THE UNIVERSITY
of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Mapping the Transcriptional Dynamics of the Endothelium During Human Embryonic Stem Cell Differentiation and Cardiac Development Using Single Cell RNA Sequencing

Ian Richard McCracken

Submitted for the degree of Doctor of Philosophy

University of Edinburgh

2021

Declaration

I declare that the work contained within this thesis is my own except where clearly stated otherwise within the author contribution statement of each results chapter. This work has not been submitted for any other degree or professional qualification and published material contained within this thesis is my own work except where clearly indicated otherwise within the text.

Ian Richard McCracken

Acknowledgements

Firstly, to my supervisors Prof Andy Baker and Prof Chris Ponting. I think it is safe to say the last four years have been incredibly rewarding, and your unwavering encouragement, guidance, and support has been instrumental to the success of the project but also to my development as a scientist and a person. Thanks to both of you, and I look forward to our paths crossing again in the future. In addition, I would particularly like to thank Prof Baker for taking a chance on me at the beginning by welcoming me into his group and opening my eyes to the world of cardiovascular research.

I would also like to thank all past and present members of the Baker lab for offering their technical expertise, advice, and friendship during my time in the Edinburgh. In particular to Amira Mahmoud, Fernando De La Cuesta, Ana-Mishel Spiroski, and Appie Beqqali for providing encouragement and support throughout all the highs and lows of my PhD journey.

I would also like to express my gratitude to Prof Jo Mountford and her group at SNBTS for all their stem cell related expertise. Especially to Dr Scott Cowan and Dr Alison Condie for their time and patience in recent years. Without the assistance from Jo and her group this project would simply not have been possible, so I am incredibly grateful to them.

I am also appreciative for the support I have received from Neil Henderson and members of his research group with all things 10X. In particular from Dr Ross

Dobie and Dr Beth Henderson for all the long hours they devoted to processing samples for my project. Thanks also to Andrea Corsinotti and other members of the CRM and QMRI cell sorting facilities as well as the CRM tissue culture staff and CVS lab management for all their help.

I would also like to acknowledge the Medical Research Council and organisers of the Precision Medicine DTP for giving me the incredible opportunity of doing this PhD and allowing me to further my scientific career.

Finally, I would like to express my thanks to all my friends and family who have supported me throughout my PhD, despite having to endure countless explanations about cells and why they really do require feeding every day, even at weekends. Although there are far too many names to mention, I would like to give special thanks to my parents, two sisters, and Ciara McDonnell for their unfaltering encouragement and patience over these years, and to whom I am eternally grateful.

Abstract

The endothelium comprises the luminal surface of all blood and lymphatic vessels and is known to first emerge from mesodermal precursors in a process known as vasculogenesis. Following formation of a primitive vascular plexus, the endothelium rapidly expands by angiogenesis and undergoes specification into arterial, venous, capillary, lymphatic, and haemogenic subtypes. Additionally, in response to intrinsic and tissue specific cues, the endothelium undergoes further specialisation to meet the individual requirements of the surrounding tissue. However, despite improving understanding of the sequence of events occurring during endothelial cell (EC) development, the transcriptional control underlying these processes remains largely uncharacterised.

The differentiation of human embryonic stem cells (hESC) to endothelium offers an appropriate model to study the transcriptional control of endothelial differentiation *in vitro*. In this study I applied high throughput single cell RNA sequencing (scRNA-seq) to an efficient 8-day hESC endothelial differentiation protocol, collecting cells at days 0, 4, 6, and 8. Flow cytometric analysis of day 8 cells revealed a population in which 66% of cells co-expressed endothelial markers CD31 and CD144. Analysis of scRNA-seq data from day 0 and 4 revealed homogeneous populations defined by expression of pluripotent and lateral mesoderm markers, respectively. In contrast, scRNA-seq analysis of cells collected at days 6 and 8 show the emergence of distinct endothelial and mesenchymal populations. Repeating scRNA-seq analysis of hESC-

endothelial differentiation with a second hESC line as well as using an alternative differentiation protocol revealed highly comparable transcriptional signatures of endothelial differentiation. However, comparison of data from hESC derived endothelial cells (hESC-EC) to equivalent scRNA-seq data from cultured and freshly isolated human organ-specific endothelial cells, demonstrated a clear distinction in their transcriptional state, thus highlighting the likely unspecified nature of hESC-EC.

Following on from these findings, I then sought to determine if mapping the dynamic transcriptional landscape of the endothelium during human cardiac development could be used to identify novel regulators of endothelial maturation and specification. CD31+/CD45- endothelial cells isolated by FACS from heart tissue obtained from 13- and 14-week fetuses were processed for scRNA-seq analysis. Subsequent dimensionality reduction and clustering of data revealed distinct endocardial, capillary, venous, arterial, and lymphatic populations each with a unique transcriptional signature. Application of trajectory inference methods predicted a minor endocardial contribution to the coronary vasculature via a venous population, prior to the subsequent arterial specification of capillary EC. Mapping differentially expressed genes over pseudotime revealed the expression of *MECOM* to coincide with the onset of arterial specification. In conjunction, gene regulatory network analysis also placed *MECOM* amongst known regulators of arterial EC specification such as *HEY1* and *SOX17*. Subsequent knockdown of *MECOM* in arterial-like hESC-EC suggested a function in repressing venous identity in arterial EC.

Together these studies provide a comprehensive map of the transcriptional landscape accompanying both hESC-EC differentiation and during human cardiac EC development. Mapping these dynamic processes offers new insight into the mechanisms underlying endothelial development, creating future opportunities for vascular regeneration for ischaemic disease.

Lastly, in response to the COVID-19 pandemic I determined the susceptibility of endothelial cells to in-vitro infection with human coronaviruses, including SARS-CoV-2. Despite demonstrating productive infection of endothelial cells by HCoV-229E, a lack of replicative infection was observed by SARS-CoV-2. This was most likely due to an absence of the expression of the SARS-CoV-2 receptor, angiotensin converting enzyme 2 (*ACE2*).

Lay Summary

The vascular system is composed of a series of interconnected blood and lymphatic vessels responsible for carrying oxygen and nutrients to every organ and tissue within the body as well as removing waste products. Therefore, the correct formation and maintenance of the vasculature is essential throughout embryonic development and adulthood. Lining the innermost layer of all blood and lymphatic vessels, endothelial cells play a crucial function in health and disease, with endothelial dysfunction being a major contributor to vascular pathology. Central to endothelial function is the capability for endothelial cells to become specialised to a given vascular bed and tissue, thus allowing them to support the specific requirements of the surrounding tissue. Recent studies using animal models have provided substantial insight into how endothelial cells are formed and subsequently become specialised during embryonic development. However, the applicability of these findings to human biology and the genes regulating these processes remain incompletely understood.

The recent emergence of single-cell RNA sequencing (scRNA-seq) technology has made it possible to characterise a mixed cell population by classifying which genes are turned on and off within individual cells. Therefore, in the work contained within this thesis we used scRNA-seq to map the changes in gene activity during human endothelial cell development.

The first part of this work used a cell culture model of endothelial cell development in which human stem cells, which have the capability to

specialise into virtually any cell type in the body, are treated to become endothelial cells in a process called endothelial differentiation. Collecting cells for scRNA-seq at different stages throughout this process allowed us to determine which genes are turned on and off as the stem cells acquire an endothelial identity. This identified the presence of a homogeneous intermediary mesoderm population prior to the emergence of separate endothelial and mesenchymal (non-endothelial) cell populations. Comparisons with a different stem cell line and an alternative method of differentiation identified several genes without a previously reported role in endothelial development to be switching on as cells acquire an endothelial identity, suggesting a function in endothelial differentiation.

In a subsequent study we isolated endothelial cells from human fetal cardiac tissue prior to performing scRNA-seq. This revealed the presence of several distinct subtypes of endothelial cells within the developing human heart. Comparison of the selection genes switched on and off allowed the identities and potential function of these endothelial subtypes to be inferred. Data analysis also identified the transition of capillary endothelial cells to arterial endothelial cells. Furthermore, we identified increasing activity of the gene *MECOM* as cells acquired an arterial identity, suggesting a regulatory function in arterial specification. Confirming this prediction, reducing the activity of the *MECOM* gene in human stem cell derived arterial-like endothelial cells resulted in an increase in activity of genes normally only turned on in endothelial cells found in veins.

Together the data presented within this thesis provides new information into the genes likely to be involved in the development and specialisation of human endothelial cells. This information may prove useful to informing future therapeutic strategies aimed to regenerate the vasculature of diseased organs with compromised blood flow.

First Author publications

McCracken IR, Dobie R, Henderson NC, Mountford JC, Riley PR, Ponting CP, Smart N, Brittan M, Baker AH. Mapping the developing human cardiac endothelium at single cell resolution identifies MECOM as a regulator of arteriovenous identity. **(Under Review)**.

McCracken IR, Saginc G, He L, Huseynov A, Daniels A, Fletcher S, Peghaire C, Kalna V, Andaloussi-Mäe M, Muhl L, Craig NM, Griffiths SJ, Haas JG, Tait-Burkard C, Lendahl U, Birdsey GM, Betsholtz C, Noseda M, Baker AH, Randi AM. Lack of Evidence of Angiotensin-Converting Enzyme 2 Expression and Replicative Infection by SARS-CoV-2 in Human Endothelial Cells. Circulation. 23;143(8):865-868 (2021).

McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BEP, Mountford JC, Caudrillier A, Henderson NC, Ponting CP, Baker AH. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. Eur Heart J. 1;41(9):1024-1036. (2020).

McCracken IR, Taylor RS, Henderson NC, Sluimer JC, Baker AH. Unravelling atherosclerotic heterogeneity by single cell RNA sequencing. Curr Opin Lipidol. 29(6):488-489 (2018).

Additional publications

Spiroski AM, **McCracken IR**, Thomson A, Magalhaes-Pinto M, Newton KJ, Miller E, Markose D, Bénézech C, Hadoke P, Brittan M, Mountford JC, Gray GA and Baker AH. Targeting cellular mechanisms of myocardial preservation in infarcted hearts with human embryonic stem cell-derived endothelial cell product injection. **(Under Review)**.

van Kuijk K, **McCracken IR**, Tillie RJHA, Wichers Schreur R, Taylor RS, Kuppe C, Jin H, Asselberghs SEJ, Dobie R, Ramachamdran P, Gijbels MJ, Temmerman L, Luyten J, Noels H, Goossens P, Wilson-Kanamori JR, Schurgers LJ, Kirkwoord PM, Mees BME, Biessen EAL, Henderson NC, Kramann R, Baker AH, Sluimer JC. Fibroblast heterogeneity revealed by single-cell sequencing in healthy and atherosclerotic arteries.**(Under Review)**.

Robinson EL, Baker AH, Brittan M, **McCracken IR**, Condorelli G, Emanueli C, Srivastava KP, Gaetano C, Thum T, Vanhaverbeke M, Angione C, Heymans S, Devaux Y, Pedrazzini T, Martelli F; EU-CardioRNA COST Action CA17129. Dissecting the transcriptome in cardiovascular disease. Cardiovasc Res. cvab117 (2021)

Spiroski AM, Sanders R, Meloni M, **McCracken IR**, Thomson A, Brittan M, Gray GA, Baker AH. The Influence of the LINC00961/SPAAR Locus Loss on Murine Development, Myocardial Dynamics, and Cardiac Response to Myocardial Infarction. Int J Mol Sci. Jan 19;22(2):969 (2021).

Presentations

“Mapping the Developing Human Cardiac Endothelium at Single Cell Resolution”. Oral presentation. BHF Regenerative Medicine Centres Symposium. Virtual conference. August 2021.

“Using Stem Cells to Regenerate the Heart”. Oral presentation. Cafe Scientifique. Glasgow, UK. December 2019.

“Dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA sequencing”. Oral presentations. BHF Regenerative Medicine Centres Symposium. Oxford, UK. June 2019.

“Dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA sequencing”. Oral presentations. University of Edinburgh Single Cell Transcriptomics Symposium. Edinburgh, UK. June 2019.

“Dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA sequencing”. Poster presentation. Basic Cardiovascular Sciences Conference (BCVS). Boston, USA. August 2019

“Assessing In-vivo Functionality and Cellular Composition of a Heterogeneous Human Embryonic Stem Cell Derived Endothelial Cell Population”. Poster presentation. Noncoding RNAs in Embryonic Development and Cell Differentiation conference. Tel Aviv, Israel. April 2018.

Abbreviations

ACE2	Angiotensin-converting enzyme 2
ADAM	A disintegrin and metalloproteinases
AEC	Arterial endothelial cell
AIC	Akaike Information Criterion
ALK1	Activin A receptor like type 1
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
AVM	Arteriovenous malformation
BBB	Blood brain barrier
BMP	Bone morphogenetic protein
CCA	Canonical correlation analysis
CHIP-seq	Chromatin immunoprecipitation followed by high-throughput DNA sequencing
CLI	Critical limb ischaemia
COVID-19	Coronavirus disease 2019
CPM	Counts per million
CV	Coefficient of variation
CVD	Cardiovascular disease
DGE	Differential gene expression
E	Embryonic day
EB	Embryoid body
EC	Endothelial cell
EHT	Endothelial-to-haematopoietic transition
EMP	Erythro-myeloid progenitors
EMT	Epithelial-to-mesenchymal transition

EndMT	Endothelial-to-mesenchymal transition
eNOS	Endothelial nitric oxide synthase
ERG	ETS-related gene
ESC	Embryonic stem cell
ETS	E26 transformation specific
ETV2	ETS Variant Transcription Factor 2
EYFP	Enhanced yellow fluorescent protein
FGF	Fibroblast growth factor
FLI1	Friend leukemia integration 1 transcription factor
FLK1	Fetal liver kinase 1
FOX	Forkhead box
GEMs	Gel beads in emulsion
GFP	Green fluorescent protein
GMP	Good manufacturing practice
GO	Gene ontology
GRN	Gene regulatory network
HCAEC	Human coronary artery endothelial cell
HCoV-229E	Human coronavirus 229E
HDLEC	Human dermal lymphatic endothelial cell
hESC	Human embryonic stem cell
hESC-CM	Human embryonic stem cell derived cardiomyocyte
hESC-EC	Human embryonic stem cell derived endothelial cell
hESC-ECP	Human embryonic stem cell derived endothelial cell product
HPAEC	Human pulmonary artery endothelial cell

hPSC	Human pluripotent stem cell
hPSC-EC	Human pluripotent stem cell derived endothelial cell
HSC	Hematopoietic stem cell
HUVEC	Human umbilical vein endothelial cell
IHH	Indian hedgehog
ISH	In situ hybridisation
KDR	Kinase insert domain receptor
LEC	Lymphatic endothelial cell
LPM	Lateral plate mesoderm
MACS	Magnetic activated cell sorting
MADs	Median absolute deviations
MAPK	Mitogen-activated protein kinase
MB	Mesenchymoangioblast
mESC	Mouse embryonic stem cell
MI	Myocardial infarction
NB-GAM	Negative binomial generalised additive model
NICD	Notch intracellular domain
NR2F2	Nuclear receptor subfamily 2 group F member 2
PCA	Principal component analysis
PDGFRA	Platelet derived growth factor receptor alpha
PI3K	Phosphatidylinositol-3-kinase
PXM	Paraxial mesoderm
RNA	Ribonucleic acid
ROS	Reactive oxygen species

RQ	Relative quantity
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SCL	Stem cell leukaemia
scRNA-seq	Single-cell RNA sequencing
SHH	Sonic hedgehog
siRNA	Small interfering RNA
SMC	Smooth muscle cell
SOM	Self-organising heatmap
SV	Sinus venosus
TF	Transcription factor
TMM	Trimmed Mean of M-value
tSNE	t-distributed stochastic neighbour embedding
TSS	Transcriptional start site
UBC	Ubiquitin C
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier
VEC	Venous endothelial cell
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
VMR	Variance to mean ratio
VWF	von Willebrand factor
Wnt	Wingless-type MMTV integration site family member
µg	Microgram
µl	Microlitre

Table of Contents

Chapter 1 - Introduction *25*

1.1. Overview **26**

1.2 Development of the endothelium **29**

1.2.1. Overview of endothelial development 29

1.2.2. Vasculogenesis 30

1.2.3. Evidence for a common clonal progenitor for endothelial and haematopoietic lineage 33

1.2.4. Development of endothelium from paraxial mesoderm 36

1.2.5. A potential haematopoietic contribution to the vascular endothelium . 37

1.3. Transcriptional control of vascular development **38**

1.3.1. ETS Transcription Factors 38

1.3.2. Forkhead box (FOX) Transcription Factors 49

1.3.3. SOX Transcription Factors 52

1.3.4. SCL (stem cell leukemia) 54

1.4. Arteriovenous Endothelial Specification **55**

1.4.1. Overview of arteriovenous specification 55

1.4.2. Molecular regulators of arteriovenous specification 56

1.4.3. Influence of shear stress in arteriovenous specification 62

1.4.4. Role of cell cycle regulation of arteriovenous specification 63

1.4.5. Arteriovenous specification in the human endothelium 64

1.5. Lymphatic endothelial cell development **66**

1.6 Human pluripotent stem cell models of endothelial development. **68**

1.6.1. Human pluripotent stem cells. 68

1.6.2. Methods of hPSC-derived endothelial cell differentiation. 70

1.7. Endothelial heterogeneity and specification **75**

1.7.1. Functional importance of organ specific endothelial heterogeneity 75

1.7.2. Regulation of further endothelial specification 76

1.7.3. Characterising EC heterogeneity using single cell RNA sequencing..... 78

1.8 Hypotheses and Aims 83

Chapter 2 - Transcriptional dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA-Seq. 84

2.1 Chapter Introduction 85

2.2 Title Page 86

2.3 Abstract..... 87

2.4 Translational Perspective 89

2.5 Introduction 90

2.6 Methods..... 93

2.7 Results 96

2.8 Discussion 106

2.9 Figure Legends 110

2.10 Supplementary Methods 121

2.11 Supplementary Legends 132

2.12 Chapter Conclusion..... 153

Chapter 3 - Mapping the developing human cardiac endothelium at single cell resolution identifies MECOM as a regulator of arteriovenous identity..... 154

3.1 Chapter Introduction 155

3.2 Title Page 156

3.3 Abstract..... 157

3.4 Translational Perspective 159

3.5 Introduction 160

3.6 Methods..... 163

3.7 Results 166

3.8 Discussion 176

3.9 Figure Legends 181

3.10 Supplementary Methods 190

3.11 Supplementary Figure Legends 203

3.12 Chapter Conclusion 210

Chapter 4 - Lack of Evidence of Angiotensin-Converting Enzyme 2 Expression and Replicative Infection by SARS-CoV-2 in Human Endothelial Cells 213

4.1 Chapter Introduction 214

4.2 Title Page 215

4.3 Main text 216

4.4 Figure Legend 220

4.5 Methods 224

4.6 Chapter Conclusion 229

Chapter 5 - General Discussion 230

5.1 Summary 231

5.2 scRNA-seq analysis of hESC-EC differentiation 233

5.3 Mapping the transcriptome of the human fetal cardiac endothelium using scRNA-seq 239

5.4 Assessment of the susceptibility of EC to direct SARS-CoV-2 infection 243

5.5 Conclusions 245

References 246

List of Figures

Figure 1.1 Development of the endothelium by vasculogenesis and angiogenesis.. 30
Figure 1.2 FLK1+ mesoderm represents a multipotent population. 33
Figure 1.3 Domain structure of ETS transcription factors. 39
Figure 1.4 Subpopulations of FLK1+ mesoderm. 42
Figure 1.5 Retinal vasculature of control and FOXO1 KO mice 51
Figure 1.6 Molecular regulation of arteriovenous specification. 56
Figure 1.7 scRNA-seq atlas of murine endothelial cell heterogeneity. 81
Figure 2.1 scRNA-seq analysis of hESC-ECP differentiation across individual time points 114
Figure 2.2 Mapping hESC-ECP differentiation using scRNA-seq 115
Figure 2.3 Pseudotime trajectories of hESC-ECP differentiation 116
Figure 2.4 Longitudinal scRNA-seq analysis of an alternative hESC-EC differentiation 117
Figure 2.5 Pseudotime analysis of hESC-EC differentiation 118
Figure 2.6 Transcriptomic comparison of hESC-EC to mature endothelial cells .. 119
Figure 2.7 Summary of findings from scRNA-seq analysis of hESC-EC differentiation 120
Figure 3.1 Mapping the human fetal heart endothelium using scRNA-seq. 184
Figure 3.2 Gene regulatory network analysis of human fetal heart endothelium. .. 185
Figure 3.3 Trajectory inference analysis of developing cardiac endothelium. 186
Figure 3.4 Comparison of the transcriptional profiles of human and mouse fetal cardiac endothelial cell populations. 187
Figure 3.5 Knockdown of MECOM in hESC-EC. 188
Figure 3.6 Summary of key findings from scRNA-seq characterisation of human fetal cardiac endothelial cells 189
Figure 4.1 Lack of ACE2 expression and productive SARS-CoV-2 replication in endothelial cells 223
Figure 5.1 Experimental design for exposure of hESC-EC to shear stress. 237
Figure 5.2 Generation of hPSC derived cardiac spheroids. 238

List of Supplementary Figures

Supplementary Figure 2.1 136
Supplementary Figure 2.2 138
Supplementary Figure 2.3 139
Supplementary Figure 2.4 140
Supplementary Figure 2.5 141
Supplementary Figure 2.6 142
Supplementary Figure 2.7 143
Supplementary Figure 2.8 144
Supplementary Figure 2.9 145
Supplementary Figure 2.10 146
Supplementary Figure 2.11 147
Supplementary Figure 2.12 148
Supplementary Figure 2.13 149
Supplementary Figure 2.14 150
Supplementary Figure 2.15 151
Supplementary Figure 2.16 152
Supplementary Figure 3.1 206
Supplementary Figure 3.2 207
Supplementary Figure 3.3 208
Supplementary Figure 3.4 209
Supplementary Figure 3.5 210
Supplementary Figure 3.6 211

List of Supplementary Tables

Supplementary Table 2.1: Antibodies used for flow cytometric analysis. . 129
Supplementary Table 2.2: Summary of sequencing of 10X libraries. 130
Supplementary Table 2.3: Summary of 10X EC data obtained from publicly available datasets. 131
Supplementary Table 3.1 Antibodies used for staining of fetal cardiac EC 201
Supplementary Table 3.2 Silencer Select siRNA..... 201
Supplementary Table 3.3 Probes used for TaqMan Assays..... 201
Supplementary Table 3.4 Antibodies used for Western Blot analysis 202

Chapter 1 - Introduction

1.1. Overview

The branched vessel networks of the blood and lymphatic vasculature penetrate nearly all tissues in the body, providing essential functions to maintain organ homeostasis throughout embryonic development and adult life. The blood vasculature is a closed network consisting of interconnected arteries, capillaries, and veins, which together function to transport gases, nutrients, cells, proteins, and signalling molecules to neighbouring tissues[1]. Conversely, the lymphatic vasculature is an open network of blind-ended vessels functioning primarily to regulate tissue fluid balance, as well as additional roles in immune cell trafficking and fat absorption[2]. Endothelial cells (EC) comprise the innermost layer of all blood and lymphatic vessels. Beyond functioning as a simple barrier between the blood and vessel wall, the single EC monolayer is poised to respond to physical and chemical signals via the production of regulators central to maintaining vascular homeostasis[1]. The physical and functional properties of EC differ depending on the vascular bed and tissue type in which they are localised[3].

The formation of new blood vessels during embryonic development occurs via two separate processes, vasculogenesis and angiogenesis[4]. Vasculogenesis refers to the de novo derivation of EC from mesodermal progenitors, whilst angiogenesis describes the emergence of new EC via the sprouting of pre-existing vessels[4]. The transcriptional control of EC development has been extensively studied using a range of model organisms including mouse and zebrafish[5]. More recently, the differentiation of embryonic stem cells to EC has

served as an additional model to study endothelial development. This includes the use of human embryonic stem cells (hESC) which has provided human relevance to previous findings from model organisms[5]. However, further characterisation of the transcriptional changes accompanying the differentiation of hESC to EC is required to allow the identification of novel regulators controlling human EC development. This information is also likely to prove useful in future strategies to guide further specification of hESC derived EC (hESC-EC) *in vitro*, as well as attempts to drive neovasculogenesis for the treatment of ischaemic disease.

Transcription factors (TF) belonging to the E26 transformation specific (ETS) family have been demonstrated to play a central role in EC differentiation and the formation of vascular networks[6, 7]. However, increasing amounts of evidence have suggested that ETS TF function cooperatively to fulfil their EC specific function, either with each other or with members of other TF families[6, 8, 9]. Furthermore, efforts to characterise the roles of individual TF in endothelial differentiation and function is further complicated by the functional redundancy existing between related ETS TF[6].

Following the formation of the primitive vascular network, the endothelium must undergo further specification during vascular remodelling to give rise to a mature network of arteries, veins, and capillaries[10]. Central to this process is the specification of EC into arterial or venous identities. Although the regulation of arteriovenous specification remains incompletely understood, several

opposing transcriptional programs have been identified to promote either arterial or venous identity[11]. This includes NR2F2 (also known and COUP-TFII) as the master transcription factor for venous EC identity[12], and Notch signalling as a positive regulator of arterial EC specification[13].

In addition to specification to arterial or venous identities, EC must undergo further specialisation in response to intrinsic or extrinsic cues to meet the individual requirements of the surrounding tissue[3]. This level additional level of vessel- and organ-specific EC specification results in a high degree of endothelial heterogeneity across the vasculature[3]. The emergence of single-cell RNA sequencing (scRNA-seq) technologies has made it possible to capture the transcriptional state of individual cells, thus allowing for cellular heterogeneity to be evaluated at single cell resolution[14]. In recent years, an increasing number of studies have applied scRNA-seq to study cellular heterogeneity within the heart and vasculature, including mapping the transcriptional profile of the endothelium across different vascular beds and tissues throughout development, adulthood, and disease[15-17]. However, studies designed to be empowered to characterise the transcriptional profile of the endothelium have predominantly utilised murine models, with many human datasets containing a limited number of EC.

In this thesis I first provide a review of the current knowledge surrounding endothelial development and further specification with a particular focus on how these processes are regulated at the transcriptional level. I then provide

scRNA-seq data mapping the transcriptional dynamics during the differentiation of hESC derived endothelial cells (hESC-EC), identifying potentially novel regulators promoting the specification of EC from a human mesodermal origin. Finally, again using scRNA-seq, I then characterise the transcriptional landscape of the developing cardiac endothelium from human fetal heart tissue, identifying the TF *MECOM* as a novel regulator of arteriovenous identity.

1.2 Development of the endothelium

1.2.1. Overview of endothelial development

The formation and homeostasis of the vascular system is an essential requirement to supporting the metabolic demands of complex multicellular organisms throughout embryonic and adult life. Defects in vascular formation have been demonstrated to result in embryonic lethality or subsequent vascular pathology[18]. Prior to development of other organ systems, the formation of the vasculature begins with the assembly of a primitive vascular plexus consisting of a single layer of endothelial cells. Vessel formation is known to occur via two distinct processes, vasculogenesis and angiogenesis[19]. Vasculogenesis, refers to the *de novo* formation of endothelial cells (EC) via the differentiation of mesodermal progenitors. On the other hand, angiogenesis describes the process by which new vessels form by the proliferation, migration, and sprouting of endothelial cells in pre-existing vessels[19] (Figure 1.1).

1.2.2. Vasculogenesis

During mammalian embryonic development, vasculogenesis is first observed in the extraembryonic yolk sac as early as the gastrulation stage, prior to later commencing within the embryo proper[20, 21]. Within the yolk sac, signals from the underlying yolk sac endoderm stimulate cells of the extra-embryonic mesoderm to form aggregates called blood islands. Endothelial cells comprise the outer layer of blood islands whilst the inner core consists of primitive haematopoietic cells[21]. Fusion of blood islands leads to the formation of a primitive capillary plexus which subsequently forms an extensive network of arteries, veins, and capillary beds via angiogenesis and vascular remodelling[22].

Figure 1.1 Development of the endothelium by vasculogenesis and angiogenesis.

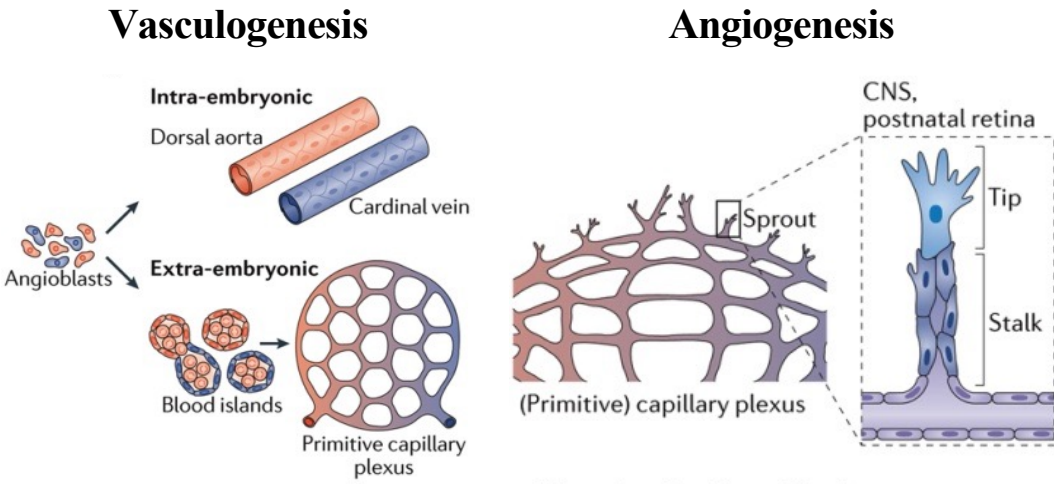


Figure 1.1: Vasculogenesis (left hand panel) within the embryo involves the fusion of endothelial progenitor cells, termed angioblasts, to form the dorsal aorta and cardinal vein. Within extraembryonic tissue, angioblasts fuse to form blood islands which subsequently remodel to form a primitive capillary plexus. The vascular network expands via angiogenesis (right hand panel) in which new vessels are formed from pre-existing vessels via vessel sprouting. Figure adapted from Potente et al, 2017.

Within the embryo proper, vasculogenesis is required for the formation of the primitive embryonic circulation, including the dorsal aortae, vitelline veins, cardinal veins, and endocardium[20, 21]. Vasculogenesis is first initiated in the cranial region of the embryo when cardiogenic mesoderm undergoes differentiation to form endocardial precursor cells, which subsequently coalesce to form the endocardial layer of the primitive heart tube[21, 23]. Similarly, outside the cardiac crescent, mesodermal derived endothelial precursor cells, termed angioblasts, coalesce to form the paired dorsal aortae[20]. Further merging of angioblasts posterior to the cardiac crescent gives rise to the vitelline veins (responsible for draining the yolk sac), shortly followed by the emergence of anterior and posterior cardinal veins[20].

Extensive studies using murine lineage tracing models and the differentiation of embryonic stem cells (ESC) have demonstrated EC to be derived from FLK1+ (fetal liver kinase 1; also known as VEGFR2 or KDR) mesoderm[4]. In mouse, Flk1+ cells first emerge at the posterior portion of the primitive streak and subsequently in the newly formed lateral mesoderm[24]. *Flk1* expression then persists in the vasculature of the embryo proper and yolk sac as well as in the endocardium[24]. Embryonic lethality is observed in *Flk1*$^{-/-}$ embryos with the absence of blood islands, embryonic vasculature, and endocardium, thus demonstrating a requirement for Flk1 in murine vasculogenesis[25]. In addition to giving rise to EC, FLK1+ mesoderm has been shown to represent a multipotent progenitor population for multiple mesodermal derived cell types (Figure 1.2). Flk1+ cells isolated from the primitive streak of E7.5 embryos can

give rise to both endothelial and haematopoietic lineages in ex vivo culture[26]. Similarly, Flk1+ cells isolated from differentiating ESC were able to undergo differentiation into either ECs, cardiomyocytes, blood cells, or smooth muscle cells (SMC) depending on the culture conditions used following replating[27-29]. Supporting these findings in ESC differentiations, Kattman et al[29] identified a comparable Flk1+ clonal population from mouse embryos with capability to form EC, cardiomyocytes, or SMC during *ex vivo* culture.

Signals derived from the endoderm are imperative for the initiation of vasculogenesis from FLK1+ mesoderm. For instance, endodermal derived fibroblast growth factor 2 (FGF2) has been shown to promote angioblast formation[30]. Endodermal derived hedgehog ligands, Sonic hedgehog (SHH) and Indian hedgehog (IHH) are also required to initiate vasculogenesis[31]. Endodermal SHH and IHH invoke expression of forkhead box F1 (FOXF1) in the mesoderm which in turn promotes expression of bone morphogenic protein 4 (BMP4)[31]. The presence of BMP4 has been shown to be required for formation of the yolk sac vasculature in an explant culture model as well as for promoting the differentiating of hESC to EC[31, 32]. Vascular endothelial growth factor (VEGF) plays a central role in vasculogenesis and angiogenesis[33]. VEGF-A is the most studied member of the VEGF family with a prominent role in vasculogenesis with expression from the endoderm and mesoderm[33]. Loss of a single Vegfa allele result in embryonic lethality with abnormal blood vessel formation[34]. FLK1 (VEGFR2) is the most important of the VEGF receptors during angiogenesis and vasculogenesis and is responsible for the

propagation of the VEGF-A signal via various canonical and non-canonical downstream pathways to promote EC differentiation, migration, and survival[35, 36]. In contrast, activation of VEGFR1 via VEGF-A results in comparatively less intracellular tyrosine kinase activity, indicating VEGFR1 functions predominantly to sequester VEGF-A from VEGFR2 binding[33, 37].

Figure 1.2 FLK1+ mesoderm represents a multipotent population.

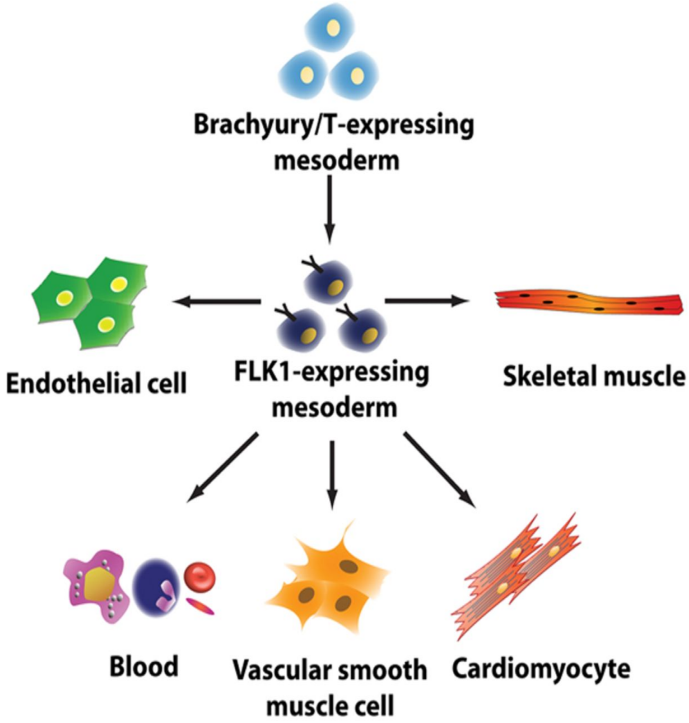


Figure 1.2: FLK1+ mesoderm arises from T/Brachyury expressing mesoderm and represents a multipotent progenitor cell population with capability to differentiate into a range of mesodermal derived cell types. Figure adapted from Park et al, 2013.

1.2.3. Evidence for a common clonal progenitor for endothelial and haematopoietic lineage

The high degree of similarity between the emergence of endothelial and haematopoietic lineages led to the view that both must originate from a

common precursor cell type, named the haemangioblast. The first evidence for the haemangioblast as a clonal precursor came from the differentiation of mouse embryonic stem cells (mESC) in embryoid bodies (EB) from which a transient mesoderm population could be isolated capable of differentiating into either endothelial or haemogenic populations[27]. Subsequent investigations identified a comparable population isolated from the primitive streak of the gastrulating mouse embryo, with potential to differentiate into blood cells or EC during *ex vivo* culture[26]. Further evidence for the existence of the haemangioblast came from fate mapping studies in zebrafish which identified bipotential cells within the ventral portion of the embryo as the origin of both endothelial and haematopoietic lineages[38]. A comparable colony forming cell type with haematopoietic and endothelial potential was also isolated from EB made from human embryonic stem cells (hESC)[39].

Despite these findings, there remains little evidence in mammals of the in presence of a haemangioblast population capable of giving rise to both lineages *in vivo*, with previous lineage tracing studies in mouse having failed to validate the existence of haemangioblasts[40, 41]. However, concerns have since been raised questioning the experimental design and interpretation of results from these studies[42]. These contrasting findings, and the fact the bipotential capability of the mammalian haemangioblast has only been observed during explant culture of isolated cells, has led to the suggestion that whilst a bipotential progenitor state capable of giving rise to both lineages

may exist, the constraints of the surrounding *in vivo* environment mean that under normal conditions only one of the two fates is ever specified[43].

Further evidence against a haemangioblast population *in vivo* was provided by increasing understanding of a specialised endothelial population, termed haemogenic endothelium, capable of giving rise to haematopoietic cells via endothelial to haemogenic transition (EHT)[42]. Lineage tracing of endothelial cells in mouse using a VE-cadherin Cre reporter line revealed the emergence of haematopoietic stem cells (HSC) from endothelium of the developing aortic-gonado-mesonephros[44]. Frame et al[45] subsequently also identified haemogenic endothelium as the source of definitive haematopoiesis in the yolk sac of the E8.25 mouse embryos. However, given that primitive erythropoiesis within blood islands is known to occur prior to the formation of the vasculature suggests that primitive erythroblasts cannot be derived from via EHT. However, investigation revealed primitive hematopoietic cells are in fact derived from aggregates of endothelial marker expressing progenitors (termed haemogenic angioblasts) localised within blood islands, as opposed to from a bona fide haemogenic endothelial population[46, 47].

Collectively, and although subject to ongoing debate, current evidence suggests a model in which all blood cells are derived from a specialised endothelial population via EHT as opposed to from a bipotential progenitor of both blood and endothelial lineages[42, 43]. However, novel experimental

strategies are required to conclusively determine the existence or absence of a mammalian haemangioblast population *in vivo*.

1.2.4. Development of endothelium from paraxial mesoderm

Although endothelium is typically described as being derived from FLK1+ lateral plate mesoderm (LPM), several studies have also determined a contribution from paraxial mesoderm (PXM)[48, 49]. Interestingly, the mesodermal origin of EC has been shown to influence their subsequent localisation within the vasculature[49]. This challenges the previous dogma that in response to specific external stimuli, a single endothelial progenitor can undergo specification to give rise to the large spectra of heterogenous EC populations observed across the vasculature[48, 49].

During gastrulation, cells of the epiblast migrate through the primitive streak whilst undergoing epithelial to mesenchymal transition (EMT). Once localised between the visceral endoderm and overlying epiblast, cells undergo commitment to either mesoderm or definitive endoderm[50]. Although the pleiotropic signalling mechanisms governing these processes are still incompletely understood, both murine and human embryonic stem cell hESC models have demonstrated an essential role for BMP4 in mesoderm induction[51-53]. Additionally, during mediolateral patterning of the mesoderm, BMP4 also acts as a lateralising signal to induce a lateral mesoderm fate, whilst FGF, canonical Wnt signalling, and Noggin (a BMP4 antagonist) are necessary for induction of PXM fate[54-56]. Following its specification, PXM

undergoes segmentation giving rise to somites containing two major populations: dermomyotome and sclerotome[57]. Dermomyotome cells, defined by expression of Pax3 in mouse, have been demonstrated to give rise to endothelial cells[49, 58]. The decision in this Pax3+ population to specify to endothelium over a myogenic fate is driven by increased Notch signalling[58]. Furthermore, a recent study from Stone et al[49], using a *Pax3* lineage tracing mouse model revealed PXM as the predominant source of lymphatic EC across multiple organs and tissues.

1.2.5. A potential haematopoietic contribution to the vascular endothelium

In recent years, the consensus that during development the entirety of the vascular endothelium is derived from mesodermal progenitors via vasculogenesis or the subsequent expansion of pre-existing vessels via angiogenesis has been challenged[59]. Plein et al[59] reported that erythro-myeloid progenitors (EMP), could undergo transition to form EC capable of integrating into the vasculature of different organs. EMP arise in the yolk sac from haemogenic endothelium via EHT and in combination with the erythroid progenitor cells from the primitive wave in blood islands, EMP can provide sufficient blood cells to support prenatal life prior to the emergence of haematopoietic stem cells (HSCs)[60, 61]. Plein and colleagues[59] used a murine model in which CRE recombinase was placed under control of a myeloid lineage specific gene promoter (*Csf1r*) to label EMP and their progeny within the developing embryo. This revealed the incorporation of labelled EC into the

vasculature of multiple organs including the liver, brain, heart, retina, and lungs[59]. Additionally, *Csf1r* labelled EMP demonstrated potential to differentiate into EC during in vitro culture[59]. However, these findings were disputed in a subsequent study, which reported an absence of labelled cells within the endothelium of the lung, liver, brain, and heart of their Csf1r reporter mouse line[62]. It therefore remains controversial whether these haemogenic progenitors can revert to their endothelial origin and incorporate into the vasculature with mesodermal derived endothelium.

1.3. Transcriptional control of vascular development

1.3.1. ETS Transcription Factors

Amongst the numerous transcription factors (TF) demonstrated to have an involvement in endothelial development and identity, none have been found to have a more central role than those of the ETS family[6, 7]. Virtually all characterised endothelial enhancers and promoters to date have been found to contain an ETS binding domain, with ETS TF binding being an essential requirement for the expression of many endothelial associated genes[8, 18, 63-65]. RT-qPCR analysis of the 28 ETS TF conserved in humans found 19 to be expressed at some level in human umbilical vein endothelial cells (HUVEC)[66].

All members of the ETS family of TF are characterised by the presence of an ETS domain approximately 85 amino acids in length. This domain is a winged helix turned motif consisting of 3 alpha helices and four anti-parallel beta sheets, with two DNA contacting arginine residues in the 3rd helix thought to

facilitate DNA binding specificity[67]. In addition to the ETS domain, a subset ETS TF have also been found to also contain a pointed domain, required to allow protein dimerization[68, 69]. Unlike the rest of the members of ETS TF family, ETV6 contains a transcriptional repressor domain instead of the usual transcriptional activating domain[70]. ELK3 was found to contain both activating and repressing domains, facilitating its capability to act as a repressor or activator for different targets (Figure 1.3) [71-73].

Figure 1.3 Domain structure of ETS transcription factors.

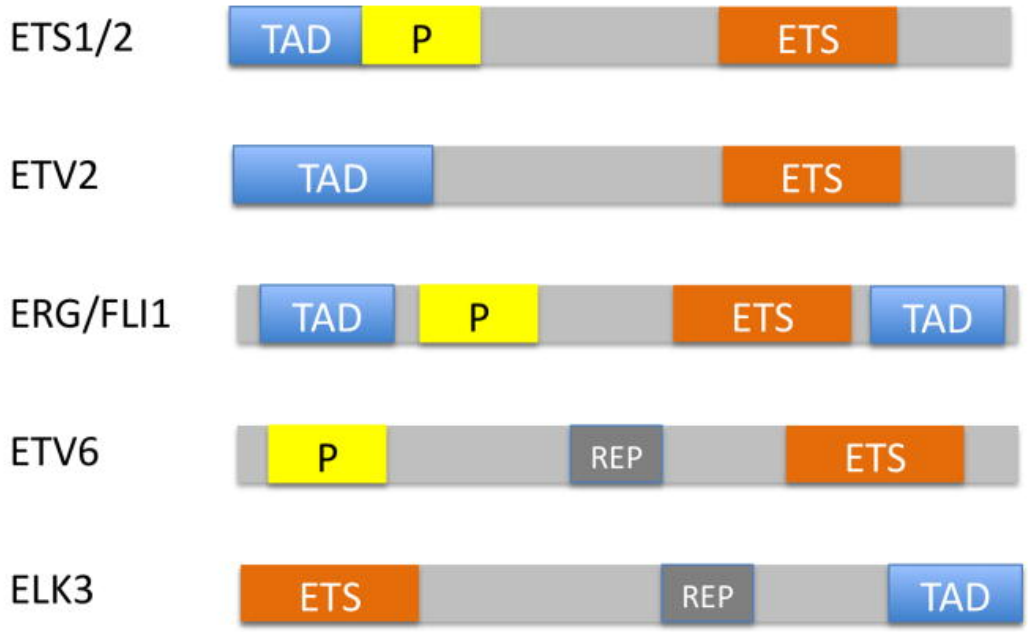


Figure 1.3: Domain structure of selection of ETS TF with function in vascular development including ETS domain (ETS), transcriptional activator domain (TAD), pointed domain (P) and repressor domain (rep). Figure adapted from Meadows et al, 2011.

The ETS domain binds to the highly conserved core DNA 5'-GGA(A/T)-3' sequence. However, ETS TF expression is not restricted to just EC and the

core DNA binding motif is not unique to endothelial gene promoters/enhancers. Therefore, it was suggested that ETS TF must function cooperatively with other proteins to enable the downstream expression of EC specific genes[66]. This idea was subsequently validated in a study by De Val et al[8] which identified FOX:ETS binding motifs in many EC specific enhancers, with joint expression of *foxc2* and *etv2* (an ETS TF) in Xenopus embryos being required to elicit ectopic EC gene expression[8]. More recently it has was also demonstrated that whilst ETS binding motifs are necessary for the enhancer activity of both arterial and venous EC genes, their selective expression in arterial or venous EC requires the binding of additional TFs to mediate activation or repression of select genes[9].

ETV2 (ETS Variant Transcription Factor 2)

Unlike for most other members of the ETS TF family, the detrimental effect of the absence of ETV2 (also known as ER71) in the developing endothelium cannot be compensated by the redundancy that exists between other ETS TF. This places ETV2 as a master regulator of EC fate[74]. Loss of *Etv2* in mouse embryos results in embryonic lethality by E9.5 with a complete absence of the endothelial lineage[75, 76]. Similar findings have been observed following morpholino mediated *etv2* knockdown in zebrafish embryos with a near absence of flk1+ EC[77].

In situ hybridisation (ISH) studies in developing mouse embryos revealed *Etv2* exhibits transient expression in the primitive streak, blood islands, and

embryonic mesoderm at E7.5 before expression is rapidly downregulated, becoming more restricted to larger vessels by E9.5[76]. By adulthood, *Etv2* expression becomes confined to the testis[78]. A similar pattern of transient expression is also observed in mESC differentiation where *Etv2* expression closely follows the increase of primitive streak marker *Brachyury* before subsequently decreasing[76]. Furthermore, loss of *Etv2* at later stages in development (E9.5), does not result in any observable vascular defects, suggesting that *Etv2* is not required following the specification of vascular progenitors[79].

Further insight into the role of ETV2 in establishing endothelial fate came from the use of a transgenic mouse model in which the *EYFP* gene (Enhanced yellow fluorescent protein) was placed under control of the *Etv2* promoter, allowing the tracing of vascular mesoderm progenitors in wild type and *Etv2* mutant backgrounds[80]. Flow cytometric analysis of EYFP+ progenitors revealed a reduction in the FKL1+/PDGFRA- mesodermal population in mutant embryos[80]. This population represents the lateral mesoderm from which endothelial and haematopoietic lineages are predominantly derived (Figure 1.4)[81, 82]. Studies using *Etv2* mutant mouse embryos and *Etv2* knockout mESC demonstrated normal formation of unpatterned FLK1+/PDGFRA+ mesoderm, despite the resultant lack of an emerging endothelial population[83]. However, the lack of *Etv2* was found to prevent the progression of unpatterned/primitive (FLK1+/PDGFRA+) mesoderm into

endothelial generating FKL1+/PDGFRA- lateral mesoderm, resulting in an absence in EC differentiation[83].

Figure 1.4 Subpopulations of FLK1+ mesoderm.

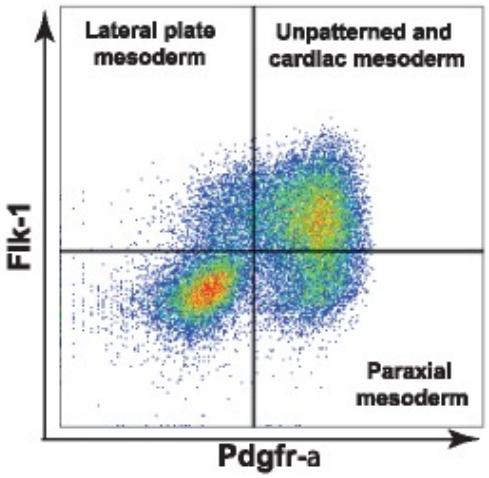


Figure 1.4: FLK1+ mesoderm undergoes further specification to form distinct subsets, each with different differentiation capabilities. Subsets can be distinguished from each other by FLK1 and Platelet Derived Growth Factor Receptor Alpha (PDGFRA). Figure adapted from Koyano-Nakagawa et al (2017).

Etv2 expression was found to be reduced in *Flk1* deficient embryos, whereas *Flk1* expression was not significantly altered in *Etv2* null embryos, suggesting *Etv2* functions downstream of *Flk1*[84]. Furthermore, overexpression of *Etv2* in *Etv2* knockout EB reinstated the capability to form KDR+/PDGFRA- lateral mesoderm, with a resulting rescue in EC differentiation[85]. This rescue effect could not be accomplished by forced overexpression of *Flk1*[85].

In addition to ETV2 being required for endothelial differentiation, the lack of *Etv2* has been shown to promote specification to alternative mesodermal fates[80, 83, 86, 87]. For instance, recent scRNA-seq of *etv2* deficient zebrafish embryos demonstrated that the lack of *etv2* causes mesodermal progenitors to undergo specification to a population defined by expression of skeletal

muscle markers[87]. Similarly, the absence of *Etv2* in murine embryos results in expansion of the myocardium via the alleviation of ETV2 inhibition on canonical Wnt signalling during mesoderm specification[86]. This ability of ETV2 to modulate Wnt signalling has led to the suggestion that ETV2 may also play a role in repressing cardiomyocyte fate during endocardial specification from *Nkx2-5*+ cardiac progenitor cells[23, 75, 86, 88].

Collectively, these studies demonstrate ETV2 performs a non-redundant role via promoting mesodermal fate towards the endothelial lineage, thus making it an essential TF for EC development and mesodermal patterning. Whilst some downstream targets of ETV2, such as *SCL* and *FLI1* have been identified[83], future work is required to fully characterise the downstream mechanisms by which ETV2 functions as a master regulator of EC differentiation. Furthermore, given the reported role for ETV2 in vascular regeneration[89], improved understanding is likely to be informative for future regenerative medicine applications.

ERG (ETS related gene)

ERG represents another well characterised ETS TF responsible for a range of functions in endothelial development and homeostasis. This includes regulation of vascular permeability, angiogenesis, cell survival, and regulation of EC differentiation (reviewed by Shah et al[90]). Like many other ETS TF, the function of ERG has been demonstrated, at least in part, to be dependent on physical and/or functional interaction with other TF. This includes other

members of the ETS family such as ETS2 and FLI1[69]. Physical binding of ERG to other TF to form heterodimers is thought to be facilitated by both the ETS and pointed domains[69]. The cooperative nature of ERG was highlighted in a study by Liu et al[91] which revealed that the dual knockout of *erg* and *fli1a* in zebrafish resulted in a severe angiogenic defect and downregulation of *cdh5*, whereas their individual knockout only caused a comparatively mild phenotype. Working either alone or cooperatively, ERG has been determined to drive the expression of multiple endothelial specific genes, including *CLDN5, CHD5, VWF, Dll4, FLK1*, and *SNAI1*[92-97]. Recently, the EC specific activity of ERG was revealed to be achieved via the regulation of lineage specific enhancers and super-enhancers capable of eliciting expression of EC specific genes[98].

ERG expression during murine embryogenesis remains confined to mesoderm derived tissue, including the blood islands at E8.5 and the developing vascular endothelium of the embryo proper where it persists throughout adulthood[99, 100]. The spatiotemporal pattern of *Erg* expression was found to closely overlap with that of *Fli1*, supporting their propensity to function cooperatively[99]. In zebrafish, *erg* is expressed within the lateral plate mesoderm during vasculogenesis before expression gradually decreasing as development proceeds[101].

ERG has been proposed to play a direct role in endothelial differentiation[102, 103]. Ectopic *erg* expression in Xenopus embryos was found to be sufficient to induce expression of a single EC marker (X-msr)[102]. shRNA knockdown of *Erg*

in EB also revealed a subsequent reduction in the proportion of the resultant endothelial population[103].

Constitutive endothelial specific deletion of *Erg* by crossing *Tie2-Cre* mice with *Erg* floxed mice resulted in embryonic lethality between E10.5 and E12.5 with two independent studies reporting defective vascular remodelling within the yolk sac and embryo as well as the presence of haemorrhages[104, 105]. Interestingly, Han et al[105] reported no observable differences between *Erg* knockout and wild type mice up until E9.5. This would appear to contradict the findings from ESC differentiation and Xenopus studies, instead suggesting that ERG is responsible for vascular remodelling of the primitive vascular network as opposed to having a non-redundant function in endothelial differentiation. A comparable haemorrhaging phenotype was also observed in the head region of zebrafish embryos following the knockdown of e*rg*[91]. Dysfunctional vascular remodelling with haemorrhages also occurred following the global knockdown of *Erg* during murine development with dysfunctional cardiac valve morphogenesis also being reported[97]. Abnormal valve formation was determined to be due to a reduction in the formation of endocardial-derived valvular mesenchyme via endothelial to mesenchymal transition (EndMT)[97]. ERG was found to promote expression of key EndMT regulators, *Snail1* and *Snai2*, by binding to their proximal promoter and intronic regions, respectively[97].

In contrast to these severe vascular defects reported in *Erg* mutant mouse embryos[97, 104, 105], Lathen et al[106] reported that global *Erg* knockout resulted in a delayed vascular phenotype with embryos surviving until between E16.5 and 3 months after birth. A proposed explanation for this discrepancy was that the substitution mutation in this knockout model was insufficient to block DNA binding activity[91]. However, future studies are required to validate the full extent of the absence of ERG during vascular development.

FLI1 (Friend Leukaemia Integration 1)

FLI1 is another key ETS TF identified to play an important role in EC development and vascular homeostasis. Expression of *Fli1* in mouse is observed in early progenitors of both endothelial and haemogenic lineages before becoming more confined to the endothelium at later stages and during adulthood[107, 108]. Abedin et al[108] noted that the expression of *Fli1* (and its close family member *ERG)* was downregulated in *Etv2* null mouse embryos, suggesting that *Fli1* is a direct target of ETV2. This was subsequently confirmed by co-transfecting a luciferase reporter construct containing the *Fli1* promoter and increasing amounts of *Etv2* expression plasmid into endothelial and non-endothelial cell types[108]. Furthermore, it was determined in mouse that following the decline in *Etv2* expression by midgestation that FLI1 itself is capable of binding to the ETS binding sites within the *Fli1* promoter to regulate its own expression[108].

Morpholino knockdown *fli1* in Xenopus embryos revealed a dramatic reduction in genes known to be involved in endothelial and haemogenic lineage development, suggesting a role for Fli1 in the initial specification of EC[109]. Further evidence for FLI1 fulfilling this role was provided by identifying that expression of *FLI1* in conjunction with increased protein kinase C activity promoted the differentiation of hESC to EC[110]. However, these findings contrast with results from studies perturbing *Fli1* in mouse and zebrafish which suggest FLI1 is dispensable for EC specification[109, 111]. Embryonic lethality in homozygous *Fli1* knockout mouse embryos was found not to occur until E12-E12.5 with embryos demonstrating dysfunctional haematopoiesis and brain haemorrhaging due to excessive apoptotic cell death of EC[108, 111]. Similarly, morpholino *fli1* knockdown in zebrafish was found not to disturb the emergence of the endothelial lineage[109].

Despite not having an apparent role in vasculogenesis, Craig et al[74] demonstrated a lack of angiogenic sprouting in the vasculature of zebrafish embryos devoid of both *etv2* and *fli1b* (one of the two *fli1* isoforms found in zebrafish). Subsequent RNA sequencing and chromatin immunoprecipitation revealed substantial overlap in the downstream targets for both TF. This suggests FLI1 functions in a redundant manner with ETV2 to regulate embryonic angiogenesis[74]. The mechanism by which FLI1 regulates angiogenesis was demonstrated to involve a cooperative function with one of its transcriptional targets, GATA2[109, 112]. Marks-Bluth et al[112] determined both GATA2 and FLI1 were capable of binding to regulatory elements controlling

the expression of *Smad1* and *Smad5,* both known regulators of angiogenic sprouting. More recently, *fli1*+ cells isolated from zebrafish embryos were found to highly express histone methyltransferase *prdm16*, with a subsequent function identified for *prdm16* in angiogenesis *in vivo* and *in vitro*[113].

