthy iPSCs (**a**). The patient samples showed no significant differences in their tube formation capacity (**b-c**). Migration was assessed through a scratch assay, where both the N819S and L175P endothelial cells showed delayed migration in the first 12 hours when compared to the healthy endothelial cells (**d**, * $p < 0.05$). mRNA expression of VCAM increased in all samples following TNF α treatment (**e**, t-test; * $p < 0.05$, *** $p = 0.00253$) but the extent of VCAM mRNA expression in the N819S and R947C samples after treatment was greater than what was seen in the control (**e**, 2-way ANOVA; ### $p = 0.0001$, ## $p = 0.0043$). A similar trend was seen for ICAM although no significant increase in expression was measured within samples following TNF α treatment except in R947C (**f**, 2-way ANOVA; * $p = 0.044$). The function of the adhesion markers was measured through the THP1 adhesion assay, where adherent fluorescently-labelled THP1 monocytes were visualized (**g**) and then quantified (**h**). There was no significant increase in adherent THP1 following TNF α treatment, although the number of THP1 cells that adhered to the TNF α -treated L175P cells was significantly greater than what was measured in the presence of the healthy endothelial cells (**h**, * $p = 0.0227$).

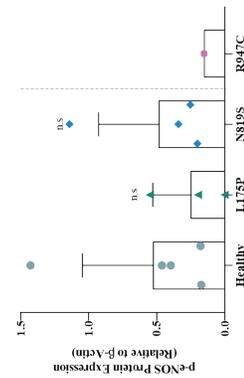
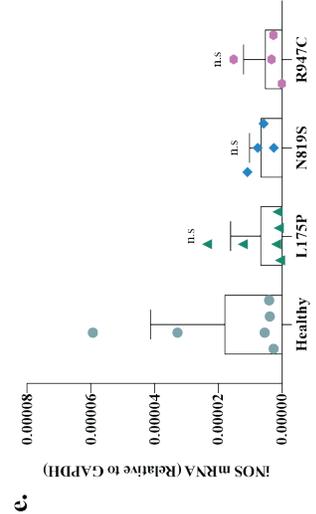
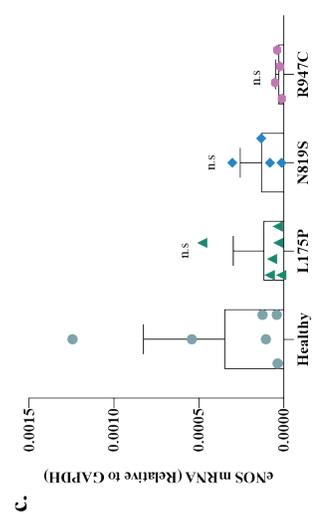
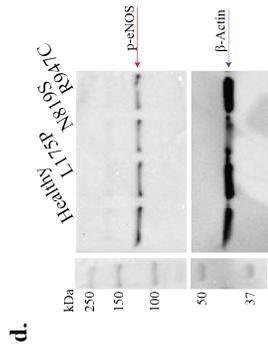
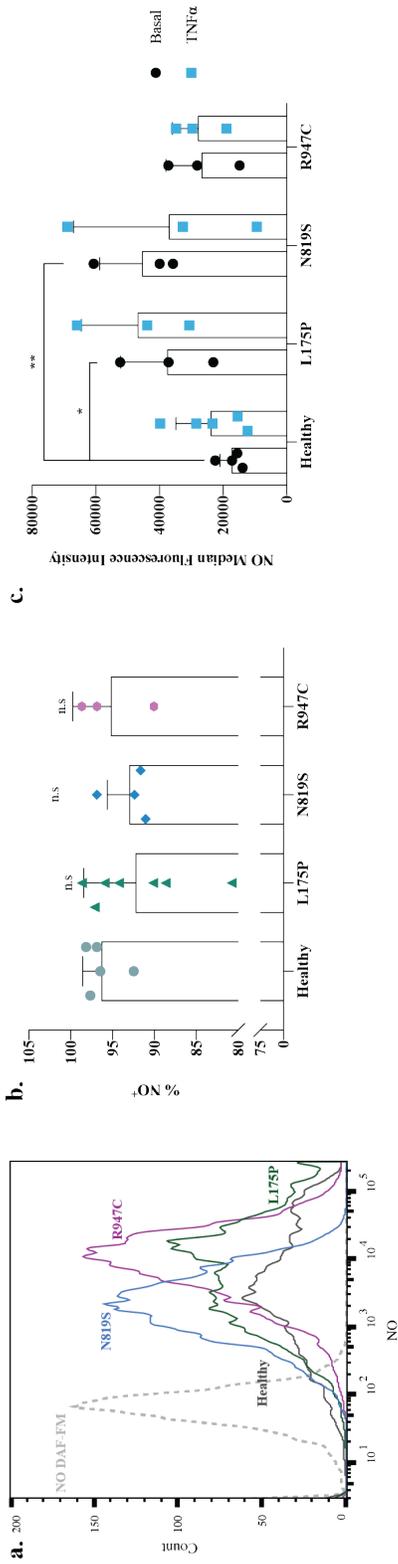
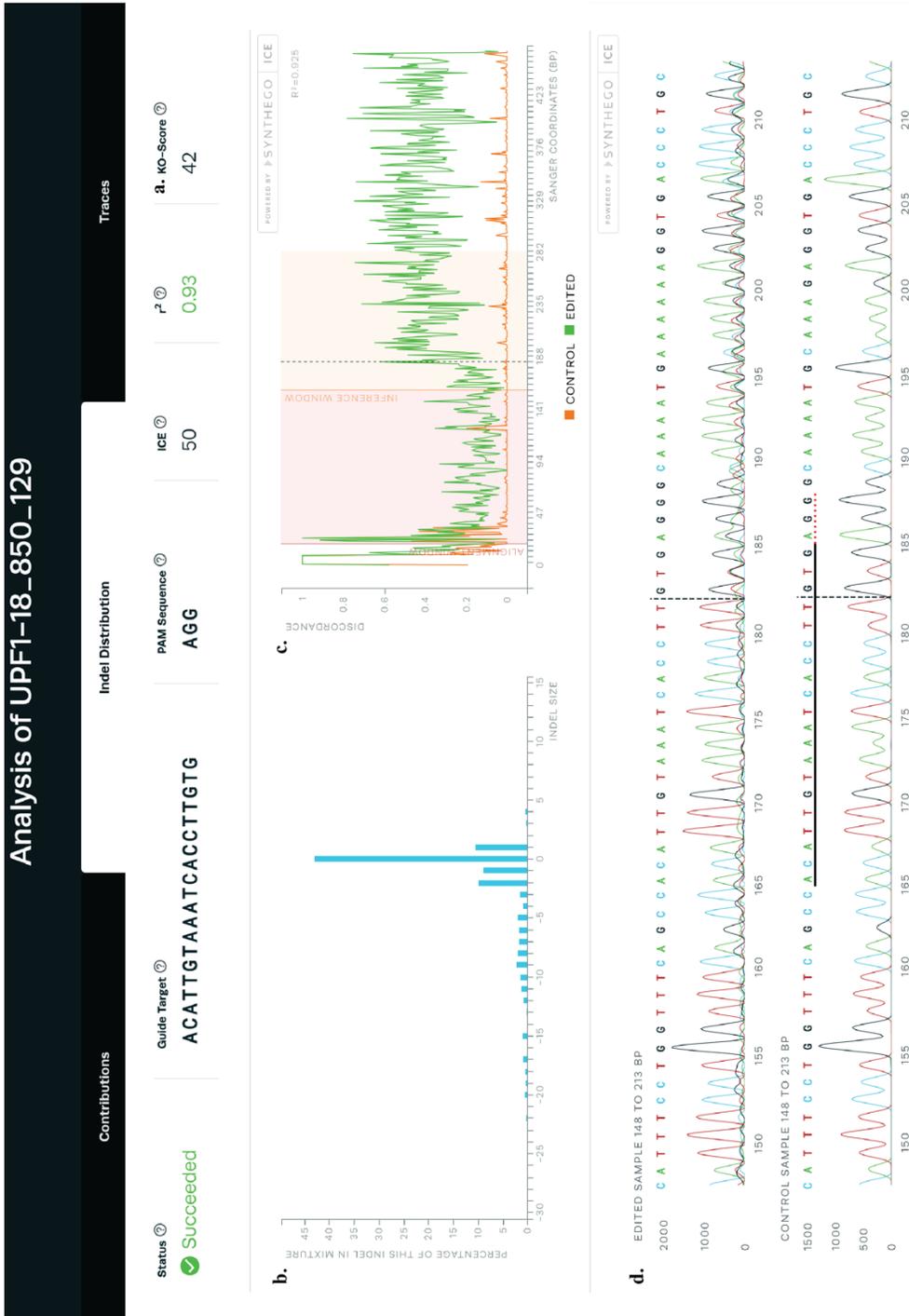
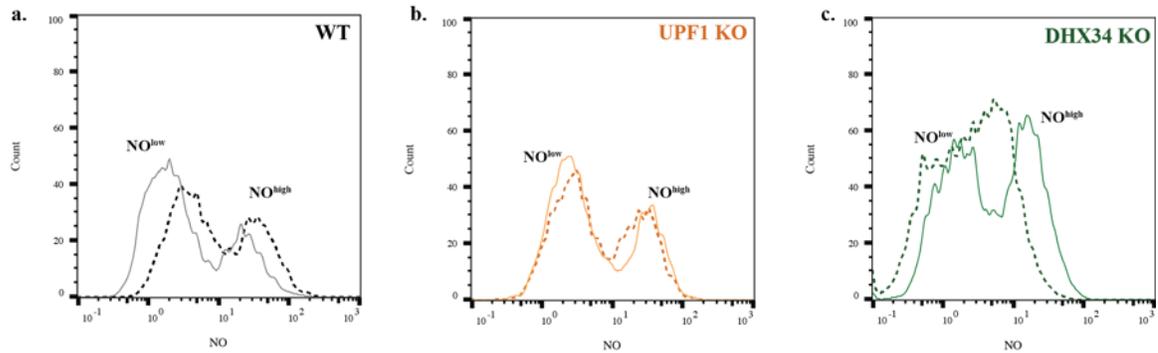


Figure 9. Differences in NO production of patient-endothelial cells is not a result of differences in eNOS. DAF-FM was used to measure NO production in patient iPSC-derived endothelial cells by adding it to the culture media and then quantifying with flow cytometry (a). There was no difference between the patient samples and healthy control when the percentage of NO⁺ cells was measured (b); however, the mean intensity fluorescence was elevated in both the L175P and N819S endothelial cells under basal conditions (c, t-test; *p=0.041 **p=0.0091). No differences were noticed within the samples due to TNF α treatment (p>0.05). There was also no difference between the patient samples and the control with respect to eNOS mRNA levels (c) or phosphorylation status (d, predicted molecular weight: p-eNOS – 140 kDa, β -Actin – 42 kDa, t-test; p>0.05, R947C only has n=1 at this time) or in the iNOS mRNA levels (e) between the patient samples and the control (t-test; p>0.05).