Role of other ETS Transcription Factors

Other members of the ETS TF family have been determined to have a role in vascular development and homeostasis. However, the large degree of functional redundancy that exists between individual TF results in their role often being masked in gene perturbation studies. This was found to be the case for ETS1 (E-twenty-six gene 1) for which several *in vitro* studies have identified a role for ETS1 in angiogenesis[114-117]. However, *Ets1* knockout mice were found not to have any vascular abnormalities, despite defects with increased T cell apoptosis and terminal B cell differentiation[118]. However, embryonic lethality was observed by E15.5 in *Ets1*$^{-/-}$*;Ets2*$^{-/-}$ mouse embryos with observable defects in embryonic angiogenesis[119]. Furthermore, double deficient mice demonstrated increased apoptosis in CD31+ EC *in vivo* and *in vitro* suggesting a redundant role for *Ets1/Ets2* in EC survival during murine embryonic angiogenesis[119]. In contrast with findings from mouse[119], *ets1* deficiency alone in zebrafish increased EC apoptosis[120]. Additionally, Chetty et al[120] recently found that *ets1*$^{-/-}$*;etv2*$^{-/-}$ mutant zebrafish embryos displayed a more severe defect in angiogenesis compared to single mutant *etv2*$^{-/-}$ embryos. The absence of *ets1* was found to further downregulate expression of *flk1* in *etv2* mutants, with the defective angiogenic phenotype being partially

rescued by constitutive activation of MAPK signalling[120]. Together this suggests that the lack of vascular phenotype in ETS1 mutants is likely due to the compensatory roles played by other ETS TF.

ETV6, a documented transcriptional repressor[70], has also been identified to have a functional role in angiogenesis via its association with transitional repressor CtBP to form a heterodimeric ETV6-CtBP complex[121]. The ETV6-CtBP complex was determined to be responsible for priming EC for angiogenic sprouting via the transcriptional repression of sprouting antagonists such as *Dll4* and *Sprouty*[121]. In line with these in vitro studies, *Etv6* knockout mouse embryos displayed dysfunctional yolk sac angiogenesis phenotype leading to embryonic lethality by E10.5[122]. More recently, Li et al[123] reported Etv6 to induce *vegfa* expression in Xenopus embryos via simultaneous activation and repression mechanisms. Etv6 was shown transcriptionally repress *foxo3*, a known repressor of *vegfa* expression[123]. Simultaneously Etv6 was also found to upregulate *klf4* which subsequently coordinated the recruitment of Etv6 to the *vegfa* promoter to induce *vegfa* expression[123].

1.3.2. Forkhead box (FOX) Transcription Factors

FOX TF are group of evolutionary conserved group of transcriptional regulators, involved in a diverse range of functions throughout development and adult life and in various pathologies. All FOX TF share a conserved ~100 residue winged helix DNA binding domain, named the forkhead box, which is capable of binding to the promoters of target genes[124]. FOX TF are grouped

into subclasses, designated A-S, based on their degree of sequence similarity outside the forkhead box domain[124].

FOXO1 is highly expressed in endothelial cells and demonstrated to be an important regulator of vascular development[125, 126]. Global *Foxo1* knockout mouse embryos die by around E11 with profound defects in vascular development with dilation of the capillary plexus in the heads of *Foxo1*$^{-/-}$ embryos[125]. More recently, Wilhelm et al[127] demonstrated that endothelial specific deletion of *Foxo1* caused hyperproliferation of EC resulting in vessel enlargement (Figure 1.5). In contrast, constitutive activation of *Foxo1* had an opposing effect with vascular networks containing fewer EC and narrowed lumen. The ability of FOXO1 to mediate EC proliferation was found to be mediated by its suppression of MYC and production of metabolite S-2-hydroxyglutarate resulting in reduced glycolysis and mitochondrial respiration[127, 128]. Together these studies identify FOXO1 as a master regulator of vascular expansion.

Transcription factors of the FOXC subclass are also involved in regulating EC development[8 129-133]. As discussed above, FOXC TF can interact with ETS TF to bind to FOX:ETS binding motifs to regulate the expression of many EC specific genes[8]. *Foxc1* mutant mice develop vascular defects including narrowing of the aortic arch leading to death in either the prenatal or perinatal period[129]. *Foxc2* null mice exhibited defects in lymphatic vasculature, consistent with recent reports revealing a role in lymphatic valve formation and

maintenance[130, 131]. Compound *Foxc1+/-;Foxc2-/-* mutants demonstrated arteriovenous malformations with the lack of induction of arterial EC markers as well as reduced sprouting of lymphatic EC from veins[132]. The subsequent identification of FOXC binding motifs in the promoters of *Hey2* and *Dll4* further indicated a direct role for FOXC TFs in inducing arterial EC specification[133].

Figure 1.5 Retinal vasculature of control and FOXO1 KO mice

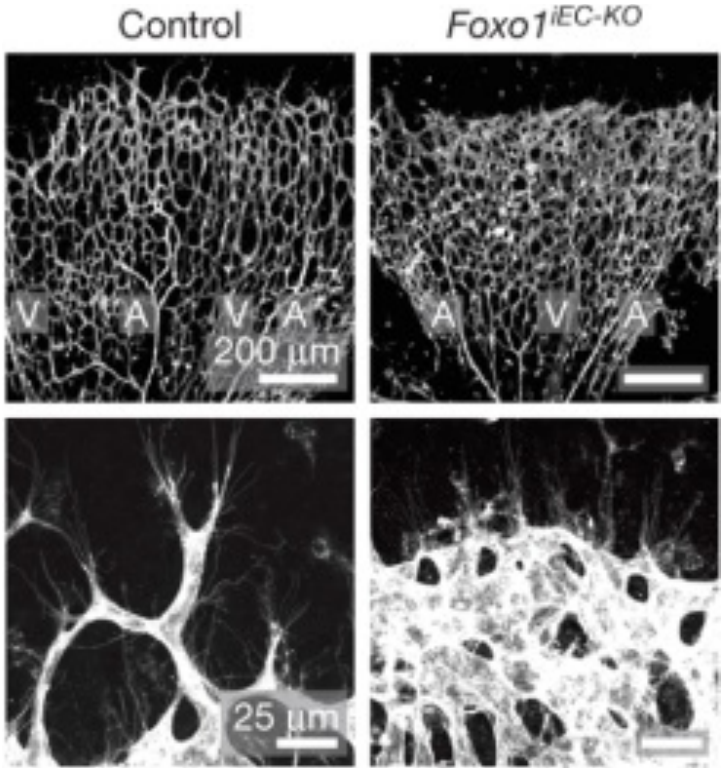


Figure 1.5: Isolectin B4 stained postnatal day 5 (P5) retinal vasculature from either control mice (left) or mice with EC specific *Foxo1* knockout (*Foxo1*$^{iEC-KO}$). EC hyperproliferation observed in retinal vasculature of *Foxo1*$^{iEC-KO}$ mice. Figure adapted from Wilhelm et al (2016).

Other members of the FOX TF family implicated in vascular development include FOXF1 and FOXH1. *Foxf1* mutant mice die by midgestation and lack expression of EC markers in the yolk sac due to a failure of mesodermal differentiation[134]. Choi et al[135] also demonstrated in zebrafish that Foxh1 can

bind to a *flk1* enhancer causing repression of *flk1* expression and a resulting negative effect on vascular formation.

1.3.3. SOX Transcription Factors

Members of the SOX TF family are characterised by the presence of a conserved HMG box which is a sequence specific DNA binding motif[136]. SOX factors are divided into subgroups designated A-H, based on their sequence conservation and gene expression profile[136]. TF belonging to the SOXF subgroup have been demonstrated to play an important role in cardiovascular development and homeostasis[137]. This group consists of SOX7, SOX17, and SOX18[137].

Single cell PCR analysis demonstrated enrichment of both *Sox7* and *Sox18* in the ETV2+/FLK1+/CD41- cells from E7.5 mouse embryos, demonstrating their early expression during the emergence of the endothelial lineage in vivo[137]. Whilst *Sox7* and *Sox18* continue to be expressed across the entirety of the vascular endothelium, *Sox17* expression remains more restricted to arterial EC[138, 139]. This expression profile conforms with the reported role of SOX17 in arterial EC specification (discussed below)[138].

In line with *in vivo* findings, culture of FLK1+ cells isolated from EB revealed the expression of *SOX7* to coincide with the emergence of EC markers, suggesting a function during vasculogenesis[140]. Whilst global or endothelial specific knockdown of *Sox7* in mouse did not block the emergence of CD31+

endothelial cells, severe vascular defects were evident as early as E8.5. This included disorganisation of the posterior portion of the dorsal aorta, suggestive of a failure of angioblasts to correctly coalesce during vasculogenesis[140]. In addition to this vasculogenic defect, angiogenesis is also dysfunctional in Sox7 knockout mice, characterised by the disorganisation of the developing vascular plexus and a lack of angiogenic remodelling[140, 141]. Interestingly the loss of *Sox7* resulted in increased expression of *Sox17* and *Sox18*, supporting findings that SOXF factors can function redundantly during angiogenesis[139, 141]. Together these studies suggest Sox7 functions to regulate both vasculogenesis and angiogenesis during vascular development.

The vascular phenotype observed in *Sox18* knockout mice has been demonstrated to vary depending on the strain of mouse used, from no defect (mixed background) to embryonic lethality and extensive subcutaneous oedema due to failure to form a functioning lymphatic system[142-144]. Whilst the role of SOX18 in lymphatic EC specification will be discussed in detail below, Hosking et al[143] determined that the strain specific variability observed in *Sox18* knockout mice is due to the compensatory roles of *Sox7* and *Sox17* which are upregulated in the absence of SOX18 to induce lymphatic EC specification. In addition to its role in lymphatic specification, SOX18 has been demonstrated to control blood vessel development via regulation by VEGF-D, with *Vegfd*$^{-/-}$/*Sox18*$^{-/-}$ mice and zebrafish embryos displaying severe arteriovenous defects[145].

1.3.4. SCL (stem cell leukemia)

The *SCL* gene (also known as *TAL1*) encodes a basic helix-loop-helix TF required for both endothelial and blood development with expression first detected in FLK1+ mesoderm of E6.5 mouse embryos[21]. Ectopic overexpression of *scl* alone in both Xenopus and zebrafish is sufficient to induce EC fate, suggesting a function in the specification of endothelial cells from mesoderm during vasculogenesis[146, 147]. However, whilst knockout of *Scl* in mouse causes embryonic lethality owing to an absence of primitive erythropoiesis, blood vessel formation does occur, suggesting *Scl* is dispensable for vasculogenesis[148]. Chimeric mouse embryos, with *Scl* expression rescued exclusively within the haematopoietic lineage, demonstrated a lack of incorporation of *Scl* deficient *Tie2*+ endothelial cells into the macrovasculature during angiogenic remodelling[149]. A comparable phenotype was also observed in zebrafish following *scl* knockdown with a disruption of angiogenesis following angioblast specification[150]. Although SCL has been shown to regulate the activation of several endothelial specific enhancers and promoters, the mechanism by which it regulates angiogenesis remains unknown[18, 151].

In addition to its role in haematopoietic and vascular development, SCL is also necessary for the suppression of the acquisition of cardiomyogenic fate in non-cardiac regions[152]. This function was demonstrated in a study from Van Handel[152] et al which identified spontaneously beating cardiomyocytes from the yolk sac endothelium in *Scl*$^{-/-}$ mouse embryos[152]. Scl was also found to be

required to prevent misspecification of the endocardium to cardiomyocytes during murine cardiogenesis[152]. Subsequent CHIP-seq analysis of FLK1+ mesodermal cells revealed SCL represses cardiomyogenic fate by binding to and repressing gene enhancers responsible for the activation of known cardiac TFs such as *Gata4* and *Hand1*[153].

1.4. Arteriovenous Endothelial Specification

1.4.1. Overview of arteriovenous specification

During the formation of the primitive vascular network, the endothelium must undergo further specification to give rise to distinct arterial and venous EC populations. This requirement for arteriovenous specification is necessary to generate distinct EC populations with specific functions, allowing them to fulfil the individual demands of the vascular bed in which they reside. The regulation of the arteriovenous specification has been subject to intensive investigation in recent years, revealing a complex process requiring the convergence of multiple signalling pathways (summarised in Figure 1.6) [11, 154].

The differing identity of arterial (AEC) and venous EC (VEC) is reflected by their expression of distinct markers, commonly used to distinguish the two populations. Two of the most widely used markers are the ligand EphrinB2 and its corresponding receptor Ephrin receptor B4 (EPHB4), which are found to be specifically expressed in AEC and VEC, respectively[155]. Importantly, binding of EphrinB2 to EPHB4 functions as a repulsive signal necessary to ensure segregation of arteries from veins during vessel formation[156].

Figure 1.6 Molecular regulation of arteriovenous specification.

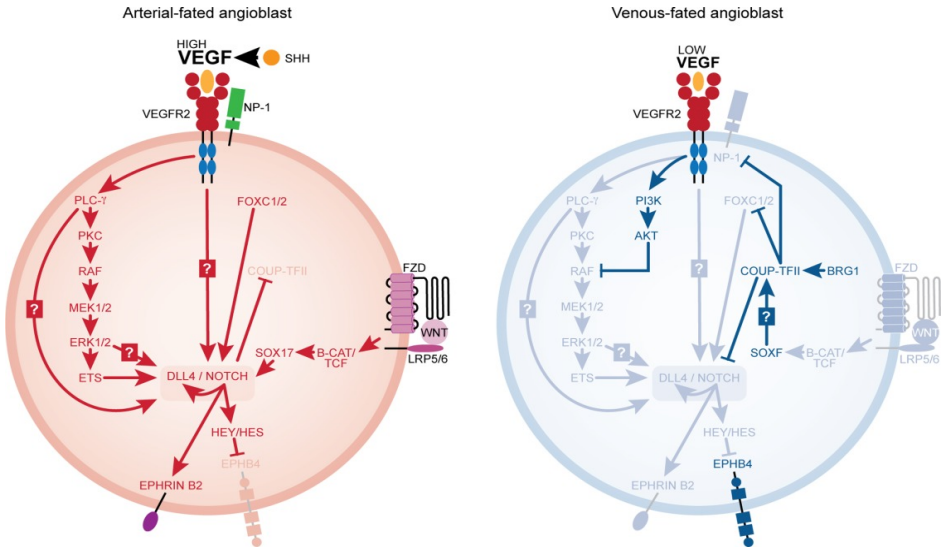


Figure 1.6: Graphical summary of molecular signalling pathways regulating arteriovenous specification in arterial-fated angioblasts (left) or venous-fated angioblasts (right). Figure adapted from Fish et al (2015).

1.4.2. Molecular regulators of arteriovenous specification

Sonic Hedgehog (Shh) and VEGF Signalling

In addition to its essential role during vasculogenesis and angiogenesis, VEGF-A signalling is also required for both arterial and venous EC specification. Whilst VEC specification requires a low VEGF-A concentration, high concentrations of VEGF-A are needed to sufficiently activate downstream Notch signalling to maintain AEC identity[157, 158].

Upstream of VEGF-A signalling, Sonic hedgehog signalling is also required for arterial specification, predominately by endodermal derived Sonic Hedgehog (Shh) inducing expression of VEGF-A from somites[159]. Loss of *shh* in zebrafish results in the absence of EphrinB2 in the dorsal aorta, whilst *shh*

overexpression leads to ectopic expression of AEC markers[159]. Microinjection of *vegfa* mRNA into *shh* mutant embryos is sufficient to rescue arterial specification, demonstrating Shh signalling acts upstream of VEGF signalling[159]. Shh can also induce Notch signalling and arterial specification by bypassing VEGF signalling via the calcitonin receptor-like receptor[160].

In arterial specification, VEGF-A signalling primarily functions to increase notch signalling to induce expression of AEC genes. The most characterised process by which this occurs is via the binding of VEGF-A to FLK1/VEGFR2 to stimulate the downstream activation of the mitogen-activated protein kinase (MAPK) pathway and the subsequent activation of Notch signalling[161]. Although the mechanism by which the MAPK/ERK signalling pathway can induce Notch activation is still unclear, several studies have suggested a role for ETS TF[95, 162].

In contrast to the activation of MAPK/ERK signalling at high VEGF-A concentrations, low concentrations of VEGF-A can repress AEC identity by inducing AKT activation via phosphatidylinositol-3-kinase (PI3K)[161]. PI3K/AKT signalling functions to inhibit AEC identity by antagonising MAPK/ERK signalling via phosphorylation of RAF1 (a critical kinase in the MAPK/ERK pathway) rendering in inactive[161, 163]. Inactivation of RAF1 by PI3K/AKT signalling results in a subsequent reduction in Notch signalling and repression of AEC identity[161]. Hong et al[161], demonstrated in zebrafish that constitutive activation of Akt in zebrafish was sufficient to promote venous specification. A

subsequent study in mouse demonstrated expression of a mutated version of RAF1, resistant to inactivation by AKT, resulted in increased induction of arterial fate *in vivo* and *in vitro*[163]. Although differing concentrations of VEGF-A have been suggested to be responsible for determining the preferential activity of either ERK/MAPK or PI3K/AKT during arteriovenous specification, it remains unknown if additional determining factors are required to modulate this balance[157].

Notch signalling

Notch signalling is widely regarded as the key pathway underlying the transcriptional state of AEC, responsible for driving the expression of AEC markers including DLL4 and Ephrin-B2[162, 164]. Disruption of Notch signalling in both mouse and zebrafish leads to arteriovenous defects with hyperproliferation of the endothelium and a reduction in arterial markers[13, 165, 166]. Furthermore, the combined loss of downstream Notch target genes, *Hey1* and *Hey2* results in arteriovenous defects characterised by a reduction in arterial markers such as Ephrin-B2[167].

Notch signalling requires activation of Notch receptors by transmembrane Notch ligands during cell-cell contact[168]. Notch 1 and 4 are specifically expressed in AEC with *Notch1*$^{-/-}$; *Notch4*$^{-/-}$ mutants exhibiting severe arteriovenous defects[169] . Ligand binding initiates the proteolytic cleavage of the Notch receptor to generate the Notch intracellular domain (NICD). The proteolytic cleavage is mediated by ADAM (A disintegrin and

metalloproteinase) proteases at the cell surface followed by γ -secretase intracellularly[168]. Upon its entry to the nucleus, NICD interacts with RBPJ converting it from a transcriptional repressor to an activator by displacing corepressors with newly recruited coactivators to invoke the transcription of target genes[168].

Although Notch signalling has been reported to induce the expression of arterial markers, the mechanism of exactly how arterial identity is established by Notch signalling remains largely unknown, although recent reports have suggested Notch dependent regulation of the cell cycle may be an important factor (discussed below)[154]. Whilst Notch signalling clearly plays a central role in AEC identity, Notch ligand and arterial marker, *Dll4*, has been found to be expressed prior to the induction of Notch signalling in mouse[20]. This finding was subsequently followed up by Wythe et al[95], demonstrating that Notch signalling is not required for initiating *Dll4* expression. Instead, it was shown that *Dll4* expression is initiated via VEGF/MAPK dependent activation of ETS TF factors including ERG and ETV2[95]. Therefore, it is suggested that Notch signalling is required for the maintenance of arterial identity as opposed to the initial specification[95].

WNT/ β-Catenin signaling

Arteriovenous specification has also been demonstrated to be modulated by WNT/β-catenin signalling with increased signalling promoting arterial fate[170]. Constitutive activation of Wnt/β-catenin signalling results in a phenotype

resembling that caused by constitutive Notch activation[170]. Wnt/β-catenin signalling is induced downstream of VEGF-A dependent activation of FLK1 and leads to the induction of *Dll4* expression via interaction with NICD at the *Dll4* promoter[170, 171]. However endothelial specific knockout of *Ctnnb1* (the gene encoding β-catenin) does not interfere with formation of the dorsal aorta or cardinal vein or the expression of *DLL4*[95]. This suggests that Wnt/β-catenin signalling does not ordinarily have a function in promoting arterial specification during development.

SOXF Transcription Factors

Induction of AEC specification has also been demonstrated to be the dependent on the action of SOXF TF[138, 162]. Sacilotto et al revealed the activation of arterial specific enhancers regulating *Dll4* expression requires binding of both RBPJ/NICD and SOXF TF[162]. Subsequent findings in zebrafish demonstrated the combined loss of *rbpj*, *sox17* and *sox18* was sufficient to ablate expression of AEC markers and prevent the formation of the dorsal aorta[162]. This extreme phenotype was not observed following the individual knockdown of each of these genes, highlighting their combinatorial role in arterial specification[162]. SOXF binding is also required to activate enhancers of *Notch1* to drive expression and promote arterial identity[172]. Unlike *Sox7* and *Sox18* which are expressed uniformly across the vasculature, *Sox17* is predominantly expressed in the AEC[138, 139]. Induced by canonical Wnt signalling, *Sox17* expression has been shown to be required for arterial

specification by the direct binding to the promoters of *Notch1, Notch4 and Dll4*[138].

TGF-β/ BMP signalling

Arteriovenous specification is also regulated by members of the TGF-β superfamily, controlling both venous and arterial fates depending on the combinations of soluble ligand and membrane bound receptor involved[11]. Activin receptor-like kinase 1 (*Alk1*) is highly expressed in AEC and its loss at embryonic or adult stages in mouse results in arteriovenous specification defects including the loss of EphrinB2[173, 174]. BMP10 signalling via ALK1 is promoted by blood flow/shear stress resulting in quiescence of AEC[175, 176]. Despite a clear phenotype in Alk1 deficient mice, the mechanism by which ALK1 signalling induces arterial fate remains largely unknown although this may involve an interaction with Notch signalling and/or cell cycle regulation[177]. Interestingly, the ALK1 coreceptor, ENG, has been shown to repress VEC identity by the suppression of *Nr2f2*[178]. Conversely, Neal et al[179], demonstrated in mouse and zebrafish that BMP signalling via Alk3 can induce expression of venous genes such as *EphB4* and *Nr2f2* via the activation of SMAD1/5.

NR2F2 (nuclear receptor subfamily 2, group F, member 2)

NR2F2 (also known as COUP-TFII) is known to positively regulate venous identity via the repression of Notch signalling[12]. You et al., demonstrated the deletion of *Nr2f2* in EC resulted in embryonic lethality by E12 and the ectopic expression of arterial markers at the expense of reduced venous marker

expression[12]. Conversely, overexpression of *Nr2f2* was found to supress Notch signalling, resulting in depleted expression of AEC markers[12].

Studies in zebrafish have demonstrated expression of *nr2f2* itself its reliant on SOXF TF, Sox7 and Sox18[180]. Therefore, given that *sox7* and *sox18* are expressed widely across the vasculature, the induction of Notch signalling is required in AEC to repress *nr2f2* expression[180]. This provides further evidence to the concept that venous is the default EC state[180, 181]. Additionally, PI3K/AKT signalling, which functions to antagonise VEGR2/MAPK signalling also promotes *Nr2f2* expression in VEC[182]. Chromatin remodelling enzyme BRG1 was also found to induce *Nr2f2* expression[183].

1.4.3. Influence of shear stress in arteriovenous specification

Initially it was believed that the shear stress generated by blood flow played a central role in arterial specification. However, given the earliest markers of arteriovenous specification are expressed prior to the initiation of blood flow[20], this would suggest that shear stress is not required for the initiation of AEV/VEC identity. This was confirmed by Le Noble et al[184], demonstrating that the induction of arterial marker EphrinB2 expression can occur in the absence of blood flow in the chick embryo yolk sac. However, studies using $Nkx^{-/-}$ mouse embryos (which never develop a heartbeat) revealed a reduction of Notch signalling in the dorsal aorta and resultant reduction and increase in AEC and VEC markers, respectively[185]. Similarly, studies using isolated or hPSC derived endothelial cells have shown an upregulation of AEC markers

in response to shear stress[186, 187]. Together this suggests that flow induced shear stress is necessary for the maintenance of the AEC signature, as opposed to initiating the initial specification event. Although the mechanism by which shear stress induces an arterial fate is still unclear, endothelial NOTCH1 has been shown to function as mechanosensor of shear stress in adult arteries[188].

1.4.4. Role of cell cycle regulation of arteriovenous specification

Recent studies have suggested an important role for the regulation of cell cycle to regulate arteriovenous specification. Reduction in EC proliferation has been found to coincide with the onset of arteriovenous specification, indicating regulation of EC proliferation may be required for AEC/VEC specification[189, 190]. Supporting this hypothesis, *in vitro* studies determined that shear stress promoted EC quiescence and AEC specification via GJA4 mediated induction of cell cycle inhibitor P27[189]. Increased *Gja4* expression was determined to be induced by increased Notch signalling, suggesting Notch may mediate AEC specification via cell cycle regulation[189].

Further evidence of a role of cell cycle regulation in arteriovenous specification was provided in a seminal study from Luo et al[190] which further contested the notion that Notch signalling directly induces AEC identity. Ablation of DLL4/Notch signalling was found to induce genes associated with cell cycle progression, including targets of MYC, a known positive regulator of cell cycle

progression[190]. Subsequent use of a mosaic mouse model determined that EC with loss of *Myc* contributed to artery formation at a higher rate compared to cells with high *Myc* levels, despite their reduced proliferation[190]. Furthermore, the reduced ability of cells with deletion of *RBPJ* to contribute to forming arteries was reverted by the additional knockdown of *MYC*[190]. Recently, use of the Fucci2aR mouse model revealed AEC exist in the late G1 phase of the cell cycle whilst VEC exist in the early G1 phase[191]. Pharmacological induction of G1 cell cycle arrest is sufficient to rescue AEC specification defect in the retinal vasculature of postnatal *Gja4* knockout mice[191]. Collectively these studies suggest that although Notch signalling plays an essential function in AEC identity, this may occur indirectly via the regulation of cell cycle progression.

1.4.5. Arteriovenous specification in the human endothelium

Given that most of our current understanding of the regulation of arteriovenous specification has originated from studies using murine of zebrafish models, the relevance to human biology remains largely undetermined. Some insight into the genes involved in this process during human endothelial development have been determined by the identification of genes with a role in the formation of arteriovenous malformations (AVMs)[11, 192]. AVMs are vascular developmental abnormalities that arise when arterial blood flow is directly connected to veins or venules without passing through an intermediatory capillary network[192]. Abnormalities during arteriovenous specification are thought, at least in part, to be a contributing factor to AVM formation[192]. This

includes the identification of mutations in the *ALK1* gene in individuals with hereditary haemorrhagic telangiectasia type 2, an autosomal dominant disorder characterised by the presence of AVMs[193]. This causal role for *ALK1* in human AVM formation along with the identified function in murine EC development would therefore suggest a conserved role in arteriovenous specification[173, 174]. Despite these studies providing a limited insight into the potential regulators of arteriovenous specification, further characterisation of this process in human EC has been largely thwarted by a lack of relevant omics datasets obtained from human embryonic and fetal tissue. The increasing popularity of single cell omics technologies, with the capability of profiling dynamical cellular processes in developing human tissues, gives further opportunities to provide needed human relevance to our understanding of the regulation of arteriovenous specification[194-196].

Despite the widespread use of primary EC isolated from either arterial or venous vascular beds in routine cell culture, their usefulness as a model for studying the regulation of arteriovenous identity has been disputed[197]. Notably, Richardson et al[197] identified a decline in venous identity in HUVECs as soon as 48 hours in culture with a further decline until 12 days. The development of increasingly refined hPSC differentiation protocols to yield EC with either arterial or venous identities may provide a more appropriate developmental model of human arteriovenous specification[198, 199].

1.5. Lymphatic endothelial cell development

In contrast to the blood vasculature, the lymphatic vasculature is an open circulatory system responsible for maintaining tissue fluid homeostasis as well as roles in immune surveillance and lipid absorption[2]. Lymphatic vessels consist of a single layer of specialised lymphatic endothelial cells (LEC) and lack the cover of perivascular mural cells such as pericytes and SMC. LEC can be distinguished from their blood EC counterparts by their expression of distinct markers including *VEGFR3* (FLT4), *LYVE1*, *PDPN*, and *PROX1*[200-203]. Despite continued controversy over the developmental origin of LEC, the current consensus is that LEC can originate via two distinct mechanisms, either from a venous origin via lymphangiogenesis[204, 205] or from non-venous origins via lymphvasculogenesis[206, 207].

Venous derived LEC first emerge from the dorsolateral region of the cardinal vein by E9.5 during murine development[203, 205, 208]. By E10.5 aggregated LEC begin to bud out ventrally from the cardinal vein giving rise to lymph sacs[204]. Subsequent expansion of lymph sacs via lymphangiogenesis generates a primitive network of maturing, interconnected lymphatic vessels expanding into embryonic tissues[204, 209]. As mentioned previously, lineage tracing using the *Pax3-Cre* transgenic mouse model revealed LEC are predominantly derived from the paraxial mesoderm as opposed to the lateral plate mesoderm[49]. This suggests that as subset of EC in the cardinal vein are already intrinsically programmed to form LEC via lymphangiogenesis[49].

Expression of the transcription factor, *Prox1*, within the VEC of the cardinal vein is sufficient to induce lymphatic EC identity[203, 208]. This was demonstrated in *Prox1*$^{-/-}$ mutant embryos which lack a lymphatic vasculature with the absence of LEC budding from the cardinal vein[208]. *Prox1* expression itself is reliant on expression of both *NR2F2* and *SOX18* in the cardinal vein, which together which bind to *PROX1* regulatory elements to initiate expression[144, 210]. In mouse, expression of *Prox1* in budding LEC induces the expression of LEC specific markers such as *Vegfr3*, *Lyve1*, and *Pdpn*[211, 212]. The exit of PROX1+ LEC from the cardinal vein is induced by VEGF-C signalling via its cognate receptor VEGFR3. This was demonstrated in *Vegfr3*$^{-/-}$ mice in which LEC specification occurs but fail to migrate out from the cardinal vein[204]. VEGF-C dependent activation of VEGFR3 is then required to form a positive feedback loop to sustain *Prox1* expression in newly formed LEC[212].

The development of the murine lineage tracing tools revealed LEC can also be derived from non-venous origins via lymphvasculogenesis[206, 207, 213, 214]. The contribution of venous and non-venous derived LEC has been shown to vary between and within the vascular beds of individual tissues. For instance, in addition to the venous derived lymphatic vessels, endothelial specific knockout of *Prox1* demonstrated a significant non-venous contribution to the dermal lymphatic vasculature[207]. The cardiac lymphatic vasculature also consists of lymphatic vessels from venous and non-venous origin[206, 213, 214]. Whilst most of the cardiac lymphatic vasculature was found to be derived from paraxial mesoderm derived Tie2+ EC in the cardinal vein, Klotz et al[206]

identified an additional contribution from the haemogenic endothelium of the yolk sac. Two subsequent studies later revealed *Isl1*+ second heart field progenitors contribute to the ventral lymphatic vasculature of the heart, providing further evidence of a heterogenous origin for the cardiac lymphatic vasculature[213, 214].

1.6 Human pluripotent stem cell models of endothelial development.

1.6.1. Human pluripotent stem cells.

Human pluripotent stem cells (hPSC) such as human embryonic stem cells (hESC) or human induced pluripotent stem cells (hiPSC) are classified as such due to their capability to differentiate into each of the three germ layers (ectoderm, mesoderm, and endoderm)[215, 216]. The differentiation potency of hPSC is only lesser to that of totipotent stem cells which have the additional capability of generating extraembryonic lineages[216].

hESC are derived from the inner cell mass of a trophoblast stage human embryo[217]. The first study describing the derivation and long-term culture of hESC came from Thompson and colleagues in 1998[218] which also demonstrated that hESC transplanted into an immunodeficient mouse model form teratomas containing cells of each of the embryonic germ layers. The capability of hESC to differentiate to somatic lineages *in vitro* in response to specific stimuli was later described by Reubinoff et al[219], paving the way for the use of hESC as a human model for cell development. Since the generation

of the earliest hESC lines, substantial progress has been made in the methods used for their derivation and culture with the latest approaches utilising xeno-free derivation and culture techniques, thus reducing experimental variability and enabling the translation of hESC-derived cell populations for therapeutic purposes[217].

In a landmark study in 2006, Takahashi and Yamanaka reported the reprograming of murine fibroblasts by the overexpression of just four factors to generate an induced pluripotent stem cell population[220]. The gene expression profile and differentiation capability of this population was found to be comparable to that of embryonic stem cells[220]. These four factors - Oct4, Sox2, Klf4, and c-Myc are now commonly referred to as the Yamanaka factors[221]. The reprogramming of human fibroblasts using a similar method to yield hiPSC was reported by two independent studies the following year[222, 223]. Along with the latest developments in gene editing methods, the capability to reprogram primary patient derived somatic cells to generate hiPSC has provided substantial opportunities for disease modelling, personalised medicine, and drug screening[221]. In addition, refinements in the methods used to generate hiPSC have circumvented the need for integration dependant viral vectors with more recent approaches utilising non-integrating vectors such as episomal DNA plasmids[224] or Sendai virus[225] to deliver reprogramming factors, thus avoiding the risk of insertional mutagenesis.

1.6.2. Methods of hPSC-derived endothelial cell differentiation.

Early methods used to differentiate hPSC to EC were reliant on the use of murine feeder layers to provide the cues needed to promote differentiation[226-228]. For instance, Kauffman et al[227] demonstrated the emergence of $CD34^+$ hematoendothelial precursors following culture of hESC on either a murine stromal cell layer or on murine yolk sac EC in the absence of FGF2. Subsequent characterisation of isolated $CD34^+$ cells demonstrated their capacity to differentiate into either EC or SMC depending on the culture conditions used[228]. Similar results were also obtained by Sone and colleagues with the generation of a hESC derived $TRA1\text{-}60^-/KDR^+$ population following 8 days of culture on a murine bone marrow stromal cell feeder layer[226]. Although these protocols provided the first evidence that hESC could differentiate into EC *in vitro*, their use for further applications was severely limited by low differentiation efficiency and inability to determine the factors needed to promote EC differentiation.

Embryoid body (EB) based differentiation methods have been extensively used to study the differentiation of human cell types including EC[229]. EBs are small aggregates of hPSC which undergo differentiation to yield cells of each of the three germ layers[230]. Although EBs undergoing spontaneous differentiation were found to give rise to hPSC-EC, the proportion was found to be as low as 2% when quantified by flow cytometric analysis[231]. However, subsequent studies demonstrated the efficiency of EC differentiation to be greatly enhanced by the addition of different combinations of cytokines to

promote mesodermal or EC commitment[232-234]. This included the use of BMP4 to promote mesoderm commitment prior to inducing endothelial specification, causing in an increased proportion of $CD144^+$ EC by day 7 of differentiation[232]. Similarly, the addition of VEGF-A throughout EB differentiation was found to result in a 4-5-fold change increase in the yield of $CD31^+/CD144^+$ hESC-EC[233]. The differentiation of hESC to EC has also been described to occur via an intermediate $KDR^{low}/C\text{-}KIT^-$ cardiac mesoderm population isolated from EBs following culture in media containing BMP4, FGF2, DKK1, and activin A[235]. Despite EC differentiation in EBs more closely recapitulating the 3D *in vivo* environment, EB-based differentiation protocols are limited by several factors including variability in EB generation, complex culture conditions, and the requirement to dissociate EBs to generate a single cell suspension for further applications. Furthermore, the efficiency of hPSC-EC differentiation in EBs is generally lower compared to the more widely used monolayer differentiation methods.

In recent years there has been an increase in the number of monolayer hPSC-EC differentiation protocols, largely replacing lower efficient EB and feeder layer dependant methods. Such methods begin with seeding hPSC on a supportive matrix such as matrigel[236-240], fibronectin[241], vitronectin[198], or laminin[242]. Matrigel is derived from the basement membrane isolated from Engelbreth-Holm-Swarm mouse tumours and consists predominantly of laminin and collagen[243]. However, despite being a cost-effective option, the animal origin and non-defined nature of matrigel make it unsuitable for clinical

translation and prone to causing variable differentiation efficiency[244]. Therefore, an increasing number of monolayer EC differentiation protocols are opting to use defined recombinant human extracellular matrix proteins such as laminin[242], fibronectin[241], or vitronectin[198]. The choice of matrix has also been shown to influence hPSC-EC differentiation with the more efficient differentiation observed using plates coated with the E8 fragment of laminin-411 compared to those coated with the E8 fragment of laminin-511[242].

In addition to the inclusion of cytokines and a suitable matrix, the addition of specific small molecules during differentiation has been shown to enhance endothelial differentiation. This includes the GSK3 inhibitor, CHIR99021, which was found to promote commitment to a brachyury expressing mesoderm prior to induction to an EC fate[236, 238, 239]. The inclusion of CHIR99021 is commonly used during the first stage of many EC differentiation protocols in combination with cytokines such as BMP4 and activin A to promote efficient commitment to a mesoderm prior to promoting an endothelial fate[236, 240, 241]. Leveraging from findings initially conducted using EB-based methods[245], the presence of a TGFβ inhibitor, SB431542, was found to minimise differentiation to mural cell types, thus increasing the yield of hPSC-EC[198]. Patsch et al[236] also demonstrated enhanced hPSC-EC differentiation following the addition of forskolin to promote the cyclic AMP mediated activation of protein kinase A during the commitment of mesoderm to an EC fate.

The resultant identity of the hPSC-EC population produced during differentiation has been found to be dependent on the mesodermal intermediate population from which it is specified. Palpant et al[246] demonstrated using high concentrations of activin A alongside low levels of BMP4 results in the generation of a cardiac mesoderm population capable of producing both cardiomyocytes and hPSC-EC with a cardiac EC like identity. Conversely, low levels activin A with a higher concentration of BMP4 produces a haemogenic mesoderm population with a propensity to form haemogenic EC and blood cells[246]. Similar findings were reported in a study from Giacomelli et al[237] which described the generation of a cardiac mesoderm population capable of specifying into either cardiomyocytes or hPSC-EC with a transcriptional signature comparable to that of primary cardiac EC. Further work is therefore required to evaluate how the inclusion of different cytokines or small molecules may alter the path of differentiation and the resultant identity of the hPSC-EC population.

Despite monolayer protocols offering a more efficient and reproducible option for hPSC-EC differentiation these methods typically still include the co-emergence of a minor population on non-EC, producing a heterogenous population of differentiated cells. However, few studies to date have looked to characterise the extent of cellular heterogeneity during hPSC-EC differentiation. Patsch et al[236] identified a population of non-EC produced during hPSC-EC differentiation characterised by expression of the SMC marker alpha smooth muscle actin. The subsequent replacement of factors

used to promote endothelial commitment with factors promoting SMC differentiation was found to bias differentiation towards a SMC fate at the expense of EC yield[236]. Similarly, flow cytometric of an hPSC-EC differentiation protocol adapted from Patsch and colleagues demonstrated the cell surface expression of mural cell markers such as CD146 and CD73 in non hESC-EC[241]. In 2018, Paik et al used scRNA-seq to characterise the cellular heterogeneity and associated transcriptional signatures of hPSC-EC[247]. However, clustering and visualisation of data revealed less than 5% of the total sequenced cells represented hPSC-EC with the majority of remaining clusters having a cardiomyocyte or hepatic identity[247]. Along with this low yield of hPSC-EC, this study failed to analyse cells at multiple timepoints during the protocol, thus preventing the characterisation of cellular heterogeneity and the underlying transcriptional cues throughout differentiation. Therefore, future studies using more efficient hPSC-EC differentiation protocols are required to characterise the cellular heterogeneity and transcriptional signatures of endothelial differentiation. This information is likely to be important for aiding further refinement of hPSC-EC differentiation protocols, as well as providing human relevance to our understanding of the transcriptional regulation of EC development.

1.7. Endothelial heterogeneity and specification

1.7.1. Functional importance of organ specific endothelial heterogeneity

In addition to endothelial identity differing by arteriovenous identity, substantial heterogeneity also exists between the vascular beds of individual organs. Early evidence for organ-specific endothelial heterogeneity was first provided using phage display libraries, demonstrating the non-uniform distribution of bound peptides between the vasculature different organs[248]. Despite recent technological advances made in the ability to characterise EC heterogeneity, the molecular regulation of the processes underpinning further EC specification remain largely uncharacterised.

Further specification of the endothelium is essential to providing the vasculature with the specific functions to fulfil the individual requirements different organs. For instance, the endothelium of the coronary vasculature highly expresses *CD36*, a fatty acid transporter required for fatty acid - dependent cardiomyocyte metabolism[249]. Further specification of EC is also required to modulate vascular permeability. This is especially important in the blood brain barrier (BBB) where endothelial cells form a physical barrier to restrict the movement of toxins, pathogens, and leukocytes into the brain by forming tight and adherens junctions to limit paracellular diffusion[250]. BBB EC also selectively express multiple transport proteins required to actively transport specific molecules across the otherwise impermeable barrier[250]. Endothelial heterogeneity can also be required within the same organ to

execute different functions. This was shown to be the case in the bone marrow vasculature where less permeable arterial vessels maintain the quiescent haemopoietic stem cell (HSC) niche by limiting exposure to reactive oxygen species (ROS)[251]. Conversely, the high permeability of sinusoidal capillaries results in increased levels of ROS promoting HSC mobilisation and differentiation[251].

Organ specific EC heterogeneity also plays an important role in regulating organ formation, homoeostasis, and regeneration via the release of tissue specific paracrine factors, termed angiocrine factors[252-254]. Angiocrine signalling fulfils an essential function in the developing lung by regulating patterning of the expanding airway epithelium via the selective expression of *FGF10*, a key positive branching mediator[253]. Angiocrine signalling is also essential during cardiac development, including to promote the compaction of the trabeculated myocardium in the ventricular wall[254]. This includes the recent identification of COL15A1 as an important angiocrine factor which is downregulated in a murine model of left ventricular non-compaction[252]. Release of COL15A1 from the newly formed coronary vascular EC promotes cardiomyocyte proliferation enabling compaction of the myocardium[252].

1.7.2. Regulation of further endothelial specification

In contrast to the transcriptional control of endothelial differentiation and arteriovenous specification, comparatively little is known about the mechanisms regulating further organ specific EC specification. These

mechanisms can either be promoted by the tissue specific environmental cues or by intrinsic cell factors[3].

The ability of the tissue microenvironment to guide endothelial specialisation was first demonstrated in isogenic transplantation experiments in which the tissue specific expression of von Willebrand factor (VWF) was shown to be induced in auricular vessels only following exposure to a myocardial environment[255]. More recent co-culture studies using hPSC derived EC have also demonstrated the ability of the microenvironment to induce the activation of organ specific EC transcriptional signatures[194, 237, 256]. Microenvironmental cues include exposure of EC to one or more specific growth factors. For example, in the heart and skeletal muscle VEGF-B is needed to prime the endothelium for fatty acid uptake[257]. Similarly, hepatic epithelial derived VEGF-A is required to maintain the identity of liver sinusoidal EC[258]. Biophysical cues such as shear stress also have a role in modulating EC heterogeneity. For instance, in the aorta high levels of shear stress are required to induce expression of endothelial nitric-oxide synthase (eNOS), a known atheroprotective gene[259].

In addition to cues from the surrounding tissue environment, epigenetic regulation of gene expression can function as an intrinsic mechanism for mediating endothelial heterogeneity[260]. Marcu et al[261] demonstrated that that EC isolated from human fetal lung, heart, liver, and kidney retained a proportion of their heterogenous expression profile following prolonged *in vitro*

culture, thus suggesting epigenetic modifications contribute, at least in part, to endothelial heterogeneity. Furthermore, You et al[262] revealed a role for histone acetylation in mediating the expression of transport protein, *MDR1*, in EC of the BBB. Despite the overall regulation of EC heterogeneity likely relying on a combination of intrinsic and extrinsic cues, further studies are required to determine the relative contribution of either mechanism in establishing tissue specific EC heterogeneity.

1.7.3. Characterising EC heterogeneity using single cell RNA sequencing

In recent years, technological advancements in single-cell RNA sequencing (scRNA-seq) have made it possible to map the transcriptional landscape of a heterogenous cellular population at single cell resolution. This has allowed for the identification of novel cell types/states and improved understanding of the transcriptional program underpinning cellular heterogeneity, including during dynamic processes such as differentiation.

scRNA-seq protocols can be divided into plate based or microfluidic-based approaches with each method having strengths and weaknesses depending on the research question being addressed[263]. Microfluidic methods such as the 10X Chromium platform offer high throughput allowing the transcriptome of thousands of cells to profiled in a single run[14]. However, high cell throughput comes at the expense of required sequencing depth, resulting in a reduction in library complexity[14, 263]. Conversely, plate-based approaches such as

Smart-Seq2, have lower cell throughput but greater library complexity, typically capturing more genes per cell than microfluidic approaches[14, 263].

Given the highly heterogenous nature of the endothelium within individual vascular beds and between different organs, scRNA-seq has been instrumental in providing novel insight into EC heterogeneity and its underlying transcriptional signatures. This was demonstrated in a study from Vanlandewijck et al[264] which used scRNA-seq to map the transcriptional landscape of EC and associated mural cells from the murine brain. Following their identification in the dataset, ordering of EC based on their transcriptional similarity was found to recapitulate the arteriovenous axis of the brain vasculature, with venous and arterial marker expressing EC localised near the opposing ends of the scale[264]. This revealed the transcriptional state of brain EC exists as a continuum along the arteriovenous axis, with gradual changes in gene expression as opposed to distinct cell states[264].

High throughput microfluidic scRNA-seq methods have also been used to in large cell atlas studies which aim to map the total cellular heterogeneity within an entire organ. The recent publication of a cell atlas of the human adult heart composed of over 480,000 cells from 14 different donors has provided a significant advancement in our understanding of cellular heterogeneity in the human heart[17]. Within this heart cell atlas, 10 distinct endothelial clusters were identified based on expression of pan-endothelial markers such as *CDH5* and *PECAM1*[17]. This included 3 capillary populations defined by expression of

capillary markers *RGCC* and *CA4*. Arterial and venous EC populations were also identified by expression of *EFNB2* and *NR2F2*, respectively. A distinct EC cluster expressing *NPR3* was annotated as representing an endocardial population[17]. In addition, analysis of the expression of ligand and receptor pairs in EC and mural cells suggested different cell-cell interactions occur between EC and SMC depending on the arteriovenous identity of EC[17].

Endothelial heterogeneity within the murine kidney vasculature has also been explored using scRNA-seq[265]. This revealed distinct transcriptional signatures for renal EC depending on their localisation in either cortical, medullary, and glomerular compartments. Additional heterogeneity was identified within individual compartments largely owing to changing identity along the arteriovenous axis[265]. Furthermore, gene regulatory network analysis (GRN) was applied to try understanding the transcriptional regulation underlying intercompartmental EC heterogeneity. This identified a putative role for *Gata5* and *Foxp1* in determining glomerular and cortical EC identity, respectively[265].

Despite an increasing number of studies using scRNA-seq to characterise EC heterogeneity within individual organs, studies using scRNA-seq to compare EC from different organs to identify organ-specific transcriptional signatures remained limited. However, a recent study from Kalucka et al[16] addressed this knowledge gap by performing scRNA-seq using EC freshly isolated in parallel from 11 different mouse tissues (Figure 1.7). Comparison of EC from different tissues revealed the existence of some tissue specific EC markers such as

Tmem100, Igfbp5, *Pglyrp1* for lung, kidney, and brain EC, respectively[16]. Subsequent GRN analysis identified the enrichment of transcription factor networks within organ specific EC. This included *Foxf2* in brain EC and *Hoxa9* in the kidney[16]. It would be interesting in future *in vitro* studies to explore if forced expression of one or more of these TF was sufficient to invoke organ specific associated gene expression. Furthermore, comparison of EC from different vascular beds (capillary, arterial, venous, and lymphatic) across different organs identified a lack of conserved capillary markers, suggesting a higher degree of inter-organ capillary EC heterogeneity between individual organs[16].

Figure 1.7 scRNA-seq atlas of murine endothelial cell heterogeneity.

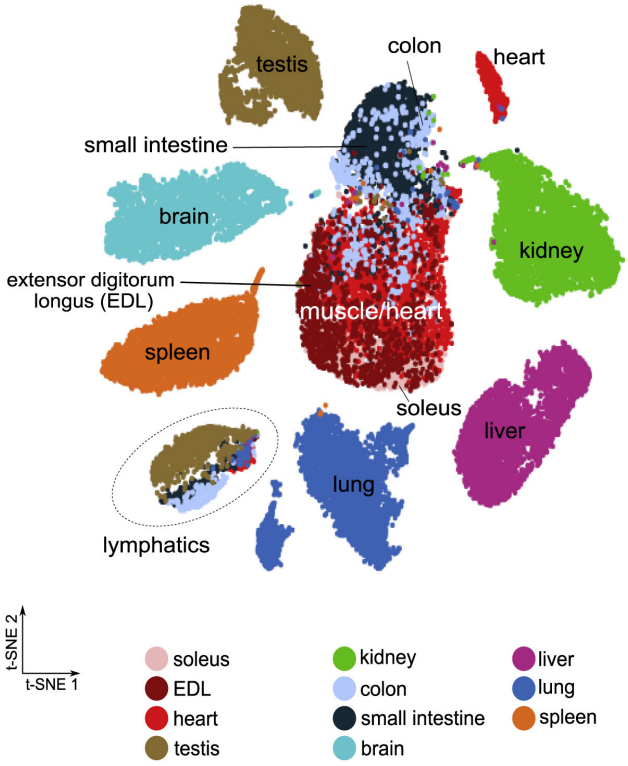


Figure 1.7: tSNE visualisation of scRNA-seq data acquired from EC isolated from 11 different mouse tissues. Figure adapted from Kalucka et al (2020).

In addition to scRNA-seq, advancements in other single cell omics methods have been made in recent years[266]. This includes methods used to characterise the epigenomic status of individual cells, including single cell methylation profiling and single cell ATAC-seq (assay for transposase-accessible chromatin using sequencing) for assessing chromatin accessibility[266]. Given the previously discussed evidence that epigenetic modifications may contribute to specifying organ specific EC identity, future studies characterising the epigenome at single cell resolution are likely to provide further insight into the regulation of EC specification[261, 262].

1.8 Hypotheses and Aims

Despite increasing knowledge of the transcriptional regulators controlling endothelial differentiation and specification, these findings are predominantly obtained from murine models. Therefore, given that the relevance to human development remains largely unknown, the following hypotheses were proposed:

- The differentiation of hESC to endothelial cells can be used as a model to study the transcriptional control of human endothelial development.
- Transcriptional changes occurring during the differentiation of hESC-derived mesoderm to endothelial cells will largely be conserved to those identified in studies of murine vasculogenesis.
- Endothelial cells derived from hESC are transcriptionally similar to human adult and fetal endothelial cells.
- Endothelial heterogeneity in the developing human heart is attributable to distinct programs of gene regulation controlled by differential transcription factor expression.
- Dynamical cellular changes in the human fetal heart, such as arterial specification, are comparable to those previously identified in mouse, aiding the identification of conserved regulators of endothelial fate.

To address these hypotheses, the work contained within this thesis aimed to apply single-cell RNA sequencing to characterise the transcriptional landscape of both hESC derived endothelial cells and cardiac endothelial cells isolated from the human fetal heart.

Chapter 2 - Transcriptional dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA-Seq.

2.1 Chapter Introduction

This aim of this project was to use single cell RNA sequencing to characterise the transcriptional changes which occur during the differentiation of hESC to endothelial cells. The main hESC differentiation protocol used in these experiments had previously been established in Baker lab, although the underlying transcriptional signatures and level of cellular heterogeneity were largely unknown.

A comprehensive characterisation of the transcriptional changes during hESC-EC differentiation was required to provide human relevance to existing understanding of the transcriptional regulation of EC development, which has largely been determined using murine or zebrafish models. Profiling the transcriptional state of hESC-EC was also necessary to compare to that of mature EC types from different vascular beds.

Finally, given the potential for the hESC-EC differentiation protocol used in this chapter to be used for cell therapy purposes, the application of scRNA-seq was used to identify the level of cellular heterogeneity at multiple timepoints. Differential gene expression analysis of angiogenic related genes was also used to provide insight into the potential mechanism by which the injection of the differentiated cell product results in the increased perfusion of ischaemic tissue.

2.2 Title Page

Transcriptional dynamics of pluripotent stem cell derived endothelial cell differentiation revealed by single cell RNA-Seq.

McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BEP, Mountford JC, Caudrillier A, Henderson NC, Ponting CP, Baker AH. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. Eur Heart J. 2020 Mar 1;41(9):1024-1036. doi: 10.1093/eurheartj/ehz351.

Author contribution statement

I.R.M. performed all hESC-EC differentiation experiments and prepared cells for processing on the 10X scRNA-seq platform. 10X library preparation was performed by R.D. and B.E.P.H. Primary culture of EC for scRNA-seq was performed by I.R.M., F.C., and A.C. Validation of findings by RT-qPCR was performed by I.R.M. Analysis of scRNA-seq data was performed by I.R.M. in conjunction with R.S.T. I.R.M. was responsible for writing the manuscript with all authors involved in discussion of results and contributing to the production of the final version of the manuscript.

2.3 Abstract

Aims: Pluripotent stem cell-derived endothelial cell products possess therapeutic potential in ischemic vascular disease. However, the factors that drive endothelial differentiation from pluripotency and cellular specification are largely unknown. The aims of this study were to use single-cell RNA sequencing (scRNA-seq) to map the transcriptional landscape and cellular dynamics of directed differentiation of human embryonic stem cell-derived endothelial cells (hESC-EC), and to compare these cells to mature endothelial cells from diverse vascular beds.

Methods and Results: A highly efficient directed 8-day differentiation protocol was used to generate a hESC-derived endothelial cell product (hESC-ECP), in which 66% of cells co-expressed CD31 and CD144. We observed largely homogeneous hESC and mesodermal populations at days 0 and 4 respectively, followed by a rapid emergence of distinct endothelial and mesenchymal populations. Pseudotime trajectory identified transcriptional signatures of endothelial commitment and maturation during the differentiation process. Concordance in transcriptional signatures was verified by scRNA-seq analysis using both a second hESC line RC11, and an alternative hESC-EC differentiation protocol. In total 105,727 cells were subjected to scRNA-seq analysis. Global transcriptional comparison revealed a transcriptional architecture of hESC-EC that differs from freshly isolated and cultured human endothelial cells and from organ-specific endothelial cells.

Conclusions: A transcriptional bifurcation into endothelial and mesenchymal lineages was identified, as well as novel transcriptional signatures underpinning commitment and maturation. The transcriptional architecture of hESC-ECP was distinct from mature and fetal human EC.

Keywords: Endothelial differentiation, embryonic stem cell, single-cell RNA-sequencing, pseudotime, RNA velocity, vascular regeneration.

2.4 Translational Perspective

Ischaemic conditions contribute substantially to global mortality and morbidity. Human embryonic stem cell derived endothelial cells (hESC-EC) offer a promising strategy to promote therapeutic angiogenesis in tissue ischaemia. Our GMP compliant protocol for the generation of hESC-EC promotes therapeutic angiogenesis *in vivo*. However, knowledge of the transcriptional dynamics of hESC-EC differentiation is currently limited. Here, we apply scRNA-seq to two hESC differentiation protocols. We characterise the transcriptional dynamics of hESC-EC, important for informing how to drive controlled angiogenesis, endothelial cell maturation, and specification. We also evaluate the cellular heterogeneity at each stage using a hESC-EC differentiation protocol that is relevant to clinical cell therapy.

2.5 Introduction

Cardiovascular disease (CVD) remains the most common cause of mortality globally despite improved pharmacological and interventional options[267]. For example, current interventions for critical limb ischaemia (CLI) are unsuccessful, with 43% of patients requiring limb amputation within five years[268]. Similarly, for myocardial infarction (MI) 25% of patients still progress to heart failure owing to the overburdening of the remaining viable myocardium[269].

Consequently, using stem cell therapies to promote vascular regeneration is a promising strategy for treating ischaemic disease. However, despite studies yielding positive results in small animal models of ischaemia, the efficacy of using such cells in human trials has been disappointing, providing limited benefit[270].

While many protocols have tested autologous cells, a number of studies have recently assessed the use of pluripotent cells in relevant animal models[271] and more recently clinically[272]. As such, derivation and production of therapeutic cells from pluripotent sources is required using an efficient and GMP (good manufacturing practice) compliant protocol. Transplantation of human pluripotent stem cell derived endothelial cells (hPSC-EC) has been demonstrated to promote therapeutic angiogenesis in a range of small animal models of ischaemia[273, 274]. However, differentiation protocols are limited by

their low efficiency resulting in substantial and uncharacterised heterogeneity within the final cell population[247].

We recently reported a rapid and efficient protocol[273] for production of a regenerative human embryonic stem cell-derived endothelial cell product (hESC-ECP), adapted from the protocol of Patsch et al[236]. The protocol consistently yields CD31+/CD144+ cells comprising ≈60% of the population, with the remaining cells expressing pericyte and mesenchymal markers and, crucially, exhibiting no residual pluripotency[273]. When injected into ischaemic muscle, a robust pro-angiogenic response was evident across a range of mouse models [273]. Despite the expression of mature endothelial cell (EC) markers by cells from the hESC-ECP, it is likely that these cells remain transcriptionally immature and are non-specified further into arterial, venous, or lymphatic lineages[273].

Comprehensive characterisation of the transcriptional dynamics throughout differentiation is a crucial step towards understanding the key molecular regulators and mechanisms governing commitment and maturation of EC. Additionally, defining the temporal changes in pluripotent and angiogenic related genes may provide insight into the safety and functional efficacy of these cells. Recent studies of cardiac and liver differentiation have demonstrated that longitudinal application of single-cell RNA sequencing (scRNA-seq) permits tracking of transcriptional changes throughout differentiation at single cell resolution[194, 275].

A recent study used scRNA-seq to attempt to understand the directed differentiation of induced pluripotent stem cells derived EC, yet the final cell population was highly heterogeneous, consisting of fewer than 10% EC and co-populated by cardiac-like and hepatic-like cells[247]. Development of droplet based scRNA-seq platforms has enabled the sequencing of thousands of cells in parallel, and the rigorous unbiased modelling of cellular heterogeneity[14]. Consequently, application of these platforms for characterising hESC-EC has clear advantages over previous studies that surveyed fewer than 100 cells[242].