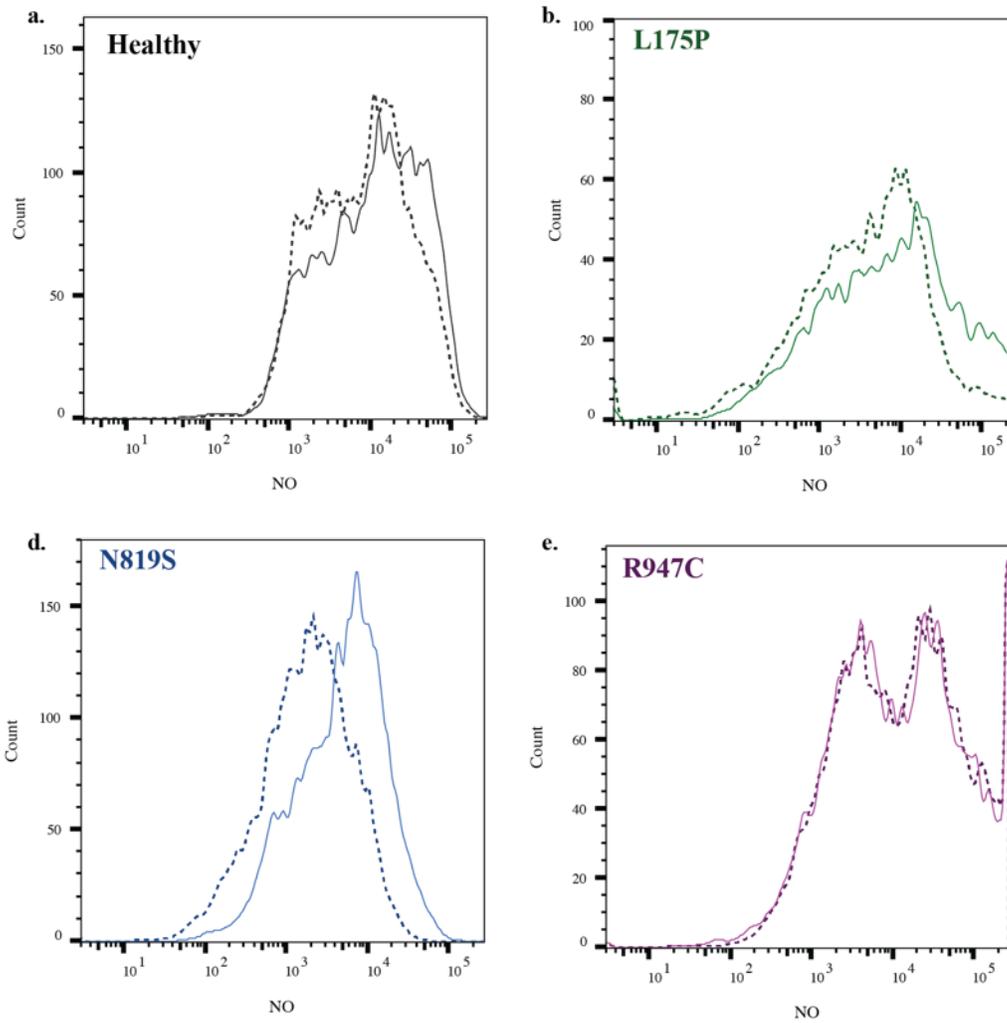
Supplementary Figure 1. Synthego analysis of the DHX34 KO pool. Synthego used CRISPR/Cas9 to induce indel formation in DHX34 such that the gene would be knocked out. The KO score of 50 (**a**) is the proportion of indels that indicate a frameshift or are >21bp in length, assuming they all fall within the coding region. The knockout process was validated on their end through analyzing the indels. The indel plot (**b**) shows the inferred distribution of indels in the entire population of genome along with the percentage of genomes that contain it. The discordance plot (**c**) shows the level of alignment per base between the wildtype and the edited sample, where the inference window represents the area around the cut site. The r^2 value of 0.71 represents how well the proposed indel distribution fits with the sequencing results. They also performed sanger sequencing (**d**) that showed mixed base calls in the edited sample as a result of the indel formation. Image taken from the Synthego ICE analysis page (<https://ice.synthego.com/#!/analyze/results/a5mpwcf66t9gsjru>).



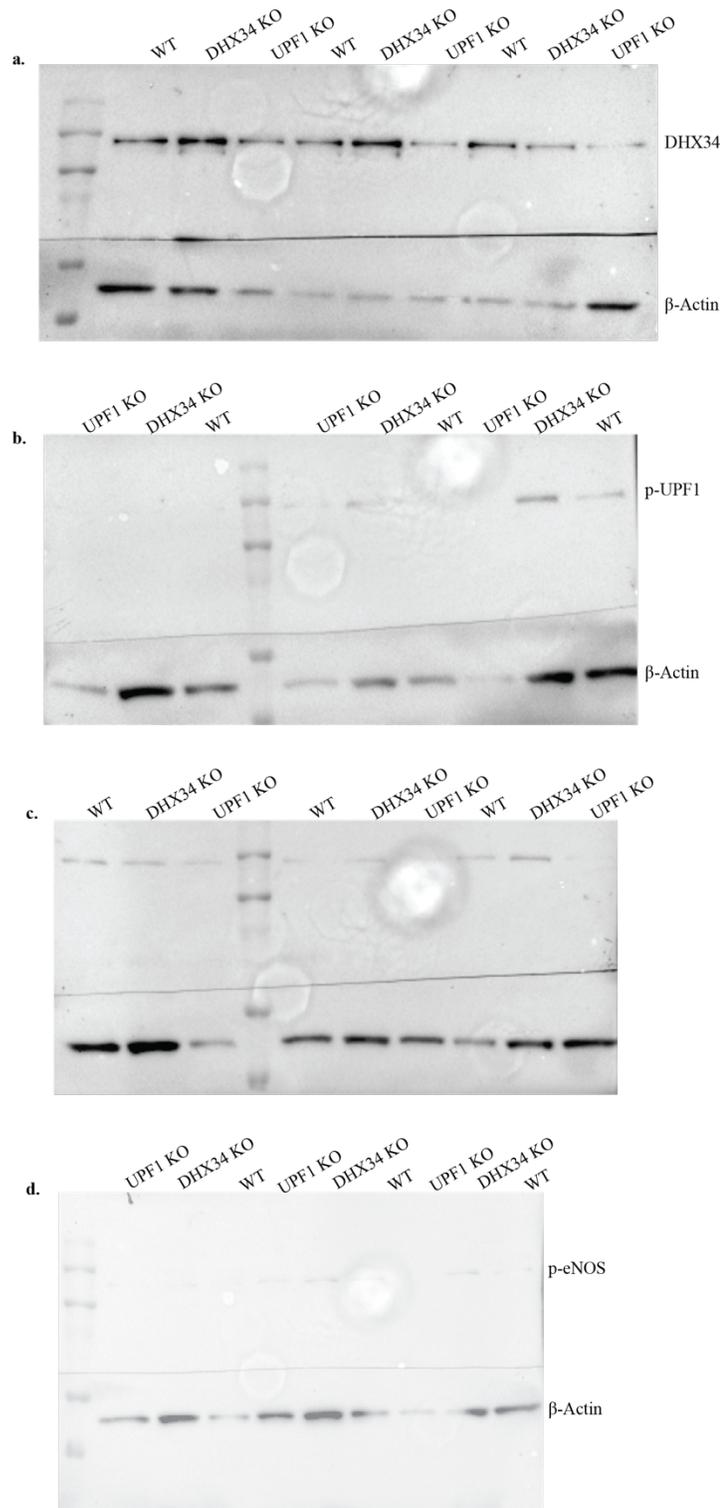
Supplementary Figure 2. Synthego analysis of the UPF1 KO pool. Synthego used CRISPR/Cas9 to induce indel formation in UPF1 such that the gene would be knocked out. The KO score of 42 (**a**) is the proportion of indels that indicate a frameshift or are >21bp in length, assuming they all fall within the coding region. The knockout process was validated on their end through analyzing the indels. The indel plot (**b**) shows the inferred distribution of indels in the entire population of genome along with the percentage of genomes that contain it. The discordance plot (**c**) shows the level of alignment per base between the wildtype and the edited sample, where the inference window represents the area around the cut site. The r^2 value of 0.95 represents how well the proposed indel distribution fits with the sequencing results. They also performed sanger sequencing (**d**) that showed mixed base calls in the edited sample as a result of the indel formation. Image taken from the Synthego ICE analysis page (<https://ice.synthego.com/#!/analyze/results/972n6bybrhhfh72q>).



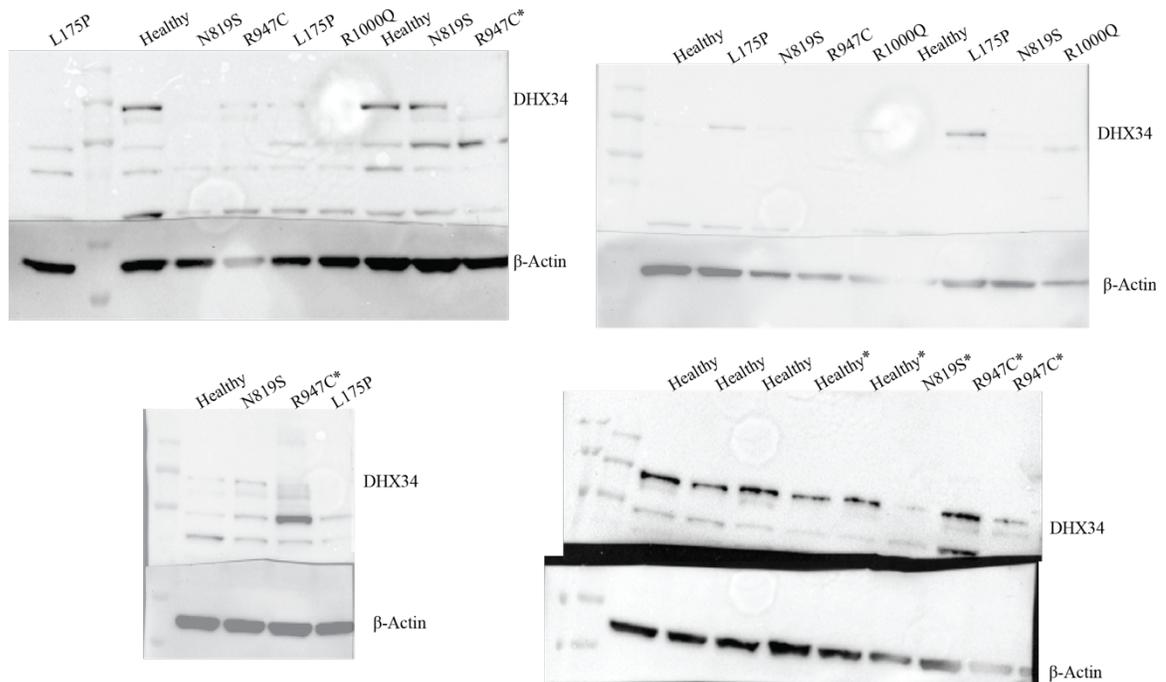
Supplementary Figure 3. Treatment with TNF α results in altered NMD signalling in the DHX34 KO pool. NO production was detected with DAF-FM and quantified using flow cytometry. The cells were measured at both basal (dotted line in each) and following treatment with TNF α for 24 hours (solid line in each). In the WT cells, treatment with TNF α resulted in a general decrease in NO as can be seen through a shift in the peaks towards the NO^{low} population (a). A similar trend was noted in the UPF1 KO pool, although not as dramatic (b). In the DHX34 KO pool, TNF α treatment resulted in the production of two distinct NO⁺ populations and a general shift towards the NO^{high} population (c), opposite to what was observed in the other samples.



Supplementary Figure 4. NO production in patient-derived endothelial cells increases following TNF α treatment. NO production was measured in patient-derived endothelial cells using DAF-FM under basal (dotted line in each plot) and TNF α -stimulated (solid lines in each plot) conditions. In all samples, there was a slight shift towards higher NO production under inflammatory conditions.

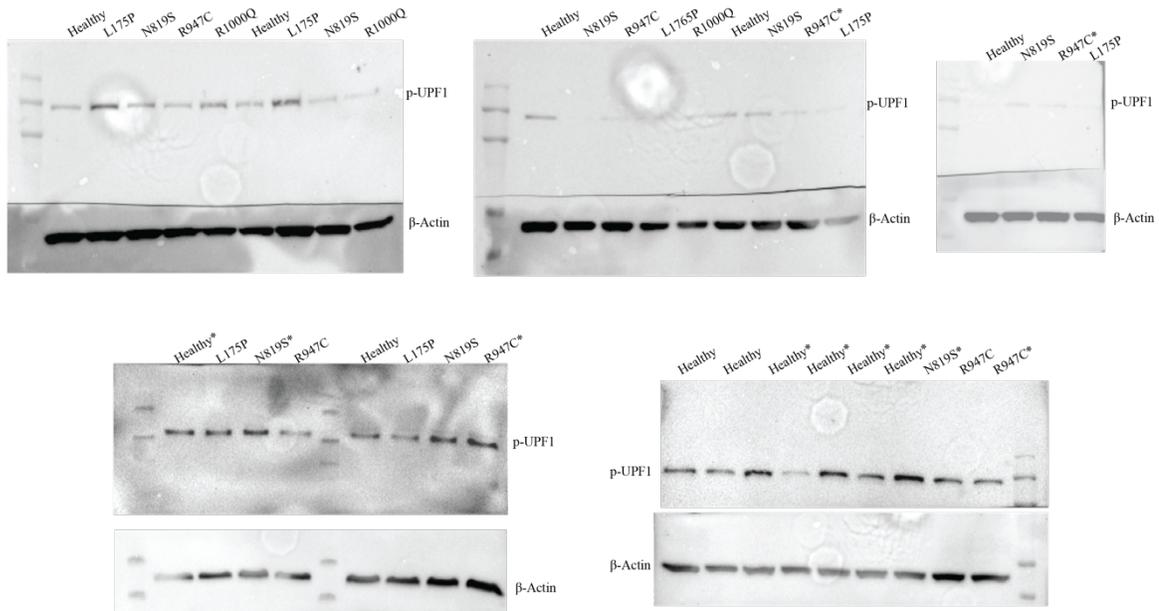


Supplementary Figure 5. Western blots with HUVEC KO pool samples. Full images of western blots performed with HUVEC protein lysates for DHX34 (**a**), p-UPF1 (**b**), e-NOS (**c**) and p-eNOS (**d**), with β -Actin as the loading control for all blots (lower band). The Kaleidoscope ladder (BioRad) was used. No alterations were made to the image other than to merge the ladder and the protein blots as they were taken with different settings.

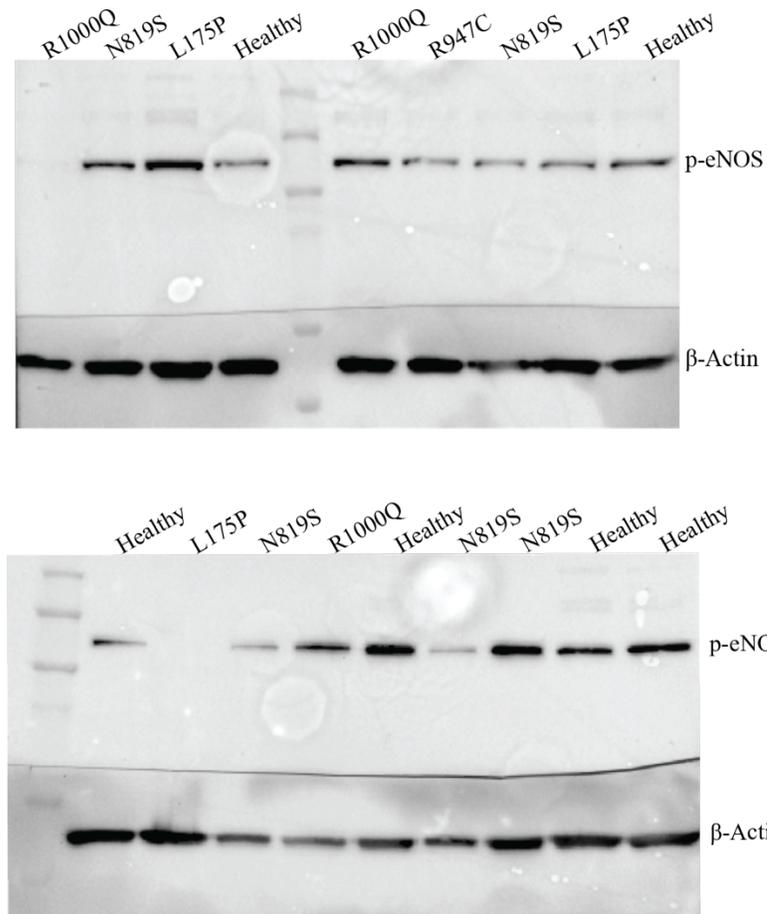


Supplementary Figure 6. DHX34 western blots with iPSC-derived endothelial cells.

Full images of western blots performed with iPSC-derived endothelial cells for DHX34 with β -Actin used as the loading control (lower band). The Kaleidoscope ladder (BioRad) was used. No alterations were made to the image other than to merge the ladder and the protein blots as they were taken with different settings. Wells are labelled with the samples loaded into the lanes. Those marked with * were from a second iPSC line derived from the patient whose data was not used for additional assays or the data was considered an outlier through outlier analysis.



Supplementary Figure 7. Phosphorylated UPF1 western blots with iPSC-derived endothelial cells. Full images of western blots performed with iPSC-derived endothelial cells for phosphorylated UPF1 (p-UPF1) with β -Actin used as the loading control (lower band). The Kaleidoscope ladder (BioRad) was used. No alterations were made to the image other than to merge the ladder and the protein blots as they were taken with different settings. Wells are labelled with the samples loaded into the lanes. Those marked with * were from a second iPSC line derived from the patient whose data was not used for additional assays or the data was considered an outlier through outlier analysis.



Supplementary Figure 8. Phosphorylated eNOS western blots with iPSC-derived endothelial cells. Full images of western blots performed with iPSC-derived endothelial cells for phosphorylated eNOS (p-eNOS) with β -Actin used as the loading control (lower band). The Kaleidoscope ladder (BioRad) was used. No alterations were made to the image other than to merge the ladder and the protein blots as they were taken with different settings. Wells are labelled with the samples loaded into the lanes. Those marked with * were from a second iPSC line derived from the patient whose data was not used for additional assays or the data was considered an outlier through outlier analysis.

Table 1. Association of NMD with disease

Gene	Affected cells	Mutation type	Associated Disease
β -globin (<i>HBB</i>)	Erythrocytes	5' PTC	β -thalassemia major
		3' PTC	β -thalassemia intermedia
von Willebrand factor (<i>vWF</i>)	Platelets and Endothelial cells	5' PTC	Type 3 von Willebrand disease
		3' PTC	Type 2A 3 von Willebrand disease
Myelin protein zero (<i>MPN</i>)	Schwann cells	5' PTC	Charcot-Marie-Tooth disease
		3' PTC	Congenital hypomyelinating neuropathy
Survival motor neuron gene (<i>SMN1</i>)	Motor neurons	5' PTC	Spinal muscular atrophy type III
		3' PTC	Spinal muscular atrophy type I
Cystic fibrosis transmembrane conductance regulator (<i>CFTR</i>)	Epithelial cells	5' PTC	Cystic fibrosis (severe)
		3' PTC	Cystic fibrosis (mild)
UPF3b*	Neurons	Loss-of-function mutations	Intellectual disability, autism, schizophrenia, attention-deficit hyperactivity disorder

Adapted from Bhuvangiri *et al.*³⁰⁸

* Not a target of NMD but a protein involved in the pathway itself

Table 2. qRT-PCR primers

Gene	Forward Primer 5' → 3'	Reverse Primer 5' → 3'
GAPDH	TCCCTGAGCTGAACGGGAAG	GGAGGAGTGGGTGTCGCTG T
VCAM	GTCTCCAATCTGAGCAGCAA	TGGGAAAAACAGA AAAGAGGTG
ICAM	AGGGTAAGGTTCTTGCCAC	TGCTATTCAAACCTGCCCTGA
NOS1	TCTCCTCCTACTCTGACTCC	TTGTGGACATTGGATAGACC
NOS2	TCTTGGTCAAAGCTGTGCTC	CATTGCCAAACGTACTGGTC
NOS3	AGGAACCTGTGTGACCCTCA	CGAGGTGGTCCGGGTATCC
DHX34	ACATGGAGCCCAGGATCA	GAAACCGCCATCCTCTACCT
ATF3	ACTTCCGAGGCAGAGACCTG	GGCCAGACAAACAGCCC
GADD45B	AGACAGGACCCCCGAAGCT	GGGCTTGCAGTCAGTCTCAC T
UPF1	ATGAGCGTGGAGGCGTACGGGCCAG CTCGCAG	ATACTGGGACAGCCCCGTC ACCCCGCCATG
SMG1	CATCAGCCGACCCAGATACTT	GTTGCAGTCCATAAGAGGA AGAA

Table 3. Western blot antibodies

Protein	Species	Concentration	Source
DHX34	Rabbit	1:500	Abcam (cat. 94989)
P-UPF1 (Ser1127)	Rabbit	1:1,000	Santa Cruz Biotechnology (cat. 166092)
eNOS	Mouse	1:500	BD Biosciences (cat. 610296)
P-eNOS (Ser1177)	Rabbit	1:250	Millipore Sigma (cat. 07-428-1)
β-actin	Mouse	1:10,000	Abcam (cat. 8224)

Table 4. Clinical presentation of patients

Patient	Sex	Age of disease-onset	# of vessels affected
L175P	Female	32	1
N819S	Male	37	2
R947C	Male	38	3
R1000Q	Female	45	1

Table 5. Sanger Sequencing Primers

Variant	Forward Primer 5' → 3'	Reverse Primer 5' → 3'
L175P	CCTGTTGCACTACCTGGACT	TGCACGCCACATGACTGAA
N819S	CTGCCAGCTCAGCCCAG	GGTTCGTGTGTTGGAGGACA
R947C	TCTGCCCCCTCTCTTTCAGT	GTGGAGTGAATGGGAAGCCA
R1000Q	GATGGGGGTGGGTTATCAGG	CCGTGAGGCAGTTGTAGGTG

CHAPTER 5: CONCLUSIONS & FUTURE DIRECTIONS

5.1 Patient-specific models as a means to better study the genetic component of disease etiology and treatment

The development of iPSC technology has created unprecedented opportunities for the disease modelling field. iPSCs generated from patients harbour the same genetic mutations as their source and therefore stand as an indispensable tool for elucidating the genetic contribution to many diseases. In this work, I've shown that with only a small number of PBMCs, iPSCs can be generated using the non-integrative Sendai virus. The patient-specific iPSCs can be differentiated towards the cell types of interest so that the mechanism through which specific mutations impact that cell type can be determined *in vitro*.

Reprogramming of human somatic cells using the Yamanaka factors has been performed for over a decade, but one concern that remains is the use of proto-oncogenes, such as cMyc and Klf4, in the reprogramming cocktail²⁴. Their role in promoting proliferation and self-renewal are important for reprogramming processes; however, the potential for these factors to be reactivated in the terminally differentiated cells remains a limitation to their clinical applications¹⁵¹. Using non-integrative approaches, such as the Sendai virus, reduces this risk substantially as there are no additional copies of the genes in the genome. Additionally, in the case of the Sendai virus, the virus itself is replication deficient, therefore the reprogramming factors get diluted over the initial passages, resulting in the lack of exogenous source of reprogramming factors after several

weeks^{175,339,340}. However, this method still remains imperfect as some groups have shown evidence that the Sendai virus impacts the iPSCs even after it cannot be detected in the cells¹⁷⁶.

Improving somatic cell reprogramming with non-integrative methods with genes that harbour no oncogenic potential will significantly improve their clinical potential in the future. Applying the method Mai, *et al.* used for Oct4 to the other Yamanaka factors may help to identify novel pathways or protein interactors that could be used in place of current reprogramming factors to the same end¹⁴⁷. Although human SCNT is not possible in Canada¹⁵, studying the mechanism through which it occurs in the mouse system could provide new insights to the process. Reprogramming with SCNT occurs very quickly (<22 hours) with factors that exist in the oocytes¹³³. Identifying the factors that are involved in SCNT reprogramming may identify alternative factors that could be used for somatic cell reprogramming in other contexts.

Another major consideration for the use of iPSCs in disease modelling is their ability to be differentiated efficiently into the cell type of interest. Many differentiation protocols exist with varying effectiveness, which are partially a result of an inconsistent standard for reporting on differentiation efficiency. In this work, I've shown a corrected method to optimize and improve upon a previously published endothelial differentiation protocol which in my hands, did not work as it had been reported. The optimized protocol worked well and allowed for the subsequent experiments to be performed; however, more work should be done to better characterize the subtype of endothelial cells that were generated. The cells were used to model CAD, which typically develops under conditions that promote arterial endothelial cell dysfunction, and therefore, using arterial endothelial

cells specifically would improve the model. For many cell types, endothelial cells included, there are multiple differences in phenotype and specialized functions across the subtypes based on the tissue and site of origin. Recently, more protocols have been developed that focus on these differences to create a more specific and homogeneous population^{218,341}. When disease modelling, these differences are important to consider as they likely play an important role in the overall clinical manifestations of the disease and as such, should not be overlooked when designing *in vitro* disease models. As the field moves forward, using more efficient differentiation protocols that include stages for specialization towards a specific subtype and applying them to 3D organoid systems will greatly improve not only the *in vitro* models, but is also a critical step when considering future regenerative medicine applications¹³².

5.2 Role of DHX34 in endothelial functions associated with early-onset coronary artery disease

NMD plays an important role in regulating the translation of cellular transcripts, particularly those with a PTC. Through limiting the production of PTC-containing transcripts, NMD protects cells from the accumulation of aberrant proteins. NMD has a role in regulating the translation of a large number of naturally occurring transcripts as well, positioning it as an important component of regular cellular function.

In 2014, Hug, *et al.* provided evidence that DHX34 plays an important role in the process of NMD by promoting the transition from the SURF complex to the DECID complex, thereby resulting in NMD activation³¹⁷. Two years later, Melero, *et al.* showed

that in the DECID complex, DHX34 acts as a scaffolding protein to further NMD activation by supporting the SMG1-mediated phosphorylation of UPF1³¹⁸. These two publications support the importance of DHX34 in the NMD pathway; however, the reliance of the multiple independent NMD pathways on DHX34 remains unknown. Additionally, although NMD is known to regulate a number of naturally occurring transcripts, very little has been done in the way of identifying these transcripts across multiple cell types. Considering the various pathways and their reliance on different NMD factors, elucidating the cell-specific differences in NMD is an important step to appreciate the role this process plays in both normal and diseased states.

In this work, I used knockout HUVEC cell pools generated by Synthego along with patient iPSC-derived endothelial cells to better understand the role of DHX34 in endothelial cells and the impact it may have on endothelial function associated with EOCAD. Although the impact on the tested NMD targets was inconclusive, there were clear functional ramifications from altering DHX34. Functionally, both the knockout cells and the patient-derived endothelial cells differed from the controls in migration, NO signaling, and monocyte adhesion, all of which are important aspects of maintaining an anti-atherosclerotic environment *in vivo*. The minimal measured impact on NMD targets in the patient cells along with the more severe functional phenotypes observed suggest that in endothelial cells, DHX34 could have an important role in endothelial function. Investigating whether this involves regulating pro-atherogenic transcripts is an important future direction of this work. Because of the minimal work that has been done on DHX34 itself, it is important to also consider that the observed changes in phenotype could also be a result of an NMD-independent function of DHX34.