In this study we sought to use scRNA-seq to dissect the cellular heterogeneity and transcriptional dynamics of hESC-EC differentiation and maturation, utilising alternative hESC lines and two different differentiation protocols. Maturity and specification of hESC-EC following their initial commitment were assessed throughout the process by comparison to foetal and adult endothelial cells from various sources. We also aim to evaluate the clinical suitability of our hESC-ECP by comparison to the products of an alternative differentiation.

2.6 Methods

hESC-EC differentiation

H9 and RC11 hESC lines were differentiated to hESC-ECP using our previously reported protocol[273]. Human ESC lines were used in accordance with the UK Stem Cell Bank Steering Committee guidelines (Project Approvals SCS11-51 and SCSC17-26). Briefly, hESC were plated on a fibronectin matrix at day 0 (d0). At d1, lateral mesoderm was induced with GSK3 inhibitor (7μM) and BMP-4 (25ng/ml) added to N2B27/Neurobasal/DMEM:F12 media. This was followed by endothelial cell induction at d4 with Forskolin (2μM) and VEGF (200ng/ml) in StemPro34 media. In the final step, cells were replated and cultured until d8 or d12 with human AB serum (1%) and VEGF (50ng/ml) in EGM-2 media. Cells were taken for scRNA-sequencing at d0 (H9 only), d4, d6, d8, and d12 (H9 only). In the alternative differentiation protocol H9 hESC were differentiated to hESC-EC as described by Zhang et al[198]. In summary, hESC were plated as single cells on a vitronectin matrix and cultured in E8 media supplemented with BMP4 (5 ng/mL), Activin A (25 ng/mL), and CHIR99021(1 μM). From d2-6, media was changed to E5 media supplemented with FGF (100 ng/mL), VEGF (50 ng/mL), SB431542 (10 μM), RESV (5 μM), and L690 (10 μM). Cells were harvested for scRNA-seq at d2, d4, and d6.

scRNA-seq library construction

Following determination of their number and viability, dissociated cells were loaded into the 10X Chromium Controller. Library construction was performed

using the Single Cell 3' Reagent Kit. To summarise, in reaction vesicles (gel beads in emulsion, GEMs), cells were lysed and barcoded oligonucleotides reverse transcribed before clean-up and cDNA amplification. The Chromium Single-Cell 3' Library Kit was then used to generate indexed sequencing libraries. Library sequencing was performed using either Illumina HiSeq 4000 or Illumina NovaSeq 6000 platforms. Recovery rates ranged from between 3,778 and 10,053 cells, with the number of reads per cell ranging between 35,630 and 119,015 (Supplementary Table 2.2).

Single cell RNA sequencing data analysis

Raw reads were processed using the 10X CellRanger pipeline to align these to the reference transcriptome (GRCh38) and to generate gene-cell count matrices. Initial QC and clustering was performed with the aid of Seurat version 2.2.1.[276] In an attempt to remove dead or falsely identified cells, as well as doublets, cells either expressing fewer than 250 genes, having a unique molecular identifier (UMI) count greater than 10,000, or having greater than 15% of reads mapping to mitochondrial genes were excluded from further analysis. Following data normalisation and scaling to remove unwanted sources of variation[276] principal component analysis (PCA) was conducted using genes with highly variable expression. Seurat graph-based clustering[276] was then applied followed by t-distributed stochastic neighbour embedding (tSNE) allowing the visualisation of identified clusters in tSNE plots. Differential gene expression (DGE) analysis was then used to identify significantly differentially expressed genes within each cluster using the Wilcoxon test for

significance (adjusted p value <0.05) and a $\log_e$(FC) greater than 0.25. Signature scores were calculated by taking the mean of the scaled and centred expression values across multiple signature genes (Supplementary Methods).

Pseudotime trajectories of differentiation were generated using the Monocle R Package v.2.6.4[277]. Pseudotime analysis proceeds on the basis that cells undergo biological processes in an asynchronous manner, and thus that cells can be ordered along a calculated trajectory to infer the transcriptional changes throughout the process. Genes that were most differentially expressed ($q<10^{-40}$) in identified clusters were used to assign pseudotime values to individual cells. Differentially expressed genes across pseudotime were identified using the differentialGeneTest command in Monocle[277]. RNA velocity of individual cells was calculated using the Velocyto R package (V0.6.0)[278]. RNA velocity analysis determined the fraction of spliced-to-unspliced reads to predict the future transcriptional state of individual cells.

All further experimental and analysis details are included in the Supplementary Material.

2.7 Results

Longitudinal scRNA-seq analysis reveals the emergence of committed endothelial and mesenchymal lineages from a homogeneous mesodermal population

Human embryonic stem cells (H9) underwent an 8-day differentiation protocol[273] via a mesodermal stage to form a hESC-EC product (Figure 2.1A). At day 8, two-thirds (65.7%) of the cells were $CD31^{+}/CD144^{+}$ with fewer than 0.05% co-expressing pluripotent markers SSEA4/TRA-181 as shown by flow cytometry (Figure 2.1B). Single-cell transcriptomics data was obtained from 21,369 cells harvested at four time points across the differentiation protocol using droplet-based microfluidic technology (Figure 2.1A).

Cellular heterogeneity was characterised at each time point using PCA. The starting hESC population (d0) was largely homogeneous: 99.7% and 95.9% of cells expressed pluripotent markers *POU5F1* (*OCT4*) and *SOX2*, respectively (Supplementary Figure 2.1A).

At day 4 (d4), 97.4% of cells expressed lateral mesoderm marker *HAND1* (Figure 2.1C). The paucity of cells expressing early and late endothelial markers (*KDR* (5.4%) and *PECAM1* (0.16%), respectively) demonstrated that this population was not pre-determined to an endothelial lineage (Figure 2.1C). In cluster d4-B, reduced expression of *MIXL1* (Figure 2.1C) coupled to a higher expression of bone marrow stromal cell marker *BST2* (Supplementary Figure 2.1B), suggested pre-disposition to a mesenchymal fate[279].

In contrast, cells at day 6 (d6) and day 8 (d8) were dichotomised into two divergent populations (Figure 2.1C). The first was characterised by expression of endothelial markers (*KDR* and *PECAM1)*, and the second by mesenchymal markers such as *ACTA2* (Figure 2.1C). Residual heterogeneity within both populations at either d6 or d8 reflects cell cycle variation, indicated by *MKI67* expression (Figure 2.1C). Cluster d8-F (3.2%) likely represented a doublet cluster owing to the expression of both mesenchymal and endothelial markers, whilst lacking unique markers (Supplementary Figure 2.1D).

By d8, fewer than 0.5% of cells expressed either *POU5F1* or *SOX2* with no cells co-expressing both markers (Supplementary Figure 2.2), confirming the absence of residual pluripotency. Mesodermal markers *MIXL1* and *HAND1* were expressed in 12% and 14% of d8 cells, respectively (Supplementary Figure 2.2) confirming a non-mesodermal population.

At day 8, hESC-EC and mesenchymal populations expressed angiogenesis associated genes such as *TGFB1*, *HIF1A,* and *FLT1* suggesting contribution to the pro-angiogenic effect (Supplementary Figure 2.3). Other pro-angiogenic markers, such as *TIE1* and *ANGPT2*, were highly expressed in the endothelial lineage whilst *TGFB2* and *FGF1* were primarily expressed within mesenchymal clusters. To investigate phenotypical changes beyond d8, we repeated the differentiation and processed an additional 13,657 H9 cells at d8 and d12. Separation of endothelial and mesenchymal populations at day 8 was replicated in both rounds of differentiation as well as in a previous replicate

(n=3) (Supplementary Figure 2.4). Distinct endothelial and mesenchymal populations were still evident at day 12 (Figure 2.1D). The proportion of endothelial cells in cycle decreased from 49.6% at d8 to 23.1% at d12 (D12-B). No cluster representing cycling cells was identified in the d12 mesenchymal population.

Genes that were differentially expressed in cluster D8-E (such as *ENPP2, TIE1*, and *CLDN5*) showed elevated expression in hESC-EC at d12 in comparison to at d8 (Figure 2.1E)[92, 280, 281]. This is consistent with the observed increase in cell confluency (Supplementary Figure 2.5). Furthermore, Gene Ontology (GO) term enrichment analysis revealed extracellular matrix organisation to be a significantly enriched term (Figure 2.1F). At each timepoint, hESC-ECs predominantly expressed arterial markers as opposed to venous or lymphatic ECs (Supplementary Figure 2.6).

Longitudinal scRNA-seq analysis reveals rapid endothelial differentiation and subsequent maturation following a bifurcation point

In order to characterise the transcriptional changes throughout differentiation, data from all time points were combined and normalised (see Methods) prior to conducting PCA and data visualisation (tSNE). Cells from d0 and d4 each formed a separate cluster, whereas d6 and d8 cells separate into endothelial or mesenchymal populations (Figure 2.2A). To trace transcriptional dynamics associated with each identified cell type, expression values of pluripotent, mesodermal, mesenchymal, and endothelial markers were used to calculate a

score for each of the four signatures for each individual cell (see Supplementary Methods). As expected, 100% of d0 cells were positive for the pluripotent signature (Figure 2.2B-C), whereas virtually no d4 cells were positive for this signature. Similarly, d4 cells were mainly positive for the mesodermal signature (Figure 2.2B-C) with enriched expression of *APLNR* (Supplementary Figure 2.7). Increased endothelial and mesenchymal signature scores for d6 and d8 within their respective populations suggested that both cell types have matured further following initial specification (Figure 2.2B-C).

The concept of RNA velocity can be used to predict future transcriptional states, taking advantage of changes in read fractions in exonic versus intronic sequence[278]. RNA velocities in Figure 2.2E are represented by the size and direction of arrows. The paucity of vectors with large magnitude and coherent direction associated with d0 and d4 cells is indicative of their equilibrium state (Figure 2.2E). Vectors associated with either endothelial or mesenchymal cell clusters indicated the direction of cell type maturity. The larger vectors for the d8 mesenchymal cell cluster, compared with the d8 endothelial cluster, is consistent with the latter reaching a steadier equilibrium state following initial maturation (Figure 2.2E).

Early (d6) and late (d8) EC and mesenchymal populations exhibit high transcriptional similarity (Figure 2.2D). Nevertheless, higher expression of transcription factor (TF), *SNAI2* was observed in the early (d6) mesenchymal

population in comparison to later mesenchymal population (d8) (Supplementary Figure 2.8). Similarly, higher expression of genes such as *SOX7* and *GJA4* at d6 defines these as markers of early endothelial specification (Supplementary Figure 2.9A-B). Conversely, higher expression of *GNG11, ESM1*, and *ANGPT2* at d8 defines these as later markers of endothelial maturation (Supplementary Material Online, Figure 2.9A-B). RT-qPCR analysis of several marker genes utilising RNA from CD144+ and CD144- populations at d6 and d8 demonstrated high concordance with scRNA-seq findings (Supplementary Figure 2.10).

To ensure findings were not limited to H9 cells, an alternative hESC line, the RC11 line, was used to reproduce findings. Cells were sampled for scRNA-seq at days 4, 6, and 8 (total of 21,972 cells) and the identical analysis applied. Strikingly, we identified equivalent mesodermal, mesenchymal, and endothelial populations in both RC11- and H9-derived cells (Figure 2.2F). A small additional cluster was identified defined by differential expression epithelial markers including *EPCAM*, *CDH1*, and *RAB25* (Supplementary Figure 2.11). This likely reflects an epithelial-to-mesenchymal transition (EMT) during mesoderm formation[282].

Pseudotime analysis is a complementary approach used to determine the path and progress of individual cells undergoing differentiation[277]. The resultant trajectory for H9 hESC-ECP differentiation (Figure 2.3A) indicated a bifurcation point giving rise to distinct endothelial and mesenchymal branches, consistent

with the divergence represented in Figure 2.2A. The majority of d4 cells had pseudotime values close to zero, implying they had yet to commence endothelial or mesenchymal differentiation (Supplementary Figure 2.12). Pseudotime analysis using RC11 cells reproduced an essentially identical bifurcation point as using H9 (Figure 2.3B). Increasing expression of known endothelial markers (Figure 2.3C) and TF (Figure 2.3D) defined the endothelial branch. Similarly, upregulation of mesenchymal markers and TF (Figure 2.3C-D) identified the mesenchymal branch (Figure 2.3A). Expression of the same markers and TF during the RC11 differentiation traced highly similar pseudotime trajectories (Figure 2.3C-D).

scRNA-seq analysis of an alternative hESC-EC differentiation protocol reveals substantial heterogeneity

A comparable longitudinal scRNA-seq data set was acquired using an alternative hESC-EC differentiation protocol, namely the five-factor protocol described by Zang et al[198]. This allowed us to compare its cellular heterogeneity and transcriptional dynamics with those from our 8-day differentiation protocol. Using the Zang et al. 6-day protocol, differentiated H9 cells yielded 51.8% CD31+/CD144+ cells by day 6 (Supplementary Figure 2.13). Cells were sampled for scRNA-seq at a mesodermal stage (d2) and two endothelial stages (d4 and d6). We sequenced the transcriptomes of 27,225 cells. Unsupervised clustering revealed the presence of 12 individual clusters (Figure 2.4A-B).

Within the d2 population, a large (22%) epithelial cluster was identified, characterised by expression of markers including *EPCAM, RAB25*, and *CLDN4* (Figure 2.4C-D). The remaining d2 cells were characterised by expression of mesodermal markers *HAND1* and *BMP4.* Interestingly, d2 mesoderm was also positive for mesendoderm markers such as *GSC* and *EOMES* (Figure 2.4C-D). Expression of a pluripotent marker, *NANOG*, was predominantly localised to the epithelial cluster (Supplementary Figure 2.14A).

By d4, 37% of cells localised to a distinct endothelial population defined by expression of classical endothelial markers (Figure 2.4A-D). The two-remaining major d4 clusters were defined by expression of mesodermal markers (Figure 2.4A-D). In contrast to d2, d4 mesoderm was predominantly negative for expression of mesendoderm markers (Figure 2.4D). The 18.5% reduction of *EOMES*+ cells in cluster 'mesoderm 2', in comparison to cluster 'mesoderm 1' likely identifies this as the later of the two d4 mesoderm populations. Interestingly, a subset of cluster 'mesoderm 2' expressed both *CD34* and *KDR* but not *PDGFRA* suggesting a more specialised mesoderm with restricted specification to a hematoendothelial lineage[283] (Supplementary Figure 2.14B). We annotated a small cluster of d4 cells that differentially express endoderm markers, including *FOXA2* and *CLDN6* as well as pluripotent marker *POU5F1*, as an early endoderm cluster with residual pluripotency (Supplementary Figure 2.14C).

At d6, a comparable *HAND1+* mesoderm population was present (Figure 2.4A-D). The large d6 endothelial cluster accounted for 43% of d6 cells (Figure 2.4A-D). 10.3% of d6 cells localised to a small cluster exclusively expressing various known haemogenic markers including *SPN1*, *SPI1*, and *RUNX1*[284] (Figure 2.4C-D). Furthermore, a small d6 cluster (10.7% of d6 cells) was annotated as being a nephrogenic population, defined by expression of *PAX2, PAX8*, and *CRYM* [285] (Figure 2.4C-D).

We next used pseudotime analysis choosing cells at d4 over d6 so as to compare early versus late mesodermal populations. As expected, the resultant trajectory (Figure 2.4E) revealed cells from the 'mesoderm 1' cluster preceding those from 'mesoderm 2'. Unexpectantly, the bifurcation point giving rise to the endodermal population appeared to emerge from the later, 'mesoderm 2' population. Endothelial and mesodermal cells populated the start of the endothelial branch, thus suggesting that endothelial commitment from mesoderm was still in progress at d4 (Figure 2.4E).

Pseudotime analysis characterises transcription factor expression throughout endothelial differentiation

We sought to characterise the expression of transcription factors over these trajectories. Using a comprehensive list of over 1600 likely human TF[286] we identified 144 that were significantly differentially expressed over pseudotime in both RC11 and H9 differentiations. Of these, several are known to be implicated in endothelial differentiation (*ERG, ETS1*, *SOX7, ELF1*[18]) and these

all resided within cluster 5 (containing 37 genes) of genes differentially expressed across pseudotime (Figure 2.5A).

The expression profile of the 37 cluster 5 TF (Figure 2.5A) was then visualised along the endothelial branch of the trajectory from the five-factor protocol (Figure 2.5B). Expression dynamics for the majority of TF showed substantial resemblance to those observed in the five-factor protocol (Figure 2.5C; Supplementary Figure 2.15). In addition, several TF with no clear reported link to endothelial differentiation also followed a similar expression profile and showed remarkable concordance across cells and protocols (Figure 2.5C).

hESC-EC remain transcriptionally distinct from mature and adult endothelial cells despite undergoing additional maturation following endothelial commitment

We then sought to evaluate the maturity of hESC-EC produced by comparison of scRNA-seq data of foetal endothelial cells and mature endothelial cells with that from H9-derived cells. First, we combined hESC-EC data from d6, d8, and d12 datasets with data obtained from mature primary cultured ECs from each of the arterial, lymphatic or venous vascular beds. PCA revealed that hESC-EC at d6, d8, and d12 cluster separately from mature EC (Figure 2.6A). Interestingly, d12 hESC-EC clustered separately from their d6 and d8 counterparts, but not visibly closer to any given mature EC. Large proportions of genes were expressed within hESC-EC but not in mature EC (d8: 7.2%; d12: 7.3%; Figure 2.6B). Differential gene expression analysis between d8 and

d12 hESC-EC revealed increasing expression levels of several ECM related genes, such as *COL4A1* and *FN1* (Supplementary Figure 2.16A). Furthermore, between d8 and d12 hESC-EC an increase in endothelial signature score was observed, suggesting of further maturation beyond d8 (Supplementary Figure 2.16B).

We then collected all publicly available 10X Chromium scRNA-seq datasets including endothelial cell data from freshly isolated human organs (Supplementary Table 2.3). Differential gene expression analysis of combined endothelial clusters (Figure 2.6C) revealed key transcriptional markers for each organ-specific endothelial phenotype. Visualisation of signature scores across all clusters demonstrated their high specificity to their corresponding cluster (Figure 2.6D). Neither the d8 nor d12 hESC-EC cluster was positive for any organ-specific signature score, thus indicating their likely unspecified nature. Furthermore, correlations between transcriptomes of endothelial cell types were higher among the d6, d8 and d12 cells, than these cells were to fetal (kidney) and mature endothelial cells (Figure 2.6E).

2.8 Discussion

In the present study, we performed the first longitudinal droplet-based scRNA-seq analysis of our highly efficient hESC-EC differentiation protocol[273]. In two different hESC lines, endothelial differentiation from a homogeneous lateral mesoderm population resulted in a bifurcation leading to the emergence of non-specified endothelial and mesenchymal populations. Each population was capable of undergoing further maturation. We performed comparable scRNA-seq analysis of an alternative hESC-EC differentiation protocol, revealing substantially more heterogeneity during differentiation, including the emergence of multiple non-endothelial populations by the protocol endpoint. Pseudotime analysis comparing different protocols and starting hES cell lines uncovered robust transcriptional signatures of endothelial differentiation. Finally, we compared the hESC-EC transcriptional identity to those of primary cultured endothelial cells and freshly isolated foetal and adult endothelial cells. Despite hESC-EC appearing to undergo additional maturation following their initial commitment, they remained transcriptionally distinct from foetal and mature endothelial cell types, with no indication of commitment to an organ-specific endothelial phenotype.

In contrast to earlier time points, at d6 we observed the emergence of two distinct populations expressing either mesenchymal or endothelial markers, a dichotomy that was confirmed by pseudotime trajectory analysis. Recent studies have reported the existence of a common mesenchyme and endothelial precursor, termed the mesenchymoangioblast (MB) derived from

APLNR+/KDR+/PDGFRA+ mesoderm[287]. MB-derived mesenchymal cells were demonstrated to emerge from an endothelial origin by endothelial-to-mesenchymal transition (EndMT) in the absence of VEGF[288]. However, other mesenchymal progenitors, called mesopheres, were found to develop in the presence of VEGF and have shared phenotypic and transcriptomic properties with MBs[289]. The observed expression of the EndMT-associated TF *SNAI2* in the early mesenchymal population may suggest induction of EndMT in the presence of VEGF[290]. However here, although the day 4 mesodermal population selectively expressed *APLNR*, it was found not to express *KDR* or *PDGFRA,* thus suggesting that mesenchymal and endothelial commitment is more likely to occur independently from a common mesoderm progenitor[283].

In contrast to our own 8-day differentiation protocol[273], analysis of the alternative, five-factor differentiation protocol[198] revealed considerably more heterogeneity, with an early mesendoderm population identified by primitive streak markers and residual pluripotency. Interestingly, the model proposed by Evseenko et al[283] shows early mesoderm can then undergo further specification to KDR+/CD34+/PDGFRA- mesoderm which has limited specification potential, forming only hematoendothelial lineages. This model may explain the unexpected haemogenic endothelial cluster in the d6 population. The transcriptional signatures identified around the haemogenic endothelial cluster demonstrated high concordance with scRNA-seq analysis of an iPSC model of Endothelial-to-hematopoietic transition (EHT)[284].

We also identified expression patterns of several known and novel TFs in all three hESC-EC pseudotime trajectories. For example, the expression profile of *BCL6B,* which encodes a transcription factor shown to be expressed within the endothelium of the developing mouse kidney[291], showed remarkable similarity across all trajectories. This suggests a potential role in endothelial differentiation and/or maturation, making this a strong candidate for further experimental validation. Additional cues such as shear stress are likely to be required to drive EC maturation[292]. Our initial attempts to subject d8 hESC-EC to shear stress resulted in detachment of cells (data not shown), possibly due to cells not being at a sufficient level of maturation to lay down the extracellular matrix necessary for effective adhesion or at the correct cell density.

In addition, recent cell tracking studies revealed that the majority of cells are lost soon after transplantation suggesting a paracrine dependence *in vivo*[273, 293]. Crucially with scRNA-seq analysis, several angiogenic factors such as *FLT1* and *HIF1A*, were highly expressed across both endothelial and mesenchymal populations in our d8 hESC-ECP. Furthermore, the non-teratogenic nature of the hESC-ECP was further demonstrated by the complete absence of cells co-expressing *SOX2* and *POU5F1.*

In conclusion, we have used a droplet based scRNA-seq platform to map the transcriptional changes throughout the differentiation of hESCs to pluripotent stem cell derived EC across different cell lines and differentiation protocols (Figure 2.7). This provides novel insight into early endothelial commitment and

maturation and will aid in the refinement of protocols that drive endothelial specification in a lineage or organ specific nature. We also provide essential evidence of the transcriptional state and heterogeneity by day 8 of our protocol, and we are developing a first-in-man cell therapy approach using day 8 differentiated cells.

2.9 Figure Legends

Figure 2.1: scRNA-seq analysis of hESC-ECP differentiation at individual time points. (A) Schematic of hESC-ECP differentiation and scRNA-seq pipeline. **(B)** Representative flow cytometric analysis of cells taken at day 8 (hESC- ECP). Cells stained for endothelial (CD31 and CD144) (left panels), pluripotent markers (TRA-181 and SSEA-4) (right panels), and their corresponding isotype controls. **(C)** tSNE plots, constructed separately, using data from cells sampled at days 4, 6, and 8 for scRNA-seq. Feature plots (right) show key marker gene expression for each population. **(D)** tSNE plot constructed from d12 data from a separate round of H9 differentiation. **(E)** Violin plots showing expression (ln(UMI+1)) of angiogenic and cell contact related genes in d8 and d12 hESC-EC. **(F)** GO term analysis conducted using genes significantly differentially expressed within d12 hESC-EC in comparison to d8 hESC-EC.

Figure 2.2: scRNA-seq analysis across all time points mapping the transcriptional changes and predicted future direction of cells undergoing hESC-ECP differentiation. (A) tSNE plot constructed using data from 21,369 cells taken at day 0, 4 ,6, and 8 during H9 to hESC-ECP differentiation. Cluster identity determined by the expression of known pluripotent (*SOX2*), mesodermal (*HAND1*), mesenchymal (*ACTA2*), and endothelial markers (*PECAM1*). **(B)** Cell signature scores shown on tSNE plot from Figure 2.2A. **(C)** Cell signature scores within identified clusters from Figure 2.2A. **(D)** PCA plot containing cells taken at the four time points during

H9 differentiation. Feature plots show key marker gene expression. **(E)** RNA velocities visualised on the PCA plot from Fig. 2D**. (F)** tSNE plot using data from d4, d6, and d8 during differentiation of RC11.

Figure 2.3: Pseudotime trajectories of hESC-ECP differentiation. Pseudotime trajectories containing cells from d4, d6, and d8 during **(A)** H9 differentiation and **(B)** RC11 differentiation. Arrows show direction of pseudotime across trajectories. Expression of known endothelial and mesenchymal **(C)** markers and **(D)** transcription factors in their respective branches for each of the two starting ES cells (H9 and RC11).

Figure 2.4: Longitudinal scRNA-seq analysis of an alternative hESC-EC differentiation protocol. (A) tSNE plot showing data of hESC-EC differentiation using an alternative protocol. Cells taken for scRNA-seq at d2, d4, and d6. **(B)** tSNE plot of alternative differentiation protocol, categorising cells according to their original dataset. **(C)** Signature scores shown on tSNE plot from Figure 2.4A. **(D)** Key marker gene expression used to characterise identified clusters. **(E)** Pseudotime trajectory containing cells from d4 clusters of the alternative differentiation protocol.

Figure 2.5: Pseudotime analysis of transcription factor expression during hESC-EC differentiation. (A) Differential gene expression variance over pseudotime of H9 differentiation within pseudotime trajectory branches. Genes used are TF that are differentially expressed in both H9 and RC11-

derived trajectories. **(B)** Gene expression variance of TF taken from cluster 5 of heatmap from Figure 2.5A in the endothelial branch of the trajectory from the five-factor protocol. **(C)** Spline plots showing the expression of novel TF from Figure 2.5B across pseudotime in RC11 and H9 hESC-EC differentiation using the original protocol, as well as for the five-factor protocol.

Figure 2.6: Transcriptomic comparison of hESC-EC to mature endothelial cells. (A) 3D PCA plot of hESC-ECs from d6, d8, and d12 datasets alongside mature EC: human dermal lymphatic ECs (HDLEC), human umbilical vein ECs (HUVEC), and human pulmonary artery ECs (HPAEC). **(B)** Venn diagrams of number of genes expressed in d6, d8, and d12 hESC-EC and the overlap of these genes with each type of mature endothelial EC. Percentage of uniquely expressed hESC-EC genes shown in red. **(C)** tSNE plot showing clustering of EC extracted from 10X datasets of human organs. **(D)** Signature scores for each EC type shown in Figure 2.6C. **(E)** Heatmap with hierarchical clustering comparing transcriptional status of each cell type by Spearman's rank correlation coefficient.

Figure 2.7: Summary of findings from scRNA-seq analysis of hESC-EC differentiation: Proposed model for the differentiation of the hESC-derived endothelial cells via two differentiation methods. Human embryonic stem cells (hESC) undergo highly efficient differentiation using the hESC-ECP differentiation method, forming a homogeneous lateral mesoderm population. Following endothelial induction, a bifurcation point exists, giving rise to early

endothelial and mesenchymal populations. Conversely, in the five-factor protocol, hESC cells undergo differentiation to an earlier mesendoderm population. Subsequent endothelial induction drives the mesendoderm to give rise to multiple unwanted cell types including nephron progenitors and haemogenic endothelium. Despite clear differences in cellular heterogeneity, transcriptional dynamics of hESC-EC differentiation has substantial commonalities between differentiation methods.

Figure 2.1 scRNA-seq analysis of hESC-ECP differentiation across individual time points

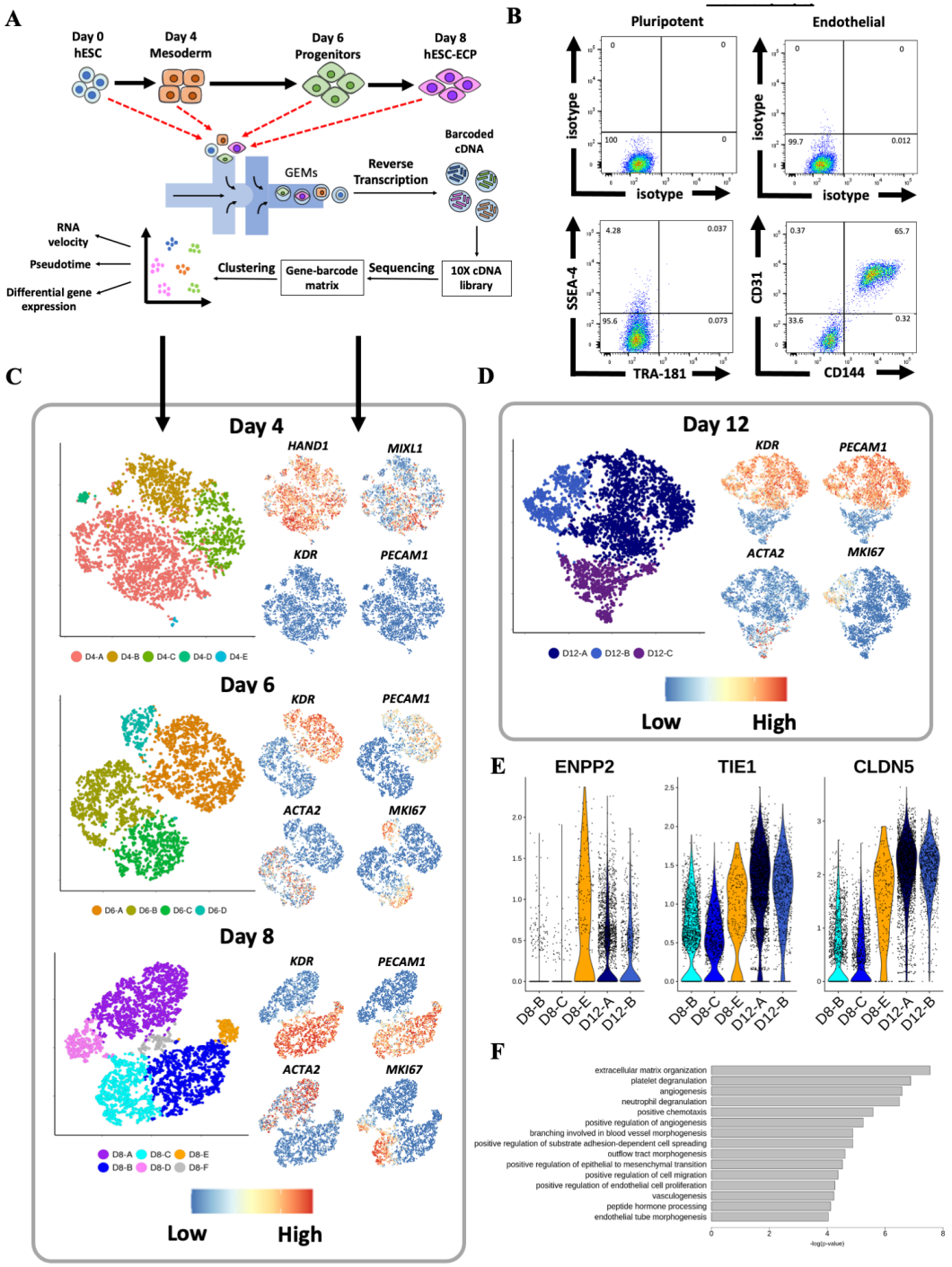


Figure 2.2 Mapping hESC-ECP differentiation using scRNA-seq

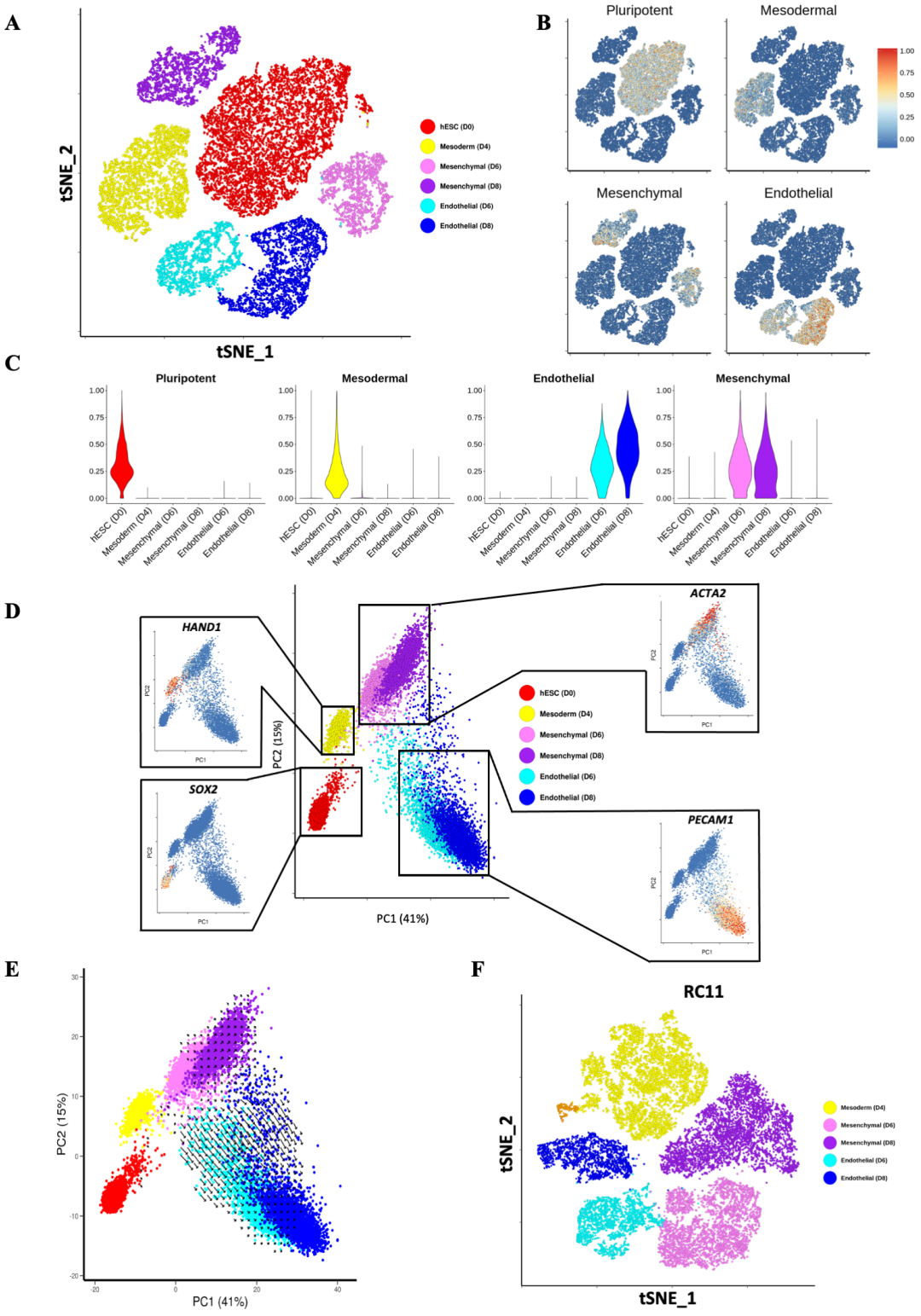


Figure 2.3 Pseudotime trajectories of hESC-ECP differentiation

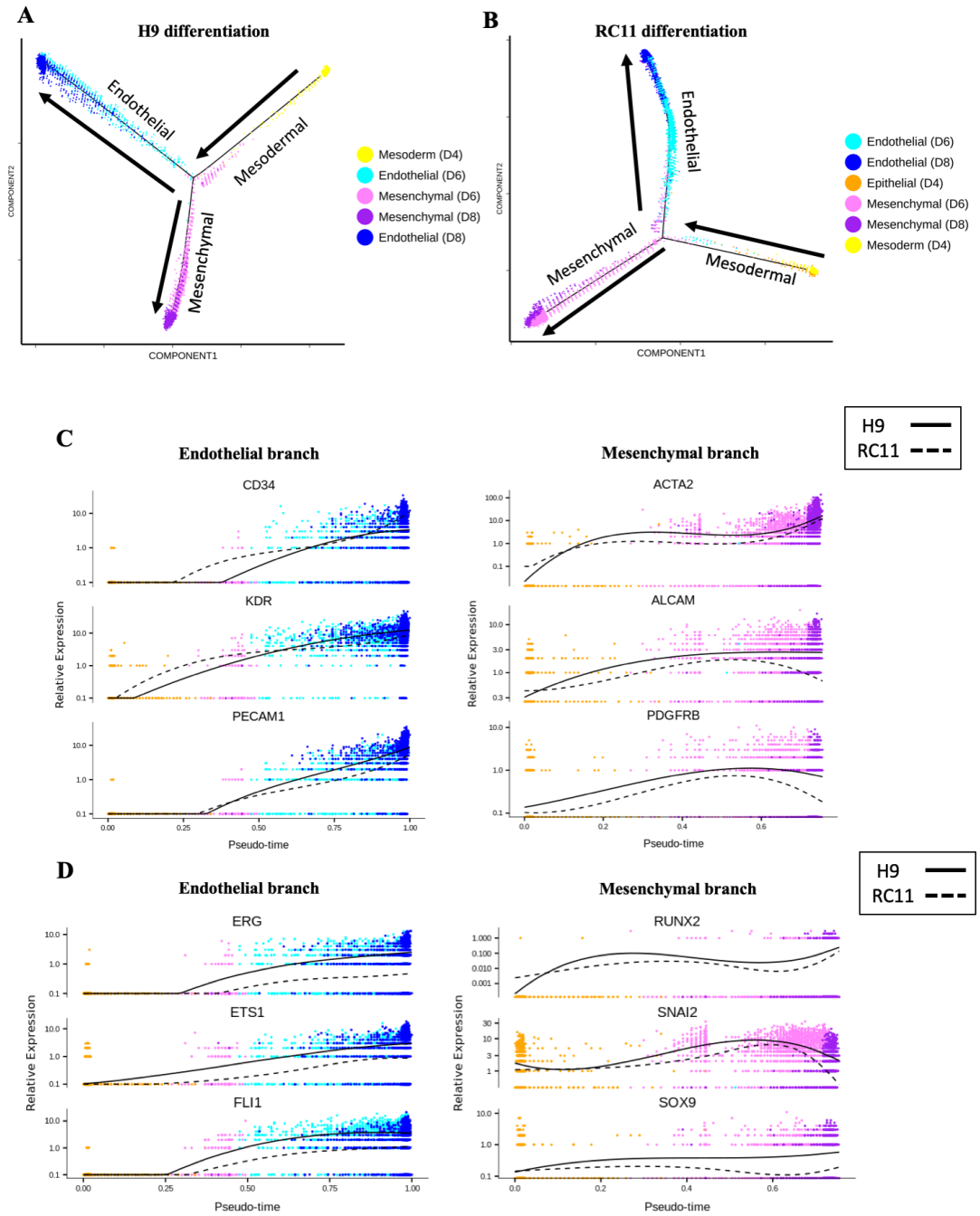


Figure 2.4 Longitudinal scRNA-seq analysis of an alternative hESC-EC differentiation

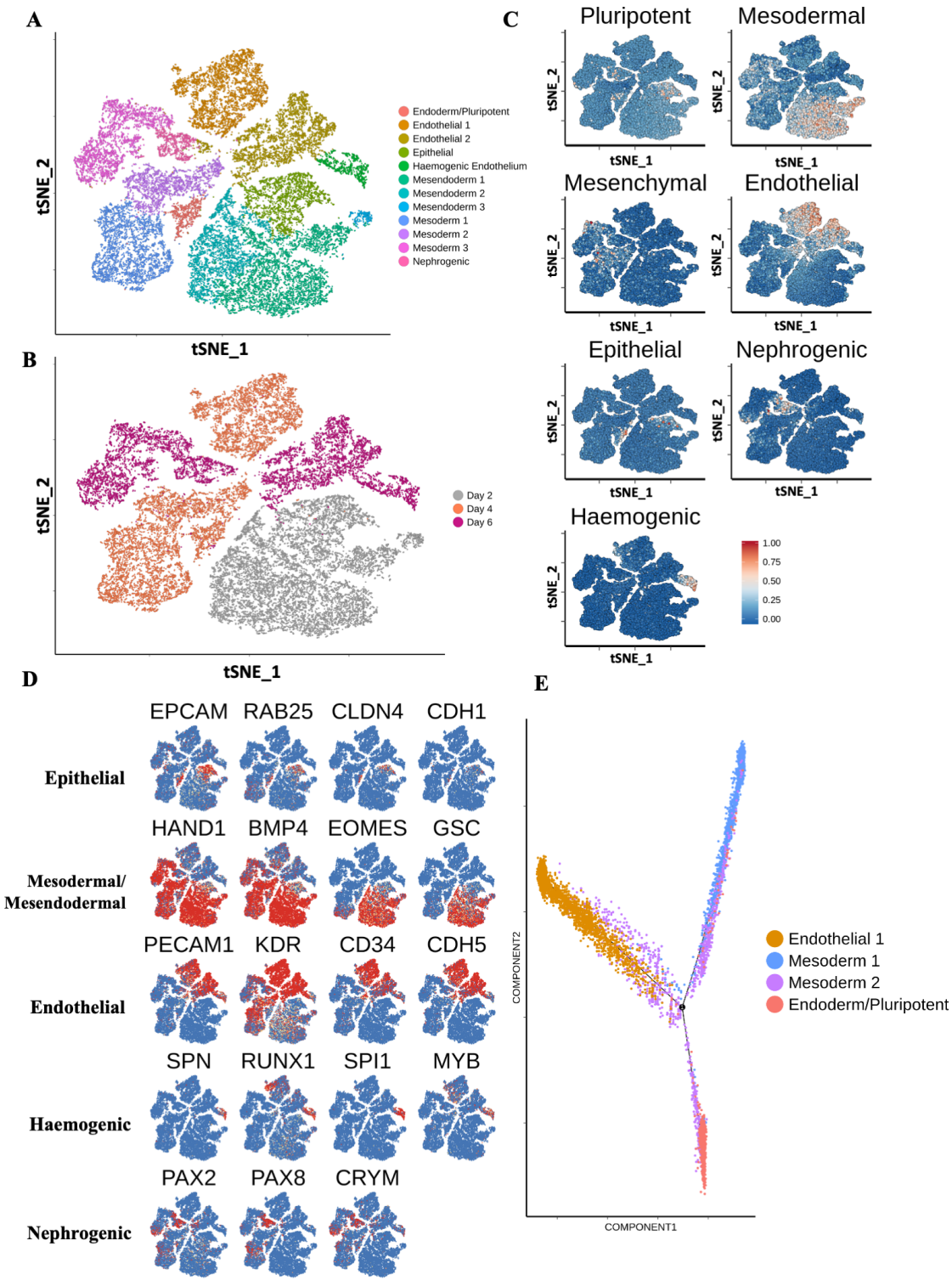


Figure 2.5 Pseudotime analysis of hESC-EC differentiation

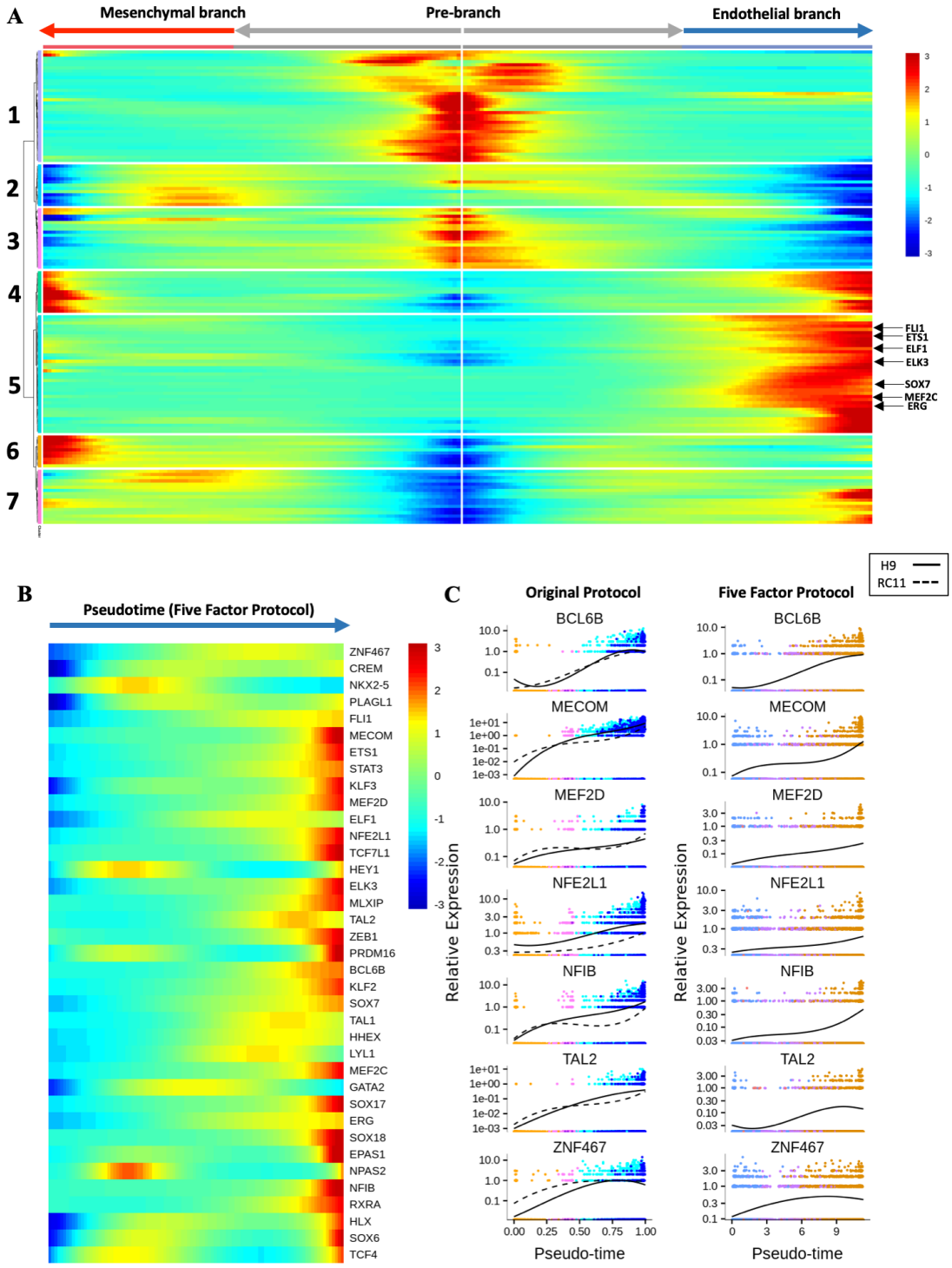


Figure 2.6 Transcriptomic comparison of hESC-EC to mature endothelial cells

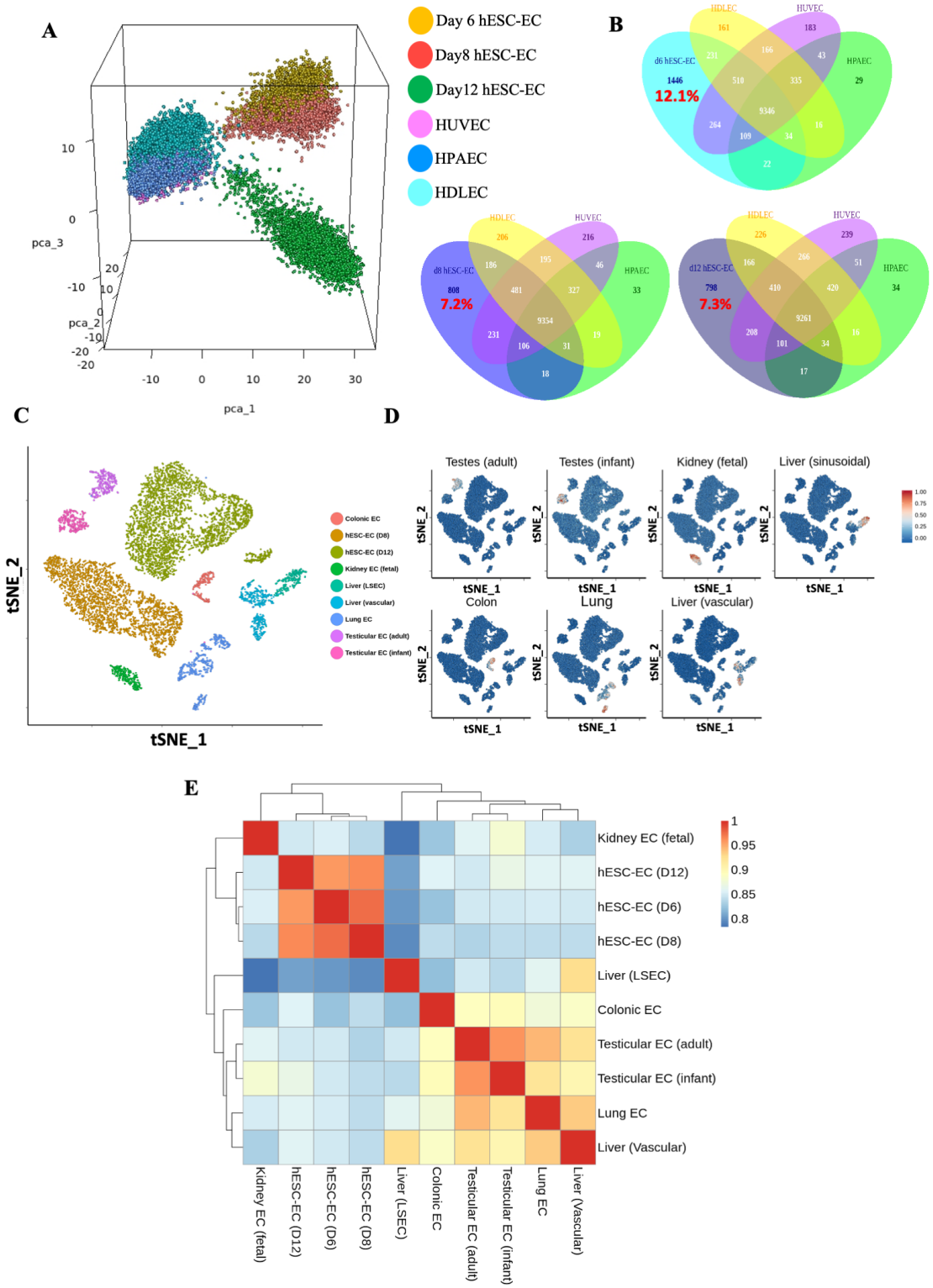
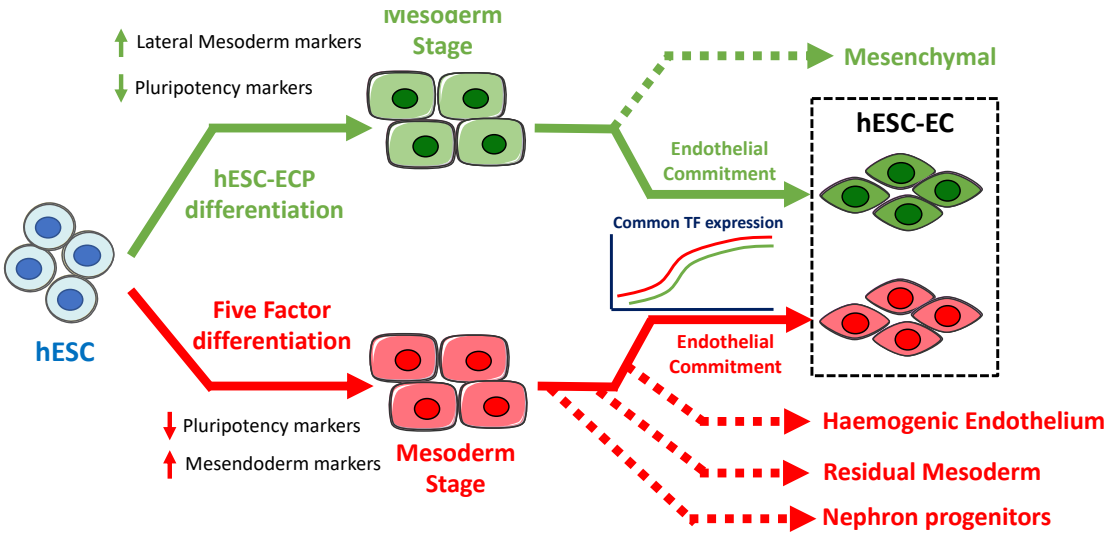


Figure 2.7 Summary of findings from scRNA-seq analysis of hESC-EC differentiation



2.10 Supplementary Methods

Differentiation of hESCs to hESC endothelial cell product (hESC-ECP).

H9 (P38-47) or RC11 (P34) hESCs were differentiated to the hESC-ECP using the previously reported protocol[273]. This involved hESCs previously maintained on a vitronectin matrix (Life Technologies, Paisley, UK), being dissociated and plated as single cells on a fibronectin matrix (R&D Systems, Minneapolis, USA) at day 0 with mTeSR1 media (STEMCELL TECHNOLOGIES, Vancouver, Canada) supplemented with 10μM ROCK inhibitor (Y27632) (Tocris, Bristol, UK). At day 1, lateral mesoderm was induced with 7μM CHIR-99021 (Sigma, St. Louis, USA) and 25ng/ml BMP-4 (R&D Systems, Minneapolis, USA) added to N2B27/Neurobasal/DMEM:F12 media (Life Technologies, Paisley, UK). Cells were then left without changing media until endothelial cell induction at day 4 using 2μM Forskolin (Sigma, St. Louis, USA) and 200ng/ml VEGF (R&D Systems, Minneapolis, USA) in Stem Pro34 media (Life Technologies, Paisley, UK). Cells were subject to complete media change at day 5, using the same media used at day 4. In the final step, day 6 cells were replated onto fresh plates without matrix and cultured with daily media changes until day 8 with human 1% hAB serum (Sigma, St. Louis, USA) and VEGF (50ng/ml) in EGM-2 media (Lonza, Basel, Switzerland). Cells taken for scRNA-seq library construction were detached using StemPro Accutase (Life Technologies, Paisley, UK) at day 0, or TrypLE Express Enzyme (Thermo Fisher Scientific, Waltham, USA) at days 4, 6, 8, and 12.

In the alternative “Five Factor” protocol H9 cells were differentiated to hESC-EC as described by Zhang et al[198]. Briefly, hESC were dissociated to single cells and seeded in vitronectin coated wells (Life Technologies, Paisley, UK) at a density of 50,000 cells/cm^2. Cells were cultured for 48 hours in the E8 media (STEMCELL TECHNOLOGIES, Vancouver, Canada) supplemented with BMP4 (5 ng/mL) (R&D Systems, Minneapolis, USA), Activin A (25 ng/mL) (STEMCELL TECHNOLOGIES, Vancouver, Canada), and CHIR99021(1 μM) (Sigma, St. Louis, USA). To improve cell survival 10 μM Y27632 (Tocris, Bristol, UK) was used for the first 24 hours. At day 2, media was changed to E5 media supplemented with FGF (100 ng/mL) (R&D Systems, Minneapolis, USA), VEGF (50 ng/mL) (R&D Systems, Minneapolis, USA), SB431542 (10 μM) (Tocris, Bristol, UK), RESV (5 μM) (STEMCELL TECHNOLOGIES, Vancouver, Canada), and L690 (10 μM) (Tocris, Bristol, UK). Media was then changed daily until day 6.

Cell staining and flow cytometry.

Day 8 cells were stained using antibodies (given in Supplementary Table 2.1) to determine the proportion of pluripotent (SSEA-4$^+$/TRA-181$^+$) and endothelial (CD31$^+$/CD144$^+$) cells within the population. Flow cytometry was conducted with the BD LDR Fortessa system (Becton Dickinson, Franklin Lakes, USA) or the Attune NxT system (Thermo Fisher Scientific, Waltham, USA) and data analysed using the FlowJo software (FlowJo LLC, Ashland, USA).

Mature endothelial cell culture.

Primary human dermal lymphatic endothelial cells (HDLECs) obtained from PromoCell (Heidelberg, Germany) were maintained with Endothelial Cell Media MV2 (PromoCell, Heidelberg, Germany) supplemented with Penicillin-Streptomycin (P/S) (100 U/ml) (Gibco, Paisley, UK). Human umbilical vein endothelial cells (HUVEC) (Lonza, Basel, Switzerland) and human pulmonary artery endothelial cells (HPAEC) (Lonza, Basel, Switzerland) were maintained using endothelial growth media (EGM-2 BulletKit) (Lonza, Basel, Switzerland) supplemented with 10% FBS (Life Technologies, Paisley, UK) and P/S (100U/ml) (Gibco, Paisley, UK). Upon harvesting for scRNA-seq, mature cells were detached using 1X Trypsin-EDTA before washing and resuspending in PBS-0.01% BSA. Cells were used for scRNA-seq between P3 and P6 and cultured at 37°C and 5% CO_2.

Transcript expression validation.

H9 cells were differentiated in biological triplicates and cells harvested at day 6 and day 8 by CD144 MACS beads (Miltenyi Biotech, Bergisch Gladbach, Germany) into CD144 positive and CD144 negative populations. Separated cell populations were then resuspended in Qiazol (Qiagen, Manchester, UK) prior to total RNA extraction using the miRNAeasy kit (Qiagen, Manchester, UK). For gene expression analysis, cDNA was produced from total RNA using the Multiscribe Reverse Transcriptase (Life Technologies, Paisley, UK). Quantitative real-time polymerase chain reaction (qRT-PCR) was performed in technical triplicates using the QuantStudio 5 Real-time PCR system (Life

Technologies, Paisley, UK) and TaqMan probes (Life Technologies, Paisley, UK). Ubiquitin C was used for normalisation (Life Technologies, Paisley, UK). Relative expression was calculated using the 2-(ΔΔCT) (RQ) method.

Single cell RNA sequencing library construction.

Prior to loading the 10X Chromium controller, cells were stained with trypan blue and viability and concentration determined using a BioRad TC20 automated cell counter (BioRad, Hercules, USA). Viability of all samples exceeded 80%. Construction of single cell RNA-sequencing libraries was conducted using the Single Cell 3' Reagent Kit (10X Genomics, Pleasanton, USA) in accordance with the manufacturer's protocol. In brief, following generation of the gel bead in emulsions (GEMs) and reverse transcription, cDNA was extracted from GEMs and amplified. cDNA clean-up was performed using a SPRIselect Reagent Kit (Beckman Coulter, Brea, USA). Indexed sequencing libraries were generated using the Chromium Single-Cell 3' Library Kit (10X Genomics, Pleasanton, USA). The key steps in this protocol included enzymatic fragmentation, end-repair, A-tailing, adaptor ligation, ligation clean-up, sample index PCR, and PCR clean-up. Recovery rate across all samples ranged from between 3,778 and 10,053 cells (Supplementary Table 2.2). Libraries were sequenced on an Illumina NovaSeq 6000 platform (Supplementary Table 2.2) except from libraries obtained from HUVECs, HPAECs, and the 'hESC-ECP day 8 replicate 2' sample which were sequenced on an Illumina HiSeq 4000. All sequencing was performed by Edinburgh Genomics (Edinburgh, UK).

Single cell RNA sequencing data analysis.

Raw reads were processed using the 10X Cell Ranger pipeline (10X Genomics) in order to perform read alignment, filtering, barcode and UMI (unique molecular identifier) counting, and subsequently generate gene-cell matrices. The Cell Ranger pre-built GRCh38 based reference transcriptome was used as the transcriptome reference. Subsequent downstream analysis was conducted using Seurat version 2.2.1 [276]. 10X Chromium datasets containing endothelial cell data (Supplementary Table 2.3) from the various human organs (testicle infant[294], testicle adult[294], lung[295], foetal kidney[296], colon[297], liver (LSEC)[298], and liver (vascular)[298]) were directly imported as count matrices. Initial recommended quality control steps performed for each data set involved the removal of any cells with a UMI count greater than 10,000 and cells expressing fewer than 250 genes [276]. Additionally, cells in which more than 15% of reads corresponded to mitochondrial genes were discarded. Data was then normalised utilising a global scaling method which normalises gene expression within each cell based on the total gene expression followed by multiplying by a scale factor of 10,000 and then $\log_e$ transforming the result. Data was then scaled using the ScaleData [276] method to regress out unwanted sources of variation such as UMI count. Most variable genes across cells were then identified using the FindVariableGenes method [276], selecting genes with an average expression (quantified as normalised ln(UMI+1)) between 0.0125 and 3 and with a log variance to mean ratio (VMR) greater than 0.5. Selected highly variable genes were then used for principal component analysis (PCA) in the scaled data. The appropriate number of principal components (PC) to

be used for graph based clustering and non-linear dimensional reduction (tSNE) was determined by choosing the PC after which the standard deviation of subsequent PCs remained approximately constant. Cells were then clustered using the Seurat graph-based clustering method[276] using this number of PCs. tSNE plots were then generated for visualising the identified clusters using the appropriate number of PCs. Differential gene expression analysis between clusters compared gene expression within cells of a given cluster to either the rest of the dataset or a specified cluster. Significantly differentially expressed markers were identified as having an adjusted (Bonferroni correction) P value <0.05 as determined using the Wilcoxon signed rank test. Additionally, differentially expressed markers expressed in fewer than 25% of the chosen cluster and with a $\log_e$ fold change (logFC) in expression less than 0.25 were not considered for further analysis. Heatmaps of individual datasets were generated using the top 10 most differentially expressed genes in a given cluster (based on logFC in expression).

Construction of Venn diagrams illustrating the number of overlapping genes between cell types was conducted using only those genes expressed within more than 5% of cells in a given cluster. All raw sequencing data have been deposited in the Gene Expression Omnibus (GEO; GSE131736).

Cell type signatures.

Cell type signature scores were generated using the following marker genes for each cell type: pluripotent (POUF51, SOX2, NANOG), mesodermal

(HAND1, T, MIXL1), mesenchymal (ACTA2, ALCAM, PDGFRB), endothelial (PECAM1, KDR, CD34, CDH5), epithelial (*EPCAM, CLDN4, RAB25, WNT6*), nephrogenic (*PAX2, PAX8, CRYM*), and haemogenic (*SPI1, RUNX1, SPN, MYB, GFI1B, GATA2*). Signature scores per cell were calculated as the geometric mean of expression values for these marker genes.

GO Term Enrichment Analysis.

The topGO R package[299] was used to perform gene ontology enrichment analysis on the genes that were differentially expressed between the hESC-EC at d8 and d12. The fisher exact test with "elim" algorithm determined which GO terms were significantly enriched[299].

Inference of cellular trajectories and directionality.

Monocle R package v.2.6.4[277] was used to calculate cellular trajectories and pseudotime dynamics on the quality-controlled cells sampled at d4, d6 and d8 from our protocol and day 4 from the five-factor protocol. Day 0 cells were removed from the pseudotime analysis as these were demonstrated to be non-transitional, as well as cells identified as likely doublets. A set of differentially expressed genes between these datasets was generated using the differentialGeneTest function in Monocle with a q-value upper threshold of $1x10^{-40}$, and this was used as the set of informative genes for the dimensionality reduction to two dimensions, performed with reduceDimension, and the pseudotime calculation performed with orderCells and default Monocle settings [277]. Pseudotime values were scaled from 0 to 1 and the value assigned

to each cell was visualised by mapping to the previously generated tSNE and PCA plots, and on Monocle trajectory plots generated with the plot_cell_trajectory. Differentially expressed genes across pseudotime were calculated by selecting the earliest 10% of cells in both mesenchymal and endothelial branches and differentially expressed genes were found using the Seurat FindMarkers command with default settings, and a value of k =4 in the subsequent hierarchical clustering [277].

The 'RNA velocity' of each cell was calculated using the velocyto R package v0.6.0 [278], which uses the proportion of reads containing unspliced intronic sequence to estimate whether genes are in steady state, and to estimate cell-specific RNA velocity and predict future cell states based on an assessment of many such genes. This has been shown to be an effective method of inferring the directionality of cell differentiation with 10X Chromium data, where typically around 23% of a library incorporate intronic sequence [278]. DropEst was used to count the reads mapping to intronic and exonic sequence, using the binary alignment files generated by Cell Ranger in the initial processing. Velocyto was then used to calculate gene-relative velocity using kNN pooling with k=25, the slope gamma determined using the full range of cellular expression and fitting gene offsets with spanning reads [278]. Aggregate velocity fields were then mapped to and displayed on the previously generated PCA plots.

Supplementary Table 2.1: Antibodies used for flow cytometric analysis.

	Antibody Name	Dilution
Pluripotent	PE Mouse anti-Human TRA-1-81 Antigen, BD Biosciences, 560161	1:100
	PE Mouse IgM, κ Isotype Control, BD Biosciences, 555584	1:1600 (final concentration matched to PE TRA-1-81)
	Alexa Fluor 647 Mouse anti-SSEA-4, BD Biosciences, 560796	1:100
	Alexa Fluor 647 Mouse IgG3, κ Isotype Control, BD Biosciences, 560803	1:4000 (final concentration matched to AlexaFluor647 SSEA-4)
Endothelial	PE Mouse anti-Human CD144, BD Biosciences, 580410	1:200
	PE Mouse IgG1, κ Isotype Control, BD Biosciences, 555749	1:200
	APC Anti-Human CD31, eBioscience, 17-0319-42	1:200
	APC Mouse IgG1 K Isotype Control, eBioscience, 17-4714-42	1:200

Supplementary Table 2.2: Summary of sequencing of 10X libraries.

Sample Name	Number of Cells	Mean Reads Per Cell	Median Genes per Cell
Day 0	8,179	56,802	3,612
Day 4	3,778	119,015	4,958
Day 6	4,127	112,128	4,782
Day 8	5,285	84,463	4,948
Day 8 replicate	4,424	75,729	4,220
HDLEC	5,135	77,891	4,439
HPAEC	7,663	35,630	3,358
HUVEC	4,282	90,251	4.992
FF day 2	10,053	70,348	4,137
FF day 4	9,298	60,494	4,088
FF day 6	7,874	59,862	2,945
RC11 day 4	6,882	76,612	4,190
RC11 day 6	6,959	80,378	3,900
RC11 day 8	8,131	66,389	4,157
H9 day 8 replicate 3	7,165	86,868	4,355
H9 day12	6,492	84,492	3,728

Supplementary Table 2.3: Summary of 10X EC data obtained from publicly available datasets.

Sample Name	Number of EC Cells	Reference
Testicular EC (infant)	337	Guo et al (2018)[294]
Testicular EC (adult)	350	Guo et al (2018)[294]
Lung EC	638	Reyfman et al (2018)[295]
Kidney EC (foetal)	308	Lindstorm et al (2018)[296]
Colonic EC	243	Kinchen et al (2018)[297]
Liver (LSEC)	273	MacParland et al (2018)[298]
Liver vascular	522	MacParland et al (2018)[298]

2.11 Supplementary Legends

Supplementary Figure 2.1: Individual time point analysis of hESC-ECP differentiation. (A) tSNE plot constructed from data using the day 0 (hESC) dataset. Feature plots show expression of key pluripotent, mesodermal, and endothelial marker genes. **(B-E)** Heatmaps of 10 most differentially expressed genes within clusters identified in cells taken at **(B)** d4, **(C)** d6, **(D)** d8, and **(E)** d12 (from a separate round of differentiation). Each column represents a single cell whilst rows represent a z-score of normalised ln(UMI+1). **(B)** Expression of BST2 in cluster d4-B shown in feature plot.

Supplementary Figure 2.2: Expression of pluripotent and mesodermal markers in hESC-ECP. hESC-ECP (d8) expression of pluripotent markers (*SOX2* and *POU5F1*) and mesodermal markers (*HAND1* and *MIXL1*) shown on day 8 tSNE plot.

Supplementary Figure 2.3: Localisation of angiogenic gene expression within hESC-ECP. Expression of angiogenic related genes in hESC-ECP (day 8) shown on day 8 tSNE plot.

Supplementary Figure 2.4: PCA comparison of scRNA-seq data from three day 8 (hESC-ECP) replicates. PCA plot of PC1 vs PC2 produced using data from day 8 of longitudinal hESC-ECP study (day8_rep1) and two other rounds of hESC-ECP differentiation (day8_rep2 and day8_rep3).

Supplementary Figure 2.5: Morphological appearance of hESC-ECP at d8 and d12. Brightfield images of hESC-ECP taken at day 8 (top) and day12 (bottom) at 10X magnification. Scale bars represent 400μm.

Supplementary Figure 2.6: Vasculature specific marker gene expression in progenitor and hESC-ECP populations. Arterial (left), lymphatic (middle), and venous (right) specific endothelial marker gene expression at d6, d8, and d12 of differentiation.

Supplementary Figure 2.7: Expression *APLNR* throughout hESC-ECP differentiation.

Supplementary Figure 2.8: Expression of *SNAI2* during hESC-ECP differentiation.

Supplementary Figure 2.9: Differential gene expression throughout hESC-ECP differentiation. (A) Heatmap of top 10 most differentially expressed genes in identified cell types during hESC-ECP differentiation until d8. **(B)** Enriched expression of differentially expressed genes during hESC-ECP differentiation.

Supplementary Figure 2.10: RT-qPCR validations of gene expression within d6 and d8 CD144 positive and CD144 negative populations during hESC-ECP differentiation. d6 and d8 of hESC-ECP differentiation cells were harvested throughout hESC-ECP differentiation and sorted via MACS into

CD144 positive and CD144 negative populations. RT-qPCR analysis used to validate marker genes expression as identified by scRNA-seq analysis. Results were obtained from three independent rounds of hESC-ECP differentiation (n=3) and shown as the mean relative quantity of gene expression +/- SEM. Statistical significance determined using paired t-test. *= $p<0.05$, **= $p<0.01$ and ***= $p<0.001$.

Supplementary Figure 2.11: Epithelial marker gene expression during differentiation of RC11 hESC to hESC-ECP. Violin plots show expression ln(UMI+1) of epithelial markers within clusters identified during the differentiation of RC11 hESC to hESC-ECP (d8).

Supplementary Figure 2.12: Histogram showing distribution of cells throughout the pseudotime trajectory. Frequency of cells across pseudotime values corresponding to each branch of the trajectory for the differentiation of H9 hESC to hESC-ECP (d8).

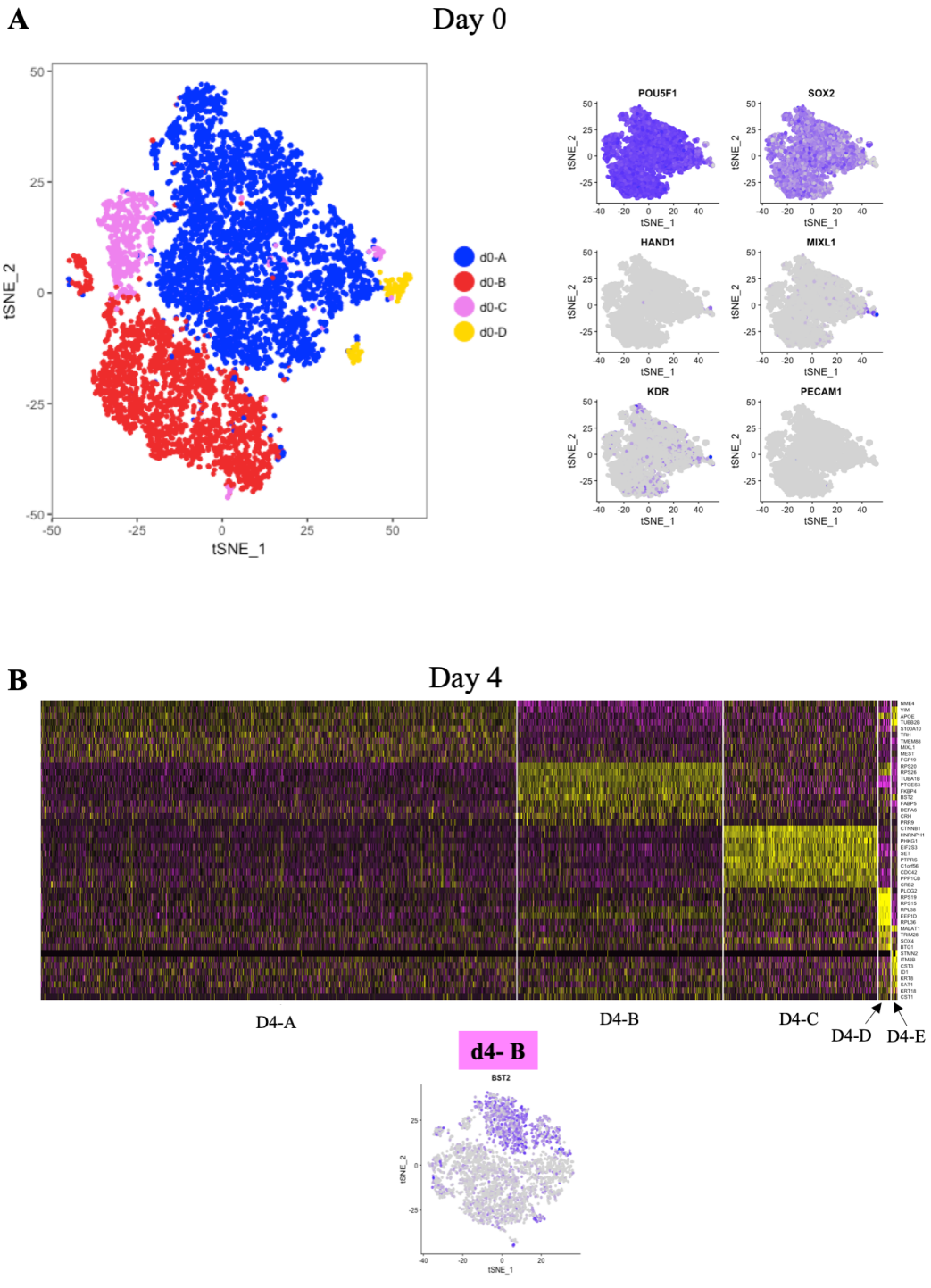
Supplementary Figure 2.13: Flow cytometric analysis of d6 cells from the Five Factor differentiation protocol. Representative flow cytometric analysis of cells taken at d6 (protocol endpoint) from the Five Factor differentiation protocol. Cells stained for endothelial markers (CD31 and CD144) (right panel) as well as the corresponding isotype controls (left panel).

Supplementary Figure 2.14: Marker gene expression during Five Factor hESC-EC differentiation protocol. (A) Expression of pluripotent markers within clusters (visualised in tSNE) identified from data of five factor differentiation protocol. **(B)** Feature plots showing co-expression of CD34 and KDR (left panel) and PDGFRA (right panel) **(C)** Feature plots from five factor protocol showing expression of endodermal markers.

Supplementary Figure 2.15: Expression of known endothelial TF over pseudotime. Expression of well characterised TF known to be involved in endothelial differentiation across pseudotime in the differentiation of H9 and RC11 cells to the hESC-ECP (original protocol – left panel), and for H9 cell differentiation to hESC-EC using the five-factor protocol (right panel).

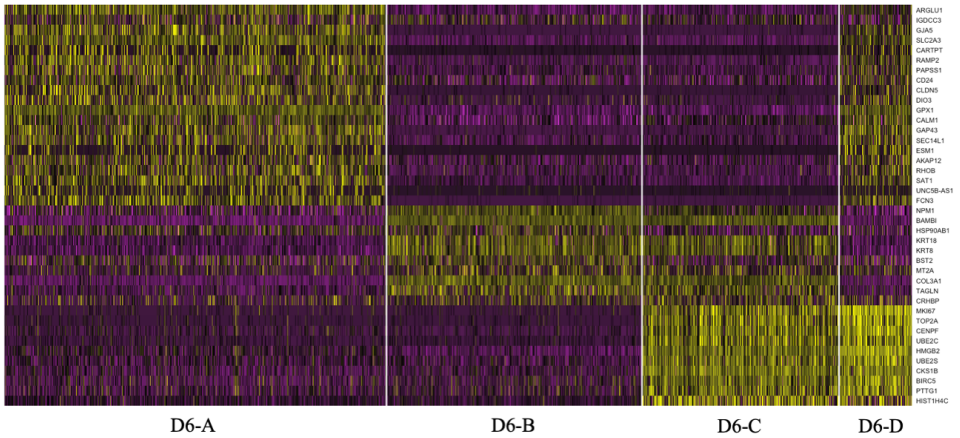
Supplementary Figure 2.16: Transcriptional changes between day 8 and day12 cells from hESC-ECP differentiation protocol. (A) Violin plots comparing expression of ECM associated genes within hESC-EC clusters identified within day 8 and day 12 samples. **(B)** Endothelial signature scores for cells within hESC-EC clusters identified within d8 and d12 datasets.

Supplementary Figure 2.1



C

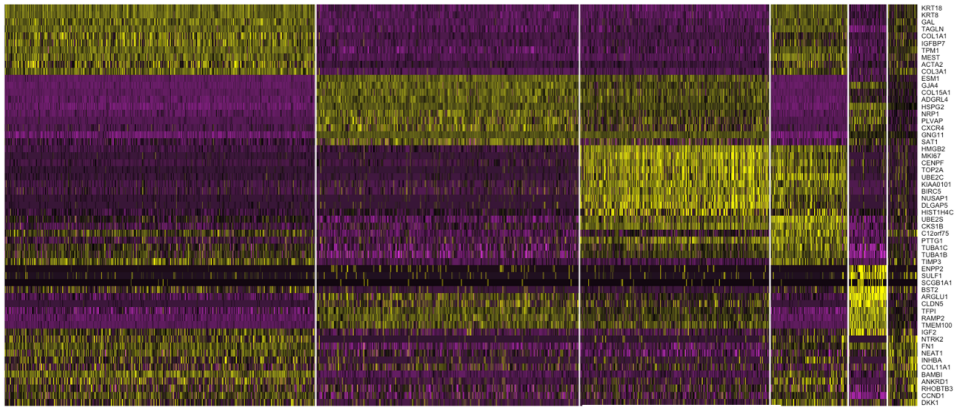
Day 6



D6-A D6-B D6-C D6-D

D

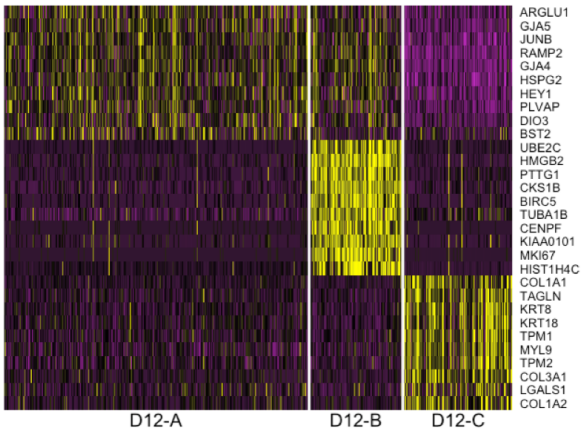
Day 8



D8-A D8-B D8-C D8-D D8-E D8-F

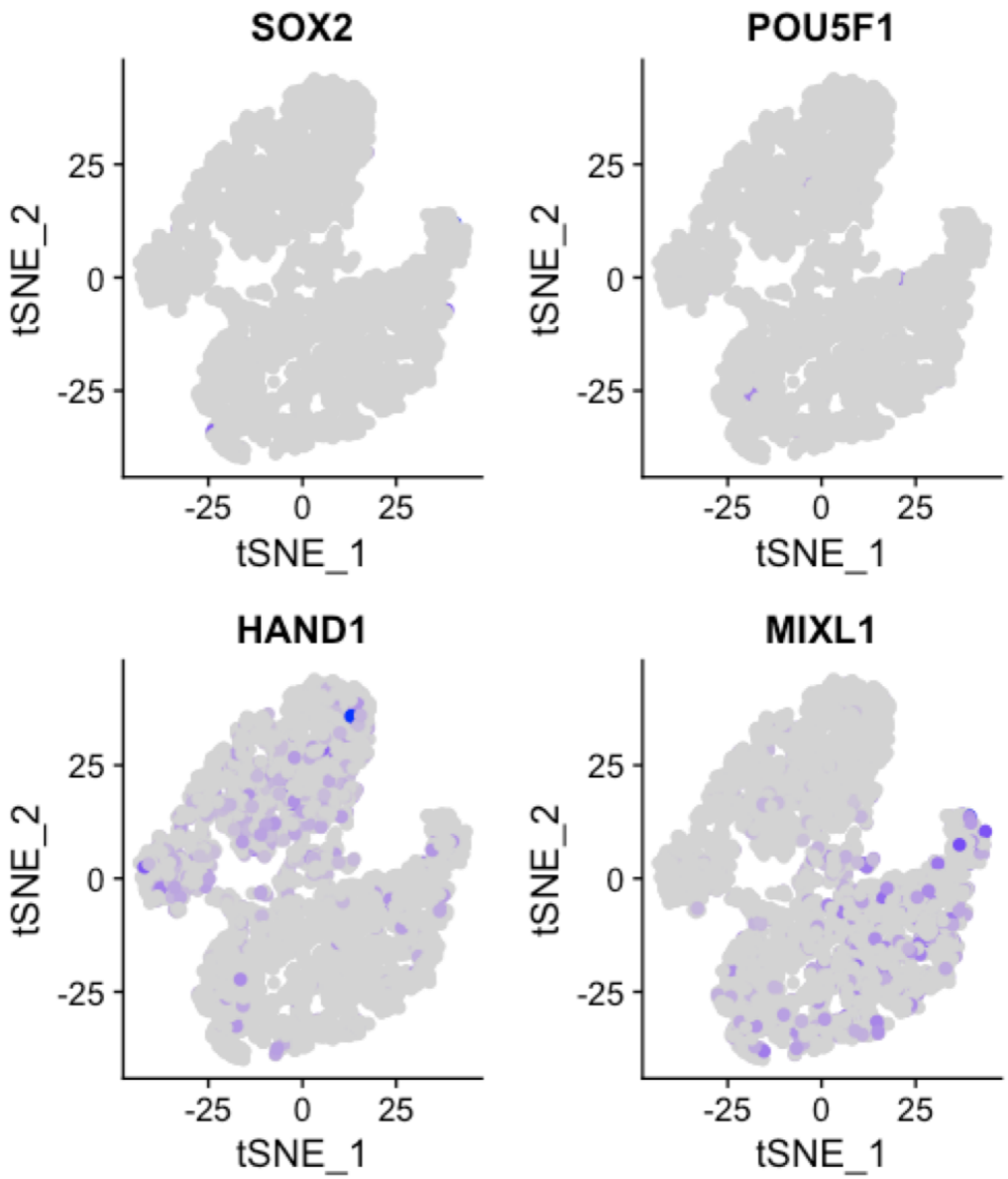
E

Day 12

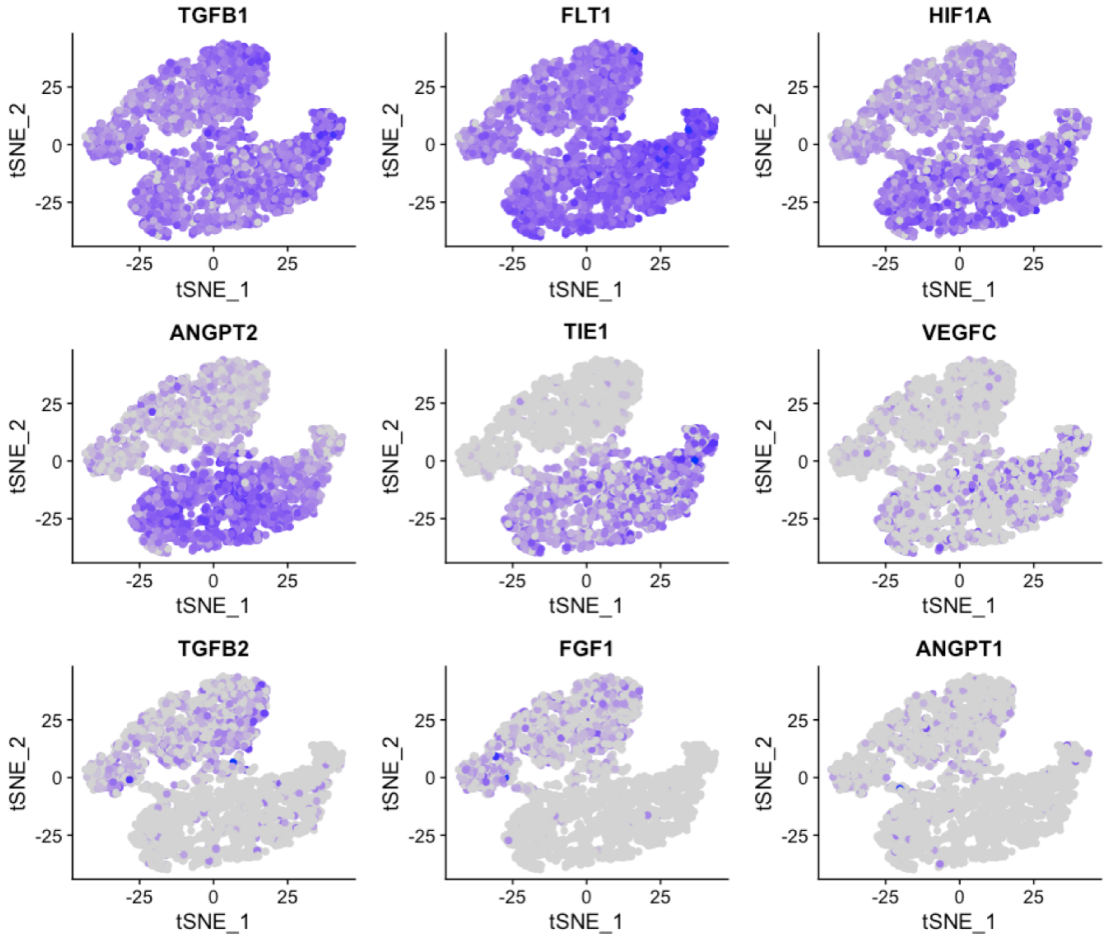


D12-A D12-B D12-C

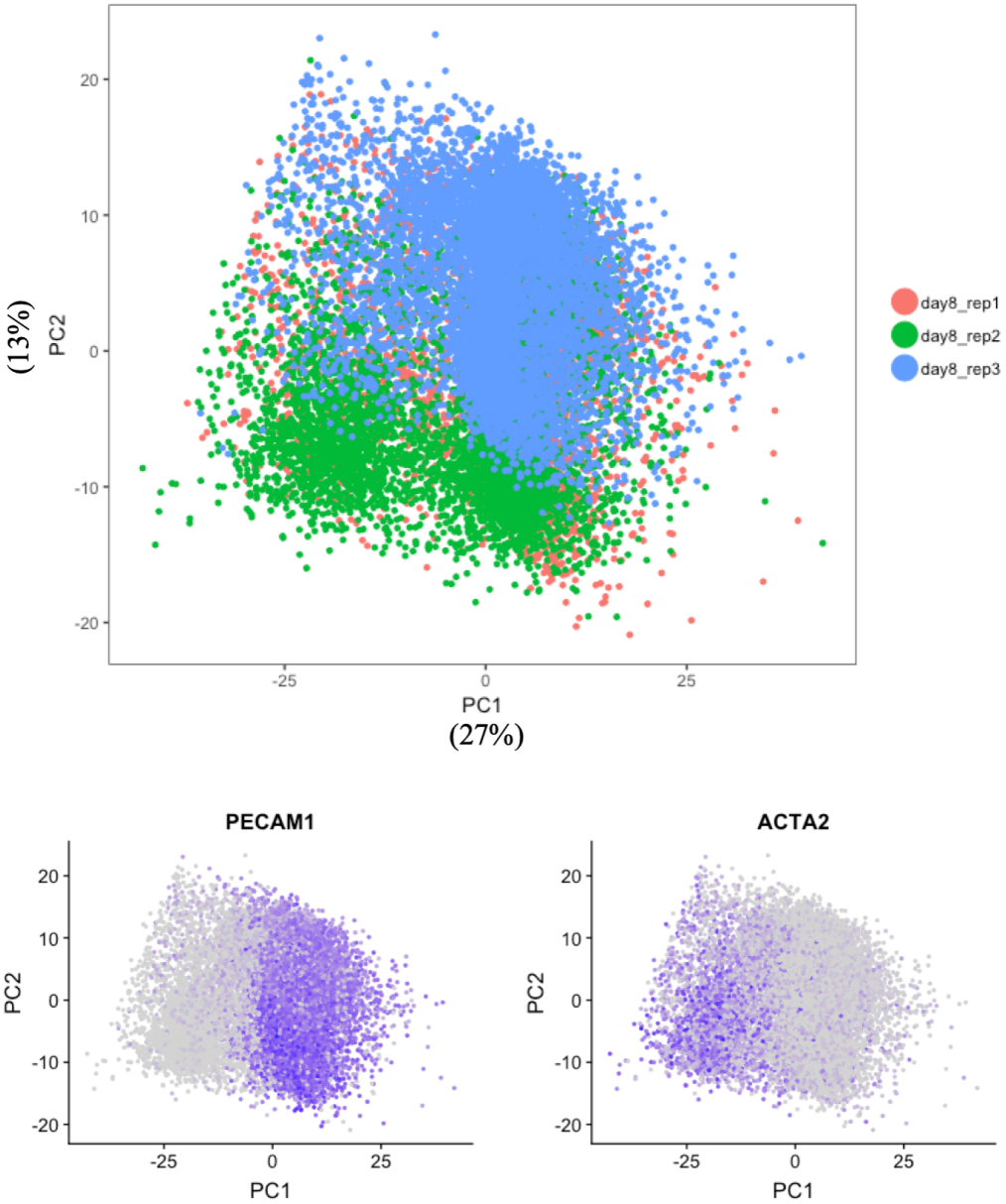
Supplementary Figure 2.2



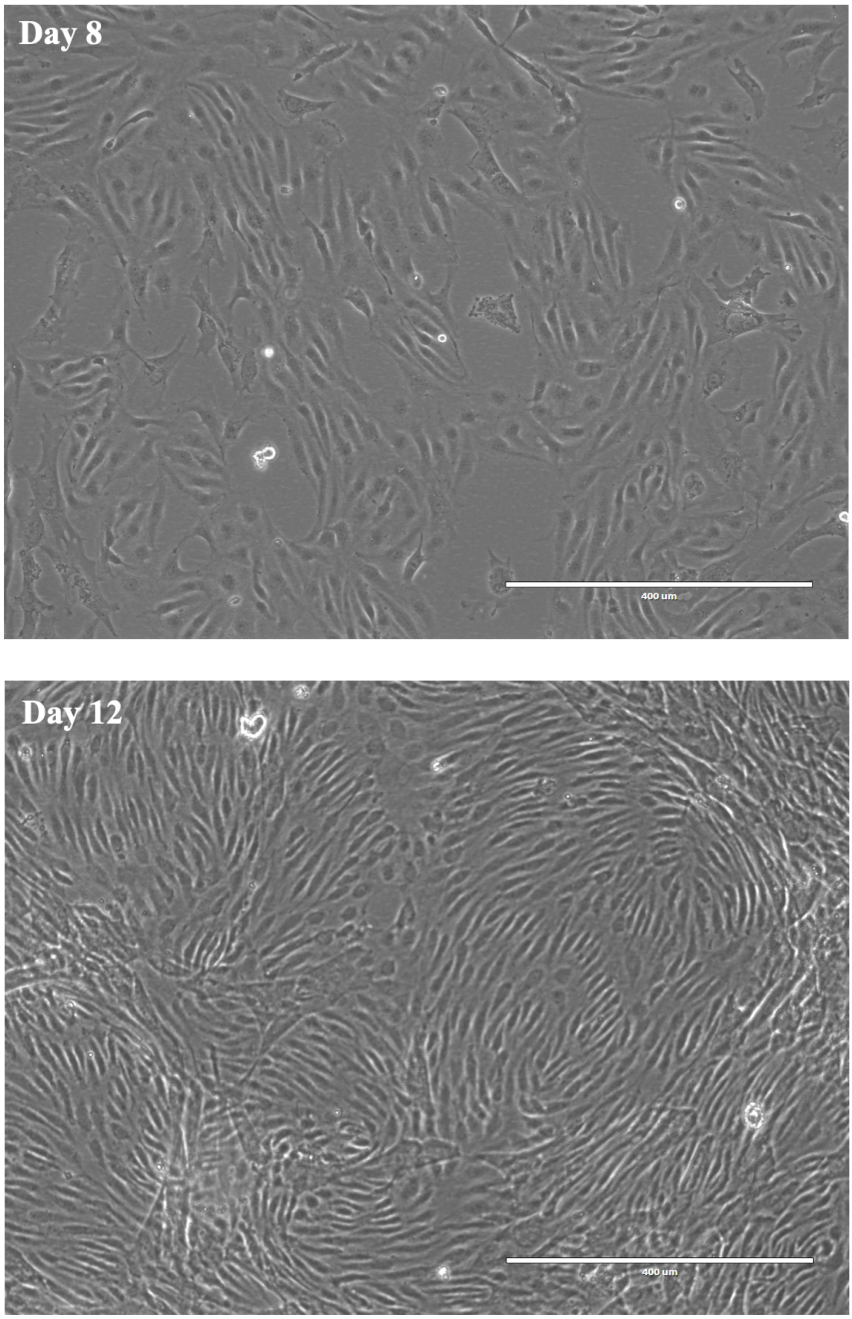
Supplementary Figure 2.3



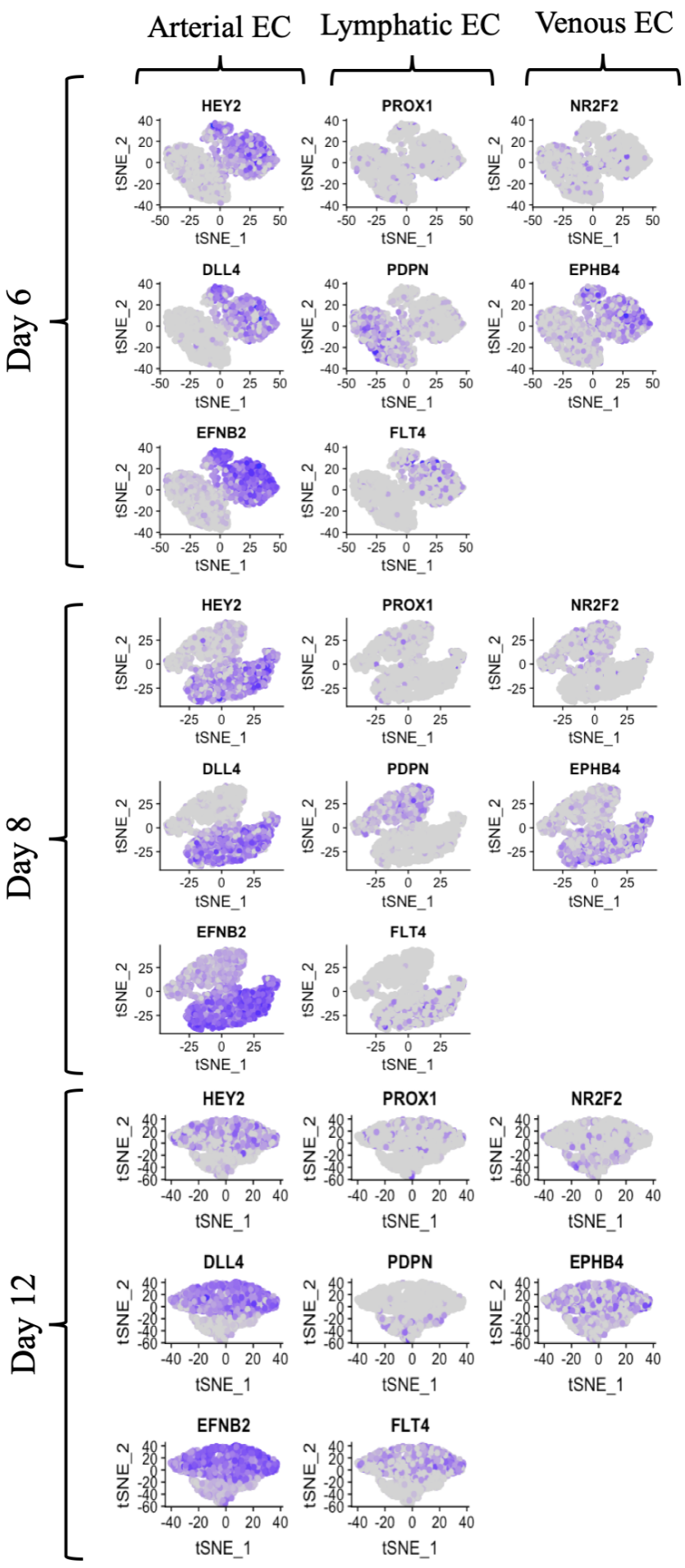
Supplementary Figure 2.4



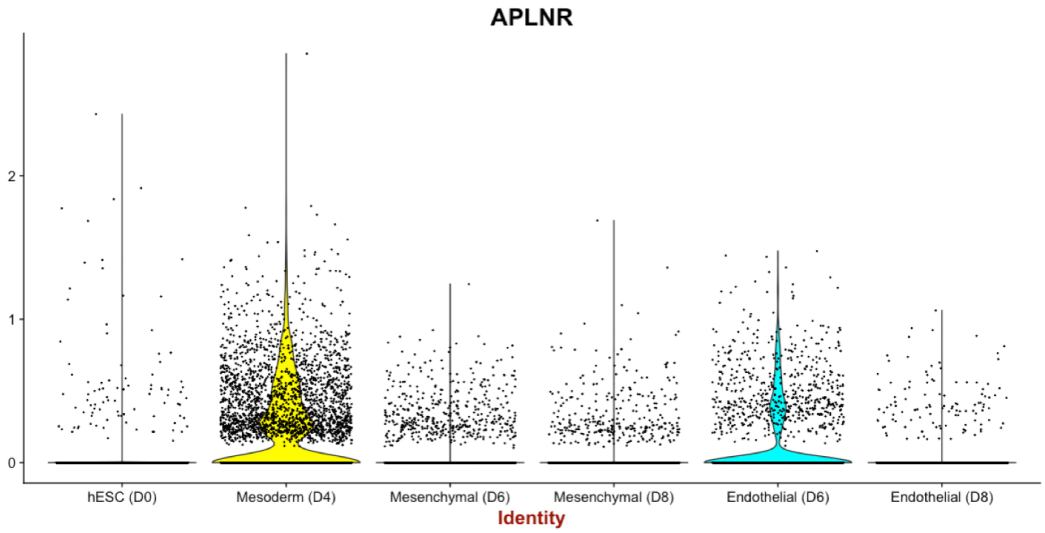
Supplementary Figure 2.5



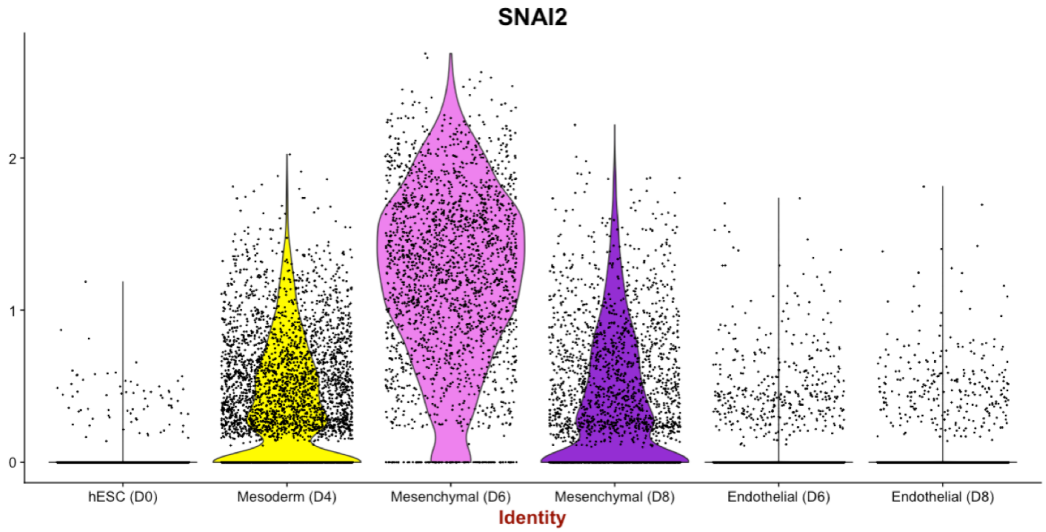
Supplementary Figure 2.6



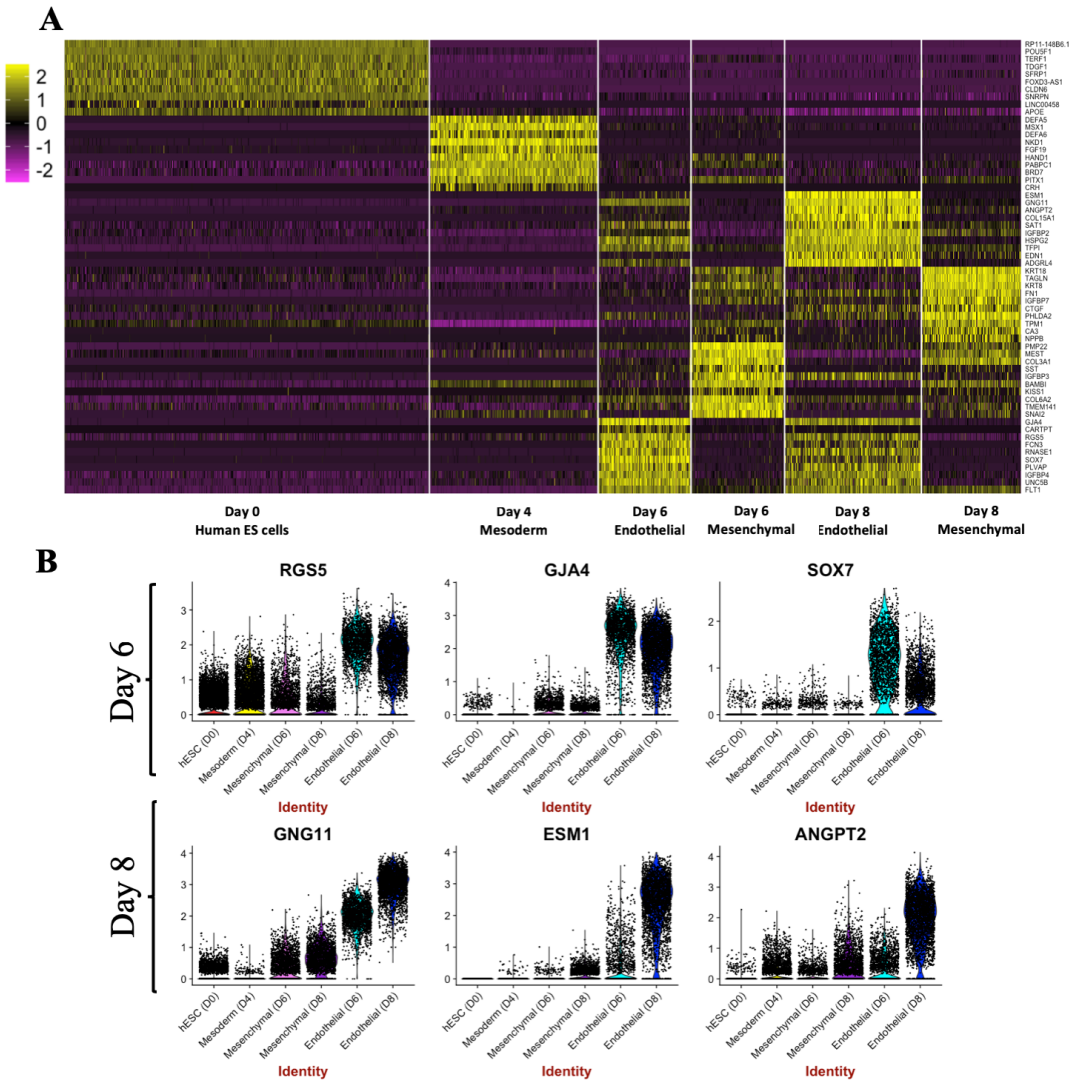
Supplementary Figure 2.7



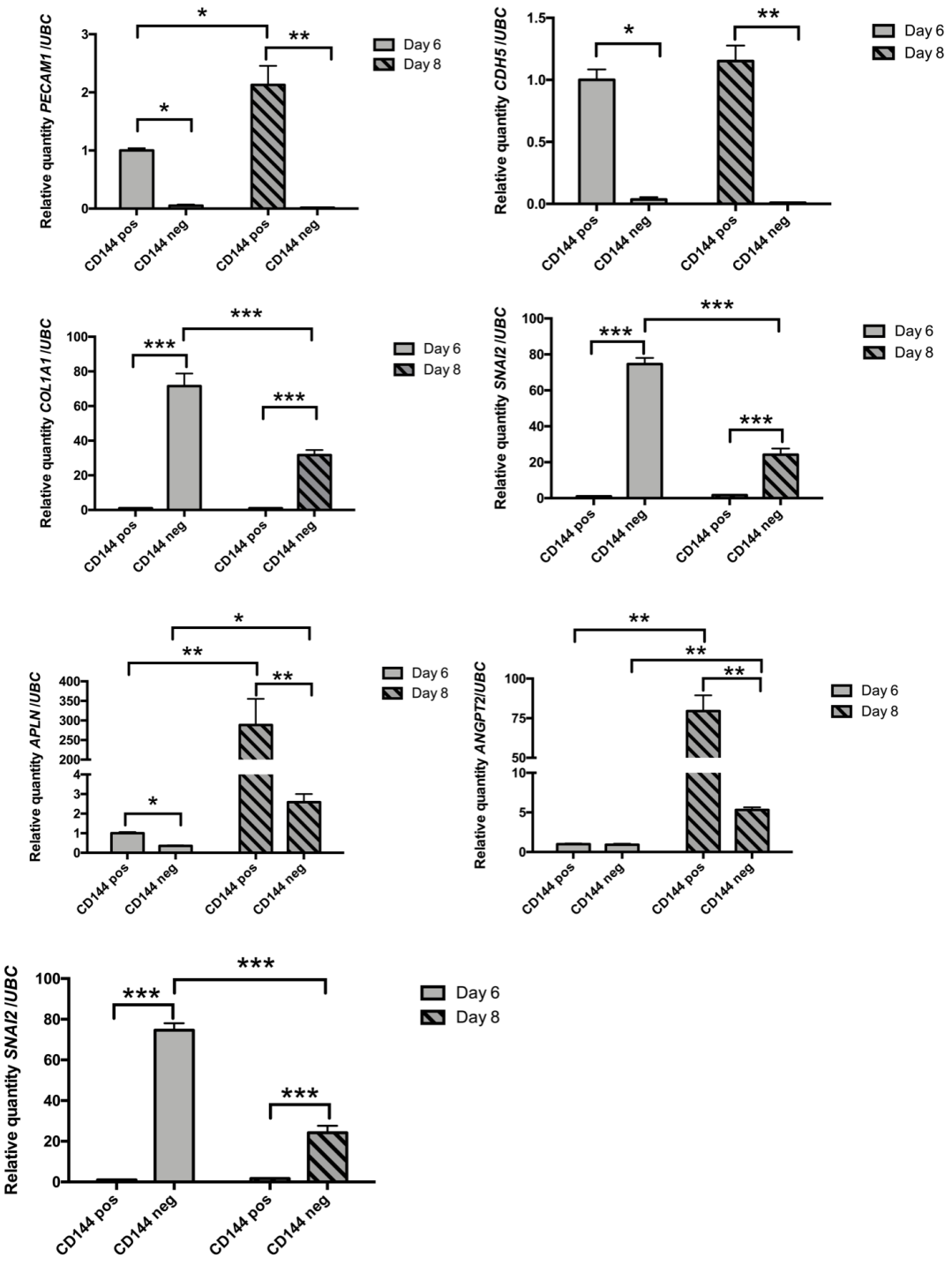
Supplementary Figure 2.8



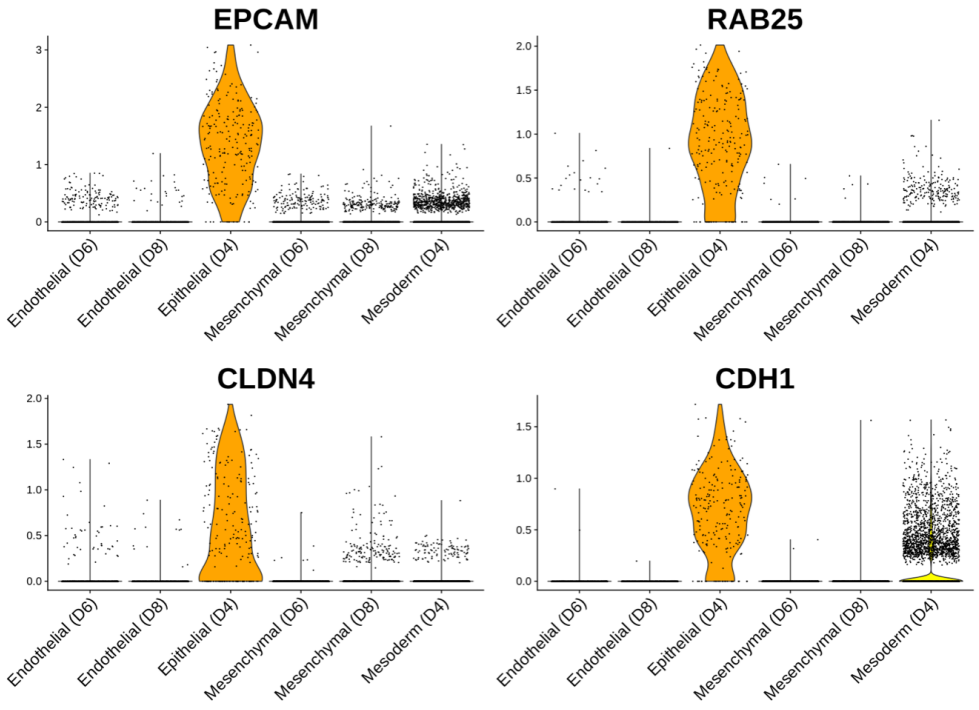
Supplementary Figure 2.9



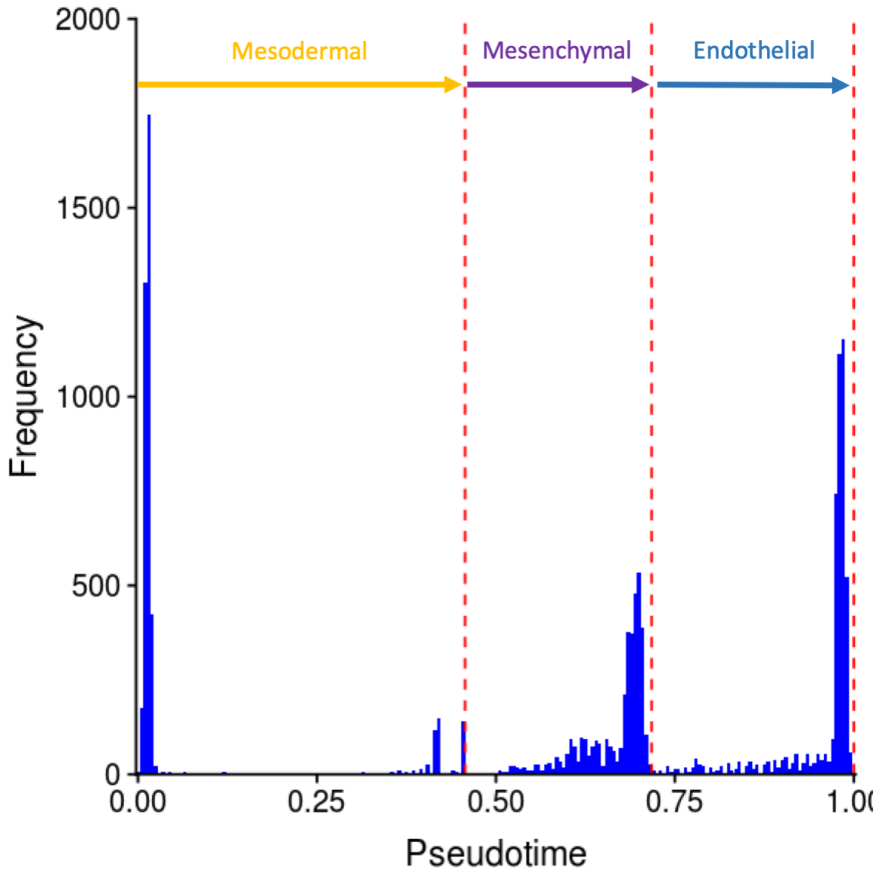
Supplementary Figure 2.10



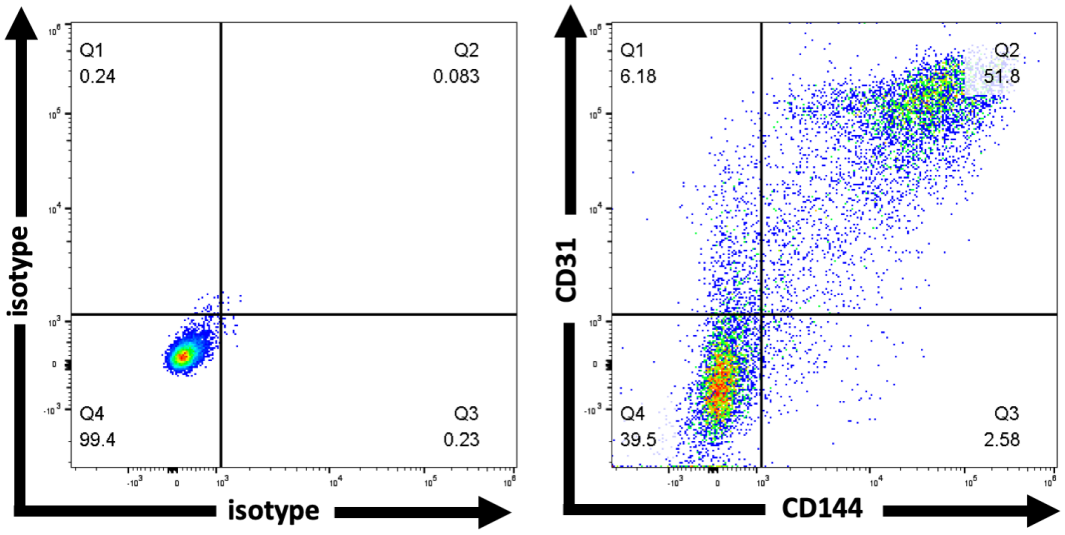
Supplementary Figure 2.11



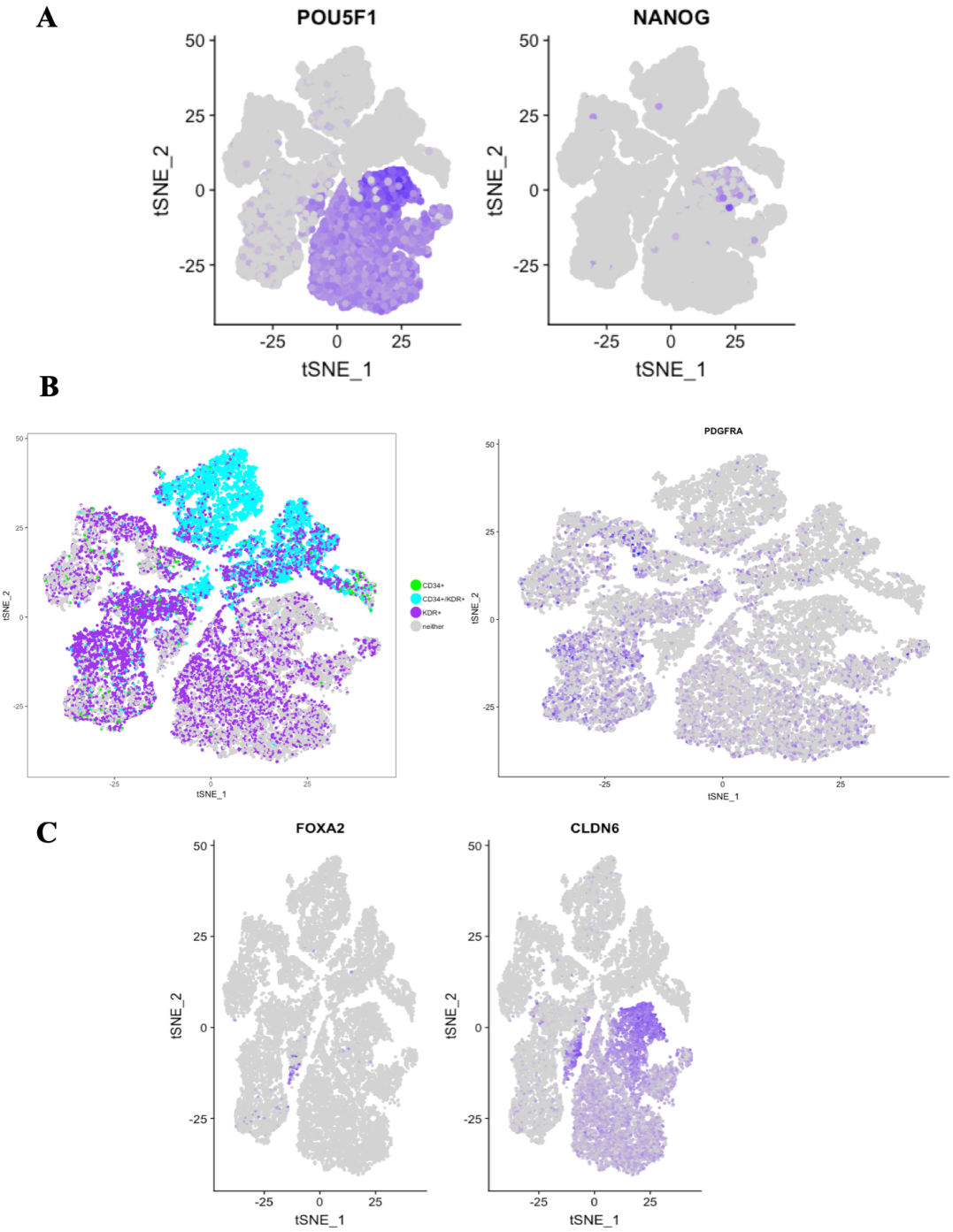
Supplementary Figure 2.12



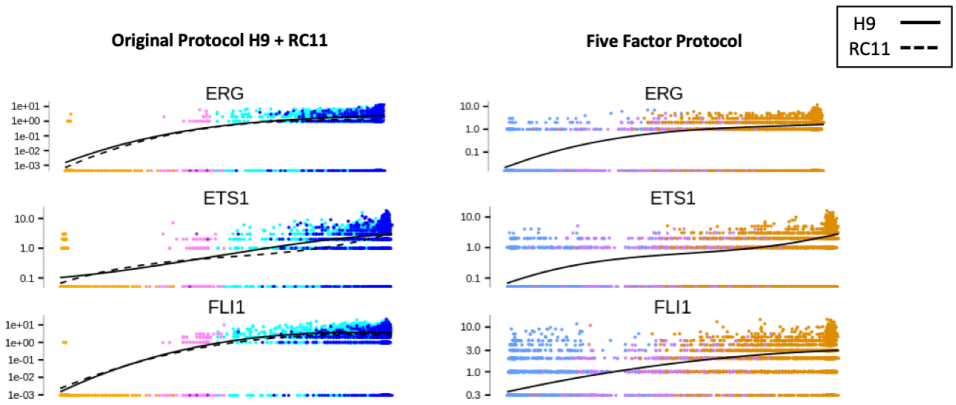
Supplementary Figure 2.13



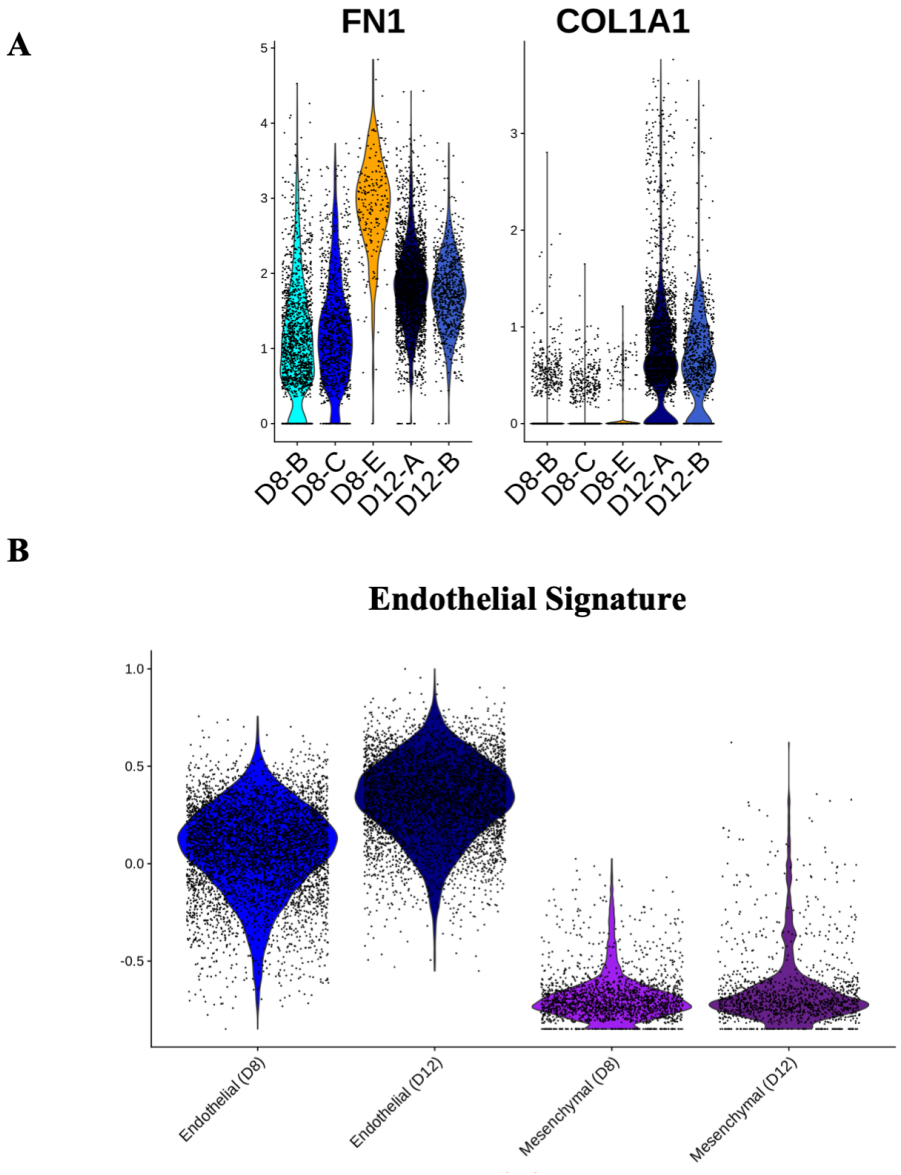
Supplementary Figure 2.14



Supplementary Figure 2.15



Supplementary Figure 2.16



2.12 Chapter Conclusion

Characterisation of two separate hESC-EC differentiation protocols using scRNA-seq revealed the existence of mesodermal cells prior to the emergence of endothelial and non-endothelial cell populations. Despite differences in the degree of cellular heterogeneity between protocols, common signatures of gene expression during differentiation were identified. These common signatures included many of those transcription factors with a previously identified role in murine vasculogenesis, including TFs belonging to the ETS family, thus suggesting a high level of conservation between human and murine EC development.