Further work is required to fully understand the mechanisms through which the observed phenotypic changes occurred. The polygenic nature of CAD is an important consideration in this study and as such, the contribution of multiple variants that the selected patients harbour should not be ignored as potential contributor to the observed phenotype. The creation of isogenic iPSC lines using CRISPR/Cas9 would be the ideal way to dissect the specific role of the variants in the observed phenotypes. Considering the availability of the patient samples, both introducing the variants into an unaffected control line as well as correcting the variants in the patient lines would provide a definitive answer to this question. With the creation of the iPSC lines, it would also be possible to test the potential role of DHX34 in additional cellular compartments that are involved in CAD development.

The DHX34 variants were in part selected because of the pre-existing association of SMG6 SNPs with an increase CAD risk³²². As has been mentioned previously, the exact mechanism through which the vast majority of SNPs identified by GWAS impact disease risk remains unknown. Using the platform shown in this work, in combination with the methods used by Gupta, *et al.*³⁰² and Lo Sardo, *et al.*²⁹¹ future studies could be done to better understand whether the SNPs associated with SMG6 do in fact impact NMD processes and/or whether they result in similarly altered endothelial function as was shown here.

5.3 Significance and future directions

Through this work, I have uncovered a previously unknown role of DHX34 in endothelial function and a potential impact on CAD development. Through the use of patient-specific iPSC lines generated from PBMCs with the Sendai virus, an endothelial model was generated which allowed for the investigation of DHX34's role in endothelial function. The impact on key anti-atherosclerotic functions of the endothelial cells suggests an important role of this protein in CAD prevention and highlights the importance of continuing work on this aspect of disease development.

Further work needs to be done to fully understand the mechanism by which these observed changes occur. Through the creation of additional healthy iPSC lines, as well as isogenic lines with the use of CRISPR/Cas9 for the specific DHX34 variants, a better idea of the mechanism will emerge. Furthermore, the use of a differentiation protocol that generates a homogeneous arterial endothelial cell population will greatly benefit this work and greatly advance the current work.

Despite the additional studies that are required to elucidate an exact mechanism and further our understanding of NMD's role in CAD, the work done here stands as an important starting point for future work. Elucidating the role of DHX34, and NMD as a whole, in CAD development could allow for the identification of a novel pathway that could potentially be targeted as a means to prevent and/or treat CAD, a leading cause of global mortality.

BIBLIOGRAPHY

1. De Los Angeles, A. *et al.* Hallmarks of pluripotency. *Nature* **525**, 469–478 (2015).
2. Thomson, J. a. *et al.* Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science (80-.)*. **282**, 1145–1147 (1998).
3. Stevens, L. C. Studies on Transplantable Testicular Teratomas of Strain 129 Mice. *J. Natl. Cancer Inst.* **20**, 1257–1275 (1958).
4. Evans, M. J. & Kaufman, M. H. Establishment in Culture of Pluripotential Cells from Mouse Embryos. *Nature* **292**, 154–156 (1981).
5. Martin, G. R. Isolation of a Pluripotent Cell Line from Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells. *Proc. Natl. Acad. Sci.* **78**, 7634–7638 (1981).
6. Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. Formation of Germ-line Chimaeras from Embryo-Derived Teratocarcinoma Cell Lines. *Nature* **309**, 255–256 (1984).
7. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. Derivation of Completely Cell Culture-Derived Mice from Early-Passage Embryonic Stem Cells. *Proc. Natl. Acad. Sci.* **90**, 8424–8428 (1993).
8. Williams, R. L. *et al.* Myeloid Leukaemia Inhibitory Factor Maintains the Developmental Potential of Embryonic Stem Cells. *Nature* **336**, 684–687 (1998).
9. Papp, B. & Plath, K. Epigenetics of Reprogramming to Induced Pluripotency. *Cell* **152**, 1324–1343 (2013).
10. Nichols, J. & Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**, 487–492 (2009).

11. Brons, I. G. M. *et al.* Derivation of Pluripotent Epiblast Stem Cells from Mammalian Embryos. *Nature* **448**, 191–195 (2007).
12. Tesar, P. J. *et al.* New Cell Lines from Mouse Epiblast Share Defining Features with Human Embryonic Stem Cells. *Nature* **448**, 196–199 (2007).
13. Robertson, J. A. Human Embryonic Stem Cell Research: Ethical and Legal Issues. *Nat. Genet.* **11**, 262–283 (2001).
14. Minister of Justice. *Assisted Human Reproduction Act.* (2012).
15. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada & Social Sciences and Humanities Research Council of Canada. *Ethical conduct for research involving humans.* (2014).
16. Gurdon, J. B., Elsdale, T. R. & Fischberg, M. Sexually Mature Individuals of *Xenopus laevis* from the Transplantation of Single Somatic Nuclei. *Nature* **182**, (1958).
17. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–76 (2006).
18. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–72 (2007).
19. National Institutes of Health. NIH Human Embryonic Stem Cell Registry.
20. Kiskinis, E. *et al.* Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell* **14**, 781–795 (2014).
21. Wainger, B. J. *et al.* Intrinsic Membrane Hyperexcitability of Amyotrophic Lateral Sclerosis Patient-Derived Motor Neurons. *Cell Rep.* **7**, 1–11 (2014).
22. McNeish, J., Gardner, J. P., Wainger, B. J., Woolf, C. J. & Eggan, K. From Dish to

- Bedside: Lessons Learned while Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. *Cell Stem Cell* **17**, 8–10 (2015).
23. Passier, R., Orlova, V. & Mummery, C. Complex Tissue and Disease Modeling using hiPSCs. *Cell Stem Cell* **18**, 309–321 (2016).
 24. Tapia, N. & Schöler, H. R. Molecular Obstacles to Clinical Translation of iPSCs. *Cell Stem Cell* **19**, 298–309 (2016).
 25. Plath, K. & Lowry, W. E. Progress in Understanding Reprogramming to the Induced Pluripotent State. *Nat. Rev. Genet.* **12**, 253–265 (2011).
 26. Hanna, J. *et al.* Direct Cell Reprogramming is a Stochastic Process Amenable to Acceleration. *Nature* **462**, 595–601 (2009).
 27. Buganim, Y. *et al.* Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell* **150**, 1209–1222 (2012).
 28. Smith, Z. D., Nachman, I., Regev, A. & Meissner, A. Dynamic Single-Cell Imaging of Direct Reprogramming Reveals an Early Specifying Event. *Nat. Biotechnol.* **28**, 521–526 (2010).
 29. Li, R. *et al.* A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts. *Cell Stem Cell* **7**, 51–63 (2010).
 30. Samavarchi-Tehrani, P. *et al.* Functional Genomics Reveals a BMP-Driven Mesenchymal-to-Epithelial Transition in the Initiation of Somatic Cell Reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
 31. Wernig, M., Meissner, A., Cassady, J. P. & Jaenisch, R. c-Myc Is Dispensable for Direct Reprogramming of Mouse Fibroblasts. *Cell Stem Cell* **2**, 10–12 (2008).

32. Nakagawa, M. *et al.* Generation of Induced Pluripotent Stem Cells Without Myc From Mouse and Human Fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2007).
33. Theunissen, T. W. & Jaenisch, R. Molecular Control of Induced Pluripotency. *Cell Stem Cell* **14**, 720–734 (2014).
34. Li, S. *et al.* Association of plasma PCSK9 levels with white blood cell count and its subsets in patients with stable coronary artery disease. *Atherosclerosis* **234**, 441–5 (2014).
35. Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells. *Cell Stem Cell* **10**, 440–454 (2012).
36. Masui, S. *et al.* Pluripotency Governed by Sox2 via Regulation of Oct3/4 Expression in Mouse Embryonic Stem Cells. *Nat. Cell Biol.* **9**, 625–635 (2007).
37. Chen, X. *et al.* Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. *Cell* **133**, 1106–1117 (2008).
38. Sridharan, R. *et al.* Role of the Murine Reprogramming Factors in the Induction of Pluripotency. *Cell* **136**, 364–377 (2009).
39. Chambers, I. *et al.* Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells. *Cell* **113**, 643–655 (2003).
40. Loh, Y.-H. *et al.* The Oct4 and Nanog Transcription Network Regulates Pluripotency in Mouse Embryonic Stem Cells. *Nat. Genet.* **38**, 431–440 (2006).
41. Boyer, L. A. *et al.* Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* **122**, 947–956 (2005).
42. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and Impediments of the

- Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* **151**, 994–1004 (2012).
43. Schwarz, B. A. *et al.* Prospective Isolation of Poised iPSC Intermediates Reveals Principles of Cellular Reprogramming. *Cell Stem Cell* **23**, 289–305.e5 (2018).
 44. Carey, B. W. *et al.* Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell* **9**, 588–598 (2011).
 45. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative Expression of Oct-3/4 Defines Differentiation, Dedifferentiation or Self-Renewal of ES Cells. **24**, 2–6 (2000).
 46. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–90 (2010).
 47. Polo, J. M. *et al.* Cell type of Origin Influences the Molecular and Functional Properties of Mouse Induced Pluripotent Stem Cells. *Nat. Biotechnol.* **28**, 848–855 (2010).
 48. Rouhani, F. *et al.* Genetic Background Drives Transcriptional Variation in Human Induced Pluripotent Stem Cells. *PLoS Genet.* **10**, e1004432 (2014).
 49. Aoi, T. *et al.* Generation of Pluripotent Stem Cells from Adult Mouse Liver and Stomach Cells. *Science (80-.).* **321**, 699–702 (2008).
 50. Aasen, T. *et al.* Efficient and Rapid Generation of Induced Pluripotent Stem Cells from Human Keratinocytes. *Nat. Biotechnol.* **26**, 1276–1284 (2008).
 51. Shi, Y. *et al.* Induction of Pluripotent Stem Cells from Mouse Embryonic Fibroblasts by Oct4 and Klf4 with Small-Molecule Compounds. *Cell Stem Cell* **3**, 568–574 (2008).

52. Singhal, N. *et al.* Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141**, 943–955 (2010).
53. Huangfu, D. *et al.* Induction of Pluripotent Stem Cells by Defined Factors is Greatly Improved by Small-Molecule Compounds. *Nat. Biotechnol.* **26**, 795–797 (2008).
54. Mikkelsen, T. S. *et al.* Dissecting Direct Reprogramming Through Integrative Genomic Analysis. *Nature* **454**, 49–55 (2008).
55. Ma, T. *et al.* Atg5-Independent Autophagy Regulates Mitochondrial Clearance and is Essential for iPSC Reprogramming. *Nat. Cell Biol.* **17**, 1379–1387 (2015).
56. Cliff, T. S. *et al.* MYC Controls Human Pluripotent Stem Cell Fate Decisions through Regulation of Metabolic Flux. *Cell Stem Cell* **21**, 502–516.e9 (2017).
57. Varum, S. *et al.* Energy Metabolism in Human Pluripotent Stem Cells and Their Differentiated Counterparts. **6**, (2011).
58. Park, I.-H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134**, 877–86 (2008).
59. Sternecker, J. L., Reinhardt, P. & Schöler, H. R. Investigating Human Disease Using Stem Cell Models. *Nat. Rev. Genet.* **15**, 625–639 (2014).
60. Gaudelli, N. M. *et al.* Programmable Base Editing of A•T to G•C in Genomic DNA Without DNA Cleavage. *Nature* **551**, 464–471 (2017).
61. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable Editing of a Target Base in Genomic DNA Without Double-Stranded DNA Cleavage. *Nature* **533**, 420–424 (2016).
62. Hu, J. H. *et al.* Evolved Cas9 Variants with Broad PAM Compatibility and High DNA Specificity. *Nature* **556**, 57–63 (2018).

63. Howden, S. E. *et al.* A Cas9 Variant for Efficient Generation of Indel-Free Knockin or Gene-Corrected Human Pluripotent Stem Cells. *Stem Cell Reports* **7**, 508–517 (2016).
64. Tidball, A. M. *et al.* Rapid Generation of Human Genetic Loss-of-Function iPSC Lines by Simultaneous Reprogramming and Gene Editing. *Stem Cell Reports* **9**, 725–731 (2017).
65. Howden, S. E. *et al.* Simultaneous Reprogramming and Gene Correction of Patient Fibroblasts. *Stem Cell Reports* **5**, 1109–1118 (2015).
66. Wen, W. *et al.* High-Level Precise Knockin of iPSCs by Simultaneous Reprogramming and Genome Editing of Human Peripheral Blood Mononuclear Cells. *Stem Cell Reports* **10**, 1821–1834 (2018).
67. World Health Organization. *Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016.* (2018).
68. Papakonstantinou, N. a, Stamou, M. I., Baikoussis, N. G., Goudevenos, J. & Apostolakis, E. Sex differentiation with regard to coronary artery disease. *J. Cardiol.* **62**, 4–11 (2013).
69. Ghattas, A., Griffiths, H. R., Devitt, A., Lip, G. Y. H. & Shantsila, E. Monocytes in coronary artery disease and atherosclerosis: where are we now? *J. Am. Coll. Cardiol.* **62**, 1541–51 (2013).
70. Gutiérrez, E. *et al.* Endothelial dysfunction over the course of coronary artery disease. *Eur. Heart J.* **34**, 3175–81 (2013).
71. Choi, B.-J. *et al.* Coronary endothelial dysfunction in patients with early coronary artery disease is associate with the increase in intracascular lipid core plaque. *Eur.*

- Heart J.* **34**, 8 (2013).
72. Nora, J. J., Lortscher, R. H., Spangler, R. D., Nora, A. H. & Kimberling, W. J. Genetic-Epidemiologic Study of Early-Onset Ischemic Heart Disease. *Circulation* **61**, 503–508 (1980).
73. Glagov, S., Weisenberg, E., Zarins, C. K., Stankunavicius, R. & Kolettis, G. J. Compensatory Enlargement of Human Atherosclerotic Coronary Arteries. *N. Engl. J. Med.* **316**, 1371–1375 (1987).
74. Ross, R. & Glomset, J. A. Atherosclerosis and the Arterial Smooth Muscle Cell. *Science (80-.)*. **180**, 1332–1339 (1973).
75. Ross, R. The Pathogenesis of Atherosclerosis — An Update. *N. Engl. J. Med.* **314**, 488–500 (1986).
76. Ross, R. Atherosclerosis-an inflammatory disease. *N. Engl. J. Med.* **340**, 115–126 (1999).
77. Fuster, V., Badimon, L., Badimon, J. J. & Chesebro, J. H. The pathogenesis of coronary artery disease and the acute coronary syndromes. *N. Engl. J. Med.* **326**, 242–250 (1992).
78. Liao, J. K. Endothelium and acute coronary syndromes. *Clin. Chem.* **8**, 1799–1808 (1998).
79. Ross, R. The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature* **362**, 801–809 (1993).
80. Klingenberg, R., Hasun, M. & Corti, R. Inflammation and atherosclerosis. 39–58 (2012). doi:10.1007/978-3-7091-0338-8
81. Cooke, J. P. & Dzau, V. J. Nitric oxide synthase: role in the genesis of vascular

- disease. *Annu. Rev. Med.* **48**, 489–509 (1997).
82. Cines, D. B. *et al.* Endothelial cells in physiology and in the pathophysiology of vascular disorders. *Blood* **91**, 3527–3561 (1998).
83. Sullivan, G. The role of inflammation in vascular diseases. *J. Leukoc. Biol.* **67**, 591–602 (2000).
84. Ozaki, M. *et al.* Overexpression of Endothelial Nitric Oxide Synthase Accelerates Atherosclerotic Lesion Formation in apoE-deficient Mice. *J. Clin. Invest.* **110**, 331–340 (2002).
85. Förstermann, U. & Sessa, W. C. Nitric Oxide Synthases: Regulation and function. *Eur. Heart J.* **33**, 829–837 (2012).
86. Ghosh, S., Karin, M., Haven, N., Diego, S. & Jolla, L. Missing Pieces in the NF- κ B Puzzle Review. *Cell* **109**, 81–96 (2002).
87. Tedgui, A. & Mallat, Z. Cytokines in atherosclerosis: pathogenic and regulatory pathways. *Physiol. Rev.* **86**, 515–81 (2006).
88. Adams, W. J. *et al.* Functional vascular endothelium derived from human induced pluripotent stem cells. *Stem Cell Reports* **1**, 105–113 (2013).
89. Kodama, T., Reddy, P., Kishimoto, C. & Krieger, M. Purification and Characterization of a Bovine Acetyl Low Density Lipoprotein Receptor. *Proc. Natl. Acad. Sci.* **85**, 9238–9242 (2006).
90. Sakakura, K. *et al.* Pathophysiology of atherosclerosis plaque progression. *Heart. Lung Circ.* **22**, 399–411 (2013).
91. Li, A. C. & Glass, C. K. The Macrophage Foam Cell as a Target for Therapeutic Intervention. *Nat. Med.* **8**, 39–50 (2002).

92. Stary, H. C. Evolution and Progression of Atherosclerotic Lesions in Coronary Arteries of Children and Young Adults. *Arteriosclerosis* **9**, 119-32 (1989).
93. Simionescu, M. & Sima, A. V. Inflammation and atherosclerosis. 19–37 (2012). doi:10.1007/978-3-7091-0338-8
94. Insull, W. The pathology of atherosclerosis: plaque development and plaque responses to medical treatment. *Am. J. Med.* **122**, S3–S14 (2009).
95. Tabas, I. Macrophage Death and Defective Inflammation Resolution in Atherosclerosis. *Nat. Rev. Immunol.* **10**, 36–46 (2010).
96. Virmani, R., Kolodgie, F. D., Burke, A. P., Farb, A. & Schwartz, S. M. Lessons From Sudden Coronary Death. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1262–1275 (2000).
97. Libby, P. Mechanisms of Acute Coronary Syndromes and Their Implications for Therapy — NEJM. *N Engl J Med* **368**, 2004–2013 (2013).
98. Childs, B. *et al.* Senescent intimal foam cells are deleterious at all stages of atherosclerosis. *Science (80-.).* **42**, 3017–3021 (2016).
99. Ho, M. *et al.* Identification of endothelial cell genes by combined database mining and microarray analysis. *Physiol. Genomics* **13**, 249–262 (2003).
100. Heusch, G. *et al.* Cardiovascular Remodelling in Coronary Artery Disease and Heart Failure. *Lancet* **383**, 1933–1943 (2014).
101. Bui, Q. T., Prempeh, M. & Wilensky, R. L. Atherosclerotic plaque development. *Int. J. Biochem. Cell Biol.* **41**, 2109–13 (2009).
102. Libby, P., Ridker, P. M. & Hansson, G. K. Progress and challenges in translating the biology of atherosclerosis. *Nature* **473**, 317–25 (2011).

103. Farb, A. *et al.* Coronary Plaque Erosion Without Rupture into a Lipid Core. A Frequent Cause of Coronary Thrombosis in Sudden Coronary Death. *Circulation* **93**, 1354–63 (1996).
104. Lampugnani, M. G. *et al.* The Molecular Organization of Endothelial Cell to Cell Junctions: Differential Association of Plakoglobin, β -catenin, and α -catenin with Vascular Endothelial Cadherin (VE-cadherin). *J. Cell Biol.* **129**, 203–217 (1995).
105. Furie, B. & Furie, B. C. Mechanisms of Thrombus Formation. *N. Engl. J. Med.* **359**, 938–949 (2008).
106. Davies, M. J. Stability and Instability: Two Faces of Coronary Atherosclerosis. *Circulation* **94**, (1995).
107. Falk, E., Shah, P. K. & Fuster, V. Coronary Plaque Disruption. *Circulation* **92**, (1995).
108. Bowie, E. J. W. & Fuster, V. Resistance to Arteriosclerosis in Pigs with von Willebrand's Disease. *Acta Med Scand* **61**, 121–130 (1980).
109. Fuster, V. *et al.* Arteriosclerosis in Normal and von Willebrand Pigs: Long-Term Prospective Study and Aortic Transplantation Study. *Circ. Res.* **51**, 587–593 (1982).
110. Davies, M. J., Bland, J. M., Hangartner, J. R. W., Angelini, A. & Thomas, A. C. Factors Influencing the Presence or Absence of Acute Coronary Artery Thrombi in Sudden Ischaemic Death. *Eur. Heart J.* **10**, 203–208 (1989).
111. Ambrose, J. A. *et al.* Angiographic Progression of Coronary Artery Disease and the Development of Myocardial Infarction. *J. Am. Coll. Cardiol.* **12**, 56–62 (1989).
112. Vaughan, C. J., Gotto, A. M. & Basson, C. T. The Evolving Role of Statins in the Management of Atherosclerosis. *J. Am. Coll. Cardiol.* **35**, 1–10 (2000).

113. Nissen, S. E. *et al.* Effect of Very High-Intensity Statin Therapy on Regression of Coronary Atherosclerosis. *Jama* **295**, 1556 (2006).
114. Erdmann, J., Kessler, T., Munoz Venegas, L. & Schunkert, H. A Decade of Genome-wide Association Studies for Coronary Artery Disease: The Challenges Ahead. *Cardiovasc. Res.* **114**, 1241–1257 (2018).
115. Abifadel, M. *et al.* Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat. Genet.* **34**, 154–6 (2003).
116. Soria, L. F. *et al.* Association Between a Specific Apolipoprotein B Mutation and Familial Defective Apolipoprotein B-100. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 587–91 (1989).
117. Lehrman, M. A. *et al.* Mutation in LDL Receptor: Alu-Alu Recombination Deletes Exons Encoding Transmembrane and Cytoplasmic Domains. *Science (80-.).* **23**, 935–939 (1985).
118. Vane, J. R., Anggard, E. E. & Botting, R. Regulatory functions of the vascular endothelium. *N. Engl. J. Med.* **325**, (1990).
119. Barbieri, S. S. & Weksler, B. B. Tobacco Smoke Cooperates with Interleukin-1 β to Alter β -catenin Trafficking in Vascular Endothelium Resulting in Increased Permeability and Induction of Cyclooxygenase-2 Expression In Vitro and In Vivo. *FASEB J.* **21**, 1831–1843 (2007).
120. *Statistics Canada. Table 13-10-0096-10 Smokers, by age group.*
121. Plutzky, J. Inflammatory pathways in atherosclerosis and acute coronary syndromes. *Am. J. Cardiol.* **88**, (2001).
122. Getz, G. S. & Reardon, C. a. Animal models of atherosclerosis. *Arterioscler.*

- Thromb. Vasc. Biol.* **32**, 1104–15 (2012).
123. von Scheidt, M. *et al.* Applications and Limitations of Mouse Models for Understanding Human Atherosclerosis. *Cell Metab.* **25**, (2016).
124. Havel, R. J., Yamada, N., Shames, D. M. & Defect, G. Watanabe Heritable Hyperlipidemic Rabbit. *Arterioscler. Suppl.* **19**, 33–38 (1983).
125. Hu, W., Polinsky, P., Sadoun, E., Rosenfeld, M. E. & Schwartz, S. M. Atherosclerotic Lesions in the Common Coronary Arteries of ApoE Knockout Mice. **14**, 120–125 (2005).
126. Schwartz, S. M., Galis, Z. S., Rosenfeld, M. E. & Falk, E. Plaque Rupture in Humans and Mice. (2007). doi:10.1161/01.ATV.0000261709.34878.20
127. Nussenblatt, R. B., Bielekova, B., Childs, R., Krensky, A. & Strober, W. Meeting the Human Immunology Challenge - National Institutes of Health Center for Human Immunology Conference , September 2009. *Ann. N. Y. Acad. Sci.* **1200**, 1–23 (2010).
128. Majesky, M. W. Developmental Basis of Vascular Smooth Muscle Diversity. 1248–1258 (2007). doi:10.1161/ATVBAHA.107.141069
129. Davies, M. J., Woolf, N., Rowles, P. M. & Pepper, J. Morphology of the Endothelium over Atherosclerotic Plaques in Human Coronary Arteries. *Heart* **60**, 459–464 (1988).
130. Ross, R., Wright, T. N., Strandness, E. & Thiele, B. Human Atherosclerosis - Cell Constitution and Characteristics of Advanced. *Am. J. Pathol.* **114**, 79–93 (1984).
131. Rashid, S. T. *et al.* Modeling Inherited Metabolic Disorders of the Liver Using Human Induced Pluripotent Stem Cells. *J. Clin. Invest.* **120**, (2010).
132. Wimmer, R. A. *et al.* Human Blood Vessel Organoids as a Model of Diabetic