Transcriptional comparison of hESC-EC to mature endothelial cells used in primary cell culture or to those freshly isolated from adult or fetal tissue revealed a transcriptional distinction, despite hESC-EC expressing known EC markers. This suggests hESC-EC may represent an early immature EC population which may be amenable to undergoing further specification to distinct EC subtypes in response to specific cues.

Expression of angiogenic related genes was identified within both the mesenchymal and endothelial populations that comprise the day 8 hESC-ECP. This indicates that both populations may contribute via a paracrine effect to induce angiogenesis following injection into ischaemic tissue.

Chapter 3 - Mapping the developing human cardiac endothelium at single cell resolution identifies MECOM as a regulator of arteriovenous identity

3.1 Chapter Introduction

The objective of this chapter was to apply scRNA-seq to map the transcriptional landscape of the endothelium in the developing human heart. Enriching for EC by FACS prior to performing scRNA-seq was needed to allow a more comprehensive characterisation of endothelial heterogeneity. This was not possible using existing scRNA-seq datasets from the human fetal heart which were limited by a comparatively low number of EC. Subsequent application of differential expression and gene regulatory network analysis methods was required to gain insight into the putative identity of identified subpopulations, as well as identifying the transcription factors most likely responsible for establishing the associated transcriptional signature. This information is likely to be important for informing future strategies to guide the further specification of hPSC-EC *in vitro*.

Current understanding of the processes which occur during coronary vascular EC formation and specification has largely originated from lineage tracing studies in murine models. Therefore, the use of trajectory inference methods to identify the transcriptional changes which occur during these dynamical processes in the human fetal cardiac endothelium provides much needed human relevance to our understanding of these events. This includes the arterial specification of the immature coronary vascular network which has been described to occur during murine cardiac development. Comparison of the transcriptional signatures between human and mouse may identify novel conserved regulators of arterial EC specification.

3.2 Title Page

Mapping the developing human cardiac endothelium at single cell resolution identifies MECOM as a regulator of arteriovenous identity

McCracken IR, Dobie R, Henderson NC, Mountford JC, Riley PR, Ponting CP, Smart N, Brittan M, Baker AH; Manuscript under review.

Author contribution statement

I.R.M. carried out fetal dissection and subsequent digestion of heart tissue prior to isolation of EC by FACS. FACS was conducted by QMRI and CRM flow cytometry facilities. 10X library preparation was performed either by R.D or the CRM flow cytometry facility. scRNA-seq data analysis was performed by I.R.M. All hESC-EC work was carried out by I.R.M. Validation of findings by RT-qPCR and western blot was performed by I.R.M. I.R.M. was responsible for writing the manuscript with all authors involved in discussion of results and contributing to the production of the final version of the manuscript.

3.3 Abstract

Aims: Coronary vasculature formation is a critical event during cardiac development, essential for heart function throughout perinatal and adult life. However, current understanding of coronary vascular development has largely been derived from transgenic mouse models. The aim of this study was to characterise the transcriptome of the human fetal cardiac endothelium using single-cell RNA sequencing (scRNA-seq) to provide critical new insights into the cellular heterogeneity and transcriptional dynamics that underpin endothelial specification within the vasculature of the developing heart.

Methods and Results: We acquired scRNA-seq data of over 10,000 fetal cardiac endothelial cells (EC), revealing divergent EC subtypes including endocardial, capillary, venous, arterial, and lymphatic populations. Gene regulatory network analyses predicted roles for *SMAD1* and *MECOM* in determining the identity of capillary and arterial populations, respectively. Trajectory inference analysis suggested an endocardial contribution to the coronary vasculature and subsequent arterialisation of capillary endothelium accompanied by increasing *MECOM* expression. Comparative analysis of equivalent data from murine cardiac development demonstrated that transcriptional signatures defining endothelial subpopulations are largely conserved between human and mouse. Furthermore, we revealed that knockdown of MECOM in human embryonic stem cell-derived EC (hESC-EC) resulted in an increase in venous EC marker expression, validating our prediction of its role in arterial EC identity.

Conclusions: scRNA-seq of the human fetal cardiac endothelium identified distinct EC populations. A predicted endocardial contribution to the developing coronary vasculature was identified, as well as subsequent arterial specification of capillary EC. Loss of *MECOM* in hESC-EC increased venous EC marker expression, suggesting a role in maintaining arterial EC identity.

Keywords: Human cardiac development, single-cell RNA sequencing, endothelial heterogeneity, coronary vasculature formation, MECOM, vascular regeneration.

3.4 Translational Perspective

Endogenous blood vessel formation in the adult heart following myocardial infarction is insufficient to support adequate survival of the remaining myocardium, often ultimately leading to heart failure. Improved understanding of the mechanisms regulating human coronary vessel formation is required to inform therapeutic strategies to reactivate developmental pathways promoting therapeutic angiogenesis in patients. We applied scRNA-seq to map the transcriptome of the endothelium of the developing human heart. We identified novel transcriptional signatures underlying the cellular heterogeneity and dynamic changes occurring within the developing cardiac endothelium. This included identifying and validating MECOM as a novel regulator of arterial EC identity which may serve as a target for therapeutic neovascularization.

3.5 Introduction

While the formation and homeostasis of the coronary vasculature is essential for heart muscle function, the molecular mechanisms underlying coronary vascular development remain incompletely understood. Previous studies using lineage tracing tools in mouse have provided much needed insight into these mechanisms, including identifying the endocardium and sinus venosus (SV) as the two major sources of coronary vascular endothelium during cardiac development[300, 301]. A third source, the proepicardium, was previously proposed to contribute a minor population of coronary endothelial cells (ECs)[302, 303], although this notion has recently been challenged[304]. Following the formation of the primitive coronary vascular plexus from these sources and onset of blood flow, subsequent remodelling occurs giving rise to the distinct EC populations present in the mature vascular bed of the fully developed heart[10, 305]. Recent studies have elegantly mapped the remodelling of the immature coronary EC plexus in mouse cardiac development, including identification of a role for the transcription factor Dach1 in potentiating developmental arterial remodelling[305-307]. However, given that these advances in our understanding of coronary vascular development primarily originate from murine lineage tracing studies, the relevance of these findings in human cardiac development remains largely unknown.

Advances in single-cell RNA sequencing (scRNA-seq) have been instrumental in enhancing our understanding of embryonic development, permitting the objective mapping of underlying transcriptional changes at single-cell

resolution. In addition, improvements in high throughput scRNA-seq platforms have facilitated the characterisation of tens of thousands of cells in parallel, thus allowing for 'atlas' studies to map the gene expression profile of entire organs during embryogenesis[308]. In recent years, such scRNA-seq studies have mapped the transcriptional profile of both murine and human heart development[196, 309-315]. In the study by Cui et al[196], scRNA-seq was conducted using cells isolated from specific regions of 18 human fetal hearts, ranging from 5 – 24 weeks gestation. Subsequent dimensionality reduction and clustering analysis revealed an EC cluster of 595 cells characterised by expression of endothelial markers such as *PECAM1*[196]. Similarly, a clear EC population was identified in a study from Suryawanshi et al[310] in which cells isolated from three healthy human fetal hearts (19-22 weeks) were processed using scRNA-seq. Both studies mapped the expression of known EC marker subtypes to allow annotation of clusters corresponding to endocardium, coronary vascular EC, and valvular EC. Nevertheless, the relatively low numbers of EC in these datasets prevented further characterisation of cardiac EC subtypes, including the identification of distinct arterial, venous, capillary, and lymphatic populations. In addition, these low EC numbers also prevented the application of methods to infer the dynamic cellular changes accompanying cardiac EC development.

While scRNA-seq studies of the developing mouse heart yielded large numbers of EC in their datasets[309, 313-315], their analysis focused on other cell types such as the cardiac conduction system, with minimal interpretation of EC

heterogeneity and potential function. These included a study from Goodyer et al[312] which analysed distinct vascular EC and endocardial cell populations from E16.5 mouse hearts.

In this study we used scRNA-seq to comprehensively map the transcriptional signature of over 10,000 human fetal cardiac ECs isolated by fluorescence activated cell sorting (FACS) from two human fetal hearts at 13- and 14-weeks' gestation. Unsupervised clustering, gene regulatory analysis, and trajectory inference methods revealed the transcriptional profile of heterogeneous EC populations and predicted dynamic cellular changes including arterial EC specification. In addition, we functionally validated MECOM as a regulator of arterial EC identity, thereby demonstrating the suitability of our novel scRNA-seq dataset to make *in silico* predictions, capable of informing future strategies to guide endothelial identity. Collectively, findings from this study complement and expand upon knowledge previously obtained from murine development, bringing insights into human EC heterogeneity and pathways determining specification of subpopulations that are essential for understanding human coronary vascular formation.

3.6 Methods

Tissue collection and study approval.

Human fetal cardiac tissue was acquired following elective termination of pregnancy. Informed written parental consent was obtained from all participants. Tissue was not collected in cases where termination of pregnancy was conducted due to an identified fetal or pregnancy abnormality. Ethical approval for the collection of fetal tissue was performed in accordance with all relevant guidelines and following study approval from the Lothian Research Ethics Committee (Study code: 08/1101/1) and the Research and Development Office (Study code: 2007/R/RM/10).

Isolation of fetal cardiac endothelial cells.

Cardiac endothelial cells were isolated from freshly collected human fetal heart tissue using a method adapted from van Beijnum et al[316]. Digestion was performed at 37°C using a digestion solution containing 9ml 0.1% collagenase II and 1ml of 2.5U/ml dispase. 75µl of 1mg/ml DNaseI was added following 20 minutes incubation prior to a further 15 minutes incubation at 37°C. Digestion was quenched by the addition of 10ml cold RPMI with 10% FCS and undigested clumps of tissue removed using a 100µM cell strainer. Red blood cell lysis was performed by incubating cells for 2 minutes in red blood cell lysis buffer at room temperature prior to neutralising with RPMI + 0.1% BSA. Cells were stained on ice for 45 minutes with APC anti-human CD31 and PE anti-human CD45 (Supplementary Table 3.1). CD31+ CD45- endothelial cells were

isolated by fluorescence activated cell sorting (FACS) with DAPI staining being used to allow exclusion of dead cells.

Single-cell RNA sequencing of fetal cardiac endothelial cells.

Sorted CD31+ CD45- endothelial cells were counted manually using a haemocytometer with trypan blue staining used to identify non-viable cells. Viability exceeded 85% for both samples. 8,000 cells were loaded onto the 10X Chromium controller and library construction conducted using the Single Cell 3' Reagent Kit (V3.1) in accordance with the manufacturer's instructions. Libraries were sequenced using the Illumina NovaSeq 6000 platform.

Single-cell RNA sequencing data analysis.

Raw sequencing data was processed using the 10X CellRanger pipeline (Version 3.1.0.) aligning reads to the GRCh38-3.0.0 genome reference. Barcodes corresponding to cells were distinguished from those corresponding to empty droplets using both the DropUtils package[317] and the default cell calling method applied within the CellRanger pipeline. Cells with a total UMI count exceeding 3 median absolute deviations (MADs) from the median value were removed from downstream analysis using the R Scater package[318]. Similarly, cells with a high proportion of counts from mitochondrial genes (>3MADs) or with a low total gene count (<2 MADs) were also excluded. Data normalisation was performed using the MultiBatchNormalisation method[319] prior to merging datasets. Normalised count data was then scaled, and principal component analysis (PCA) applied using genes with the most

variable expression across the combined dataset[320]. Following batch correction using Harmony[321], non-supervised clustering was performed, and data visualised using Uniform Manifold Approximation and Projection (UMAP)[320]. Significantly differentially expressed genes (DEGs) within individual clusters were identified using the Wilcoxon signed rank test (Bonferroni corrected p value <0.05) and a minimum $\log_e$(fold change) threshold of 0.3[320]. Additionally, only DEGs expressed in more than 30% of cells within their corresponding cluster were retained for further analysis. Enriched metagene signatures were identified using the R package SCRAT v1.0.0[194]. Gene regulatory analysis was performed using the standard R SCENIC (Single Cell rEgulatory Network Inference and Clustering) workflow[322]. RNA velocity analysis was conducted using the python package scVelo[323] with the stochastic model being applied to predict the direction and magnitude of cellular dynamics. Trajectory inference tool Slingshot was performed using the standard workflow[324]. Genes significantly differentially expressed over pseudotime were identified using the TradeSeq package[325] with the top 2,000 most variably expressed genes in the dataset being used to fit the negative binomial generalised additive model (NB-GAM).

Human embryonic stem cell derived endothelial cell (hESC-EC) differentiation and siRNA mediated MECOM knockdown.

Human ESC lines were used in accordance with the UK Stem Cell Bank Steering Committee guidelines (Project Approvals SCS11-51 and SCSC17-26). H9 hESC were differentiated to hESC-EC as previously described[241, 326].

Small interfering RNA (siRNA) -mediated knockdown of MECOM was performed using day 7 hESC-EC using predesigned siRNA at a final concentration of 5nM (Supplementary table 3.2). After 6 hours, transfection media was replaced with EGM-2 media supplemented with 1 % human AB serum and 50 ng/ml VEGF-A. At day 10, CD144+ hESC-EC were isolated by magnetic activated cell sorting (MACS) and cell pellets stored at -80°C for subsequent isolation of RNA and protein.

All further experimental details are included in the Supplementary Material.

3.7 Results

Identification of distinct cardiac endothelial populations

Single-cell RNA sequencing (scRNA-seq) was performed on CD31+ CD45- cardiac endothelial cells isolated by FACS from heart ventricular tissue obtained from two human fetuses at 13 and 14 weeks of gestation (Figures 3.1A and 3.1B). At this developmental stage all major structures in the heart have formed, including the coronary vasculature. However, studies from equivalent timepoints in murine development (E15.5.-E17.5) have revealed extensive remodelling occurring within the established coronary vasculature producing a mature vascular bed containing heterogeneous EC populations[10, 306, 327]. Following quality control, unsupervised clustering, and UMAP visualisation of transcriptomic data from 10,267 cells, 11 distinct clusters (numbered 0-10) were revealed, each with expression of typical pan-endothelial cell (EC) markers (Figures 3.1C; Supplementary Figure 3.1A).

Leukocyte marker *PTPRC* was absent across all clusters (Supplementary Figure 3.1A). Annotating cells by sample demonstrated successful integration of datasets, with each cluster containing cells from both samples (Supplementary Figure 3.1B).

Expression of *NPR3*, a known endocardial marker, was localised to cluster 4 (Figure 3.1D; Supplementary Figure 3.1C)[328]. Clusters 0, 1, 2, 5, and 6 were defined by expression of capillary marker *RGCC*[16], whilst arterial marker, *HEY1*[167], was expressed predominantly in clusters 2, 3, and 9 (Figure 3.1D; Supplementary Figure 3.1C). Metagene analysis revealed 5 major signatures (A-E) indicating 5 key populations within the data (Figure 3.1E). Signature A was enriched in clusters 2, 3, and 9, and included genes involved in arterial EC function such as *JAG1, DLL4,* and *HEY1* (Figure 3.1E). Furthermore, Gene Ontology (GO) term enrichment analysis using signature A genes identified 'artery morphogenesis' and 'positive regulation of Notch signalling pathway' as significantly enriched terms (Figure 3.1F). Notch signalling is known to be required for arterial EC specification[329].

Signature B was predominantly enriched in cluster 4 and contained known endocardial markers *CDH11* and *NPR3* (Figure 3.1E and 3.1F)[328, 330]. Lower levels of signature B enrichment were also observed in clusters 7, 8, and 9 (Figure 3.1E). Clusters 0, 1, and 5 were enriched for signature C which included capillary EC marker genes such as *CA4* and *RGCC* (Figure 3.1E; Supplementary Figure 3.1D)[16, 331]. Cluster 10 was enriched for signature D for

which GO term analysis returned terms relating to lymphatic EC (Figure 3.1E; Figure S 3.1D). Lymphatic EC (LEC) markers[2], *LYVE1, FLT4, PROX1*, and *PDPN*, were differentially expressed in cluster 10 (Supplementary Figure 3.1G). Signature E was selective for cluster 6 with GO term analysis identifying enriched terms relating to proliferation (Supplementary Figure 3.1E). In line with this, categorising cells according to their predicted cell cycle phase revealed that 77 % of cells in cluster 6 were predicted to be in the G2M phase of rapid growth, whereas 95 % of cells in cluster 5 were in S phase (Supplementary Figure 3.1F). The remaining cells in the dataset were predominantly in the G1 phase (79%), with only a small proportion predicted to be in G2M (4%) and S phases (17%) (Supplementary Figure 3.1F). For clusters not associated with a metagene signature, analysis of differentially expressed genes (DEGs) revealed enrichment of *NR2F2* in cluster 7, suggesting a venous EC identity (Supplementary Figure 3.1G)[12]. A valvular identity of cluster 8 was supported by its differential expression of *NFATC1* and *BMP4*, both with known roles in valvulogenesis (Supplementary Figure 3.1G)[332, 333]. Collectively, these analyses demonstrate that each major subtype of EC within the heart (endocardial, venous, lymphatic, capillary, arterial, proliferating, and valvular EC) is represented by one or more of the 10 clusters.

Gene regulatory network analysis of fetal cardiac endothelial cells

Global differential gene expression analysis revealed markers for subpopulations of arterial and capillary cardiac EC (Figure 3.2A). Notably, cluster 2, an arterial EC population, is more closely correlated with capillary

EC clusters (0 and 1) than with the other two minor arterial clusters (3 and 9) (Supplementary Figure 3.2A). Together with the co-expression of both arterial and capillary markers, this suggested that cluster 2 represents an arterial microvascular population. In contrast, arterial clusters 3 and 9 differentially expressed genes associated with ECM organisation (*FBLN5, ELN, FBN1*) and shear stress (*KLF4)* suggesting a macrovascular identity (Supplementary Figures 3.2B and 3.2C).

Expression of fatty acid translocase encoding *CD36* was absent from macrovascular and endocardial populations, corresponding with a previous report of microvasculature-restricted expression (Figures 3.2A and 3.2B)[249]. Differential gene expression analysis revealed heterogeneity between the two major capillary clusters, 0 and 1. Cluster 0 was defined by differential expression of aminotransferase encoding gene, *INMT*, whilst *KIT* expression was enriched in cluster 1 (Figures 3.2A and 3.2B). Selective expression of *KIT* (also known as *C-KIT*) in cluster 1 was accompanied by upregulated expression of the transcription factor (TF) *SMAD1* (Supplementary Figures 3.2D and 3.2E).

Gene regulatory network (GRN) analysis was applied to identify gene modules, known as regulons, predicted to be controlled by an individual TF, giving insight into the likely transcriptional regulators of EC heterogeneity. Visualisation of differentially expressed TFs and their predicted targets within the GRN largely recapitulated the data structure observed following

unsupervised clustering (Figure 3.2C). Genes differentially expressed in LEC (cluster 10) localised together in the GRN and included TFs such as *PROX1*, a known master regulator of LEC identity[208], as well as *HOXD9* and *TBX1* (Figure 3.2C). A set of differentially expressed endocardial TFs (cluster 4) was evident, including *GATA6, MEIS2*, and *FOXC1*, as well as *GATA4*, known to be implicated in endocardial cushion development[334, 335]. *MECOM* and *MAFF* were located amongst known regulators of arterial EC specification such as *HEY1* and *SOX17* (Figure 3.2C) [138, 167]. A distinct cluster of genes differentially expressed in *KIT1*$^{+}$ capillary EC (cluster 1) included *SOX4* and *SMAD1* (Figure 3.2C). Enrichment of *SOX4* and *SMAD1* regulons was also observed in *KIT1*+ capillary EC (Figure 3.2D).

Trajectory analysis predicts an endocardial contribution to the developing coronary vasculature and potential regulators of subsequent Arterial Specification

Several recent studies using murine models of coronary vascular development have provided insight into the origin of the coronary endothelium and the dynamic changes that occur during its subsequent remodelling[10, 336]. Endocardial derived vessels vascularise the heart from the inside-out contributing to vessels of the interventricular septum and inner myocardial wall[300, 337, 338]. Conversely, sinus venosus (SV) derived vessels populate the outer ventricular free walls of the heart from the outside-in[328, 337, 338]. Following the formation of the primitive coronary vascular plexus, EC undergo further remodelling to form a functional network of veins, arteries, and capillaries[306].

We used trajectory inference methods to determine whether these processes could be identified during human cardiac development and to characterise their accompanying transcriptional changes. Given that these dynamic changes are known to originate from microvascular EC, we excluded the two previously identified arterial macrovascular clusters (clusters 3 and 9) from the dataset and performed secondary clustering of the remaining cells. The distinct LEC cluster was also excluded prior to re-clustering. The same marker genes used for annotating the complete dataset were used for the annotation of re-clustered data (Supplementary Figures 3.3A and 3.3B).

RNA velocity analysis[323], which utilises the ratio of spliced to unspliced transcripts to infer the direction and magnitude of cellular transitions, was first used to gain an overview of the pseudotemporal dynamics of the fetal cardiac endothelium (Figure 3.3A). We identified a proportion of the endocardial cluster with velocity vectors indicating a probable transition towards a venous identity (Figure 3.3A). Evidence for this transition was further supported by venous EC associated genes such as *PLVAP* and *NR2F2*, having positive residuals/velocities in endocardial cells (Supplementary Figure 3.3C). In turn, venous EC were subsequently predicted to transition to *INMT*+ capillary EC. This predicted transition of endocardium to coronary vascular EC concurs with studies that identified the endocardium as a significant source of EC for the coronary vasculature, including at later stages in cardiac development[339]. Velocity analysis also predicted a likely transition of both capillary EC

populations towards an arterial EC fate (Figure 3.3A). This is an agreement with previous reports of developmental arterial remodelling in mouse[305-307].

In addition to the RNA velocity analysis, we also independently applied the trajectory inference tool Slingshot[324], which yielded a comparable interpretation (Figure 3.3B). Identification of the top 200 genes with most variable expression over pseudotime revealed 4 temporal patterns of expression, arranged in modules 1 - 4 (Figures 3.3C and 3.3D). Module 2 genes were expressed early in pseudotime with their reduction in expression occurring in conjunction with the loss of endocardial identity (Figure 3.3D). These included known markers of endocardium such as *CDH11* and *NPR3* as well as the TFs *DKK3* and *GATA6* (Figure 3.3E and Supplementary Figure 3.3D). TFs *CEBPD* and *FOS* were identified in module 1 along with *NR2F2*, a known regulator of venous EC specification[12] (Supplementary Figure 3.3D). Module 4 genes demonstrated peak expression within *IMNT*+ capillary EC and included TFs *TCF15* and *MEOX1*. Interestingly, despite not being identified within module 4, expression of *DACH1* was found to peak within the *IMNT*$^{+}$ capillary cluster before gradually decreasing again within the arterial population (Supplementary Figure 3.3E). Previous studies have identified *Dach1* as a driver of developmental arterial remodelling in murine cardiac development[306, 307].

The predicted transition of capillary EC to arterial EC was defined by increased expression of module 3 genes. This included *HEY1*, known to mediate arterial

EC specification[167]. Interestingly, module 3 also contained the TF *MECOM*, earlier predicted by GRN analysis to underlie arterial EC identity (Figures 3.2C and 3.3C-E; Supplementary Figure 3.3D).

Comparison with murine coronary developmental gene expression reveals conserved markers of cardiac endothelial cell populations

Our current understanding of cardiac vascular development is derived predominantly from murine models. Consequently, we next compared the transcriptional profiles of developing fetal human and embryonic mouse cardiac EC. For this comparison, a publicly available mouse embryonic heart scRNA-seq dataset[312] was used due to its good representation of cardiac EC and because its embryonic stage (E16.5) corresponded with the later developmental stage of our human fetal heart data (13-14 weeks)[327].

Dimensionality reduction revealed successful integration of mouse and human cardiac EC (Supplementary Figure 3.4A). Unlike our observation in the human heart, no distinct populations of *KIT*+ or *INMT*+ capillary populations were observed in the mouse data (Supplementary Figure 3.4B). Clusters were therefore merged to represent the major subtypes of EC within the heart (endocardial, venous, lymphatic, capillary, arterial, proliferating, and valvular EC). Genes found to be amongst the most significantly differentially expressed in the same population in both human and mouse were classified as conserved markers (Figure 3.4A). *MECOM* and *UNC5B* were among genes with enriched arterial EC expression in both species (Figures 3.4A-B). In agreement with

previous findings in mouse[328], *NPR3* expression was highly specific to the endocardial population with minimal expression in valvular EC[328]. A lack of clear conserved venous EC markers was observed, with partial overlap of DEG in some lymphatic and endocardial populations (Figures 3.4A-B). Interestingly, whilst the known LEC TF, *PROX1*, was expressed in both valvular EC and LEC in both species, *PTX3* and a LEC marker, *LYVE1*, were found to be highly LEC specific (Figures 3.4A-B). Species-specific markers were also identified for each EC dataset (Figure 3.4A and Supplementary Figure 3.4C). A human endocardial marker, *NPCC*, described elsewhere as specifically defining human fetal cardiac endocardium[310], was not enriched in the corresponding mouse population (Figure 3.4C).

MECOM is required in arterial-like hESC-EC to suppress venous EC identity

Given the *in-silico* predictions of a role for MECOM in arterial fate and enriched MECOM expression in arterial EC for both human and mouse, suggesting an evolutionarily conserved role in arterial EC, we sought to validate its role in determining human arterial EC identity. Our previous scRNA-seq based characterisation of our 8 day hESC-EC differentiation protocol demonstrated its suitability as an *in vitro* model of human EC development[326]. Additionally, we determined that after acquisition of an early EC identity by day 6, hESC-EC assume a clear arterial-like EC transcriptional signature by day 8, characterised by expression of arterial markers such as *SOX17* and *DLL4* (Figure 3.5A). Expression of venous (*NR2F2, EPHB4*) and lymphatic markers

(*PROX1*) in hESC-EC was low by day 8 of the differentiation (Figure 3.5A). Notably, in agreement with its arterial EC specificity, *MECOM* was specifically expressed in hESC-EC at days 6 and 8 of differentiation (Figure 3.5A).

Using hESC-EC as a developmental model for arterial EC specification, we next determined whether small interfering RNA (siRNA)-mediated MECOM knockdown in hESC-EC resulted in changes to their arteriovenous identity (Figure 3.5B). Significant knockdown (>50%) of MECOM was observed at the RNA and protein level in hESC-EC 72 hours after siRNA transfection (Figures 3.5C-D; Supplementary Figure 3.5A). Knockdown of MECOM did not result in significant changes to the expression of arterial EC markers, including *DLL4, HEY1, JAG1,* and *JAG2* (Figure 3.5E; Supplementary Figure 3.5B). However, increased expression of venous markers *NR2F2* and *EPBH4* was observed following knockdown of MECOM (Figure 3.5E). Upregulation (6.13-fold change) of NR2F2 (also known as COUP-TFII) was also validated at the protein level (Figure 3.5F). Importantly, reduction of MECOM did not result in altered expression of the pan-endothelial marker, *CDH5,* suggesting that the changes observed to arteriovenous markers are not due to a loss of general EC identity (Supplementary Figure 3.5B).

3.8 Discussion

In this study we comprehensively mapped the transcriptional landscape of the developing human fetal heart endothelium using scRNA-seq. In contrast to previous studies, isolation of fetal cardiac EC by FACS prior to performing scRNA-seq empowered this study to identify the full extent of EC heterogeneity. This included identifying distinct endocardial, valvular, venous, capillary, and arterial EC populations each expressing a separate transcriptional signature.

Gene regulatory network analysis identified the TFs most likely responsible for establishing the observed endothelial cell heterogeneity. Application of trajectory inference methods to microvascular ECs was used to map the cellular dynamics accompanying coronary vascular EC development. This revealed a small proportion of endocardial cells that appeared to transition to a vascular EC identity *via* a venous EC population. In addition, capillary EC were predicted to be undergoing specification to assume an arterial EC identity, defined by increasing expression of the TF, *MECOM*. Comparison of our human fetal heart EC data with E16.5 murine cardiac EC scRNA-seq data demonstrated the existence of several conserved, as well as species-specific, markers for each of the major cardiac EC populations. This included identifying *NPR3* and *MECOM* as conserved markers of endocardial and arterial populations, respectively. Finally, we demonstrated that loss of MECOM in arterial-like hESC-EC resulted in the specific increased expression of venous EC markers, suggesting a function to repress venous identity in arterial EC.

In contrast to the capillary EC cluster defined by differential expression of methyltransferase encoding gene, *INMT*, gene regulatory network analysis revealed several regulons enriched within the *KIT*+ capillary population. This included the SMAD1 regulon. BMP/SMAD1 signalling has been demonstrated to promote angiogenesis whilst KIT/C-KIT has been shown to mediate neovascularisation in retinal microvascular endothelial cells in response to hypoxia[340, 341]. This suggests that SMAD1 may mediate angiogenesis within hypoxic regions in the developing heart wall, although this will require further investigation to verify.

Our in-silico findings suggested a transition of endocardium to coronary vascular endothelium. This is consistent with previous findings from murine lineage-tracing studies, in which a proportion of the endocardium gives rise to coronary vascular EC, either *via* angiogenic sprouting[300] or *via* the segregation of endocardial trabeculae protruding into the myocardium during compaction[339]. However, it is currently unknown whether cells assume a venous identity upon their conversion from endocardium to coronary vascular endothelium, as observed in our trajectory inference analysis. Although studies in mouse have demonstrated a significant proportion of coronary vascular EC to be derived from venous cells of the SV, this is thought to occur much earlier in cardiac development than the comparative gestational age of the human fetal samples used in this analysis[301, 328]. Characterisation of the transcriptional regulation underlying the two methods of endocardial to

vascular endothelial transition may also provide further context to the cellular dynamics predicted in this study.

Trajectory inference analysis also revealed subsequent arterial specification of capillary EC. This finding corroborates observations from murine studies in which a primitive coronary vascular capillary plexus undergoes specification to form coronary arteries[306]. In addition to finding upregulation of known mediators of arterial specification such as *HEY1*[167] and *SOX17*[138], *MECOM* was also identified as having a role in the establishment of an arterial EC identity. Furthermore, enriched arterial expression of MECOM was also observed in coronary EC from E16.5 mouse hearts, suggesting an evolutionary conserved function.

The role of MECOM in arterial EC identity was validated by siRNA mediated knockdown in arterial-like hESC-EC. Previous work from Li et al[342] demonstrated that MECOM acts upstream of Notch signalling during zebrafish nephrogenesis. Given the importance of Notch signalling in arterial EC specification, this suggested that MECOM may alter arteriovenous identity by regulating Notch signalling. Whilst reduction in MECOM expression did not alter the expression of arterial EC markers, including those belonging to the Notch pathway, an increase in venous EC marker expression was observed. This included the TF, NR2F2, which is known to establish venous identity in part *via* repression of Notch signalling[12]. This suggests that MECOM may be required to supress venous identity during arterial EC specification. Previous

studies from mouse have hypothesised that a venous identity is the default state for EC, with venous identity needing to be repressed via Notch signalling during arterial specification[180, 181]. Our finding that MECOM knockdown altered venous marker (NR2F2 and EPHB4) expression, without changes to expression of Notch signalling genes, thus suggests that additional factors are required to supress venous identity during human arterial EC specification. Recently, increasing evidence has suggested induction of EC quiescence is required for arterial specification to occur[190, 191, 305]. It would therefore be interesting to determine whether perturbation of MECOM in arterial EC causes changes in cell cycle state leading to modulation of arteriovenous identity.

Although this study is the most comprehensive of its type to date, due to limited sample availability its data provides only a snapshot of a narrow developmental window (13-14 weeks). Whilst trajectory inference methods permit the dynamical changes to be characterised within individual datasets, inclusion of fetal samples from a wider range gestational ages would provide a more comprehensive understanding of human coronary vascular development, especially at earlier stages.

In summary, we have used a high throughput scRNA-seq platform to comprehensively map the transcriptional landscape of the human fetal heart endothelium at 13-14 weeks. This study complements studies using murine models of cardiovascular development by providing novel insight into endothelial cell heterogeneity within the developing human heart, as well as

the dynamical changes accompanying coronary vasculature formation. In addition to helping understand the mechanisms giving rise to congenital coronary vascular abnormalities, this information may prove valuable in future strategies to guide coronary vascular formation for the treatment of coronary vascular disease.

3.9 Figure Legends

Figure 3.1: Mapping the human fetal heart endothelium using scRNA-seq.

(A) Schematic of experimental design for mapping the human fetal heart endothelium using 10X scRNA-seq. **(B)** Representative FACS gating strategy used to isolate viable CD31+ CD45- endothelial cells. **(C)** UMAP visualisation of clusters identified in scRNA-seq data from cardiac endothelial cells isolated from human fetal heart (n=2). **(D)** Feature plots showing expression of key marker genes defining distinct endothelial populations. **(E)** Metagene analysis of fetal heart scRNA-seq data visualised in self organised maps for total dataset (left) and subpopulations of endothelial cell (right). Radar plots show enrichment of each metagene signature in individual clusters. **(F)** GO term enrichment analysis conducted using genes from metagenes' signatures A (left) and F (right).

Figure 3.2: Gene regulatory network analysis of human fetal heart endothelium.

(A) Heatmap of differentially expressed cluster genes: expression of top 10 differentially expressed genes for each cluster identified in the complete dataset. Genes were grouped according to the cluster in which they were differentially expressed. **(B)** Violin plots: expression of *CD36, RGCC, INMT*, and *KIT* across clusters identified in fetal heart endothelial cell dataset. **(C)** Gene regulatory network constructed using SCENIC analysis. Transcription factors and target genes shown as squares or circles, respectively. Genes are

coloured based on the cluster in which they were differentially expressed. White nodes represent gene targets that were not differentially expressed. **(D)** Violin plots showing enrichment/AUC score of *SOX4* and *SMAD1* regulons across identified clusters.

Figure 3.3: Trajectory inference analysis of developing cardiac endothelium.

(A) RNA velocity analysis of microvascular cardiac endothelium. The RNA velocity field was superimposed onto a UMAP visualisation of microvascular cardiac endothelial cells. **(B)** Slingshot trajectory demonstrating pseudotemporal cellular dynamics. **(C)** Heatmap of 200 genes found to be most differentially expressed across pseudotime of trajectory from Panel B. Genes grouped into modules by k-means clustering (k=4). **(D)** Smoothing spline curves show average scaled gene expression of genes within modules identified in Panel A. **(E)** Feature plots showing expression of selected genes from modules 2 (*DKK3* and *GATA6*), 3 (*MECOM* and *HEY1*), or 4 (*TCF15* and *MEOX1*).

Figure 3.4: Comparison of the transcriptional profiles of human and mouse fetal cardiac endothelial cell populations.

(A) Heatmaps showing expression of either conserved (left) or human-specific (right) markers for each subpopulation of fetal cardiac endothelium. **(B)** Expression of selected conserved markers in EC populations in human and mouse. **(C)** Expression of markers identified as being human-specific.

Figure 3.5: Knockdown of MECOM in hESC-EC.

(A) Expression of arterial (*SOX17, DLL4, MECOM*), venous (*NR2F2, EPHB4*), and lymphatic markers (*PROX1*) across hESC-EC differentiation. **(B)** Experimental design for siRNA mediated knockdown of MECOM in hESC-EC. **(C)** qRT-PCR quantification of MECOM expression following siRNA knockdown (n = 4 biological replicates). **(D)** Quantification of MECOM protein abundance following siRNA knockdown using siRNA MECOM 1 (n = 4 biological replicates) and siRNA MECOM 2 (n = 3 biological replicates) **(E)** qRT-PCR quantification of known markers of arterial (*DLL4*, *HEY1, JAG1, JAG2*) and venous (*EPHB4, NR2F2*) endothelial cells after MECOM siRNA knockdown (n = 4 biological replicates). **(F)** Western blot quantification of NR2F2 abundance following knockdown of MECOM in hESC-EC (n = 4 biological replicates).

Figure 3.6: Summary of key findings from scRNA-seq characterisation of human fetal cardiac endothelial cells.

Figure 3.1 Mapping the human fetal heart endothelium using scRNA-seq.

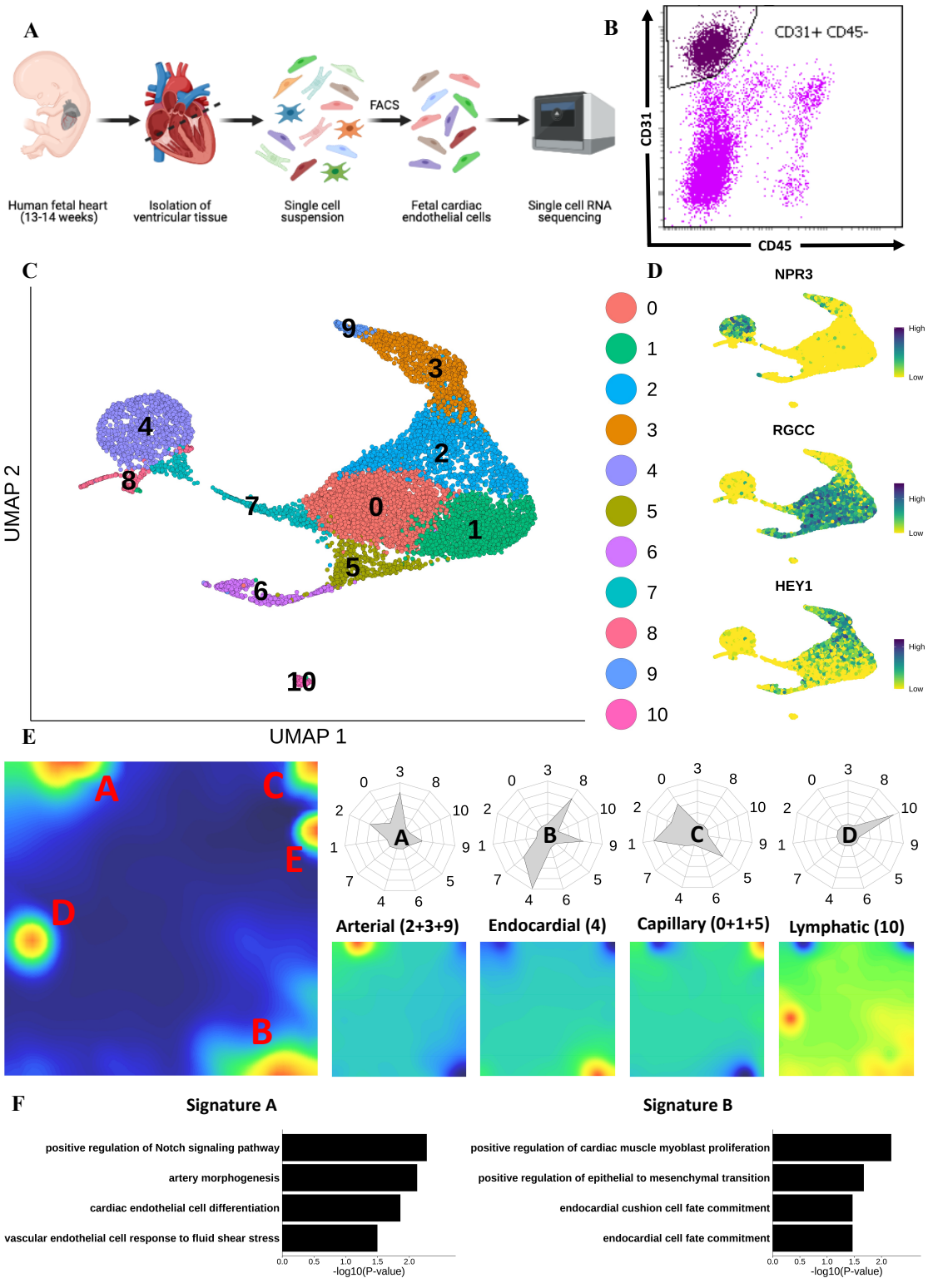


Figure 3.2 Gene regulatory network analysis of human fetal heart endothelium.

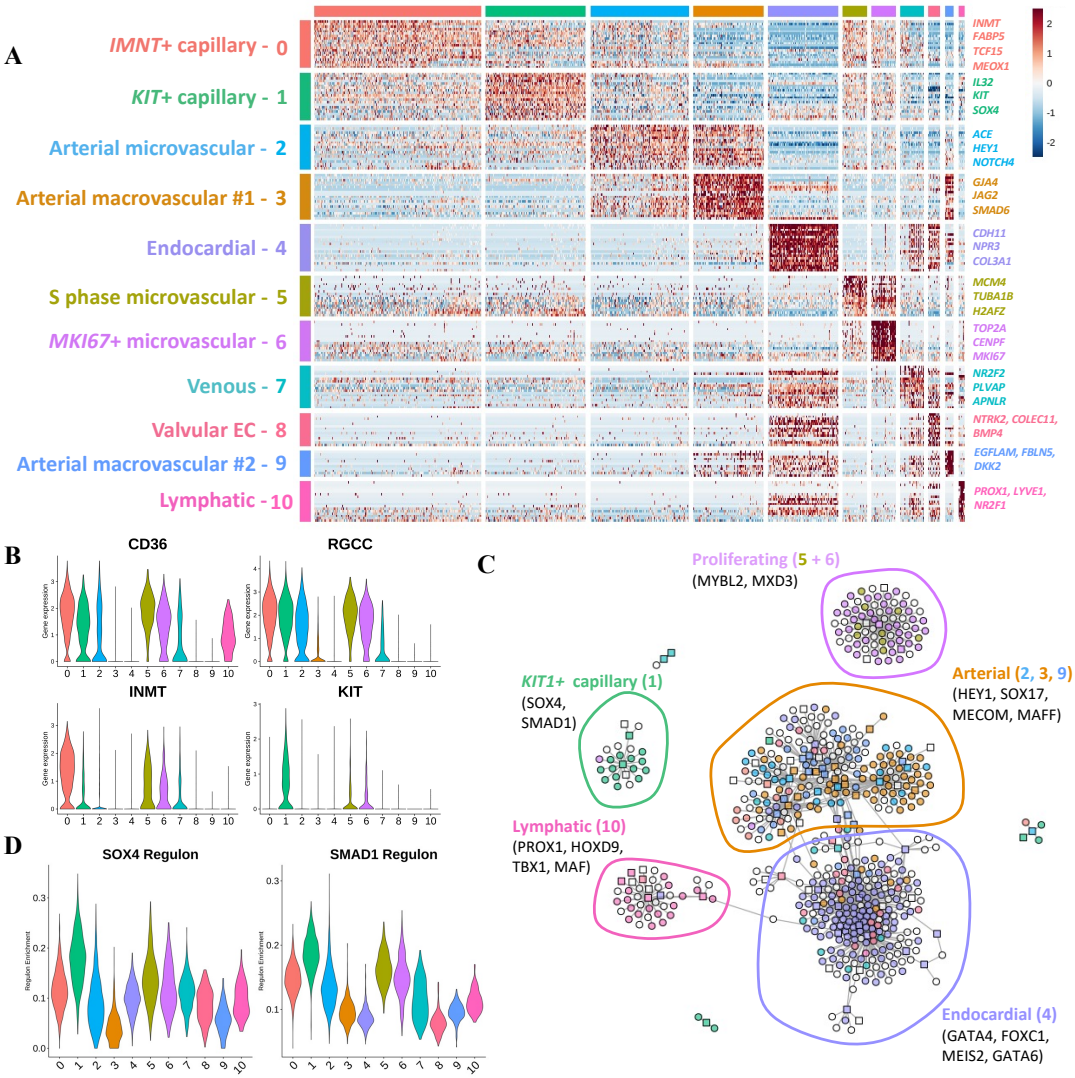


Figure 3.3 Trajectory inference analysis of developing cardiac endothelium.

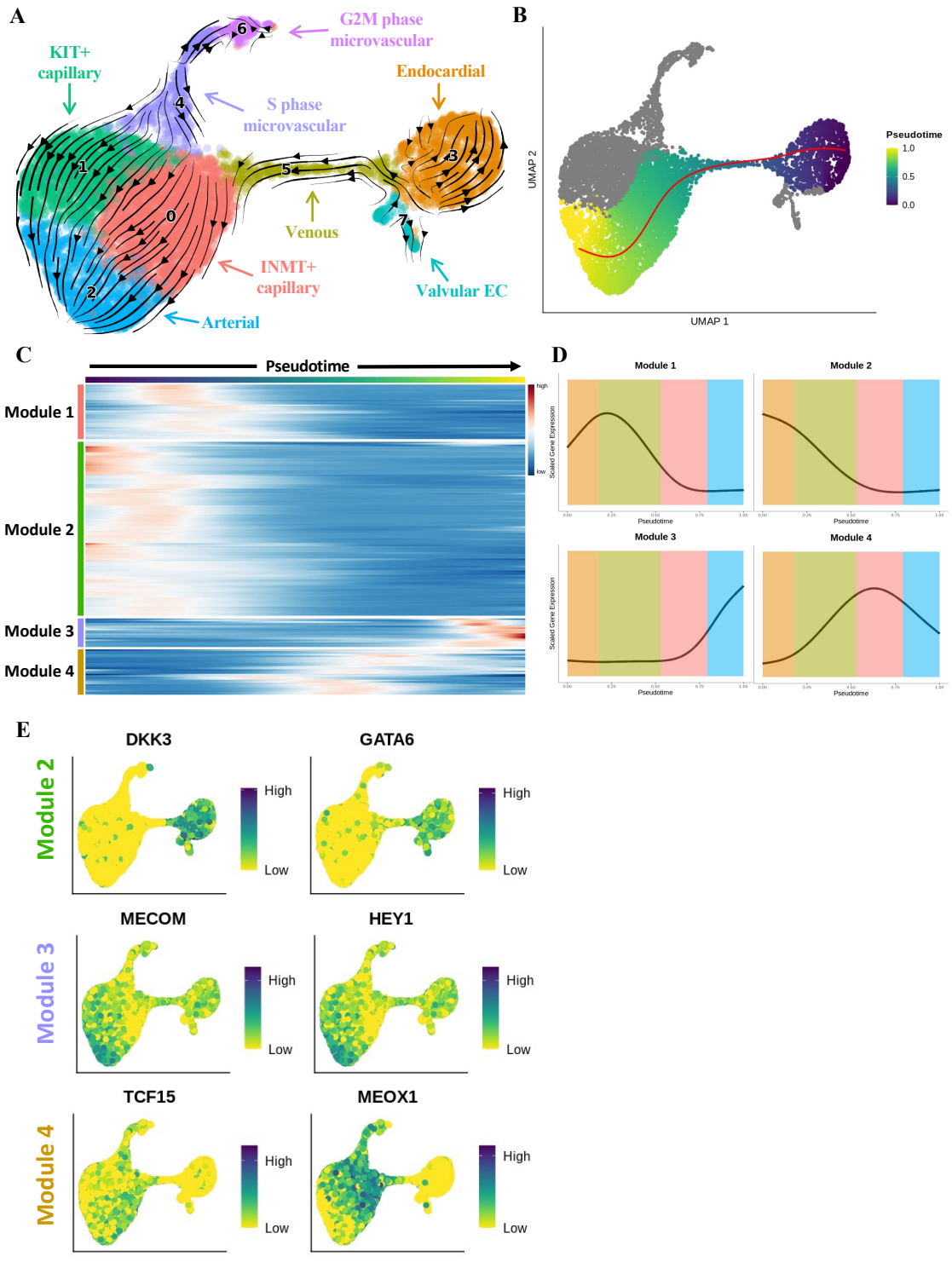


Figure 3.4 Comparison of the transcriptional profiles of human and mouse fetal cardiac endothelial cell populations.

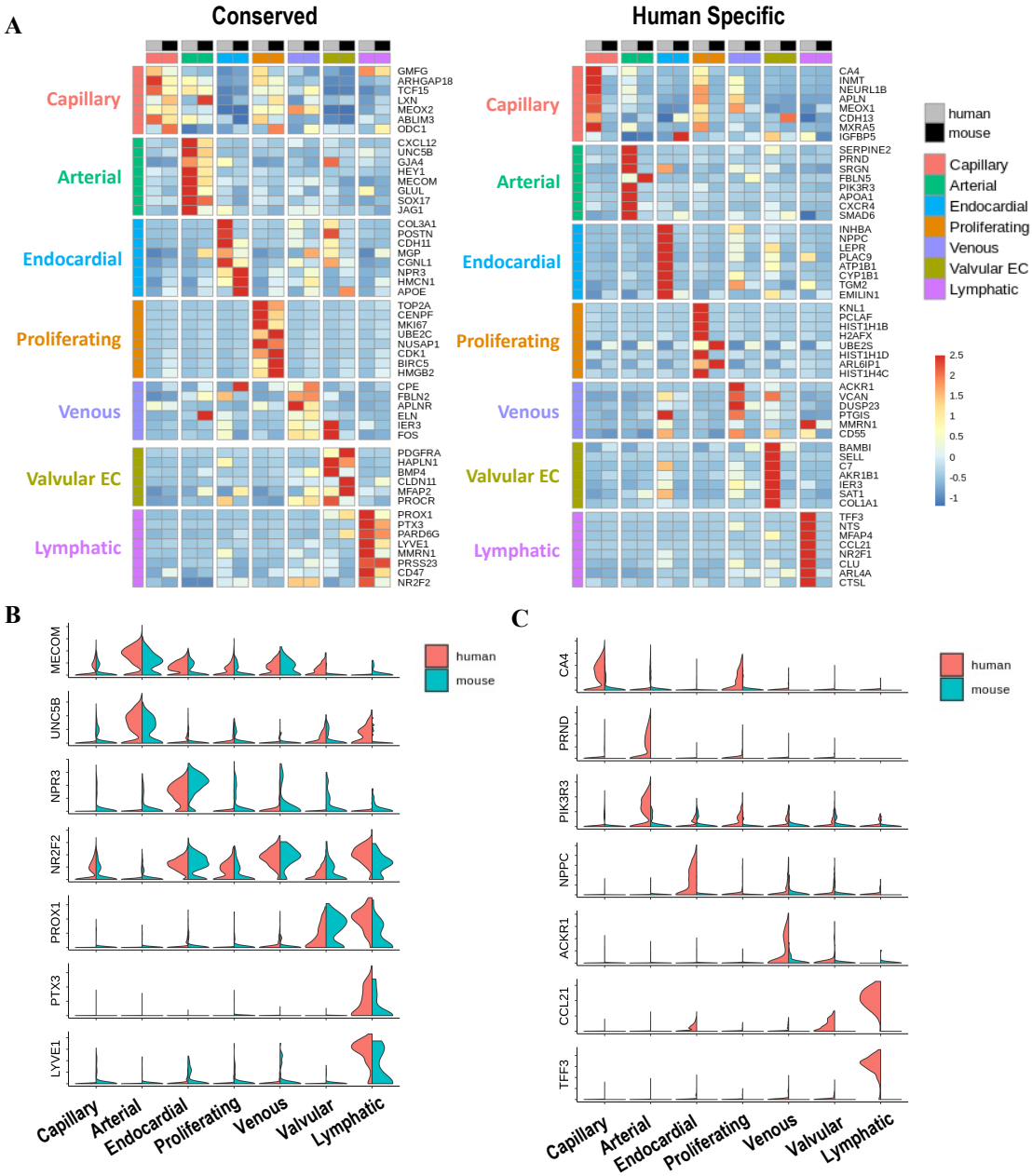


Figure 3.5 Knockdown of MECOM in hESC-EC.