- Vasculopathy. *Nature* (2019). doi:10.1038/s41586-018-0858-8
133. Egli, D. *et al.* Reprogramming Within Hours Following Nuclear Transfer into Mouse but Not Human Zygotes. *Nat. Commun.* **2**, 410–488 (2011).
 134. Tachibana, M. *et al.* Human Embryonic Stem Cells Derived by Somatic Cell Nuclear Transfer. *Cell* **153**, 1228–1238 (2013).
 135. Tachibana, M. *et al.* Towards Germline Gene Therapy of Inherited Mitochondrial Diseases. *Nature* **493**, 627–631 (2013).
 136. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (80-.)*. **318**, 1917–20 (2007).
 137. Heng, J. C. D. *et al.* The Nuclear Receptor Nr5a2 Can Replace Oct4 in the Reprogramming of Murine Somatic Cells to Pluripotent Cells. *Cell Stem Cell* **6**, 167–174 (2010).
 138. Eguchi, A. *et al.* Reprogramming Cell Fate with a Genome-scale Library of Artificial Transcription Factors. *Proc. Natl. Acad. Sci.* **113**, E8257–E8266 (2016).
 139. Gao, Y. *et al.* Replacement of Oct4 by Tet1 during iPSC Induction Reveals an Important Role of DNA Methylation and Hydroxymethylation in Reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
 140. Long, Y., Wang, M., Gu, H. & Xie, X. Bromodeoxyuridine Promotes Full-Chemical Induction of Mouse Pluripotent Stem Cells. *Cell Res.* **25**, 1171–1174 (2015).
 141. Hou, P. *et al.* Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. *Science (80-.)*. **341**, 651–654 (2013).
 142. Tan, F., Qian, C., Tang, K., Abd-Allah, S. M. & Jing, N. Inhibition of Transforming Growth Factor β (TGF- β) Signaling Can Substitute for Oct4 Protein in

- Reprogramming and Maintain Pluripotency. *J. Biol. Chem.* **290**, 4500–4511 (2015).
143. Anokye-Danso, F. *et al.* Highly Efficient miRNA-Mediated Reprogramming of Mouse and Human Somatic Cells to Pluripotency. *Cell Stem Cell* **8**, 376–388 (2011).
144. Shu, J. *et al.* Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. *Cell* **153**, 963 (2013).
145. Montserrat, N. *et al.* Reprogramming of Human Fibroblasts to Pluripotency with Lineage Specifiers. *Cell Stem Cell* **13**, 341–350 (2013).
146. Miyoshi, N. *et al.* Reprogramming of Mouse and Human Cells to Pluripotency Using Mature MicroRNAs. *Cell Stem Cell* **8**, 633–638 (2011).
147. Mai, T. *et al.* NKX3-1 is Required for Induced Pluripotent Stem Cell Reprogramming and Can Replace OCT4 in Mouse and Human iPSC Induction. *Nat. Cell Biol.* **20**, 900–908 (2018).
148. González, F., Boué, S. & Belmonte, J. C. I. Methods for Making Induced Pluripotent Stem Cells: Reprogramming à la Carte. *Nat. Rev. Genet.* **12**, 231–242 (2011).
149. Yamanaka, S. Induced pluripotent stem cells: Past, present, and future. *Cell Stem Cell* **10**, 678–684 (2012).
150. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–7 (2007).
151. Hanna, J. *et al.* Treatment of Sickle Cell Anemia Mouse Model with iPS Cells Generated from Autologous Skin. *Science (80-.).* **318**, 1920–1923 (2007).
152. Kajii, K. *et al.* Virus-Free Induction of Pluripotency and Subsequent Excision of Reprogramming Factors. *Nature* **458**, 771–775 (2009).
153. Feng, B., Ng, J. H., Heng, J. C. D. & Ng, H. H. Molecules that Promote or Enhance

- Reprogramming of Somatic Cells to Induced Pluripotent Stem Cells. *Cell Stem Cell* **4**, 301–312 (2009).
154. Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. Induced Pluripotent Stem Cells Generated Without Viral Integration. *Science (80-.)*. 945–949 (2008).
155. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of Mouse Induced Pluripotent Stem Cells Without Viral Vectors. *Science (80-.)*. **319**, 447–451 (2008).
156. Gonzalez, F. *et al.* Generation of Mouse-Induced Pluripotent Stem Cells by Transient Expression of a Single Nonviral Polycistronic Vector. *Proc. Natl. Acad. Sci.* **106**, 8918–8922 (2009).
157. Wen, W. *et al.* Enhanced Generation of Integration-free iPSCs from Human Adult Peripheral Blood Mononuclear Cells with an Optimal Combination of Episomal Vectors. *Stem Cell Reports* **6**, 873–884 (2016).
158. Miura, K. *et al.* Variation in the Safety of Induced Pluripotent Stem Cell Lines. *Nat. Biotechnol.* **27**, 743–745 (2009).
159. Vangipuram, M., Ting, D., Kim, S., Diaz, R. & Schüle, B. Skin Punch Biopsy Explant Culture for Derivation of Primary Human Fibroblasts. *J. Vis. Exp.* **77**, 9–11 (2013).
160. Zhang, X. B. Cellular Reprogramming of Human Peripheral Blood Cells. *Genomics, Proteomics Bioinforma.* **11**, 264–274 (2013).
161. Kim, J. B. *et al.* Oct4-Induced Pluripotency in Adult Neural Stem Cells. *Cell* **136**, 411–419 (2009).

162. Zhu, S. *et al.* Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. *Cell Stem Cell* **7**, 651–655 (2010).
163. Giorgetti, A. *et al.* Generation of Induced Pluripotent Stem Cells from Human Cord Blood Using OCT4 and SOX2. *Cell Stem Cell* **5**, 353–357 (2009).
164. Eminli, S. *et al.* Differentiation Stage Determines Potential of Hematopoietic Cells for Reprogramming into Induced Pluripotent Stem Cells. *Nat. Genet.* **41**, 968–976 (2009).
165. Loh, Y. *et al.* Generation of induced pluripotent stem cells from human blood. *Blood* **113**, 5476–5479 (2009).
166. Okabe, M. *et al.* Definitive Proof for Direct Reprogramming of Hematopoietic Cells to Pluripotency. *Blood* **114**, 1764–1767 (2009).
167. Stadtfeld, M., Brennand, K. & Hochedlinger, K. Reprogramming of Pancreatic β Cells into Induced Pluripotent Stem Cells. *Curr. Biol.* **18**, 890–894 (2008).
168. Loh, Y. *et al.* Reprogramming of T cells from human peripheral blood. *Cell Stem Cell* **7**, 15–19 (2010).
169. Seki, T. *et al.* Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* **7**, 11–4 (2010).
170. Hanna, J. *et al.* Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency. *Cell* **133**, 250–264 (2008).
171. Staerk, J. *et al.* Reprogramming of peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell* **7**, 20–24 (2010).
172. Warren, C. R. *et al.* Induced Pluripotent Stem Cell Differentiation Enables Functional Validation of GWAS Variants in Metabolic Disease. *Cell Stem Cell* **20**,

- 547–557.e7 (2017).
173. Seki, T., Yuasa, S. & Fukuda, K. Generation of induced pluripotent stem cells from a small amount of human peripheral blood using a combination of activated T cells and Sendai virus. *Nat. Protoc.* **7**, 718–28 (2012).
 174. Crennell, S., Takimoto, T., Portner, A. & Taylor, G. Crystal Structure of the Multifunctional Paramyxovirus Hemagglutinin-Neuraminidase. *Nat. Struct. Biol.* **7**, 1068–1074 (2000).
 175. Ban, H. *et al.* Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc. Natl. Acad. Sci.* **108**, 14234–9 (2011).
 176. Choi, J. *et al.* A Comparison of Genetically Matched Cell Lines Reveals the equivalence of Human iPSCs and ESCs. *Nat. Biotechnol.* **33**, 1173–1181 (2015).
 177. Cheng, L. *et al.* Low Incidence of DNA Sequence Variation in Human Induced Pluripotent Stem Cells Generated by Nonintegrating Plasmid Expression. *Cell Stem Cell* **10**, 337–344 (2012).
 178. Bock, C. *et al.* Reference Maps of Human ES and iPS Cell Variation Enable High-Throughput Characterization of Pluripotent Cell Lines. *Cell* **144**, 439–452 (2011).
 179. Tsankov, A. M. *et al.* A qPCR ScoreCard Quantifies the Differentiation Potential of Human Pluripotent Stem Cells. *Nat. Biotechnol.* **33**, 1182–1192 (2015).
 180. Bidy, B. A. *et al.* Single-cell Mapping of Lineage and Identity in Direct Reprogramming. *Nature* **564**, 219–224 (2018).
 181. Lian, X. *et al.* Efficient differentiation of human pluripotent stem cells to endothelial progenitors via small-molecule activation of WNT signaling. *Stem Cell Reports* **3**,

- 804–816 (2014).
182. Rafii, S., Butler, J. M. & Ding, B. Sen. Angiocrine Functions of Organ-Specific Endothelial Cells. *Nature* **529**, 316–325 (2016).
 183. Ghesquiere, B., Wong, B. W., Kuchnio, A. & Carmeliet, P. Metabolism of Stromal and Immune Cells in Health and Disease. *Nature* **2**, (2014).
 184. Carmeliet, P. & Jain, R. K. Molecular Mechanisms and Clinical Applications of Angiogenesis. *Nature* **473**, 298–307 (2011).
 185. Kuosmanen, S. M., Kansanen, E., Sihvola, V. & Levonen, A.-L. MicroRNA Profiling Reveals Distinct Profiles for Tissue-Derived and Cultured Endothelial Cells. *Sci. Rep.* **7**, 10943 (2017).
 186. Vittet, D. *et al.* Embryonic stem cells differentiate in vitro to endothelial cells through successive maturation steps. *Blood* **88**, 3424–3431 (1996).
 187. Herbert, S. P. & Stainier, D. Y. R. Molecular control of endothelial cell behaviour during blood vessel morphogenesis. *Nat. Rev. Mol. Cell Biol.* **12**, 551–564 (2011).
 188. Wu, M.D, K. K. & Thiagarajan, M.D, P. Role of Endothelium in Thrombosis and Hemostasis. *Annu. Rev. Med.* **47**, 315–331 (1996).
 189. Traub, O. & Berk, B. C. Laminar shear stress: Mechanisms by which endothelial cells transduce an atheroprotective force. *Arterioscler. Thromb. Vasc. Biol.* **18**, 677–685 (1998).
 190. Gimbrone, M. A. & Garcia-Cardena, G. Endothelial Cell Dysfunction and the Pathobiology of Atherosclerosis. *Circ. Res.* **118**, 620–636 (2016).
 191. Paolinelli, R., Corada, M., Orsenigo, F. & Dejana, E. The Molecular Basis of the Blood Brain Barrier Differentiation and Maintenance. Is it Still a Mystery?

- Pharmacol. Res.* **63**, 165–171 (2011).
192. Tse, D., Ph, D. & Stan, R. V. Morphological Heterogeneity of Endothelium. *Semin. Thromb. Hemost.* **36**, 236–245 (2010).
193. Radomski, M. ., Palmer, R. M. . & Moncada, S. Endogenous nitric oxide inhibits human platelet adhesion to vascular endothelium. *Lancet* **330**, 1057–1058 (1987).
194. Ignarro, L. J. *et al.* Endothelium-derived relaxing factor produced and released from artery and vein is nitric oxide (endothelium-dependent relaxation/vascular smooth muscle/cyclic GMP). *Proc. Natl. Acad. Sci.* **84**, 9265–9269 (1987).
195. Rodgers, G. M. & Conn, M. T. Homocysteine, an atherogenic stimulus, reduces protein C activation by arterial and venous endothelial cells. *Blood* **75**, 895–901 (1990).
196. Marx, N., Bourcier, T., Sukhova, G. K., Libby, P. & Plutzky, J. PPAR γ Activation in Human Endothelial Cells Increases Plasminogen Activator Inhibitor Type-1 Expression: PPAR γ as a Potential Mediator in Vascular Disease. *Arterioscler. Thromb. Vasc. Biol.* **19**, 546–551 (1999).
197. Mack, J. J. *et al.* NOTCH1 is a mechanosensor in adult arteries. *Nat. Commun.* **8**, 1–18 (2017).
198. Li, Y. *et al.* Glutaredoxin 1 mediates the protective effect of steady laminar flow on endothelial cells against oxidative stress-induced apoptosis via inhibiting Bim. *Sci. Rep.* **7**, 1–11 (2017).
199. Chan, Y. *et al.* The cell-specific expression of endothelial nitric-oxide synthase: A role for DNA methylation. *J. Biol. Chem.* **279**, 35087–35100 (2004).
200. Ferreri, D. M., Minnear, F. L., Yin, T., Kowalczyk, A. P. & Vincent, P. A. N-