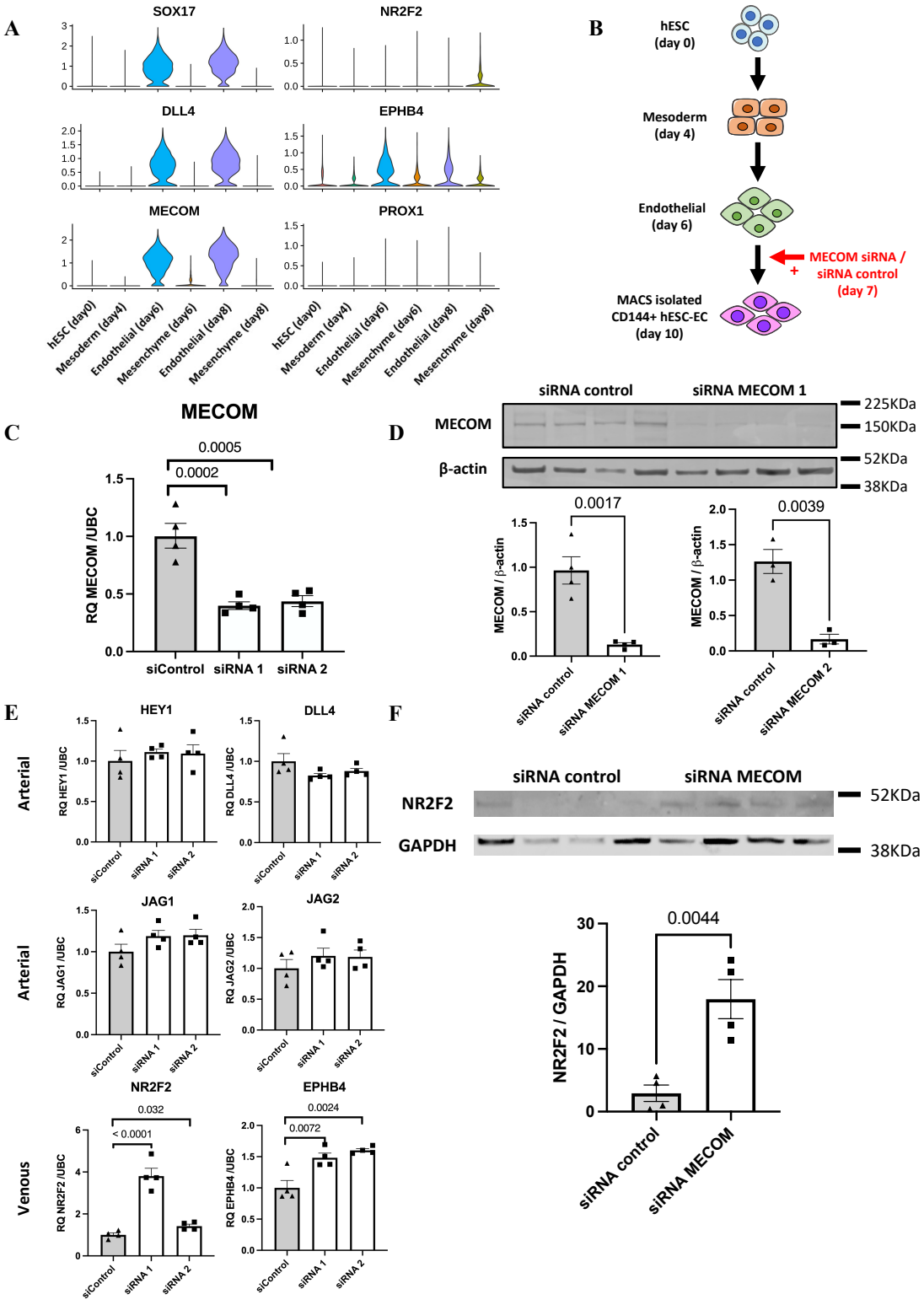
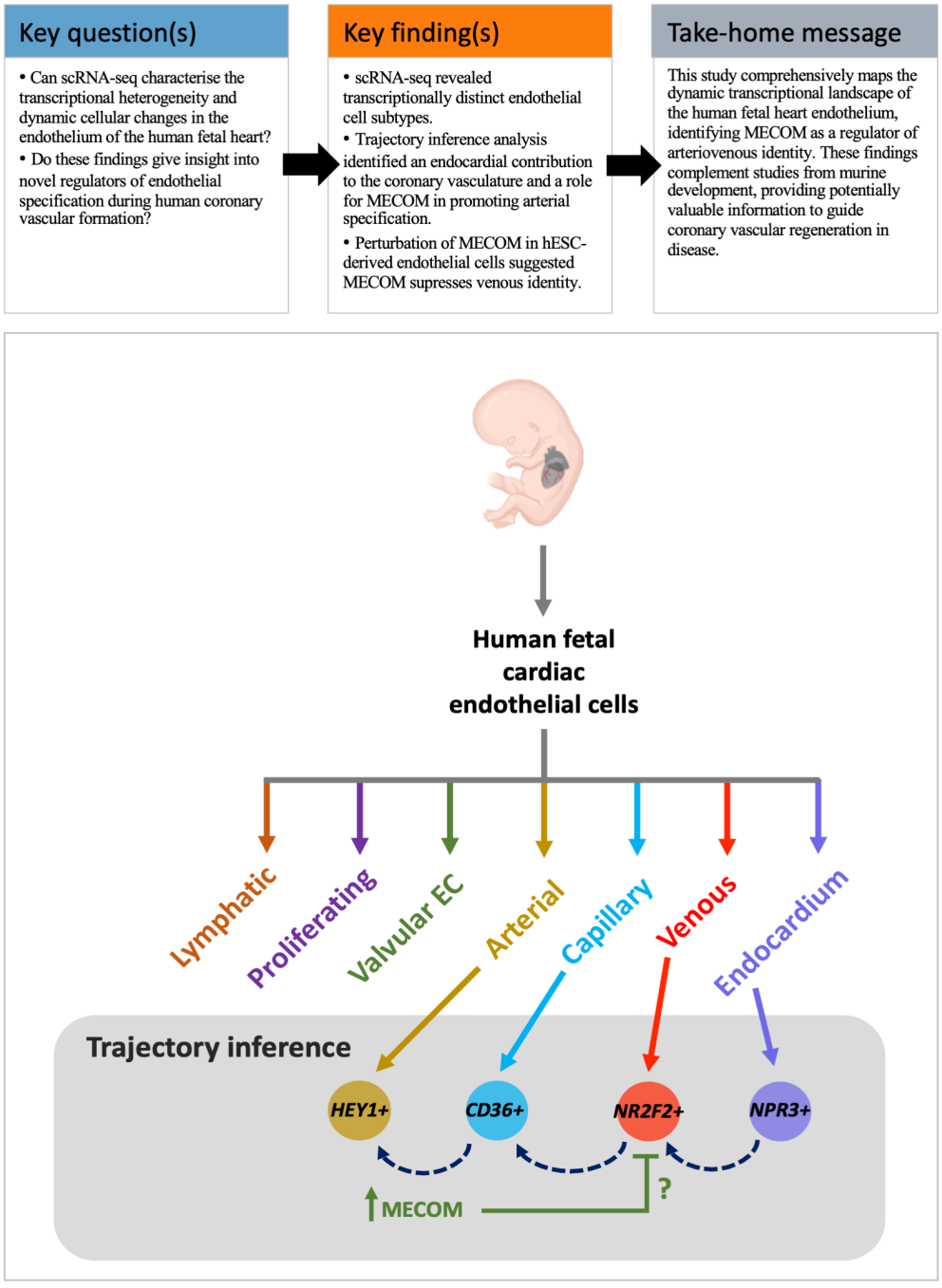


Figure 3.6 Summary of key findings from scRNA-seq characterisation of human fetal cardiac endothelial cells



3.10 Supplementary Methods

Isolation of endothelial cells from fetal cardiac tissue.

Cardiac endothelial cells were isolated from human fetal heart tissue using a method adapted from van Beijnum et al[316]. Fetal heart tissue was kept on ice prior to digestion to minimise RNA degradation. Ventricular tissue was finely minced prior incubating with continuous agitation at 37 °C in a digestion solution consisting of 9 ml 0.1 % collagenase II (Thermo Fisher, Paisley, UK) and 1 ml dispase (2.5U/ml) (Thermo Fisher). After 20 minutes incubation, 75 µl of 1 mg/ml DNaseI (Roche, Basel, Switzerland) was added prior to incubation for a further 15 minutes. Digestion was then quenched by placing the sample on ice and adding 5 ml cold RPMI with 10 % FCS (Life Technologies, Paisley, UK). Undigested clumps of tissue were then excluded by passing the sample through a 100 µm cell strainer and rinsing with a further 5 ml of cold RPMI + 10 % FCS. Cells were then collected by centrifugation (500 g for 5 minutes at room temperature) and 2 ml red blood cell lysis buffer (Thermo Fisher) added for 2 minutes at room temperature before neutralising with cold 15 ml RPMI + 0.1 % BSA. Cells were then pelleted by centrifugation (500 g for 5 minutes at 4 °C). Cells resuspended in RPMI + 0.1 % BSA were then stained on ice with APC anti-human CD31 and PE anti-human CD45 (Supplementary Table 3.1) for 45 minutes before resuspending again in RPMI + 0.1 % BSA. DAPI was added prior to sorting to allow exclusion of dead cells. CD31+ CD45- endothelial cells were isolated by fluorescence activated cell sorting (FACS) using the BD FACS Aria III Fusion cytometer (Becton Dickinson, Franklin Lakes, USA).

Single-cell RNA sequencing library construction.

Cell viability and concentration were determined by staining a 10 µl aliquot of FACS sorted endothelial cells 1:1 with trypan blue and loading onto a haemocytometer and counting manually. Viability of cells from both samples exceeded 85 %. 8,000 cells were loaded onto the 10X Chromium Controller (10X Genomics, Pleasanton, USA) according to the manufacturer's protocol. Library construction was conducted using the Single Cell 3' Reagent Kit (V3.1) in accordance with the manufacturer's instructions (10X Genomics, Pleasanton, USA). In brief, following generation of the gel bead in emulsions (GEMs) and reverse transcription, cDNA was extracted from GEMs and amplified. cDNA clean-up was performed using a SPRIselect Reagent Kit (Beckman Coulter, Brea, USA). Indexed sequencing libraries were generated using the V3.1 Chromium Single-Cell 3' Library Kit (10X Genomics, Pleasanton, USA). The key steps in this protocol included enzymatic fragmentation, end-repair, A-tailing, adaptor ligation, ligation clean-up, sample index PCR, and PCR clean-up. Libraries were sequenced by Genewiz (USA) using the Illumina NovaSeq 6000 platform.

Single-cell RNA sequencing data processing.

Raw sequencing data was processed using the 10X CellRanger V3.1.0 pipeline (10X Genomics, Pleasanton, USA) to generate filtered and non-filtered count matrices. RNA-seq reads were aligned to the GRCh38-3.0.0 genome reference downloaded from 10X Genomics. In addition to the default cell calling method applied within the CellRanger pipeline, the DropUtils

package was used with a FDR threshold of 0.01 to identify barcodes corresponding to cells *versus* those from empty droplets[317]. Only data from cells identified using both methods were used for downstream analysis. Quality control was performed using the R Scater package[318]. Cells with a total UMI count exceeding 3 median absolute deviations (MADs) from the median value or with total gene count lower than 2 MADs from the median value were removed from downstream analysis. In addition, cells with a high proportion of counts from mitochondrial genes (>3 MADs) were also excluded. Prior to merging individual datasets, data normalisation was performed using the MultiBatchNormalisation method[319] to minimise batch effects between datasets. The top 2,000 most variable genes were subsequently identified using the standard Seurat FindVariableFeatures method. Mitochondrial and ribosomal genes were excluded from the list of variable genes. Normalised data was then scaled using the standard ScaleData function[320] and principal component analysis conducted using the previously identified list of variable genes. Harmony batch correction was implemented to integrate both datasets, following the standard workflow[321]. Cells were clustered and visualised in two dimensions using Uniform Manifold Approximation and Projection (UMAP) following the standard Seurat workflow and utilising the corrected harmony reduction embedding values[320]. Fetal cardiac EC were assigned cell cycle stage using the list of cell cycle genes from Tirosh et al[343] and implementation of the 'CellCycleScoring' function[320]. Differential gene expression analysis was performed using the standard FindMarkers method[320]. Significantly differentially expressed genes were determined using the Wilcoxon signed

rank test with Bonferroni correction (adjusted p value < 0.05). A $\log_e$ fold change threshold of 0.3 was used when identifying differentially expressed genes (DEGs). Furthermore, DEGs were only considered further if expressed in more than 30 % of cells within their corresponding cluster. DEGs were then ranked based on their log fold change in expression. A complete list of DEGs is provided in Supplementary File 2.

Metagene signature analysis.

The R package SCRAT v1.0.0 [194] was used to map the transcriptional signatures underlying the heterogeneity observed in the data. Normalised count data was used as the input data to construct self-organising maps (SOM) for the entire dataset and individual clusters. Enriched metagene signatures within individual cells were identified by applying a threshold retaining only metagene signatures with a signature score within the top 5% of signature scores (threshold score = max cell signature score x 0.95). Radar plots were constructed to visualise the proportion of cells within each cluster that exceeded the threshold for each metagene signature. GO term analysis was conducted using the R package topGO [299] using all detected genes as the background and the Fisher exact test with the "elim" algorithm to determine relevant significantly enriched GO terms.

Gene regulatory network analysis.

Gene regulatory network analysis performed using the standard R SCENIC (Single Cell rEgulatory Network Inference and Clustering) workflow [322]. Briefly,

the log normalised count matrix and accompanying cluster annotations were used as the input to GENIE3 [344] to infer potential transcription factor targets. A gene filter was first applied to remove genes with zero counts or those expressed in fewer than 10 cells. The Spearman's rank correlation coefficient matrix for the input expression matrix was calculated using the default 'runCorrelation' command. Results from GENIE3 were then used to generate co-expression modules before performing cis-regulatory motif analysis for each transcription factor to identify predicted regulons [322]. Motif enrichment analysis was performed using the two provided human RcisTarget databases [322] which provide rankings for motifs up to 500 bp upstream of the transcriptional start site (TSS) and for those in the 20 kbp region around the TSS. Regulons within cells were then scored using AUCell. Motifs annotated to the corresponding transcription factor and with a normalised enrichment score (NES) greater than 3.0 were retained for downstream analysis. Regulon enrichment scores, represented by AUCell values, were visualised using violin plots for each of the previously identified clusters. The gene regulatory network was constructed using the R package iGraph including only differentially expressed transcription factors and their predicted targets with a Genie3 weighting exceeding 0.01.

Trajectory inference analysis.

Clusters corresponding to lymphatic (*PROX1*$^{+}$*, LYVE1*$^{+}$*)* and macrovascular (*FBLN5*$^{+}$, *ELN*$^{+}$) endothelial cells were removed from the dataset prior to applying trajectory inference methods. RNA velocity analysis using the python

package scVelo[323] was applied using the stochastic model to determine the direction and magnitude of the inferred cellular dynamics. Identified trajectories were confirmed using the R package Slingshot[324] following the standard pipeline and with the endocardial cluster assigned as the starting cluster in accordance with findings from RNA velocity analysis. Genes differentially expressed across pseudotime were then identified using the TradeSeq package using the standard workflow[325]. Briefly, a negative binomial generalised additive model (NB-GAM) was fitted for the 2,000 genes identified as having the most variable expression within the dataset. The optimal number of knots used when fitting the NB-GAM was determined as 7 by using the 'Evaluate K' function and selecting the number of knots after which the Akaike Information Criterion (AIC) value remained approximately constant [325]. The 200 genes most differentially expressed across pseudotime were identified using the 'associationTest' function in which significantly differentially expressed genes ($p<0.05$) were then ranked using the Wald test statistic[325].

Multispecies comparison.

Human fetal cardiac EC scRNA-seq data was compared to equivalent 10X scRNA-seq data from E16.5 mouse hearts published previously by Goodyer et al[312]. Following downloading count matrices from the Gene Expression Omnibus (GSE132658), mouse data was processed as described above to identify cell clusters visualised using in 2D using UMAP. Mouse EC were identified in silico using the SingleR annotation package[345] following the standard pipeline, as well as by differential expression of endothelial markers

including Pecam1, Kdr, and Cdh5. Mouse data was then filtered to exclude non-EC from downstream analysis. Human orthologs for mouse genes were identified using the BioMart package[346] to allow marker gene expression comparison between species. Count data from mouse and human datasets were then processed in parallel to normalise gene expression and identify variable genes using the standard 'NormalizeData' and 'FindVariableFeatures' functions[320]. Individual datasets from human and mouse were then integrated using the standard Seurat integration method, using the first 30 dimensions from canonical correlation analysis (CCA) to identify integration anchors prior to running the 'IntegrateData' function[320]. Mouse EC populations equivalent to those identified above in human fetal EC data were classified using the 'TranferData' function and the human EC dataset as an integrated reference[320].

The coefficient of variation (CV) was then calculated from normalised expression values for individual genes in human and mouse EC datasets, retaining only genes with a CV between the 5th and 95th percentiles. This step was necessary prior to differential gene expression analysis to exclude genes with low variability across all cells (housekeeper genes) as well as highly variable genes with sporadic expression within a low number of cells. Differentially expressed genes (DEGs) for each EC population in mouse and human datasets were then identified using the 'FindAllMarkers' [320] function as described above. Marker genes were classified as conserved if found to be

differentially expressed within the same EC population in both mouse and human datasets.

Human embryonic stem cell (hESC) maintenance and endothelial differentiation.

H9 hESC (passage number 38 - 47) were maintained in StemPro hESC SFM (Thermo Fisher) on vitronectin coated plates (human recombinant vitronectin; Thermo Fisher). Cells were mechanically passaged at ~90 % confluence at a ratio of 1:6 – 1:8 using the StemPro EZ passage tool (Life Technologies). hESC lines were used in accordance with the UK Stem Cell Bank Steering Committee guidelines (Project Approvals SCS11-51 and SCSC17-26). hESC were differentiated to endothelial cells as previously described[241, 326]. hESC at ~90 % confluence were dissociated to single cells using StemPro Accutase (Life Technologies) before resuspending in mTeSR1 media (Stemcell Technologies, Vancouver, Canada) supplemented with 10 μM Y27632 (Tocris, Bristol, UK). Cells were then replated in T25 flasks (12,000 cells/ cm^2) coated 1 hour previously with fibronectin (5μg/cm^2). 24 hours later (day 1) media was changed to N2B27 media (comprising of 25ml DMEM/F12/GlutaMAX medium + 25 ml Neurobasal medium + 250μl GlutaMAX (100X) + 1ml B27 supplement + 500μl N2 supplement + 0.5 mL β-mercaptoethanol; Thermo Fisher) supplemented with 7μM CHIR-99021 (Sigma, St. Louis, USA) and 25 ng/ml BMP4 (R&D Systems, Minneapolis, USA). Cells were then left for 3 days without media change. At days 4 and 5, media was changed for StemPro34 media supplemented with 200 ng/ml VEGF-A (R&D) and 2μM forskolin

(Sigma). At day 6 cells were dissociated using 1X TrypLE Express (Life Technologies) and replated on uncoated flasks (40,000 cells/cm^2) in EGM2 media (Lonza, Basel, Switzerland) containing 50 ng/ml VEGF-A and 1% human AB serum (Sigma).

MECOM siRNA knockdown

Small interfering RNA (siRNA) -mediated knockdown of MECOM in day 7 hESC-EC was performed using predesigned SilencerSelect siRNA (Thermo Fisher) at a final concentration of 5 nM (Supplementary Table 3.2). To reduce the likelihood of siRNA specific off-target effects, two independent siRNA were used in parallel to knockdown MECOM. siRNA transfection was performed by combining 2 ml of diluted Lipofectamine RNAiMAX reagent (12 µl in 2 ml Opti-MEM media) (Thermo Fisher) with 2 ml diluted siRNA (diluted in 2 ml Opti-MEM media to achieve a final concentration of 5 nM) and allowing complexes to form for 15 minutes at room temperature. Following washing cells with Opti-MEM media, siRNA-lipid complexes were added, and cells incubated at 37 °C / 5 % CO_2. After 6 hours, transfection media was replaced with EGM-2 media supplemented with 1 % human AB serum and 50 ng/ml VEGF-A. Media was refreshed 48 hours later at day 9 before harvesting CD144+ cells at day 10 for RNA and protein. Day 10 cells were dissociated using 1X TrypLE Express and pelleted by centrifugation at 300 g for 3 minutes. Cells were washed once in DPBS-/- (minus calcium and magnesium) before resuspending in 80 µl MACS buffer (DPBS, 0.05 % BSA, 2 mM EDTA) and 20 µl CD144 MACS beads (Miltenyi Biotech, Bergisch Gladbach, Germany). Following staining at 4 °C for

15 minutes, CD144+ cells were isolated using MS columns (Miltenyi Biotech) according to manufacturer's instructions.

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was isolated from CD144+ hESC-EC using the miRNAeasy kit (Qiagen) following manufacturer's instructions. Reverse transcription was then conducted with the Multiscribe Reverse Transcriptase kit (Thermo Fisher) using 400 ng of total RNA. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed using predesigned TaqMan probes (Supplementary Table 3.3) and the QuantStudio 5 Real-time PCR system (Thermo Fisher) with a programme of 10 minutes at 95 °C followed by 40 cycles of 15 seconds at 95 °C and 60 seconds at 60 °C. Ubiquitin C was used as the housekeeper gene for normalisation (Thermo Fisher) and relative quantification calculated using the $2^{-(\Delta\Delta CT)}$ (RQ) method.

Western Blot

CD144+ hESC-EC were pelleted (300 g, 3 minutes) and resuspended in cold RIPA buffer containing 1X cOmplete protease inhibitor cocktail (Sigma). After centrifugation at 10,000 g for 15 minutes at 4°C the supernatant was collected, and protein quantified using the Pierce BCA Protein Assay Kit (Life Technologies). Following adding 6X laemmli SDS reducing buffer (Thermo Fisher) and denaturing at 95 °C for 10 minutes, protein was separated on a 4-12 % Bolt Bis-Tris 1mm gel (Thermo Fisher) before transferring to a nitrocellulose membrane (Life Technologies). Membranes were blocked in 4%

milk for 1 hour at room temperature and incubated in primary antibody solution diluted overnight at 4 °C with continuous agitation (Supplementary Table 3.4). After washing 3 times in TBST, membranes were stained with secondary LiCor antibodies (Lincoln, USA) diluted in 4 % milk (Supplementary Table 3.4). Following 3 further washes in TBST, membranes were imaged using the LiCor CLx imaging system. Uncropped blot images are provided in Supplementary Figure 3.6.

Supplementary Table 3.1 Antibodies used for staining of fetal cardiac EC

Target	Antibody	Supplier	Category Number	Dilution
CD31 / PECAM1	APC anti-human CD31	eBioscience	17-0319-42	1:100
CD45 / PTPRC	PE anti-human CD45	Biolegend	4303144	1:100

Supplementary Table 3.2 Silencer Select siRNA

	ID number	Category Number
siRNA MECOM 1	s4873	4392420
siRNA MECOM 2	s4872	4392420
Negative control siRNA	NA	4390843

Supplementary Table 3.3 Probes used for TaqMan Assays

Target	ID number
MECOM	Hs00602795_m1
NR2F2	Hs00819630_m1
EPHB4	Hs00174752_m1
DLL4	Hs00184092_m1
HEY1	Hs01114113_m1
UBC	Hs01871556_s1

Supplementary Table 3.4 Antibodies used for Western Blot analysis

	Target	Supplier	Category Number	Host Species	Dilution
Primary	MECOM	Abcam	ab124934	Rabbit	1:500
	NR2F2	Abcam	ab211777	Rabbit	1:500
	β -actin	Santa Cruz Bio	sc-69879	Mouse	1:2000
	GAPDH	Abcam	ab8245	Mouse	1:2000
Secondary	680RD anti-Rabbit IgG	Licor	926-68071	Goat	1:10,000
	800CW anti-Mouse IgG	Licor	926-32210	Goat	1:15,000

3.11 Supplementary Figure Legends

Supplementary Figure 3.1: scRNA-seq of human fetal cardiac endothelium.

(A) Feature plots showing expression of endothelial cell markers. **(B)** UMAP visualisation of cells from each fetal heart sample after implementing batch correction. **(C)** Violin plots showing expression of *NPR3* (endocardial), *RGCC* (capillary), and *HEY1* (arterial). **(D)** GO term enrichment analysis using genes from metagene signatures C (left) and D (right). **(E)** Transcriptional signature of proliferating EC: Enrichment of signature E within proliferating EC demonstrated in SOM (left) and radar plot (middle). GO term analysis using signature E genes returns terms relating to cell proliferation. **(F)** Bar plot showing predicted proportion of cardiac EC in G1, G2M, and S phase of the cell cycle. **(G)** Violin plots of key marker genes defining subpopulation of fetal cardiac EC.

Supplementary Figure 3.2: Profiling the transcriptional profile of subpopulations of fetal cardiac EC.

(A) Heatmap showing the correlation of overall transcriptional profile of identified clusters. **(B)** Violin plots showing expression of extracellular matrix related genes (*FBLN5, ELN, FBN1*) and shear stress associated gene *KLF4*. **(C)** GO term analysis using genes differentially expressed in macrovascular EC populations (clusters 3 and 9). **(D)** Violin plots showing expression of *SMAD1*. **(E)** Heatmap showing expression of top 10 differentially expressed

TFs for each cluster identified in the total dataset. TFs grouped by the cluster for which they were identified as differentially expressed in.

Supplementary Figure 3.3: Profiling the transcriptional signature of subpopulations of fetal cardiac endothelial cells.

(A) UMAP visualisation of microvascular EC. **(B)** Violin plots of key EC population markers shown in clusters identified in Figure S3A. **(C)** Ratio of spliced to unspliced transcripts identified for *PLVAP* and *NR2F2*. **(D)** Spline plots showing normalised expression of *NPR3, CDH11, NR2F2, CEBPD*, *FOS,* and *MECOM* across pseudotime. **(E)** Spline plot of normalised expression of *DACH1* over pseudotime.

Supplementary Figure 3.4: Comparison of the endothelial cell transcriptome in the developing human and mouse heart.

(A) UMAP visualisation of integrated mouse and human cardiac EC data after batch correction. **(B)** Feature plots of *INMT* and *KIT* expression in human and mouse datasets. **(C)** Heatmaps showing expression of mouse-specific markers for each subpopulation of cardiac endothelium.

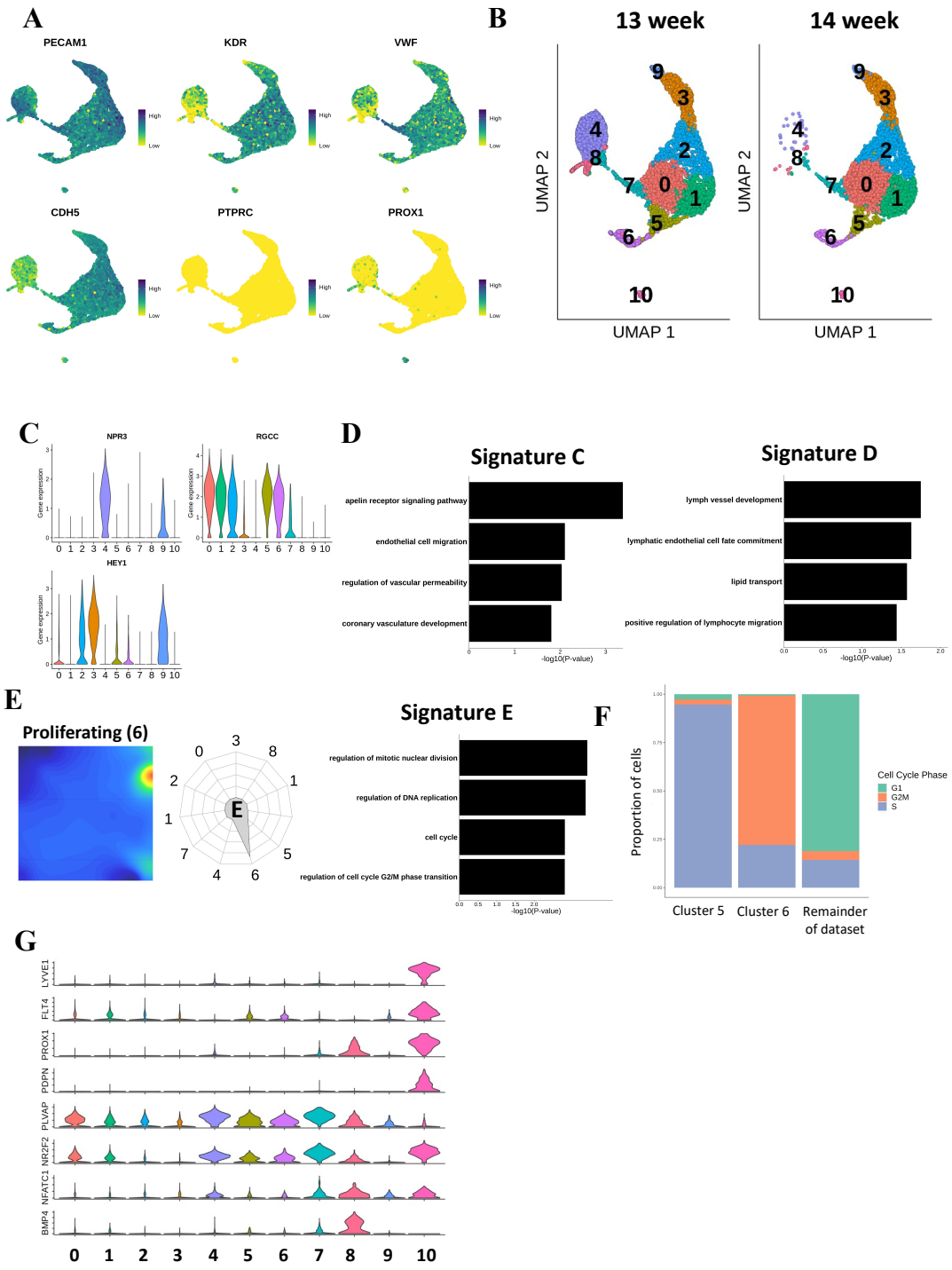
Supplementary Figure 3.5: siRNA mediated knockdown of MECOM in hESC-EC.

(A) Quantification of MECOM protein abundance following knockdown with siRNA *MECOM* 2 (n=3 biological replicates). **(B)** Expression of pan-endothelial marker, *CDH5*, 72 hours after MECOM knockdown (n = 4 biological replicates).

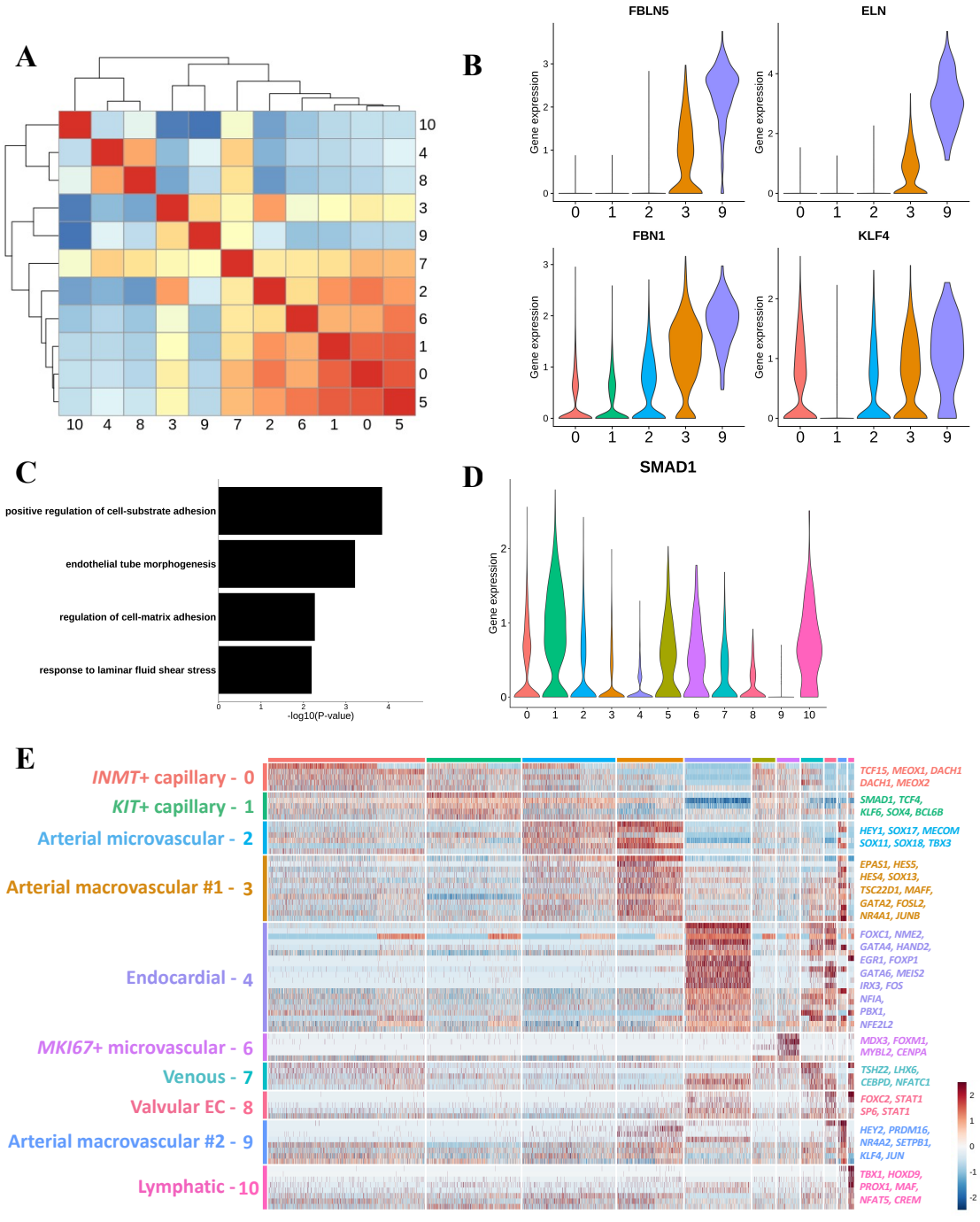
Supplementary Figure 3.6: Uncropped western blot images from hESC-EC MECOM knockdown.

Uncropped western blot images demonstrating abundance of MECOM following knockdown using **(A)** siRNA MECOM 1. **(B)** siRNA MECOM 2. **(C)** NR2F2 abundance after MECOM knockdown using siRNA MECOM 1.

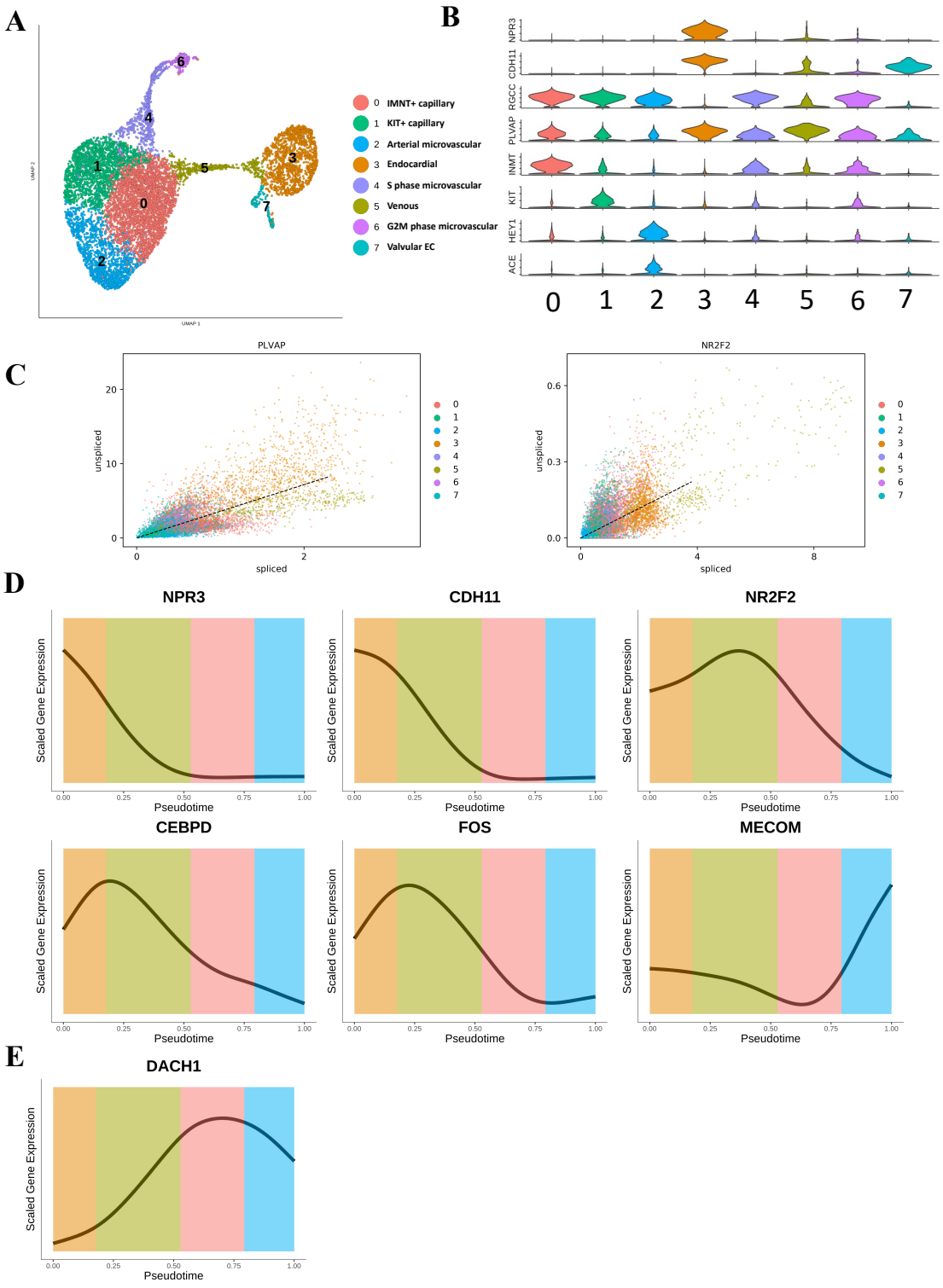
Supplementary Figure 3.1



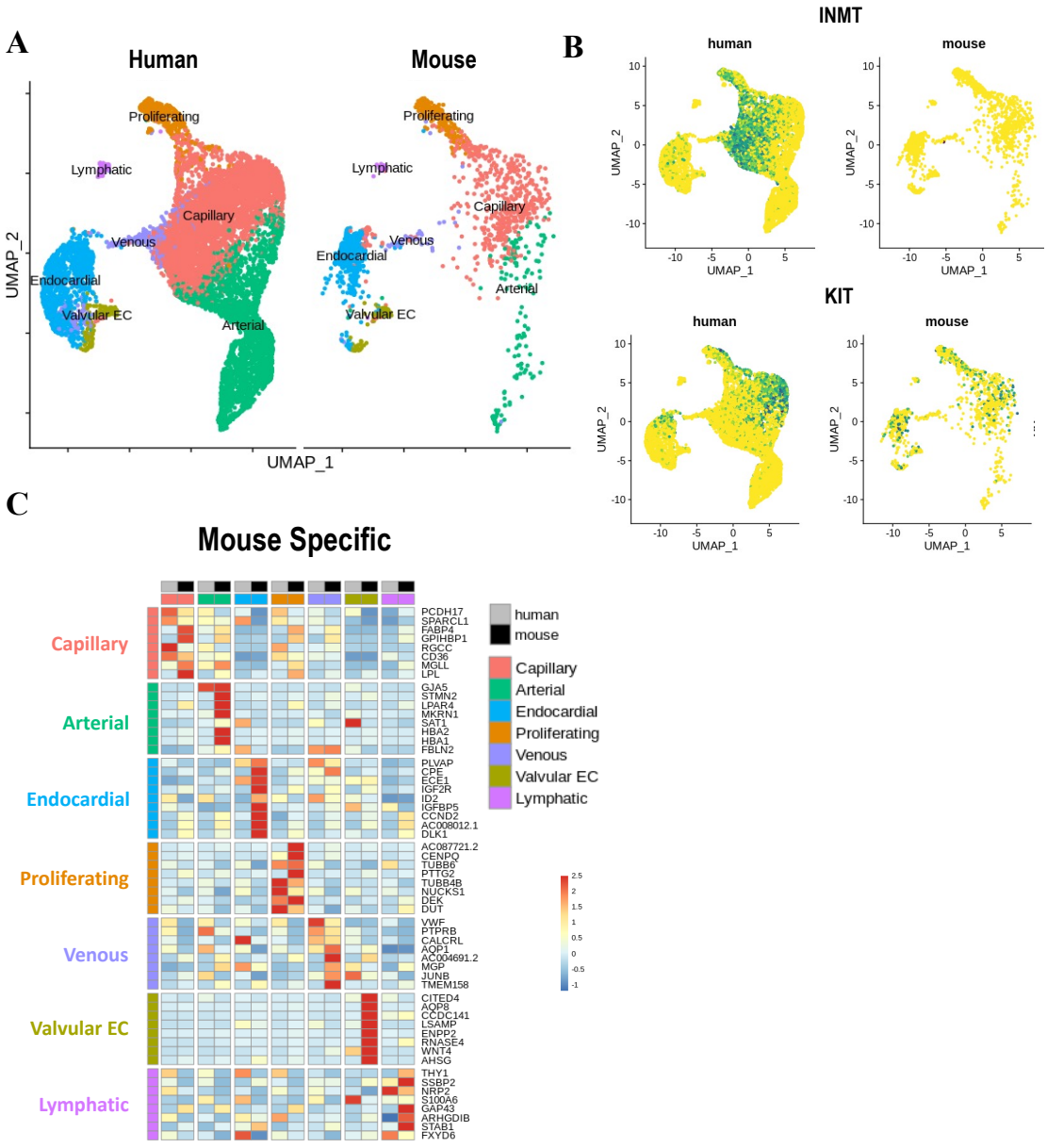
Supplementary Figure 3.2



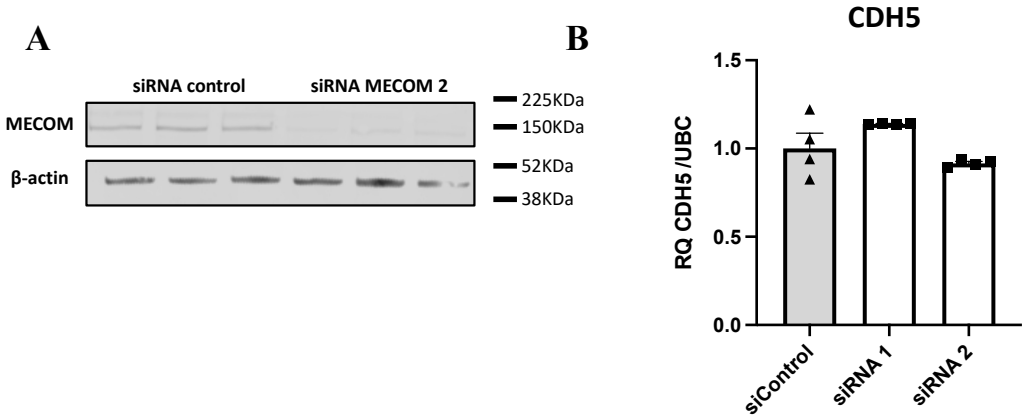
Supplementary Figure 3.3



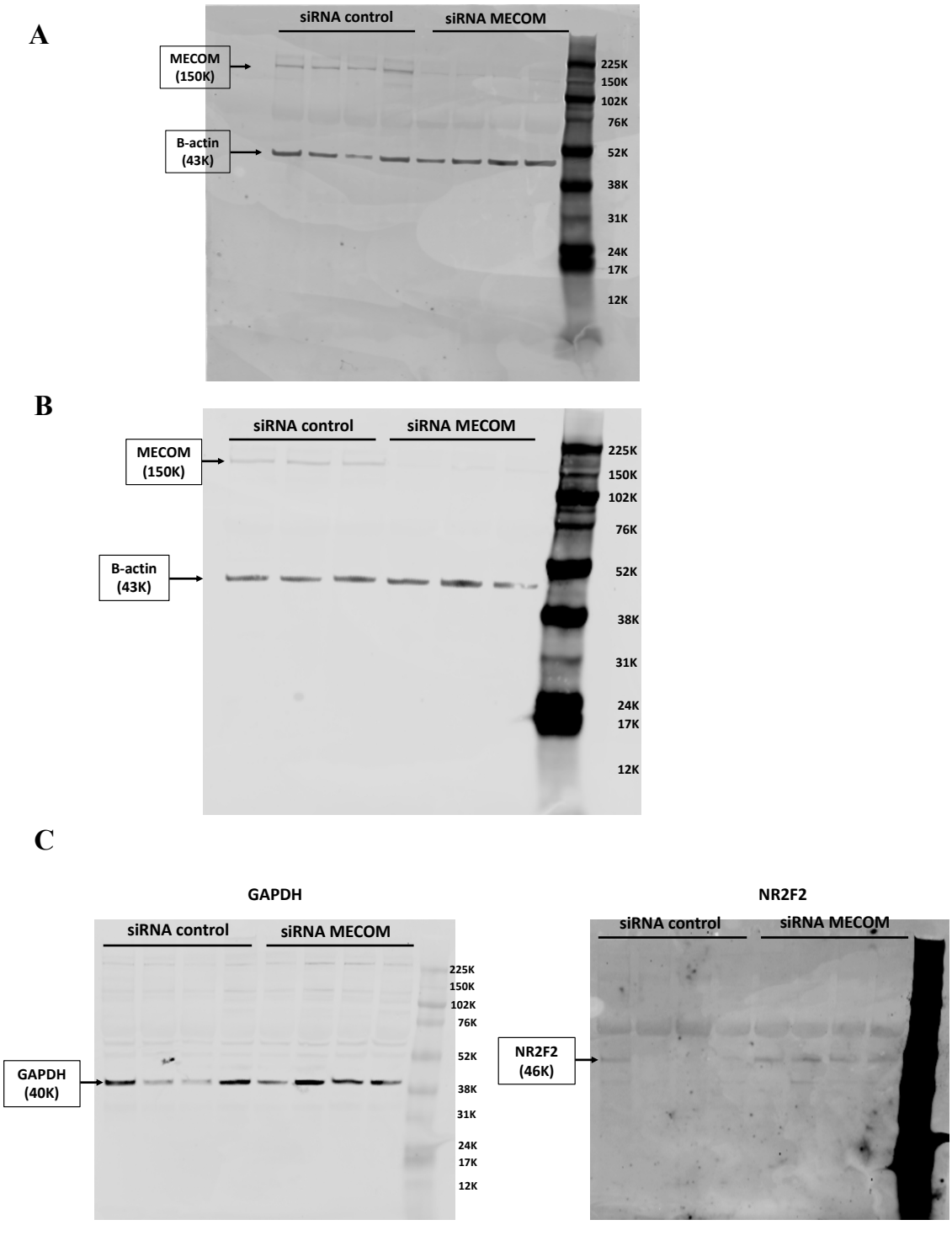
Supplementary Figure 3.4



Supplementary Figure 3.5



Supplementary Figure 3.6



3.12 Chapter Conclusion

In this chapter, scRNA-seq was applied to characterise the transcriptional landscape of the endothelium in the human fetal heart. This permitted the identification of distinct endothelial subtypes each expressing its own transcriptional signature. This included arterial, capillary, endocardial, venous, and lymphatic EC populations. Gene regulatory network analysis was applied to determine the TFs predicted to be responsible for underlying the unique transcriptional identity of EC subpopulations. A putative role for *MECOM* and *SMAD1* was identified in arterial and capillary populations, respectively. Subsequent trajectory inference analysis predicted an endocardial contribution to the coronary vasculature as well as the subsequent arterial specification of capillary EC, which was accompanied by increasing expression of *MECOM*.

Finally, a hESC differentiation protocol, characterised in chapter 2, yielding arterial-like hESC-EC was used as a developmental model of arterial EC specification to characterise the role of MECOM in determining arteriovenous identity. Knockdown of MECOM was found to result in an increase in non-arterial EC marker expression. This suggests MECOM may be required in arterial EC to suppress expression of non-arterial genes. Collectively, findings presented in this chapter provide new insight into the likely transcriptional regulators of human EC specification which are required to form the heterogenous endothelial population found within the mature coronary vascular network.

Chapter 4 - Lack of Evidence of Angiotensin-Converting Enzyme 2 Expression and Replicative Infection by SARS-CoV-2 in Human Endothelial Cells

4.1 Chapter Introduction

Following the emergence of the COVID-19 pandemic in 2020 and the subsequent temporary shutdown on all non-COVID related lab work, I was provided with the opportunity to repurpose my experience of analysing transcriptomic data and endothelial cell culture experiments to help determine the effect of SARS-CoV-2 on the endothelium. During the early stages of the pandemic, several largely unsubstantiated reports suggested that direct SARS-CoV-2 infection of EC may be a major contributor to COVID-19 pathology. It was therefore considered important to determine if publicly available transcriptomic datasets demonstrated expression of the SARS-CoV-2 receptor, *ACE2*, in the endothelium and if direct infection of primary EC can occur *in vitro.*

4.2 Title Page

Lack of Evidence of Angiotensin-Converting Enzyme 2 Expression and Replicative Infection by SARS-CoV-2 in Human Endothelial Cells

McCracken IR, Saginc G, He L, Huseynov A, Daniels A, Fletcher S, Peghaire C, Kalna V, Andaloussi-Mäe M, Muhl L, Craig NM, Griffiths SJ, Haas JG, Tait-Burkard C, Lendahl U, Birdsey GM, Betsholtz C, Noseda M, Baker AH, Randi AM. Lack of Evidence of Angiotensin-Converting Enzyme 2 Expression and Replicative Infection by SARS-CoV-2 in Human Endothelial Cells. Circulation. 2021 Feb 23;143(8):865-868.

Author contribution statement

Analysis of bulk and single-cell RNA sequencing datasets was caried out by G.S., L.H., and A.H. C.P. performed studies determining the effect of cytokine stimulation on *ACE2* expression in EC. G.M.B. conducted analysis is epigenomic data. Culture of endothelial cells for viral replication assays was performed by I.R.M. Viral replication assay using hCoV-229E was carried out by I.R.M. with assistance from S.J.G. Exposure of EC to SARS-CoV-2 was performed in the BSL-3 tissue culture facility by A.D., S.F., and N.M.C. A.M.R. and G.M.B. were responsible for writing the manuscript with all authors involved in discussion of results and contributing to the production of the final version of the manuscript.

4.3 Main text

A striking feature of severe forms of coronavirus disease 2019 (COVID-19), the current pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is severe endothelial injury with micro- and macrothrombotic disease in the lung and other organs, including the heart. This has led to speculation that viral infection may damage the endothelium through 2 mechanisms: indirectly, via neighbourhood effects, circulating mediators, and immune mechanisms or directly by viral infection of endothelial cells (ECs).

To support the hypothesis of direct viral damage of ECs via virus-induced infection, the cells should express the main receptor for SARS-CoV-2, ACE2 (angiotensin-converting enzyme 2), a metalloprotease component of the renin–angiotensin hormone system and a critical regulator of cardiovascular homeostasis[347]. Indeed, several recent review articles propose that SARS-CoV-2 binding to ACE2 on ECs is the mechanism through which the virus may cause direct endothelial damage and endothelialitis[347]. However, expression of ACE2 in ECs has not been convincingly demonstrated to support this assumption, nor has there been sufficient evidence to support a direct infection of ECs by SARS-CoV-2.

To address the questions of ACE2 expression in human ECs and of the ability of SARS-CoV-2 to infect the endothelium, we interrogated transcriptomic and epigenomic data on human ECs and studied the interaction and replication of

SARS-Cov-2 and its viral proteins with ECs in vitro. The data, analytic methods, and study materials will be maintained by the corresponding author and made available to other researchers on reasonable request.

Analysis of RNA sequencing was carried out on ENCODE data from ECs from arterial, venous, and microvascular beds, in comparison with epithelial cells from respiratory, gastrointestinal, and skin sources. Very low or no basal *ACE2* expression was found in ECs compared with epithelial cells (Figure 4.1A and 4.1B). Moreover, in vitro exposure of ECs to inflammatory cytokines reported as elevated in the plasma of patients with severe COVID-19 failed to upregulate *ACE2* expression (Figure 4.1C).

Publicly available single-cell RNA sequencing of human organ donor hearts[17] showed that, while *ACE2* sequence reads are abundant in pericytes, they are rare in ECs (Figure 4.1D). Of 100,579 ECs, only 468 (0.47%) were ACE2+, and in the majority (424), only a single *ACE2* transcript was detected. This could reflect true low and rare endothelial *ACE2* expression but also contamination from adherent pericyte fragments, a common confounder in vascular single-cell RNA sequencing data[348]. If such fragments contributed to the *ACE2* transcripts observed in certain ECs, we would expect to detect other pericyte transcripts in the same cells. Among the top-50 gene transcripts enriched in *ACE2*+ versus *ACE2*− ECs, we noticed several known pericyte markers, including *PDGFRB, ABCC9, KCNJ8*, and *RGS5* (Figure 4.1E). Comparison of transcript abundance across the 3 major vascular and

mesothelial cells showed that the top-50 gene transcripts were expressed at the highest levels in pericytes (Figure 4.1E). This suggests that the rare occurrence of *ACE2* transcripts in human heart ECs is likely caused by pericyte contamination. Similar conclusions have been reached previously in mouse tissues[348].

Analysis of the chromatin landscape at the *ACE2* gene locus in human umbilical vein ECs using data from ENCODE further supports this concept. The histone modification mark H3K27me3, which indicates repressed chromatin, was enriched at the *ACE2* transcription start site; conversely, promoter, enhancer, and gene body activation marks (H3K27ac, H3K4me1, H3K4me2, H3K4me3, and H3K36me3), RNA polymerase-II and deoxyribonuclease I hypersensitivity were absent or low, suggesting that *ACE2* is inactive in ECs. In marked contrast, the adjacent gene *BMX*, an endothelial-restricted nonreceptor tyrosine kinase, displays an epigenetic profile consistent with active endothelial expression (Figure 4.1F). Thus, transcriptomic and epigenomic data indicate that *ACE2* is not expressed in human ECs.

Other cell surface molecules have been suggested as possible receptors for the virus, but their role in supporting SARS-CoV-2 cell infection remains to be demonstrated. We therefore tested directly whether EC could be capable of supporting coronavirus replication *in vitro*. Productive levels of replication in primary human cardiac and pulmonary ECs were observed for the human

coronavirus 229E GFP (green fluorescent protein) reporter virus[349], which uses CD13 as its receptor, demonstrating directly that human ECs can support coronavirus replication in principle (Figure 4.1G). However, when cells were exposed to SARS-CoV-2, replication levels were extremely low for ECs, even after exposure to very high concentrations of virus compared with more permissive VeroE6 cells (Figure 4.1H). The observed low levels of SARS-CoV-2 replication in ECs are likely explained by viral entry via a non-ACE2–dependent route, attributable to exposure to extremely high concentrations of virus in these experiments (multiplicity of infection, 10 and 100).

These data indicate that direct endothelial infection by SARS-CoV-2 is not likely to occur. The endothelial damage reported in severely ill patients with COVID-19 is more likely secondary to infection of neighbouring cells and/or other mechanisms, including immune cells, platelets and complement activation, and circulating proinflammatory cytokines. Our hypothesis is corroborated by recent evidence that plasma from critically ill and convalescent patients with COVID-19 causes EC cytotoxicity[350]. These findings have implications for therapeutic approaches to tackle vascular damage in severe COVID-19 disease.

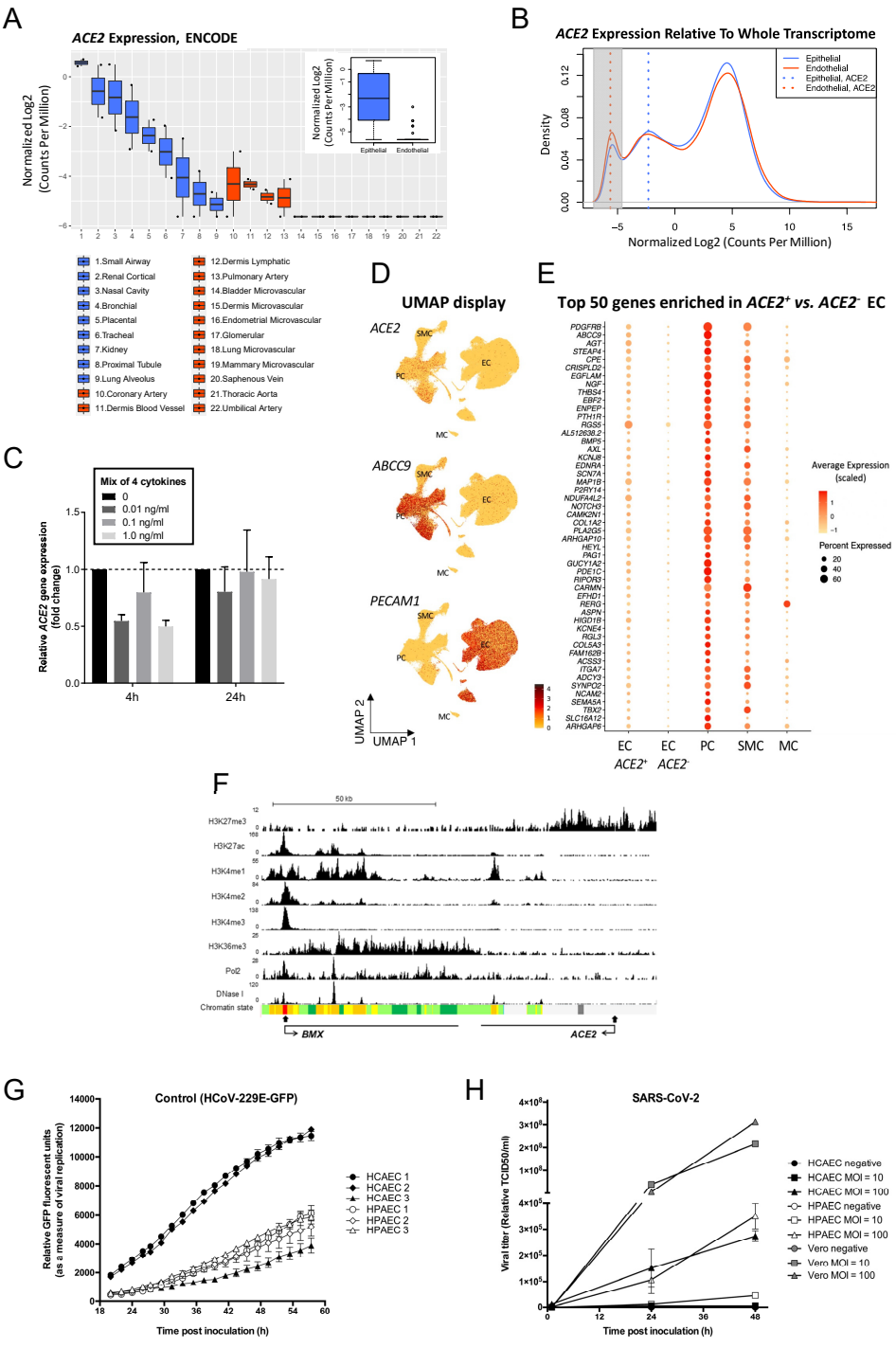
4.4 Figure Legend

Figure 4.1: Analysis of *ACE2* expression in human ECs and of coronavirus replication in primary human ECs. **A** and **B**, Comparison of *ACE2* expression in human primary epithelial cells and ECs using total RNA sequencing data from the ENCODE database shows low or absent expression in ECs. **A**, The difference of *ACE2* expression in epithelial cells and ECs is shown in boxplots with individual, as well as grouped, samples (inner boxplot). Each dot represents a single sample (n=2 per cell type). **B**, Transcriptome profiles of epithelial cells and ECs are shown in a density plot, using the median of all samples per group (n=19 360 genes). *ACE2* expression in each group is marked with a dotted line: *ACE2* expression in ECs (red) overlaps with the peak for nonexpressed transcripts (highlighted in grey), while *ACE2* expression in epithelial cells (blue) is to the right, indicating detectable expression. Median *ACE2* expression in endothelial cells equals -5.6 in log2 CPM, which corresponds with 0 raw read counts, signifying undetectable *ACE2* expression in the majority of ECs. Expression values in all plots are represented as log2-transformed CPM, normalized by trimmed mean of M-value (blue: epithelial, red: endothelial). **C**, *ACE2* expression is not regulated by inflammatory cytokines in HUVECs. qPCR analysis of *ACE2* mRNA expression in HUVECs treated with a mix of 4 cytokines/chemokines (TNF-α, IL1-β, IL8, and IL6/IL6R chimeric protein) for 4 hours or 24 hours at 0, 0.01, 0.1, or 1.0 ng/mL. Data are normalized to *GAPDH* and presented as mean±SEM of 3 independent experiments. **D** and **E**, Very low-level, rare, and likely contaminating *ACE2* transcripts are seen in ECs. **D**, *ACE2* transcript

reads are detected preferentially in PCs. UMAP landscapes of publicly available human heart data sets[17] include 100 579 ECs, 77 856 PCs, 16 242 SMCs, and 718 MCs (https://www.heartcellatlas.org/). *ACE2* transcript reads are detected preferentially in the PC cluster (enriching for *ABCC9*) and are rare in the EC cluster (enriching for *PECAM1*). E, PC transcripts are enriched together with *ACE2* in 0.47% of ECs. Dot plot displaying the abundance of top-50 transcripts enriched *ACE2*+ versus *ACE2*- ECs, across cell types indicated in D. (The Wilcoxon rank-sum tests with Bonferroni-corrected P values are < 1E-60 for each). **F**, Epigenetic profiling indicates that the *ACE2* gene is inactive in ECs. ChIP-seq binding profiles in HUVECs for histone modifications, RNA Pol2 enrichment, and deoxyribonuclease I hypersensitivity. The x axis represents the genomic position, the transcription start sites are indicated by closed arrow, and the direction of transcription is indicated by open arrows; the y axis shows ChIP-seq signal in reads per million per base pair (rpm/bp). The bottom row represents the chromatin state segmentation. Color key: active promoter, red; enhancers, yellow; transcriptional elongation, green; repressed, grey. **G** and **H**, Coronavirus replication in primary human cardiac and pulmonary ECs shows limited replication of SARS-CoV-2. G, Viral replication curves in human pulmonary (HPAEC) and cardiac (HCAEC) endothelial cells following infection with control HCoV-229E GFP (green fluorescent protein) reporter virus (MOI=0.6). Virus replication was measured via GFP fluorescence every 2 hours from 20 to 58 hours postinoculation. Mean±SEM of 3 technical replicates are shown at each time point for each biological replicate. **H**, Viral growth curves in HPAEC (n=3), HCAEC (n=3), and

nonendothelial Vero cells (n=1) after infection with SARS-CoV-2 at MOI=10 or 100. Supernatant were collected at 1, 24, and 48 hours post infection and virus copy number quantified by RT-qPCR detection of the SARS-CoV-2 N3 gene. ChIP-seq indicates chromatin immunoprecipitation sequencing; CPM, counts per million; EC, endothelial cell; HUVEC, human umbilical vein endothelial cell; IL, interleukin; MC, mesothelial cell; MOI, multiplicity of infection; PC, pericyte; qPCR, quantitative polymerase chain reaction; RT, real time; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SMC, smooth muscle cell; TCID, tissue culture infectivity dose; TNF, tumor necrosis factor; and UMAP, Uniform Manifold Approximation and Projection.

Figure 4.1 Lack of *ACE2* expression and productive SARS-CoV-2 replication in endothelial cells



4.5 Methods

Whole-transcriptome analysis with total RNA sequencing (RNA-Seq)

RNA-seq data files with gene quantifications for primary human epithelial and endothelial cells were downloaded from the ENCODE database (www.encodeproject.org)6, using Bioconductor package “ENCODEExplorer” and R. For consistency, only total RNAseq data generated by the same source, “Thomas Gingeras, CSHL”, and aligned to GrCh38 reference genome were included in the analysis. Cell types with one replicate were excluded. Raw counts were converted to counts per million (CPM), filtered based on expression, log2 transformed and normalized with Trimmed Mean of M-value (TMM) method using Bioconductor package “EdgeR”. Only genes with more than 0.3 counts per million (CPM) in at least three samples were kept for the analysis (19360 genes).

The accession numbers for the ENCODE files are ENCFF592BLV, ENCFF788JHJ, ENCFF235TGN, ENCFF401NRN, ENCFF207VFL, ENCFF623ZJJ, ENCFF620THF, ENCFF699QEJ, ENCFF797YZO, ENCFF985PDI, ENCFF110UGQ, ENCFF764AOQ, ENCFF325DPM, ENCFF564EGR, ENCFF233LYV, ENCFF378BDR, ENCFF555QVG, ENCFF972UYG, ENCFF511TST, ENCFF711SNV, ENCFF037WEH, ENCFF498AYE, ENCFF577VBY, ENCFF604VVJ, ENCFF177SUW, ENCFF456RET, ENCFF145DRX, ENCFF224ZRP, ENCFF176IZX, ENCFF674SRB, ENCFF580CHD, ENCFF709VCC, ENCFF747XTG, ENCFF756RKP, ENCFF060LPA, ENCFF262OBL, ENCFF091SWU,

ENCFF224YSC, ENCFF819IDA, ENCFF894GLT, ENCFF231GYQ, ENCFF836KPM, ENCFF361WEZ, ENCFF592KDP.

Endothelial cell cytokine treatments and real-time quantitative polymerase chain reaction

Pooled human umbilical vein endothelial cells (HUVEC; Lonza C2519A) were cultured in Endothelial Cell Growth Medium-2 media (EGM-2) (Lonza). HUVEC were plated at a density of 10,000 cells/96-well plate in EGM-2 supplemented with recombinant human TNF-α (300-01A, PeproTech EC Ltd), IL1-β (200-01B, PeproTech EC Ltd), IL8 (72aa, monocytes derived, 200-08M, PeproTech EC Ltd) and IL-6/IL-6R Alpha Protein Chimera (8954-SR, R&D Biotechne) at 0, 0.01, 0.1 or 1.0 ng/ml for 4 or 24 hours. Total RNA was isolated by using the RNeasy kit (Qiagen) and reverse transcribed into cDNA using Superscript III Reverse Transcriptase (Invitrogen). Quantitative real-time PCR was performed using PerfeCTa SYBR Green Fastmix (Quanta Biosciences) on a Bio-Rad CFX96 system. Gene expression values were normalized to *GAPDH* expression. Sequences of the human primers used in this study are listed below:

GAPDH forward: 5'-CAAGGTCATCCATGACAACTTTG-3' and GAPDH reverse: 5'-GGGCCATCCACAGTCTTCTG-3'

ACE2 forward: 5'-ACAGTCCACACTTGCCCAAAT-3' and ACE2 reverse: 5'-TGAGAGCACTGAAGACCCATT-3'.

Human heart single cell RNAseq data analysis

Human heart single cell raw counts data and the cell annotation meta data were obtained from a recent published study[17]. The data were processed using Seurat package (version: 3.1.1) in R[320]. The endothelial cells which contain *ACE2* reads were grouped as *ACE2*+ EC cells, and the rest as *ACE2*- EC cells. The genes enriched in *ACE2*+ EC cells were identified using FindMarkers function in Seurat. It applies a Wilcoxon Rank Sum test and then performs Bonferroni-correction using all genes in the dataset. The corrected p value < 0.05 was used as cutoff for significance. To visualize the top 50 genes among different cell groups, the DotPlot function was applied. The size of the dot represents percentage of cells expressing the gene and color represents the scaled average expression levels.

Epigenetic analysis of chromatin state

Genome-wide ChIP-seq data for histone modifications and DNase I hypersensitivity in HUVEC were obtained from the ENCODE[351]/Broad Institute under GEO accession number GSE29611. ChIP-seq reads were mapped to the human reference genome GRCh37/hg19. Tracks were visualized using the UCSC Genome Browser database (https://genome.ucsc.edu). HUVEC chromatin-state discovery and genome annotation was obtained using ChromHMM7[352] from ENCODE.

Viral replication assays

HCoV-229E

Stocks of HCoV-229E GFP reporter virus described by Cervantes-Barragan et al[349] were generated following cultivation on the HUH7 cell line. HPAEC (Lonza) and HCAEC (Promocell) were seeded in 96 well plates (2.1x104 cells/cm2) 24 hours prior to infection with HCoV-229E-GFP at MOI = 0.6. Virus inoculum was then replaced 1 hour later with 100ul fresh media before incubation at 34°C/5%CO2. GFP fluorescence was measured every 2 hours from 20 to 58 hours post inoculation using the ClarioStar plate reader (BMG).

SARS-CoV-2

A sample from a confirmed COVID-19 patient was collected by a trained healthcare professional using combined nose-and-throat swabbing. The sample was stored in virus transport medium prior to cultivation and isolation on Vero E6 (ATCC CRL-1586) cells. Samples were obtained anonymised by coding, compliant with Tissue Governance for the South East Scotland Scottish 279 Academic Health Sciences Collaboration Human Annotated BioResource (reference no. SR1452). Virus sequence was confirmed by Nanopore sequencing according to the ARCTIC network protocol (https://artic.network/ncov-2019), amplicon set V3, and validated against the patient isolate sequence. For the virus isolate used in this project is EDB-2 (Scotland/EDB1827/2020, UK lineage 109, B 1.5) genetic stability was observed up to 5 passages on Vero E6 cells, with particular attention to the S1/S2 furin cleavage site. HCAEC (Promocell) and HPAEC (Lonza) were

seeded in a 12-well plate (2.0x104 cells/cm2) one day prior to infection with SARS-CoV-2 EDB-2, passage 2, at MOI = 100 or 10 (assessed by endpoint titration on Vero E6 cells). At 0, 24 and 48 hours post infection (hpi) 10 µl of supernatant was lysed in VL buffer (20mM Tris-HCl pH 7.5, 300mM NaCl, 2.5% IGEPAL CA-630, 1:200 RNAsin Plus). The lysate was analysed by RT-qPCR, using GoTaq 1-Step RT-qPCR (Promega) with the SARS-CoV-2 CDC N3 primers (F – GGGAGCCTTGAATACACCAAAA, R – TGTAGCACGATTGCAGCATTG) at 350nM each to determine viral copy number. To assess viral copy numbers, resulting Cts were analysed against a standard curve of lysate with known TCID50/ml value as well as an RNA template.

4.6 Chapter Conclusion

Results in this chapter demonstrated an absence of *ACE2* expression, both in EC from culture and in those freshly isolated from adult human tissue. Although productive levels of viral infection occurred following exposure of HPAEC and HCAEC to HCoV-229E, an absence of productive infection was observed following incubation with SARS-CoV-2. The absence of direct EC infection by SARS-CoV-2 is likely due to an absence of *ACE2* expression, although this would require further experimental validation. Collectively, this study suggests that the EC dysfunction that can occur in patients with COVID-19 is likely to happen via an indirect mechanism as opposed to by direct infection by SARS-CoV-2. This finding provides important information for the development of future therapeutic strategies for the treatment of COVID-19.

Chapter 5 - General Discussion

5.1 Summary

The development of the endothelium and its underlying regulation have been subject to intensive research over recent years, with studies primarily relying on transgenic animal models of vascular development. However, the transcriptional regulators governing the individual stages of this process remain largely uncharacterised, especially in relation to their relevance to human biology. In this thesis I apply scRNA-seq to map the transcriptional signatures accompanying human endothelial cell development. This work builds upon previous findings from model organisms, providing new insight into the transcriptional regulation of EC differentiation and specification, with important relevance to future efforts of driving therapeutic neovasculogenesis in disease.

I first utilised a highly efficient eight-day hESC-EC differentiation to perform a longitudinal scRNA-seq study to map the cellular heterogeneity and transcriptional dynamics accompanying endothelial differentiation. This identified the emergence of a homogeneous mesodermal population by day 4 prior to the specification of distinct endothelial and mesenchyme populations. We then performed a comparable study using an alternative hESC-EC differentiation protocol. This revealed substantially more heterogeneity at each stage, including the unexpected emergence of putative nephrogenic and haemogenic EC populations. Despite clear differences between protocols, several conserved transcriptional signatures of endothelial differentiation were identified, including for TF with no previous reported role in EC differentiation.

Transcriptional comparison of hESC-EC to cultured and freshly isolated human EC revealed that hESC-EC remained transcriptionally distinct to their mature counterparts, likely reflecting their relative immaturity.

In a subsequent study, we performed scRNA-seq of the human fetal cardiac endothelium to investigate EC heterogeneity within the developing heart as well as to predict the dynamical cellular changes known to occur during murine cardiac development. This revealed distinct endothelial populations defined by differential expression of venous, arterial, endocardial, capillary, valvular, and lymphatic EC populations. Trajectory inference analysis then revealed an endocardial contribution to the coronary vasculature prior to the specification of capillary EC to arterial EC. These analyses identified the TF *MECOM* as having a central role in arterial EC identity. This function was validated via *MECOM* knockdown experiments in arterial-like hESC-EC, suggesting *MECOM* is necessary to suppress venous identity in arterial EC.

Finally, in light of the COVID-19 pandemic, we investigated the susceptibility of EC to direct infection with SARS-CoV-2. RNA sequencing analysis revealed an absence of the expression of SARS-CoV-2 receptor, *ACE2,* in EC. Furthermore, despite observing productive hCoV-229E viral replication in EC, a lack of viral replication was observed following the exposure of EC *in vitro* to SARS-CoV-2.

Despite the novelty of these findings and their important contribution to the field of human endothelial biology, several limiting factors within each major part of the study should be considered in future experiments. Importantly, given the prominence of *in vitro* models of EC development within these studies, the *in vivo* relevance of these findings will be discussed within this chapter.

5.2 scRNA-seq analysis of hESC-EC differentiation

Within the study described in chapter 2 of this thesis we used the differentiation of hESC to endothelial cells as an *in vitro* model of human endothelial development. We utilised high throughput scRNA-seq at multiple timepoints throughout the eight-day protocol to map the cellular heterogeneity and transcriptional signature of cells as they underwent directed endothelial differentiation. Although previous studies have applied scRNA-seq to characterise hESC-EC differentiation, these studies were severely limited by either the chosen scRNA-seq platform only profiling a low number of cells[242] or by the non-directed nature of the differentiation protocol used[247]. Collectively, our study captured the transcriptome of over 105,000 cells, firmly placing it as the largest and most comprehensive study of the transcriptional dynamics during hESC-EC differentiation. Furthermore, comparison of hESC-EC differentiation across multiple hESC lines and differentiation protocols allowed for the robust identification of conserved transcriptional signatures. Amongst the well characterised TF such as *ERG, ETS1, SOX7* with a known function in EC, we also identified TF with no current reported role in EC

differentiation to have increasing expression accompanying the assumption of hESC-EC identity. This included *BCL6B* which currently has no reported function in EC biology/development despite having previously been described to be expressed in the kidney endothelium of developing mouse embryos[291]. Given its remarkable similarity in expression across different hESC lines and differentiation protocols, BCL6B would be a strong candidate for further investigation to establish if it has a functional role in EC differentiation. Although outside the scope of this thesis, this could be tested *in vitro* by perturbing the expression of *BLCL6B* during hESC-EC differentiation and quantifying the effect on the efficiency of EC differentiation at day 6 by flow cytometry. The recent emergence of CRISPR/Cas9 genome editing methodologies have permitted the selective perturbation gene expression in hPSC-EC to evaluate the resultant effects on EC biology and/or differentiation[353-355]. Generation of a knockout hESC line using CRISPR/Cas9 editing to fully ablate *BCL6B* expression would provide an advantage over more conventional methods such as siRNA mediated knockdown when assessing the resulting effect on EC differentiation.

Although our study provides the most comprehensive profile of the transcriptional landscape during hESC-EC differentiation to date, it is important to highlight that this study failed to capture the transcriptional profile of cells at the specific timeframe at which they underwent commitment to the endothelial lineage from their mesoderm progenitors. Our data therefore captures the transcriptional state of the mesoderm population just before

committing to an EC fate (day4) and then hESC-EC population shortly after commitment from mesoderm (day6). Consequently, we were unable to pinpoint the earliest transcriptional changes accompanying the emergence of the EC population. This included for *ETV2* which as discussed previously is a key ETS TF responsible for priming mesodermal fate prior to the emergence of the endothelial lineage[83, 88]. The absence of this transitioning mesodermal population in our data also had important consequences when performing trajectory inference analysis, with an insufficient number of cells around this timepoint to accurately place the predicted bifurcation point. The effect of this limitation was largely mitigated by performing a comparable study using an alternative differentiation protocol in which EC commitment proceeds in a more asynchronous manner. However, this should be considered prior to selecting gene candidates for functional validation. Future studies may opt to use a narrower timeframe between establishment of the mesoderm and emergence of the hESC-EC to more accurately map the earliest transcriptional changes during EC differentiation.

We observed by day 8 of the differentiation that hESC-EC had acquired a transcriptional profile most closely aligning with that of arterial EC as opposed to markers of venous or lymphatic EC. This included enriched expression of arterial markers *DLL4* and *SOX17* within identified hESC-EC populations. Given the importance of VEGF-A concentration in determining either venous or arterial EC specification, it would be interesting to determine if hESC-EC differentiation could be guided towards a venous identity by reducing the

concentration of VEGF-A at days 4 and 5. Previous studies have demonstrated the use of lower concentrations of VEGF-A during embryoid body based hPSC-EC differentiation results in an increase in expression of venous markers *NR2F2* and *EPHB4*[199].

We demonstrated that the overall transcriptional state of our hESC-EC was distinct from that of both cultured human EC and EC freshly isolated from various fetal or adult tissues. Notably, the hESC-EC transcriptional profile did not more closely resemble that of either of the mature organ-specific EC types. This likely reflects the unspecified and comparatively immature nature of day 8 hESC-EC. As discussed previously, further EC specification is partially reliant on specific extrinsic signals from the surrounding tissue microenvironment. In addition to the results provided in chapters 2, 3, and 4, we also aimed to investigate if recapitulating certain environmental cues *in vitro* would be sufficient to promote the further maturation and/or specification of hESC-EC. Firstly, we used a peristaltic pump to expose hESC-EC to physiological levels of shear stress to determine if this could promote hESC-EC to acquire a more mature transcriptional signature. Comparable methods have been used previously to investigate the effects of shear stress on both human primary EC and hPSC-EC[189, 356, 357]. However, in our hands hESC-EC detached from the gelatin coated slides upon exposure to shear stress (Figure 5.1). Future studies should investigate if alternative coating matrices and/or lower levels of shear stress can prevent hESC-EC detachment, allowing for the resulting transcriptional changes to be explored.

Figure 5.1 Experimental design for exposure of hESC-EC to shear stress.

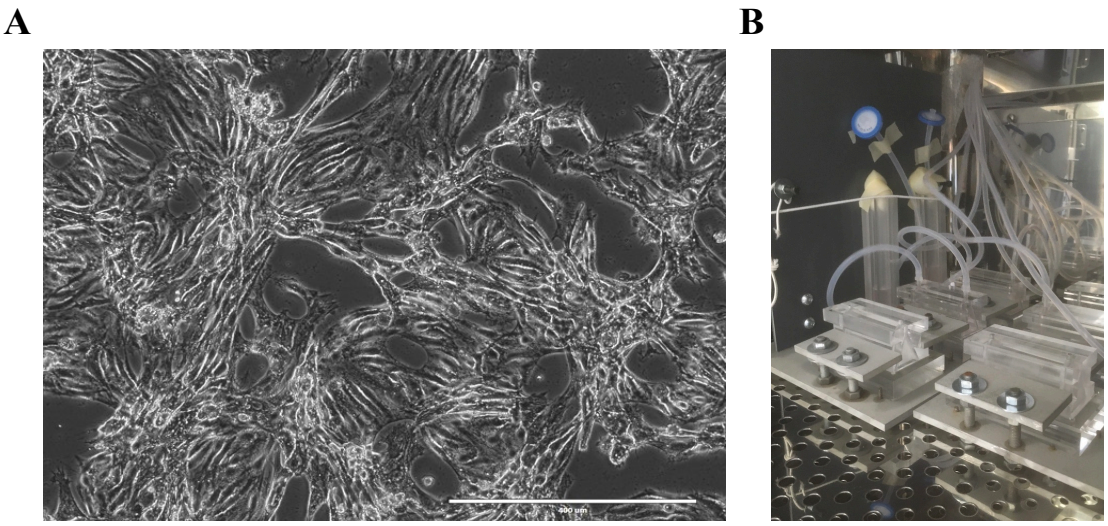


Figure 5.1: (A) Brightfield image of hESC-EC on gelatin coated coverslips prior to subjecting to shear stress. **(B)** Setup of equipment used to place hESC-EC under shear stress.

In addition to trying to expose the cells to shear stress to drive hESC-EC maturation, we also attempted to co-culture hESC-EC with hESC derived cardiomyocytes (hESC-CM) in a spheroid-based culture system. Previous studies have effectively co-cultured either primary cardiac EC or hPSC-EC with hPSC-CM using various culture techniques including spheroid and hanging drop-based methods[237, 358, 359]. These studies primarily focused on the resulting effects on hPSC-CM maturation meaning the effects of co-culture on hESC-EC maturation and specification remain unknown. However, our attempts to generate cardiac spheroids were impaired by inconsistencies in the efficiency of hESC-CM differentiation. Furthermore, whilst spheroids constructed using hESC-CM alone were able to form compact spheroids undergoing spontaneous contraction (Figure 5.2A), spheroids formed using a

mixed population of hESC-EC and hESC-CM at a 1:5 ratio did not retain any viable hESC-EC after 14 days in culture. In agreement with images from brightfield microscopy showing increased cell debris around the edge of the spheroid (Figure 5.2B) This suggests that the preferential adhesion between hESC-CM resulted in a lack of hESC-EC incorporation into the forming spheroid.

Figure 5.2 Generation of hPSC derived cardiac spheroids.

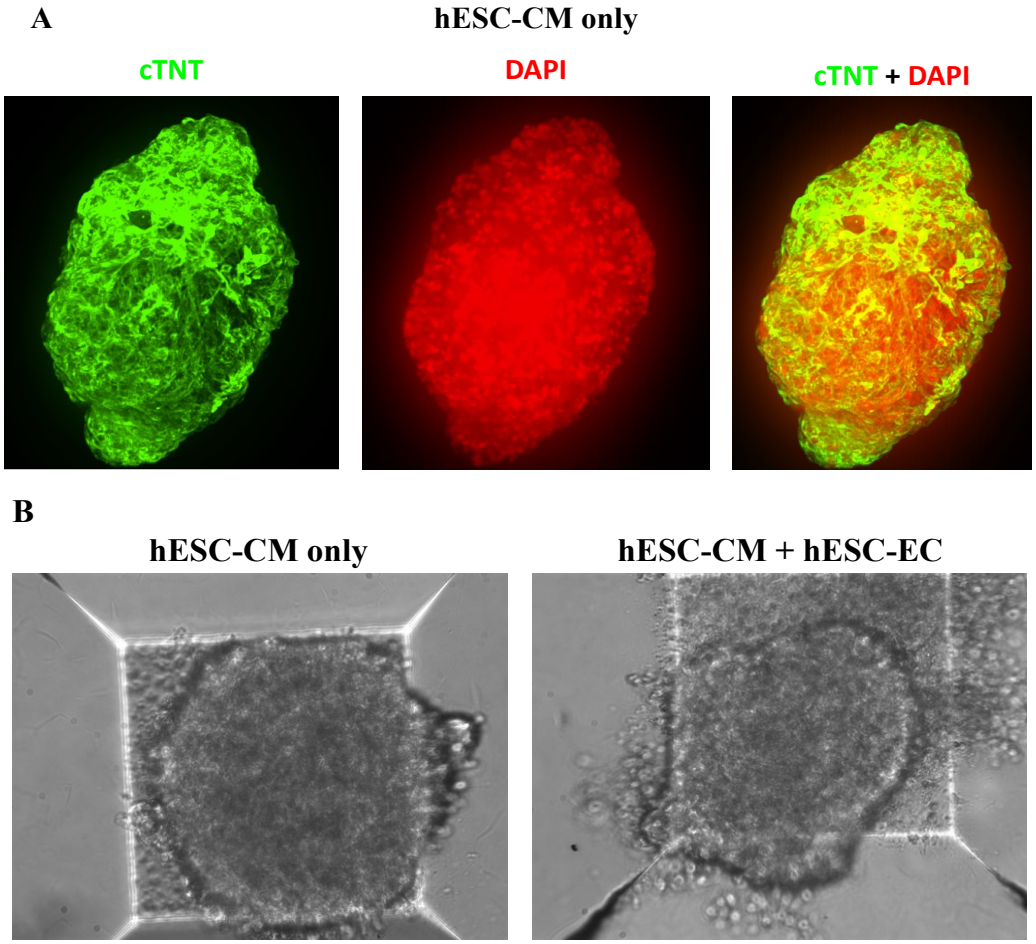


Figure 5.2: (A) Confocal images of cardiac spheroids constructed using cells from day15 of hESC-CM differentiation (~70-80% cTNNT2+). **(B)** Brightfield images of cardiac spheroids generated using either only cells from hESC-CM differentiation (hESC-CM only) or with the addition of CD144+ hESC-EC (at a 1:4 hESC-CM: hESC-CM ratio).

It therefore remains undetermined if co-culture of hESC-EC with hESC-CM is sufficient to induce expression of genes specifically expressed in cardiac EC. This finding would be important to inform future methods aimed to generate tissue specific hPSC-EC for future regenerative medicine approaches.

5.3 Mapping the transcriptome of the human fetal cardiac endothelium using scRNA-seq

The current body of knowledge surrounding the formation of the coronary vasculature is predominantly derived from transgenic mouse models, meaning the relevance of existing understanding to human development remains undetermined. The results presented in chapter 3 of this thesis aimed to address this current gap in knowledge by characterising the transcriptional profile of the human fetal cardiac endothelium using scRNA-seq. Although previous studies have used scRNA-seq to map the transcriptome of the fetal heart[196, 310, 311], the paucity of EC in these datasets have limited efforts to comprehensively profile the heterogeneity and underling dynamical cellular changes in endothelium the fetal heart. In contrast to these studies, we opted to isolate fetal cardiac EC by FACS prior to performing scRNA-seq, enabling for the transcriptional profile of over 10,000 individual EC to be captured. Subsequent unsupervised clustering and UMAP visualisation revealed the presence of distinct EC subtypes corresponding to capillary, venous, arterial, endocardial, valvular, and lymphatic populations, each expressing a separate transcriptional signature.

Gene regulatory network analysis, conducted to identify TF responsible for underpinning the transcriptional state of identified EC subpopulations, revealed an enrichment of the *SMAD1* regulon within a capillary population defined by differential expression of *KIT.* Although not explored further in this thesis, it would first be important to validate the co-localisation of *KIT* and *SMAD1* expression by in-situ hybridisation (ISH) prior to further investigation. Subsequent studies may also include determining if the overexpression or knockdown of *SMAD1* in hESC-EC leads to alterations in *KIT* expression, thus validating our *in-silico* prediction. Furthermore, given the reported role of KIT in hypoxia[340] it would be interesting to ascertain if *SMAD1* expression is also altered following the exposure of EC to a hypoxic environment *in vitro,* and if this results in differences in angiogenic capability, determined using the tube formation assay.

Trajectory inference analysis identified a contribution of the endocardium to the coronary vasculature via a population defined by enriched expression of venous EC markers. Although the endocardium has been demonstrated to provide a substantial contribution to the coronary vasculature during murine cardiac development[300, 328, 339], it is not known if endocardial cells first acquire a venous identity during this transition. Given our finding that *NPR3* expression is highly specific to the endocardium during mouse and human cardiac development, the previously characterised *Npr3* Cre mouse line[328, 360] could be used to determine if endocardial cells transitioning to vascular endothelium, either by angiogenesis[300] or by segregation of trabeculae[339], express markers

of venous EC. Additionally, recent published studies have described methods for the derivation of hPSC derived endocardial cells[361-363]. Whilst these studies have primarily used these cells as a model of valvulogenesis by inducing EndMT, their suitability as a human *in vitro* model of coronary vascular formation is yet to be explored. The recent identification of VEGF-B as a positive regulator of the transition of endocardial cells to coronary vascular EC[364] may provide a method to recapitulate this process using hPSC derived endocardial cells, thus allowing the accompanying transcriptional changes to be characterised.

Trajectory inference analysis also demonstrated the transition of capillary EC to arterial EC within the human fetal heart. This finding aligns closely with murine studies also focusing on a comparatively late stage of murine cardiac development[10, 301, 307]. Both gene regulatory network analysis and the detection of significantly differentially expressed genes across pseudotime identified the TF *MECOM* as a key regulator of arterial EC identity. Subsequent perturbation of *MECOM* expression in arterial-like hESC-EC demonstrated that although knockdown of *MECOM* does not alter arterial marker expression, it does result in an upregulation of venous markers, *NR2F2* and *EPHB4*. This suggests MECOM is required during arterial specification to repress venous identity. Despite identifying this functional role, the mechanism by which MECOM influences arteriovenous specification remains unknown and should be subject to further investigation. Advances in chromatin immunoprecipitation assays with sequencing (ChIP-Seq) methodologies have provided insight into

the likely transcriptional targets of numerous transcription factors with a demonstrated role in EC biology[98, 365-367]. Notably, these studies included using ChIP-Seq to identify a direct role for NR2F2 in binding to various vein-specific DNA regulatory elements and preventing the activation of an enhancer of arterial marker, *HEY2*, in venous EC[365]. Although not within the scope of this thesis, performing a MECOM ChIP-Seq study in hESC-EC would help identify the direct transcriptional targets of MECOM, thus giving insight into the mechanism by which venous identity is repressed. In addition to these *in vitro* studies, it would also be interesting to determine if the loss of *MECOM* in mouse embryos using an endothelial specific knockout strategy would result in arteriovenous defects, and if these are specific to coronary vasculature. A similar approach has been used previously for other regulators of arteriovenous specification to demonstrate an *in vivo* function[12, 167, 174]. Future studies may also investigate the effect of MECOM overexpression on coronary artery differentiation during murine cardiac development or following myocardial infarction during adulthood. This approach was used recently by Raftrey et al[307], revealing a role for Dach1 in potentiating coronary artery formation.

As discussed in chapter 3, although this study provided substantial insight into the transcriptional regulation of human coronary vascular development, due to limited sample availability our scRNA-seq data was derived from just two fetal samples. In addition to the low number of sample replicates, data was obtained from fetuses at similar gestational ages (13 and 14 weeks) meaning

transcriptional changes over a wider developmental period could not be evaluated. Whilst logistically challenging, following a similar experimental procedure using fetal tissue from earlier developmental stages would provide further insight into the early events of human coronary vasculature formation.

5.4 Assessment of the susceptibility of EC to direct SARS-CoV-2 infection

In response to COVID-19 pandemic we sought to determine if SARS-CoV-2 could contribute to disease pathology via the direct infection the endothelium. Given ACE2 was identified as the main receptor required for the cellular entry of SARS-CoV-2 we first used publicly available RNA sequencing datasets to assess the levels of *ACE2* expression in endothelial cells. Data from bulk RNA sequencing revealed a lack of *ACE2* expression across a range of EC from multiple sources. Similarly, scRNA-seq data from the human heart demonstrated an absence of *ACE2* expression with pericyte contamination likely explaining the rare expression of *ACE2* in cells expressing EC markers.

Viral replication assays using human pulmonary or coronary artery EC demonstrated minimal levels of replicative infection following exposure to SARS-CoV-2 compared to permissive VeroE6 cells. However, high levels of viral replication were observed using a replication competent human coronavirus 229E GFP reporter virus, demonstrating the capability of EC to support coronavirus replication *in vitro*. These findings contradicted previous reports of direct endothelial SARS-CoV-2 infection in post-mortem kidney

samples[368]. However, findings from this earlier study were later disputed suggesting the structures observed in electron microscopy were rough endoplasmic reticulum rather than SARS-CoV-2 particles[369]. Recently published findings from Liu et al[370] also reported EC infection by SARS-CoV-2 using a mouse model expressing human *ACE2*. A possible explanation for this contradictory finding is that the human adenovirus type 5 vector used to express human ACE2 (Ad5-hACE2) also transduced EC, therefore making them permissive to SARS-CoV-2 infection. Endothelial tropism for Ad5 has previously been described[371]. Finally, although our data suggests the lack of ACE2 expression is the most likely cause for the absence of direct EC infection by SARS-CoV-2, further evidence for this could be provided by determining if forced expression of *ACE2* in EC is sufficient to enable productive SARS-CoV-2 replication.

5.5 Conclusions

Collectively, the results contained within this thesis describe the transcriptional dynamics of human endothelial cells throughout their differentiation from pluripotent stem cells and during coronary vascular development. Whilst these results offer important insight into the transcriptional regulation of these processes, further investigation is required to determine the functional relevance of these findings. The identification of MECOM as a regulator of arteriovenous specification highlights the suitability of scRNA-seq data to be used to inform hypothesis-driven approaches of guiding endothelial identity. Furthermore, coupled with the recent advances in single cell omics, improving gene editing and high throughput screening technologies provide new opportunity to identify novel regulators of human endothelial cell differentiation and specification.

References

1. Krüger-Genge A, Blocki A, Franke RP, Jung F. Vascular Endothelial Cell Biology: An Update. Int J Mol Sci. 2019 Sep 7;20(18).
2. Oliver G, Kipnis J, Randolph GJ, Harvey NL. The Lymphatic Vasculature in the 21(st) Century: Novel Functional Roles in Homeostasis and Disease. Cell. 2020 Jul 23;182(2):270-96.
3. Potente M, Mäkinen T. Vascular heterogeneity and specialization in development and disease. Nat Rev Mol Cell Biol. 2017 Aug;18(8):477-94.
4. Park C, Kim TM, Malik AB. Transcriptional regulation of endothelial cell and vascular development. Circ Res. 2013 May 10;112(10):1380-400.
5. Qiu J, Hirschi KK. Endothelial Cell Development and Its Application to Regenerative Medicine. Circ Res. 2019 Aug 2;125(4):489-501.
6. Craig MP, Sumanas S. ETS transcription factors in embryonic vascular development. Angiogenesis. 2016 Jul;19(3):275-85.
7. Meadows SM, Myers CT, Krieg PA. Regulation of endothelial cell development by ETS transcription factors. Semin Cell Dev Biol. 2011 Dec;22(9):976-84.
8. De Val S, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, et al. Combinatorial Regulation of Endothelial Gene Expression by Ets and Forkhead Transcription Factors. Cell. 2008;135(6):1053-64.
9. Neal A, Nornes S, Louphrasitthiphol P, Sacilotto N, Preston MD, Fleisinger L, et al. ETS factors are required but not sufficient for specific patterns of enhancer activity in different endothelial subtypes. Developmental Biology. 2021;473:1-14.
10. Sharma B, Chang A, Red-Horse K. Coronary Artery Development: Progenitor Cells and Differentiation Pathways. Annu Rev Physiol. 2017 Feb 10;79:1-19.
11. Fish JE, Wythe JD. The molecular regulation of arteriovenous specification and maintenance. Dev Dyn. 2015 Mar;244(3):391-409.
12. You L-R, Lin F-J, Lee CT, DeMayo FJ, Tsai M-J, Tsai SY. Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. Nature. 2005;435(7038):98-104.
13. Lawson ND, Scheer N, Pham VN, Kim CH, Chitnis AB, Campos-Ortega JA, et al. Notch signaling is required for arterial-venous differentiation during embryonic vascular development. Development. 2001 Oct;128(19):3675-83.
14. Haque A, Engel J, Teichmann SA, Lonnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. 2017 Aug 18;9(1):75.
15. Chavkin NW, Hirschi KK. Single Cell Analysis in Vascular Biology. Front Cardiovasc Med. 2020;7:42.
16. Kalucka J, de Rooij LPMH, Goveia J, Rohlenova K, Dumas SJ, Meta E, et al. Single-Cell Transcriptome Atlas of Murine Endothelial Cells. Cell. 2020;180(4):764-79.e20.
17. Litviňuková M, Talavera-López C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. Nature. 2020;588(7838):466-72.
18. De Val S, Black BL. Transcriptional control of endothelial cell development. Dev Cell. 2009 Feb;16(2):180-95.

19. Poole TJ, Coffin JD. Vasculogenesis and angiogenesis: two distinct morphogenetic mechanisms establish embryonic vascular pattern. J Exp Zool. 1989 Aug;251(2):224-31.
20. Chong DC, Koo Y, Xu K, Fu S, Cleaver O. Stepwise arteriovenous fate acquisition during mammalian vasculogenesis. Dev Dyn. 2011 Sep;240(9):2153-65.
21. Drake CJ, Fleming PA. Vasculogenesis in the day 6.5 to 9.5 mouse embryo. Blood. 2000;95(5):1671-9.
22. Garcia MD, Larina IV. Vascular development and hemodynamic force in the mouse yolk sac. Front Physiol. 2014;5:308.
23. Misfeldt AM, Boyle SC, Tompkins KL, Bautch VL, Labosky PA, Baldwin HS. Endocardial cells are a distinct endothelial lineage derived from Flk1+ multipotent cardiovascular progenitors. Dev Biol. 2009 Sep 1;333(1):78-89.
24. Yamaguchi TP, Dumont DJ, Conlon RA, Breitman ML, Rossant J. flk-1, an flt-related receptor tyrosine kinase is an early marker for endothelial cell precursors. Development. 1993 Jun;118(2):489-98.
25. Shalaby F, Rossant J, Yamaguchi TP, Gertsenstein M, Wu X-F, Breitman ML, et al. Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature. 1995;376(6535):62-6.
26. Huber TL, Kouskoff V, Joerg Fehling H, Palis J, Keller G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature. 2004;432(7017):625-30.
27. Choi K, Kennedy M, Kazarov A, Papadimitriou JC, Keller G. A common precursor for hematopoietic and endothelial cells. Development. 1998 Feb;125(4):725-32.
28. Yamashita J, Itoh H, Hirashima M, Ogawa M, Nishikawa S, Yurugi T, et al. Flk1-positive cells derived from embryonic stem cells serve as vascular progenitors. Nature. 2000;408(6808):92-6.
29. Kattman SJ, Huber TL, Keller GM. Multipotent flk-1+ cardiovascular progenitor cells give rise to the cardiomyocyte, endothelial, and vascular smooth muscle lineages. Dev Cell. 2006 Nov;11(5):723-32.
30. Cox CM, Poole TJ. Angioblast differentiation is influenced by the local environment: FGF-2 induces angioblasts and patterns vessel formation in the quail embryo. Dev Dyn. 2000 Jun;218(2):371-82.
31. Astorga J, Carlsson P. Hedgehog induction of murine vasculogenesis is mediated by Foxf1 and Bmp4. Development. 2007;134(20):3753-61.
32. Boyd NL, Dhara SK, Rekaya R, Godbey EA, Hasneen K, Rao RR, et al. BMP4 promotes formation of primitive vascular networks in human embryonic stem cell-derived embryoid bodies. Exp Biol Med (Maywood). 2007 Jun;232(6):833-43.
33. Patel-Hett S, D'Amore PA. Signal transduction in vasculogenesis and developmental angiogenesis. Int J Dev Biol. 2011;55(4-5):353-63.
34. Carmeliet P, Ferreira V, Breier G, Pollefeyt S, Kieckens L, Gertsenstein M, et al. Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. Nature. 1996 Apr 4;380(6573):435-9.

35. Simons M, Gordon E, Claesson-Welsh L. Mechanisms and regulation of endothelial VEGF receptor signalling. Nature Reviews Molecular Cell Biology. 2016;17(10):611-25.
36. Apte RS, Chen DS, Ferrara N. VEGF in Signaling and Disease: Beyond Discovery and Development. Cell. 2019 Mar 7;176(6):1248-64.
37. Fong GH, Rossant J, Gertsenstein M, Breitman ML. Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium. Nature. 1995 Jul 6;376(6535):66-70.
38. Vogeli KM, Jin S-W, Martin GR, Stainier DYR. A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. Nature. 2006;443(7109):337-9.
39. Kennedy M, D'Souza SL, Lynch-Kattman M, Schwantz S, Keller G. Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood. 2007;109(7):2679-87.
40. Ueno H, Weissman IL. Clonal Analysis of Mouse Development Reveals a Polyclonal Origin for Yolk Sac Blood Islands. Developmental Cell. 2006;11(4):519-33.
41. Padrón-Barthe L, Temiño S, Villa del Campo C, Carramolino L, Isern J, Torres M. Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. Blood. 2014;124(16):2523-32.
42. Lacaud G, Kouskoff V. Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. Experimental Hematology. 2017;49:19-24.
43. Amaya E. The hemangioblast: a state of competence. Blood. 2013;122(24):3853-4.
44. Zovein AC, Hofmann JJ, Lynch M, French WJ, Turlo KA, Yang Y, et al. Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. Cell Stem Cell. 2008;3(6):625-36.
45. Frame JM, Fegan KH, Conway SJ, McGrath KE, Palis J. Definitive Hematopoiesis in the Yolk Sac Emerges from Wnt-Responsive Hemogenic Endothelium Independently of Circulation and Arterial Identity. STEM CELLS. 2016;34(2):431-44.
46. Lancrin C, Sroczynska P, Stephenson C, Allen T, Kouskoff V, Lacaud G. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature. 2009;457(7231):892-5.
47. Pearson S, Lancrin C, Lacaud G, Kouskoff V. The Sequential Expression of CD40 and Icam2 Defines Progressive Steps in the Formation of Blood Precursors from the Mesoderm Germ Layer. STEM CELLS. 2010;28(6):1089-98.
48. Pardanaud L, Luton D, Prigent M, Bourcheix LM, Catala M, Dieterlen-Lievre F. Two distinct endothelial lineages in ontogeny, one of them related to hemopoiesis. Development. 1996 May;122(5):1363-71.
49. Stone OA, Stainier DYR. Paraxial Mesoderm Is the Major Source of Lymphatic Endothelium. Dev Cell. 2019 Jul 22;50(2):247-55 e3.
50. Bardot ES, Hadjantonakis A-K. Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. Mechanisms of Development. 2020;163:103617.

51. Winnier G, Blessing M, Labosky PA, Hogan BL. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. 1995 Sep 1;9(17):2105-16.
52. Zhang P, Li J, Tan Z, Wang C, Liu T, Chen L, et al. Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. Blood. 2008;111(4):1933-41.
53. Richter A, Valdimarsdottir L, Hrafnkelsdottir HE, Runarsson JF, Omarsdottir AR, Ward-van Oostwaard D, et al. BMP4 promotes EMT and mesodermal commitment in human embryonic stem cells via SLUG and MSX2. Stem Cells. 2014 Mar;32(3):636-48.
54. Row RH, Pegg A, Kinney BA, Farr GH, 3rd, Maves L, Lowell S, et al. BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix transcription factor activity. Elife. 2017.
55. Martin Benjamin L, Kimelman D. Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. Developmental Cell. 2012;22(1):223-32.
56. Tonegawa A, Takahashi Y. Somitogenesis controlled by Noggin. Dev Biol. 1998 Oct 15;202(2):172-82.
57. Tani S, Chung U-i, Ohba S, Hojo H. Understanding paraxial mesoderm development and sclerotome specification for skeletal repair. Experimental & Molecular Medicine. 2020;52(8):1166-77.
58. Mayeuf-Louchart A, Lagha M, Danckaert A, Rocancourt D, Relaix F, Vincent SD, et al. Notch regulation of myogenic versus endothelial fates of cells that migrate from the somite to the limb. Proceedings of the National Academy of Sciences. 2014;111(24):8844.
59. Plein A, Fantin A, Denti L, Pollard JW, Ruhrberg C. Erythro-myeloid progenitors contribute endothelial cells to blood vessels. Nature. 2018;562(7726):223-8.
60. Palis J, Yoder MC. Endothelial Cells Transition to Blood Cells but Probably Not Back Again. Circ Res. 2020 Oct 23;127(10):1233-5.
61. Ottersbach K. Endothelial-to-haematopoietic transition: an update on the process of making blood. Biochem Soc Trans. 2019 Apr 30;47(2):591-601.
62. Feng T, Gao Z, Kou S, Huang X, Jiang Z, Lu Z, et al. No Evidence for Erythro-Myeloid Progenitor-Derived Vascular Endothelial Cells in Multiple Organs. Circ Res. 2020 Oct 23;127(10):1221-32.
63. Korhonen J, Lahtinen I, Halmekytö M, Alhonen L, Jänne J, Dumont D, et al. Endothelial-Specific Gene Expression Directed by the tie Gene Promoter In Vivo. Blood. 1995;86(5):1828-35.
64. Kappel A, Schlaeger TM, Flamme I, Orkin SH, Risau W, Breier G. Role of SCL/Tal-1, GATA, and Ets transcription factor binding sites for the regulation of Flk-1 expression during murine vascular development. Blood. 2000;96(9):3078-85.
65. Pimanda JE, Chan WYI, Donaldson IJ, Bowen M, Green AR, Göttgens B. Endoglin expression in the endothelium is regulated by Fli-1, Erg, and Elf-1 acting on the promoter and a –8-kb enhancer. Blood. 2006;107(12):4737-45.

66. Hollenhorst PC, Jones DA, Graves BJ. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. Nucleic Acids Res. 2004;32(18):5693-702.
67. Wei GH, Badis G Fau - Berger MF, Berger Mf Fau - Kivioja T, Kivioja T Fau - Palin K, Palin K Fau - Enge M, Enge M Fau - Bonke M, et al. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. EMBO J. 2010;29(13)(1460-2075 (Electronic)):2147-60.
68. Klambt C. The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. Development. 1993;117(1):163-76.
69. Carrère S, Verger A, Flourens A, Stehelin D, Duterque-Coquillaud M. Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. Oncogene. 1998;16(25):3261-8.
70. Lopez RG, Carron C, Oury C, Gardellin P, Bernard O, Ghysdael J. TEL is a sequence-specific transcriptional repressor. J Biol Chem. 1999 Oct 15;274(42):30132-8.
71. Heo SH, Cho JY. ELK3 suppresses angiogenesis by inhibiting the transcriptional activity of ETS-1 on MT1-MMP. Int J Biol Sci. 2014;10(4):438-47.
72. Chen YH, Layne MD, Chung SW, Ejima K, Baron RM, Yet SF, et al. Elk-3 is a transcriptional repressor of nitric-oxide synthase 2. J Biol Chem. 2003 Oct 10;278(41):39572-7.
73. Zheng H, Wasylyk C, Ayadi A, Abecassis J, Schalken JA, Rogatsch H, et al. The transcription factor Net regulates the angiogenic switch. Genes Dev. 2003 Sep 15;17(18):2283-97.
74. Craig MP, Grajevskaja V, Liao HK, Balciuniene J, Ekker SC, Park JS, et al. Etv2 and fli1b function together as key regulators of vasculogenesis and angiogenesis. Arterioscler Thromb Vasc Biol. 2015 Apr;35(4):865-76.
75. Ferdous A, Caprioli A, Iacovino M, Martin CM, Morris J, Richardson JA, et al. Nkx2-5 transactivates the Ets-related protein 71 gene and specifies an endothelial/endocardial fate in the developing embryo. Proc Natl Acad Sci U S A. 2009 Jan 20;106(3):814-9.
76. Lee D, Park C, Lee H, Lugus JJ, Kim SH, Arentson E, et al. ER71 Acts Downstream of BMP, Notch, and Wnt Signaling in Blood and Vessel Progenitor Specification. Cell Stem Cell. 2008;2(5):497-507.
77. Sumanas S, Lin S. Ets1-related protein is a key regulator of vasculogenesis in zebrafish. PLoS Biol. 2006 Jan;4(1):e10.
78. Brown TA, McKnight SL. Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. Genes Dev. 1992 Dec;6(12b):2502-12.
79. Kataoka H, Hayashi M, Kobayashi K, Ding G, Tanaka Y, Nishikawa S-I. Region-specific Etv2 ablation revealed the critical origin of hemogenic capacity from Hox6-positive caudal-lateral primitive mesoderm. Experimental Hematology. 2013;41(6):567-81.e9.

80. Rasmussen TL, Kweon J, Diekmann MA, Belema-Bedada F, Song Q, Bowlin K, et al. ER71 directs mesodermal fate decisions during embryogenesis. Development. 2011 8/17/2021;138(21):4801-12.
81. Sakurai H, Era T, Jakt LM, Okada M, Nakai S, Nishikawa S. In vitro modeling of paraxial and lateral mesoderm differentiation reveals early reversibility. Stem Cells. 2006 Mar;24(3):575-86.
82. Kataoka H, Takakura N, Nishikawa S, Tsuchida K, Kodama H, Kunisada T, et al. Expressions of PDGF receptor alpha, c-Kit and Flk1 genes clustering in mouse chromosome 5 define distinct subsets of nascent mesodermal cells. Development, Growth & Differentiation. 1997;39(6):729-40.
83. Kataoka H, Hayashi M, Nakagawa R, Tanaka Y, Izumi N, Nishikawa S, et al. Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRα+ primitive mesoderm. Blood. 2011 8/17/2021;118(26):6975-86.
84. Rasmussen TL, Shi X, Wallis A, Kweon J, Zirbes KM, Koyano-Nakagawa N, et al. VEGF/Flk1 signaling cascade transactivates Etv2 gene expression. PLoS One. 2012;7(11):e50103.
85. Rasmussen TL, Martin CM, Walter CA, Shi X, Perlingeiro R, Koyano-Nakagawa N, et al. Etv2 rescues Flk1 mutant embryoid bodies. Genesis. 2013 Jul;51(7):471-80.
86. Liu F, Kang I, Park C, Chang LW, Wang W, Lee D, et al. ER71 specifies Flk-1+ hemangiogenic mesoderm by inhibiting cardiac mesoderm and Wnt signaling. Blood. 2012 Apr 5;119(14):3295-305.
87. Chestnut B, Casie Chetty S, Koenig AL, Sumanas S. Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. Nature Communications. 2020;11(1):2796.
88. Koyano-Nakagawa N, Garry DJ. Etv2 as an essential regulator of mesodermal lineage development. Cardiovascular Research. 2017;113(11):1294-306.
89. Park C, Lee TJ, Bhang SH, Liu F, Nakamura R, Oladipupo SS, et al. Injury-Mediated Vascular Regeneration Requires Endothelial ER71/ETV2. Arterioscler Thromb Vasc Biol. 2016 Jan;36(1):86-96.
90. Shah AV, Birdsey GM, Randi AM. Regulation of endothelial homeostasis, vascular development and angiogenesis by the transcription factor ERG. Vascular Pharmacology. 2016;86:3-13.
91. Liu F, Patient R. Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. Circ Res. 2008 Nov 7;103(10):1147-54.
92. Yuan L, Le Bras A, Sacharidou A, Itagaki K, Zhan Y, Kondo M, et al. ETS-related gene (ERG) controls endothelial cell permeability via transcriptional regulation of the claudin 5 (CLDN5) gene. J Biol Chem. 2012 Feb 24;287(9):6582-91.
93. Birdsey GM, Dryden NH, Amsellem V, Gebhardt F, Sahnan K, Haskard DO, et al. Transcription factor Erg regulates angiogenesis and endothelial apoptosis through VE-cadherin. Blood. 2008 Apr 1;111(7):3498-506.
94. Schwachtgen JL, Janel N, Barek L, Duterque-Coquillaud M, Ghysdael J, Meyer D, et al. Ets transcription factors bind and transactivate the core promoter of the von Willebrand factor gene. Oncogene. 1997 Dec 18;15(25):3091-102.