- cadherin levels in endothelial cells are regulated by monolayer maturity and p120 availability. *Cell Commun. Adhes.* **15**, 333–349 (2008).
201. Siekmann, A. F. & Lawson, N. D. Notch signalling limits angiogenic cell behaviour in developing zebrafish arteries. *Nature* **445**, 781–784 (2007).
 202. Stainier, D. Y., Weinstein, B. M., Detrich, H. W., Zon, L. I. & Fishman, M. C. Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. *Development* **121**, 3141–3150 (1995).
 203. Jakobsson, L. *et al.* Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nat. Cell Biol.* **12**, 943–953 (2010).
 204. Bussmann, J. *et al.* Arteries provide essential guidance cues for lymphatic endothelial cells in the zebrafish trunk. *Development* **137**, 2653–2657 (2010).
 205. Liao, W. *et al.* The zebrafish gene cloche acts upstream of a flk-1 homologue to regulate endothelial cell differentiation. *Development* **124**, 381–389 (1997).
 206. Hwangbo, C. *et al.* Modulation of endothelial bone morphogenetic protein receptor type 2 activity by vascular endothelial growth factor receptor 3 in pulmonary arterial hypertension. *Circulation* **135**, 2288–2298 (2017).
 207. Kalén, M. *et al.* Combination of reverse and chemical genetic screens reveals angiogenesis inhibitors and targets. *Chem. Biol.* **16**, 432–41 (2009).
 208. Mahmoud, M. M. *et al.* TWIST1 integrates endothelial responses to flow in vascular dysfunction and atherosclerosis. *Circ. Res.* **119**, 450–462 (2016).
 209. Deloukas, P. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet.* **45**, 25–33 (2013).
 210. Bhasin, M. *et al.* Bioinformatic identification and characterization of human

- endothelial cell-restricted genes. *BMC Genomics* **11**, 342 (2010).
211. Craige, S. M. *et al.* PGC-1 α dictates endothelial function through regulation of eNOS expression. *Sci. Rep.* **6**, 38210 (2016).
212. Pasceri, V., Willerson, J. T. & Yeh, E. T. H. Direct proinflammatory effect of C-reactive protein on human endothelial cells. *Circulation* **102**, 2165–2168 (2000).
213. Long, L. *et al.* Selective enhancement of endothelial BMPR-II with BMP9 reverses pulmonary arterial hypertension. *Nat. Med.* **21**, 777–785 (2015).
214. Wang, K.-C. *et al.* Flow-dependent YAP/TAZ activities regulate endothelial phenotypes and atherosclerosis. *Proc. Natl. Acad. Sci.* **113**, 11525–11530 (2016).
215. Gaitzsch, E. *et al.* Double-stranded DNA induces a prothrombotic phenotype in the vascular endothelium. 1–12 (2017). doi:10.1038/s41598-017-01148-x
216. Barzilai, S. *et al.* Leukocytes breach endothelial barriers by insertion of nuclear lobes and disassembly of endothelial actin filaments. *Cell Rep.* **18**, 685–699 (2017).
217. Su, X. *et al.* Post-transcriptional regulation of TNF-induced expression of ICAM-1 and IL-8 in human lung microvascular endothelial cells: an obligatory role for the p38 MAPK-MK2 pathway dissociated with HSP27. *Biochim. Biophys. Acta* **1783**, 1623–31 (2008).
218. Zhang, J. *et al.* Functional characterization of human pluripotent stem cell-derived arterial endothelial cells. *Proc. Natl. Acad. Sci.* **114**, E6072–E6078 (2017).
219. Schwartz, J. N., Kong, Y., Hackel, D. B. & Bartel, A. G. Comparison of Angiographic and Postmortem Findings in Patients with Coronary Artery Disease. *Am J Cardiol* **36**, 174–178 (1975).
220. Mann, J. & Davies, M. J. Mechanisms of progression in native coronary artery

- disease : role of healed plaque disruption Mechanisms of progression in native coronary artery disease : role of healed plaque disruption. *Heart* **82**, 265–268 (1999).
221. Austin, G. E., Ratliff, N. B., Hollman, J., Tabei, S. & Phillips, D. F. Intimal proliferation of smooth muscle cells as an explanation for recurrent coronary artery stenosis after percutaneous transluminal coronary angioplasty. *J. Am. Coll. Cardiol.* **6**, 369–375 (1985).
222. Qiu, C. *et al.* The critical role of SENP1-mediated GATA2 deSUMOylation in promoting endothelial activation in graft arteriosclerosis. *Nat. Commun.* **8**, 15426 (2017).
223. Baudin, B., Bruneel, A., Bosselut, N. & Vaubourdolle, M. A protocol for isolation and culture of human umbilical vein endothelial cells. *Nat. Protoc.* **2**, 481–485 (2007).
224. Montgomery, K. & Osborn, L. Activation of endothelial-leukocyte adhesion molecule 1 (ELAM-1) gene transcription. *Proc. Natl. Acad. Sci.* **88**, 6523–6527 (1991).
225. Blume, C. *et al.* AMP-activated protein kinase impairs endothelial actin cytoskeleton assembly by phosphorylating vasodilator-stimulated phosphoprotein. *J. Biol. Chem.* **282**, 4601–4612 (2007).
226. Bompais, H. *et al.* Human endothelial cells derived from circulating progenitors display specific functional properties compared with mature vessel wall endothelial cells. *Blood* **103**, 2577–2584 (2004).
227. Lieb, W. *et al.* Genome-wide association study for endothelial growth factors. *Circ. Cardiovasc. Genet.* **8**, 389–97 (2015).

228. Chimen, M. *et al.* Homeostatic regulation of T cell trafficking by a B cell–derived peptide is impaired in autoimmune and chronic inflammatory disease. *Nat. Med.* **21**, 467–475 (2015).
229. Calabró, P., Willerson, J. T. & Yeh, E. T. H. Inflammatory cytokines stimulated C-reactive protein production by human coronary artery smooth muscle cells. *Circulation* **108**, 1930–2 (2003).
230. Fong, A. *et al.* Fractalkine and CX3CR1 mediate a novel mechanism of leukocyte capture, firm adhesion, and activation under physiologic flow. *J. Exp. Med.* **188**, 1413–1419 (1998).
231. Lonza. *Amara HUVEC Nucleofector Kit.* (2009).
232. Fang, J. S. *et al.* Shear-induced Notch-Cx37-p27 axis arrests endothelial cell cycle to enable arterial specification. *Nat. Commun.* **8**, (2017).
233. Leisegang, M. S. *et al.* Long noncoding RNA MANTIS facilitates endothelial angiogenic function. *Circulation* **136**, 65–79 (2017).
234. Fernandez-Alonso, R. *et al.* P73 is required for endothelial cell differentiation, migration and the formation of vascular networks regulating VEGF and TGF β signaling. *Cell Death Differ.* **22**, 1287–1299 (2015).
235. Abrahimi, P. *et al.* Efficient gene disruption in cultured primary human endothelial cells by CRISPR/Cas9. *Circ. Res.* **117**, 121–128 (2015).
236. Martin-Ramirez, J., Hofman, M., van den Biggelaar, M., Hebbel, R. P. & Voorberg, J. Establishment of outgrowth endothelial cells from peripheral blood. *Nat. Protoc.* **7**, 1709–15 (2012).
237. Lin, Y., Weisdorf, D. ., Solovey, A. & Hebbel, R. P. Origins of circulating

- endothelial cells and endothelial outgrowth from blood. *J. Clin. Invest.* **105**, 71–77 (2000).
238. Peichev, M. *et al.* Expression of VEGFR-2 and AC133 by circulating human CD34⁺ cells identifies a population of functional endothelial precursors. *Blood* **95**, 952–958 (2000).
239. Wang, J. *et al.* Analysis of the storage and secretion of von Willebrand factor in blood outgrowth endothelial cells derived from patients with von Willebrand disease. **121**, 2762–2773 (2016).
240. Hill, J. & Zalos, G. Circulating endothelial progenitor cells, vascular function, and cardiovascular risk. *N. Engl. J. Med.* 593–600 (2003). at <http://www.nejm.org/doi/full/10.1056/nejmoa022287>
241. Schmidt-Lucke, C. *et al.* Reduced Number of Circulating Endothelial Progenitor Cells Predicts Future Cardiovascular Events: Proof of Concept for the Clinical Importance of Endogenous Vascular Repair. *Circulation* **111**, 2981–2987 (2005).
242. Smith, A. Mouse embryo stem cells; their identification, propagation and manipulation. *Semin. Cell Biol.* **3**, 385–399 (1992).
243. Doetschman, T., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. The in vitro development of blastocyst-derived embryonic stem cell lines: Formation of visceral yolk sac, blood islands and myocardium. *Development* **87**, 27–45 (1985).
244. Risau, W. *et al.* Vasculogenesis and angiogenesis in embryonic stem cell-derived embryoid bodies. *Development* **102**, 471–478 (1988).
245. Doetschman, T., Shull, M., Kier, A. & Coffin, J. Embryonic stem cell model systems for vascular morphogenesis and cardiac disorders. *Hypertension* **22**, 618–629

- (1993).
246. Young, P., Baumhueter, S. & Lasky, L. The sialomucin CD34 is expressed on hematopoietic cells and blood vessels during murine development. *Blood* **85**, 96–105 (1995).
247. Wang, R., Clark, R. & Bautch, V. Embryonic stem cell-derived cystic embryoid bodies form vascular channels: An in vitro model of blood vessel development. *Development* **114**, 303–316 (1992).
248. Wang, L. *et al.* Endothelial and hematopoietic cell fate of human embryonic stem cells originates from primitive endothelium with hemangioblastic properties. *Immunity* **21**, 31–41 (2004).
249. Bhatia, M. *et al.* Quantitative analysis reveals expansion of human hematopoietic repopulating cells after short-term ex vivo culture. *J. Exp. Med.* **186**, 619–624 (1997).
250. Levenberg, S., Golub, J. S., Amit, M., Itskovitz-Eldor, J. & Langer, R. Endothelial cells derived from human embryonic stem cells. *Proc. Natl. Acad. Sci.* **99**, 4391–6 (2002).
251. Ferreira, L. S. *et al.* Vascular progenitor cells isolated from human embryonic stem cells give rise to endothelial and smooth muscle like cells and form vascular networks in vivo. *Circulation* **101**, 286–94 (2007).
252. Kennedy, M., D’Souza, S. L., Lynch-Kattman, M., Schwantz, S. & Keller, G. Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. *Blood* **109**, 2679–87 (2007).
253. Wang, Z. Z. *et al.* Endothelial cells derived from human embryonic stem cells form

253. durable blood vessels in vivo. *Nat. Biotechnol.* **25**, 317–8 (2007).
254. James, D., Nam, H., Seandel, M. & Nolan, D. Expansion and maintenance of human embryonic stem cell-derived endothelial cells by TGF inhibition is Id1 dependent. *Nat. Biotechnol.* **28**, 161–167 (2010).
255. Sahara, M. *et al.* Manipulation of a VEGF-Notch signaling circuit drives formation of functional vascular endothelial progenitors from human pluripotent stem cells. *Cell Res.* **24**, 820–841 (2014).
256. Patsch, C. *et al.* Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat. Cell Biol.* **17**, 994–1003 (2015).
257. Nguyen, M. T. X. *et al.* Differentiation of human embryonic stem cells to endothelial progenitor cells on laminins in defined and xeno-free systems. *Stem Cell Reports* **7**, 802–816 (2016).
258. Zhang, S., Dutton, J. R., Su, L., Zhang, J. & Ye, L. The influence of a spatiotemporal 3D environment on endothelial cell differentiation of human induced pluripotent stem cells. *Biomaterials* **35**, 3786–3793 (2014).
259. Tsang, K. M. *et al.* Embryonic stem cell differentiation to functional arterial endothelial cells through sequential activation of ETV2 and NOTCH1 signaling by HIF1 α . *Stem Cell Reports* **9**, 796–806 (2017).
260. Wigle, J. T. & Oliver, G. Prox1 function is required for the development of the murine lymphatic system. *Cell* **98**, 769–778 (1999).
261. Zhang, L. *et al.* VEGFR-3 ligand-binding and kinase activity are required for lymphangiogenesis but not for angiogenesis. *Cell Res.* **20**, 1319–1331 (2010).
262. Karkkainen, M. J. *et al.* Vascular endothelial growth factor C is required for

- sprouting of the first lymphatic vessels from embryonic veins. *Nat. Immunol.* **5**, 74–80 (2004).
263. Gauvrit, S. *et al.* HHEX is a Transcriptional Regulator of the VEGFC/FLT4/PROX1 Signaling Axis During Vascular Development. *Nat. Commun.* **9**, (2018).
264. Krebs, L. T. *et al.* Notch Signaling is Essential for Vascular Morphogenesis in Mice. 1343–1352 (2000).
265. Rossant, J. & Howard, L. Signaling Pathways in Vascular Development. *Annu. Rev. Cell Dev. Biol.* **18**, 541–73 (2002).
266. Jain, R. K. Molecular regulation of vessel maturation. *Nat. Med.* **9**, 685–93 (2003).
267. Atkins, G. B., Jain, M. K. & Hamik, A. Endothelial Differentiation - Molecular Mechanisms of Specification and Heterogeneity. *Arterioscler. Thromb. Vasc. Biol.* **31**, 1476–1485 (2011).
268. Gerety, S. S. & Anderson, D. J. Cardiovascular EphrinB2 Function is Essential for Embryonic Angiogenesis. **1410**, 1397–1410 (2002).
269. Adams, R. H. *et al.* Roles of EphrinB Ligands and EphB Receptors in Cardiovascular Development : Demarcation of Arterial/Venous Domains, Vascular Morphogenesis, and Sprouting Angiogenesis. *Genes Dev.* **13**, 295–306 (1999).
270. Wang, H. U., Chen, Z. & Anderson, D. J. Molecular Distinction and Angiogenic Interaction between Embryonic Arteries and Veins Revealed by ephrin-B2 and Its Receptor Eph-B4. **93**, 741–753 (1998).
271. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. & Keller, G. Haemangioblast Commitment is Initiated in the Primitive Streak of the Mouse Embryo. *Nature* **432**, 625–630 (2004).