95. Wythe JD, Dang LT, Devine WP, Boudreau E, Artap ST, He D, et al. ETS factors regulate Vegf-dependent arterial specification. Dev Cell. 2013 Jul 15;26(1):45-58.
96. Meadows SM, Salanga MC, Krieg PA. Kruppel-like factor 2 cooperates with the ETS family protein ERG to activate Flk1 expression during vascular development. Development. 2009 Apr;136(7):1115-25.
97. Vijayaraj P, Le Bras A, Mitchell N, Kondo M, Juliao S, Wasserman M, et al. Erg is a crucial regulator of endocardial-mesenchymal transformation during cardiac valve morphogenesis. Development. 2012 8/18/2021;139(21):3973-85.
98. Kalna V, Yang Y, Peghaire CR, Frudd K, Hannah R, Shah AV, et al. The Transcription Factor ERG Regulates Super-Enhancers Associated With an Endothelial-Specific Gene Expression Program. Circ Res. 2019 Apr 26;124(9):1337-49.
99. Vlaeminck-Guillem V, Carrere S, Dewitte F, Stehelin D, Desbiens X, Duterque-Coquillaud M. The Ets family member Erg gene is expressed in mesodermal tissues and neural crests at fundamental steps during mouse embryogenesis. Mechanisms of Development. 2000;91(1):331-5.
100. Mohamed AA, Tan SH, Mikhalkevich N, Ponniah S, Vasioukhin V, Bieberich CJ, et al. Ets family protein, erg expression in developing and adult mouse tissues by a highly specific monoclonal antibody. J Cancer. 2010 Oct 25;1:197-208.
101. Ellett F, Kile BT, Lieschke GJ. The role of the ETS factor erg in zebrafish vasculogenesis. Mechanisms of Development. 2009;126(3):220-9.
102. Baltzinger M, Mager-Heckel A-M, Remy P. Xl erg: Expression pattern and overexpression during development plead for a role in endothelial cell differentiation. Developmental Dynamics. 1999;216(4-5):420-33.
103. Nikolova-Krstevski V, Yuan L, Le Bras A, Vijayaraj P, Kondo M, Gebauer I, et al. ERG is required for the differentiation of embryonic stem cells along the endothelial lineage. BMC Developmental Biology. 2009;9(1):72.
104. Birdsey Graeme M, Shah Aarti V, Dufton N, Reynolds Louise E, Osuna Almagro L, Yang Y, et al. The Endothelial Transcription Factor ERG Promotes Vascular Stability and Growth through Wnt/β-Catenin Signaling. Developmental Cell. 2015;32(1):82-96.
105. Han R, Pacifici M, Iwamoto M, Trojanowska M. Endothelial Erg expression is required for embryogenesis and vascular integrity. Organogenesis. 2015;11(2):75-86.
106. Lathen C, Zhang Y, Chow J, Singh M, Lin G, Nigam V, et al. ERG-APLNR axis controls pulmonary venule endothelial proliferation in pulmonary veno-occlusive disease. Circulation. 2014 Sep 30;130(14):1179-91.
107. Thompson MA, Ransom DG, Pratt SJ, MacLennan H, Kieran MW, Detrich HW, et al. The cloche and spade tail Genes Differentially Affect Hematopoiesis and Vasculogenesis. Developmental Biology. 1998;197(2):248-69.
108. Abedin MJ, Nguyen A, Jiang N, Perry CE, Shelton JM, Watson DK, et al. Fli1 acts downstream of Etv2 to govern cell survival and vascular homeostasis via positive autoregulation. Circ Res. 2014 May 23;114(11):1690-9.
109. Liu F, Walmsley M, Rodaway A, Patient R. Fli1 Acts at the Top of the Transcriptional Network Driving Blood and Endothelial Development. Current Biology. 2008;18(16):1234-40.

110. Zhao H, Zhao Y, Li Z, Ouyang Q, Sun Y, Zhou D, et al. FLI1 and PKC co-activation promote highly efficient differentiation of human embryonic stem cells into endothelial-like cells. Cell Death & Disease. 2018;9(2):131.
111. Spyropoulos Demetri D, Pharr Pamela N, Lavenburg Kim R, Jackers P, Papas Takis S, Ogawa M, et al. Hemorrhage, Impaired Hematopoiesis, and Lethality in Mouse Embryos Carrying a Targeted Disruption of the Fli1Transcription Factor. Molecular and Cellular Biology. 2000;20(15):5643-52.
112. Marks-Bluth J, Khanna A, Chandrakanthan V, Thoms J, Bee T, Eich C, et al. SMAD1 and SMAD5 Expression Is Coordinately Regulated by FLI1 and GATA2 during Endothelial Development. Mol Cell Biol. 2015 Jun;35(12):2165-72.
113. Matrone G, Xia B, Chen K, Denvir MA, Baker AH, Cooke JP. Fli1 + cells transcriptional analysis reveals an Lmo2–Prdm16 axis in angiogenesis. Proceedings of the National Academy of Sciences. 2021;118(31):e2008559118.
114. Chen Z, Fisher RJ, Riggs CW, Rhim JS, Lautenberger JA. Inhibition of vascular endothelial growth factor-induced endothelial cell migration by ETS1 antisense oligonucleotides. Cancer Res. 1997 May 15;57(10):2013-9.
115. Oda N, Abe M, Sato Y. ETS-1 converts endothelial cells to the angiogenic phenotype by inducing the expression of matrix metalloproteinases and integrin beta3. J Cell Physiol. 1999 Feb;178(2):121-32.
116. Harel S, Sanchez V, Moamer A, Sanchez-Galan JE, Abid Hussein MN, Mayaki D, et al. ETS1, ELK1, and ETV4 Transcription Factors Regulate Angiopoietin-1 Signaling and the Angiogenic Response in Endothelial Cells. Frontiers in Physiology. 2021;12:1176.
117. Harris TA, Yamakuchi M, Kondo M, Oettgen P, Lowenstein CJ. Ets-1 and Ets-2 regulate the expression of microRNA-126 in endothelial cells. Arterioscler Thromb Vasc Biol. 2010 Oct;30(10):1990-7.
118. Muthusamy N, Barton K, Leiden JM. Defective activation and survival of T cells lacking the Ets-1 transcription factor. Nature. 1995 Oct 19;377(6550):639-42.
119. Wei G, Srinivasan R, Cantemir-Stone CZ, Sharma SM, Santhanam R, Weinstein M, et al. Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. Blood. 2009;114(5):1123-30.
120. Casie Chetty S, Sumanas S. Ets1 functions partially redundantly with Etv2 to promote embryonic vasculogenesis and angiogenesis in zebrafish. Developmental Biology. 2020;465(1):11-22.
121. Roukens MG, Alloul-Ramdhani M, Baan B, Kobayashi K, Peterson-Maduro J, van Dam H, et al. Control of endothelial sprouting by a Tel–CtBP complex. Nature Cell Biology. 2010;12(10):933-42.
122. Wang LC, Kuo F, Fujiwara Y, Gilliland DG, Golub TR, Orkin SH. Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. Embo j. 1997 Jul 16;16(14):4374-83.
123. Li L, Rispoli R, Patient R, Ciau-Uitz A, Porcher C. Etv6 activates vegfa expression through positive and negative transcriptional regulatory networks in Xenopus embryos. Nature Communications. 2019;10(1):1083.
124. Golson ML, Kaestner KH. Fox transcription factors: from development to disease. Development. 2016 Dec 15;143(24):4558-70.

125. Furuyama T, Kitayama K, Shimoda Y, Ogawa M, Sone K, Yoshida-Araki K, et al. Abnormal angiogenesis in Foxo1 (Fkhr)-deficient mice. J Biol Chem. 2004 Aug 13;279(33):34741-9.
126. Paik J-H, Kollipara R, Chu G, Ji H, Xiao Y, Ding Z, et al. FoxOs Are Lineage-Restricted Redundant Tumor Suppressors and Regulate Endothelial Cell Homeostasis. Cell. 2007;128(2):309-23.
127. Wilhelm K, Happel K, Eelen G, Schoors S, Oellerich MF, Lim R, et al. FOXO1 couples metabolic activity and growth state in the vascular endothelium. Nature. 2016;529(7585):216-20.
128. Andrade J, Shi C, Costa ASH, Choi J, Kim J, Doddaballapur A, et al. Control of endothelial quiescence by FOXO-regulated metabolites. Nature Cell Biology. 2021;23(4):413-23.
129. Winnier GE, Kume T, Deng K, Rogers R, Bundy J, Raines C, et al. Roles for the Winged Helix Transcription Factors MF1 and MFH1 in Cardiovascular Development Revealed by Nonallelic Noncomplementation of Null Alleles. Developmental Biology. 1999;213(2):418-31.
130. Petrova TV, Karpanen T, Norrmén C, Mellor R, Tamakoshi T, Finegold D, et al. Defective valves and abnormal mural cell recruitment underlie lymphatic vascular failure in lymphedema distichiasis. Nature Medicine. 2004;10(9):974-81.
131. Norden PR, Sabine A, Wang Y, Demir CS, Liu T, Petrova TV, et al. Shear stimulation of FOXC1 and FOXC2 differentially regulates cytoskeletal activity during lymphatic valve maturation. eLife. 2020;9:e53814
C1 - eLife 2020;9:e.
132. Seo S, Fujita H, Nakano A, Kang M, Duarte A, Kume T. The forkhead transcription factors, Foxc1 and Foxc2, are required for arterial specification and lymphatic sprouting during vascular development. Developmental Biology. 2006;294(2):458-70.
133. Hayashi H, Kume T. Foxc transcription factors directly regulate Dll4 and Hey2 expression by interacting with the VEGF-Notch signaling pathways in endothelial cells. PLoS One. 2008 Jun 11;3(6):e2401.
134. Mahlapuu M, Ormestad M, Enerback S, Carlsson P. The forkhead transcription factor Foxf1 is required for differentiation of extra-embryonic and lateral plate mesoderm. Development. 2001;128(2):155-66.
135. Choi J, Dong L, Ahn J, Dao D, Hammerschmidt M, Chen J-N. FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. Developmental Biology. 2007;304(2):735-44.
136. She Z-Y, Yang W-X. SOX family transcription factors involved in diverse cellular events during development. European Journal of Cell Biology. 2015;94(12):547-63.
137. Lilly AJ, Lacaud G, Kouskoff V. SOXF transcription factors in cardiovascular development. Seminars in Cell & Developmental Biology. 2017;63:50-7.
138. Corada M, Orsenigo F, Morini MF, Pitulescu ME, Bhat G, Nyqvist D, et al. Sox17 is indispensable for acquisition and maintenance of arterial identity. Nat Commun. 2013;4:2609.

139. Zhou Y, Williams J, Smallwood PM, Nathans J. Sox7, Sox17, and Sox18 Cooperatively Regulate Vascular Development in the Mouse Retina. PLoS One. 2015;10(12):e0143650.
140. Lilly AJ, Mazan A, Scott DA, Lacaud G, Kouskoff V. SOX7 expression is critically required in FLK1-expressing cells for vasculogenesis and angiogenesis during mouse embryonic development. Mech Dev. 2017 Aug;146:31-41.
141. Kim K, Kim IK, Yang JM, Lee E, Koh BI, Song S, et al. SoxF Transcription Factors Are Positive Feedback Regulators of VEGF Signaling. Circ Res. 2016 Sep 16;119(7):839-52.
142. Pennisi D, Bowles J, Nagy A, Muscat G, Koopman P. Mice Null for Sox18 Are Viable and Display a Mild Coat Defect. Molecular and Cellular Biology. 2000 2021/08/23;20(24):9331-6.
143. Hosking B, François M, Wilhelm D, Orsenigo F, Caprini A, Svingen T, et al. Sox7 and Sox17 are strain-specific modifiers of the lymphangiogenic defects caused by Sox18 dysfunction in mice. Development. 2009;136(14):2385-91.
144. François M, Caprini A, Hosking B, Orsenigo F, Wilhelm D, Browne C, et al. Sox18 induces development of the lymphatic vasculature in mice. Nature. 2008;456(7222):643-7.
145. Duong T, Koltowska K, Pichol-Thievend C, Le Guen L, Fontaine F, Smith KA, et al. VEGFD regulates blood vascular development by modulating SOX18 activity. Blood. 2014;123(7):1102-12.
146. Gering M, Rodaway AR, Göttgens B, Patient RK, Green AR. The SCL gene specifies haemangioblast development from early mesoderm. Embo j. 1998 Jul 15;17(14):4029-45.
147. Mead PE, Kelley CM, Hahn PS, Piedad O, Zon LI. SCL specifies hematopoietic mesoderm in Xenopus embryos. Development. 1998;125(14):2611-20.
148. Shivdasani RA, Mayer EL, Orkin SH. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature. 1995 Feb 2;373(6513):432-4.
149. Visvader JE, Fujiwara Y, Orkin SH. Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. Genes Dev. 1998 Feb 15;12(4):473-9.
150. Dooley KA, Davidson AJ, Zon LI. Zebrafish scl functions independently in hematopoietic and endothelial development. Developmental Biology. 2005;277(2):522-36.
151. Deleuze V, Chalhoub E, El-Hajj R, Dohet C, Le Clech M, Couraud P-O, et al. TAL-1/SCL and Its Partners E47 and LMO2 Up-Regulate VE-Cadherin Expression in Endothelial Cells. Molecular and Cellular Biology. 2007;27(7):2687-97.
152. Van Handel B, Montel-Hagen A, Sasidharan R, Nakano H, Ferrari R, Boogerd Cornelis J, et al. Scl Represses Cardiomyogenesis in Prospective Hemogenic Endothelium and Endocardium. Cell. 2012;150(3):590-605.
153. Org T, Duan D, Ferrari R, Montel-Hagen A, Van Handel B, Kerényi MA, et al. Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. Embo j. 2015 Mar 12;34(6):759-77.
154. Marziano C, Genet G, Hirschi KK. Vascular endothelial cell specification in health and disease. Angiogenesis. 2021;24(2):213-36.

155. Wang HU, Chen ZF, Anderson DJ. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. Cell. 1998 May 29;93(5):741-53.
156. Füller T, Korff T, Kilian A, Dandekar G, Augustin HG. Forward EphB4 signaling in endothelial cells controls cellular repulsion and segregation from ephrinB2 positive cells. J Cell Sci. 2003 Jun 15;116(Pt 12):2461-70.
157. Casie Chetty S, Rost MS, Enriquez JR, Schumacher JA, Baltrunaite K, Rossi A, et al. Vegf signaling promotes vascular endothelial differentiation by modulating etv2 expression. Developmental Biology. 2017;424(2):147-61.
158. Akeson A, Herman A, Wiginton D, Greenberg J. Endothelial cell activation in a VEGF-A gradient: Relevance to cell fate decisions. Microvascular Research. 2010;80(1):65-74.
159. Lawson ND, Vogel AM, Weinstein BM. sonic hedgehog and vascular endothelial growth factor Act Upstream of the Notch Pathway during Arterial Endothelial Differentiation. Developmental Cell. 2002;3(1):127-36.
160. Wilkinson RN, Koudijs MJ, Patient RK, Ingham PW, Schulte-Merker S, van Eeden FJ. Hedgehog signaling via a calcitonin receptor-like receptor can induce arterial differentiation independently of VEGF signaling in zebrafish. Blood. 2012 Jul 12;120(2):477-88.
161. Hong CC, Peterson QP, Hong J-Y, Peterson RT. Artery/Vein Specification Is Governed by Opposing Phosphatidylinositol-3 Kinase and MAP Kinase/ERK Signaling. Current Biology. 2006;16(13):1366-72.
162. Sacilotto N, Monteiro R, Fritzsche M, Becker PW, Sanchez-Del-Campo L, Liu K, et al. Analysis of Dll4 regulation reveals a combinatorial role for Sox and Notch in arterial development. Proc Natl Acad Sci U S A. 2013 Jul 16;110(29):11893-8.
163. Deng Y, Larrivée B, Zhuang ZW, Atri D, Moraes F, Prahst C, et al. Endothelial RAF1/ERK activation regulates arterial morphogenesis. Blood. 2013;121(19):3988-96.
164. Grego-Bessa J, Luna-Zurita L, del Monte G, Bolós V, Melgar P, Arandilla A, et al. Notch signaling is essential for ventricular chamber development. Dev Cell. 2007 Mar;12(3):415-29.
165. Duarte A, Hirashima M, Benedito R, Trindade A, Diniz P, Bekman E, et al. Dosage-sensitive requirement for mouse Dll4 in artery development. Genes Dev. 2004 Oct 15;18(20):2474-8.
166. Krebs LT, Shutter JR, Tanigaki K, Honjo T, Stark KL, Gridley T. Haploinsufficient lethality and formation of arteriovenous malformations in Notch pathway mutants. Genes Dev. 2004 Oct 15;18(20):2469-73.
167. Fischer A, Schumacher N, Maier M, Sendtner M, Gessler M. The Notch target genes Hey1 and Hey2 are required for embryonic vascular development. Genes Dev. 2004 Apr 15;18(8):901-11.
168. Hofmann JJ, Iruela-Arispe ML. Notch signaling in blood vessels: who is talking to whom about what? Circ Res. 2007 Jun 8;100(11):1556-68.
169. Krebs LT, Xue Y, Norton CR, Shutter JR, Maguire M, Sundberg JP, et al. Notch signaling is essential for vascular morphogenesis in mice. Genes Dev. 2000 Jun 1;14(11):1343-52.

170. Corada M, Nyqvist D, Orsenigo F, Caprini A, Giampietro C, Taketo MM, et al. The Wnt/beta-catenin pathway modulates vascular remodeling and specification by upregulating Dll4/Notch signaling. Dev Cell. 2010 Jun 15;18(6):938-49.
171. Yamamizu K, Matsunaga T, Uosaki H, Fukushima H, Katayama S, Hiraoka-Kanie M, et al. Convergence of Notch and beta-catenin signaling induces arterial fate in vascular progenitors. J Cell Biol. 2010 Apr 19;189(2):325-38.
172. Chiang IK, Fritzsche M, Pichol-Thievend C, Neal A, Holmes K, Lagendijk A, et al. SoxF factors induce Notch1 expression via direct transcriptional regulation during early arterial development. Development. 2017 Jul 15;144(14):2629-39.
173. Urness LD, Sorensen LK, Li DY. Arteriovenous malformations in mice lacking activin receptor-like kinase-1. Nat Genet. 2000 Nov;26(3):328-31.
174. Tual-Chalot S, Mahmoud M, Allinson KR, Redgrave RE, Zhai Z, Oh SP, et al. Endothelial depletion of Acvrl1 in mice leads to arteriovenous malformations associated with reduced endoglin expression. PLoS One. 2014;9(6):e98646.
175. Corti P, Young S, Chen CY, Patrick MJ, Rochon ER, Pekkan K, et al. Interaction between alk1 and blood flow in the development of arteriovenous malformations. Development. 2011 Apr;138(8):1573-82.
176. Laux DW, Young S, Donovan JP, Mansfield CJ, Upton PD, Roman BL. Circulating Bmp10 acts through endothelial Alk1 to mediate flow-dependent arterial quiescence. Development. 2013 Aug;140(16):3403-12.
177. Larrivée B, Prahst C, Gordon E, del Toro R, Mathivet T, Duarte A, et al. ALK1 signaling inhibits angiogenesis by cooperating with the Notch pathway. Dev Cell. 2012 Mar 13;22(3):489-500.
178. Mancini ML, Terzic A, Conley BA, Oxburgh LH, Nicola T, Vary CP. Endoglin plays distinct roles in vascular smooth muscle cell recruitment and regulation of arteriovenous identity during angiogenesis. Dev Dyn. 2009 Oct;238(10):2479-93.
179. Neal A, Nornes S, Payne S, Wallace MD, Fritzsche M, Louphrasitthiphol P, et al. Venous identity requires BMP signalling through ALK3. Nat Commun. 2019 Jan 28;10(1):453.
180. Swift MR, Pham VN, Castranova D, Bell K, Poole RJ, Weinstein BM. SoxF factors and Notch regulate nr2f2 gene expression during venous differentiation in zebrafish. Developmental Biology. 2014;390(2):116-25.
181. Wolf K, Hu H, Isaji T, Dardik A. Molecular identity of arteries, veins, and lymphatics. Journal of Vascular Surgery. 2019;69(1):253-62.
182. Chu M, Li T, Shen B, Cao X, Zhong H, Zhang L, et al. Angiopoietin receptor Tie2 is required for vein specification and maintenance via regulating COUP-TFII. Elife. 2016 Dec 22;5.
183. Davis RB, Curtis CD, Griffin CT. BRG1 promotes COUP-TFII expression and venous specification during embryonic vascular development. Development. 2013 8/27/2021;140(6):1272-81.
184. le Noble F, Moyon D, Pardanaud L, Yuan L, Djonov V, Matthijsen R, et al. Flow regulates arterial-venous differentiation in the chick embryo yolk sac. Development. 2004 Jan;131(2):361-75.

185. Hwa JJ, Beckouche N, Huang L, Kram Y, Lindskog H, Wang RA. Abnormal arterial-venous fusions and fate specification in mouse embryos lacking blood flow. Scientific Reports. 2017;7(1):11965.
186. Buschmann I, Pries A, Styp-Rekowska B, Hillmeister P, Loufrani L, Henrion D, et al. Pulsatile shear and Gja5 modulate arterial identity and remodeling events during flow-driven arteriogenesis. Development. 2010;137(13):2187-96.
187. Sivarapatna A, Ghaedi M, Le AV, Mendez JJ, Qyang Y, Niklason LE. Arterial specification of endothelial cells derived from human induced pluripotent stem cells in a biomimetic flow bioreactor. Biomaterials. 2015;53:621-33.
188. Mack JJ, Mosqueiro TS, Archer BJ, Jones WM, Sunshine H, Faas GC, et al. NOTCH1 is a mechanosensor in adult arteries. Nature Communications. 2017;8(1):1620.
189. Fang JS, Coon BG, Gillis N, Chen Z, Qiu J, Chittenden TW, et al. Shear-induced Notch-Cx37-p27 axis arrests endothelial cell cycle to enable arterial specification. Nature Communications. 2017;8(1):2149.
190. Luo W, Garcia-Gonzalez I, Fernández-Chacón M, Casquero-Garcia V, Sanchez-Muñoz MS, Mühleder S, et al. Arterialization requires the timely suppression of cell growth. Nature. 2021;589(7842):437-41.
191. Chavkin NW, Genet G, Poulet M, Genet N, Marziano C, Vasavada H, et al. Endothelial Cell Cycle State Determines Propensity for Arterial-Venous Fate. bioRxiv. 2020:2020.08.12.246512.
192. Atri D, Larrivée B, Eichmann A, Simons M. Endothelial signaling and the molecular basis of arteriovenous malformation. Cell Mol Life Sci. 2013 Sep 28.
193. Johnson DW, Berg JN, Baldwin MA, Gallione CJ, Marondel I, Yoon SJ, et al. Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. Nat Genet. 1996 Jun;13(2):189-95.
194. Camp JG, Sekine K, Gerber T, Loeffler-Wirth H, Binder H, Gac M, et al. Multilineage communication regulates human liver bud development from pluripotency. Nature. 2017;546(7659):533-8.
195. Tyser RCV, Mahammadov E, Nakanoh S, Vallier L, Scialdone A, Srinivas S. Single-cell transcriptomic characterization of a gastrulating human embryo. Nature. 2021 Dec;600(7888):285-9.
196. Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, et al. Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. Cell Rep. 2019 Feb 12;26(7):1934-50.e5.
197. Richardson MR, Cahan P, Lai X, Prasain N, Yoder MC, Witzmann FA. Proteomic Analysis Reveals GLUT1 to be a Novel Discriminating Marker of Human Arterial Endothelium In vivo and loss of Venous Identity in Cell Culture. Open Journal of Proteomics and Genomics. 2016;1(1):001-7.
198. Zhang J, Chu LF, Hou Z, Schwartz MP, Hacker T, Vickerman V, et al. Functional characterization of human pluripotent stem cell-derived arterial endothelial cells. Proc Natl Acad Sci U S A. 2017 Jul 25;114(30):E6072-E8.
199. Rufaihah AJ, Huang NF, Kim J, Herold J, Volz KS, Park TS, et al. Human induced pluripotent stem cell-derived endothelial cells exhibit functional heterogeneity. Am J Transl Res. 2013;5(1):21-35.

200. Kaipainen A, Korhonen J, Mustonen T, van Hinsbergh VW, Fang GH, Dumont D, et al. Expression of the fms-like tyrosine kinase 4 gene becomes restricted to lymphatic endothelium during development. Proc Natl Acad Sci U S A. 1995 Apr 11;92(8):3566-70.
201. Banerji S, Ni J, Wang SX, Clasper S, Su J, Tammi R, et al. LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph-specific receptor for hyaluronan. J Cell Biol. 1999 Feb 22;144(4):789-801.
202. Breiteneder-Geleff S, Soleiman A, Kowalski H, Horvat R, Amann G, Kriehuber E, et al. Angiosarcomas express mixed endothelial phenotypes of blood and lymphatic capillaries: podoplanin as a specific marker for lymphatic endothelium. Am J Pathol. 1999 Feb;154(2):385-94.
203. Wigle JT, Oliver G. Prox1 function is required for the development of the murine lymphatic system. Cell. 1999 Sep 17;98(6):769-78.
204. Karkkainen MJ, Haiko P, Sainio K, Partanen J, Taipale J, Petrova TV, et al. Vascular endothelial growth factor C is required for sprouting of the first lymphatic vessels from embryonic veins. Nat Immunol. 2004 Jan;5(1):74-80.
205. Srinivasan RS, Dillard ME, Lagutin OV, Lin FJ, Tsai S, Tsai MJ, et al. Lineage tracing demonstrates the venous origin of the mammalian lymphatic vasculature. Genes Dev. 2007 Oct 1;21(19):2422-32.
206. Klotz L, Norman S, Vieira JM, Masters M, Rohling M, Dube KN, et al. Cardiac lymphatics are heterogeneous in origin and respond to injury. Nature. 2015 Jun 04;522(7554):62-7.
207. Martinez-Corral I, Ulvmar MH, Stanczuk L, Tatin F, Kizhatil K, John SW, et al. Nonvenous origin of dermal lymphatic vasculature. Circ Res. 2015 May 8;116(10):1649-54.
208. Wigle JT, Harvey N, Detmar M, Lagutina I, Grosveld G, Gunn MD, et al. An essential role for Prox1 in the induction of the lymphatic endothelial cell phenotype. Embo j. 2002 Apr 2;21(7):1505-13.
209. Veikkola T, Jussila L, Makinen T, Karpanen T, Jeltsch M, Petrova TV, et al. Signalling via vascular endothelial growth factor receptor-3 is sufficient for lymphangiogenesis in transgenic mice. Embo j. 2001 Mar 15;20(6):1223-31.
210. Srinivasan RS, Geng X, Yang Y, Wang Y, Mukatira S, Studer M, et al. The nuclear hormone receptor Coup-TFII is required for the initiation and early maintenance of Prox1 expression in lymphatic endothelial cells. Genes Dev. 2010 Apr 1;24(7):696-707.
211. Hong YK, Harvey N, Noh YH, Schacht V, Hirakawa S, Detmar M, et al. Prox1 is a master control gene in the program specifying lymphatic endothelial cell fate. Dev Dyn. 2002 Nov;225(3):351-7.
212. Srinivasan RS, Escobedo N, Yang Y, Interiano A, Dillard ME, Finkelstein D, et al. The Prox1-Vegfr3 feedback loop maintains the identity and the number of lymphatic endothelial cell progenitors. Genes Dev. 2014 Oct 1;28(19):2175-87.
213. Maruyama K, Miyagawa-Tomita S, Mizukami K, Matsuzaki F, Kurihara H. Isl1-expressing non-venous cell lineage contributes to cardiac lymphatic vessel development. Dev Biol. 2019 Aug 15;452(2):134-43.

214. Lioux G, Liu X, Temiño S, Oxendine M, Ayala E, Ortega S, et al. A Second Heart Field-Derived Vasculogenic Niche Contributes to Cardiac Lymphatics. Dev Cell. 2020 Feb 10;52(3):350-63.e6.
215. Liu G, David BT, Trawczynski M, Fessler RG. Advances in Pluripotent Stem Cells: History, Mechanisms, Technologies, and Applications. Stem Cell Rev Rep. 2020 Feb;16(1):3-32.
216. Zakrzewski W, Dobrzyński M, Szymonowicz M, Rybak Z. Stem cells: past, present, and future. Stem Cell Res Ther. 2019 Feb 26;10(1):68.
217. Lee JE, Lee DR. Human embryonic stem cells: derivation, maintenance and cryopreservation. Int J Stem Cells. 2011 Jun;4(1):9-17.
218. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998 Nov 6;282(5391):1145-7.
219. Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol. 2000 Apr;18(4):399-404.
220. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006 Aug 25;126(4):663-76.
221. Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: a decade of progress. Nat Rev Drug Discov. 2017 Feb;16(2):115-30.
222. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007 Nov 30;131(5):861-72.
223. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007 Dec 21;318(5858):1917-20.
224. Okita K, Matsumura Y, Sato Y, Okada A, Morizane A, Okamoto S, et al. A more efficient method to generate integration-free human iPS cells. Nat Methods. 2011 May;8(5):409-12.
225. 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 Jpn Acad Ser B Phys Biol Sci. 2009;85(8):348-62.
226. Sone M, Itoh H, Yamahara K, Yamashita JK, Yurugi-Kobayashi T, Nonoguchi A, et al. Pathway for differentiation of human embryonic stem cells to vascular cell components and their potential for vascular regeneration. Arterioscler Thromb Vasc Biol. 2007 Oct;27(10):2127-34.
227. Kaufman DS, Hanson ET, Lewis RL, Auerbach R, Thomson JA. Hematopoietic colony-forming cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A. 2001 Sep 11;98(19):10716-21.
228. Hill KL, Obrtlikova P, Alvarez DF, King JA, Keirstead SA, Allred JR, et al. Human embryonic stem cell-derived vascular progenitor cells capable of endothelial and smooth muscle cell function. Exp Hematol. 2010 Mar;38(3):246-57.e1.

229. Williams IM, Wu JC. Generation of Endothelial Cells From Human Pluripotent Stem Cells. Arterioscler Thromb Vasc Biol. 2019 Jul;39(7):1317-29.
230. Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, et al. Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol Med. 2000 Feb;6(2):88-95.
231. Levenberg S, Golub JS, Amit M, Itskovitz-Eldor J, Langer R. Endothelial cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A. 2002 Apr 2;99(7):4391-6.
232. Goldman O, Feraud O, Boyer-Di Ponio J, Driancourt C, Clay D, Le Bousse-Kerdiles MC, et al. A boost of BMP4 accelerates the commitment of human embryonic stem cells to the endothelial lineage. Stem Cells. 2009 Aug;27(8):1750-9.
233. Nourse MB, Halpin DE, Scatena M, Mortisen DJ, Tulloch NL, Hauch KD, et al. VEGF induces differentiation of functional endothelium from human embryonic stem cells: implications for tissue engineering. Arterioscler Thromb Vasc Biol. 2010 Jan;30(1):80-9.
234. Boulberdaa M, Scott E, Ballantyne M, Garcia R, Descamps B, Angelini GD, et al. A Role for the Long Noncoding RNA SENCR in Commitment and Function of Endothelial Cells. Mol Ther. 2016 May;24(5):978-90.
235. Yang L, Soonpaa MH, Adler ED, Roepke TK, Kattman SJ, Kennedy M, et al. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. Nature. 2008 May 22;453(7194):524-8.
236. Patsch C, Challet-Meylan L, Thoma EC, Urich E, Heckel T, O'Sullivan JF, et al. Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. Nat Cell Biol. 2015;17(8):994-1003.
237. Giacomelli E, Bellin M, Sala L, van Meer BJ, Tertoolen LG, Orlova VV, et al. Three-dimensional cardiac microtissues composed of cardiomyocytes and endothelial cells co-differentiated from human pluripotent stem cells. Development. 2017 Mar 15;144(6):1008-17.
238. Gu M. Efficient Differentiation of Human Pluripotent Stem Cells to Endothelial Cells. Curr Protoc Hum Genet. 2018 Jul 6:e64.
239. Lian X, Bao X, Al-Ahmad A, Liu J, Wu Y, Dong W, et al. Efficient differentiation of human pluripotent stem cells to endothelial progenitors via small-molecule activation of WNT signaling. Stem Cell Reports. 2014 Nov 11;3(5):804-16.
240. Palpant NJ, Pabon L, Friedman CE, Roberts M, Hadland B, Zaunbrecher RJ, et al. Generating high-purity cardiac and endothelial derivatives from patterned mesoderm using human pluripotent stem cells. Nat Protoc. 2017 Jan;12(1):15-31.
241. MacAskill MG, Saif J, Condie A, Jansen MA, MacGillivray TJ, Tavares AAS, et al. Robust Revascularization in Models of Limb Ischemia Using a Clinically Translatable Human Stem Cell-Derived Endothelial Cell Product. Mol Ther. 2018 Jul 5;26(7):1669-84.
242. Ohta R, Niwa A, Taniguchi Y, Suzuki NM, Toga J, Yagi E, et al. Laminin-guided highly efficient endothelial commitment from human pluripotent stem cells. Sci Rep. 2016 Nov 2;6:35680.

243. Kleinman HK, McGarvey ML, Hassell JR, Star VL, Cannon FB, Laurie GW, et al. Basement membrane complexes with biological activity. Biochemistry. 1986 Jan 28;25(2):312-8.
244. Aisenbrey EA, Murphy WL. Synthetic alternatives to Matrigel. Nat Rev Mater. 2020 Jul;5(7):539-51.
245. James D, Nam HS, Seandel M, Nolan D, Janovitz T, Tomishima M, et al. Expansion and maintenance of human embryonic stem cell-derived endothelial cells by TGFbeta inhibition is Id1 dependent. Nat Biotechnol. 2010 Feb;28(2):161-6.
246. Palpant NJ, Pabon L, Roberts M, Hadland B, Jones D, Jones C, et al. Inhibition of β-catenin signaling respecifies anterior-like endothelium into beating human cardiomyocytes. Development. 2015 Sep 15;142(18):3198-209.
247. Paik DT, Tian L, Lee J, Sayed N, Chen IY, Rhee S, et al. Large-Scale Single-Cell RNA-Seq Reveals Molecular Signatures of Heterogeneous Populations of Human Induced Pluripotent Stem Cell-Derived Endothelial Cells. Circ Res. United States2018.
248. Arap W, Kolonin MG, Trepel M, Lahdenranta J, Cardó-Vila M, Giordano RJ, et al. Steps toward mapping the human vasculature by phage display. Nature Medicine. 2002;8(2):121-7.
249. Son NH, Basu D, Samovski D, Pietka TA, Peche VS, Willecke F, et al. Endothelial cell CD36 optimizes tissue fatty acid uptake. J Clin Invest. 2018 Oct 1;128(10):4329-42.
250. Zhao Z, Nelson AR, Betsholtz C, Zlokovic BV. Establishment and Dysfunction of the Blood-Brain Barrier. Cell. 2015 Nov 19;163(5):1064-78.
251. Itkin T, Gur-Cohen S, Spencer JA, Schajnovitz A, Ramasamy SK, Kusumbe AP, et al. Distinct bone marrow blood vessels differentially regulate haematopoiesis. Nature. 2016;532(7599):323-8.
252. Rhee S, Paik DT, Yang JY, Nagelberg D, Williams I, Tian L, et al. Endocardial/endothelial angiocrines regulate cardiomyocyte development and maturation and induce features of ventricular non-compaction. European Heart Journal. 2021 8/31/2021.
253. Lazarus A, Del-Moral PM, Ilovich O, Mishani E, Warburton D, Keshet E. A perfusion-independent role of blood vessels in determining branching stereotypy of lung airways. Development. 2011 Jun;138(11):2359-68.
254. Kim H, Wang M, Paik DT. Endothelial-Myocardial Angiocrine Signaling in Heart Development. Frontiers in Cell and Developmental Biology. 2021;9:1760.
255. Aird WC, Edelberg JM, Weiler-Guettler H, Simmons WW, Smith TW, Rosenberg RD. Vascular bed-specific expression of an endothelial cell gene is programmed by the tissue microenvironment. J Cell Biol. 1997 Sep 8;138(5):1117-24.
256. Neal EH, Marinelli NA, Shi Y, McClatchey PM, Balotin KM, Gullett DR, et al. A Simplified, Fully Defined Differentiation Scheme for Producing Blood-Brain Barrier Endothelial Cells from Human iPSCs. Stem Cell Reports. 2019;12(6):1380-8.
257. Hagberg CE, Falkevall A, Wang X, Larsson E, Huusko J, Nilsson I, et al. Vascular endothelial growth factor B controls endothelial fatty acid uptake. Nature. 2010;464(7290):917-21.
258. Walter TJ, Cast AE, Huppert KA, Huppert SS. Epithelial VEGF signaling is required in the mouse liver for proper sinusoid endothelial cell identity and

hepatocyte zonation in vivo. Am J Physiol Gastrointest Liver Physiol. 2014 May 15;306(10):G849-62.
259. Won D, Zhu S-N, Chen M, Teichert A-M, Fish JE, Matouk CC, et al. Relative Reduction of Endothelial Nitric-Oxide Synthase Expression and Transcription in Atherosclerosis-Prone Regions of the Mouse Aorta and in an in Vitro Model of Disturbed Flow. The American Journal of Pathology. 2007;171(5):1691-704.
260. Aird WC. Endothelial cell heterogeneity. Cold Spring Harb Perspect Med. 2012 Jan;2(1):a006429.
261. Marcu R, Choi YJ, Xue J, Fortin CL, Wang Y, Nagao RJ, et al. Human Organ-Specific Endothelial Cell Heterogeneity. iScience. 2018 Jun 29;4:20-35.
262. You D, Wen X, Gorczyca L, Morris A, Richardson JR, Aleksunes LM. Increased MDR1 Transporter Expression in Human Brain Endothelial Cells Through Enhanced Histone Acetylation and Activation of Aryl Hydrocarbon Receptor Signaling. Mol Neurobiol. 2019 Oct;56(10):6986-7002.
263. Baran-Gale J, Chandra T, Kirschner K. Experimental design for single-cell RNA sequencing. Briefings in Functional Genomics. 2018;17(4):233-9.
264. Vanlandewijck M, He L, Mäe MA, Andrae J, Ando K, Del Gaudio F, et al. A molecular atlas of cell types and zonation in the brain vasculature. Nature. 2018;554(7693):475-80.
265. Dumas SJ, Meta E, Borri M, Goveia J, Rohlenova K, Conchinha NV, et al. Single-Cell RNA Sequencing Reveals Renal Endothelium Heterogeneity and Metabolic Adaptation to Water Deprivation. J Am Soc Nephrol. 2020 Jan;31(1):118-38.
266. Kashima Y, Sakamoto Y, Kaneko K, Seki M, Suzuki Y, Suzuki A. Single-cell sequencing techniques from individual to multiomics analyses. Experimental & Molecular Medicine. 2020;52(9):1419-27.
267. Collaborators GMaCoD. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015 Jan 10;385(9963):117-71.
268. Howard DP, Banerjee A, Fairhead JF, Hands L, Silver LE, Rothwell PM. Population-Based Study of Incidence, Risk Factors, Outcome, and Prognosis of Ischemic Peripheral Arterial Events: Implications for Prevention. Circulation. 2015 Nov 10;132(19):1805-15.
269. Minicucci MF, Azevedo PS, Polegato BF, Paiva SA, Zornoff LA. Heart failure after myocardial infarction: clinical implications and treatment. Clin Cardiol. 2011 Jul;34(7):410-4.
270. Tompkins BA, Balkan W, Winkler J, Gyongyosi M, Goliasch G, Fernandez-Aviles F, et al. Preclinical Studies of Stem Cell Therapy for Heart Disease. Circ Res. United States: 2018 American Heart Association, Inc.; 2018. p. 1006-20.
271. Liu YW, Chen B, Yang X, Fugate JA, Kalucki FA, Futakuchi-Tsuchida A, et al. Human embryonic stem cell-derived cardiomyocytes restore function in infarcted hearts of non-human primates. Nat Biotechnol2018. p. 597-605.
272. Menasche P, Vanneaux V, Hagege A, Bel A, Cholley B, Cacciapuoti I, et al. Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: first clinical case report. Eur Heart J. 2015;36(30):2011-7.

273. MacAskill MG, Saif J, Condie A, Jansen MA, MacGillivray TJ, Tavares AAS, et al. Robust Revascularization in Models of Limb Ischemia Using a Clinically Translatable Human Stem Cell-Derived Endothelial Cell Product. Mol Ther. United States: 2018 The Authors. Published by Elsevier Inc; 2018.
274. Rufaihah AJ, Huang NF, Jame S, Lee JC, Nguyen HN, Byers B, et al. Endothelial cells derived from human iPSCS increase capillary density and improve perfusion in a mouse model of peripheral arterial disease. Arterioscler Thromb Vasc Biol. 2011 Nov;31(11):e72-9.
275. Loh KM, Chen A, Koh PW, Deng TZ, Sinha R, Tsai JM, et al. Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell. 2016 Jul 14;166(2):451-67.
276. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol2018. p. 411-20.
277. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014 Apr;32(4):381-6.
278. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature2018. p. 494-8.
279. Roson-Burgo B, Sanchez-Guijo F, Del Canizo C, De Las Rivas J. Insights into the human mesenchymal stromal/stem cell identity through integrative transcriptomic profiling. BMC Genomics. 2016 Nov 21;17(1):944.
280. Nam SW, Clair T, Kim YS, McMarlin A, Schiffmann E, Liotta LA, et al. Autotaxin (NPP-2), a metastasis-enhancing motogen, is an angiogenic factor. Cancer Res. 2001 Sep 15;61(18):6938-44.
281. Chan B, Yuan HT, Ananth Karumanchi S, Sukhatme VP. Receptor tyrosine kinase Tie-1 overexpression in endothelial cells upregulates adhesion molecules. Biochem Biophys Res Commun2008. p. 475-9.
282. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2014 Mar;15(3):178-96.
283. Evseenko D, Zhu Y, Schenke-Layland K, Kuo J, Latour B, Ge S, et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc Natl Acad Sci U S A. 2010 Aug 3;107(31):13742-7.
284. Guibentif C, Ronn RE, Boiers C, Lang S, Saxena S, Soneji S, et al. Single-Cell Analysis Identifies Distinct Stages of Human Endothelial-to-Hematopoietic Transition. Cell Rep. United States: 2017 The Authors. Published by Elsevier Inc; 2017. p. 10-9.
285. Narlis M, Grote D, Gaitan Y, Boualia SK, Bouchard M. Pax2 and pax8 regulate branching morphogenesis and nephron differentiation in the developing kidney. J Am Soc Nephrol2007. p. 1121-9.
286. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. Cell. United States: 2018 Elsevier Inc; 2018. p. 650-65.
287. Slukvin, II, Kumar A. The mesenchymoangioblast, mesodermal precursor for mesenchymal and endothelial cells. Cell Mol Life Sci. 2018.

288. Vodyanik MA, Yu J, Zhang X, Tian S, Stewart R, Thomson JA, et al. A mesoderm-derived precursor for mesenchymal stem and endothelial cells. Cell Stem Cell. 2010 Dec 3;7(6):718-29.
289. Gertow K, Hirst CE, Yu QC, Ng ES, Pereira LA, Davis RP, et al. WNT3A promotes hematopoietic or mesenchymal differentiation from hESCs depending on the time of exposure. Stem Cell Reports. 2013;1(1):53-65.
290. Cooley BC, Nevado J, Mellad J, Yang D, St Hilaire C, Negro A, et al. TGF-beta signaling mediates endothelial-to-mesenchymal transition (EndMT) during vein graft remodeling. Sci Transl Med. 2014 Mar 12;6(227):227ra34.
291. Yu J, Valerius MT, Duah M, Staser K, Hansard JK, Guo JJ, et al. Identification of molecular compartments and genetic circuitry in the developing mammalian kidney. Development. 2012 May;139(10):1863-73.
292. Atkins GB, Jain MK. Role of Kruppel-like transcription factors in endothelial biology. Circ Res. 2007;100(12):1686-95.
293. Wysoczynski M, Khan A, Bolli R. New Paradigms in Cell Therapy: Repeated Dosing, Intravenous Delivery, Immunomodulatory Actions, and New Cell Types. Circ Res. 2018 Jul 6;123(2):138-58.
294. Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, et al. The adult human testis transcriptional cell atlas. Cell Res. 2018 Dec;28(12):1141-57.
295. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am J Respir Crit Care Med. 2018 Dec 15.
296. Lindstrom NO, Guo J, Kim AD, Tran T, Guo Q, De Sena Brandine G, et al. Conserved and Divergent Features of Mesenchymal Progenitor Cell Types within the Cortical Nephrogenic Niche of the Human and Mouse Kidney. J Am Soc Nephrol. 2018 Mar;29(3):806-24.
297. Kinchen J, Chen HH, Parikh K, Antanaviciute A, Jagielowicz M, Fawkner-Corbett D, et al. Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. Cell. 2018 Oct 4;175(2):372-86 e17.
298. MacParland SA, Liu JC, Ma XZ, Innes BT, Bartczak AM, Gage BK, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun. 2018 Oct 22;9(1):4383.
299. A A, J R. *topGO: Enrichment Analysis for Gene Ontology* . R package version 2.34.0. 2018.
300. Wu B, Zhang Z, Lui W, Chen X, Wang Y, Chamberlain AA, et al. Endocardial Cells Form the Coronary Arteries by Angiogenesis through Myocardial-Endocardial VEGF Signaling. Cell. 2012;151(5):1083-96.
301. Red-Horse K, Ueno H, Weissman IL, Krasnow MA. Coronary arteries form by developmental reprogramming of venous cells. Nature. 2010;464(7288):549-53.
302. Cai C-L, Martin JC, Sun Y, Cui L, Wang L, Ouyang K, et al. A myocardial lineage derives from Tbx18 epicardial cells. Nature. 2008;454(7200):104-8.
303. Katz TC, Singh MK, Degenhardt K, Rivera-Feliciano J, Johnson RL, Epstein JA, et al. Distinct compartments of the proepicardial organ give rise to coronary vascular endothelial cells. Dev Cell. 2012 Mar 13;22(3):639-50.

304. Lupu IE, Redpath AN, Smart N. Spatiotemporal Analysis Reveals Overlap of Key Proepicardial Markers in the Developing Murine Heart. Stem Cell Reports. 2020 May 12;14(5):770-87.
305. Su T, Stanley G, Sinha R, D'Amato G, Das S, Rhee S, et al. Single-cell analysis of early progenitor cells that build coronary arteries. Nature. 2018;559(7714):356-62.
306. Chang AH, Raftrey BC, D'Amato G, Surya VN, Poduri A, Chen HI, et al. DACH1 stimulates shear stress-guided endothelial cell migration and coronary artery growth through the CXCL12-CXCR4 signaling axis. Genes Dev. 2017 Jul 1;31(13):1308-24.
307. Raftrey B, Williams I, Rios Coronado PE, Fan X, Chang AH, Zhao M, et al. Dach1 Extends Artery Networks and Protects Against Cardiac Injury. Circ Res. 2021 Sep 17;129(7):702-16.
308. Potter SS. Single-cell RNA sequencing for the study of development, physiology and disease. Nat Rev Nephrol. 2018 Aug;14(8):479-92.
309. DeLaughter DM, Bick AG, Wakimoto H, McKean D, Gorham JM, Kathiriya IS, et al. Single-Cell Resolution of Temporal Gene Expression during Heart Development. Dev Cell. 2016 Nov 21;39(4):480-90.
310. Suryawanshi H, Clancy R, Morozov P, Halushka MK, Buyon JP, Tuschl T. Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. Cardiovasc Res. 2020 Jul 1;116(8):1446-57.
311. Asp M, Giacomello S, Larsson L, Wu C, Fürth D, Qian X, et al. A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. Cell. 2019 Dec 12;179(7):1647-60.e19.
312. Goodyer WR, Beyersdorf BM, Paik DT, Tian L, Li G, Buikema JW, et al. Transcriptomic Profiling of the Developing Cardiac Conduction System at Single-Cell Resolution. Circ Res. 2019 Aug 2;125(4):379-97.
313. Li G, Tian L, Goodyer W, Kort EJ, Buikema JW, Xu A, et al. Single cell expression analysis reveals anatomical and cell cycle-dependent transcriptional shifts during heart development. Development. 2019 Jun 14;146(12).
314. Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, et al. Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. Dev Cell. 2016 Nov 21;39(4):491-507.
315. de Soysa TY, Ranade SS, Okawa S, Ravichandran S, Huang Y, Salunga HT, et al. Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. Nature. 2019 Aug;572(7767):120-4.
316. van Beijnum JR, Rousch M, Castermans K, van der Linden E, Griffioen AW. Isolation of endothelial cells from fresh tissues. Nat Protoc. 2008;3(6):1085-91.
317. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, Marioni JC, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biology. 2019;20(1):63.
318. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics. 2017;33(8):1179-86.
319. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nature Biotechnology. 2018;36(5):421-7.

320. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019 Jun 13;177(7):1888-902 e21.
321. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods. 2019;16(12):1289-96.
322. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017 Nov;14(11):1083-6.
323. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nature Biotechnology. 2020;38(12):1408-14.
324. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018 Jun 19;19(1):477.
325. Van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeys Y, et al. Trajectory-based differential expression analysis for single-cell sequencing data. Nature Communications. 2020;11(1):1201.
326. McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BEP, et al. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. European Heart Journal. 2020;41(9):1024-36.
327. Krishnan A, Samtani R, Dhanantwari P, Lee E, Yamada S, Shiota K, et al. A detailed comparison of mouse and human cardiac development. Pediatr Res. 2014 Dec;76(6):500-7.
328. Zhang H, Pu W, Li G, Huang X, He L, Tian X, et al. Endocardium Minimally Contributes to Coronary Endothelium in the Embryonic Ventricular Free Walls. Circ Res. 2016 Jun 10;118(12):1880-93.
329. Kume T. Specification of arterial, venous, and lymphatic endothelial cells during embryonic development. Histol Histopathol. 2010 May;25(5):637-46.
330. Zhou J, Bowen C, Lu G, Knapp Iii C, Recknagel A, Norris RA, et al. Cadherin-11 expression patterns in heart valves associate with key functions during embryonic cushion formation, valve maturation and calcification. Cells Tissues Organs. 2013;198(4):300-10.
331. Gillich A, Zhang F, Farmer CG, Travaglini KJ, Tan SY, Gu M, et al. Capillary cell-type specialization in the alveolus. Nature. 2020;586(7831):785-9.
332. Wu B, Wang Y, Lui W, Langworthy M, Tompkins KL, Hatzopoulos AK, et al. Nfatc1 coordinates valve endocardial cell lineage development required for heart valve formation. Circ Res. 2011 Jul 8;109(2):183-92.
333. McCulley DJ, Kang JO, Martin JF, Black BL. BMP4 is required in the anterior heart field and its derivatives for endocardial cushion remodeling, outflow tract septation, and semilunar valve development. Dev Dyn. 2008 Nov;237(11):3200-9.
334. Seo S, Kume T. Forkhead transcription factors, Foxc1 and Foxc2, are required for the morphogenesis of the cardiac outflow tract. Developmental Biology. 2006;296(2):421-36.

335. Laforest B, Nemer M. GATA5 interacts with GATA4 and GATA6 in outflow tract development. Dev Biol. 2011 Oct 15;358(2):368-78.
336. Lupu I-E, De Val S, Smart N. Coronary vessel formation in development and disease: mechanisms and insights for therapy. Nature Reviews Cardiology. 2020;17(12):790-806.
337. Chen HI, Sharma B, Akerberg BN, Numi HJ, Kivelä R, Saharinen P, et al. The sinus venosus contributes to coronary vasculature through VEGFC-stimulated angiogenesis. Development. 2014;141(23):4500-12.
338. Sharma B, Ho L, Ford GH, Chen HI, Goldstone AB, Woo YJ, et al. Alternative Progenitor Cells Compensate to Rebuild the Coronary Vasculature in Elabela- and Apj-Deficient Hearts. Developmental Cell. 2017;42(6):655-66.e3.
339. Tian X, Hu T, Zhang H, He L, Huang X, Liu Q, et al. De novo formation of a distinct coronary vascular population in neonatal heart. Science. 2014;345(6192):90.
340. Kim KL, Seo S, Kim JT, Kim J, Kim W, Yeo Y, et al. SCF (Stem Cell Factor) and cKIT Modulate Pathological Ocular Neovascularization. Arterioscler Thromb Vasc Biol. 2019 Oct;39(10):2120-31.
341. Icli B, Wara AK, Moslehi J, Sun X, Plovie E, Cahill M, et al. MicroRNA-26a regulates pathological and physiological angiogenesis by targeting BMP/SMAD1 signaling. Circ Res. 2013 Nov 8;113(11):1231-41.
342. Li Y, Cheng CN, Verdun VA, Wingert RA. Zebrafish nephrogenesis is regulated by interactions between retinoic acid, mecom, and Notch signaling. Dev Biol. 2014 Feb 1;386(1):111-22.
343. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016 Apr 8;352(6282):189-96.
344. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. PLoS One. 2010 Sep 28;5(9).
345. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nature Immunology. 2019;20(2):163-72.
346. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols. 2009;4(8):1184-91.
347. Evans PC, Rainger GE, Mason JC, Guzik TJ, Osto E, Stamataki Z, et al. Endothelial dysfunction in COVID-19: a position paper of the ESC Working Group for Atherosclerosis and Vascular Biology, and the ESC Council of Basic Cardiovascular Science. Cardiovasc Res. 2020 Dec 1;116(14):2177-84.
348. He L, Mäe MA, Sun Y, Muhl L, Nahar K, Liébanas EV, et al. Pericyte-specific vascular expression of SARS-CoV-2 receptor ACE2 – implications for microvascular inflammation and hypercoagulopathy in COVID-19 patients. bioRxiv. 2020:2020.05.11.088500.
349. Cervantes-Barragan L, Züst R, Maier R, Sierro S, Janda J, Levy F, et al. Dendritic cell-specific antigen delivery by coronavirus vaccine vectors induces long-lasting protective antiviral and antitumor immunity. mBio. 2010;1(4):e00171-10.

350. Rauch A, Dupont A, Goutay J, Caplan M, Staessens S, Moussa M, et al. Endotheliopathy Is Induced by Plasma From Critically Ill Patients and Associated With Organ Failure in Severe COVID-19. Circulation. 2020 Nov 10;142(19):1881-4.
351. Consortium E. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 Sep 6;489(7414):57-74.
352. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc. 2017 Dec;12(12):2478-92.
353. Gu M. Using Patient-Specific Induced Pluripotent Stem Cells to Understand and Treat Pulmonary Arterial Hypertension. Molecular Mechanism of Congenital Heart Disease and Pulmonary Hypertension. 2020:131-8.
354. Kang H, Mesquitta W-T, Jung HS, Moskvin OV, Thomson JA, Slukvin II. GATA2 Is Dispensable for Specification of Hemogenic Endothelium but Promotes Endothelial-to-Hematopoietic Transition. Stem Cell Reports. 2018;11(1):197-211.
355. Jung HS, Uenishi G, Park MA, Liu P, Suknuntha K, Raymond M, et al. SOX17 integrates HOXA and arterial programs in hemogenic endothelium to drive definitive lympho-myeloid hematopoiesis. Cell Reports. 2021;34(7):108758.
356. Adams WJ, Zhang Y, Cloutier J, Kuchimanchi P, Newton G, Sehrawat S, et al. Functional vascular endothelium derived from human induced pluripotent stem cells. Stem Cell Reports. 2013;1(2):105-13.
357. Huang Y, Chen X, Che J, Zhan Q, Ji J, Fan Y. Shear Stress Promotes Arterial Endothelium-Oriented Differentiation of Mouse-Induced Pluripotent Stem Cells. Stem Cells Int. 2019;2019:1847098.
358. Giacomelli E, Meraviglia V, Campostrini G, Cochrane A, Cao X, van Helden RWJ, et al. Human-iPSC-Derived Cardiac Stromal Cells Enhance Maturation in 3D Cardiac Microtissues and Reveal Non-cardiomyocyte Contributions to Heart Disease. Cell Stem Cell. 2020 Jun 4;26(6):862-79.e11.
359. Campbell M, Chabria M, Figtree GA, Polonchuk L, Gentile C. Stem Cell-Derived Cardiac Spheroids as 3D In Vitro Models of the Human Heart Microenvironment. Methods in Molecular Biology: Springer New York; 2019. p. 51-9.
360. Tang J, Zhang H, He L, Huang X, Li Y, Pu W, et al. Genetic Fate Mapping Defines the Vascular Potential of Endocardial Cells in the Adult Heart. Circ Res. 2018 Mar 30;122(7):984-93.
361. Mikryukov AA, Mazine A, Wei B, Yang D, Miao Y, Gu M, et al. BMP10 Signaling Promotes the Development of Endocardial Cells from Human Pluripotent Stem Cell-Derived Cardiovascular Progenitors. Cell Stem Cell. 2020.
362. Miao Y, Tian L, Martin M, Paige SL, Galdos FX, Li J, et al. Intrinsic Endocardial Defects Contribute to Hypoplastic Left Heart Syndrome. Cell Stem Cell. 2020 Oct 1;27(4):574-89.e8.
363. Cheng L, Xie M, Qiao W, Song Y, Zhang Y, Geng Y, et al. Generation and characterization of cardiac valve endothelial-like cells from human pluripotent stem cells. Communications Biology. 2021;4(1):1039.
364. Räsänen M, Sultan I, Paech J, Hemanthakumar KA, Yu W, He L, et al. VEGF-B Promotes Endocardium-Derived Coronary Vessel Development and Cardiac Regeneration. Circulation. 2021 Jan 5;143(1):65-77.

365. Sissaoui S, Yu J, Yan A, Li R, Yukselen O, Kucukural A, et al. Genomic Characterization of Endothelial Enhancers Reveals a Multifunctional Role for NR2F2 in Regulation of Arteriovenous Gene Expression. Circ Res. 2020 Mar 27;126(7):875-88.
366. Reyes-Palomares A, Gu M, Grubert F, Berest I, Sa S, Kasowski M, et al. Remodeling of active endothelial enhancers is associated with aberrant gene-regulatory networks in pulmonary arterial hypertension. Nature Communications. 2020;11(1):1673.
367. Liu F, Li D, Yu YY, Kang I, Cha MJ, Kim JY, et al. Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. EMBO Rep. 2015 May;16(5):654-69.
368. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. Lancet. 2020 May 2;395(10234):1417-8.
369. Goldsmith CS, Miller SE, Martines RB, Bullock HA, Zaki SR. Electron microscopy of SARS-CoV-2: a challenging task. Lancet. 2020 May 30;395(10238):e99.
370. Liu F, Han K, Blair R, Kenst K, Qin Z, Upcin B, et al. SARS-CoV-2 Infects Endothelial Cells In Vivo and In Vitro. Front Cell Infect Microbiol. 2021;11:701278.
371. Nicklin SA, Von Seggern DJ, Work LM, Pek DC, Dominiczak AF, Nemerow GR, et al. Ablating adenovirus type 5 fiber-CAR binding and HI loop insertion of the SIGYPLP peptide generate an endothelial cell-selective adenovirus. Mol Ther. 2001 Dec;4(6):534-42.