272. Yang, L. *et al.* Human Cardiovascular Progenitor Cells Develop from a KDR+ Embryonic-Stem-Cell-Derived Population. *Nature* **453**, 524–528 (2008).
273. Carlos, T. M. & Harlan, J. M. Leukocyte-endothelial adhesion molecules. *Blood* **84**, 2068–2101 (1994).
274. de Almeida, P. . . *et al.* Transplanted Terminally Differentiated Induced Pluripotent Stem Cells are Accepted by Immune Mechanisms Similar to Self-Tolerance. *Nat. Commun.* **5**, 1–12 (2014).
275. Oswald, J. *et al.* Mesenchymal stem cells can be differentiated into endothelial cells in vitro. *Stem Cells* **22**, 377–384 (2004).
276. Choi, K.-D. *et al.* Hematopoietic and endothelial differentiation of human induced pluripotent stem cells. *Stem Cells* **27**, 559–67 (2009).
277. Wu, Y. T. *et al.* Defining Minimum Essential Factors to Derive Highly Pure Human Endothelial Cells from iPS/ES Cells in an Animal Substance-Free System. *Sci. Rep.* **5**, 1–9 (2015).
278. Kim, D. N., Scott, R. F., Schmee, J. & Thomas, W. A. Endothelial Cell Denudation, Labelling Indices and Monocyte Attachment in Advanced Swine Coronary Artery Lesions. *Atherosclerosis* **73**, 247–257 (1988).
279. Civelek, M., Manduchi, E., Riley, R. J., Stoeckert, C. J. & Davies, P. F. Coronary Artery Endothelial Transcriptome In Vivo. *Circ. Cardiovasc. Genet.* **4**, 243–252 (2011).
280. Tzima, E. *et al.* A mechanosensory complex that mediates the endothelial cell response to fluid shear stress. *Nature* **437**, 426–31 (2005).
281. Hajra, L. *et al.* The NF-kappa B signal transduction pathway in aortic endothelial

- cells is primed for activation in regions predisposed to atherosclerotic lesion formation. *Proc. Natl. Acad. Sci.* **97**, 9052–9057 (2000).
282. Alp, N. J. & Channon, K. M. Regulation of Endothelial Nitric Oxide Synthase by Tetrahydrobiopterin in Vascular Disease. *Arterioscler. Thromb. Vasc. Biol.* **24**, 413–420 (2004).
283. Collins, T., Read, M. A., Neish, A. S. & Wiiiitley, M. Z. Transcriptional molecules : regulation of endothelial cell adhesion enhancers and cytokine-inducible. *FASEB J.* **9**, 899–909 (1995).
284. Plutzky, J. The vascular biology of atherosclerosis. *Am. J. Med.* **115**, 55–61 (2003).
285. Gertler, M. M., York, N., Garn, S. M., Ph, D. & White, P. D. Young Candidates for Coronary Heart Disease. *Jama* **147**, 621–625 (1951).
286. Boomsma, D., Busjahn, A. & Peltonen, L. Classical Twin Studies and Beyond. *Nat. Rev. Genet.* **3**, 872–882 (2002).
287. The Wellcome Trust Case Control Consortium. Genome-wide Association Study of 14,000 Cases of Seven Common Diseases and 3,000 Shared Controls. *Nature* **447**, 661–678 (2007).
288. Helgadottir, A. *et al.* A Common Variant on Chromosome 9p21 Affects the Risk of Myocardial Infarction. *Science (80-.).* **317**, 1491–1494 (2007).
289. McPherson, R. *et al.* A Common Allele on Chromosome 9 Associated with Coronary Heart Disease. *Science (80-.).* **316**, 1488–1491 (2007).
290. Samani, N. J. *et al.* Genomewide Association Analysis of Coronary Artery Disease. *N. Engl. J. Med.* **357**, 443–453 (2007).
291. Lo Sardo, V. *et al.* Unveiling the Role of the Most Impactful Cardiovascular Risk

- Locus through Haplotype Editing. *Cell* **175**, 1796–1810.e20 (2018).
292. Khera, A. V. & Kathiresan, S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat. Rev. Genet.* **18**, 331–344 (2017).
293. Morton, N. E., Yee, S., Elston, R. C. & Lew, R. Discontinuity and Quasi-Continuity: Alternative Hypotheses of Multifactorial Inheritance. *Clin. Genet.* **1**, 81–94 (1970).
294. Thériault, S. *et al.* Polygenic Contribution in Individuals With Early-Onset Coronary Artery Disease. *Circ. Genomic Precis. Med.* **11**, e001849 (2018).
295. Khera, A. V. *et al.* Genome-wide Polygenic Scores for Common Diseases Identify Individuals with Risk Equivalent to Monogenic Mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
296. Hajek, C. *et al.* Coronary Heart Disease Genetic Risk Score Predicts Cardiovascular Disease Risk in Men, Not Women. *Circ. Genomic Precis. Med.* **11**, e002324 (2018).
297. Sundström, J. *et al.* Multilocus Genetic Risk Scores for Coronary Heart Disease Prediction. *Arterioscler. Thromb. Vasc. Biol.* **33**, 2267–2272 (2013).
298. Thanassoulis, G. *et al.* A Genetic Risk Score Is Associated With Incident Cardiovascular Disease and Coronary Artery Calcium. *Circ. Cardiovasc. Genet.* **5**, 113–121 (2012).
299. Tikkanen, E., Havulinna, A. S., Palotie, A., Salomaa, V. & Ripatti, S. Genetic Risk Prediction and a 2-Stage Risk Screening Strategy for Coronary Heart Disease. *Arterioscler. Thromb. Vasc. Biol.* **33**, 2261–2266 (2013).
300. Montori, V. M. *et al.* Incorporating a Genetic Risk Score Into Coronary Heart Disease Risk Estimates. *Circulation* **133**, 1181–1188 (2016).
301. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The Long-Range Interaction

- Landscape of Gene Promoters. *Nature* **489**, 109–113 (2012).
302. Gupta, R. M. *et al.* A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. *Cell* **170**, 522–533.e15 (2017).
303. Turner, A. W., Wong, D., Dreisbach, C. N. & Miller, C. L. GWAS Reveal Targets in Vessel Wall Pathways to Treat Coronary Artery Disease. *Front. Cardiovasc. Med.* **5**, 1–10 (2018).
304. Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* **12**, 745–755 (2011).
305. Barnett, I. J., Lee, S. & Lin, X. Detecting Rare Variant Effects Using Extreme Phenotype Sampling in Sequencing Association Studies. *Genet. Epidemiol.* **37**, 142–151 (2013).
306. Peloso, G. M. *et al.* Phenotypic Extremes in Rare Variant Study Designs. *Eur. J. Hum. Genet.* **24**, 924–930 (2016).
307. Frischmeyer, P. a & Dietz, H. C. Nonsense-mediated mRNA decay in health and disease. *Hum. Mol. Genet.* **8**, 1893–1900 (1999).
308. Bhuvanagiri, M., Schlitter, A. M., Hentze, M. W. & Kulozik, A. E. NMD: RNA biology meets human genetic medicine. *Biochem. J.* **430**, 365–377 (2010).
309. Hall, G. & Thein, S. Nonsense codon mutations in the terminal exon of the β -globin gene are not associated with a reduction in β -mRNA accumulation: a mechanism for the phenotype of dominant β -thalassemia. *Blood* **83**, 2031–2037 (1994).
310. Lou, C.-H. *et al.* Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate. *Stem cell reports* **6**, 844–857 (2016).
311. Mendell, J. T., Sharifi, N. a, Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C.

- Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat. Genet.* **36**, 1073–1078 (2004).
312. Peccarelli, M. & Kebaara, B. W. Regulation of natural mRNAs by the nonsense-mediated mRNA decay pathway. *Eukaryot. Cell* **13**, 1126–1135 (2014).
313. Maquat, L. E. Nonsense-mediated mRNA decay in mammals. *J. Cell Sci.* **118**, 1773–1776 (2005).
314. Popp, M. W. & Maquat, L. E. Leveraging rules of nonsense-mediated mRNA decay for genome engineering and personalized medicine. *Cell* **165**, 1319–1332 (2016).
315. Hug, N., Longman, D. & Cáceres, J. F. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* **44**, 1483–1495 (2015).
316. Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W. & Kulozik, A. E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747 (2008).
317. Hug, N. & Cáceres, J. F. The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell Rep.* **8**, 1845–1856 (2014).
318. Melero, R. *et al.* The RNA helicase DHX34 functions as a scaffold for SMG1-mediated UPF1 phosphorylation. *Nat. Commun.* **7**, 1–12 (2016).
319. Medghalchi, S. M. *et al.* Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* **10**, 99–105 (2001).
320. Shum, E. Y. *et al.* The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. *Cell* **165**, 382–395 (2016).

321. Tarpey, P. *et al.* Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat. Genet.* **39**, 1127–1133 (2007).
322. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333–338 (2011).
323. Lou, C. H. *et al.* Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. *Cell Rep.* **6**, 748–764 (2014).
324. Hug, N. & Cáceres, J. F. The RNA Helicase DHX34 Activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell Rep.* **8**, 1845–1856 (2014).
325. Mourtada-Maarabouni, M. & Williams, G. T. Growth arrest on inhibition of nonsense-mediated decay is mediated by noncoding RNA GAS5. *Biomed Res. Int.* **2013**, (2013).
326. Huang, L. RNA Homeostasis Governed by Cell Type-Specific and Branched Feedback Loops Acting on NMD. *Mol. Cell* **43**, 950–961 (2011).
327. Wang, J., Gudikote, J. P., Olivas, O. R. & Wilkinson, M. F. Boundary-Independent Polar Nonsense-Mediated Decay. *EMBO Rep.* **3**, 274–279 (2002).
328. Adcock, I. M., Brown, C. R., Kwon, O. & Barnes, P. J. Oxidative stress induces NFκB dna binding and inducible nos mRNA in human epithelial cells. *Biochemical and Biophysical Research Communications* **199**, 1518–1524 (1994).
329. Jung, D. H. *et al.* Involvement of ATF3 in the negative regulation of iNOS expression and NO production in activated macrophages. *Immunol. Res.* **62**, 35–45

- (2015).
330. Arnaoutova, I., George, J., Kleinman, H. K. & Benton, G. The Endothelial Cell Tube Formation Assay on Basement Membrane Turns 20: State of the Science and the Art. *Angiogenesis* **12**, 8 (2009).
 331. Cybulsky, M. I. *et al.* A Major Role for VCAM-1, but Not ICAM-1, in Early Atherosclerosis. *J. Clin. Invest.* **107**, 1255–62 (2001).
 332. Hu, S. *et al.* Vascular Semaphorin 7A Upregulation by Disturbed Flow Promotes Atherosclerosis Through Endothelial β 1 Integrin. *Arterioscler. Thromb. Vasc. Biol.* **38**, 335–343 (2018).
 333. Gehring, N. H. *et al.* Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol. Cell* **20**, 65–75 (2005).
 334. Chan, W.-K. *et al.* An alternative branch of the nonsense-mediated decay pathway. *EMBO J.* **26**, 1820–1830 (2007).
 335. Guo, Q. *et al.* ‘Cold shock’ Increases the Frequency of Homology Directed Repair Gene Editing in Induced Pluripotent Stem Cells. *Sci. Rep.* **8**, 1–11 (2018).
 336. Rivera-Torres, N., Banas, K., Bialk, P., Bloh, K. M. & Kmiec, E. B. Insertional Mutagenesis by CRISPR/Cas9 Ribonucleoprotein Gene Editing in Cells Targeted for Point Mutation Repair Directed by Short Single-Stranded DNA Oligonucleotides. *PLoS One* **12**, e0169350 (2017).
 337. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing Homology-Directed Genome Editing by Catalytically Active and Inactive CRISPR-Cas9 Using Asymmetric Donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).

338. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 1–18 (2016).
339. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Japan Acad. Ser. B* **85**, 348–362 (2009).
340. Kudva, Y. *et al.* Transgene-free disease-specific induced pluripotent stem cells from patients with type 1 and type 2 diabetes. *Stem Cells Transl. Med.* **1**, 451–461 (2012).
341. Sriram, G., Tan, J. Y., Islam, I., Rufaihah, A. J. & Cao, T. Efficient differentiation of human embryonic stem cells to arterial and venous endothelial cells under feeder- and serum-free conditions. *Stem Cell Res. Ther.* **6**, 1–17 (2015).